

# **Bewertung analytischer Planungsmethoden zur Erhöhung der Prognosesicherheit von Rahmenterminplänen bei Bauprojekten**

Zur Erlangung des akademischen Grades einer

**DOKTORIN DER INGENIEURWISSENSCHAFTEN  
(Dr.-Ing.)**

von der KIT-Fakultät für  
Bauingenieur-, Geo- und Umweltwissenschaften  
des Karlsruher Instituts für Technologie (KIT)

genehmigte

**DISSERTATION**

von

**Svenja Lauble (geb. Oprach), M.Sc.**  
aus Bad Säckingen

Tag der mündlichen Prüfung:	28.08.2023
Hauptreferent:	Prof. Dr.-Ing. Dipl.-Kfm. Shervin Haghsheno
Korreferent:	Prof. Dr. rer. pol. Frank Schultmann



# Kurzfassung

In einem Bauprojekt werden in frühen Projektphasen erstmals Dauern für das Projekt durch den Bauherrn prognostiziert und festgehalten. Diese bilden die Grundlage für die weitere (Termin-) Planung. Abweichungen dieser geplanten Dauern haben weitreichende Auswirkungen auf das Projekt und seine Umwelt. Diese Auswirkungen umfassen: Erstens folgt eine Unzufriedenheit seitens der Nutzer oder Betreiber, da vertraglich vereinbarte Zwischen- oder Endtermine nicht eingehalten werden können. Durch diese nicht eingehaltenen Termine können weitere Anforderungen der Stakeholder an den Schnittstellen des Projektes zu seiner Umwelt nicht erfüllt werden. So ist beispielsweise ein verzögerter Produktionsstart, Einzug der Mieter oder eine verzögerte Nutzung mit hohen Kosten verbunden. Drittens können Folgeverzögerungen innerhalb des Projektes entstehen. Viertens ist eine verlässliche Planung der Ressourcen erschwert. Fünftens nimmt der Aufwand und damit die verbundenen Kosten für notwendige Änderungen während der Planung und Ausführung zu. Sechstens können Abweichungen die Projektatmosphäre negativ beeinflussen und Konflikte können entstehen. Zuletzt beeinflussen Abweichungen die Reputation und damit schließlich auch die wirtschaftliche Kraft des jeweiligen Unternehmens.

Die Praxis zeigt hingegen anhand vieler Fallbeispiele die Herausforderungen einer möglichst genauen Prognose in frühen Projektphasen: Viele Informationen sind unbekannt und systematische Planungsfehlschlüsse erschweren die Prognose. Mit der Zunahme der Rechenleistung und verfügbaren Daten in einem Bauprojekt stehen analytische Planungsmethoden immer stärker im Fokus der Forschung. Mit diesen werden Daten objektiv ausgewertet, um systematische Planungsfehlschlüsse reduzieren zu können. Aufgrund der Unsicherheit in frühen

Phasen stellt sich jedoch die Frage des Potenzials analytischer Planungsmethoden mit begrenzt verfügbaren Daten zum Bauprojekt. Die vorliegende Arbeit hat daher das Ziel, den Einsatz analytischer Planungsmethoden zur Reduktion der Abweichungen von in frühen Phasen eines Bauprojektes geplanten Dauern zu untersuchen.

Mithilfe einer explorativen Studie werden erstens Strategien zur Reduktion von Abweichungen geplanter Dauern statistisch untersucht. Zweitens werden ausgewählte analytische Planungsmethoden anhand von Beispieldatensätzen im Hinblick auf ihre Prognosegenauigkeit verglichen. Hier werden die Ergebnisse der Untersuchung zur Prognosegenauigkeit den Erwartungen der Projektbeteiligten an ein zufriedenstellendes Niveau von Abweichungen der frühen Planung gegenübergestellt. Drittens wird anhand eines Beispieldatensatzes die Fehlschätzung analytischer Planungsmethoden der Expertenintuition gegenübergestellt. Mit den Ergebnissen der Arbeit erhalten Bauherren in frühen Phasen eines Bauprojektes Hilfestellungen zur Planung von Dauern in frühen Projektphasen, zur gezielten Anwendung analytischer Methoden und zur Projektdokumentation.

# Abstract

In early project phases, durations of a construction project are forecast and fixed by the client or their representatives for the first time. The effort for changes increases during the project and deviations have far-reaching effects. Firstly, this is followed by dissatisfaction on the part of clients or operators because contractually agreed interim or final deadlines cannot be met. Secondly, the non-compliance with these deadlines prevents the fulfillment of stakeholder requirements regarding the project's interfaces with its environment. For instance, a delayed start of production, tenant occupancy, or delayed utilization entail high costs. Thirdly, consequential delays can occur within the project. Fourth, reliable planning of resources is more difficult. Fifth, the effort and thus the associated costs for necessary changes during planning and execution increase. Sixth, deviations can negatively influence the project atmosphere and conflicts can arise. Finally, deviations affect the reputation and thus ultimately the economic strength of the respective company.

Planning as accurately as possible is therefore the subject of numerous research studies. In practice, however, many case studies show the challenges during predictions in early project phases: Much information is unknown and systematic planning errors make planning difficult. With the increase in computing power and available data in a construction project, analytical planning methods are therefore increasingly the focus of research. With these, data can be evaluated objectively, and systematic planning errors can be reduced. However, due to the uncertainty in early project phases, the question of the potential of analytical planning methods with limited available data arises. Therefore, this thesis aims to investigate the use of analytical planning methods to reduce deviations of planned durations in early project phases of a construction project.

Firstly, strategies to reduce deviations of planned durations are statistically investigated by means of an explorative study. Secondly, selected analytical planning methods are compared in terms of their forecasting accuracy using sample data sets. The results are contrasted with the expectations of project participants for a satisfactory level regarding deviations in the early planning. In a third part, the misestimation of analytical planning methods to expert intuition is compared using a sample data set. With the results of the work, clients of a construction project or their representatives receive assistance for planning durations in early project phases, for a targeted application of analytical methods and for their project documentation.

# Inhaltsverzeichnis

<b>Kurzfassung</b> . . . . .	<b>i</b>
<b>Abstract</b> . . . . .	<b>iii</b>
<b>Abbildungsverzeichnis</b> . . . . .	<b>xi</b>
<b>Tabellenverzeichnis</b> . . . . .	<b>xv</b>
<b>Abkürzungen und Symbole</b> . . . . .	<b>xix</b>
<b>Hinweis</b> . . . . .	<b>xxiii</b>
<b>1 Einleitung</b> . . . . .	<b>1</b>
1.1 Ausgangslage und Problemstellung . . . . .	1
1.2 Stand der Forschung . . . . .	4
1.3 Ziele der Arbeit und Forschungsfragen . . . . .	6
1.4 Vorgehensweise der Arbeit . . . . .	9
<b>2 Grundlagen analytischer Planungsmethoden in frühen Phasen eines Bauprojektes</b> . . . . .	<b>15</b>
2.1 Vorbemerkung . . . . .	15
2.2 Begriffe des Projektmanagements in Bauprojekten . . . . .	16
2.3 Terminplanung in frühen Projektphasen . . . . .	19
2.3.1 Terminplanung zur Projektstrukturierung . . . . .	19
2.3.2 Die Rahmenterminplanung . . . . .	21
2.3.3 Zuständigkeiten in der Rahmenterminplanung . . . . .	25
2.3.4 Verständnis der prognostizierten Dauern . . . . .	26
2.4 Analytische Planungsmethoden . . . . .	27

2.4.1	Expertenintuition und Analytik . . . . .	27
2.4.2	Wettbewerbsvorteil durch analytische Planungsmethoden .	31
2.4.3	Stufen analytischer Planungsmethoden . . . . .	32
2.5	Abweichungen von geplanten Dauern . . . . .	34
2.5.1	Auswirkungen von Abweichungen . . . . .	34
2.5.2	Komplexität von Bauprojekten . . . . .	37
2.5.3	Unsicherheit als Ursache . . . . .	41
2.5.4	Ansätze zur Reduktion von Unsicherheiten . . . . .	44
2.5.5	Systematische Planungsfehlschlüsse als Ursache für Abweichungen . . . . .	46
2.5.6	Ansätze zur Reduktion von systematischen Planungsfehlschlüssen . . . . .	50
2.6	Schlussfolgerungen und Anforderungen an analytische Planungsmethoden zur Rahmenterminplanung . . . . .	53
<b>3</b>	<b>Klassifizierung bestehender Forschungsarbeiten zu analytischen Planungsmethoden für die Terminplanung . . .</b>	<b>61</b>
3.1	Vorbemerkung . . . . .	61
3.2	Methodisches Vorgehen der Literaturrecherche . . . . .	62
3.3	Klassifizierung der analytischen Planungsmethoden . . . . .	66
3.3.1	Methoden der statistischen Analyse . . . . .	66
3.3.2	Hochrechnungen . . . . .	69
3.3.3	Prädiktive Modelle . . . . .	74
3.3.4	Optimierung . . . . .	77
3.4	Evaluierung der kategorisierten Planungsmethoden und Aufzeigen der Forschungslücke . . . . .	80
3.5	Klassifizierung relevanter Projektmerkmale . . . . .	83
3.6	Zusammenfassung und kritische Betrachtung . . . . .	87
<b>4</b>	<b>Explorative Studie zu wirksamen Ansätzen zur Reduktion von Abweichungen geplanter Dauern . . . . .</b>	<b>91</b>
4.1	Vorbemerkung . . . . .	91
4.2	Methodisches Vorgehen der explorativen Studie . . . . .	92
4.2.1	Übersicht . . . . .	92
4.2.2	Schritt 1: Item-Generierung . . . . .	93



---

4.2.3	Schritt 2: Durchführung der Erhebung . . . . .	98
4.2.4	Schritt 3: Reduktion der ursprünglichen Items . . . . .	100
4.2.5	Schritt 4: Weitere Konstruktanalyse . . . . .	100
4.2.6	Schritt 5: Wiederholung . . . . .	101
4.3	Ergebnisse der explorativen Studie . . . . .	102
4.3.1	Ergebnisse der deskriptiven univariaten Analyse . . . . .	102
4.3.2	Ergebnisse der explorativen Faktoranalyse . . . . .	106
4.3.3	Validitätsüberprüfung des Systemkonstrukts . . . . .	110
4.4	Zusammenfassung und kritische Betrachtung . . . . .	114
<b>5</b>	<b>Grundlagen ausgewählter Methoden der Künstlichen</b>	
	<b>Intelligenz . . . . .</b>	<b>117</b>
5.1	Vorbemerkung . . . . .	117
5.2	Künstliche Intelligenz . . . . .	118
5.3	Maschinelles Lernen . . . . .	121
5.3.1	Ziele des Maschinellen Lernens . . . . .	121
5.3.2	Lernen . . . . .	122
5.3.3	Aufgaben . . . . .	126
5.3.4	Verlustfunktion . . . . .	128
5.4	Künstliche Neuronale Netze . . . . .	131
5.5	Deep Learning . . . . .	133
5.6	Ausgewählte Methoden der erklärbaren Künstlichen Intelligenz . . . . .	133
5.6.1	Ziele der erklärbaren Künstlichen Intelligenz . . . . .	133
5.6.2	Entscheidungsbäume . . . . .	135
5.6.3	SHAP-Framework . . . . .	139
5.7	Zusammenfassung . . . . .	141
<b>6</b>	<b>Vergleich und Bewertung ausgewählter analytischer</b>	
	<b>Planungsmethoden . . . . .</b>	<b>143</b>
6.1	Vorbemerkung . . . . .	143
6.2	Herleitung der Anforderungen an die Prognosegenauigkeit . . . . .	144
6.2.1	Ziele der Umfrage . . . . .	144
6.2.2	Methodisches Vorgehen der Umfrage . . . . .	145
6.2.3	Ergebnisse der Umfrage . . . . .	147

6.2.4	Zusammenfassung der Anforderungen an die Prognosegenauigkeit . . . . .	152
6.3	Bewertung der Prognosegenauigkeit . . . . .	153
6.3.1	Methodisches Vorgehen zur Modellanwendung . . . . .	153
6.3.2	Beschreibung der verwendeten Datensätze . . . . .	156
6.3.3	Bewertung der ausgewählten Datensätze . . . . .	161
6.3.4	Ergebnisse der Datenanalyse . . . . .	165
6.3.5	Evaluierung relevanter Merkmale in den Modellen . . . . .	169
6.4	Zusammenfassung und kritische Betrachtung . . . . .	172
<b>7</b>	<b>Bewertung der Expertenintuition im Vergleich zu ausgewählten analytischen Planungsmethoden . . . . .</b>	<b>177</b>
7.1	Vorbemerkung . . . . .	177
7.2	Methodisches Vorgehen zur Modellanwendung . . . . .	179
7.3	Beschreibung des ausgewählten Datensatzes . . . . .	183
7.4	Bewertung des ausgewählten Datensatzes . . . . .	187
7.5	Ergebnisse der Datenanalyse . . . . .	190
7.6	Zusammenfassung und kritische Betrachtung . . . . .	198
<b>8</b>	<b>Handlungsempfehlungen für die Praxis . . . . .</b>	<b>201</b>
8.1	Empfehlungen zur Mensch-Maschinen-Kollaboration . . . . .	201
8.2	Zusammenfassung der Empfehlungen in Handlungsfelder . . . . .	204
<b>9</b>	<b>Schlussbetrachtung . . . . .</b>	<b>209</b>
9.1	Zusammenfassung . . . . .	209
9.2	Kritische Würdigung und Ausblick . . . . .	213
	<b>Literaturverzeichnis . . . . .</b>	<b>217</b>
<b>A</b>	<b>Anhang zu Kapitel 3 . . . . .</b>	<b>249</b>
<b>B</b>	<b>Anhang zu Kapitel 4 . . . . .</b>	<b>263</b>
B.1	Übersicht der Steuerungsfragen im Interview . . . . .	263
B.2	Online-Umfrage . . . . .	264
<b>C</b>	<b>Anhang zu Kapitel 6 . . . . .</b>	<b>277</b>

C.1	Umfrage zu Anforderungen an die Prognosegenauigkeit . . . . .	277
C.2	Datensätze aus Teheran, Iran, und San Francisco, USA . . . . .	281
C.3	Externe Marktbeschreibende Daten . . . . .	284
C.4	Python Code . . . . .	287
<b>D</b>	<b>Anhang zu Kapitel 7 . . . . .</b>	<b>335</b>
D.1	Datensatz aus New York, USA . . . . .	335
D.2	Python Code . . . . .	336



# Abbildungsverzeichnis

1.1	Einsparmöglichkeiten im Projektverlauf (in Anlehnung an Brandenberger und Ruosch 1991, S. 26, Stark 2006, S. 35) . . . . .	2
1.2	Beispielhafter Rahmenterminplan nach AHO (2014, S. 120) . . . . .	4
1.3	Aufbau der Arbeit . . . . .	11
2.1	Zusammenhang zwischen Bauprojekten und seinen Projektmerkmalen (vgl. Mittag und Schüller 2020, S. 23) . . . . .	17
2.2	Wissenstreppe nach North (2016, S. 23) . . . . .	31
2.3	Analytische Methoden und Wettbewerbsvorteil Davenport und Harris (vgl. 2007, S. 8) . . . . .	33
2.4	Wirkungskette und Folgerung der Anforderungen an die Terminplanung in frühen Phasen eines Bauprojektes . . . . .	56
3.1	Vorgehen der Literaturrecherche . . . . .	63
3.2	Vergleich ausgewählter multivariater Regressionen verschiedener Autoren . . . . .	72
3.3	Anzahl der identifizierten Literaturquellen in den Kategorien . . . . .	81
3.4	Anzahl verfügbarer Publikationen nach Jahr in der entsprechenden Kategorie . . . . .	82
3.5	Klassifizierung der Einflussmerkmale mit RapidMiner . . . . .	84
3.6	Auswertung der beeinflussenden Merkmale auf Basis der Literaturrecherche mit RapidMiner und Tableau. . . . .	85
4.1	Univariate Ergebnisse zu den Teilnehmern der Studie und ihrer Erfahrung . . . . .	103
4.2	Univariate Ergebnisse zu der Anzahl dokumentierter Projekte und häufig dokumentierter Datenpunkte . . . . .	105
4.3	Univariate Ergebnisse der Methoden zur Definition von Projektdauern . . . . .	106

4.4	Qualitätsbezogene Validitätsbewertung des identifizierten Frameworks mit den Signifikanzcodes: '***' 0,001 '**' 0,01 '*' 0,05; graue, gestrichelte Pfeile kennzeichnen nicht signifikante Beziehungen . . . . .	110
4.5	Korrelationsmonoplot für Kriterien zur Konstruktvalidität (Items des Konstrukts in fetter Schrift, andere Items in grauer Schrift) . . . . .	113
5.1	KI, ML, DL (vgl. Copeland 2016, Denk 2017) . . . . .	118
5.2	Das Programmierparadigma des ML im Vergleich zur klassischen Programmierung . . . . .	122
5.3	Regression und Klassifikation . . . . .	127
5.4	Beispiel eines Clustering . . . . .	128
5.5	Idealisierter Test- und Trainingsverlauf . . . . .	130
5.6	Verschiedene KNN . . . . .	132
5.7	Vereinfachte Darstellung eines Entscheidungsbaums (vgl. Myles et al. 2004) . . . . .	136
5.8	Beispielhafte Darstellung eines SHAP-Frameworks (vgl. Lundberg 2018) . . . . .	140
6.1	Zusammenfassung des Vorgehens in Kapitel 6 . . . . .	144
6.2	Ergebnisse der Umfrage: Rolle des Unternehmens aufgeteilt in Auftraggeber, Auftragnehmer und Beratungen . . . .	148
6.3	Ergebnisse der Umfrage: Unternehmensgröße anhand Mitarbeiteranzahl und Umsatz . . . . .	149
6.4	Ergebnisse der Umfrage zur Zufriedenheit bei Abweichungen des jeweiligen Terminplans . . . . .	150
6.5	Schritte bei der Anwendung eines Konzeptes des Maschinellen Lernens . . . . .	154
6.6	Boxlplot zum Vergleich des Datensatzes 1 (Teheran, Iran) und Datensatz 2 (San Francisco, USA) in Quartalen . . . . .	159
6.7	Verteilung der Datensätze: Linke Abbildung Teheran, Iran in Quartalen; rechte Abbildung San Francisco, USA in Tagen (Normalverteilung vor (blau) und nach (schwarz) der log-Transformation) . . . . .	160
6.8	Entscheidungsbaum eines Random Forest (auf Basis von Datensatz zwei - San Francisco, USA) . . . . .	169

---

6.9	SHAP-Framework zu Datensatz eins (Teheran, Iran) . . . . .	170
6.10	SHAP-Framework zu Datensatz zwei (San Francisco, USA) . . . . .	171
7.1	Methodisches Vorgehen zum Vergleich der Vorhersagefehler der Expertenintuition mit dem trainierten Modell sowie Analyse ihrer situativen Treiber . . . . .	180
7.2	Boxplot zur Dauer des Datensatzes aus New York, USA . . . . .	185
7.3	Histogramm über die Projektdauer (in Tagen) und Anzahl der Bauprojekte . . . . .	186
7.4	Boxplot zum Vergleich des absoluten Vorhersagefehlers in Tagen zwischen der Expertenintuition und dem trainierten Modell CatBoost . . . . .	191
7.5	Shapely-Darstellung für Datensatz eins bis vier (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren) .	195
8.1	Ein beispielhaftes Modell zur Zusammenarbeit zwischen Experte und maschinellem Lernmodell (vgl. Bernard et al. 1991) .	203
8.2	Die vier Handlungsfelder zum Einsatz von Entscheidungsbäumen als Teil maschineller Lernmodelle zur Planung von Dauern in frühen Projektphasen eines Bauprojektes (angelehnt an Österle und Winter 2000, S. 12) . . .	205
9.1	Zusammenfassung der Ergebnisse . . . . .	212





# Tabellenverzeichnis

2.1	Kategorisierung von Rahmenterminplänen durch Maßnahmen und Puffer . . . . .	37
2.2	Zusammenfassung der Ansätze zur Reduktion von Unsicherheit (vgl. Brashers 2001) und systematischen Planungsfehlschlüssen (vgl. Siemiatycki 2009) . . . . .	55
3.1	Systematische Literaturlauswahl . . . . .	65
4.1	Struktur des Interviewleitfadens für die Experteninterviews . . . . .	94
4.2	Übersicht der verwendeten Items (* Verwendung für univariate Ergebnisse) . . . . .	96
4.3	EFA mit Varimax-Rotation und Kaiser-Normalisierung; Hauptgewichtungen in schwarzer und Kreuzgewichtungen in grauer Schrift . . . . .	107
4.4	Matrix der Spearman-Rangkorrelationskoeffizient mit 95 % Konfidenzintervall, p-Wert und Kovarianzen zur Zeitabweichung (Item 17) . . . . .	112
6.1	Vergleich von Auftraggeber (AG), Auftragnehmer (AN) und Beratungen bei der Zufriedenheit von Abweichungen zwischen der Rahmenterminplanung und der Bauausführung . . . . .	151
6.2	Pearson Korrelationskoeffizient und p-Wert für den Zusammenhang zwischen der Berufserfahrung und der Zufriedenheit bei Abweichungen (in kursiv moderate Korrelation) . . . . .	152
6.3	Beschreibung der verwendeten Datensätze (Zur besseren Vergleichbarkeit sind für Datensatz 2 die Dauern in Quartale umgerechnet. Es wird angenommen, dass 90 Tage einem Quartal entsprechen) . . . . .	159

6.4	Vergleich der Prognosegüte der Entscheidungsbäume anhand des Performance-Indikators MAPE (* ohne externe Daten; ** mit externen Daten; in fett bestes Ergebnis pro Datensatz) . . . . .	166
6.5	Vergleich der Prognosegüte zwischen KNN, dem CatBoost und linearen Regressionsmodellen repräsentiert durch die jeweiligen Autoren mit dem Performance-Indikator MAPE (* mit externen Daten; in fett beste Ergebnisse des Datensatzes zwei) . . . . .	167
7.1	Beschreibung des verwendeten Datensatzes New York, USA . . . . .	185
7.2	Übersicht über die Anzahl der Projekte je Kategorie . . . . .	187
7.3	Vergleich des absoluten Vorhersagefehlers (MAE) in Tagen zwischen der Expertenintuition und dem trainierten Modell CatBoost (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren), in fett die Ergebnisse mit dem geringsten Vorhersagefehler je Zeile . . . . .	191
7.4	Vergleich der Fälle, in denen das trainierte Modell oder der Experte die bessere Prognose durchgeführt hat (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren) . . . . .	192
7.5	Vergleich der präferierten Fälle für die Expertenintuition oder der CatBoost je Projektkategorie (in Prozent) (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren) . . . . .	194

---

7.6	Vergleich der Präferenz des maschinellen Lernmodells und der Expertenintuition anhand der „Phase zu Beginn“ in Datensatz drei . . . . .	196
A.1	Literaturübersicht der gefilterten Veröffentlichungen (1/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	250
A.2	Literaturübersicht der gefilterten Veröffentlichungen (2/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	251
A.3	Literaturübersicht der gefilterten Veröffentlichungen (3/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	252
A.4	Literaturübersicht der gefilterten Veröffentlichungen (4/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	253
A.5	Literaturübersicht der gefilterten Veröffentlichungen (5/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	254
A.6	Literaturübersicht der gefilterten Veröffentlichungen (6/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	255
A.7	Literaturübersicht der gefilterten Veröffentlichungen (7/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	256
A.8	Literaturübersicht der gefilterten Veröffentlichungen (8/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	257
A.9	Literaturübersicht der gefilterten Veröffentlichungen (9/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	258
A.10	Literaturübersicht der gefilterten Veröffentlichungen (10/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	259

A.11	Literaturübersicht der gefilterten Veröffentlichungen (11/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	260
A.12	Literaturübersicht der gefilterten Veröffentlichungen (12/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen) . . . . .	261
B.1	Steuerungsfragen im Interviewleitfaden . . . . .	263
B.2	Steuerungsfragen im Interviewleitfaden . . . . .	264
C.1	Datensatz 1: Teheran, Iran (1/2) . . . . .	281
C.2	Datensatz 1: Teheran, Iran (2/2) . . . . .	282
C.3	Datensatz 2: San Francisco, USA . . . . .	283
C.4	Externe Daten zur Beschreibung der wirtschaftlichen und politischen Lage (1) . . . . .	285
C.5	Externe Daten zur Beschreibung der wirtschaftlichen und politischen Lage (2) . . . . .	286
D.1	Datensatz von Investitionsprojekten in New York City Office of Operations (2018) . . . . .	336

# Abkürzungen und Symbole

## Abkürzungen

<b>AHO</b>	Ausschuss der Verbände und Kammern der Ingenieure und Architekten für die Honorarordnung
<b>AACE</b>	Association for the Advancement of Cost Engineering
<b>ADDSS</b>	Activity Duration Decision Support System
<b>ANOVA</b>	Analysis of Variance oder Varianzanalyse
<b>BAI</b>	Bruttoanlageinvestitionen
<b>BIP</b>	Bruttoinlandsprodukt
<b>DACH</b>	Deutschland, Österreich, Schweiz
<b>DL</b>	Deep Learning
<b>EFA</b>	Explorative Faktoranalyse
<b>GBR</b>	Gradient Boosting Regression
<b>GDPR</b>	General Data Protection Regulation
<b>GOSS</b>	Gradient-based One-Side Sampling oder einseitiges Samplingverfahren
<b>HOAI</b>	Honorarordnung für Architekten und Ingenieure
<b>I</b>	Item

<b>IN</b>	Interviewer
<b>IP</b>	Interviewpartner
<b>IPA</b>	Integrierte Projektabwicklung
<b>IRR</b>	Iranische Rial
<b>KI</b>	Künstliche Intelligenz
<b>KIT</b>	Karlsruher Institut für Technologie
<b>KNN</b>	Künstliche Neuronale Netze
<b>MAE</b>	Mean Absolute Error oder mittlerer absoluter Fehler
<b>MAPE</b>	Mean Absolute Percentage Error oder mittlerer absoluter prozentualer Fehler
<b>max.</b>	maximal
<b>MATLAB</b>	MATrix LABoratory
<b>ML</b>	Machine Learning oder Maschinelles Lernen
<b>N/A</b>	Not available bzw. nicht verfügbar
<b>NEDO</b>	National Economic Development Office
<b>OECD</b>	Organisation for Economic Co-operation and Development
<b>PMI</b>	Projektmanagement-Institut
<b>PT</b>	Befragter im Pre-Test
<b>RIBA</b>	Royal Insitute of British Architects
<b>RII</b>	Relative Importance Index
<b>SHAP</b>	SHapley Additive exPlanations
<b>SIA</b>	Schweizer Ingenieur- und Architektenverein

<b>SVM</b>	Support Vector Machine oder Stützvektormethode
<b>TMB</b>	Institut für Technologie und Management im Baubetrieb
<b>UK</b>	United Kingdom
<b>USA</b>	United States of America
<b>USD</b>	US Dollar
<b>V</b>	Variable
<b>WoS</b>	Web of Science
<b>XAI</b>	eXplainable Artificial Intelligence
<b>XGBoost</b>	eXtreme Gradient Boosting

## **Symbole**

$A$	Gewicht
$a_m$	Achsenabschnitt bzw. Geradenneigung
$\alpha(p)$	Wert zur Prognose mit und ohne das Merkmal
$B$	Konstante zur Sensitivität der Zeit zu den Kosten
$C$	Kosten des Vertrages (in Millionen Dollar)
$f$	Funktion
$\hat{f}$	approximierte Funktion
$f(\mathcal{X})$	Wertebereich
$H_{et}$	Vorhersagefehler der Expertenintuition
$i$	spezifisches Merkmal
$k$	Parameter der Kreuzvalidierung

$K$	Konstante für die Projektperformance
$L$	Verlustfunktion
$m^2$	Quadratmeter
$M_{et}$	Vorhersagefehler des trainierten Modells
$m_j$	Merkmal
$m_{jk}$	Merkmalsträger
$N$	Teilnehmer
$n$	Anzahl Nennungen
$\theta_i$	Shapely-Wert
$P$	Zufallsverteilung
$\mathbb{R}$	Menge reeller Zahlen
$r_{sp}$	Spearman-Rangkoeffizient
$S, F$	Teilmenge der Merkmale
$T$	Zeit (in Arbeitstagen)
$w$	Bewertung bzw. Gewichtung
$\Delta w_j$	Änderungen der Gewichtungen
$\mathcal{X}$	Menge an Eingaben
$x$	Eingabe
$x_i, y_i, (z_i)$	Datenpunkt im zwei- (drei-) dimensionalem Raum
$X_{n+1}$	Unbeschriftete Merkmalsvektoren
$y$	Ausgabe
$\hat{y}$	Vorausgesagter Wert



# Hinweis

Bei personenbezogenen Bezeichnungen bezieht sich das generische Maskulinum auf weibliche, männliche und diverse Personen. Zum Zweck der besseren Lesbarkeit wird auf die geschlechtsspezifische Schreibweise sowie auf eine Mehrfachbezeichnung verzichtet. Alle personenbezogenen Bezeichnungen sind somit geschlechtsneutral zu verstehen.



# 1 Einleitung

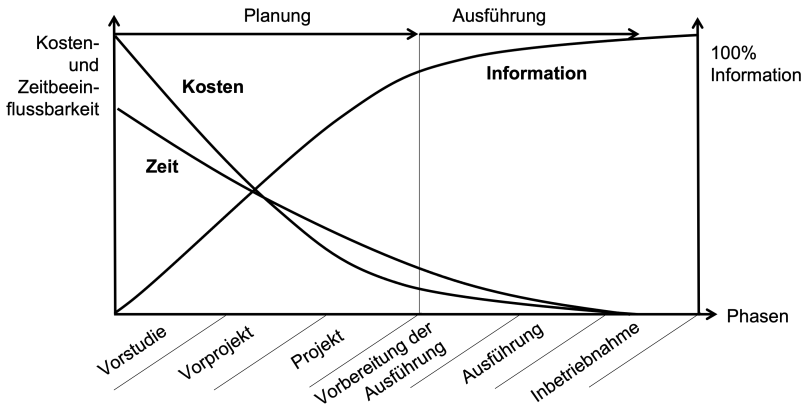
## 1.1 Ausgangslage und Problemstellung

In Bauvorhaben dient die erste Terminplanung in frühen Projektphasen als Ausgangsbasis für das gesamte Projekt und definiert damit wesentliche Parameter. Die Terminplanung wird darauf aufbauend sukzessive verfeinert. So werden auf Basis prognostizierter Projektphasen einzelne Vorgänge detailliert und erste Prognosen werden festgeschrieben. Die Definition und Kommunikation der ersten Prognose zu geplanten Dauern einzelner Projektphasen ist damit Grundlage für:

- Investitionsentscheidungen,
- Ressourcenplanung,
- Definition von vertraglich vereinbarten Synchronisationspunkten und die Koordination mit Partnern,
- Projektatmosphäre: Das Projektteam arbeitet entweder unter Termindruck oder mit Zeitpuffern (vgl. Rogel und Osebold 2013) und
- Definition des Projekterfolgs.

Oft können jedoch die prognostizierten Dauern der einzelnen Projektphasen eines Bauprojektes nicht eingehalten werden (vgl. Potts 2005, Magnussen und Olsson 2006). Die Folge sind Abweichungen zwischen der geplanten und der realisierten Dauer. So zeigen auch bekannte Bauprojekte, wie der Bau des Eurotunnels, des Opernhauses in Sydney (vgl. Hall 1982) oder zahlreiche Olympiastadien (vgl.

Flyvbjerg et al. 2016) eine deutliche Abweichung zu ihrem ursprünglichen Zeitplan. Diese Abweichungen haben direkte wirtschaftliche Auswirkungen auf die beteiligten Unternehmen und beeinflussen auch die Volkswirtschaft durch die nicht zeitgerechte Bereitstellung von Bauwerken und Infrastruktur. Mit steigendem Termindruck zum Ende einer Projektphase werden u. a. Kosten durch Änderungen, die Qualität, Konflikte, Behinderungen (vgl. Braimah und Ndekugri 2008) sowie die Arbeitssicherheit beeinflusst. Die Nutzer bzw. Kunden sind in Folge der fehlenden Zielerreichung unzufrieden. Auch eine Unterschreitung der geplanten Dauern verursacht Kosten durch gebundene Ressourcen und nicht genutzte Potenziale im Projekt. So definiert für viele Beteiligte die termingerechte Fertigstellung den Erfolg des Bauprojekts (vgl. Latham 1994, Rwelamila und Hall 1995, Lim und Mohamed 1999, Cox et al. 2003). Auch der 1983 veröffentlichte, „NEDO-Bericht für schnelleres Bauen für die Industrie“ sieht die Einhaltung der Zeit als Grundlage für den Erfolg.



**Abbildung 1.1:** Einsparmöglichkeiten im Projektverlauf (in Anlehnung an Brandenberger und Ruosch 1991, S. 26, Stark 2006, S. 35)

Änderungen an den festgelegten Dauern sind häufig mit hohen Folgekosten verbunden und die Möglichkeit zur Beeinflussung nimmt mit dem Projektverlauf

ab. Die Planung in den frühen Projektphasen ist demnach für die Einhaltung von Terminen und Kosten äußerst relevant. Jedoch liegen zu diesem Zeitpunkt noch wenig Informationen über das jeweilige Projekt vor. Dies zeigt Abbildung 1.1 in schematischer Form.

Dabei sei ein Projekt entlang der Planung und Realisierung beispielhaft nach den Projektphasen von Brandenberger und Ruosch (1991, S. 26) unterteilt. Während Möglichkeiten der Beeinflussung von Terminen und Kosten zu Beginn des Projekts noch hoch sind, existiert nur eine geringe Informationsmenge. Während der Projektausführung nimmt die vorhandene Informationsmenge zu und die Kosten für Änderungen im Programm wachsen exponentiell an (vgl. Stark 2006, S. 35). Die erste Terminplanung eines Bauprojektes, mit der Termine und Kosten noch am stärksten beeinflusst werden können, erfolgt oft Monate oder Jahre vor der eigentlichen Realisierung der jeweiligen Projektphase und muss daher viele Unsicherheiten betrachten. Trotz der geringen Informationsdichte und hohen Unsicherheit in einer frühen Planungsphase besteht für Bauherren und deren Projektpartner das Ziel, möglichst valide Aussagen über Dauern einzelner Projektphasen zu treffen. Damit soll ein stabiler Projektablauf gewährleistet und die generelle zeitliche Machbarkeit überprüft werden. Eine höhere Planungsgenauigkeit zu Beginn reduziert das Risiko von Kostenerhöhungen. Die erste zeitliche Planung der Projektphasen bildet daher für das Projekt und alle daran Beteiligten eine wichtige Planungs- und Erfolgsgrundlage.

International existiert kein einheitliches Rahmenwerk für eine erste Terminplanung und diese wird zudem regional unterschiedlich verstanden. In Deutschland wird die Planung von Dauern einzelner Projektphasen erstmals in der sogenannten „Rahmenterminplanung“ festgehalten. Die Rahmenterminplanung dient in der vorliegenden Arbeit als zu referenzierender Detaillierungsgrad für eine Terminplanung in frühen Phasen eines Bauprojektes. Die Rahmenterminplanung wird durch das Projektmanagement, das den Bauherrn, Projektentwickler oder Investoren vertritt oder einem beauftragten Projektsteuerer erstellt. Diese planen die wesentlichen Projektphasen der Planung und Realisierung sowie Tätigkeiten nach der Fertigstellung des Bauwerks, wie z. B. der Einzug mit Nutzungsbeginn

und die erste Nutzungsphase (vgl. Würfele et al. 2017, S. 2f.). Mit diesen Phasen soll der Zeithorizont des Gesamtprojektes dargestellt werden. Die Dauern der einzelnen Phasen erstrecken sich oft über mehrere Monaten oder werden auf einer Quartalsbasis ermittelt (vgl. Greiner et al. 2005). Ein beispielhafter Rahmen-terminplan mit neun Projektphasen ist in Abbildung 1.2 dargestellt. Dabei stellen die einzelnen Spalten Monate dar und sind zu Quartalen zusammengefasst.

Nr.	Phase	IV	I	II	III	IV	I	II	III	IV	I	II
1	Grundlagenermittlung	■										
2	Vorplanung	■	■									
3	Entwurfsplanung		■	■	■							
4	Genehmigungsplanung		■	■	■	■						
5	Ausführungsplanung		■	■	■	■	■					
6	Vergabeverfahren		■	■	■	■	■	■				
7	Baurealisierung			■	■	■	■	■	■	■	■	■
8	Einrichtung										■	■
9	Mängelbeseitigung										■	■

Abbildung 1.2: Beispielhafter Rahmenterminplan nach AHO (2014, S. 120)

Aufgrund des Potenzials und der Herausforderungen einer möglichst genauen Prognose der Dauern einzelner Projektphasen, wie im Rahmenterminplan dargestellt, existieren zahlreiche Publikationen in diesem Forschungsbereich. Diese haben das Ziel, die Ursachen etwaiger Planungsabweichungen zu untersuchen, aber auch Methoden zu identifizieren, die in der Planung unterstützen.

## 1.2 Stand der Forschung

Neben fehlenden Informationen in frühen Projektphasen beschreiben Kahneman und Tversky (1977) als weiteren Grund für Abweichungen systematische Planungsfehlschlüsse. Sie zeigen auf, dass Menschen bei der Planung von Projekten tendenziell die Zeit unterschätzen, die benötigt wird, um das Projekt erfolgreich abzuschließen. Dies kann durch externe oder interne Faktoren bedingt sein. Die Ursache für diese Planungsfehlschlüsse liegt nach Kahneman und Tversky (1977) darin, dass Planer häufig eine „interne Sicht“ einnehmen und aktive

Prozesse stärker gewichten, anstatt das aktuelle Projekt mit Projektergebnissen historischer Projekte zu vergleichen. Der Einsatz moderner analytischer Techniken kann eine Strategie zur Reduktion der Abweichungen sein (vgl. Siemiatycki 2009) und durch die Verwendung von Daten als objektive Grundlage den menschlichen Prognosefehler verringern.

In den 1980er Jahren wurden Expertensysteme entwickelt (z.B. Sheh 2017, S. 629, Biran und Cotton 2017, S. 1), die dokumentiertes Wissen zur Terminplanung einbeziehen. Beispielhaft seien hier Gray (1986), Hendrickson et al. (1987), Chevallier und Russell (2001), Navinchandra et al. (1988) oder Darwiche et al. (1988) zu nennen. Bei diesen übergeben die Programmierer dem System Regeln (vgl. House 2016, S. 18). Mit diesen Regeln wird versucht, eine komplexe Aufgabenstellung auf Kriterien zu reduzieren, die der Algorithmus nacheinander prüfen kann (vgl. Kaplan und Haenlein 2019, S. 18). Die Komplexität eines Bauprojektes erfordert jedoch die Entwicklung hoch spezialisierter Expertensysteme und benötigt daher eine Vielzahl an Informationen über das Projekt. Diese Expertensysteme sind daher in späteren Projektphasen zu finden, in denen die Informationsmenge hoch ist. Sie beziehen die Unsicherheit der frühen Projektphasen nicht mit ein und werden daher im Folgenden nicht weiter betrachtet.

Eines der ersten analytischen Modelle für eine Prognose der Phase „Bauausführung“ wurde 1960 durch den australischen Forscher Bromilow (1969) entwickelt. Dieses Modell verwendet als Datengrundlage die Kosten, die durch den Bauherrn bezahlt werden, eine allgemeine Projektperformance, die länderspezifisch zu definieren ist, sowie einen Faktor, der die Abhängigkeit der Zeit zu den Kosten angibt. Auf dem Modell von Bromilow aufbauend wurden weitere Mehrfaktoren-Modelle erstellt (vgl. Chan und Kumaraswamy 1997, Chan 1999, Choudhury und Rajan 2003, Ogunsemi und Jagboro 2006). Des Weiteren beziehen Zhang und Koreisha (2015), Agapiou et al. (1998) oder Albogamy et al. (2014) Wahrscheinlichkeiten mit ein, um die Prognose zu verbessern.

Mit der Zunahme dokumentierter Bauprojekte (vgl. Bilal et al. 2016, S. 1) und Möglichkeiten zur Datenauswertung werden Modelle zunehmend komplexer. So

finden beispielhaft „Künstliche Neuronale Netze“ (KNN) Anwendung bei Petrusseva et al. (2012), Lam und Olalekan (2016) oder Bhokha und Ogunlana (1999). Mit der Verwendung von Heuristiken werden Lösungen verbessert. Hier finden beispielhaft „genetische Algorithmen“ (vgl. Zheng et al. 2004), der „Ameisenalgorithmus“ (vgl. Zhang und Ng 2012) oder das „Simulated Annealing“ (vgl. Kumar und Abdullah 2011) Anwendung.

Trotz der intensiven Forschungstätigkeiten sind Abweichungen in Bauprojekten immer noch ein weit verbreitetes Problem in der Praxis (vgl. Potts 2005, Magnusson und Olsson 2006). Folglich stellt sich die Frage, ob die Forschungsansätze den Anforderungen der Praxis genügen oder ob es geeignetere Ansätze zur Reduktion von Abweichungen gibt.

## 1.3 Ziele der Arbeit und Forschungsfragen

Ziel dieser Arbeit ist die Untersuchung von Ansätzen zur Reduktion der Abweichungen von in frühen Projektphasen mithilfe analytischer Planungsmethoden geplanter Dauern.

Mit den Ergebnissen erhalten so Unternehmen Hilfestellungen in der Erfassung relevanter Merkmale und der Anwendung analytischer Planungsmethoden zur Rahmenterminplanung. Folgekosten bei einer Über- oder Unterschreitung der geplanten Dauer sowie notwendige Anpassungen des Terminrahmens können für Unternehmen reduziert werden. Ebenso entsteht ein volkswirtschaftlicher Nutzen durch die Sicherheit in Terminen und der Ressourcenplanung.

Die Arbeit liefert damit folgende neue Erkenntnisse:

- Übersicht und Bewertung des wirtschaftlichen Vorteils bestehender analytischer Planungsmethoden in der Forschung sowie eine Identifikation von Forschungslücken,
- Identifikation und Analyse von Anforderungen an analytische Planungsmethoden für die Rahmenterminplanung,



- Übersicht des Status Quo der Datenerfassung und Planung der Rahmentermine in der Praxis im Untersuchungsraum Deutschland,
- Identifikation von strategischen Treibern zur Reduktion von Abweichungen mit einem statistisch signifikanten Einfluss in der Praxis,
- Bewertung der Prognosegenauigkeit ausgewählter analytischer Methoden zur Reduktion von Abweichungen anhand beispielhafter Datensätze, auch im Vergleich zur Expertenintuition,
- Sammlung und Bewertung relevanter Merkmale für den Einsatz analytischer Planungsmodelle sowie der situativen Bevorzugung analytischer Planungsmethoden oder einer Expertenintuition.

Die Hauptforschungsfrage wird wie folgt definiert:

**Inwieweit sind analytische Planungsmethoden zur Reduktion der Abweichungen von in frühen Projektphasen eines Bauprojektes geplanten Dauern dienlich?**

Hieraus leiten sich folgende drei Forschungsfragen ab:

**1. Forschungsfrage:**

Wie werden derzeit Rahmenterminpläne für Bauprojekte im Untersuchungsraum Deutschland durch das Projektmanagement erstellt und welche Ansätze zur Reduktion von Abweichungen geplanter Dauern sind statistisch signifikant? (Fragetyp: Deskription)

**2. Forschungsfrage:**

Welche ausgewählten Modelle analytischer Planungsmethoden sind in der frühen Projektphase zur Prognose von Dauern dienlich? (Fragetyp: Gestaltung)

Hieraus ergeben sich weitere untergeordnete Fragen:

- Welche Anforderungen hinsichtlich ihrer Gestaltung und der durch die Praxis geforderten Prognosegenauigkeit haben in frühen Projektphasen analytische Planungsmethoden zur Planung von Dauern?

- Wie werden diese Anforderungen in bestehenden Forschungsarbeiten umgesetzt?
- Wie hoch ist die Prognosegenauigkeit der ausgewählten Methoden in beispielhaften Datensätzen?

### 3. Forschungsfrage:

Inwiefern lässt sich im Vergleich zur Planung durch Experten mit den ausgewählten Modellen analytischer Planungsmethoden in einer frühen Planungsphase die Prognosegenauigkeit von Dauern eines Bauprojektes unterstützen und verbessern? (Fragetyp: Prognose)

Auch hieraus ergeben sich untergeordnete Fragen:

- Wie unterscheidet sich die Prognosegenauigkeit der Expertenintuition zu den ausgewählten analytischen Planungsmethoden in einem beispielhaften Datensatz im direkten Vergleich?
- Welche Projektmerkmale haben Einfluss auf die Differenzierung?

In der ersten Forschungsfrage wird der Untersuchungsraum Deutschland gewählt. Diese Begrenzung wird aufgrund der international nicht einheitlichen Definition der ersten Terminplanung vorgenommen. Die Definitionen der Honorarordnung für Architekten und Ingenieure (HOAI) (z.B. HOAI 2018) und des Ausschusses der Verbände und Kammern der Ingenieure und Architekten für die Honorarordnung (AHO) (z. B. AHO 2014) dienen daher als Rahmenwerke zur Rahmenterminplanung als erste zeitliche Planung eines Bauprojektes, die in Deutschland Anwendung finden. In der ersten Forschungsfrage wird der Untersuchungsraum Deutschland gewählt. Diese Begrenzung wird aufgrund der international nicht einheitlichen Definition der ersten Terminplanung vorgenommen. Die Definitionen der Honorarordnung für Architekten und Ingenieure (HOAI) (z. B. HOAI 2018) und des Ausschusses der Verbände und Kammern der Ingenieure und Architekten für die Honorarordnung (AHO) (z. B. AHO 2014) dienen daher als Rahmenwerke zur Rahmenterminplanung als erste zeitliche Planung

eines Bauprojektes, die in Deutschland Anwendung finden. Für Forschungsfrage zwei und drei werden an beispielhaften Datensätzen Analysen durchgeführt. Folgende Eingrenzungen sind für diese Datenanalysen festzuhalten:

- Die Datenanalysen sind begrenzt auf die verfügbaren Datensätze. Dies beinhaltet die Größe, Detaillierung der Dauer und dokumentierte Merkmale im Datensatz.
- Es werden entsprechend dem verfügbaren Datensatz beispielhafte Phasen eines Rahmenterminplans untersucht.
- In den Analysen werden frei zugängliche Datensätze öffentlicher Bauprojekte betrachtet. Diese sind zum einen durch ihre freie Zugänglichkeit nachvollziehbar und zum anderen als repräsentativ angesehen. Sie fassen Projekte mit unterschiedlichen Merkmalen zusammen, die von unterschiedlichen Bauunternehmen durchgeführt wurden.
- Es werden nicht die Sichtweisen unterschiedlicher Beteiligter an den Ergebnissen der Datenanalysen untersucht, sondern lediglich allgemeine unternehmerische Vorteile durch optimierte Prognosen bewertet.

Der Untersuchungsraum der Forschungsfrage eins muss daher nicht zwingend den Datenanalysen der Forschungsfragen zwei und drei entsprechen. Vielmehr dienen die Ergebnisse der ersten Forschungsfrage als Vergleichsrahmen für die aktuelle Dokumentation von Bauprojekten in Deutschland, sodass andere und frei zugängliche Datensätze in diesen Rahmen eingeordnet werden können.

## **1.4 Vorgehensweise der Arbeit**

Die Vorgehensweise der Arbeit erfolgt durch wissenschaftliche Untersuchungsmethoden, um Forschungsfragen systematisch und strukturiert zu beantworten. Der Aufbau der Arbeit ist in Abbildung 1.3 schematisch dargestellt.

In Kapitel 2 werden relevante Grundlagen und Definitionen zum Projektmanagement, der Rahmenterminplanung und analytischen Planungsmethoden dargelegt. Strategien zur Reduktion von Abweichungen als Ursache der Unsicherheit und systematischen Planungsfehlschlüssen werden zusammengefasst. Diese Strategien bilden die Grundlage für die erste Forschungsfrage. Zuletzt werden Anforderungen an analytische Planungsmethoden zur Rahmenterminplanung identifiziert.

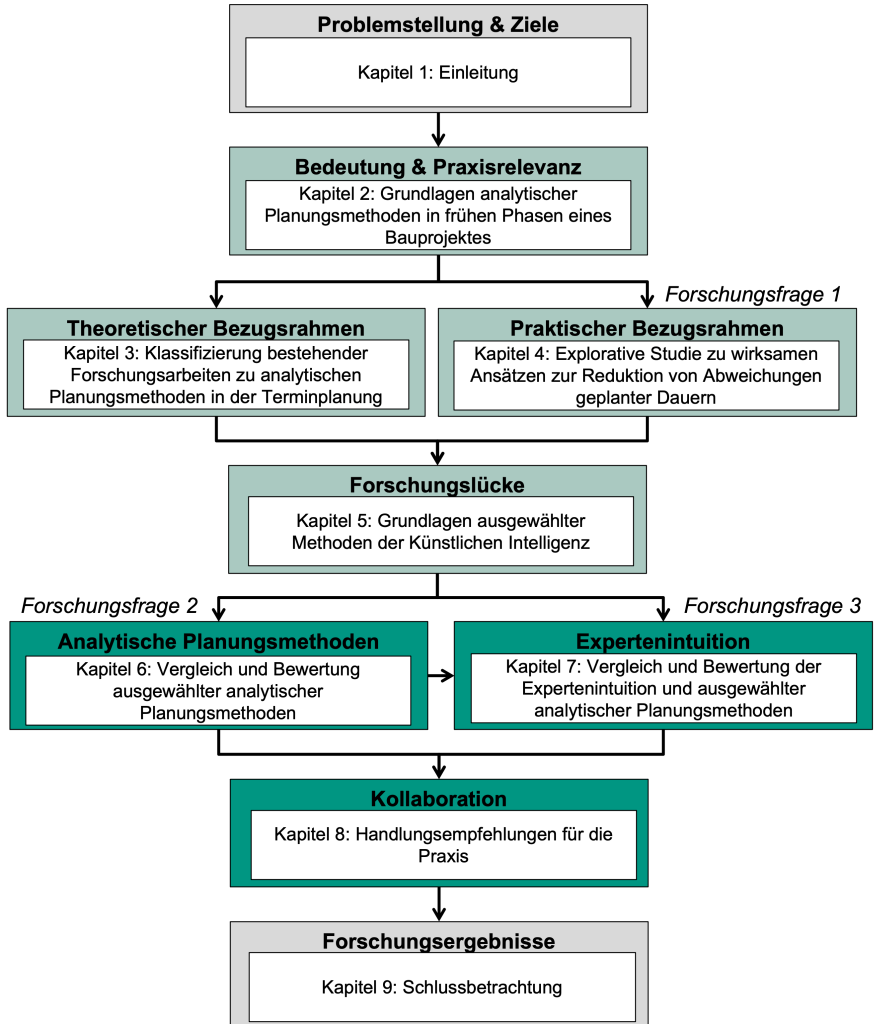


Abbildung 1.3: Aufbau der Arbeit

Mithilfe einer systematischen Literaturrecherche werden im nächsten Schritt in Kapitel 3 bestehende deutsch- und englischsprachige Untersuchungen über Planungsmethoden und relevante Merkmale in der Terminplanung von Bauprojekten erfasst. Die identifizierten Planungsmethoden werden entsprechend ihrem unternehmerischen Vorteil bewertet, Forschungslücken identifiziert und relevante Merkmale zur Identifikation von Schwerpunkten kategorisiert. Die Ergebnisse der Literaturrecherche sind eine relevante Grundlage zur Beantwortung der zweiten Forschungsfrage.

Zur Beantwortung der ersten Forschungsfrage wird in Kapitel 4 eine explorative Studie im Untersuchungsraum Deutschland durchgeführt. Hierfür wird auf Basis von Interviews eine explorative Faktoranalyse (EFA) nach Hinkin (1998) durchgeführt. Univariate Ergebnisse spiegeln den aktuellen Stand der Datenerfassung und die Verwendung von Methoden zur Rahmenterminplanung wider. Zur Untersuchung bivariater Ergebnisse werden strategische Treiber im Zusammenhang zur Planungsqualität gesetzt.

Auf den Ergebnissen aus Kapitel 3 und 4 aufbauend werden der Themenbereich der Künstlichen Intelligenz (KI) und insbesondere ausgewählte Methoden der erklärbaren KI (engl. „eXplainable Artificial Intelligence“ (XAI)) in Kapitel 5 erläutert. Als Teil inhärent erklärbarer Modelle werden Entscheidungsbäume des Maschinellen Lernens als relevant identifiziert und zur Analyse relevanter Merkmale wird das SHAP-Framework (SHapley Additive exPlanations) verwendet.

Die Beantwortung der zweiten Forschungsfrage findet in Kapitel 6 statt. Im ersten Teil werden mit einer Umfrage Zielwerte für zufriedenstellende Abweichungen der Terminplanung von Bauprojekten ermittelt. Darauf aufbauend wird die Prognosegenauigkeit verschiedener analytischer Planungsmethoden miteinander verglichen und bewertet. Abschließend lassen sich relevante Merkmale aus diesen Modellen extrahieren, die zur Datenerfassung relevant sind. Auszüge dieses Kapitels sind auch in der Veröffentlichung Lauble und Hagshsheno (2023) zu finden.

Die dritte Forschungsfrage wird in Kapitel 7 beantwortet. Die ausgewählten analytischen Planungsmethoden werden mit der Prognosegenauigkeit der Expertenintuition verglichen. Im Anschluss können auch hier relevante Merkmale identifiziert werden, die eine Bevorzugung der Prognose durch den Experten oder analytische Planungsmethoden aufzeigen.

Mit Kapitel 8 werden die Erkenntnisse der vorherigen Kapitel in konkrete Handlungsempfehlungen für die Praxis überführt. In Kapitel 9 sind die Ergebnisse zu den jeweiligen Forschungsfragen zusammengefasst und es wird ein Ausblick auf weiterführende Forschung gegeben.





## **2 Grundlagen analytischer Planungsmethoden in frühen Phasen eines Bauprojektes**

### **2.1 Vorbemerkung**

In diesem Kapitel werden relevante Grundlagen für diese Arbeit definiert und erläutert. Dies beinhaltet Begriffe zum Projekt und dessen Management (Kapitel 2.2) sowie zur Terminplanung in frühen Projektphasen (Kapitel 2.3). Mit der Rahmenterminplanung werden erstmals Dauern in einem Bauprojekt bestimmt. Diese prognostizierten Dauern bilden die Grundlage für die weitere Terminplanung. Darauf aufbauend werden der Begriff der analytischen Planungsmethoden (Kapitel 2.4) und die Auswirkungen von Abweichungen von Dauern erläutert (Kapitel 2.5). Die Ursache der Planabweichungen kann zum einen durch die Unsicherheit in frühen Phasen und zum anderen durch systematische Planungsfehl-schlüsse begründet sein. Beide Ursachen und mögliche Ansätze zu deren Reduktion werden erläutert. Diese Grundlagen resultieren in Anforderungen an die Planung von Dauern in frühen Projektphasen (Kapitel 2.6). Bestehende analytische Methoden können so bezüglich ihrer Anwendbarkeit im folgenden Kapitel 3 überprüft werden.

## 2.2 Begriffe des Projektmanagements in Bauprojekten

### Projekte

Die Dauern in frühen Phasen eines Bauprojektes werden im Rahmen des Projektmanagements prognostiziert (vgl. Kochendörfer et al. 2010, S. 8ff.).

Ein Projekt ist ein „Vorhaben, das im Wesentlichen durch die Einmaligkeit der Bedingungen in ihrer Gesamtheit gekennzeichnet ist, wie z. B. Zielvorgabe, zeitliche, finanzielle, personelle und andere Begrenzungen, projektspezifische Organisation.“ (DIN69901-5 2009, S. 11)

Das Projektmanagement Institut (PMI) hebt neben der Einmaligkeit der Bedingungen und Begrenzungen weiter den Mehrwert des Projektes hervor. Es definiert ein Projekt als „[...] ein zeitlich begrenztes Unternehmen, das unternommen wird, um ein einmaliges Produkt, eine Dienstleistung oder ein Ergebnis zu erzeugen.“ (PMBOK 2018, S. 4)

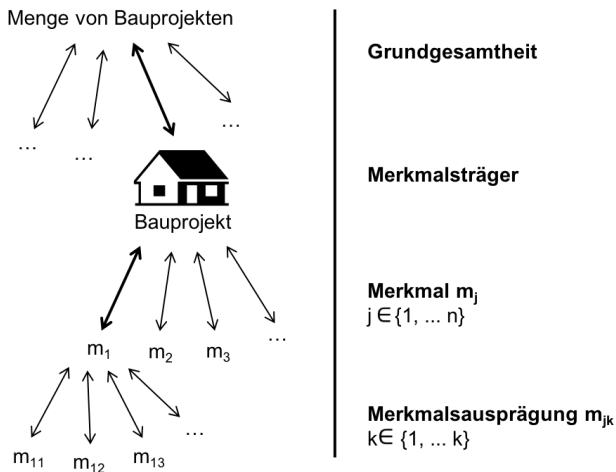
Die Dauern einzelner Phasen können eine Bedingung an das Projekt sein oder sich aufgrund der Projektmerkmale und -attribute ergeben.

### Projektmerkmale und -attribute

Das zu erstellende Produkt, das Bauwerk, sowie seine Rahmenbedingungen beschreiben ein Bauprojekt und können daher als Projektmerkmale bezeichnet werden. Beschreibende Merkmale eines Projektes können u. a. der Projektumfang, die Projektgröße, Merkmale zur Komplexität oder die Anzahl der beteiligten Personen sein (vgl. DIN69901-5 2009, S. 0). Die Ausprägungen der Merkmale werden als Attribute der Merkmale bezeichnet.

Dabei stehen nach Hall (1956) Inhalt und Bedingungen in direkter Abhängigkeit. Dies resultiert in einem System aus Projektmerkmalen mit gegenseitigen Wechselbeziehungen.

Ein einzelnes Bauprojekt kann schließlich als ein Merkmalsträger aus einer Grundgesamtheit vieler Bauprojekte bezeichnet werden (siehe Abbildung 2.1).



**Abbildung 2.1:** Zusammenhang zwischen Bauprojekten und seinen Projektmerkmalen (vgl. Mittag und Schüller 2020, S. 23)

Die Beschreibung dieser Projektmerkmale und seiner Attribute kann mittels verschiedener Skalen stattfinden (vgl. Mittag und Schüller 2020, S. 22ff.):

- Nominal-Skala: Kategorien, Zustände oder „Namen von Dingen“ (z. B. der Gebäudetyp)
- Binär-Skala: Nominale Merkmale mit genau zwei Zuständen (z. B. Neubau „ja“ oder „nein“ bzw. 1 oder 0)
- Ordinal-Skala: Werte, die in einer Reihenfolge zueinander stehen, ohne dass die Abstände eine Bedeutung haben (z. B. der Ampelstatus eines Projektes „grün“, „gelb“, „rot“ bzw. 1,2,3)

- Ratio-Skala: Bewertbar zum Nullpunkt durch ihre Reihenfolge und Abstände (z. B. Bruttogrundfläche in  $m^2$ , Investition in €, benötigte Tage zur Bauausführung)
- Intervall-Skalen (z. B. die Verortung des Bauprojektes liegt innerhalb der Postleitzahlen 80000-90000).

Die Gesamtheit dieser Merkmale und ihrer Attribute macht ein Vorhaben einmalig und damit zu einem Projekt. Wiederkehrende Zusammenhänge zwischen den einzelnen Attributen der Merkmale können jedoch mit einer entsprechend großen Datenmenge aufgezeigt werden. Weiter dienen einzelne Merkmale während der Projektdurchführung als Messgrößen zum geplanten Projekterfolg. Messgrößen zum Projekterfolg können u. a. die Einhaltung des zeitlichen und finanziellen Rahmens sowie Sicherheit- und Gesundheitsaspekte sein. Durch das Aufdecken von Zusammenhängen zwischen einzelnen Merkmalen und ihren Ausprägungen lassen sich die Messgrößen gezielt beeinflussen. Es ist jedoch zu erwähnen, dass einige Projektmerkmale nur subjektiv und daher schwer einheitlich zu erfassen sind (z. B. Vertrauen oder Kommunikation).

### **Projektmanagement**

Das Projektmanagement ist verantwortlich für die Erfüllung der Projektziele und Anforderungen. Ein wichtiger Teil davon sind die Merkmale und -ausprägungen des Projekts, die die im Einklang mit den spezifischen Projektanforderungen stehen müssen.

Die DIN69901-5 (2009) definiert Projektmanagement als „die Gesamtheit der Managementaufgaben, -organisation, -techniken und -mittel zur Initiierung, Definition, Planung, Kontrolle und zum Abschluss von Projekten.“ Das Projektmanagement wendet Methoden an, um das abstrahierte Bauwerk als Modell zu simulieren.

Methoden stellen eine „bestimmte standardisierte Vorgehensweise, die einem Gestaltungsprinzip zugeordnet ist und zur Erreichung von Unternehmenszielen eingesetzt wird“ (Verein Deutscher Ingenieure e.V. 2012, S. 6) dar. Der Duden

(2022a) beschreibt eine Methode weiter als ein „auf einem Regelsystem aufbauendes Verfahren zur Erlangung von Erkenntnissen oder praktischen Ergebnissen.“

Modelle hingegen beschreiben nicht die Vorgehensweise, sondern die vereinfachte Abbildung des zu planenden Produkts. Als Modell wird „eine vereinfachte Nachbildung eines existierenden oder gedachten Systems mit seinen Prozessen in einem anderen begrifflichen oder gegenständlichen System“ verstanden, das „sich hinsichtlich der untersuchungsrelevanten Eigenschaften nur innerhalb eines vom Untersuchungsziel abhängigen Toleranzrahmens vom Vorbild unterscheidet“ (Verein Deutscher Ingenieure e.V. 2014). Der Duden (2022b) beschreibt ein Modell als eine „[...] veranschaulichende Ausführung eines vorhandenen oder noch zu schaffenden Gegenstandes [...]“. Mit diesem Modell sind im Anschluss Simulationen möglich, die die Planung und Steuerung des Projektes unterstützen.

## **2.3 Terminplanung in frühen Projektphasen**

### **2.3.1 Terminplanung zur Projektstrukturierung**

Um die Messgrößen innerhalb eines Projektes besser planen und steuern zu können, wird das Projekt häufig in verschiedene Detaillierungsebenen zerlegt. Die Zerlegung des Projektes ist auch als Projektstrukturplan bekannt. „Beim [...] Projektstrukturplan [...] erfolgt die hierarchische Zerlegung der Projektaufgabe in Teilaufgaben und die Beschreibung der Beziehung zwischen den Teilaufgaben.“ (Schulte-Zurhausen 1995, S. 373) Ergebnis des Projektstrukturplans sind verschiedene Terminpläne eines Bauprojektes in einer unterschiedlichen Detaillierung. Detaillierende Ebenen unter einem Projekt sind Phasen und weiter Vorgänge. Zwischen Start- und Endzeitpunkt eines Projektes können verschiedene Phasen und Meilensteine verortet werden.

Eine Phase ist ein „zeitlicher Abschnitt des Projektablaufs, der sachlich gegenüber anderen Abschnitten abgegrenzt ist“ (Motzel und Möller 2017, S. 173).

Ein Vorgang ist die zeitliche Durchführung einer Aufgabe. Es ist „ein Ablaufelement, das ein Geschehen, eine Handlung, eine Tätigkeit oder einen Arbeitsschritt von bestimmter Dauer beschreibt und dessen Anfang und Ende definiert sind“ (Motzel und Möller 2017, S. 282).

Weiter kann die Sequenz mehrerer Vorgänge als Prozess definiert werden. Der Prozess ist demnach eine „Abfolge von zusammengehörigen und / oder zusammen betrachteten, in Beziehung und / oder Wechselwirkung stehenden [Vorgänge] [...], die einen zeitlichen Beginn und ein Ende haben und durch die Eingaben in Ergebnisse umgewandelt werden“ (Motzel und Möller 2017, S. 236).

Ein Meilenstein hingegen ist ein „Ereignis von besonderer Bedeutung in einem größeren Ganzen“. Er zeigt neben Start- und Endzeitpunkt des Projektes für den Projekterfolg relevante Zeitpunkte auf (Motzel und Möller 2017, S. 156). Die definierten Meilensteine werden häufig als Ausgangsbasis zur Bewertung von Abweichungen verwendet.

Es existiert keine standardisierte Definition der Terminpläne auf den unterschiedlichen Detaillierungsebenen. Im Folgenden sind die Definitionen von Gralla (2011), Greiner et al. (2005) und Ahrens et al. (2004) vorgestellt.

Gralla (2011) unterscheidet fünf Detaillierungsebenen. Diese teilen sich gemäß der erstellten Pläne auf in: Rahmenterminplan (Ebene eins), Generalterminpläne (Ebene zwei), Grobterminpläne (Ebene drei), Steuerungsterminpläne und Koordinationsterminplan (Ebene vier) und Feinterminpläne (Ebene fünf). Der Detaillierungsgrad ist in Ebene eins am niedrigsten und steigt an, bis in Ebene fünf der höchste Detaillierungsgrad erreicht ist. (vgl. Gralla 2011, S. 194ff.)

Greiner et al. (2005) unterscheiden den Rahmenterminplan, Generalterminplan, Grobterminplan und den Detailterminplan (vgl. Greiner et al. 2005, S. 120).

In der Ablaufplanung nach Ahrens et al. (2004) existieren drei Planungsebenen: Unterschieden werden der Gesamtrealisierungsplan (Ebene eins), der Rahmenterminplan (Ebene zwei) sowie der Detailablaufterminplan (Ebene drei). Der Detaillierungsgrad steigt dabei von Ebene eins bis drei an.

Mit dem Vergleich der drei Darstellungen zeigt sich eine Projektstrukturierung in unterschiedlichen Detaillierungsebenen ohne eine einheitliche Definition. Bei dem Vergleich der drei Referenzen zeigt sich, dass für die Rahmenterminplanung als Grundlage für das weitere Projekt kein einheitliches Verständnis bezüglich der Detaillierung vorliegt. Aus diesem Grund werden im Folgenden Definitionen bezüglich der Rahmenterminplanung miteinander verglichen.

### 2.3.2 Die Rahmenterminplanung

Die erste Terminplanung findet auf der obersten Ebene der Projektstrukturierung statt. Dabei meint die Planung die „[...] systematische Erarbeitung eines erfolgsversprechenden Weges zu einem vorgegebenen Ziel unter Berücksichtigung gegebener Bedingungen [...]“ (Motzel und Möller 2017, S. 175f.). Zu den Bedingungen zählen die grundlegenden Zielgrößen des Projektmanagements: Kosten, Zeit und Qualität. Die Anpassung einer dieser drei Grundziele geschieht häufig zulasten einer anderen Zielgröße. Hohe Qualitätsansprüche sind bspw. mit einem entsprechenden Kostenzuwachs verbunden. Somit gilt es, ein Gleichgewicht von Terminen, Kosten und Qualität zu schaffen. Dies veranschaulicht die Bedeutung einer optimierten Terminierung von Projektzeiten. (vgl. Jakoby 2019, S.86)

Die erste Projektstrukturierung bzw. Terminplanung findet in Deutschland in der Regel mit der Rahmenterminplanung statt (vgl. Kessel et al. 2015, S. 253) und stellt daher im Folgenden die zeitliche Planung in frühen Projektphasen dar. Sie wird mit folgenden Definitionen erläutert, um ein einheitliches Verständnis zu schaffen.

In der Honorarordnung für Architekten und Ingenieure (HOAI 2018) wird folgende Definition der Terminplanung im Rahmen der Vorplanung, der Leistungsphase 2, gegeben. Diese ist, wenn auch nicht explizit als solche bezeichnet, mit der Rahmenterminplanung gleichzusetzen:

„Erstellen eines Terminplans mit den wesentlichen Vorgängen des Planungs- und Bauablaufs. Dabei sind etwaige Terminvorstellungen

des [...] [Bauherren] hinsichtlich ihrer Realisierbarkeit zu überprüfen und müssen ggf. korrigiert werden.“ (HOAI 2018, S. 97)

Der Ausschuss der Verbände und Kammern der Ingenieure und Architekten für die Honorarordnung (AHO) definiert in Heft 9 der AHO die Rahmenterminplanung wie folgt und zeigt dabei die hohe Granularität auf (vgl. zur Definition einer Phase nach Motzel und Möller (2017, S. 173)):

„Die Rahmenterminplanung dient der generellen Projektplanung. Sie umfasst relativ wenige, aber übergeordnete Vorgänge [...]. Die Zeitdimension sind Monate oder Quartale.“ (AHO 2014, S. 200)

Greiner et al. (2005) bezeichnen den Rahmenterminplan ebenfalls als eine Beschreibung der Inhalte auf Monats- oder Quartalsebene. Bei Ahrens et al. (2004) umfasst der Rahmenterminplan hingegen unabhängig von der Dauer etwa 30 bis 40 Vorgänge, die je nach Komplexität je Teilprojekt angegeben werden. Gralla (2011) hingegen nennt 15 bis 30 Vorgänge je Projekt. Nach Greiner et al. (2005) und Ahrens et al. (2004) wird der Rahmenterminplan zu Projektbeginn erstellt, nach Gralla (2011) spätestens in der Grundlagenermittlung.

Diederichs et al. (2020) nennen für einen Rahmenterminplan pro Projekt bzw. Teilprojekt und für den Hochbau mindestens folgende Meilensteine, die beispielhaft zu sehen sind:

- Beginn der Projektentwicklung,
- Entscheidung zum Planungsauftrag,
- Baueingabe,
- Baugenehmigung,
- Baubeginn,
- Fertigstellung des Rohbaus und der wetterfesten Gebäudehülle,
- Fertigstellung der technischen Gebäudeausrüstung,



- Fertigstellung des Innenausbaus,
- Baufertigstellung,
- Beginn der Abnahme-, Inbetriebnahme- und Übergabephase,
- Beginn der Nutzungsphase.

Eine dem Rahmenterminplan vergleichbare Planungsdarstellung ist ebenfalls in internationalen Standards der Normierung und Empfehlungen zu finden. Im Folgenden werden beispielhaft die britische Norm, die schweizerische Norm und die internationale Empfehlung des „Association for the Advancement of Cost Engineering“ (AACE) genannt:

In Großbritannien wird der Rahmenterminplan in der Phase 0, der strategischen Definition, im Phasenplan des „Royal Institute of British Architects RIBA“ (RIBA) erläutert. Ziel ist es hier, das Geschäftsmodell des Kunden, strategische Analysen sowie grundlegende Bedingungen zu analysieren. Der Rahmenterminplan (engl. „Project Program“) ist das Ergebnis. Weiter wird vorgeschlagen, bereits durchgeführte Bauprojekte zu analysieren, um Erfahrungen in der Rahmenterminplanung aufnehmen zu können. (vgl. RIBA 2013)

Auch in der schweizerischen Norm SIA (Schweizer Ingenieur und Architektenverein) zur Definition der Grundleistungen von Architekten und Ingenieuren, ist die Rahmenterminplanung in der strategischen Planung zur Bedürfnisformulierung und Ermittlung von Lösungsstrategien zu finden. Der Gesamtleiter des Bauprojektes hat die Aufgaben, übergeordnete Ziele, Rahmenbedingungen und Anforderungen des Bauherrn in einem Bericht zusammenzufassen. Ergebnis ist hieraus ein „Entscheidungsorientierter Ablauf- und Terminplan“ mit einer Bewertung der Grobabläufe für verschiedene Lösungsstrategien. (vgl. SIA 2014, S. 20)

In der „AACE International Recommended Practice No. 37R-06“ werden fünf Ebenen der Terminplanung beschrieben. Dabei wird in Ebene 0 das gesamte Projekt erstmals durch eine Dauer von Start bis Ende sowie durch seine wesentlichen projekt-beschreibenden Meilensteine (z. B. Start und Ende der Bauausführung)

beschrieben (vgl. Douglas 2010). Diese Ebene entspricht der Rahmenterminplanung

HOAI, RIBA, SIA und die Empfehlung der AACE haben gemeinsam, dass die einmaligen Bedingungen des Vorhabens aufgenommen werden und Lösungsalternativen hinsichtlich deren Einhaltung analysiert werden.

Zusammengefasst wird eine Rahmenterminplanung nach unterschiedlichen Standards und Empfehlungen als erste strategische Planung des Bauprojekts durchgeführt (vgl. Kessel et al. 2015, S. 253) und entspricht der größten Detaillierungsebene in der Terminplanung. Mit der Rahmenterminplanung werden spezifische Bedingungen sowie der zeitliche Rahmen definiert (siehe auch Abschnitt 2.2). Aufgrund der unterschiedlichen Verwendung des Begriffs bei verschiedenen Standards, Normierungen und Autoren wird daher angenommen, dass der Rahmenterminplan zu Projektbeginn und spätestens in der Machbarkeitsstudie erstellt wird und maximal 40 Vorgänge auf Monats- oder Quartalsebene umfasst.

Die Rahmenterminplanung dient damit als Grundlage für:

1. Die Überprüfung der Realisierbarkeit des Bauprojektes (vgl. HOAI 2018, S. 97),
2. Die Bewertung von Entscheidungen anhand der gesetzten Ziele, z. B. Investitionsentscheidungen (Greiner et al. vgl. 2009, S. 121, Gralla 2011, S. 197),
3. Die Ressourcenplanung (Greiner et al. vgl. 2009, S. 121, Gralla 2011, S. 197),
4. Das Schaffen eines einheitlichen Verständnisses zwischen Bauherr und dessen Projektpartnern durch die Visualisierung (vgl. Ahrens et al. 2004, S. 369),
5. Die Abschätzung des Gesamtumfangs (vgl. Ahrens et al. 2004, S. 369),
6. Die Grundlage für Verträge und die Koordination mit Partnern,
7. Die Kontrolle der gesetzten Ziele / Definition des Projekterfolgs.

### 2.3.3 Zuständigkeiten in der Rahmenterminplanung

Die nachfolgend beschriebenen Rollen sind für die Rahmenterminplanung zuständig.

**Bauherr:** Grundsätzlich ist die übergeordnete Terminplanung für das Gesamtprojekt Aufgabe des Bauherrn (vgl. Würfele et al. 2017, S. 25). Bauherr ist derjenige, der selbst oder durch Dritte ein Bauvorhaben für eigene oder für fremde Rechnung erstellt oder erstellen lässt. Die oberste Aufgabe des Bauherrn ist es, die Projektziele zu formulieren und das Projekt zu finanzieren. Häufig fällt die Verantwortung der Rahmenterminplanung in seinen Bereich. Dieser Bereich wird jedoch oft an eine Projektleitung abgegeben und als „besondere Leistung“ verstanden (Diederichs 2006, S. 217-218, Kochendörfer et al. 2010, S. 56-62). Entscheidet sich der Bauherr gegen einen externen Projektleiter, fallen auf ihn die Aufgaben des Projektleiters zurück.

**Projektleitung:** Der Projektleiter ist eine Funktion im Rahmen des Projektmanagements. Er hat die Entscheidungs- und Durchsetzungskompetenz für den Fall, dass ihm die Verantwortung übertragen wird. Nachdem der Bauherr seine Ziele formuliert hat, obliegt es dem Projektleiter, deren Erfüllung möglich zu machen. Für die Erstellung des Rahmenterminplans ist er hauptverantwortlich. (vgl. Kochendörfer et al. 2010, S. 8ff.)

**Totalunternehmer / Generalübernehmer:** Wenn gleich zu Beginn des Projekts ein Generalübernehmer oder ein Totalunternehmer engagiert wird, übernimmt dieser einen großen Teil der Aufgaben, die ansonsten der Projektleiter innehat bzw. vergibt diese Aufgabe weiter (vgl. Greiner et al. 2009, S. 38). Wird ein solcher aber erst in einer späteren Phase der Planung herangezogen, ist er an der Rahmenterminplanung noch nicht beteiligt.

**Baubetreuer / Bauträger:** In Deutschland kann ein Baubetreuer für und auf Rechnung des Bauherrn Aufgaben bei der Vorbereitung oder Durchführung eines Bauvorhabens übernehmen. Dies kann in wirtschaftlicher oder technischer Hinsicht (Vollbetreuung) oder anteilig (Teilbetreuung) erfolgen.

Folglich kann die Rahmenterminplanung beim Bauherren liegen, aber auch durch andere Rollen, wie der Projektleitung, einem Generalübernehmer oder Totalunternehmer, Baubetreuer oder Bauträger durchgeführt werden. Lagert der Bauherr diese Aufgabe aus, muss zur Überprüfung der Ziele die Rahmenterminplanung möglichst nachvollziehbar sein. Wird diese durch den Bauherrn selbst durchgeführt, sollte eine entsprechend große Expertise vorliegen.

### **2.3.4 Verständnis der prognostizierten Dauern**

Die Darstellung der prognostizierten Dauern kann unterschiedlich verstanden werden (vgl. Cleden 2017):

- Sie kann Grundlage für einen Basisterminplan sein, der vertraglich bindend ist.
- Die Dauer kann als Nominalwert verstanden werden, der erreicht wird, wenn die Prognose korrekt ist.
- Sie kann das repräsentative Ergebnis des ausgewählten Szenarios sein. Demnach ist es der Plan, der am wahrscheinlichsten eintritt. Dieser ist kontinuierlich zu aktualisieren.

In einem Projekt sind die prognostizierten Dauern als Basisterminplan, als Nominalwert und als repräsentatives Ergebnis aufzufassen. Das repräsentative Ergebnis ist als Arbeitsmedium eines Planers zu verstehen. Ohne vertraglich geregelte Termine und einen Nominalwert existiert jedoch kein Ziel, das anzustreben ist. Umgekehrt zeigt alleine der Nominalwert, dass das Ergebnis nicht unbedingt realistisch sein muss.

Die Phasen eines Bauprojektes werden zuerst in einem Rahmenterminplan als ersten Ablaufplan festgehalten. Das Ergebnis der Planung ist vertraglich bindend (Basisterminplan). Der Weg hin zur vertraglichen Festlegung von Phasen und zugehörigen Meilensteinen kann aber mit Nominalwerten und wahrscheinlichen Szenarien ergänzt werden.

Als Vorgehensweise zur Rahmenterminplanung existieren allgemein zwei Möglichkeiten.

1. Entweder kennt der Bauherr bzw. seine Vertretung die zeitlichen Ziele des Vorhabens. Dementsprechend muss überprüft werden, welche Projektanforderungen in dem definierten Zeitrahmen zu realisieren sind.
2. Oder es sind bereits Projektanforderungen definiert, zu denen ein zeitlicher Rahmen bestimmt werden muss.

Häufig werden diese Vorgehensweisen in einem iterativen Prozess kombiniert, um die Planung zu erarbeiten. Bei beiden Vorgehensweisen können allgemein entweder die Expertenintuition und / oder analytische Planungsmethoden verwendet werden. Beide Vorgehensweisen sind im Folgenden beschrieben.

## **2.4 Analytische Planungsmethoden**

### **2.4.1 Expertenintuition und Analytik**

Aufgrund der Unsicherheit zum Zeitpunkt der Rahmenterminplanung werden die Begriffe Intuition und Analytik, insbesondere die datengetriebene Analytik, abgegrenzt sowie deren situationsbedingte Treiber zusammengefasst. Die Entscheidungsfindung kann entweder in intuitive und unbewusste Denkprozesse oder in analytische, rationale und bewusste Beziehungsüberlegungen eingeteilt werden (vgl. Kahneman 2003). Damit unterscheidet sich die Analytik durch ein wissenschaftsorientiertes Vorgehen und kann auch als Prognose verstanden werden (vgl. Duden 2022c). Die beiden Entscheidungsfindungsprozesse werden im Folgenden näher erläutert.

#### **Expertenintuition**

Braun und Benz (2015, S. 40) fassen die Intuition wie folgt zusammen: „So wird allgemein angenommen, dass die Genese von Intuition auf nicht-bewussten, d. h.

willentlich nicht steuerbaren Informationsverarbeitungsprozessen basiert, wobei davon ausgegangen wird, dass die im Rahmen derartiger Prozesse verarbeiteten Informationen der Erfahrungs- bzw. Wissensbasis des jeweiligen Entscheiders entstammen.“ Scherm et al. (2016, S. 306) bezeichnen die Intuition weiter als „Fähigkeit zum Umgang mit Mehrdeutigkeit, die reproduktive ebenso wie produktive Elemente enthält“.

Gigerenzer (2015) plädiert für den Einsatz von intuitiven Entscheidungen, die auf Basis von „Faustregeln“ und dem Bauchgefühl gemacht werden. Huang (2019) spezifiziert die Anwendung von Intuition auf Situationen mit einer sehr hohen Unsicherheit. „Bei der Projektschätzung kann eine intuitive Schätzung mit minimalem Aufwand eine grobe Vorstellung des zu erwartenden Aufwands in einer sehr frühen Planungsphase liefern“ (Jakoby 2019, S. 176). Hier ist eine intuitive Entscheidung besser als eine aufgeschobene oder nicht stattfindende Entscheidung.

Dennoch unterliegen intuitive Entscheidungen der Gefahr von situativen Verzerrungen (engl. „Bias“). Diese sollten während der Entscheidungsfindung wahrgenommen und einbezogen werden. (vgl. Scherm et al. 2016, S. 303)

So führt der „Self Serving Bias“ dazu, dass persönliche Eigenschaften infolge vergangener Erfolge zu hoch eingeschätzt werden. Der „Confirmation Trap“ beschreibt die Verzerrung, bei der Informationen, die nicht mit den eigenen Überzeugungen übereinstimmen, aus der Entscheidung ausgeschlossen werden. Zuletzt können nach dem „Overconfidence Effect“ Personen nicht einschätzen, ob die Erfahrungen auf einer kleinen Stichprobe mit unzuverlässigen Daten beruht. Sie vertrauen tendenziell einer kleinen Stichprobe mit einstimmigen Daten mehr als einer größeren Stichprobe mit nicht einstimmigen Daten. Das kann zu einer fehlerhaften Erkennung von Mustern führen. (vgl. Kahneman und Tversky 1977)

Aus diesem Grund vertreten u. a. die Autoren Kahneman (2003, S. 1449ff.), Bonabeau (2003), Davenport und Harris (2007, S. 13ff.) und Huang (2019) die Meinung, dass Intuition lediglich gerechtfertigt ist, wenn nicht genügend Informationen verfügbar sind und eine schnelle Entscheidung getroffen werden muss.

Sobald weitere Informationen vorliegen, sollte die angepasste, verbesserte Lösung unter Berücksichtigung der zusätzlichen Informationen getroffen werden. Menschen nutzen die Intuition, da sie nicht über die Fähigkeiten verfügen, alle Informationen aufzunehmen, zu speichern und optimal zu verwerten. Rechnerische Kapazitäten können dies jedoch ausführen und vor allem in komplexen Situationen bei der objektiven Entscheidungsfindung unterstützen. Dieses Forschungsgebiet wird im Weiteren unter der datenbasierte Analytik zusammengefasst.

### **Datenbasierte Analytik**

Die Einbindung analytischer, datenbasierter Methoden in der Praxis findet nur vereinzelt Anwendung (vgl. Erdis 2013). Die datenbasierte Analytik gehört zum Gebiet „Data Science“, das vor allem Verfahren aus den Bereichen Statistik und Künstliche Intelligenz nutzt (vgl. Langs und Wazir 2019, S. 141). Nach Davenport und Harris (2007, S. 7) umfasst die Analytik folgende vier Stufen: den Zugriff auf Daten, deren statistische Analysen, darauf aufbauende Hochrechnungen und voraussagende Modelle sowie Modelle zur Optimierung. Dabei sind die genannten Stufen aufeinander aufbauend und der unternehmensinterne Mehrwert bzw. Wettbewerbsvorteil zu anderen Unternehmen nimmt je Stufe zu. Bei allen Methoden wird abgestrebt, Zusammenhänge in den vorliegenden Daten zu identifizieren und ein objektives Ergebnis zu erzielen. Ein datenbasiertes Analysemodell liefert objektive Ergebnisse, die durch das Anwenden von Regeln frei von Verzerrungen, reproduzierbar und bei Abweichungen wieder anpassbar sind (vgl. Bari et al. 2014, S. 10ff., Scherm et al. 2016, S. 308). Durch den Einsatz datenbasierter Analytik können weiter unbekannte und auch schwache Zusammenhänge in Daten erfolgreich aufgedeckt werden.

Es kann in eine menschengestützte und in eine computergestützte Analytik unterschieden werden (vgl. Davenport und Harris 2007, S. 7). Die computergestützte datenbasierte Analytik eignet sich vor allem für große Datensätze. Denn es ist für Menschen schwierig, große Datenmengen zuverlässig zu bearbeiten (vgl. Bari et al. 2014, S. 51). Hier können Zeit- und Personalressourcen gespart werden. Gleichzeitig werden Fehler reduziert, die durch den Menschen erzeugt werden

(vgl. Bari et al. 2014, S. 12). So können datenbasierte Modelle vor allem auch bei geringeren Erfahrungswerten, wie bei Berufseinsteigern, im entsprechenden Themengebiet unterstützen. Es werden nicht mehr allein die gesammelten Erfahrungen und Reputation von Experten benötigt, um einen Wert aus Daten zu generieren. (vgl. Bari et al. 2014, Pfeffer und Sutton 2006)

Im Gegensatz zu den genannten Vorteilen ist ein gewisser Mehraufwand für das Datenmanagement erforderlich (vgl. Niebler und Lindner 2019, S. 6f.). Daten müssen aufbereitet, strukturiert und aktualisiert werden. Je genauer die Datenbasis ist, desto genauer ist das Ergebnis der analytischen Entscheidung auf Basis dieser Datengrundlage (vgl. Bari et al. 2014, S. 16). Aus diesem Grund ist auch bei analytischen Entscheidungen der Mensch mit seinem Erfahrungsschatz ein relevanter Faktor, der Ergebnisse nachvollziehen und unter den aktuellen Randbedingungen richtig einordnen kann (vgl. Davenport und Harris 2007, S. 14).

### **Situationsbedingte Treiber für Expertenintuition und Analytik**

Aus den wissenschaftlichen Grundlagen zur Intuition und datenbasierten Analytik, lassen sich eine Vielzahl potenzieller Treiber aufnehmen. So kann u. a. die Erfahrungstiefe des Einzelnen, die Planung im Team oder in Einzelarbeit, die als relevant identifizierten Merkmale des zu planenden Projektes, das Datenmanagement historischer Projekte (Anzahl dokumentierte Projekte und ihrer Attribute), die verwendeten Methoden zur Prognose der Phasen und ihrer Dauer sowie Merkmale der Organisation und Person ausschlaggebend sein.

Angesichts der Stärke datenbasierter Analysen bei Vorliegen zuverlässiger und ausreichender Daten ist es wichtig, auch die Bedingungen zu beschreiben, unter denen intuitive Entscheidungen zu zufriedenstellenden Ergebnissen führen können. Aus den Arbeiten von Kahneman und Klein (2009, S. 515ff.) sowie Agor (1986, S. 49ff.) lässt sich zusammenfassen, dass intuitive Entscheidungen:

- nur von Experten praktiziert werden dürfen,
- unter stabilen Umweltbedingungen bei der Vorhersage von Routineaufgaben zuverlässige Ergebnisse liefern können und



- in komplexen und unklaren Entscheidungssituationen mit begrenzten oder widersprüchlichen Daten ausreichend Erfahrung zur Vorhersage erfordern.

Es ist es für intuitive Vorhersagen wichtig, echte Experten zu haben, die die gegebene Umgebung lange studiert haben und gleichzeitig ein Talent zum Erkennen und Anwenden von Mustern haben (vgl. Kahneman und Klein 2009, S. 522f.).

## 2.4.2 Wettbewerbsvorteil durch analytische Planungsmethoden

Ist es möglich, mit analytischen Planungsmethoden Wissen zu generieren, kann sich hieraus ein Wettbewerbsvorteil ergeben. Über die Wissenstreppe nach North (2016, S. 6) erfolgt eine Unterscheidung der Begriffe Zeichen, Daten, Information, Wissen und Kompetenz (siehe Abbildung 2.2).

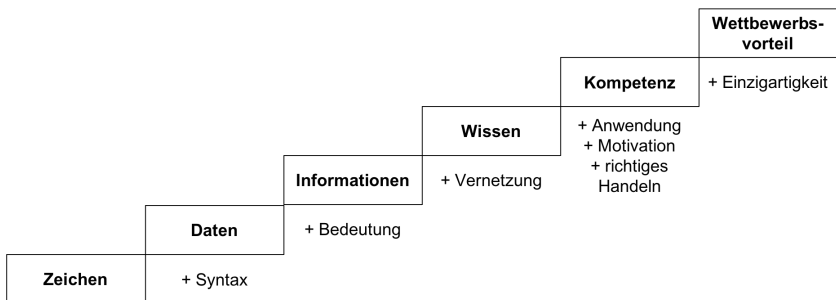


Abbildung 2.2: Wissenstreppe nach North (2016, S. 23)

Zeichen sind Buchstaben, Zahlen oder Sonderzeichen, wie z. B. „45“. Durch eine Syntax oder Regeln werden diese Zeichen zu Daten (vgl. North 2016, S. 5). Eine Zahlenfolge, wie „45, 33, 50“ bildet einen Datensatz. Durch das Hinzufügen einer Bedeutungsebene werden diese Daten zu Informationen, z. B. durch das Anhängen der Einheit. Die sich ergebende Information wäre dann: „45, 33, 50“ [Tage]. Erst durch den Menschen, der die Informationen verarbeitet und mit

Erfahrungen verknüpft, wird daraus Wissen (vgl. Probst et al. 2010, S. 6). Wissen ist somit auch kulturellen Interpretationen unterworfen. Datenbanken stellen damit einen Teilbereich des Wissens in Form von Informationen zur Verfügung (vgl. North 2016, S. 37). Der Mensch wendet dieses Wissen an. Hat das Unternehmen hierdurch ein Alleinstellungsmerkmal, stellt diese Kompetenz einen Wettbewerbsvorteil dar (vgl. North 2016, S. 36). So erkannte bereits 1597 der englische Philosoph Francis Bacon „Wissen ist Macht“ (García 2001, S. 1).

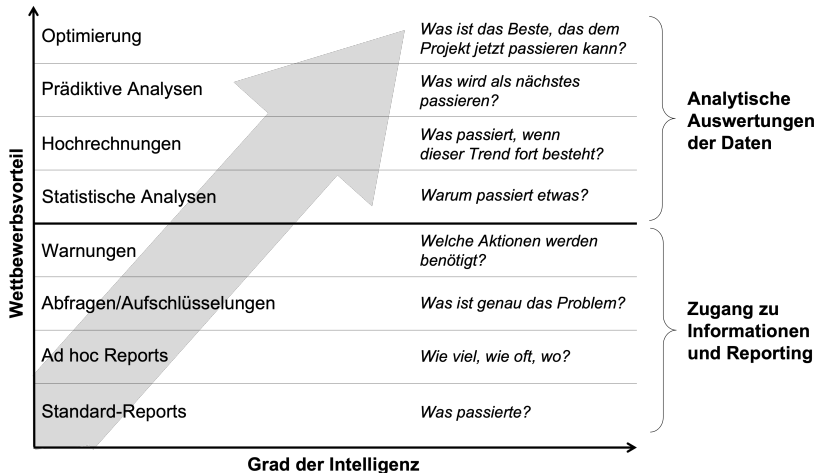
Gemäß einer Umfrage von Autodesk zu Datenvorteilen mit 3.916 Teilnehmern aus der Bauwirtschaft hat sich das Volumen von Informationen in den letzten drei Jahren bezogen auf das Jahr 2020 verdoppelt (vgl. Thomas und Bowman 2021). Dies bildet die Grundlage eines verbesserten Wissensmanagements. Mit der zunehmenden Informationsmenge wird jedoch auch die richtige Dokumentation, Interpretation und Anwendung der Daten (auch Datenstrategie genannt) immer wichtiger.

Oft scheidet die Umsetzung eines effizienten Wissensmanagements in der Praxis an der Organisation und Systematik, obwohl Mitarbeiter offen für die Anwendung sind (vgl. Ribeiro und Ferreira 2010, S. 369, Forcada et al. 2013, S. 88, Zuofa et al. 2015, S. 94). Nach Bullinger et al. (1998, S. 23) kann dies in der Personalfuktuation, in inkonsistenten Daten sowie mangelhaften Kommunikationsflüssen und fehlenden Mechanismen zum Wissensaustausch mit entsprechenden Schnittstellen begründet sein.

### **2.4.3 Stufen analytischer Planungsmethoden**

Davenport und Harris (2007, S. 8) beziehen sich in ihrem Werk „Competing on Analytics - The New Science of Winning“ auf die Vernetzung von Informationen (vgl. siehe Wissenstreppe, North 2016, S. 23) und unterscheiden hier verschiedene analytische Methoden. Demnach wird über den Zugang zu Informationen, ihrem Reporting und durch analytische Auswertungen ein Wettbewerbsvorteil generiert. Abbildung 2.3 zeigt insgesamt acht aufeinander aufbauende Stufen der

Datenintegration mit zunehmender intelligenter Nutzung und einem resultierenden Wettbewerbsvorteil. Je höher die intelligente Datennutzung, desto höher sei der Wettbewerbsvorteil.



**Abbildung 2.3:** Analytische Methoden und Wettbewerbsvorteil  
Davenport und Harris (vgl. 2007, S. 8)

Die acht Stufen reichen von einer einfachen historischen Abfrage von Einflussfaktoren (Standard-Berichte), über ihre quantitative Auswertung (Ad hoc Berichte), dem Aufzeigen von Abhängigkeiten der einzelnen Faktoren (Abfragen / Aufschlüsselungen) über Warnungen bis hin zu statistischen Analysen, Hochrechnungen, prädiktiven Analysen und Optimierungsvorhersagen.

Zur Bewertung der analytischen Methoden sind die obersten vier Stufen zu betrachten. Die Ebenen zum Informationszugang und Reporting stellen „deskriptive Methoden“ dar. Hier werden während des Bauprojektes historische Daten (z. B. in Form von regelmäßigen Berichten) dokumentiert und beschrieben. So werden im laufenden Bauprojekt Aussagen über den bisherigen Projektverlauf und zu aktuellen Herausforderungen aufgezeigt. Dabei werden diese Daten für eine Planung neuer Projekte nicht verwendet.

Eine Einteilung analytischer Methoden kann nach Davenport und Harris (2007, S. 8) wie folgt kategorisiert werden:

1. **Statistische Zusammenhänge** zeigen den Zusammenhang zwischen einzelnen Merkmalen und der Ergebnisvariablen auf Basis einer Stichprobe auf. Die Modellparameter können mithilfe der Signifikanz und Robustheit getestet werden und verdeutlichen damit die Relevanz des Merkmals (vgl. Bzdok et al. 2018). Es wird analysiert: „Warum passiert etwas?“ bzw. „Welche Merkmale haben einen Einfluss?“
2. **Hochrechnungen** beziehen Methoden der Stochastik, die offensichtliche Zusammenhänge mathematisch-statistisch ausdrücken, ein. Es werden logische Wenn-Dann-Beziehungen formuliert. Damit wird die Frage geprüft: „Was passiert, wenn der Trend bestehen bleibt?“
3. **Prädiktive Methoden** analysieren Muster in Datensätzen inklusive unvorhergesehener Ereignisse und übertragen diese auf neue Datensätze. Somit können Lösungen auf die Frage „Was wird als Nächstes passieren?“ gefunden werden.
4. **Optimierungen** analysieren aufbauend auf den vorherigen Stufen optimale Ergebnisse und können damit Entscheidungen beeinflussen. Die Frage mit dem höchsten Wettbewerbsvorteil analytischer Methoden lautet „Was sollte ich machen?“ bzw. „Was ist das Beste, das dem Projekt jetzt passieren kann?“

## 2.5 Abweichungen von geplanten Dauern

### 2.5.1 Auswirkungen von Abweichungen

Abweichungen stellen positive oder negative Differenzen einer Planung dar. Dabei stellen positive Differenzen eine zeitliche Überschreitung der geplanten Dauer dar. Die Projekte benötigen länger als geplant. Negative Differenzen zeigen

eine Unterschreitung der geplanten Dauer auf und Projekte sind vor dem geplanten Ende eines Zieltermins der jeweiligen Projektphase fertig. Die Terminalsicherheit ist zusammen mit der Kostensicherheit nach Haghsheno (2004, S. 10) ein wichtiges Nebenziel des Bauherrn. Dabei kann eine fehlende Terminalsicherheit Auswirkungen auf die Kostensicherheit haben.

Abweichungen haben weitreichende Folgen. Eine zu optimistische Planung führt zu einem Zeitdruck innerhalb einzelner Phasen bzw. zu einem Zeitmangel. Dies kann folgende Auswirkungen auf das Projekt haben (vgl. Rogel und Osebold 2013, S. 232, Braimah und Ndekugri 2008):

- Konflikt,
- Demotivation bei Projektbeteiligten,
- Gefahr der Überarbeitung,
- Qualitätsdefizite (Fehler),
- erhöhte Nacharbeit,
- erhöhte Kosten (zur Fehlerbehebung),
- unzufriedene Kunden,
- fehlende Entscheidungen für Folgeprozesse.

Werden aufgrund der optimistischen Planung weiter einzelne Phasen und damit Vertragsfristen zu eng gestaffelt, birgt dies ein Potenzial für Behinderungen (vgl. Langen 2009, Eschenbruch und Rintelen 2010).

Auf der anderen Seite führt in Bauprojekten auch eine pessimistische Planung, eine Planung mit Zeitpuffern, zu negativen Folgen für das Projekt (vgl. Rogel und Osebold 2013, S. 232): Nach dem „Parkinson Law“ (Parkinson und Lancaster 1958) wird der in einem Projekt eingeplante Zeitpuffer immer aufgebraucht. Nach Goldratt (2002) sind „Sicherheitspolster von 200% [...] eher die Norm als die Ausnahme“. Demnach werden bei einem vollständigen Aufbrauchen der eingeplanten Puffer unnötige Kapitalbindungs- und Personalkosten fällig. Dies kann

ebenfalls zu Konflikten führen und Potenziale werden nicht genutzt. Nach Parkinson und Lancaster (1958) ist eine Unterschreitung der geplanten Dauer nicht üblich. Werden dennoch entgegen dem Gesetz nach Parkinson Projektdauern unterschritten, ist durch festgelegte Vertragsfristen mit Wartezeiten zu rechnen. Auch sind unnötige Kosten zu erwarten, z. B. für Sicherung und Wartung der Bauwerke, obwohl ein Betrieb erst zu einem späteren Vertragszeitpunkt stattfindet.

Demnach führt eine Überschreitung der geplanten Dauern („unrealistischen Optimismus“) (vgl. Kahneman und Tversky 1977), kann aber auch eine pessimistische Projektplanung zu negativen wirtschaftlichen, sozialen und politischen Auswirkungen führen. Eine möglichst hohe Qualität der Planung ist daher notwendig.

Die Qualität der Planung kann ex post durch folgende Aspekte beurteilt werden:

- die Planungsgenauigkeit als Differenz zwischen der Plan- und der Ist-Dauer einer Projektphase,
- den Aufwand an Maßnahmen, um bei Abweichungen zurück zum Planstand zu kommen.
- und die Summe der zeitlichen Puffer, die entweder aufgebraucht oder zur Optimierung der Dauer nicht genutzt werden.

Terminpläne können demnach entsprechend ihrer Güte kategorisiert werden. Im Folgenden wird eine Einstufung vorgenommen (siehe Tabelle 2.1). Dabei ist Typ vier eines Rahmenterminplans zu bevorzugen.

„Typ 1 Rahmenterminplan“: Die Abweichung von der geplanten Dauer innerhalb einer Projektphase kann (trotz Maßnahmen) nicht aufgeholt werden. Dies führt zu Verschiebungen von Phasen sowie der Verlängerung der Gesamtdurchlaufzeit.

„Typ 2 Rahmenterminplan“: Die geplante Dauer von Phasen bzw. des Projektes kann eingehalten werden. Es sind jedoch Pufferzeiten eingeplant, die aufgebraucht werden. Dies führt zu einer vergleichsweise längeren Durchlaufzeit als theoretisch möglich.

„Typ 3 Rahmenterminplan“: Es herrscht Zeitdruck in den Phasen bzw. dem Projekt. Maßnahmen mit wirtschaftlichen Auswirkungen müssen zur Zielerreichung aufgebracht werden.

„Typ 4 Rahmenterminplan“: Die Projektphasen werden ohne Verschiebungen eingehalten. Es sind keine Puffer enthalten und keine Maßnahmen mit wirtschaftlichen Folgen müssen eingeführt werden.

**Tabelle 2.1:** Kategorisierung von Rahmenterminplänen durch Maßnahmen und Puffer

		Maßnahmen	
		notwendig	nicht notwendig
Puffer	enthalten	Typ 1 Rahmenterminplan	Typ 2 Rahmenterminplan
	nicht enthalten	Typ 3 Rahmenterminplan	Typ 4 Rahmenterminplan

## 2.5.2 Komplexität von Bauprojekten

Die Begriffe „Komplexität“ und „Unsicherheit“ sind eng miteinander verknüpft. Daher werden beide im Folgenden erläutert.

Baccarini (1996) gilt als Ausgangspunkt für die Definition der Projektkomplexität (vgl. Geraldini et al. 2011, S. 972). Nach Baccarini (1996, S. 201) sind Bauprojekte komplex und werden im Laufe der Zeit immer komplexer. Er definiert Komplexität als die Variation und Interdependenz von Elementen im Projekt.

Edmonds (1999, S. 72) stellt in der Definition die Schwierigkeit der Modellbeschreibung heraus: „Komplexität ist die Eigenschaft eines Modells, die es schwierig macht, sein Gesamtverhalten in einer gegebenen Sprache zu formulieren, selbst wenn man einigermaßen vollständige Informationen über seine atomaren Komponenten und ihre Beziehungen untereinander hat.“ Vidal und Marle

(2008, S. 1101) erweitern die Definition von (Edmonds 1999, S. 72) um: „Ihre Triebkräfte sind Faktoren, die mit der Projektgröße, der Projektvielfalt, der Projektverflechtung und dem Projektkontext zusammenhängen.“

Zu der Definition von Baccarini (1996) und Edmonds (1999), Vidal und Marle (2008) werden weitere Aspekte hinzugefügt: Unsicherheit (vgl. Williams 1999), Dynamik, Variabilität und Veränderungen im System oder seiner Umgebung (vgl. Ribbers und Schoo 2002, Xia und Lee 2005), Tempo (vgl. Pinto und Morris 2004, Williams 2005) sowie soziale und politische Aspekte (vgl. Geraldi und Adlbrecht 2007, Maylor et al. 2013).

Komplizierte Systeme grenzen sich im Vergleich zur Komplexität durch ein stabiles Umfeld ab. Durch das sich häufig ändernde und unsichere Umfeld von Bauprojekten werden diese als komplex eingeordnet. Die Komplexität von Bauprojekten lässt sich zum einen durch das System selbst und zum anderen durch die internen Aufgaben der Projektbeteiligten begründen.

## Systemkomplexität

Die folgenden Unterpunkte beschreiben Merkmale der Systemkomplexität.

**Individualprodukte mit geringem Vorfertigungsgrad:** Ein Großteil der Bauwerke sind Individualprodukte (vgl. Berner et al. 2013, S. 3307), die sich in Inhalt und Funktionalität unterscheiden sowie mit einem geringen Vorfertigungsgrad erstellt werden. Der Grad der Individualisierung nimmt mit der Zunahme an Produkteigenschaften und Verfahren noch weiter zu. So zeigen Böhle et al. (2015) auf, „dass je komplexer technische und organisatorische Systeme werden, desto anfälliger werden sie auch für eine Vielzahl nicht exakt bestimmbarer Einflüsse“. Aus der hohen Individualisierung der Bauwerke und Wertschöpfung auf dem Baufeld ergeben sich nach Seyffert (2011, S. 8) projektspezifische Rahmenbedingungen für das System eines Bauprojektes.

**Multiagenten-Umgebung:** Bei der Realisierung von Bauprojekten ist üblicherweise eine Vielzahl an unterschiedlichen Beteiligten (Agenten) involviert. Diese



reichen von Bauherrn zu diversen Projektbeteiligten als Auftragnehmer des Bauherrn bzw. von dessen Projektpartnern. Bei größeren Bauvorhaben sind nicht selten eine dreistellige Zahl von Organisationen involviert. Zum Beispiel waren vor etwa 60 Jahren etwa fünf ausführende Firmen und ein Architekt für die Planung und Ausführung eines Hauses zuständig. Heute sind dies bereits ca. 25 Büros bzw. Betriebe (vgl. Grote 2000). Die Aggregation aus verschiedenen Fachdisziplinen erzeugt weiteres Unsicherheitspotenzial (vgl. Rogel und Osebold 2013, S. 218). Auch ist ein fortführender Trend der Fragmentierung zu erkennen: Die Baubranche besteht in Deutschland bezogen auf die Anzahl der Betriebe zu 99,6 % aus kleinen und mittleren Unternehmen (KMU) mit weniger als 250 Mitarbeitern (vgl. Bundesamt 2018b). Dieser Trend verhält sich umgekehrt zu anderen Branchen: Der Anteil der Betriebe mit weniger Mitarbeitern stieg über die letzten 20 Jahre sowohl gemessen an Anzahl der Beschäftigten pro Unternehmen als auch am Umsatz (vgl. Bundesamt 2018a). Durch die Vielzahl und Varietät an Beteiligten in der Multiagenten-Umgebung sind Systemzusammenhänge innerhalb eines Projektes jeweils neu aufzubauen. Wissen, Erfahrungen und Fähigkeiten der Individuen, Institutionen sowie Unternehmen sind divers verteilt und können je nach Projektzusammensetzung zu einer unterschiedlichen Ausgangslage führen (vgl. Hoffmann 2017, S. 137).

**Einmaligkeit der Bedingungen:** Das Projekt findet sich an einem einmaligen Ort der Wertschöpfung wieder. Dieser ist mit jedem Projekt an neue Rahmenbedingungen geknüpft. So können beispielhaft die Bodenverhältnisse, der Grad der Urbanisierung oder die dezentrale Verortung der Projektbeteiligten genannt werden. Auch sind die Systemzusammenhänge einmalig. (vgl. Rogel und Osebold 2013, S. 218)

**Lange Projektlaufzeit:** Bauprojekte sind häufig durch lange Projektlaufzeiten gekennzeichnet. Das System baut sich über die Zeit auf, Projektinhalte werden kontinuierlich entwickelt und unterstehen einem dynamischen Prozess, der durch externe Einflüsse (z. B. die Marktlage) bestimmt werden kann. Als Folge werden Rückschlüsse bzw. Wirkungszusammenhänge häufig übersehen.

## Aufgabenkomplexität

Die Aufgabenkomplexität kann nach Russell und Norvig (vgl. 2012, S. 69ff.) in teilweise beobachtbare, sequentielle und dynamische Aufgaben unterteilt werden (Oprach et al. 2019b):

**Teilweise beobachtbare Aufgaben:** Bauprojekte weisen im Gegensatz zu einer vollständig beobachtbaren Umgebung aufgrund der hohen Anzahl von Einflussfaktoren und nicht erfassbaren manuellen Prozessen oft Lücken in der Dokumentation auf. Dies hat zur Folge, dass explizites Wissen nur unvollständig dokumentiert wird. Auch der NIST-Bericht (National Institute of Standards and Technology) von 2004 (vgl. Gallaher et al. 2004, S. 2ff.) stellt fest, dass die meisten Beteiligten bei der Umstellung auf elektronische Systeme zurückhaltend sind. Implizites oder persönliches Wissen ist im Bauwesen daher sehr verbreitet. Diese Art von Wissen ist schwer zu artikulieren und basiert auf Erfahrung, Intuition, Gefühlen und subjektiven Ansichten. Die verbale Übertragung von Expertenwissen bzw. vagen Erinnerungen resultiert in Unschärfe und / oder Missverständnissen (vgl. Rogel und Osebold 2013, S. 218). Eine vollständige Dokumentation über die Projektlaufzeit ist derzeit nicht gegeben. Elektronische Systeme auf Basis von IoT-Systemen streben jedoch eine kontinuierliche Erfassung von Daten an.

**Sequentielle Aufgaben:** In sogenannten episodischen Umgebungen, in denen Phasen unabhängig voneinander sind, haben die aktuellen Entscheidungen keinen direkten Einfluss auf die folgenden Aufgaben. Im Gegensatz dazu haben in sequentiellen Umgebungen Entscheidungen einen direkten Einfluss auf alle weiteren Prozesse. Kurzfristige Aktionen können zu langfristigen Auswirkungen führen. So kann eine Entscheidung in der Bauplanung weitreichende Auswirkungen auf den Ausführungsprozess haben. Folgen von Entscheidungen müssen demnach abgesehen werden.

**Dynamische Aufgaben:** Ein statisches Umfeld ist eines, in dem sich das Umfeld während der Entscheidungsfindung nicht verändert. Aufgrund der starken Fragmentierung von Bauprojekten müssen regelmäßig parallele Entscheidungen

getroffen werden, die sich wechselseitig beeinflussen. Man spricht daher in Bauprojekten von einer dynamischen Aufgabenumgebung.

## 2.5.3 Unsicherheit als Ursache

### Unsicherheit

Ergebnis der System- und Aufgabenkomplexität sind unvollständige Informationen über das Bauprojekt. Dies wird als Unsicherheit beschrieben.

Unsicherheit wird definiert als „die Differenz zwischen der Menge an Informationen, die zur Erfüllung einer bestimmten Aufgabe erforderlich sind, und der Menge an Informationen, über die die Organisation bereits verfügt“ (vgl. Galbraith 1973, S. 5). Je weiter man in die Zukunft plant, desto mehr Informationen fehlen und ein Einhalten von deterministischen Terminen ist mit Unsicherheiten behaftet. Dies sagt auch das Unschärfe-Prinzip nach Rescher (1982) aus. Hier nimmt mit einer zunehmenden Bestimmtheit einer Prognose deren Sicherheit ab.

In der Entscheidungstheorie werden unterschiedliche Situationen, in denen der Eintritt von künftigen Entwicklungen nicht mit Sicherheit vorausgesagt werden kann, in Risiko, Ungewissheit und Unwissenheit unterteilt. (vgl. Rogel und Osebold 2013)

**Risiko:** Risiko bezeichnet den Unkenntnisgrad, „[...] bei dem für das Eintreten zukünftiger Ereignisse objektive Wahrscheinlichkeiten vorliegen.“ Die Bestimmung bzw. Berechnung des Risikos ist somit möglich. (Rogel und Osebold 2013, S. 212)

Zusammenfassend lässt sich der Begriff Risiko als Möglichkeit einer Zielverfehlung interpretieren, wobei potenzielle positive Abweichungen Chance und eventuelle negative Abweichungen Gefahr genannt werden. (vgl. Girmscheid 2006, S. 548f.)

**Ungewissheit:** Ungewissheit ist ein Unkenntnisgrad, „[...] bei dem für das Eintreten zukünftiger Ereignisse zwar keine objektiven Wahrscheinlichkeiten vorliegen

(Risiko), aber subjektive Wahrscheinlichkeiten gebildet werden können.“ Es ist keine statistische Datenbasis vorhanden, um dazugehörige Eintrittswahrscheinlichkeiten zu bilden (vgl. Rogel und Osebold 2013, S. 212).

**Unwissenheit:** Unwissenheit liegt vor, wenn selbst die Auswirkungen der untersuchten Handlungsalternativen nicht vollständig bekannt sind. (vgl. Rogel und Osebold 2013, S. 212)

Die unvollständigen Informationen zum Zeitpunkt der Rahmenterminplanung führen häufig zu einer Ungewissheit oder sogar Unwissenheit in der Planung.

Um die Ungewissheit bzw. Unwissenheit zu begründen, existieren drei Denkschulen. In diesen kann die Begründung der Unsicherheit durch unternehmensinterne Ursachen oder durch externe und den Markt bestimmende Faktoren definiert sein. In der „klassischen Sichtweise“ ist die externe Umwelt die Ursache der Unsicherheit. Unvorhergesehene Ereignisse, Turbulenzen und Veränderungen definieren u. a. die Unsicherheit. Diese beeinflussen Entscheidungen, Strukturen und die jeweilige Leistung des Unternehmens. Hingegen sind in der „Übergangssichtweise“ interne und externe Ursachen zu finden. Interne Ursachen sind in Abhängigkeiten, in der Informationsverarbeitung und in Entscheidungen zu finden. Extern wird diese definiert durch fehlendes Wissen, Veränderungen und Komplexität. In der „Prozesssichtweise“ sind insbesondere interne Gründe der Unsicherheit zu finden. Die Wahrnehmung von Zuständen, Abhängigkeiten, kognitiven Prozessen und Veränderungen sind interne Ursachen. (vgl. Jauch und Kraft 1986)

Die externen Ursachen können mit der Systemkomplexität und die internen Ursachen mit der Aufgabenkomplexität verglichen werden.

### **Folgen für die Datengrundlage**

Die Ungewissheit bzw. Unwissenheit spiegelt sich in der Datengrundlage wider. Die Datengrundlage kann wie folgt beschrieben werden:

**Heterogene und unvollständige Datengrundlage.** Die hohe Fragmentierung der Bauwirtschaft ist u. a. ein Grund für die geringe Digitalisierung und Aufbereitung der Baudaten. Shen et al. (2010) erläutern, dass v. a. kleine Unternehmen Probleme mit der Digitalisierung aufgrund mangelnder Kapazitäten und unzureichenden Kapitals haben. Ergebnis sind komplexe Baudatenstrukturen. Hierbei sind zunächst die stark heterogenen Datenformate (Korrespondenzen, Terminpläne, Verträge, Bilder, Bestellungen etc.) (vgl. Aouad et al. 1999) zu nennen. Weiter führen die zahlreichen Projektbeteiligten verschiedener Disziplinen in voneinander abhängigen Projektphasen zu einer heterogenen Nutzung von Software und Hardware (vgl. Shen et al. 2010). Durch die Vielzahl möglicher Kombinationen an Projektbeteiligten werden Arbeitsweisen und -dokumente in der Regel projektspezifisch aufgebaut und beschrieben. Eine Standardisierung von Prozessen und Produkten ist unternehmensübergreifend kaum möglich und bleibt innerhalb der Unternehmen eine große Herausforderung. Die Fragmentierung führt zudem zu einer dezentralen Speicherung der Projektinformationen. Teile des projektspezifisch gesammelten Wissens bleiben implizit bestehen und werden durch persönlichen Erfahrungsaustausch in neue Projekte übertragen. Bedingt durch die Vielzahl möglicher Einflussfaktoren ist die Dokumentation des Projektes häufig unvollständig und in einer geringen Qualität (vgl. Bouchlaghem et al. 2004, Maniyka et al. 2011). Vor allem die langen Projektlaufzeiten führen zu einer langsamen Geschwindigkeit bei der meist manuell erstellten Dokumentation, sodass auf Teile der gesammelten Daten erst Monate oder Jahre nach der Realisierung zugegriffen werden kann.

**Unvollständige und nicht eindeutige Beschreibung von Zusammenhängen.** Die Örtlichkeit eines Bauvorhabens kann beispielsweise die Eigenschaften des Bauwerks beeinflussen oder die Art der Kommunikation hat Einfluss auf die Einhaltung des Terminplans. Sind relevante Zusammenhänge zwischen Projektmerkmalen identifiziert, kann diese Informationsgrundlage die Planung neuer Projekte verbessern. Die Beschreibung dieser Zusammenhänge ist jedoch oft unvollständig und nicht eindeutig, da häufig keine Standards für die Dokumentation dieser Zusammenhänge existieren (vgl. Oprach et al. 2019a). Die Anzahl der

Merkmale, die ein Bauprojekt beschreiben und ihrer Zusammenhänge, entsprechen einem komplexes System. Relevante Zusammenhänge können bei abgeschlossenen Projekten nur mit komplexen analytischen Planungsmethoden identifiziert werden.

## 2.5.4 Ansätze zur Reduktion von Unsicherheiten

Um die Ungewissheit bzw. Unwissenheit zu reduzieren, können folgende Ansätze unterstützen. Nach Jauch und Kraft (1986) existieren vier Strategien, um mit einer strategischen Unsicherheit umzugehen: Die Reduktion der internen und externen Unsicherheit sowie jeweils deren Simulationen. Allaire und Firsirotu (1989) fasst dies in drei konkrete Handlungsfelder zusammen:

1. **Vorhersagen treffen und sich auf unvorhergesehene Situationen vorbereiten („predict-and-prepare“):** Es sind verschiedene Pläne zu erstellen, unvorhergesehene Ereignisse sind schnell zu identifizieren, zu analysieren und Maßnahmen sind zu ergreifen. Jauch und Kraft (1986) unterscheidet hier zur Unterstützung der Vorhersage in interne und externe Simulationen der Unsicherheiten. Dieses Vorgehen empfiehlt auch Kato und Ahern (2008) und versteht es als ein „learning by doing“.
2. **Eine Prognose erstellen und Unsicherheiten eliminieren („don’t predict the future, control it“):** Dies kann u. a. durch eine Weitergabe der Risiken geschehen (z. B. vertraglich fixierte Dauern mit Strafen) oder dem Aufteilen von Risiken (z. B. durch eine integrierte Projektabwicklung (IPA)) (vgl. Hagsheno et al. 2022). Mit dieser Strategie werden insbesondere interne Unsicherheiten reduziert.
3. **Kapazitäten aufbauen, um sich flexibel anzupassen („Be ready for whatever it is“):** Hierfür sollte die Produktpalette des Unternehmens erweitert werden, um auf verschiedene Szenarien vorbereitet zu sein. Mit dieser proaktiven und ressourcenbindenden Gestaltung des Arbeitsumfeldes kann auf externe Unsicherheiten reagiert werden.

Von diesen drei Strategien ist nach Allaire und Firsirotu (1989) die erste („predict-and-prepare“) tendenziell bei einer vergleichsweise geringeren Unsicherheit zu wählen als die zweite und dritte Strategie („don‘t predict the future, control it“ und „Be ready for whatever it is“).

Cleden (2017) unterstützt die Strategien von Allaire und Firsirotu (1989) und nennt hierbei spezifische Methoden. So nennt Cleden (2017) Vorhersagen als eine relevante Strategie in Form von Prognosemodellen oder Wissensgrafiken. Um sich auf unvorhergesehene Ereignisse vorzubereiten, sind Methoden zur Antizipation der Ereignisse relevant. Dazu gehören Entscheidungsbäume, Szenariestudien, mehrere Hochrechnungen oder ein vorwärts- und rückwärtsorientiertes Durchdenken der Planung. Neben der Planung nennt Cleden (2017) den Umgang mit eingetroffenen unvorhergesehenen Ereignissen: Das Unternehmen muss agil sein und schnell aus Folgen der Unsicherheiten lernen. Weiter ist die Ursache der Unsicherheit zu finden und zu analysieren. Zuletzt sind das Mindset und die Kreativität der Planer relevant, um aus etwaigen Fehlplanungen zu lernen. So kann ein Problem bewusst aus verschiedenen Perspektiven betrachtet werden oder es können „Was wäre wenn...“ Szenarien analysiert werden.

Demnach stehen hinter diesen Strategien und Methoden der Bedarf nach (vgl. Brashers 2001):

**(1) Dem Sammeln von Informationen:** Informationen sind zu sammeln, um Szenarien zu untersuchen und das Wissen der Planer zu prüfen. Dabei ist auf eine konsistente Datenaufnahme zu achten, bei der auch verschiedene Quellen einbezogen werden sollten. Auf der anderen Seite ist Brashers (2001) der Meinung, dass die Unsicherheit durch Informationen erhöht wird, wenn nicht erklär-bare Muster identifiziert werden. Diese Muster (vgl. „Rauschen“ in Kapitel 5.3.4) sind zu vermeiden.

**(2) Akzeptanz von Unsicherheit und Vertrauen in Projektpartner:** Es sollte sich der chronischen Unsicherheit angepasst und diese akzeptiert werden. Das Ergebnis der Vorhersage kann demnach eher als Anhaltspunkt gesehen werden, das „gut genug“ ist. Hierfür ist ein hohes Maß an Vertrauen in sich selber und die Projektpartner notwendig.

**(3) Einer Unterstützung durch das Management:** Es ist eine soziale Unterstützung notwendig, indem den Planern Vertrauen gegeben wird und Fehlplanungen nicht stigmatisiert werden. Weiter ist eine Unterstützung in der Informationsgewinnung und ihrer Auswertung bereitzustellen.

**(4) Der Vermittlung zwischen Informationsgebern und -nutzern:** Häufig sind Informationen unvollständig, falsch oder fehlerhaft dokumentiert. Kognitive Fähigkeiten sind notwendig und unterschiedlich verteilt, um diese Herausforderungen in der Informationsverarbeitung zu identifizieren und mit ihnen umgehen zu können. Die Kosten zur Gewinnung der „richtigen“ Informationen sind daher oft zu hoch. Dies führt zu einem Dilemma zwischen denen, die Informationen geben und denen, die Informationen auswerten.

## 2.5.5 Systematische Planungsfehlschlüsse als Ursache für Abweichungen

### Planungsfehlschlüsse

Als weiteren Grund für Planabweichungen können die systematischen Planungsfehlschlüsse genannt werden.

Nach Kahneman und Tversky (1977) unterschätzen Personen bei der Planung von Projekten tendenziell die erforderliche Zeit, um ein Projekt abzuschließen. Dieses geschieht selbst dann, wenn sie Erfahrungen mit vergleichbaren Projekten in der Vergangenheit gesammelt haben. Die Erfahrungen müssten zu dem Schluss führen, dass die Vorhersage den tatsächlich benötigten Aufwand unterschätzt und mehr Zeit für das Projekt benötigt werden würde. Dieses Phänomen wird „systematischer Planungsfehlschluss“ (engl. „Planning Fallacy“) genannt und ist seitdem in einer Vielzahl von Studien beobachtet worden: Bei akademischen und nichtakademischen Projekten (vgl. Buehler et al. 1994), dem Zusammenbau eines Computers (vgl. Byram 1997), dem Schreiben eines Aufsatzes (vgl. Koole und van't Spijker 2000) oder auch großen Bauprojekten (vgl. Flyvbjerg et al.



2016). Der Planungsfehlschluss existiert daher ungeachtet des Planungsgegenstandes.

Grundsätzlich scheinen Personen bei Vorhersagen eher an positivere Ergebnisse zu glauben als als realistischerweise gerechtfertigt wäre. Dies wird erstmals als „unrealistischer Optimismus“ bezeichnet (vgl. Weinstein 1980).

Im Folgenden wird auf das Konzept des unrealistischen Optimismus eingegangen.

### **Unrealistischer Optimismus**

Die Literatur zu Planungsfehlschlüssen verwendet in ihren Studien zentrale Begriffe wie *Optimismusverzerrung* (vgl. Flyvbjerg 2008), *Überoptimismus* (vgl. Lovallo 2003) und *unrealistischer Optimismus* (vgl. Kruger und Evans 2004). Diese werden dabei häufig synonym verwendet. Die nachfolgende Definition eines unrealistischen Optimismus soll die Begriffsverwendung auf ein einziges systematisches Konzept reduzieren.

Unter dem unrealistischen Optimismus versteht man, dass Menschen die Zukunft positiver einschätzen, als sie überhaupt sein kann. Diese Tendenz zu positiven Ergebnissen wird als unrealistischer Optimismus bezeichnet (vgl. Shepperd 2013). Nach ihrem Konzept ist weiter zwischen unrealistischem absolutem Optimismus und unrealistischem relativem Optimismus zu unterscheiden. Unrealistischer absoluter Optimismus existiert aufgrund der Vorstellung von Personen, dass negative Ergebnisse seltener eintreten würden als objektiv gerechtfertigt ist. Unrealistischer relativer Optimismus existiert, wenn Personen ihr eigenes Risiko verglichen mit dem einer anderen Person als geringer einschätzen als objektiv gerechtfertigt ist. Beides kann zudem auf individueller Ebene oder einer Gruppenebene auftreten. Auf individueller Ebene schätzt eine Person also ihr eigenes Risiko absolut oder ihr Risiko im Verhältnis zu einem Vergleichsstandard relativ, bspw. einem Durchschnittsrisiko, ein. Auf Gruppenebene schätzt jede Person einer Gruppe ihr eigenes Risiko (absolut) oder ihr Risiko im Verhältnis zu einer Vergleichsgruppe (relativ) ein.

Bezogen auf systematische Planungsfehlschlüsse zeigt sich ein unrealistischer absoluter Optimismus (vgl. Shepperd 2013, S. 4). Schätzt der Projektleiter eines Bauvorhabens bspw. den Fertigstellungszeitpunkt seines Projektes, stellt dies zunächst eine persönliche absolute Risikoschätzung dar. Der objektive Standard könnte in diesem Fall die durchschnittliche Dauer aller ähnlichen Bauvorhaben sein. Der unrealistische Optimismus würde in diesem Fall der vorteilhafte Unterschied zwischen der Schätzung des Projektleiters und der objektiven Durchschnittsdauer aller ähnlichen Bauvorhaben sein. Der vorteilhafte Unterschied stellt eine Unterschätzung des Fertigstellungsdatums dar.

### **Erklärungsansätze für Planungsfehlschlüsse**

Es gibt zwei große Denkschulen, die die Ursachen von Abweichungen analysieren (vgl. Love et al. 2015a): die *Psychostrategen* und die *Evolutionstheoretiker*.

Die *Psychostrategen* sehen Gründe für Abweichungen vor allem in strategischer Fehlinterpretation und optimistischer Voreingenommenheit bei den geplanten Handlungen (z. B. Flyvbjerg et al. 2002, Kahneman und Tversky 1977, Siemiatycki 2009). Die optimistische Verzerrung ist das Ergebnis einer Unterschätzung der Projektrisiken und einer Überschätzung des politischen, wirtschaftlichen und strategischen Nutzens (vgl. Flyvbjerg 2006). Bei diesem Ansatz werden die eigenen Handlungen und die der Organisation als voll verantwortlich für die Planungsfehler angesehen. Beispielhaft seien vier Ansätze der Psychostrategen genauer vorgestellt und miteinander verglichen.

Kahneman und Tversky (1977) führen Planungsfehlschlüsse auf das Auslassen von Informationen aus vergleichbaren vergangenen Projekten bei der Aufwandschätzung zurück. Die Autoren argumentieren, dass Menschen ihre Projekte als einzigartig wahrnehmen und deshalb zu wenig Informationen vergleichbarer Projekte in ihre Schätzungen einbeziehen würden.

Auch Buehler et al. (1994) sind der Auffassung, dass relevante Informationen nicht berücksichtigt werden. Dies geschehe jedoch nicht der Einzigartigkeit des

Projektes wegen. Negative Auswirkungen in vergangenen Projekten würden Akteure viel mehr als externe Einflüsse einordnen und daraus resultierend nicht für einen Bestandteil des Projektaufwands halten.

Im Gegensatz zu diesen beiden Erklärungsansätzen ist Roy et al. (2005) nicht der Meinung, dass bestimmte relevante Informationen unterschlagen werden, sondern die Erinnerungen an vergangene Projektaufwände seien verzerrt.

Kruger und Evans (2004) führen Planungsfehlschlüsse darauf zurück, dass Personen normalerweise vielschichtige Projekte nicht ausreichend in ihre Subkomponenten zerlegen. Projektspezifische Informationen müssten noch viel stärker betrachtet werden. Dieser Ansatz steht damit in Kontrast zu den Annahmen von Kahneman und Tversky (1977) und Buehler et al. (1994), externe Informationen vergleichbarer Projekte stärker zu berücksichtigen.

Neben kognitiven scheinen zudem motivationsbezogene Ursachen eine wichtige Rolle bei unrealistischem absolutem Optimismus zu spielen (vgl. Kunda 1990). So können bspw. monetäre Anreize den unrealistischen Optimismus verstärken (vgl. Buehler et al. 1994).

Im Vergleich dazu konzentrieren sich *Evolutionstheoretiker* auf projektinterne Änderungen. Das heißt, sie glauben, dass Planabweichungen das Ergebnis von Änderungen des Umfangs und der Definition zwischen Beginn und Abschluss des Projekts sind (z. B. Odeyinka et al. 2012, Love et al. 2011, Ahiaga-Dagbui und Smith 2014). Diese Änderungen bezeichnen sie als Pathogene bzw. Krankheitserreger. Kategorien von Krankheitserregern sind: Die bewusste Praxis der Menschen, die Art der Aufgaben, die Umstände, die Konventionen, die Organisationsstruktur und ihre Systeme, die Eigenschaften der Branche und die technischen Werkzeuge (vgl. Busby und Hughes 2004). Nach Busby und Hughes (2004) ist die Praxis das am häufigsten beobachtete Pathogen. Häufig sind sich die Projektteilnehmer der Folgen ihres Handelns nicht bewusst. Beispiele hierfür sind aktives Versagen, Fehler, Auslassungen, Nacharbeiten, Änderungsaufträge, Zwischenfälle, Reklamationen und Änderungen des Projektumfangs. Daher lassen sich Zeit, Kosten und Qualität eines Projekts in der Anfangsphase kaum mit dem abgeschlossenen Projekt vergleichen.

Beide Ansätze nehmen unabhängig voneinander eine zu starke Vereinfachung des Problems vor. Jede Erklärung muss in einem „ausgewogenen“ Ansatz verwendet werden, um die Ursachen und Mechanismen für Abweichungen zu definieren. Die externen politischen, wirtschaftlichen, psychologischen und verwaltungstechnischen Faktoren können zu internen Problemursachen führen (vgl. Love et al. 2015b). Verschiedene Arbeiten gehen von einem deterministischen Ansatz aus, um die Beziehung zwischen Ursachen und ihren Auswirkungen zu analysieren (z. B. Hinze et al. 1992, Ahiaga-Dagbui und Smith 2014). Love et al. (2015b) zeigen jedoch den zusätzlichen Bedarf an Daten zu Wahrscheinlichkeiten und Auswirkungen bei der Analyse von Ursache-Wirkungs-Beziehungen für ein besseres Verständnis auf. Die Anzahl der beteiligten Parteien und die Komplexität der Kausalketten machen diesen Prozess aber schwierig (vgl. Love et al. 2015b). Daher werden im folgenden Abschnitt die Auswirkungen dieser Ansätze auf die Planungsqualität betrachtet und nicht die Auswirkungen von Ursachen.

## 2.5.6 Ansätze zur Reduktion von systematischen Planungsfehlschlüssen

Da es eine enge Verbindung zwischen beiden Denkschulen gibt, schließen sich die möglichen Strategien nicht gegenseitig aus. Mögliche Lösungsansätze zur Vermeidung von Planungsfehlschlüssen sind die fünf unten aufgeführten Ansätze von Siemiatycki (2009):

**(1) Verbesserung der Leistungsüberwachung, der Berichterstattung und des Informationsaustauschs:** Den meisten der von Siemiatycki (2009) analysierten Studien zufolge sind Leistungsüberwachung, Berichtswesen und Informationsaustausch die gängigsten Ansätze zur Verringerung von Abweichungen und zur Steigerung der Planungsqualität. Dazu gehören durchdacht gesetzte Zielvorgaben für Erwartungen, die Datenerhebung, ihre Analyse und Verbreitung (vgl. State North Carolina 2008). Kahneman und Tversky (1977) unterstützen diese Strategie. Sie schlagen vor, sich auf die Planung sowie auf die Verbreitung von Informationen anstatt auf singuläre Daten zu konzentrieren, um die strategische

Fehlinterpretation und den Optimismus reduzieren zu können. Singuläre Daten beschreiben die Informationen oder Daten, die sich auf das zu planende Projekt beziehen, wie z. B. die Anforderungen der Kunden, die Form des Gebäudes und die verfügbaren Ressourcen. Distributive Daten sind dokumentierte Daten von bereits abgeschlossenen Projekten. Es gibt im Allgemeinen zwei Möglichkeiten, um distributive Daten zu gewinnen. Zum einen kann man auf das intuitive Wissen eines Pools von Experten zurückgreifen. Zum anderen kann man auf eine große Anzahl von Projektdaten zur analytischen Auswertung zurückgreifen. Die Planung mit einer Datenbank liefert objektive Ergebnisse ohne Verzerrungen und bietet die Möglichkeit der Reproduzierbarkeit (vgl. Bari et al. 2014, Scherm et al. 2016). Ein klarer Nachteil ist der hohe Aufwand für die Verwaltung der Daten. Daten müssen aufbereitet, strukturiert und aktualisiert werden. Außerdem sind die Ergebnisse nicht immer korrekt. Das ist ein Problem, denn je ungenauer die Datenbasis ist, desto ungenauer ist auch das Ergebnis der analytischen Entscheidung auf dieser Datenbasis (vgl. Bari et al. 2014). Es stellt sich daher die Frage: Welche Größe und Datenbankqualität vorhanden sein muss, um den Planungsprozess ausreichend zu unterstützen?

**(2) Anwendung modernster Techniken:** Zur Planung von Dauern können verschiedene analytische Methoden eingesetzt werden. Nach Davenport und Harris (2007) bieten verschiedene analytische Methoden einen wirtschaftlichen Vorteil gegenüber Wettbewerbern. Diese Methoden können in die in Kapitel 2.4.3 vier vorgestellten Stufen eingeteilt werden: Der Zugang zu Daten, deren statistische Analysen und Benchmarking, Hochrechnungen und Vorhersagemodelle sowie Modelle zur Optimierung. In Kapitel 3 wird auf diese Methoden weiter eingegangen. Insbesondere werden Methoden der Künstlichen Intelligenz (KI) als moderne Techniken zunehmend eingesetzt. Budgetierung, Zeitdruck und fehlendes Fachwissen schränken den Einsatz von State-of-the-Art-Techniken jedoch oft ein. Es stellt sich daher die Frage, welche Techniken derzeit in den Unternehmen eingesetzt werden und ob das analytische Niveau dieser Techniken einen Einfluss auf die Planungsqualität hat.

**(3) Erhöhung der Vollständigkeit und Genauigkeit der frühen Pläne:** Um analytische Techniken nutzen zu können, müssen frühe Pläne vollständig sein

und strikt eingehalten werden. Daher schlägt Anderson et al. (2007) vor, Prozesse teilweise zu standardisieren und damit die Vollständigkeit der Pläne zu unterstützen. Alternativ zur Standardisierung kann eine erhöhte Informationstransparenz unterstützen. Frühe Pläne müssen so vollständig wie möglich sein und alle relevanten Informationen enthalten, um eine genaue Planung zu ermöglichen. Diese Informationen müssen von den Projektbeteiligten eingeholt werden. Häufig ist die Unvollständigkeit auf politischen, wirtschaftlichen und psychologischen Druck zurückzuführen. Neben einer Erhöhung der Transparenz könnte laut Siemiatycki (2007) und Flyvbjerg (2007) eine zusätzliche Lösung darin bestehen, die Gemeinschaft stärker einzubeziehen und professionelle Prüfungsgremien oder unabhängige Prüfer einzubeziehen. Funkhouser (2011) stellt fest, dass eine verbesserte Leistung der Prüfer oft mit der Einhaltung von Vorschriften verbunden ist. Folglich müssen für eine gute Planung relevante Informationen definiert und als Standard festgelegt werden, um die Qualität und Vergleichbarkeit der Dokumentation zu erhöhen.

**(4) Stärkung der Rechenschaftspflicht und Verantwortung für Fehler und Überschreitungen:** Bei diesem Ansatz werden positive Anreize für Einzelpersonen oder Unternehmen geschaffen, die ihre Zielvorgaben erreichen, während diejenigen, die ihre Leistungserwartungen nicht erfüllen, bestraft werden. Im Allgemeinen gibt es zwei mögliche Ansätze. Erstens wird die Leistung der Unternehmen als Präqualifikationssystem für den Auftrag bewertet. Zweitens werden diejenigen, die Prognosen manipulieren, sanktioniert, um ihre Verantwortung zu erhöhen (vgl. Flyvbjerg et al. 2002). Allerdings gibt es in der Forschung gemischte Erkenntnisse über die Wirksamkeit der Erhöhung der Rechenschaftspflicht und der Verantwortung. Dubnick (2005) verweist auf den Begriff „Rechenschaftsparadoxon“, was bedeutet, dass zunehmende Bemühungen um eine Leistungsverbesserung durch die Rechenschaftspflicht eher den gegenteiligen Effekt haben. Anstatt die gewünschte Leistungsverbesserung voranzutreiben, wirkt die Rechenschaftspflicht eher als „Bremsklotz“, indem sie Verbesserungen entweder verlangsamt oder stoppt. Dubnick (2005) schlägt stattdessen vor, Erwartungen zu setzen und das Vertrauen zu fördern. Dies erfordert Fähigkeiten und

Fertigkeiten des Managementteams und seiner Mitarbeiter. Es sind weitere Untersuchungen erforderlich, um die Auswirkungen dieses Ansatzes für die Anfangsphase von Bauprojekten zu analysieren.

**(5) Verbesserung der Managementfähigkeiten der Mitarbeiter:** Diese Strategie konzentriert sich auf die speziellen Fähigkeiten, Fertigkeiten und Instrumente des Managements und der Mitarbeiter. Dazu gehören die wirksame, klare und deutliche Formulierung von Vorschlägen, die Verwaltung von Ausschreibungen und die Ermittlung des besten Wertes sowie die Ausarbeitung von Verträgen, um Risiken zu verringern und Konflikte zu lösen. Laut Bedard (1991) halten sich Experten konsequenter an die Unternehmensstandards und verfügen über ein höheres Konsenswissen als Neulinge. Es gibt jedoch auch andere Studien mit widersprüchlichen und nicht schlüssigen Ergebnissen, wenn Experten zu allgemeine Informationen (vgl. Frederick und Libby 1986), zu vereinfachte Versuchsaufgaben (vgl. Abdolmohammadi und Wright 1987) und schlecht definierte Qualitätsstandards (vgl. Biggs 1985) einbeziehen. Daher sollten die Auswirkungen des Niveaus von Fachwissen und Kenntnissen in der Projektplanung in Bezug auf die Planungsqualität für dieses spezielle Thema analysiert werden.

## 2.6 Schlussfolgerungen und Anforderungen an analytische Planungsmethoden zur Rahmenterminplanung

Dieses Unterkapitel fasst wesentliche Grundlagen zur Prognose von Dauern eines Bauprojekts in frühen Projektphasen zusammen und bildet damit die Ausgangsbasis für Anforderungen an analytische Planungsmethoden.

Das Ergebnis der Prognose von Dauern wird erstmals in der Rahmenterminplanung festgehalten. Neben diesem Basisterminplan sind analytische, datenbasierte Planungsmethoden als Nominalwerte und repräsentative Ergebnisse als Planungsebene einzubeziehen.

Die Rahmenterminplanung beinhaltet das „Erstellen eines Terminplans mit den wesentlichen Vorgängen des Planungs- und Bauablaufs. Dabei sind etwaige Terminvorstellungen des Bauherrn hinsichtlich ihrer Realisierbarkeit zu überprüfen und müssen ggf. korrigiert werden“ (vgl. HOAI 2018, S. 97). Die Zuständigkeit der Rahmenterminplanung kann dabei vom Bauherrn auf verschiedene Akteure übertragen werden. Dazu gehören ein u. a. externe Projektmanager bzw. Projektsteuerer, Generalunternehmer, Totalunternehmer, Baubetreuer oder Bauträger, wenn diese entsprechend eingebunden werden. Wird die Rahmenterminplanung durch den Bauherrn selbst durchgeführt, ist eine entsprechende Expertise für eine möglichst valide Planung notwendig. Es sollte eine genaue Dokumentation historischer Bauprojekte als Vergleichsgrundlage vorliegen. Wird diese Planung ausgelagert, sollte der Rahmenterminplan möglichst einfach und bezüglich stark beeinflussender Projektmerkmale nachvollziehbar dargestellt werden. Die Darstellungsweise wird durch die Komplexität des Bauprojektes sowie seiner Aufgaben beeinflusst. Aus dieser Komplexität resultieren fehlende Informationen zum Zeitpunkt der Rahmenterminplanung, was als Planungsunsicherheit verstanden wird. Zusätzlich ist auf systematische Planungsfehlschlüsse zu reagieren. Diese bezeichnen eine generelle Unterschätzung der geplanten Dauern auf Basis verzerrter Erfahrungswerte.

Tabelle 2.2 fasst die wesentlichen Ansätze zur Reduktion der Unsicherheit und systematischer Planungsfehlschlüsse als Ursachen von Abweichungen zusammen.

Bei einem Vergleich beider Spalten ist zu erkennen, dass zur Reduktion von Unsicherheiten und systematischer Planungsfehlschlüsse analytische Planungsmethoden und soziale Kompetenzen notwendig sind. Die analytischen Planungsmethoden umfassen das Sammeln von Informationen bzw. die Vervollständigung früher Pläne mit Informationen, den Informationsaustausch, die Berichterstattung und Leistungsüberwachung sowie die Anwendung modernster Techniken (vgl. Siemiatycki 2009). Bei den sozialen Kompetenzen vertritt Brashers (2001) die Auffassung einer Entlastung der Planer durch Vertrauen und eine aktive Unterstützung durch das Management in der Planung. Siemiatycki (2009) hingegen



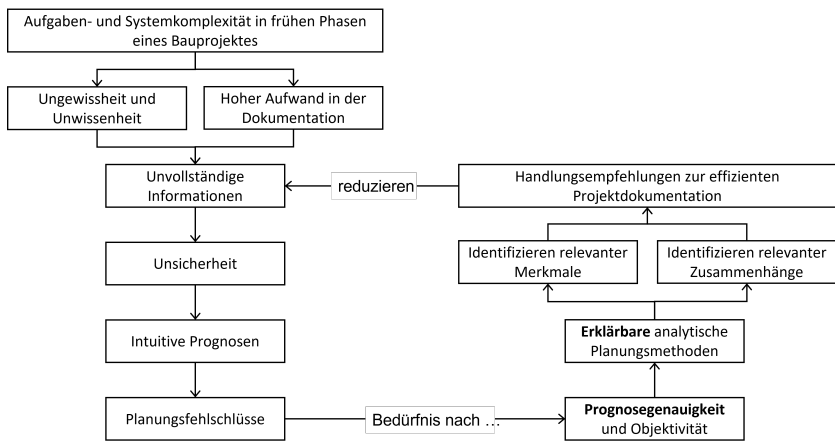
**Tabelle 2.2:** Zusammenfassung der Ansätze zur Reduktion von Unsicherheit (vgl. Brashers 2001) und systematischen Planungsfehlschlüssen (vgl. Siemiatycki 2009)

<b>Ansätze zur Reduktion der Unsicherheit</b>	<b>Ansätze zur Reduktion systematischer Planungsfehlschlüsse</b>
<b>Analytische Planungsmethoden</b>	
(1) Sammeln von Informationen	Verbesserung der Leistungs- (1) überwachung, der Berichterstattung und des Informationsaustauschs
	(2) Anwendung modernster Techniken
	(3) Erhöhung der Vollständigkeit und Genauigkeit der frühen Pläne
<b>Soziale Kompetenzen</b>	
(2) Akzeptanz der Unsicherheit und Vertrauen in die Projektpartner	Stärkung der Rechenschaftspflicht und (4) Verantwortung für Fehler und Überschreitungen
(3) Unterstützung durch das Management	(5) Verbesserung der Managementfähigkeiten der Mitarbeiter
(4) Vermittlung zwischen Informationsgebern und -nutzern	

vertritt die Meinung, dass Planer in der Rechenschaft ihrer eigenen Planung stehen sollten und ihre eigenen Fähigkeiten erweitern sollten.

Abbildung 2.4 zeigt die Wirkungskette auf, aus der Planungsfehlschlüsse bzw. zeitliche Abweichungen folgern. Auf Basis der Aufgaben- und Systemkomplexität sind zum einen einige Informationen zum Zeitpunkt der Rahmenterminplanung nicht verfügbar. Es besteht eine Ungewissheit und Unwissenheit. Zum anderen ist der Aufwand einer vollständigen und zentralen Projektdokumentation nach Abschluss einzelner Projektphasen sehr aufwändig. Durch Schnittstellen zwischen Organisationen und Softwaresystemen werden Informationen häufig dezentral abgelegt. Folge sind unvollständige Informationen, die eine Unsicherheit in der Planung folgern. Häufig finden daher intuitive Prognosen Anwendung

in der Rahmenterminplanung, die die Entstehung von Planungsfehlschlüssen unterstützen und Abweichungen sind die Folge. Da diese Abweichungen weitreichende Auswirkungen wie. z. B. auf Folgeprozesse oder die Projektatmosphäre existiert das Bedürfnis nach genauen und objektiven Prognoseergebnissen. Um diesen Abweichungen entgegenzuwirken, liegt daher der Forschungsschwerpunkt dieser Arbeit auf erklärbaren analytischen Planungsmethoden. Mit diesen können zum einen relevante Projektmerkmale und Zusammenhänge zwischen diesen Merkmalen aufgedeckt werden. Auf dieser Basis lassen sich Handlungsempfehlungen für eine effiziente Projektdokumentation ableiten, um zukünftig die Informationsgrundlage gezielt zu verbessern und unvollständige Informationen, die in der Rahmenterminplanung relevant sind, zu reduzieren.



**Abbildung 2.4:** Wirkungskette und Folgerung der Anforderungen an die Terminplanung in frühen Phasen eines Bauprojektes

Diese Wirkungskette folgert zwei Anforderungen für analytische Planungsmethoden:

1. **Erklärbarkeit:** Um die unzureichende Datengrundlage zu verbessern, müssen für die frühe Terminplanung relevante Merkmalsausprägungen

identifiziert werden können. Dies resultiert in der Anforderung der Erklärbarkeit. Kann der Einfluss der Merkmalsausprägungen auf die Zeit erklärbar aufgewiesen werden, wird die Datengrundlage verbessert und die Reduktion der Unsicherheit wird unterstützt. So folgt aus der häufigen Auslagerung der Aufgaben des Bauherrn an das Projektmanagement die Anforderung an nachvollziehbare Planungsmethoden. Mit einer erklärbaren Planung kann die Akzeptanz analytischer Planungsmethoden und das Vertrauen in diese erhöht werden. Mit der geschaffenen Akzeptanz und dem Vertrauen wird der Einsatz analytischer Planungsmethoden gestärkt. Die sozialen Kompetenzen werden unterstützt.

2. **Prognosegenauigkeit:** Die Pläne der frühen Phasen sollten möglichst genau sein. Hierfür ist es erforderlich, die „richtigen“ Informationen zu sammeln (datenzentrische Perspektive) und die „richtigen“ modernsten Techniken anzuwenden. Die richtigen Informationen zu dokumentieren, wird durch die Anforderung der Erklärbarkeit unterstützt.

Auch nach Autoren wie Chen und Guestrin (2016, S. 1) haben beide Anforderungen, die Erklärbarkeit und die Prognosegenauigkeit, einen gleichbedeutenden Einfluss auf die Akzeptanz und das Vertrauen in Empfehlungen.

Beide Anforderungen werden im Weiteren genauer ausgeführt.

### **Prognosegenauigkeit als Anforderung**

Eine Projektdurchführung in der geplanten Zeit bedeutet für viele Stakeholder den Erfolg eines Bauprojektes (vgl. Latham 1994, Rwelamila und Hall 1995, Lim und Mohamed 1999, Cox et al. 2003). Insbesondere bei öffentlichen Bauprojekten wird durch deren hohe Sichtbarkeit eine Terminalsicherheit als Erfolg gewertet.

Die Planungsqualität ist demnach vor allem durch die Abweichung der geplanten zu ihrer realisierten Dauer definiert. Aber auch durchgeführte Maßnahmen zur Zielerreichung, zeitliche Puffer und daraus resultierende soziale Spannungen sind Aspekte zur Beurteilung der Planungsqualität.

Anzumerken ist, dass aufgebrauchte Puffer bei einer Analyse historischer Projekte schwer nachzuvollziehen sind. Auch werden in der Regel soziale Spannungen in Projekten nicht oder unzureichend dokumentiert, die auf einen Zeitmangel hindeuten können. Maßnahmen zur zeitlichen Zielerreichung können u. a. über Kostenaufwendungen nachvollzogen werden. Hier ist jedoch eine saubere Dokumentation notwendig, um diese Position u. a. gegenüber Kostenfehlplanungen oder Preissteigerungen abgrenzen zu können. Aus diesen Gründen werden in späteren Datenanalysen insbesondere geplante Dauern und ggf. ihre Abweichungen zur Prognose verwendet.

Diese Abweichungen werden als Prognosegenauigkeit definiert. Die Prognosegenauigkeit kann mit den in Kapitel 5.3.4 beschriebenen Performance-Indikatoren bewertet werden.

### **Erklärbarkeit als Anforderung**

Um die Verlässlichkeit des Modells und das Vertrauen des Anwenders in das Modell zu erhöhen, werden erklärbare Modelle benötigt (vgl. Ribeiro et al. 2016, Hoffman et al. 2018, Anjomshoe et al. 2019). „To make [...] decision[s], users need to be confident that the model will perform well on real-world data, according to the metrics of interest“ (deutsche Übersetzung: Um [...] Entscheidungen treffen zu können, müssen die Nutzer darauf vertrauen können, dass das Modell bei realen Daten in Bezug auf die gewünschten Metriken gut abschneidet). (Ribeiro et al. 2016, S. 1)

Erklärungen in der Planung der Dauern eines Bauprojektes in frühen Phasen sind wichtig, da sie (vgl. Hoffman et al. 2018, S. 198):

- Die Grundlage für die darauffolgenden Prozesse sind,
- Die Kooperation unterstützen (vgl. Brézillon 1997, S. 7). Kooperationen wiederum fördern Erklärbarkeit. „Explanations improve cooperation, cooperation permits the production of relevant explanations“ (deutsche

Übersetzung: Erklärungen verbessern die Zusammenarbeit, Zusammenarbeit ermöglicht die Bereitstellung von relevanten Erklärungen) (Brézillon 1997, S. 7),

- Das Lernen und Verstehen eigener Entscheidungen unterstützen,
- Rahmenbedingungen der Planung aufzeigen können,
- Ausschlüsse transparent und nachvollziehbar aufzeigen (vgl. Miller et al. 2017),
- Auslöser und Gründe für ein unerwartetes Verhalten identifizieren.

Erklärungen stellen visuelle oder textuelle Artefakte dar, die das Ziel haben, ein qualitatives Verständnis über die Zusammenhänge und die Prognose zu erhalten (Ribeiro et al. 2016, S. 1). „[...] explaining consists in exposing something in such a way that it is understandable for the receiver of the explanation - so that he/she improves his / her knowledge about the object of the explanation - and satisfactory in that it meets the receiver's expectations“ (deutsche Übersetzung: [...] Erklären besteht darin, etwas so darzustellen, dass es für den Empfänger der Erklärung verständlich ist - so dass er sein Wissen über den Gegenstand der Erklärung verbessert - und dass es zufriedenstellend ist, da es die Erwartungen des Empfängers erfüllt) (Lacave und Díez 2004, S. 3). Aus diesen Artefakten müssen Anwender eigene Aussagen ableiten können. Erklärbarkeit ist daher eng mit dem Begriff der Interpretierbarkeit verbunden (vgl. Biran und Cotton 2017, Anjomshoae et al. 2019). Die Interpretierbarkeit setzt ein Vorwissen über Kategorien, Beziehungen und Aktivitäten beim Anwender voraus (vgl. Langley et al. 2017, S. 1). Systeme sind interpretierbar, wenn die Prozesse von einem Menschen durch Selbsteinsichten und Erklärungen verstanden werden können (vgl. Biran und Cotton 2017). Die Erklärung des Modells und die Entscheidung zur Auswahl der Dauer eines Bauprojektes bzw. einzelner Phasen sind damit abhängig vom Umfeld und dem jeweiligen Vorwissen des Anwenders.

Mit fortschreitender Entwicklung analytischer Methoden werden auch die mathematischen Modelle immer umfangreicher und komplexer. Entsprechend werden Erklärungen in der Entscheidungsfindung immer wichtiger. (vgl. Voigt und Von dem Bussche 2017)

„Models are more complex and less interpretable than ever... Justifying [their] decisions will only become more crucial“ (deutsche Übersetzung: Die Modelle sind komplexer und weniger interpretierbar als je zuvor... Die Rechtfertigung [ihrer] Entscheidungen wird noch wichtiger werden) (Biran und Cotton 2017, S. 4) So unterstützt und bestätigt auch das Gesetz der allgemeinen Datenschutzbestimmungen (engl. „General Data Protection Regulation“ (GDPR)) das Recht für Erklärungen in maschinell produzierten Ergebnissen (vgl. Carey 2018).

Um aktuelle analytische Planungsmethoden zur Prognose von Dauern mit dem Ziel der Genauigkeit und Erklärbarkeit bewerten zu können, wird im folgenden Kapitel eine Literaturrecherche durchgeführt. Hier werden bestehende analytische Planungsmethoden nach ihrer analytischen Stufe sowie ihrer Untersuchungshäufigkeit kategorisiert. Als Ergebnis der Analyse werden Forschungslücken zu den definierten Anforderungen identifiziert.

# **3 Klassifizierung bestehender Forschungsarbeiten zu analytischen Planungsmethoden für die Terminplanung**

## **3.1 Vorbemerkung**

Es existieren eine Vielzahl an Forschungsarbeiten mit dem Ziel, die Planung von Dauern eines Bauprojekts in frühen Projektphasen möglichst valide gestalten zu können. In diesen Arbeiten sind analytische Methoden beschrieben, mit denen Basis relevanter Merkmale und ihrer Ausprägungen die Dauer einzelner Projektphasen bestimmt werden solln. Damit wird das Ziel verfolgt, Bauprojekte formal zu beschreiben und identifizierte Zusammenhänge auf neue Bauprojekte zur Erhöhung der Planungsgenauigkeit zu übertragen. In den vergangenen Jahren haben sich durch die zunehmende digitale Vernetzung von Datenquellen (u. a. Softwarelösungen, „Internet of Things“) und die zunehmende Rechenleistung neue Möglichkeiten ergeben, eine Vielzahl an Daten auszuwerten und in diesen Mustern zu erkennen. So werden auch in Bauprojekten mehr und mehr Daten dokumentiert und können weiter mit externen Datenquellen verknüpft werden (vgl. Bilal et al. 2016, S. 1). Die Relevanz steigt, Wissen in diesen wachsenden Datenmengen sichtbar zu machen, um wirtschaftliche Vorteile erzielen zu können (siehe Kapitel 2.5.1). Damit haben die Anforderungen der Erklärbarkeit und Planungsgenauigkeit einen direkten Einfluss auf die Wirtschaftlichkeit der am Bauprojekt beteiligten Unternehmen.

In diesem Kapitel wird untersucht, welche analytischen, datenbasierten Planungsmethoden derzeit zur Unterstützung der Rahmenterminplanung in der Forschung betrachtet werden. Hierfür wird das methodische Vorgehen der Literaturrecherche (Kapitel 3.2) beschrieben. Weiter werden diese entsprechend ihrem angestrebten Wettbewerbsvorteil kategorisiert (Kapitel 3.3) und evaluiert (Kapitel 3.4). Abschließend werden relevante Projektmerkmale in den Planungsmodellen identifiziert (Kapitel 3.5) und die wesentlichen Erkenntnisse der Literaturrecherche mit der aufgedeckten Forschungslücke zusammengefasst (Kapitel 3.6). Analytische Methoden für die Aufnahme von Prognosen auf Basis von Erfahrungen der Projektbeteiligten sind damit nicht Fokus dieser Arbeit. Diese können, wie in Kapitel 2.5.5 beschrieben, Grundlage zur Reduktion von Abweichungen sein.

Mit den Ergebnissen kann eine strukturierte Zusammenfassung bestehender analytischer datenbasierter Methoden zur Rahmenterminplanung von Bauprojekten in Abhängigkeit vom angestrebten Wettbewerbsvorteil sowie der Projektmerkmale aufgezeigt werden. Innerhalb der Literaturrecherche konnte keine Veröffentlichung gefunden werden, die bestehende Planungsmethoden entsprechend ihrem angestrebten Wettbewerbsvorteil kategorisiert. Lediglich Kanoglu (2003) und Erdis (2013) kategorisieren Arbeiten nach inhaltlichen Aspekten, ohne eine Zuweisung des wirtschaftlichen Vorteils der jeweiligen Methoden.

## **3.2 Methodisches Vorgehen der Literaturrecherche**

Mithilfe einer systematischen Literaturrecherche wird relevante Literatur zu analytischen Methoden in der Rahmenterminplanung von Bauprojekten im deutsch- sowie englischsprachigem Raum gefiltert. Die hieraus identifizierten Inhalte stellen den Stand der Forschung als Grundlage für die vorliegende Arbeit dar. Die Literaturrecherche erfolgt in drei Phasen:

1. In der ersten Phase werden entsprechend der Zielsetzung Suchwörter definiert.



2. Anschließend werden mithilfe dieser Suchwörter Veröffentlichungen gefiltert. Durch eine vorwärts- und rückwärtsorientierte Suche werden diese Quellen erweitert: In der als relevant deklarierten Literatur wird nach Zitierungen mit inhaltlichen Überdeckungen (rückwärtsorientierte Suche) sowie nach Zitierungen der Quelle selbst (vorwärtsorientierte Suche) gesucht.
3. In der dritten Phase wird ein nach Davenport und Harris (2007) geeignetes Kategorisierungssystem aufgebaut, das den Zusammenhang zwischen der verwendeten Methode und dem angestrebten Wettbewerbsvorteil aufzeigt. Die gefilterten analytischen datengbasierten Methoden aus der Literatur werden entsprechend kategorisiert, um Forschungsschwerpunkte aufdecken zu können. Dieser Teil stellt das Ergebnis des Kapitels und der Literaturrecherche dar. So können abschließend Herausforderungen und Empfehlungen abgeleitet werden.

Die drei Phasen sind in Abbildung 3.1 zusammengefasst dargestellt und im Folgenden beschrieben.

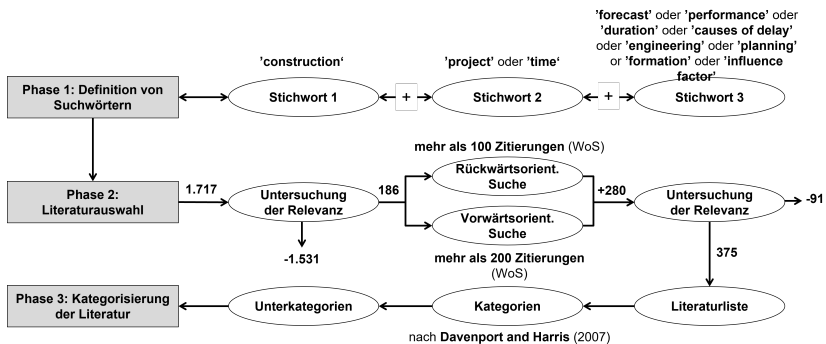


Abbildung 3.1: Vorgehen der Literaturrecherche

## Phase 1: Definition von Suchwörtern

Mit der Definition der Suchwörter werden zwei Ziele verfolgt: Zum einen sollen analytische, datenbasierte Methoden identifiziert werden, die die Dauern von Projektphasen prognostizieren. Zum anderen sind Plan-Ist-Abweichungen relevant, um ein Verständnis über relevante Komponenten zur Bestimmung der Dauern zu erhalten. Zur Sichtung der relevanten Literatur wird daher eine Kombination von drei Suchwörtern gewählt mit:

- 'construction' und
- 'project' oder 'time' und
- 'forecast' oder 'performance' oder 'duration' oder 'causes of delay' oder 'engineering' oder 'planning' oder 'formation' oder 'influence factor'.

In der deutschsprachigen Literaturrecherche wird nach gleichen Suchwörtern in ergänzender Literatur gesucht. Über verschiedene Plattformen <sup>1)</sup> wird durch die Suchwörter in Titel und Kurzzusammenfassung (Abstract) nach Veröffentlichungen in insgesamt 91 unterschiedlichen Journals gesucht. Es werden Dopplungen entfernt, die sich aus den Überschneidungen der Netzwerke ergeben. Insgesamt werden mit den genannten Suchwörtern 1.717 Veröffentlichungen gefiltert.

## Phase 2: Literatúrauswahl

Die identifizierten Veröffentlichungen werden durch eine Bewertung der Titel in relevant und nicht relevant eingeordnet. Nicht relevante Titel werden in Folge ausgeschlossen. Titel sind als nicht relevant definiert, wenn sie keinen Bezug zur Baubranche aufweisen. In einer zweiten Bewertungsanalyse der Kurzzusammenfassung wird nochmals in relevant und nicht relevant unterteilt. Kurzzusammenfassungen gelten als nicht relevant, wenn sie entweder keinen Bezug zur Planung der Dauern in Bauprojekten haben oder nicht zugänglich über den universitären

---

<sup>1)</sup> [www.emerald.com](http://www.emerald.com), [www.sciencedirect.com](http://www.sciencedirect.com), [www.ascelibrary.org](http://www.ascelibrary.org), [www.tandfonline.org](http://www.tandfonline.org), [www.onlinelibrary.wiley.com](http://www.onlinelibrary.wiley.com), [www.pmi.org](http://www.pmi.org), [www.springerlink.com/journals](http://www.springerlink.com/journals)

Zugang bzw. Open-Access-Zugang sind. Wieder wird Literatur ausgeschlossen, wenn diese keinen Bezug zur Baubranche aufweist. Die Literaturliste umfasst nun 186 Veröffentlichungen.

Diese Liste wird durch eine vorwärts- und rückwärtsorientierte Suche erweitert. Hierzu wird die Anzahl der Zitierungen nach dem „Web of Science“ (WoS) als Referenzwert für die Relevanz der Quellen hinzugezogen. Die rückwärtsorientierte Suche orientiert sich an Quellen mit mehr als 100 Zitierungen, bei der vorwärtsgerichteten Literaturrecherche mit mehr als 200 Zitierungen. Somit werden weitere 280 Quellen hinzugefügt. Auch hier wird nicht relevante Literatur nach den oben genannten Kriterien ausgeschlossen (91 Quellen). Es ergibt sich eine Literaturliste mit 375 relevanten Quellen. Tabelle 3.1 fasst das Vorgehen der systematischen Literatúrauswahl zusammen. Die Literaturliste ist nicht als eine vollständige Übersicht der Literatur zur Terminplanung von Bauprojekten zu verstehen, sondern kann als eine umfassende Sammlung verschiedener Autoren betrachtet werden. Die identifizierte Literaturübersicht ist in Anhang A zu finden.

**Tabelle 3.1:** Systematische Literatúrauswahl

	Anzahl der Quellen
1. Literaturliste anhand der definierten Suchwörter	1.717
2. Davon relevante und frei-zugängliche Literatur	186
3. Ergänzende Quellen durch eine systematische vorwärts- und rückwärtsgerichtete Literaturrecherche	+280
4. Von 3. als nicht-relevant eingeordnet	-91
Finale Literaturliste	<b>375</b>

### **Phase 3: Systematische Kategorisierung der Literatur**

Die systematische Kategorisierung der Literatur hat das Ziel, einen Überblick über die in der Forschung verwendeten analytischen Methoden und deren angestrebtem Wettbewerbsvorteil zu erhalten. Hierfür wird die Systematik nach Davenport und Harris (2007, S. 8) verwendet. Die identifizierte Literatur wird den jeweiligen aufeinander aufbauenden Stufen der vorgestellten Systematik zugeordnet, beschrieben sowie ihr angestrebter Wettbewerbsvorteil anhand der Prognosegenauigkeit analysiert. Die Stufen lauten: statistische Analysen, Hochrechnungen, prädiktive Modelle und Optimierungen.

## **3.3 Klassifizierung der analytischen Planungsmethoden**

### **3.3.1 Methoden der statistischen Analyse**

Der „Relative Importance Index (RII)“ sowie „Korrelationen und Sensitivitätsanalysen“ sind zur Planung der Dauern sehr häufig zu finden. Mit diesen Methoden wird eine Diagnose erstellt, welche Merkmale und Merkmalsausprägungen einen Einfluss auf die Dauern von Bauvorhaben haben.

#### **Relative Importance Index**

Im Folgenden werden aufgrund der Vielzahl der Veröffentlichungen in diesem Bereich Quellen aus der Literaturliste mit mehr als 100 Zitierungen betrachtet, die den RII anwenden. Der RII wertet die Relevanz einzelner Merkmale auf Basis von Umfrageergebnissen aus. Die bewerteten Merkmale werden in eine Rangfolge gebracht. Der RII wird hier als Teil analytischer Methoden verstanden, da Erfahrungen quantifiziert und dieser über die Summe der Umfrageteilnehmer als Referenzwert angesehen werden kann. Häufig wird eine Bewertung von eins bis vier (vgl. Assaf und Al-Hejji 2006, El-Razek et al. 2008, Kaliba et al. 2009, Meng 2012) oder eins bis fünf (vgl. Frimpong et al. 2003, Sambasivan und Soon 2007,

Sweis et al. 2008, Doloji et al. 2012), auch Likert-Skala genannt, abgegeben. Zur Berechnung des RII wird je Einflussfaktor die Summe der Bewertungen  $w$  durch die Teilnehmer der Umfrage gebildet und diese durch das Produkt der höchsten Gewichtung der Bewertungsskala  $A$  und der Anzahl an Teilnehmern  $N$  geteilt (vgl. Chan und Kumaraswamy 1997, S. 57):

$$RII = \frac{\sum w}{A * N} \quad (3.1)$$

Lediglich Ogunlana et al. (1996) verwenden hingegen einen offenen Interviewleitfaden, in dem Teilnehmer mögliche Ursachen für einen Zeitverzug nennen sollen. Des Weiteren schränkt Shubham (2013) die Antwortmöglichkeit auf „Ja“ und „Nein“ ein. Werden unterschiedliche Gruppen befragt, so kann ebenfalls der Spearman Rangkoeffizient (vgl. El-Razek et al. 2008, Le-Hoai et al. 2008) oder eine einseitige Varianzanalyse (engl. „Analysis of Variance“ (ANOVA)) (vgl. Sweis et al. 2008) berechnet werden, um die Übereinstimmung von Aussagen der einzelnen Gruppen bewerten zu können.

Aus der Analyse der ausgewählten Literatur ist zu erkennen, dass die Autoren einen Fokus auf die Örtlichkeit des Bauvorhabens und teilweise auf das Produkt setzen. So wird ein RII in der identifizierten Literaturrecherche in folgenden Regionen berechnet: Ägypten (vgl. El-Razek et al. 2008), Ghana (vgl. Frimpong et al. 2003), Hong Kong (vgl. Chan und Kumaraswamy 1997), Indien (vgl. Doloji et al. 2012), Jordanien (vgl. Odeh und Battaineh 2002, Sweis et al. 2008), Kuwait (vgl. Koushki et al. 2005), Malaysia (vgl. Sambasivan und Soon 2007), Nigeria (vgl. Elinwa und Mangvwat 2001), Saudi Arabien (vgl. Assaf und Al-Hejji 2006), Thailand (vgl. Ogunlana et al. 1996), UK (vgl. Meng 2012), Vietnam (vgl. Le-Hoai et al. 2008), Zambia (vgl. Kaliba et al. 2009) und Entwicklungsländer im südostasiatischen Raum (vgl. Shubham 2013). Zum anderen legen einzelne Studien einen Fokus auf spezifische Bauwerkstypen, wie auf den Hochbau allgemein (vgl. Ogunlana et al. 1996, Koushki et al. 2005), den Bau von Wohnungen (vgl. Sweis et al. 2008), die Wasserversorgung (vgl. Frimpong et al. 2003) oder den Straßenbau (vgl. Kaliba et al. 2009).

Bei einem Großteil der Studien werden Bauherren, Projektmanager, Planer und Bauunternehmen befragt. Die Anzahl der Teilnehmer reicht von 26 Teilnehmern

(vgl. Elinwa und Mangvwat 2001) bis 450 (vgl. Koushki et al. 2005). Beim Bauherren (und dem Projektmanagement) werden häufig die folgenden wichtigsten Gründe für zeitliche Abweichungen genannt: Späte Änderungen (vgl. Elinwa und Mangvwat 2001, Odeh und Battaineh 2002, Koushki et al. 2005, Assaf und Al-Hejji 2006, Sweis et al. 2008, El-Razek et al. 2008, Kaliba et al. 2009), geringe Erfahrungen der Bauherren in Bauprojekten (vgl. Koushki et al. 2005), eine Fehlplanung des Termin- und Kostenrahmens (vgl. Elinwa und Mangvwat 2001, Shubham 2013) sowie eine langsame Entscheidungsfindung (vgl. Chan und Kumaraswamy 1997). Planer übergeben durch den hohen Zeitdruck unvollständige Pläne (vgl. Ogunlana et al. 1996, Elinwa und Mangvwat 2001, Sambasivan und Soon 2007). Bei Bauunternehmen sind die am meisten genannten Gründe: ein unzureichendes Baustellenmanagement und die Baustellenkoordination (vgl. Chan und Kumaraswamy 1997, Frimpong et al. 2003, Sambasivan und Soon 2007, Le-Hoai et al. 2008, Sweis et al. 2008, Doloi et al. 2012, Shubham 2013), Material- und Personalengpässe (vgl. Ogunlana et al. 1996, Frimpong et al. 2003, Kaliba et al. 2009, Shubham 2013) sowie die geringe Erfahrung der Bauunternehmen (vgl. Odeh und Battaineh 2002, Sambasivan und Soon 2007) oder Probleme mit dem Baugrund (vgl. Chan und Kumaraswamy 1997). Weiter wurden Finanzierungsprobleme einzelner Parteien (vgl. Elinwa und Mangvwat 2001, Odeh und Battaineh 2002, Frimpong et al. 2003, Koushki et al. 2005, Sweis et al. 2008, El-Razek et al. 2008, Le-Hoai et al. 2008) und eine geringe Integration der Projektbeteiligten (vgl. Meng 2012) aufgezählt.

Aus dem RII sind keine direkten Prognosen ableitbar, es können lediglich Muster verdeutlicht werden. Diese Muster sind durch den gesetzten regionalen oder inhaltlichen Fokus nicht allgemein auf andere Bauprojekte übertragbar. Die identifizierten Faktoren bieten Folgeprojekten lediglich Anhaltspunkte, auf die in der Planung und Durchführung von Bauvorhaben zu achten sind. Dennoch zeigt sich, wie oben beschrieben, dass viele der genannten Gründe bereits in einer frühen Phase des Bauprojektes verankert sind: Der Bauherr hat entsprechend den Ergebnissen der Literaturrecherche Schwierigkeiten, einzelne Anforderungen an

das Bauvorhaben zu benennen und deren Auswirkungen für die folgende Planung aufzeigen zu können. Dies hat u. a. späte Änderungen, einen unrealistischen Termin- und Kostenplan oder langsame Entscheidungen zu Folge.

### **Korrelationen und Sensitivitätsanalysen**

Weitere neben dem RII verwendete Methoden der Diagnose sind „Korrelationen und Sensitivitätsanalysen“, die in die statistischen Analysen einzuordnen sind. Eine Korrelation zeigt allgemein einen negativen und positiven linearen Zusammenhang zwischen zwei Variablen auf. Häufig verwendet werden der Spearman und der Pearson Korrelationskoeffizient. Diese liegen zwischen Minus eins und Plus eins. Je näher der Korrelationskoeffizient an Minus eins bzw. Plus eins liegt, desto stärker kann auf einen negativen bzw. positiven Zusammenhang geschlossen werden. Die Zusammenhänge beweisen jedoch keine Kausalität. Walker (1995) erstellt Korrelationen zwischen einzelnen Faktoren und der Dauer der Baurealisierung. So zeigt er auf, dass die Performance des Baumanagements einen Einfluss auf die Dauer hat. Kaming et al. (1997) hingegen analysieren durch die Bildung von Korrelationen den Einfluss einzelner Faktoren auf den Zeitverzug. Die Sensitivitätsanalyse zeigt weiter die Empfindlichkeit der Dauer bei einer Variation der Einflussfaktoren auf. So kann u. a. bewertet werden, wie empfindlich die Dauer zur Realisierung des Bauwerks bei einer Variation der Kosten ist (vgl. Kumaraswamy und Chan 1995). Bromilow (1969) sowie Kumaraswamy und Chan (1995) führen eine Sensitivitätsanalyse durch und können einen Zeit-Kosten-Zusammenhang identifizieren. Korrelationen und Sensitivitätsanalysen sind durch die quantitative Bewertung häufig als Vorstudien für Hochrechnungen zu finden.

### **3.3.2 Hochrechnungen**

Hochrechnungen umfassen Methoden der Stochastik und beziehen sich damit auf Modelle der Statistik und der Wahrscheinlichkeitsrechnung. Mit diesen werden

Aussagen zur Grundgesamtheit auf Basis einer Stichprobe gesucht, um Zusammenhänge zwischen den beschreibenden Daten (Merkmalen) und Ergebnisvariablen aufzuzeigen. Hochrechnungen formulieren damit die Abhängigkeit zwischen den Merkmalen und ihren Merkmalsausprägungen zu der Ergebnisvariablen, der Dauer von Projektphasen. Dabei wird davon ausgegangen, dass diese Zusammenhänge über die Zeit konstant sind und der Trend fortbesteht. „Multivariate Regressionen“, „Zeitreihenanalysen“, „Fuzzy-Systeme“ oder „Methoden der bayesschen Statistik“ sind häufig verwendete Beispiele angewandeter Methoden für Hochrechnungen.

## Multivariate Regression

Ziel dieser Methode ist es, eine Funktion zu finden, deren Funktionswerte möglichst genau den Mittelwerten einer abhängigen Variablen (Ausgabe,  $y$ ) entsprechen in Abhängigkeit von einer oder mehrerer unabhängiger Variablen (Eingabe,  $x$ ) entsprechen (vgl. Quadrianto und Buntine 2010, S. 838).

Bei einer linearen Regression wird dazu eine Ausgleichsgerade  $y = a_0 + a_1x$  gesucht, sodass die untersuchten Datenpaare  $(x_i, y_i)$  möglichst genau mit der Gerade dargestellt werden können. Verfügt eine Problemstellung über mehrere unabhängige Eingabevariablen, so spricht man von multipler oder multivariater Regression (vgl. Russell und Norvig 2012, S. 839). So ist es auch möglich, nicht lineare Funktionen zu approximieren, bspw. Polynome. Das Ausgleichspolynom  $m$ -ter Ordnung hat die allgemeine Formel

$$y = a_0 + a_1x + a_2x^2 + \dots + a_mx^m \quad (3.2)$$

mit den zu bestimmenden Parametern  $a_0, \dots, a_m$  (vgl. Kruse et al. 2015, S. 481 f.).

Der Wertebereich von  $f(\mathcal{X})$  ist in diesem Fall kontinuierlich und ein Element von  $Y \in f(\mathcal{X})$  wird als Antwort bezeichnet (vgl. Knox 2018, S. 4). Die algebraische Methode der kleinsten Quadrate dient zur Anpassung der Regressionsgeraden (zum Beispiel von Ertel (2016, S. 479 ff.) beschrieben). Diese Aufgabe, eine



passende Ausgleichsfunktion zu finden, übernehmen beim Maschinellen Lernen (ML, siehe Kapitel 5) die angewendeten Algorithmen.

Multivariate Regressionsmodelle finden in der identifizierten Literatur im Bereich der Hochrechnungen die häufigste Anwendung. Eines der ersten Modelle für eine Schätzung der Zeit zur Bauwerksrealisierung wurde 1969 durch den australischen Forscher Bromilow (1969) entwickelt, mit:

$$T = K * C^B. \quad (3.3)$$

Dabei ist  $T$  die Zeit (in Arbeitstagen),  $K$  ist eine Konstante für die Projektperformance,  $C$  die Projektkosten laut Vertrag (in Millionen Dollar) und  $B$  eine Konstante, die die Sensitivität der Zeit im Verhältnis zu den Kosten wiedergibt. Auf dem Modell von Bromilow aufbauend sind weitere Mehrfaktoren-Modelle zu finden. Zeitschätzungsmodelle mit dem Faktor Kosten als Basis sind daher am weitesten verbreitet. Abbildung 3.2 zeigt Graphen ausgewählter Hochrechnungen für Projekte im Hochbau. So ergibt sich die Zeit zur Realisierung eines Bauwerks nach Ireland (1983) aus 25 Projekten im Hochbau in Australien mit  $T = 219C^{0,47}$ . Chan hingegen kalkuliert die Zeit zur Bauwerkserstellung aus 110 Projekten mit  $T = 152C^{0,29}$  in Hongkong (Chan 1999) und aus 51 öffentlichen Bauprojekten mit  $T = 269C^{0,32}$  in Malaysia (Chan 2001). Weiter ergibt sich nach Ogunsemi und Jagboro (2006) die Bauzeit in Nigeria auf Basis von 87 Bauprojekten mit  $T = 563C^{0,262}$ . Zu erkennen ist dabei die große Varianz der Formeln. Deutlich wird dies u. a. zwischen Choudhury und Rajan (2003) mit  $T = 18.96C^{0,39}$  (55 Wohngebäude in Texas) und Le-Hoai und Lee (2009) mit  $T = 369C^{0,158}$  (34 Wohngebäude in Korea). Bei einer Million Dollar Investitionskosten würde ein Bauprojekt in Texas 4.148 Arbeitstage dauern, in Korea hingegen lediglich 3.274 Tage. Dies entspricht etwa einem Unterschied von 25 %.

Viele Fallbeispiele zeigen weiter durch ihre starken Abweichungen zwischen dem realen und prognostizierten Wert, dass die Kosten nur einen Anteil der Prognosegenauigkeit an der jeweiligen Phase haben (vgl. Flyvbjerg et al. 2002, Potts 2005, Magnussen und Olsson 2006). Um den Schätzungsfehler zu reduzieren,

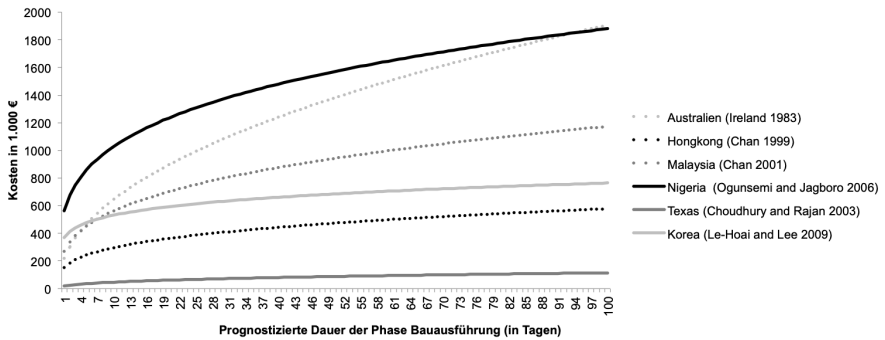


Abbildung 3.2: Vergleich ausgewählter multivariater Regressionen verschiedener Autoren

nehmen einzelne Autoren weitere Faktoren neben den Baukosten in die Regression mit auf. Irfan et al. (2011) nehmen so bei der Erstellung einer multivariaten Regression neben den Kosten weitere Merkmale, wie den Projekt- sowie Vertragstyp mit auf und können dadurch die Prognosegenauigkeit erhöhen.

### Box Jenkins Methode

Die Box Jenkins Methode bezieht historische Schätzungsfehler in einer Zeitreihenanalyse mit ein, um Vorhersagen für einen korrigierten Trend zu treffen (vgl. Zhang und Koreisha 2015). Es wird dabei von Schätzungsfehlern ausgegangen, die in den folgenden Hochrechnungen mit einbezogen werden. Regressionen sind aufgrund des begrenzten Datensatzes sehr sensitiv zur Zielvariable und können mit der Box Jenkins Methode robuster gestaltet werden (vgl. Zhang und Koreisha 2015, S. 4827). So kann trotz einer Volatilität die Konstanz des darunterliegenden Prozesses erkannt werden (vgl. Agapiou et al. 1998). Lu und AbouRizk (2009) wenden die Box Jenkins Methode bspw. zur Bestimmung der Dauer für die Bauwerksrealisierung bei großen Infrastrukturprojekten an und erstellen daraus eine Vorlage für MATLAB (MATrix LABoratory).

## Fuzzy-Systeme

Um die Unsicherheiten in einem Bauvorhaben darstellen zu können, werden Fuzzy-Systeme hinzugezogen. „Bei der Planung von Projekten aus dem wirklichen Leben müssen die Planer oft zeitliche Zwänge berücksichtigen (fixierte Meilensteine aufgrund externer Ereignisse) und eine größere Bandbreite an Vorangbedingungen.“ (vgl. Bonnal et al. 2014, S. 115) „Fuzzy-Systeme“ wurden im Jahre 1965 durch den iranischen Systemtheoretiker Lotfi Zadeh erfunden (vgl. Zadeh 1965). Sie dienen der Modellierung komplexer Systeme mit einer Akzeptanz der Unschärfe, bei der auch Zwischengrößen zwischen „null“ und „eins“ bzw. „ja“ und „nein“ möglich sind. So können z. B. formal die Zwischengrößen zwischen einem großen und kleinen Bauprojekt ausgedrückt werden. Die Unschärfe wird durch sprachliche Schlussfolgerungen und Aussagen modelliert. „Fuzzy logic is more like human thinking because it is based on degrees of truth and uses linguistic variables“ (deutsche Übersetzung: Die unscharfe Logik ist dem menschlichen Denken ähnlicher, da sie auf Wahrheitsgraden basiert und linguistische Variablen verwendet) (Kayacan und Khanesar 2016, S. 14). Die Eingabe in ein Fuzzy-System ist immer eindeutig (engl. „crisp“). Mithilfe von Fuzzy-Werten wird die Eingabe unscharf umformuliert (engl. „Fuzzifizier“). Auf dieser Basis kann ein Experte Entscheidungen mit Wenn-Dann-Regeln treffen. Das Ergebnis wird anschließend in einen scharfen Wert, wie z. B. die Projektdauer umgewandelt (engl. „Defuzzifizier“) (vgl. Kayacan und Khanesar 2016, S. 18).

Die erste Veröffentlichung, die sich mit einer Fuzzy-Terminplanung befasste, stammt von Chanas und Kamburowski (1981). Chanas und Kamburowski (1981) führen eine „Program Evaluation and Review Technique“ (PERT) ein, mit dem Ziel Unsicherheiten abzubilden. 1994 veröffentlichten Wu und Hadipriono (1994) ein simples Computerprogramm zur Anwendung von Fuzzy-Systemen. Mit ihrem Programm „Activity Duration Decision Support System“ (ADDSS) muss ein Experte lediglich markieren, welche Merkmale für seine Baustelle relevant sind und sie mit linguistischen Werten wie „sehr gut“, „ziemlich gut“, „unentschieden“, „ziemlich schlecht“ usw. beschreiben (vgl. Wu und Hadipriono 1994, S. 170). Das ADDSS-Programm zeigt, dass „die Verwendung von Fuzzy-Logik zur Bestimmung von Aktivitätsdauern machbar ist“ (Wu und Hadipriono

1994, S. 178). Hapke et al. (1997) befassen sich in mehreren Forschungsbeiträgen mit der Anwendung der Fuzzy-Logik in der Terminplanung. Darunter ist auch ihr Artikel aus dem Jahr 1997, in dem sie die Fuzzy-Logik über Probleme mit unscharfen Daten recherchierten. Folgend ist die Fuzzy-Logik in weiteren wissenschaftlichen Veröffentlichungen zu finden, wie z. B. bei Kumar und Reddy (2005).

### **Monte-Carlo-Simulation**

Die Monte-Carlo-Simulation ist Teil der Bayesschen Statistik und basiert damit auf Wahrscheinlichkeitsverteilungen. Es wird auf Basis des Vorwissens und der Datengrundlage die Plausibilität einer Aussage aufgezeigt. Um eine möglichst genaue Wahrscheinlichkeitsverteilung zu erhalten, werden eine große Datenmenge und damit auch hohe Rechenleistungen benötigt. Mit der Monte-Carlo-Simulation wird eine große Anzahl gleicher Zufallsexperimente generiert. So kann mithilfe der Wahrscheinlichkeitstheorie ein schwer lösbares Problem simuliert werden. Albogamy et al. (2014) zeigen mithilfe einer Monte-Carlo-Simulation wahrscheinliche Risiken in einer frühen Planungsphase auf, die Auswirkungen auf Zeit und Kosten haben können. Moret und Einstein (2016) modellieren mit einer Monte-Carlo-Simulation Unsicherheiten auf die Kosten und Zeit beim Bau von Eisenbahninfrastrukturprojekten.

Die Ergebnisse der genannten Methoden sind auf einem fortlaufenden Trend aufgebaut. Die folgenden „prädiktiven Methoden“ haben das weitere Ziel, unbekannte Muster zu erkennen und auf neue Datensätze zu übertragen, die den weiteren Verlauf beeinflussen können.

### **3.3.3 Prädiktive Modelle**

Die Relevanz prädiktiver Analysen nimmt in einem dynamischen Umfeld zu. Unbekannte Muster müssen erkannt werden, die im weiteren Verlauf von Bedeutung sein können. Es wird demnach kein einzelnes Ergebnis berechnet, sondern eine Auswahl von Zielzuständen und jeweiligen Treibern für diese Zustände. Unter

prädiktiven Methoden ist die Übertragung der gefundenen Muster aus den Trainingsdaten auf neue Datensätze zu verstehen.

Prädiktive Methoden umfassen, aufbauend auf den vorherigen Kategorien, „Methoden des Maschinellen Lernens“. Im Vergleich zu Methoden der Stochastik konzentrieren sich ML-Methoden auf die Vorhersage des Ergebnisses, indem mithilfe von allgemein anwendbaren Lernalgorithmen Muster in Datensätzen gefunden werden. Durch die Aufteilung in Test- und Trainingsdaten werden Vorhersagen zu zukünftigen Daten getroffen. (vgl. Bzdok et al. 2018)

Auf den Themenbereich der KI mit dem Teilbereich ML wird in Kapitel 2.4 weiter eingegangen. Mit bekannten Beispieldaten wird hier auf eine bestimmte Ausgabe und die damit einhergehende Interpretation trainiert. Anschließend können die gefundenen Regeln verwendet werden, um neue Ausgaben zu erstellen. (vgl. Kirste und Schürholz 2019, S. 25)

Vor allem sind in der Literaturrecherche „KNN“ (siehe nachfolgend) zur Prognose der Dauern in Bauprojekten zu finden.

### **Künstliche Neuronale Netze**

Die Funktionsweise von KNN wird in Kapitel 5.4 genauer erläutert. In bestehenden Veröffentlichungen zur Terminplanung von Bauprojekten wird häufig zwischen der Prognosegenauigkeit von Hochrechnungen und von KNN verglichen. Hier zeigen die in diesem Kapitel genannten Autoren Dissanayaka und Kumaraswamy (1999), Chen und Huang (2006), Petrusseva et al. (2012), Lam und Olalekan (2016) eine deutlich bessere Aussagekraft der KNN als Teil der prädiktiven Analysen auf.

Petruseva et al. (2012) erstellen eine multivariate Regression sowie ein KNN auf Basis von 75 Bauprojekten in Bosnien Herzegowina und elf beeinflussenden Faktoren (u. a. Verwendung, Baujahr, beauftragte Zeit, Bauzeit real, Unterschied (Zeit), beauftragte Kosten, Baukosten real, Unterschied (Kosten)). Vergleicht man die Prognosegenauigkeit beider Modelle, so zeigt hier das KNN bessere Ergebnisse. Auch Dissanayaka und Kumaraswamy (1999) und Chen und

Huang (2006) zeigen ähnliche Ergebnisse auf. Beide empfehlen im Vergleich zu multivariaten Regressionen die Anwendung von KNN zur Vorhersage der Projektdauern bei Bauvorhaben. Lam und Olalekan (2016) vergleichen die Güte von KNN zu der Box Jenkins Methode bei Bauprojekten in Hongkong. Auch hier werden mit dem KNN bessere Vorhersageergebnisse erzielt. Pewdum et al. (2009) erstellen weiter ein KNN auf Basis von 51 Straßenbauprojekten in Thailand, Mensah et al. (2016) auf Basis von 22 Straßenbauprojekten und Bhokha und Ogunlana (1999) ausgehend von 136 Hochbauprojekten in Bangkok. KNN werden in den selektierten Veröffentlichungen außerdem angewendet, um einzelne Prozesse innerhalb der Baurealisierung vorherzusagen und betrachten damit eine detailliertere Stufe der Terminplanung (u. a. Mauerarbeiten, Bodenlegen, Malerarbeiten (Naik und Kumar 2013), Rohbauaktivitäten (Golizadeh et al. 2015)).

### **Maschinelle Lernmethoden im Ensemble**

Weiter werden Maschinelle Lernmethoden kombiniert. Sie bilden ein sogenanntes Ensemble. Eine der verwendeten Methoden im Ensemble in der untersuchten Literatur ist dabei immer ein KNN. Zu den weiteren Methoden gehören u. a. Entscheidungsbäume und die Stützvektormethode (engl. „Support Vector Machines“ (SVM)). SVMs bilden die Daten auf einer hochdimensionalen Repräsentation ab. In dieser Repräsentation wird eine Hyperebene in ein möglichst breites Band, in dem keine Datenpunkte liegen, als Entscheidungsgrenze zwischen einzelne Klassen gelegt (vgl. Chollet 2018, S. 37). Dieses Band wird durch Datenpunkte an dessen Rändern begrenzt (Stützvektoren), die den Algorithmus maßgeblich prägen (vgl. Morik 2018, S. 23). SVMs benötigen wenige Stützvektoren, was sie zu kompakten Modellen mit geringem Speicherverbrauch macht, der Rechenaufwand beim Lernen ist jedoch vergleichsweise hoch (vgl. VanderPlas 2016, S. 420).

Auf Entscheidungsbäume wird in Kapitel 5.6.2 weiter eingegangen. Erdis (2013) vergleicht den Einfluss einer neuen Gesetzgebung in der Türkei auf die Zeit- und Kosteneinhaltung mithilfe von zwei Datensätzen. Der erste Datensatz zur alten Gesetzgebung beinhaltet 878 Projekte und der zweite mit der neuen Gesetzgebung 575 öffentliche Bauprojekte. Die Gesetzgebung hat das Ziel, öffentliche

Ressourcen effizienter anzuwenden. Maschinelle Lernmethoden zeigen den Vorteil der neuen Gesetzgebung auf. Hier wurden ein Entscheidungsbaum, ein KNN und SVM als Ensemble verwendet, um die Prognosegüte weiter zu verbessern. Wang und Chan (2012) erstellen ein Ensemble aus einem KNN und SVM, um die geplante Zeit zur Baurealisierung in einer frühen Planungsphase validieren zu können. Beide Autoren zeigen auf, dass ein Ensemble bessere Lösungen liefert als ein KNN allein.

### 3.3.4 Optimierung

Zur analytischen Optimierung der Terminplanung werden den Anwendern aktiv Handlungsalternativen vorgestellt, um die bestehende Situation zu optimieren oder einzelne Handlungen auszuschließen. So können den Anwendern auf Basis der verfügbaren Daten Wechselwirkungen zwischen Merkmalsausprägungen und der Dauer oder Kosten dargestellt werden. „Heuristiken“ sind dabei ein Teilgebiet des Operation-Research mit dem Ziel der Optimierung. Sie finden zur Rahmenterminplanung aufgrund der vielen Unsicherheiten Anwendung. Unter einem akzeptablen Rechenaufwand finden sie eine gute, aber nicht unbedingt die optimale, Lösung (vgl. Nickel et al. 2014). Während sich heuristische Methoden auf die Anwendung von Optimierungstechniken und entscheidungstheoretischen Ansätzen konzentrieren, sind ML-Methoden empirisch ausgerichtet und verwenden regelhafte relationale Strukturen und Zusammenhänge (vgl. Bockmayr und Radermacher 1993, S. 249).

Eine besondere Form der Heuristik ist die „Metaheuristik“. Sie beschreibt eine problemunspezifische Heuristik, die daher eine breit gefächerte Anwendung findet. Sie greift auf die Idee der lokalen Suche zurück und versucht ausgehend von einer existierenden zulässigen Lösung in der näheren Umgebung, auch Nachbarschaft genannt, eine bessere Lösung zu finden (vgl. Nickel et al. 2014, S. 218). Im Folgenden werden die wichtigsten Metaheuristiken näher erläutert.

## Genetische Algorithmen

Genetische Algorithmen sind zurückzuführen auf die von Charles Darwin entwickelte Evolutionstheorie unter dem Grundsatz „Survival of the fittest“. Es werden nur die jeweiligen Iterationen bzw. Generationen mit den besten Lösungen übernommen. Ein besonderes Merkmal der genetischen Algorithmen ist ihre problemunspezifische Anwendungsmöglichkeit, wodurch genetische Algorithmen in vielen Themengebieten eingesetzt werden. Beispielhaft sei der Algorithmus von Zheng et al. (2004) zu nennen. Zheng et al. (2004) wenden einen genetischen Algorithmus an, um für die beiden Zielvariablen Zeit und Kosten ein Minimum zu finden und damit in Entscheidungen zu unterstützen.

## Ameisenalgorithmen

Der Ameisenalgorithmus ist, wie auch die genetischen Algorithmen, eine populationsbasierte Suchtechnik, die zum Lösen von schwierigen kombinatorischen Optimierungsproblemen verwendet wird. Der erste Ameisenalgorithmus, auch als „Ant System“ bekannt, wurde von Dorigo und Gambardella (1997) eingeführt, um das „Travelling Salesman Problem“ (TSP) zu lösen. Hierbei wird das natürliche Verhalten der Ameisen bei der Futtersuche modelliert. Sobald eine der Ameisen eine Futterquelle gefunden hat, hinterlässt sie ein Pheromon, um die anderen Ameisen ihres Stammes darauf aufmerksam zu machen und ihr zu folgen. Der Pheromonspiegel steigt mit Anzahl der Ameisen und Wege werden attraktiver. (vgl. Dorigo und Gambardella 1997, S. 53f.)

Bezogen auf den Optimierungsprozess repräsentieren die Ameisen zulässige Startlösungen und ordnen den einzelnen Zielfunktionswerten durch den Pheromonspiegel eine bestimmte Gewichtung zu, die iterativ aktualisiert wird.

El-Gafy (2007), Zhang und Ng (2012) und Kalhor et al. (2011) wenden den Ameisen-Algorithmus an, um die Kosten und Dauer eines Bauprojektes zu optimieren. Kalhor et al. (2011) kombinieren außerdem den Ameisen-Algorithmus mit dem Fuzzy-System, um unscharfe Parameter miteinbeziehen zu können.



## **Tabusuche**

Die Tabusuche (engl. „Tabu search“) untersucht ausgehend von einer zulässigen Lösung die Nachbarschaft (vgl. Nickel et al. 2014, S. 219). Dabei werden Nachbarschaftslösungen durch vordefinierte Suchregeln in jeder Iteration neu erstellt und auf ihre Zulässigkeit geprüft. Die Suchregeln zielen auf die Veränderung einer einzelnen Variable der ursprünglichen Lösung ab. Innerhalb der Nachbarschaft wird dann die Lösung mit dem besten Zielfunktionswert für die weitere Tabusuche ausgewählt. Um nicht in einem lokalen Optimum zu bleiben, wird die Tabu-Liste eingeführt. Diese beinhaltet zulässige und bereits untersuchte Lösungen, die für eine gewisse Anzahl an Iterationen ausgeschlossen werden (vgl. Nickel et al. 2014, S. 220). Die Tabusuche endet, sobald eine Lösung keine zulässige Nachbarschaft mehr besitzt, eine bestimmte Anzahl an Iterationen durchgeführt oder ein bestimmter Zielfunktionswert erreicht wird. Hier sei die Tabusuche von Jung et al. (2016) zu nennen. Die Autoren erstellen eine Tabusuche, um den Einfluss des Wetters auf mögliche Verzögerungen der Baurealisierung in einer frühen Planungsphase zu analysieren.

## **Simulated Annealing**

Das „Simulated Annealing“-Verfahren beschreibt den werkstofflichen Abkühlungsprozess von Metallen beim Glühen (vgl. Nickel et al. 2014, S. 218). Es wurde erstmals von Kirkpatrick et al. (1983) in ihrer Ausarbeitung „Optimization by Simulated Annealing“ eingeführt und wird heute hauptsächlich bei komplexen kombinatorischen Optimierungsproblemen angewendet. Das Grundprinzip des Algorithmus verläuft wie folgt: Solange eine hohe Temperatur vorliegt, ist auch die Wahrscheinlichkeit hoch, eine schlechtere Lösung zu akzeptieren. Mit dem Sinken der Temperatur nimmt auch die Wahrscheinlichkeit ab, eine schlechtere Lösung zu akzeptieren (vgl. Bettemir und Sonmez 2014, S. 3). Somit können lokale Optima bis zu einer gewissen Wahrscheinlichkeit (bzw. der Temperatur) verlassen werden (vgl. Kumar und Abdullah 2011, S. 491). Kumar und Abdullah (2011) verwenden eine Kombination aus Simulated Annealing und Fuzzy-System, um die Dauer und Kosten eines Bauprojektes unter verschiedenen Risikostufen zu optimieren. Bei einem hohen Risiko werden

in dieser Analyse demnach auch hohe Dauern und Kosten akzeptiert.

Neben Heuristiken können für die Bauausführung weitere optimierende Modelle gefunden werden. So wenden Fischer und Aalami (1996) ein System zur Optimierung der Bauwerksrealisierung an. Hier werden Prozesse schrittweise verfeinert, um Verfahren und deren Produkte bereits in der Planung zu beeinflussen. Ziel ist eine bewusste Verfahrensauswahl, um einen optimalen Bauablauf zu erstellen. Die erstellten Regeln benötigen detaillierte Informationen (z.B. über einzelne Aktivitäten während der Baurealisierung), keine Unsicherheiten und können nicht mit einer wachsenden Datenmenge umgehen. Dieses Modell ist damit hoch spezialisiert.

### **3.4 Evaluierung der kategorisierten Planungsmethoden und Aufzeigen der Forschungslücke**

Die Klassifizierungssystematik nach Davenport und Harris (2007) bildet sich aus den vier Stufen der statistischen Analyse, Hochrechnungen, prädiktiven Methoden und Optimierungen (siehe auch Abbildung 2.3 in Kapitel 2.4). Damit gewinnt die analytische Methode an Intelligenz und angestrebtem Wettbewerbsvorteil für beteiligte Unternehmen.

Die Zuordnung der identifizierten Veröffentlichungen in die einzelnen Kategorien zeigt eine quantitative Verteilung dieser und damit den Fokus der aktuellen Forschung auf. Dabei ist die Literaturliste als nicht vollständig anzusehen, sondern zeigt vielmehr Tendenzen auf. Wie im Histogramm in Abbildung 3.3 zu sehen ist, nimmt die Anzahl an Veröffentlichungen mit zunehmender Intelligenz und dem angestrebten Wettbewerbsvorteil ab. Mögliche Gründe hierfür könnte eine geringe Expertise in der Baubranche im Bereich prädiktiver Methoden und

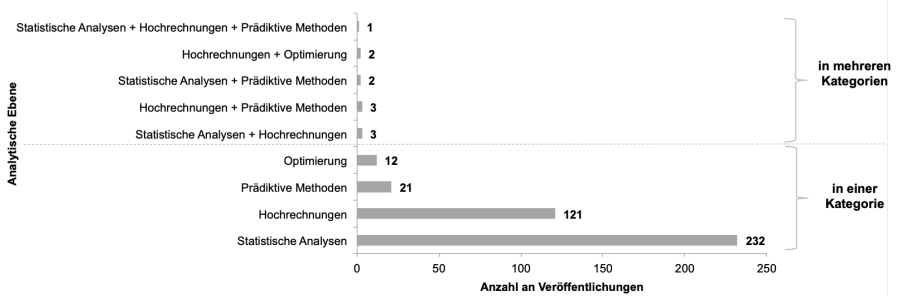


Abbildung 3.3: Anzahl der identifizierten Literaturquellen in den Kategorien

Optimierungen sein. Auch ein fehlender Druck für Veränderung durch eine gute Marktsituation oder geringe digitalisierte und strukturierte Datensätze können Ursachen sein.

Bei der Zuteilung der 375 Veröffentlichungen in die vier Kategorien sind elf Veröffentlichungen in mehreren Kategorien zu finden. Aus diesem Grund entspricht die Summe in den vier Kategorien nicht den eingangs selektierten 375 Veröffentlichungen, sondern 386. Insbesondere werden in diesen elf Veröffentlichungen die prädiktiven Methoden mit Methoden mit einer geringeren Intelligenz, den statistischen Analysen und Hochrechnungen, verglichen. In diesen Veröffentlichungen wird die Prognosegenauigkeit und damit der angestrebte Wettbewerbsvorteil mit der Analytikstufe nach Davenport und Harris (2007) bestätigt.

Abbildung 3.4 zeigt die kumulierte Anzahl der Veröffentlichungen je Kategorie. Hier kann kein klarer Trend hin zu einer Zunahme von Methoden mit einer hohen analytischen Stufe beobachtet werden. Obwohl die verfügbare Rechnerleistung und die Datenmenge in Bauprojekten stetig zunimmt, steigen die Anzahl der Veröffentlichungen zu Methoden mit einem niedrigen Reifegrad weiter stark. So nimmt die Anzahl statistischer Analysen am stärksten zu. Die Entwicklung der Forschungen zu Methoden mit einem hohen Reifegrad verhält sich nahezu konstant und die kumulierte Anzahl nimmt etwa linear zu.

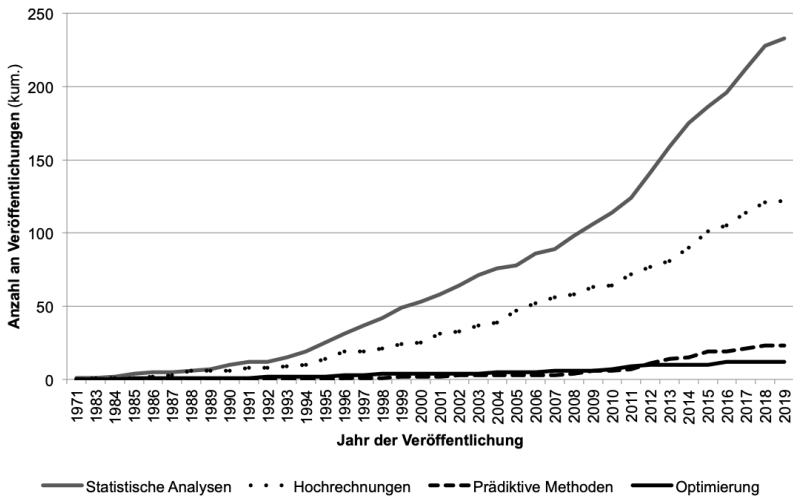


Abbildung 3.4: Anzahl verfügbarer Publikationen nach Jahr in der entsprechenden Kategorie

Weiter kann in den Veröffentlichungen nach dem Fokus der Prognose auf eine einzelne Phase oder auf das Gesamtprojekt unterschieden werden. Die meisten Veröffentlichungen konzentrieren sich auf die Phase der Baurealisierung (307 Veröffentlichungen), auf das Gesamtprojekt (58 Veröffentlichungen) und auf vorgelagerte Projektphasen, wie den Einkauf (vgl. Migliaccio und Shrestha 2009, Dixit et al. 2018) oder die Planung (vgl. Yang und Wei 2010, Liu et al. 2012). Die Phase der Bauwerksrealisierung scheint durch die hohe Varianz zwischen Bauprojekten von besonderem Interesse zu sein.

Mit der Kategorisierung bestehender Forschungsarbeiten kann aufgezeigt werden, dass analytische Planungsmethoden mit einem hohen Grad an Intelligenz vergleichsweise in geringem Umfang untersucht werden. Trotz der Datenzunahme in der Baubranche ist kein zunehmender Trend in diesem Bereich erkennbar ist.

## 3.5 Klassifizierung relevanter Projektmerkmale

Für das Ziel möglichst genaue Prognosen durchführen zu können, ist es essenziell Merkmale mit einem hohen Einfluss auf die Zielvariable zu bestimmen. Es werden hier Zusammenhänge zwischen Merkmalen und der Dauer gesucht. Herbsman und Ellis (1990), Chan und Kumaraswamy (2002), Nasirzadeh und Nojedehi (2013) sowie Abd El-Karim et al. (2017) stellen in ihren Publikationen Übersichten über beeinflussende Merkmale im Bauwesen zusammen. Obwohl die als relevant identifizierten Merkmale Gemeinsamkeiten aufweisen, unterscheiden sie sich aufgrund der Vielzahl möglicher Merkmale in vielen Aspekten. Da die Auswahl relevanter Merkmale entscheidend für die Prognosegenauigkeit der Dauer ist, werden in diesem Unterkapitel relevante Merkmale in der selektierten Literatur identifiziert, kategorisiert und hinsichtlich ihrer Anzahl an Nennungen in Publikationen bewertet. Auf diese Weise kann anschließend die Relevanz der Merkmale in den zugrundeliegenden Datensätzen bewertet werden. Die Anzahl der Nennungen dient dabei als Näherungswert für die Relevanz dieser Merkmale. Die Ergebnisse sind ebenfalls in Oprach und Budau (vgl. 2019) zu finden.

Mithilfe der identifizierten Literatur (siehe Kapitel 3.2) können enthaltene Merkmale gefiltert und analysiert werden.

Hierfür wird ebenfalls ein Strukturierungsansatz gewählt, um die identifizierten Faktoren zu kategorisieren und sie nach der Anzahl ihrer Nennungen zu klassifizieren. Es werden die folgenden fünf Kategorien für eine Erstellung einer Taxonomie bauspezifischer Einflussfaktoren gewählt (5P Framework):

1. Preis: Kosten, Projekt des Investors/eigene Investition, Zahlungsrhythmus der abgeschlossenen Arbeit, Gehalt, Angebotsart usw.
2. Position: geografische Lage, verfügbare Baufläche, Bodenbeschaffenheit usw.

3. Position: Neubau/Umbau, Projekttyp, Bruttogeschossfläche (BGF), Anzahl der Stockwerke, Gebäudehöhe, usw.
4. Personal: Größe der Unternehmen, Anzahl der Bauleiter, Erfahrung, Fähigkeiten, Unterstützung durch den Kunden, Risikoverhalten, Führungsverständnis usw.
5. Prozess: Anzahl der Feedbackschleifen, Kommunikation, Detaillierung des Entwurfs, Einsatz von Informationstechnologie, Qualität der vorgelagerten Prozesse usw.

### Auswertung des 5P-Frameworks

Aus den in Kapitel 2 identifizierten Literaturquellen werden die als relevant identifizierten Merkmale aus der Kurzzusammenfassung selektiert und in eine Datei übertragen. Diese Merkmalsliste wird mithilfe von RapidMiner ausgewertet. RapidMiner ist eine Software zum ML. Mithilfe von algorithmischen Bausteinen des Text-Minings ist die Extraktion einzelner Wörter und die Analyse nach ihrer Häufigkeit möglich (siehe Abbildung 3.5).

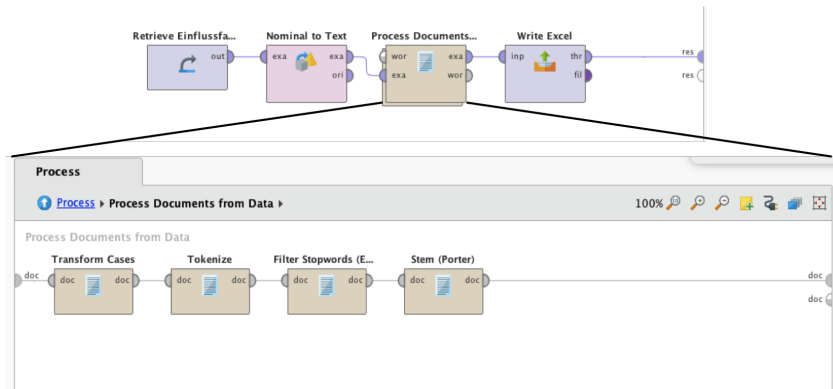
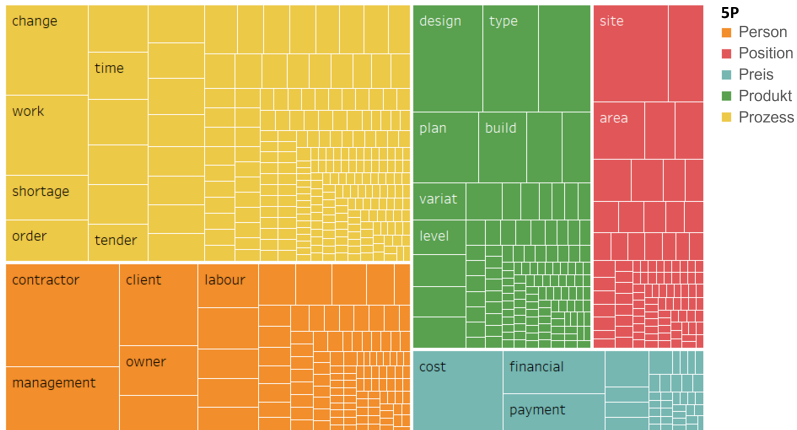


Abbildung 3.5: Klassifizierung der Einflussmerkmale mit RapidMiner

Diese Ergebnisliste wird mit Tableau visualisiert (siehe Abbildung 3.6). Tableau wird zur Datenvisualisierung und zum Reporting verwendet. Auf Basis der Anzahl der Nennungen werden in Abbildung 3.6 die Kategorien Prozess, Person, Position, Preis und Produkt skaliert.



**Abbildung 3.6:** Auswertung der beeinflussenden Merkmale auf Basis der Literaturrecherche mit RapidMiner und Tableau.

Die Anzahl der Nennungen ist je Kategorie wie folgt verteilt: 1.067 in der Kategorie Prozess, 707 in der Kategorie Person, 653 in der Kategorie Produkt, 394 in der Kategorie Position und 251 in der Kategorie Preis. Die Kategorien Prozess und Personen decken damit die Mehrheit mit 58 % der Nennungen ab.

Dabei sind in den Kategorien die drei am häufigsten genannten Merkmale je Kategorie wie folgt genannt:

- Prozess:
  1. Änderungen (z. B. Anzahl, Umfang) mit 76 Nennungen
  2. Arbeitseffizienz (z. B. Anzahl Arbeiter) mit 68 Nennungen

3. Kürzungen (z. B. der Arbeiter, Materialien, Maschinen) mit 38 Nennungen
- Person:
    1. Auftragnehmer (z. B. Expertise, Größe) mit 121 Nennungen
    2. Baumanagement (z. B. Expertise, Größe) mit 77 Nennungen
    3. Bauherr (z. B. Expertise, Größe) mit 67 Nennungen
  - Produkt:
    1. Design (z. B. funktional, Anzahl Stockwerke) mit 78 Nennungen
    2. (Produkt-)Typ (z. B. Wohnungen, Infrastruktur) mit 62 Nennungen
    3. Material (z. B. Holzbauweise, Massivbauweise) mit 58 Nennungen
  - Platzierung bzw. Örtlichkeit:
    1. Baustelle (z. B. Fläche) mit 76 Nennungen
    2. Wetter (z. B. regnerisch) mit 35 Nennungen
    3. Gebiet (z. B. urban) mit 31 Nennungen
  - Preis:
    1. Kosten als Aufwände (z. B. initial geschätzte Kosten für das Bauprojekt) mit 78 Nennungen
    2. Finanzen mit 47 Nennungen
    3. Preise als Einnahmen (z. B. Verkaufspreise) mit 41 Nennungen

Es muss erwähnt sein, dass einzelne Nennungen auch Überkategorien bilden oder Synonyme zu anderen Merkmalen darstellen können und die Beschreibungen der Merkmale daher nicht trennscharf erfolgt ist. So kann das Merkmal „Baustellenzugang“ ebenfalls Teil des Merkmales „Baustelle“ sein oder das Merkmal „Finanzen“ deckt als Synonym zu „Preis“ alle weiteren Merkmale dieser Kategorie



ab. Die Anzahl der Nennungen stellen daher eine Minimalanzahl und Näherung dar.

Weiter analysieren viele der einbezogenen Veröffentlichungen Abweichungen und einige dieser Informationen sind zum Zeitpunkt der frühen Terminplanung von Projektphasen bzw. in der Rahmenterminplanung noch nicht bekannt. So können bspw. Fähigkeiten und Wissen vieler beteiligter Rollen noch nicht vollständig bewertet sowie detaillierte Arbeitsprozesse, Änderungen und Kürzungen zum Plan nicht vorhergesehen werden. Insbesondere sind zur Rahmenterminplanung Informationen zum Produkt (z. B. Geschossfläche, Anzahl Stockwerke), zum Platz bzw. Örtlichkeit (z. B. Baustellenfläche, Zugänglichkeit, politische und wirtschaftliche Rahmenbedingungen) und finanzielle Aspekte (z. B. Preise) vorhanden. In den Gruppierungen Personen und Prozesse können dennoch einzelne Merkmale bereits in die Prognose aufgenommen werden (z. B. Auftragnehmer, das Projektabwicklungsmodell etc.). Hier besteht insbesondere die Herausforderung in einer eindeutigen und wiederholbaren Dokumentation subjektiver Eindrücke (z. B. Erfahrungen). Abbildung 3.6 verdeutlicht daher die Komplexität und Unsicherheit bei der Bestimmung der Dauern in frühen Projektphasen. Die hier durch diese Merkmale ermittelte Dauer ist daher in erster Linie als Näherungswert zu verstehen.

## **3.6 Zusammenfassung und kritische Betrachtung**

Die Ergebnisse dieses Kapitels zeigen mithilfe einer systematischen Literaturrecherche analytische Methoden zur Prognose von Dauern von Bauprojekten in frühen Projektphasen, die anschließend nach ihrem angestrebten Wettbewerbsvorteil kategorisiert sind. Hierfür wird das Modell von Davenport und Harris (2007) verwendet. Die verwendeten Kategorien sind statistische Analysen, Hochrechnungen, prädiktive Methoden und Methoden zur Optimierung. Die Kategorien sind aufeinander aufbauend und gewinnen mit ihrem steigenden Grad an Intelligenz bzw. an angestrebtem Wettbewerbsvorteil für Unternehmen.

Die Ergebnisse zeigen weiter, dass eine Vielzahl an Veröffentlichungen zu analytischen Methoden mit einer geringen analytischen Stufe existiert und die Anzahl an Veröffentlichungen hier weiter wächst. Da bei Publikationen mit einer hohen analytischen Stufe kein zunehmender Trend erkennbar ist, sollte hier verstärkt Forschung getätigt werden, sodass sich die Prognosegenauigkeit und damit auch der wirtschaftliche Vorteil erhöht. Die Analyse hinsichtlich der Dauern in Abhängigkeit relevanter Merkmale zeigt zudem die Herausforderung einer genauen Prognose von Dauern in frühen Projektphasen auf.

Die durchgeführte Literaturrecherche stellt eine Übersicht der als relevant identifizierten Literatur dar, die jedoch nicht als vollständig zu betrachten ist. Die Liste sollte demnach weitergeführt und ergänzt werden, um einen noch umfassenderen Blick zu erhalten. Weiter ist kritisch anzumerken, dass viele relevante Projektmerkmale zum Zeitpunkt der Planung in frühen Projektphasen noch nicht bekannt sind. Eine frühe Einbindung aller relevanten Projektpartner (z. B. durch eine Integrierte Projektabwicklung (IPA)) kann demnach die Planungsgenauigkeit unterstützen. Auch sollten weitere Forschungsarbeiten zur einheitlichen und projektübergreifenden Dokumentation der Merkmale vorgenommen werden. Mit der Aufnahme zum Zeitpunkt der Rahmenterminplanung bekannter Merkmale kann die Rahmenterminplanung insbesondere als Annäherung an das Ziel verstanden werden. Die Nachvollziehbarkeit und Erklärbarkeit beeinflussender Merkmale auf die Dauern ist daher von einer zentralen Bedeutung in der Untersuchung dieser Merkmale.

Nach den definierten Anforderungen der Prognosegenauigkeit und Erklärbarkeit (siehe Kapitel 2.6) genügen die in der Forschung bisher betrachteten prädiktiven und optimierenden Methoden nicht aus. Hier kommen vor allem „Black Box“-Algorithmen zur Anwendung. Es wird lediglich das Ergebnis ohne die zugrundeliegenden Entscheidungsprozesse mit den dafür beeinflussenden Merkmalen und Merkmalsausprägungen aufgezeigt. Um Auswirkungen von jeweiligen Entscheidungen in einer frühen Phase besser zu verstehen, ist eine transparente und verständliche Darstellung des Prozesses für die Planungsbeteiligten wichtig. Im Bereich der prädiktiven Analysen werden zwar Entscheidungsbäume angewendet, die erklärbar sind. Dies erfolgt jedoch nur, um das Ensemble mit dem KNN

zu verbessern oder den Einfluss eines einzelnen Merkmals und nicht die Summe aller Merkmale für eine Entscheidung darzustellen. Planungsbeteiligte werden demnach nicht aktiv in den Prozess zur Bestimmung der Dauern eingebunden. Sie besitzen durch die hohe Komplexität der analytischen Methoden über ein geringes Verständnis über den Prozess sowie die beeinflussenden Merkmale. Dies kann die Akzeptanz der Ergebnisse reduzieren.

Nach der Systematik von Davenport und Harris (2007) sollte aus den genannten Gründen auf der Stufe der prädiktiven Methoden begonnen werden, erklär-bare Prognosemodelle zu untersuchen. Aufbauend auf diesen Lösungen können dann im folgenden Schritt optimierende erklärende Ansätze zur Rahmentermin-planung hinzugezogen werden. Mit eben solchen Erklärungen beschäftigt sich das Gebiet „eXplainable Artificial Intelligence“ (XAI). Die Grundlagen dieser hierzu werden in Kapitel 5 vorgestellt. Im folgenden Kapitel wird vorerst in einer explorativen Studie untersucht, inwiefern prädiktive Methoden in der Praxis Anwendung finden. So stellt sich die Frage, ob neben analytischen Planungsmethoden weitere strategische Ansätze existieren, die ebenfalls einen relevanten Einfluss auf die Prognose haben.



# **4 Explorative Studie zu wirksamen Ansätzen zur Reduktion von Abweichungen geplanter Dauern**

## **4.1 Vorbemerkung**

In Kapitel 2.5.6 (‘Ansätze zur Reduktion von systematischen Planungsfehlschlüssen’) konnte aufgezeigt werden, dass die Anwendung analytischer Planungsmethoden zur Verbesserung der Planung und damit der Reduktion von Planabweichungen dienlich sein kann. Ziel dieses Kapitels ist es, mithilfe einer Studie das aktuelle Vorgehen zur Erstellung von Rahmenterminplänen in Deutschland aufzunehmen sowie die aufgeführten Ansätze zur Reduktion von Abweichungen auf ihre Wirksamkeit zu überprüfen. Die Ansätze werden hierfür in ihre einzelnen Elemente unterteilt und durch ein methodisches Vorgehen mit der Planungsqualität in Verbindung gesetzt sowie statistisch überprüft. Es wird überprüft, ob analytische Methoden als Strategie zur Reduktion von Abweichungen auf Basis einer entsprechend großen Datengrundlage dienlich sind und Elemente anderer Ansätze ebenfalls die Planungsqualität beeinflussen. Damit wird die Beantwortung der ersten Forschungsfrage angestrebt: Wie werden derzeit Rahmenterminpläne für Bauprojekte im Untersuchungsraum Deutschland durch das Projektmanagement erstellt und welche Ansätze zur Reduktion von Abweichungen geplanter Dauern sind statistisch signifikant?

Hierfür wird zunächst die Forschungsmethodik zur Identifizierung eines Systemkonstrukts relevanter Ansätze mit dem Ziel der Reduktion von Abweichungen geplanter Dauern vorgestellt (Kapitel 4.2). Die explorative Studie wird nach univariaten und bivariaten Gesichtspunkten ausgewertet (Kapitel 4.3). Die Ergebnisse der explorativen Studie werden anschließend ausgewertet und abschließend werden wesentliche Punkte zusammengefasst sowie ein Ausblick auf weitere Forschungsmöglichkeiten gegeben (Kapitel 4.4).

Die in Kapitel 2.5.6 beschriebenen Strategien zur Reduzierung von Planungsabweichungen bilden die Grundlage für die folgende Untersuchung.

## **4.2 Methodisches Vorgehen der explorativen Studie**

### **4.2.1 Übersicht**

Um die Einflüsse von Ansätzen zur Reduzierung von Abweichungen der Dauern zu identifizieren und damit ihre Wirksamkeit zu validieren, wird eine zweistufige Umfrage in Organisationen der Baubranche mit einer anschließenden Explorativen Faktoranalyse (EFA) durchgeführt. Da Märkte einen sozialen, wirtschaftlichen und ökologischen Rahmen mit Technologien, Rechtssystemen, Vorschriften und Projektabwicklungssystemen definieren (vgl. Love und Ahiaga-Dagbui 2018), liegt der Fokus dieser Umfrage auf Deutschland. Die EFA dient dem schrittweisen Auffinden von zusammenhängenden Konstrukten einzelner Elemente (Faktoren), die zu Beginn unbekannt sind. Die Elemente bilden sich aus dem aktuellen Prozess der Rahmenterminplanung der befragten Unternehmen. Die Ergebniskonstrukte können so der Planungsqualität gegenübergestellt werden und korrelierende Zusammenhänge zeigen wirksame Ansätze auf. Damit ist die Studie zugleich induktiv und deduktiv.

Insbesondere orientiert sich das Vorgehen an Hinkin (1998):

1. Item-Generierung,
2. Durchführung der Umfrage,
3. Reduktion der ursprünglichen Items,
4. weitere Konstruktvalidität,
5. Wiederholung der Umfrage.

Dabei wird der letzte Schritt für die weitere Forschung ausgelassen und in dieser Arbeit nicht umgesetzt. Hier sollte mit einer bestätigenden Faktoranalyse die externe Konsistenz des identifizierten Systemskonstrukts überprüft werden. Das Vorgehen nach Hinkin (1998) wird mit seinen 1.766 Zitierungen nach dem Web of Science (WoS) (Zugriff des Zeitpunkts am 18.07.2022) als ein relevantes und wissenschaftlich anerkanntes Vorgehen angesehen.

## 4.2.2 Schritt 1: Item-Generierung

Items beschreiben ein Frage-Antwort-Konstrukt (vgl. Kallus 2016, S. 18). Zur Item-Generierung werden die Elemente der Ansätze zur Reduktion von Abweichungen als Eingangsgröße sowie die Planungsqualität als Ausgangsgröße verwendet. Die Planungsqualität ist hier definiert als Zeitabweichungen, Kundenzufriedenheit, Maßnahmen und Puffer (vgl. Kapitel 2.6).

Um relevante Items zu identifizieren, ggf. zu ergänzen und vorab zu validieren, wird ein teil-standardisierter Leitfaden mit elf Fragen zur Befragung von Experten in der Rahmenterminplanung erstellt (siehe Tabelle 4.1). Es werden insgesamt sieben Experteninterviews (und ein Pre-Test) durchgeführt. Im folgenden Abschnitt werden die Interviewpartner als IP1 bis IP7 bezeichnet. Zur Auswahl der Experten werden drei Bewertungskriterien festgelegt. Diese sind Größe, Branche und Rolle in der Rahmenterminplanung des Unternehmens. Die Interviews mit den ausgewählten Experten erfolgten persönlich, am Telefon oder im Video- bzw. Audiochat. Die Interviews werden mittels eines Diktiergeräts aufgenommen und

anschließend transkribiert. Die vollständige Transkription der Interviews ist auf der Datenplattform Mendeley<sup>2)</sup> zu finden.

**Tabelle 4.1:** Struktur des Interviewleitfadens für die Experteninterviews

	<b>Strategie und Planungsqualität</b>	<b>Erklärung</b>
(1)	Einleitung	Begrüßung und Erfassung der Unternehmensdaten
(2)	Wissensdatenbank	Beschreiben Sie die Software, Dokumente und Informationen als relevante Wissensdatenbanken.
(3)	Prozess der Überwachung und Berichterstattung	Beschreiben Sie den Prozess der Steuerung, nachdem das Projekt begonnen hat.
(4)	Motivation durch Anreize	Beschreiben Sie Anreize als Motivation für eine gute Planungsqualität.
(5)	Fähigkeiten und Erfahrungen	Beschreiben Sie die erforderlichen Fähigkeiten und Erfahrungen.
(6)	Planungsprozess	Beschreiben Sie den Planungsprozess eines Rahmenterminplans.
(7)	Zeitbedarf	Beschreiben Sie die benötigte Zeit.
(8)	Benötigte Information	Beschreiben Sie die benötigten Informationen als Grundlage für die Planung.
(9)	Qualität	Beschreiben Sie die Qualität der Planung (Abweichungen, Zufriedenheit, Maßnahmen, Puffer).
(10)	Hürden und Probleme	Beschreiben Sie die Hürden und Probleme bei der Planung.
(11)	Andere	Haben Sie wahrheitsgemäß geantwortet? Haben Sie weitere Informationen oder Ideen?

Aus den Interviews leiten sich zwei wesentliche Ergebnisse ab:

<sup>2)</sup> <https://data.mendeley.com/datasets/gsk9xxfhth/1>



1. Aufgrund der langen Dauer von Bauprojekten gibt es so gut wie keine Rückkopplungsschleifen bezüglich der Überwachung und Berichterstattung. Die Interviews zeigen, dass die Befragten keine statistischen Daten erheben, um rückwirkend oder prädiktiv Bauprojekte auswerten zu können. Der Schwerpunkt von Strategie (1) „Sammeln von Informationen bzw. Verbesserung der Leistungsüberwachung, der Berichterstattung und des Informationsaustauschs“ liegt daher auf der Wissensdatenbank.
2. Die Interviews legen nahe, dass es in deutschen Unternehmen der Baubranche keine internen Anreize für eine gute Planungsqualität gibt. Außerdem gibt es in Deutschland keine öffentlichen Vorschriften zur Anreizsetzung.

Sowohl der Prozess der Überwachung und Berichterstattung als auch die Stärkung der Rechenschaftspflicht und der Verantwortung für Fehler sowie Überschreitungen sind wichtige Ansätze, die in Zukunft in Betracht gezogen werden sollten. Aufgrund der fehlenden Daten in der untersuchten Region wird eine statistische Bewertung ihrer Auswirkungen ausgeschlossen.

Tabelle 4.2 zeigt die aus den Experteninterviews abgeleiteten Items. Die mit einem Sternchen markierten Items werden für die univariate Analyse verwendet.

**Tabelle 4.2:** Übersicht der verwendeten Items (\* Verwendung für univariate Ergebnisse)

Item Beschreibung	Item Nummer (I)	
Eigenes Wissen zur Rahmenterminplanung	01	*
Repräsentativität der Rahmenterminplanung	02	*
Bedeutung der Rahmenterminplanung	03	*
Zufriedenheit	04	
Anzahl eingebundener Experten	05	
Grad an Teamwork	06	
Anzahl Phasen	07	
Analytische Methoden zur Phasendefinition	08	
Analytische Methoden zur Bestimmung der Dauer	09	
Grad der Fixierung / Einbindung von Risiken	10	*
Verfügbare Informationen zum Projekt	11	
Verfügbare dokumentierte Daten	12	*
Anzahl dokumentierter Projekte	13	
Grad der digitalisierten Dokumentation	14	
Verfügbare Informationen abgeschlossener Projekte	15	
Filtermöglichkeiten in der Datenbank	16	*
Zeitabweichungen	17	
Maßnahmen	18	
Zeitpuffer	19	
Gründe für Abweichungen	20	*
Grad der Neuartigkeit	21	
Investment	22	
Repräsentativität des letzten Projektes	23	
Funktion	24	*
Erfahrungen in dieser Funktion	25	
Erfahrungen in der Baubranche	26	
Anzahl involvierter Projekte in der Rahmenterminplanung	27	
Anzahl Bauprojekte des Unternehmens im letzten Jahr	28	
Art des Unternehmens	29	*
Spezialisierung des Unternehmens	30	*
Anzahl Angestellte	31	
Jährlicher Umsatz	32	
Standort in Deutschland	33	*
Anmerkungen an der Umfrage	34	*

Auf Basis der durchgeführten Interviews wird eine Umfrage entwickelt. Hierfür werden die Items und Skalen nach allgemeinen Richtlinien (vgl. Hinkin 1998) angepasst. So werden zu jedem Block nach Harvey et al. (1985) mindestens vier Items abgefragt, um die Homogenität innerhalb des Konstruktes gewährleisten zu können. So kann trotz einer Item-Reduktion sichergestellt werden, dass alle Items das gleiche Konstrukt messen. Die Validität der Umfrage hinsichtlich des Aufbaus und Inhalts wird mit drei verschiedenen Vorgehensweisen überprüft: Es wird erstens ein Peer-Review mit fünf Forschern im Bereich Baumanagement, zweitens mit einem Praktiker aus dem Bauwesen und drittens mit einer Person außerhalb des Baumanagements durchgeführt. Im Pre-Test werden einzelne Fragen angepasst, umbenannt oder weggelassen. Die Befragung beinhaltet insgesamt 34 Items (I01-I34) inklusive Items zur Qualität des Ergebnisses (I17-I20) sowie dem Screening (I01, I02). Weiter wird durch den Pre-Test ermittelt, dass es die schnellste und einfachste Methode ist, sich bei der Bewertung der Items jeweils an das letzte abgeschlossene Projekt zu erinnern. Dadurch wird ein Anker gesetzt, sodass die Verzerrung aufgrund von Erinnerungsfehlern reduziert werden kann.

Bei dieser Befragung werden keine einheitlichen Antwortskalen für alle Items verwendet, hingegen werden Item-spezifische Antwortskalen ausgewählt: Bei Items, die Einstellungen oder die Zustimmung bzw. Ablehnung abfragen, wird eine sieben-stufige Likert-Skala verwendet (vgl. Baur und Blasius 2019, S. 847ff., Kallus 2016, S. 43). Likert-Skalen bezeichnen fünf- bzw. siebenstufige Zustimmungsskalen zur Einstellungsmessung (vgl. Kallus 2016, S. 43). Items, die retrospektiv das spezielle Vorgehen oder vorhandene Informationen bei der Rahmenterminplanung abfragen, sind in der Befragung als Multiple Choice Items aufgeführt. Um die Items vergleichbar zu machen, wird ein begründetes Scoringmodell angewendet. Dabei sind den Items nach einem einheitlichen Bewertungsschlüssel Zahlen zugeordnet. Somit sind die Items wieder messbar. Der messbare Wert kann z. B. durch das Abzählen der Anzahl an ausgewählten Auswahlmöglichkeiten bei einem Multiple Choice Item gebildet werden. (vgl. Kallus 2016, S. 41 und 54)

Die Anordnung bzw. Reihenfolge der Items innerhalb der Online-Befragung ist nach den Empfehlungen von Baur und Blasius (2019, S. 816) erstellt.

### **4.2.3 Schritt 2: Durchführung der Erhebung**

Die Umfrage wird an Experten im Bauprojektmanagement in Deutschland verteilt und ist im Mai 2020 online verfügbar gewesen. Sie beinhaltet Fragen zum Screening, dem Prozess, den benötigten Daten, der Qualität und demografische Daten der Teilnehmer.

Potenzielle Teilnehmer der Umfrage werden online über verschiedene Kanäle rekrutiert: Die Befragung wird per E-Mail durch das Institut für Technologie und Management im Baubetrieb (TMB) des Karlsruher Instituts für Technologie (KIT) verteilt. Hier werden die Mitglieder des Deutschen Verbands für Projektmanagement in der Bau- und Immobilienwirtschaft e.V. (DVP e.V.) angeschrieben. Die E-Mail enthielt ein Anschreiben, das, wie nach Häder (2015, S. 2444f.) empfohlen, folgende Themen beinhaltete: Den Nutzen der Online-Befragung, die Notwendigkeit der Teilnahme, die Vorstellung des forschenden Instituts, die Dauer der Befragung, Hinweise auf die Vertraulichkeit mit den Ergebnissen sowie eine Erklärung, wie die Person in die Zielgruppe gelangt ist. Auch wird um Weiterleitung der Befragung an weitere Wissensträger gebeten. Zusätzlich wird die Befragung in verschiedene Gruppen der sozialen Plattformen Xing und LinkedIn gepostet. Daneben wird die Befragung auch an persönliche Kontakte verteilt. Nach zwei Wochen wurde eine Erinnerungs-E-Mail versendet. Auch wenn mit der Verteilung via E-Mail und sozialen Medien eine breite Zielgruppe erreicht werden kann, kann nicht gewährleistet werden, dass die Stichprobe repräsentativ ist.

Es haben an der Umfrage insgesamt 131 Teilnehmer teilgenommen. Damit wird die benötigte Stichprobe nach Rummel (1970) erreicht. Hier wird empfohlen, im Vergleich zu den Items die vier-fache Teilnehmeranzahl zu akquirieren. Diese Empfehlung entspricht einer Teilnahme von mindestens 128 Teilnehmern, bei den 34 Items abzüglich der Screening-Fragen ( $32 \times 4 = 128$  Teilnehmer).

Die Teilnehmer konnten in ihrer Geschwindigkeit teilnehmen. Im Median hat die Teilnahme etwa zwölf Minuten gedauert. Insgesamt haben 21 Teilnehmer an unterschiedlichen Stellen im Fragebogen abgebrochen. Dies entspricht einer Abbruchquote von 16 %.

Zusätzlich werden drei Screening-Fragen eingebaut, um Teilnehmer mit einem zu geringen Kenntnisstand ausschließen zu können. Die erste und zweite Kontrollfrage war „Führen Sie selbst die Rahmenterminplanung von Bauprojekten durch? - Ja oder Nein.“ und „Befindet sich der Standort Ihrer Organisation in Deutschland? - Ja oder Nein.“ Bei einer Beantwortung mit „Nein“ bei einer der beiden Fragen werden die Antworten ausgeschlossen. Die dritte Screening-Frage war „Wie repräsentativ war das Vorgehen bei der Rahmenterminplanung in Ihrem letzten abgeschlossenen Projekt im Vergleich zu Ihrer sonstigen Vorgehensweise?“ Wurde die Frage auf der Likert-Skala nicht positiv, also mit einem nicht repräsentativen Vorgehen, beantwortet, so wird auch dieser Teilnehmer ausgeschlossen. Durch diesen Fragenblock konnten insgesamt acht Teilnehmer ausgeschlossen werden.

Damit haben insgesamt 102 Teilnehmer im vordefinierten Expertenkreises vollständig die Umfrage abgeschlossen.

Die Daten der Umfrage werden univariat und bivariat ausgewertet. Univariate Analysen betrachten die Ergebnisse einer Frage bzw. eines Items, bivariate Analysen untersuchen den Zusammenhang zweier Items. So werden bei den univariaten Untersuchungen demografische Analysen, Analysen zu Praktiken des Datenmanagements und zu eingesetzten analytischen Methoden durchgeführt. Die bivariaten Analysen werden im Zuge der Schritte drei und vier nach Hinkin (1998) weiter untersucht.

### 4.2.4 Schritt 3: Reduktion der ursprünglichen Items

Auf Basis der Umfrage wird eine EFA durchgeführt. Es werden sechs Items schrittweise ausgeschlossen, da sie keine großen Einflüsse auf andere Komponenten haben. Dies wird nach Ford et al. (1986) allgemein mit einer Bewertung von  $\geq 0,4$  definiert. Anforderungen an die Kommunalität ( $< 0,4$ ) und Wechselwirkungen ( $\geq 0,4$ ) auf andere Komponenten haben jedoch keine Auswirkungen auf die Auswahl der Items. Die Kommunalität kann zwischen null und eins liegen. Sie beschreibt, wie gut potenzielle Untergruppen von Items den Inhalt der Variablen widerspiegeln. Eine geringe Kommunalität bedeutet eine schlechte Repräsentation der Informationen des Items in der potenziellen Untergruppe. Hohe Wechselwirkungen zwischen einzelnen Items zeigen hingegen die Möglichkeit auf, die Untergruppen um diese Items zu reduzieren.

Ergebnis sind vier identifizierte Konstrukte. Die kumulierte Varianz beträgt hierbei 78,5 % und ist damit höher als die geforderte kumulierte Varianz nach Hinkin (1998) von 60 %. Weiter wird empfohlen das Cronbachs-Alpha als Kennzahl der Konsistenz und internen Validität des Systemkonstrukts zu berechnen. Zur Berechnung des Cronbachs-Alpha werden mindestens drei Items je Konstrukt benötigt. Die Berechnung ist daher lediglich für ein Konstrukt (Konstrukt 1: Unternehmensgröße) möglich. Das Cronbachs-Alpha beträgt hier 96,4 % und übersteigt damit das benötigte Ziel von explorativen Studien von 70 % (vgl. Cortina 1993).

### 4.2.5 Schritt 4: Weitere Konstruktanalyse

Ziel ist es nun, Abhängigkeiten zwischen den entwickelten Konstrukten und den Variablen der Planungsqualität zu identifizieren. Hierfür können nach Hinkin (1998) Korrelations- oder Regressionsanalysen verwendet werden. Da ordinale Werte auf Basis der Umfrage vorliegen, wird der Rangkorrelationskoeffizient  $r_{s,p}$  nach Spearman verwendet (vgl. Cramer und Kamps 2020, S. 113ff., Krickhahn 2017, S. 102ff.).

Weiter gibt der p-Wert als Signifikanzwert an, ob das Ergebnis statistisch signifikant oder rein zufällig entstanden ist. Je kleiner der p-Wert ist, desto unwahrscheinlicher ist die Nullhypothese. Die Nullhypothese beschreibt die Annahme, die statistisch zu überprüfen ist. Bei einem p-Wert von  $p = 0,05$  können somit 5 % der Varianzen durch die Nullhypothese und somit dem Zufall beschrieben werden. (vgl. Schiefer und Schiefer 2018, S. 73f.)

Die Kovarianz gibt im Vergleich zur Korrelation das erwartete Verhalten zweier Variablen an, ohne dabei eine Beziehung aufzuzeigen. Die Kovarianz kann daher unterstützend die Stärke der Korrelation aufzeigen.

Zur Visualisierung dieser Beziehungen wird ein Monoplot verwendet. Dieses ist mit den Ergebnissen der durchgeführten Umfrage in der Abbildung 4.5 in Abschnitt 4.3.3 visualisiert. Hier werden die Abhängigkeiten zwischen allen Variablen aufgezeigt. Dabei ist dies eine Annäherung des multidimensionalen Raums auf zwei Dimensionen. Die Dimensionalität wird demnach reduziert. Das Monoplot ist wie folgt zu verstehen: Kleine Winkel zwischen zwei Variablen bedeuten eine positive Korrelation; Variablen, die in 90 Grad zueinander stehen, bedeuten keine Korrelation; ein Winkel nahe bei 180 Grad bedeutet eine negative Korrelation. Es werden demnach Faktoren gesucht, die möglichst weit weg von einem 90 Grad Winkel zur Planungsqualität der Rahmenterminplanung liegen. Weiter bedeuten Faktoren im Zentrum, dass diese schlecht repräsentiert sind und sollten zur Wertung der Variablen nicht aufgenommen werden. Faktoren nahe dem Rand des äußeren Kreises sind hingegen gut repräsentiert.

Das Monoplot sowie die durchgeführte Regressionsanalyse geben Aufschluss über die relevanten Komponenten.

## **4.2.6 Schritt 5: Wiederholung**

Bis heute wurde die Umfrage nur einmal durchgeführt und beruht daher auf einer Stichprobe. Die Wiederholung mit einer unabhängigen zweiten Stichprobe unterstützt die Auswertung der externen Validität und allgemeingültigen Aussage der

Ergebnisse (vgl. Stone 1978, Hinkin 1998). Mit diesem Schritt kann die Struktur nochmals überprüft werden (siehe Confirmatory Factor Analysis, Schritt vier nach Hinkins). Dieser Schritt bleibt zukünftigen Arbeiten und anderen Forschern vorbehalten, um eine Voreingenommenheit des Forschers auszuschließen.

## 4.3 Ergebnisse der explorativen Studie

### 4.3.1 Ergebnisse der deskriptiven univariaten Analyse

Die Ergebnisse der Umfrage sind als Datensatz auf der Datenplattform Mendeley verfügbar <sup>3)</sup>. Es wird, wie bereits beschrieben, in univariate und bivariate Analysen unterschieden.

Im Folgenden werden die Ergebnisse der Umfrage dargestellt. Sie geben Aufschluss über die Teilnehmer der empirischen Studie, ihre Erfahrung, das Datenmanagement des jeweiligen Unternehmens sowie verwendete Methoden zur Bestimmung der Rahmenterminpläne.

**Teilnehmer der empirischen Studie:** Die Studie enthält alle Organisationsgrößen: Kleinst- (20 %) und Kleinunternehmen (22 %), mittelständische (19 %) und Großunternehmen (38 %) (vgl. Commission 2003). Hier werden Unternehmen mit weniger als 10 Mitarbeiter und einem Jahresumsatz von unter 2 Mio. EUR als Kleinstunternehmen definiert. Andere Unternehmen mit weniger als 50 Mitarbeiter und einem Jahresumsatz von unter 10 Mio. EUR werden als kleines Unternehmen definiert. Bei anderen Unternehmen mit weniger als 250 Mitarbeiter und einem Jahresumsatz von unter 50 Mio. EUR wird von einem mittleren Unternehmen gesprochen. Großunternehmen sind damit Unternehmen mit mehr als 250 Mitarbeiter oder einem Jahresumsatz von mehr als 50 Mio. EUR. Großunternehmen sind in der Umfrage überdurchschnittlich repräsentiert. Da jedoch mit

---

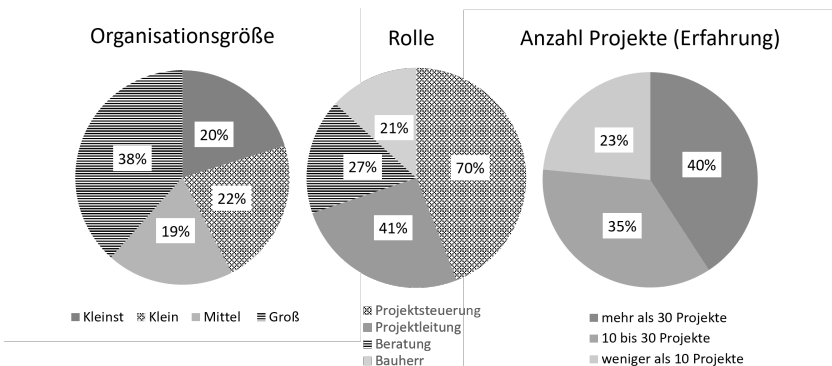
<sup>3)</sup> <https://data.mendeley.com/datasets/gsk9xxfhth/1>



der Zunahme der Unternehmensgröße auch eine zunehmende Professionalisierung angenommen werden kann, sind die Ergebnisse zum Vorgehen in der Praxis zur Rahmenterminplanung als besonders aussagekräftig anzusehen. Bei der Abfrage der Attribute der Teilnehmer war eine Mehrfachauswahl möglich. In der Studie ist vorwiegend die Organisationsart Projektmanagement (69 %) vertreten. Aber auch die Arten Ingenieurbüro (23 %), Beratungsunternehmen (23 %) und Baumanagement (22 %) befinden sich vermehrt in dieser Studie. Weiter haben verschiedene Arbeitsrollen an der Studie teilgenommen: Projektsteuerung (70 %), Projektleitung (41 %), Beratung (27 %) und Bauherr (21 %).

**Erfahrung:** Die Teilnehmer haben im arithmetischen Mittel 15 Jahre Erfahrung in diesen Rollen. Der Mittelwert der gesamten Berufserfahrung in der Baubranche liegt bei 21 Jahren. Die Teilnehmer haben zu 35 % bereits zehn bis 30 sowie zu 40 % mehr als 30 Projekte bearbeitet. 23 % haben erst weniger als zehn Projekte bearbeitet. Damit kann angenommen werden, dass die Erfahrung der Umfrageteilnehmer ausreichend fundiert ist, um das Vorgehen der Praxis zu beschreiben.

Die Verteilung der Teilnehmer der Studie bezüglich der Organisationsgröße und Rolle sowie ihrer Erfahrung sind in der folgenden Abbildung 4.1 dargestellt.



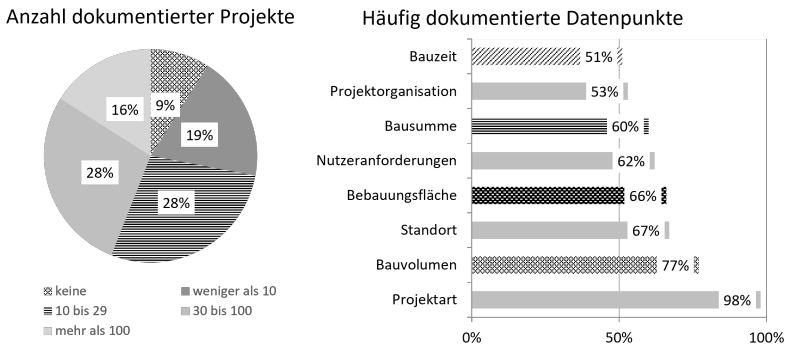
**Abbildung 4.1:** Univariate Ergebnisse zu den Teilnehmern der Studie und ihrer Erfahrung

**Datenmanagement:** Nur 9 % der Befragten verfügen über keine dokumentierten Daten zu bereits durchgeführten Projekten, 19 % verfügen über weniger als zehn, 28 % über zehn bis 29 und weitere 28 % über 30 bis 100 dokumentierte Projekte. Nur 16 % verfügen über mehr als 100 dokumentierte Projekte. Es werden im Mittelwert 14 Datenpunkte je Projekt dokumentiert, wobei es keinen Datenpunkt gibt, der bei allen Teilnehmern dokumentiert wird. Am häufigsten werden die Datenpunkte: Projektart, Bauvolumen, Bausumme sowie aktualisierte Start- und Endzeitpunkte der Vorgänge gespeichert.

Demgegenüber sind zu dem zu planenden Projekt zum Zeitpunkt der Rahmen-terminplanung im Mittelwert 9,4 Datenpunkte vorhanden. Aber auch hier gibt es keinen einzigen Datenpunkt, der bei allen Projekten der Studie vorliegt. Bei mehr als der Hälfte der Teilnehmer liegen folgende dokumentierte Merkmale zum Zeitpunkt der Rahmenterminplanung vor: Projektart (98 %), Bauvolumen (77 %), Standort (67 %), Bebauungsfläche (66 %), Nutzeranforderungen (62 %), Bausumme (60 %), Projektorganisation (53 %) und verfügbare Bauzeit (51 %). Es ist anzumerken, dass angenommen wird, dass die fehlenden Informationen dennoch vorliegen können, jedoch an dezentralen Orten. Weiter dokumentiert keiner der Teilnehmer marktbeschreibende Merkmale oder bezieht diese in die Planung mit ein. Kenngrößen, wie die Inflation während des Projektes, die politische Stabilität des Standortes oder die klimatische Lage könnten durchaus relevante Einflussgrößen auf die Dauer sein.

Die Ergebnisse zum bestehenden Datenmanagement sind in Abbildung 4.2 visualisiert.

**Verwendete Methoden:** Eine Betrachtung der verwendeten analytischen Methoden zur Bestimmung der Vorgänge im Rahmenterminplan zeigt, dass hierbei nur die analytische Methode des Datenzugriffs angewendet wird. Im Besonderen führen die Teilnehmer vorwiegend einen Projektvergleich mit bereits durchgeführten Projekten durch. Zusätzlich basiert die Bestimmung der Vorgänge stark auf der Intuition der Teilnehmer, indem auf Basis der persönlichen Erfahrung vorgegangen wird. 55 % der Teilnehmer verwenden zur Bestimmung der notwendigen Vorgänge ein kombiniertes Vorgehen aus intuitiven sowie analytischen Vorgehensweisen.



**Abbildung 4.2:** Univariate Ergebnisse zu der Anzahl dokumentierter Projekte und häufig dokumentierter Datenpunkte

Zur Bestimmung der notwendigen Dauern der Vorgänge werden verschiedene analytische Methoden angewendet (Mehrfachauswahl möglich): Auch hier führen die meisten Befragten (80 %) einen Datenzugriff in Form eines Projektvergleichs durch. 33 % der Teilnehmer wenden hier auch „statistische Analysen“ an und verwenden Benchmarks zur Bestimmung der Dauern. 16 % führen mathematische Formeln für Hochrechnungen durch und nur ein einziger Teilnehmer verwendet Methoden der KI zur Prognose der Dauern. Zudem wird auch hier intuitiv auf Basis der persönlichen Erfahrung (85 %) sowie der Erfahrung von zusätzlichen Experten (52 %) vorgegangen. Zur Bestimmung der Dauern der Phasen gehen 83 % der Befragten kombiniert vor und lediglich 10 % gehen rein intuitiv und 7 % gehen rein analytisch vor.

Die Ergebnisse zu angewendeten Methoden in der Bestimmung der jeweiligen Dauer sind in Abbildung 4.3 visualisiert.

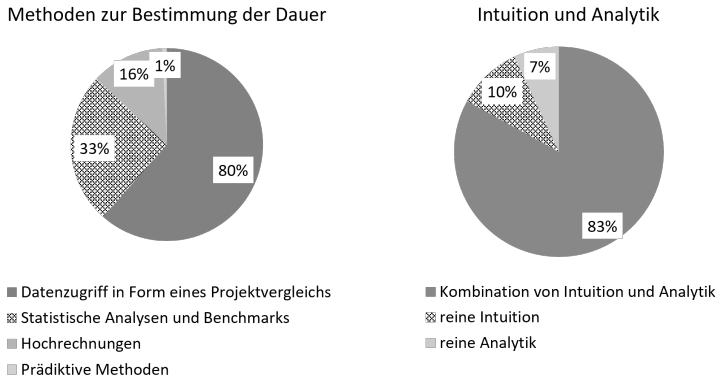


Abbildung 4.3: Univariate Ergebnisse der Methoden zur Definition von Projektdauern

### 4.3.2 Ergebnisse der explorativen Faktoranalyse

Ziel der Umfrage ist es zu verstehen, welche Ansätze bzw. Elemente dieser Ansätze zur Reduktion von Abweichungen auf die Planungsqualität der Rahmen-terminplanung wirken. Hierfür werden mithilfe einer EFA die Elemente in einem Modell zu vier Kategorien zusammengefasst. Das Modell enthält dabei acht Items und umfasst die Kategorien Unternehmensgröße, Wissensdatenbank, Wissen zum Projekt, Erfahrung und Fähigkeiten. Tabelle 4.3 zeigt die Ergebnisse der EFA sowie bestimmende Kennzahlen auf.

Bei der Varimax-Rotation wird die Varianz der quadrierten Faktorladungen pro Konstrukt maximiert, um die Aussagekraft der erstellten Konstrukte bewerten zu können. Hierfür werden in einem ersten Schritt mit der Kaiser-Normalisierung die quadrierten Faktorladungen durch die Kommunalitäten dividiert. Die Kommunalität beschreibt, wieviel Varianz die Variable an allen Faktoren erklärt. Die Einzigartigkeit je Item hingegen beschreibt die nicht aufgeklärte Varianz der Variable. Der Eigenwert eines Konstrukts entspricht der von ihm bei allen Items zusammen erklärten Varianz. Eigenwerte größer als eins (nach der Kaiser-Normalisierung) unterstützen die theoretischen Unterscheidungen der Konstrukte. Die erstellten vier Konstrukte zeigen Eigenwerte größer eins auf und lassen

**Tabelle 4.3:** EFA mit Varimax-Rotation und Kaiser-Normalisierung; Hauptgewichtungen in schwarzer und Kreuzgewichtungen in grauer Schrift

Item	Konstrukt	1	2	3	5	Eigenwert	Kommunalität
Anzahl Projekte pro Jahr in der Organisation (I28)...	Unternehmensgröße	-0,795	0,107	0,039	0,172	0,322	0,678
Anzahl an Angestellten (I31)...		-0,975	0,113	0,036	0,121	0,005	0,995
Jährlicher Umsatz (I32)...		-0,886	0,148	-0,058	0,221	0,138	0,862
Anzahl dokumentierter Projekte (I13)...	Wissensdatenbank	-0,040	0,983	-0,023	-0,098	0,005	0,995
Verfügbare Information zum Projekt (I11)...	Projektwissen	-0,017	0,041	-0,922	0,033	0,005	0,995
Erfahrungen in der Bauwirtschaft (I26)...	Erfahrung / Fähigkeiten	0,158	0,053	-0,056	-0,792	0,341	0,659
Erfahrungen in der Position (I25)...		0,303	0,129	-0,148	-0,929	0,005	0,995
	Eigenwert	3,308	2,175	3,459	2,835		
	Anteil der erklärten Varianz	0,221	0,145	0,231	0,189		
	Kum. Anteil der erklärten Varianz		0,365	0,596	0,785		

sich daher unterscheiden. Weiter sollte der Anteil der kumulierten erklärten Varianz möglichst hoch sein. Hierbei wird das akzeptable Minimum von 60 % mit 78,5 % erreicht (vgl. Hinkin 1998, S. 112).

### Konstrukt 1: Unternehmensgröße

Die Unternehmensgröße kann durch die Anzahl der Mitarbeiter, die durchgeführten Projekte oder den Umsatz beschrieben werden. Sind diese Items groß, könnte tendenziell von einer Professionalisierung des Vorgehens ausgegangen werden. Es existiert ein größerer Erfahrungsschatz in Form von Personen oder Daten.

### Konstrukt 2: Wissensdatenbank

Bezüglich der Dokumentation historischer Projekte konnten in den Interviews unterschiedliche Vorgehensweisen identifiziert werden. So dokumentiert Interviewpartner drei durchgeführte Projekte lediglich in PowerPoint (IP3), andere Organisationen verwalten eine unternehmenszentrale Datenbank (IP6). Der Zugriff auf dokumentierte Attribute historisch durchgeführter Projekte zur Terminplanung ist relevant, wie Interviewpartner sechs beschreibt:

„Und da haben wir [...] dann unsere Datenbank der abgeschlossenen Projekte, wo wir dann immer nachvollziehen können oder benchmarken können.“

Auf dem Datenzugriff sind Projektvergleiche, Benchmarks, Hochrechnungen, Vorhersagen oder Optimierungen möglich (vgl. Davenport und Harris 2007). Nach Kahneman und Tversky (1977) wird die Vorhersage von Projektdauern auf Basis von „singulären“ Informationen oder aber „distributiven“ Informationen getroffen. Singuläre Informationen sind ausschließlich auf das Projekt bezogene Informationen. Distributive Informationen hingegen sind Erfahrungen aus anderen vergangenen Projekten. In Schätzungen greifen Personen verstärkt auf singuläre Informationen zurück, da sie ihre Projekte tendenziell als einzigartig wahrnehmen und vernachlässigen distributive Informationen.

Nach Buehler et al. (1994) führt das ungerechtfertigte Bevorzugen von Planbasierten Informationen zu Planungsfehlschlüssen. Projektverlängernde Attribute werden tendenziell als externe Einflüsse wahrgenommen und sind nicht Bestandteil des Projektaufwandes. Sie werden daher weniger oder gar nicht in die Schätzung zukünftiger Projekte einbezogen.

Aufgrund der Nichtberücksichtigung distributiver Informationen kommt es häufig zu Planungsfehlschlüssen.

### **Konstrukt 3: Projektwissen**

Das Wissen zum aktuellen Projekt ist die Grundlage der Terminplanung. So stellt auch Interviewpartner eins (IP1) die Aufgabe der Informationsgewinnung zur Bestimmung eines der Projektattribute in den Vordergrund:

„Wichtig ist es [...] so viele [Informationen] wie möglich zu sammeln.“

So wird auch in der Literatur beschrieben, dass ein Zerlegen eines komplexen Problems in seine Subkomponenten bzw. Attribute die Planungsqualität unterstützt. Wird ein vielschichtiges Projekt nicht oder nicht ausreichend in seine Subkomponenten heruntergebrochen, kann dies zu einer Fehleinschätzung der

Projektdauern führen. Es wird nach Kruger und Evans (2004) zu einer Unterschätzung der Projektbestandteile führen, nach Kahneman und Tversky (1977) und Buehler et al. (1994) zu einer Überschätzung.

#### **Konstrukt 4: Erfahrungen und Fähigkeiten**

Das Einbeziehen von Erfahrungen und Fähigkeiten in der Rahmenterminplanung reicht nach den Interviews von einer Planung bei einem Verantwortlichen (IP3) bis zu einer Planung im Team (IP2). Interviewpartner zwei nennt als die größte Hürde der Rahmenterminplanung ein mangelhaftes Einschätzungsvermögen. So bezieht Interviewpartner sieben (IP7) lediglich Wissen von Experten ein, dokumentierte Daten werden vollständig missachtet:

„Wissens- und Datenbanken, wo wir Termine abgreifen, sind alles Architekten oder Ingenieure, zum Großteil mit je mehreren Jahren Berufserfahrung. Die wissen, wie lange es dauert. Da brauche ich keine Wissens- und Datenbank.“

Die Qualität einer Lösung steigt mit den Schätzungen der an der Planung beteiligten Personen. Eine Gruppe von Experten kann folglich zu einer besseren Lösung kommen als eine Einzelperson (vgl. Jakoby 2019, S. 180). Intuitive Schätzungen verschiedener Personen können sehr unterschiedlich ausfallen und besitzen somit Unsicherheiten und einen großen Streubereich (vgl. Jakoby 2019, S. 175). Bona-beau (2003) zeigt hingegen den Nachteil auf, dass die Verwendung von Intuition zwar zu Erfolg, aber genauso gut auch zu Misserfolg führen kann. Die Intuition sei dabei kein Mittel, um die Komplexität eines Problems zu bewerten, sondern um sie zu ignorieren.

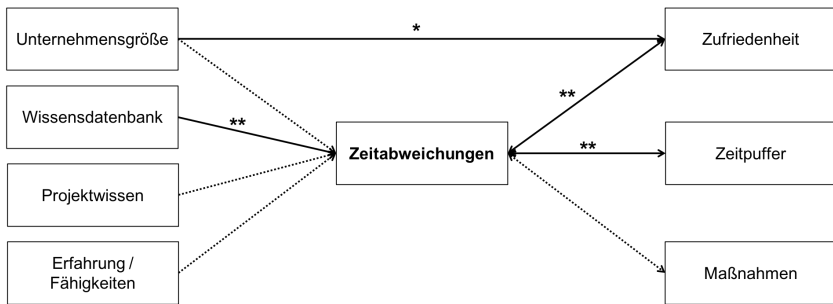
An dieser Stelle ist anzumerken, dass verwendete analytische Methoden, wie einfache Hochrechnungen oder die Anwendung von Methoden der KI, nicht Teil des Systems aus den vorgestellten vier Konstrukten sind und daher nicht repräsentiert sind.

### 4.3.3 Validitätsüberprüfung des Systemkonstrukts

Das beschriebene Systemkonstrukt wird mit der EFA auf seine inhaltliche Konsistenz und interne Validität überprüft. Im Folgenden werden im Rahmen der weiteren Konstruktanalyse Abhängigkeiten der vier identifizierten Konstrukte zu der Planungsqualität überprüft. Die Konstrukte, die sich aus den Elementen der vorgestellten Ansätze bilden, werden nun mit der Planungsqualität des zuletzt abgeschlossenen Projektes gegenübergestellt. Die Planungsqualität beinhaltet:

- Planabweichungen (in %),
- Besondere Maßnahmen zur Planeinhaltung,
- In der Planung enthaltene Zeitpuffer (in %),
- Die allgemeine Zufriedenheit mit der Planung.

Abbildung 4.4 zeigt die identifizierten statistischen Zusammenhänge zwischen den Konstrukten und der Planungsqualität.



**Abbildung 4.4:** Qualitätsbezogene Validitätsbewertung des identifizierten Frameworks mit den Signifikanzcodes: '\*\*\*\*' 0,001 '\*\*\*' 0,01 '\*\*' 0,05; graue, gestrichelte Pfeile kennzeichnen nicht signifikante Beziehungen

Dabei wird ein Spearman's Rangkorrelationskoeffizient von  $\geq 0,25$  als statistisch signifikant angenommen und mit dem p-Wert das Signifikanzniveau überprüft. Es



sind vor allem statistisch signifikante Korrelationen zur Zeitabweichung zu finden. Die Zeitabweichung ist mit der Größe der Datenbank historischer Projekte korreliert. Der Zugriff auf in die Subkomponenten zerlegten Projekte ist nach den Ergebnissen der Studie ein Erfolgsfaktor zur Reduktion der Abweichungen. Dieses Ergebnis wird durch die Studie von Cao et al. (2019) bestätigt. Die Autoren konnten in der branchenübergreifenden Studie nachweisen, dass eine statistisch positive Verbindung zwischen den Fähigkeiten der Informationsverarbeitung und dem Wettbewerbsvorteil des jeweiligen Unternehmens mittels der Effektivität der Entscheidungsfindung existiert.

Weiter zeigen die Ergebnisse einen Zusammenhang zwischen der Unternehmensgröße und der Zufriedenheit sowie zwischen der Zufriedenheit und den Zeitabweichungen. Die mit der Unternehmensgröße verbundene Professionalität spielt demnach eine Rolle. Die Abweichungen sind somit ebenfalls beeinflusst durch die Zufriedenheit und Zeitpuffer.

An dieser Stelle sei auf das „Parkinson’s Law“ verwiesen (Parkinson und Lancaster 1958): Es beschreibt den Effekt, dass in einem Projekt Zeitpuffer immer aufgebraucht werden. Diese Zeitpuffer werden auch dann aufgebraucht, wenn das Projekt den Puffer tatsächlich gar nicht zur Fertigstellung benötigt. Den Ergebnissen der Umfrage zufolge werden Zeitpuffer entweder nicht aufgebraucht oder sogar über das geplante Maß hinaus beansprucht. Die Planung von Zeitpuffern hat demnach einen positiven Einfluss auf Planungsfehlschlüsse und sollte daher in der Planung vermieden werden.

Tabelle 4.4 gibt zusätzlich die einzelnen Kennwerte der Analyse wieder. Hier wird die statistische Signifikanz mit dem Spearman Rangkorrelationskoeffizient  $r_{sp}$ , sowie der p-Wert und die Kovarianz dargestellt. Werte nahe bei Plus oder Minus eins weisen auf einen hohen Zusammenhang hin.

Um die Abhängigkeiten zwischen den Variablen besser zu verstehen und diese zu visualisieren, wird ein Korrelationsdiagramm in Form eines Monoplots erstellt (siehe Abbildung 4.5). Items der identifizierten Konstrukte und der Planungsqualität sind im Vergleich zu den anderen Items hier in schwarz und fett dargestellt.

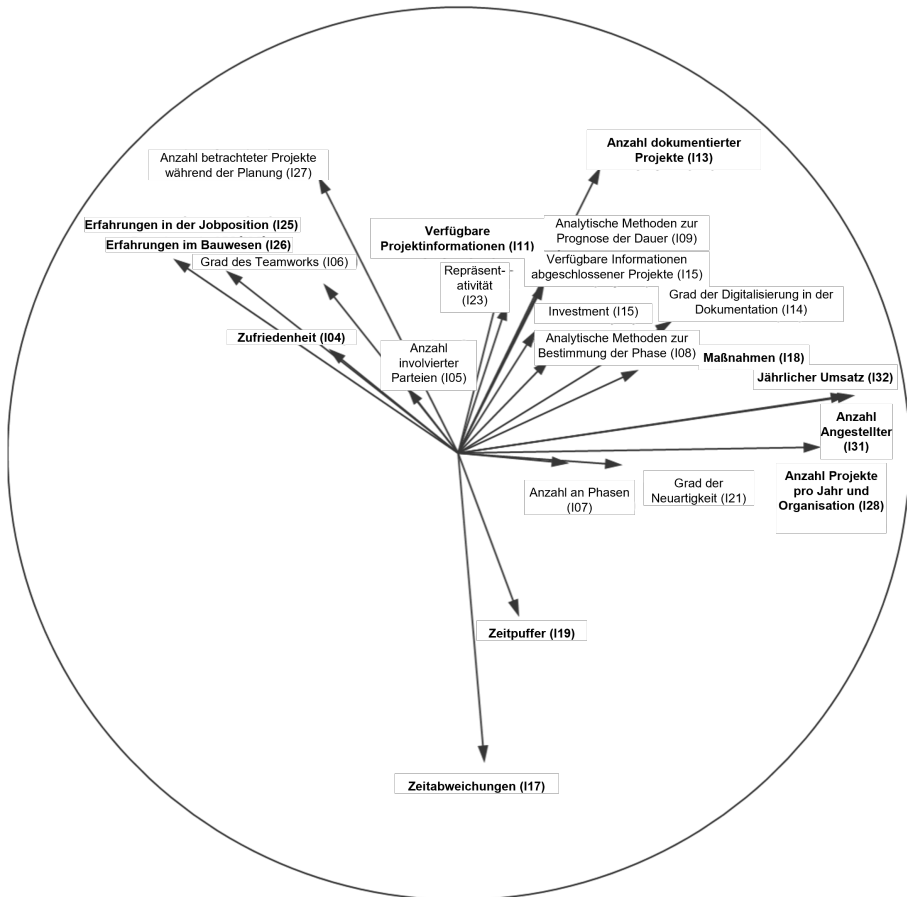
**Tabelle 4.4:** Matrix der Spearman-Rangkorrelationskoeffizient mit 95 % Konfidenzintervall, p-Wert und Kovarianzen zur Zeitabweichung (Item 17)

Zeitabweichung (I17)	Spearman's rs	p-Wert	Kovarianz
Dokumentierte Projekte (I13)	-0,301	0,0150	-0,3305
Dokumentierte Informationen zu abgeschlossenen Projekten (I15)	-0,326	0,0080	0,2137
Zufriedenheit (I04)	-0,346	0,0048	-0,2560
Zeitpuffer (I18)	0,376	0,0020	0,2248

Dieses ist entsprechend der Erläuterungen in Schritt vier aus Abschnitt 4.2.5 zu verstehen.

Das Korrelationsdiagramm spiegelt ebenfalls die bereits erkannte Korrelation zwischen der Wissensdatenbank (I13) zu den Zeitabweichungen wider. Es lässt sich zeigen, dass die Erfahrung (I25, I26) radial weiter entfernt ist und daher ein geringer Zusammenhang anzunehmen ist (I25:  $r_{sp} = -0,159$ ; I26:  $r_{sp} = -0,187$ ). Weiter ist die Unternehmensgröße nicht korreliert zu den Zeitabweichungen, da diese nahezu in einem 90 Grad Winkel zueinander stehen. Als Variable zur Messung der Planungsqualität ist die Zeitabweichung im Vergleich zu der Zufriedenheit, den Maßnahmen und eingeplanten Zeitpuffern am repräsentativsten. Es können weiter folgende Aussagen abgeleitet werden:

- Je höher die Zufriedenheit (I04) mit der Rahmenterminplanung, desto geringer die Zeitpuffer (I19).
- Je höher die Zufriedenheit (I04) mit der Rahmenterminplanung, desto mehr Erfahrung in der Baubranche (I25, I26) sowie in der aktuellen Rolle existieren und desto öfter sind die Personen in die Rahmenterminplanung (I27) involviert.
- Je höher die Zufriedenheit (I04) mit der Rahmenterminplanung, desto größer das Unternehmen (I28, I31, I32).
- Je mehr Maßnahmen (I18) zur Zielerreichung eingesetzt werden, desto mehr Daten sind zum aktuellen Projekt sowie historischen Projekten verfügbar (I11, I13, I15).



**Abbildung 4.5:** Korrelationsmonoplot für Kriterien zur Konstruktvalidität (Items des Konstrukts in fetter Schrift, andere Items in grauer Schrift)

- Je mehr Maßnahmen (I18) zur Zielerreichung eingesetzt werden, desto größer ist das Unternehmen (I28, I31, I32).

Die Analyse des entwickelten Frameworks und der Planungsqualität zeigt Abhängigkeiten zu den dokumentierten distributiven Daten auf. Wird ein komplexes

Problem in seine Subkomponenten zerlegt, kann die Planungsqualität erhöht werden (vgl. Kahneman und Tversky 1977, Buehler et al. 1994, Kruger und Evans 2004). Dabei zeigt die Umfrage Vorteile bei einem Zugriff auf dokumentierte Daten im Vergleich zu Erfahrungen auf. Diese Datenbasis kann als Grundlage analytischer, datenbasierter Methoden dienen.

## **4.4 Zusammenfassung und kritische Betrachtung**

Abweichungen von den geplanten Dauern kennzeichnen eine große Anzahl von Bauprojekten. Siemiatycki (2009) schlägt fünf Ansätze zur Reduzierung von Planungsfehlschlüssen und damit von Zeitabweichungen vor. Um die Auswirkungen dieser Ansätze zu messen, wird eine explorative Studie in Organisationen der Baubranche in Deutschland durchgeführt. Auf der Grundlage von Interviews wird eine Umfrage erstellt, durchgeführt und statistisch ausgewertet. Ergebnis der explorativen Befragung mit 131 Teilnehmern ist ein 4-Faktoren-Framework, das die Elemente der Ansätze zuordnet. Bei diesen Faktoren handelt es sich um (1) die Unternehmensgröße, (2) das Wissen aus historischen Projekten, (3) das Projektwissen und (4) die Erfahrung und das Wissen der Planungsteilnehmer. Dieses Systemkonstrukt wird mit mehreren statistischen Parametern validiert und verifiziert.

Die univariaten Analysen haben gezeigt, dass die Planer kaum moderne Techniken verwenden. Es werden insbesondere einfache Datenzugriffe verwendet. Die Planer verwenden am häufigsten eine Kombination aus einfachen, analytischen und intuitiven Methoden. Hier bestätigt zwar die Forschung, dass ein Zusammenspiel beider Bereiche, Intuition und Analytik, die größten Vorteile bietet, Methoden mit einer hohen analytischen Stufe sollten jedoch eine breitere Anwendung finden. Die Analyse des entwickelten Systemkonstrukts zeigt, dass es eine statistische Korrelation zwischen der Projektdatenbank und den Zeitabweichungen gibt. Die erste Strategie, (1) das Sammeln von Informationen bzw.

die Verbesserung der Leistungsüberwachung, des Berichtswesens und des Informationsaustauschs, hat demnach den größten Einfluss auf die Reduzierung der Zeitabweichungen. Die Ergebnisse der analysierten Studien unterstützen die Ergebnisse von Siemiatycki (2009). Auch Ahiaga-Dagbui und Smith (2014, S. 682) weisen auf „eine wachsende Anerkennung von Informationen als wichtiges Wettbewerbsinstrument“ hin. So kann die Anzahl der Projekte in der Datenbank sowie die Art und Weise, wie diese dokumentiert werden, ein Indikator für die Professionalität des jeweiligen Unternehmens sein. Sie konnte als manifeste Variable identifiziert werden. Latente Variablen, wie die Intelligenz der einzelnen Planer oder die Unternehmenskultur werden nicht überprüft und können einen Einfluss auf die Zufriedenheit mit der Planung haben.

Die Erhebung konzentriert sich ausschließlich auf den Wirtschaftsstandort Deutschland. Es kann nicht davon ausgegangen werden, dass sie auf andere Standorte übertragbar ist. Weiter konzentriert sich die Stichprobe nur auf eine Datenerhebung. Nach dem fünften Schritt des methodischen Vorgehens von Hinkin (1998) sollte die Befragung und Analyse ein zweites Mal durchgeführt werden. Die Umfrageergebnisse können auch zur erneuten Validierung der vorgeschlagenen Struktur verwendet werden. Es muss auch erwähnt werden, dass einzelne Ansätze bereits durch die Ergebnisse der Interviews nicht weiter betrachtet werden, da eine statistische Auswertung dieser Methoden auf Grundlage der Interviews ausgeschlossen ist. Nichtsdestoweniger kann auf der Grundlage der Interviews gezeigt werden, dass die Qualität des Planungsprozesses gesteigert werden kann, wenn Anreize gesetzt werden. Auch die Fokussierung auf moderne Techniken für die Berichterstattung und Überwachung von Rahmenterminplänen, trotz langer Feedback-Schleifen, erhöht die Planungsqualität weiter.

Mit der durchgeführten explorativen Studie konnte aufgrund des geringen Einsatzes analytischer Methoden der Einfluss dieser auf Planungsabweichungen nicht untersucht werden. Es bleibt zu untersuchen, ob Planungsabweichungen mit den verfügbaren Daten in Unternehmen und Planungsmethoden einer hohen analytischen Stufe weiter reduziert werden können. Aufgrund der Relevanz des aufgezeigten Datenmanagements und der Forschungslücke im Bereich analytischer Planungsmethoden bei prädiktiven und optimierenden Verfahren sollten diese in

der Forschung eine besondere Berücksichtigung finden. Die Ergebnisse der univariaten Analyse zeigen diesen Aspekt: In der Praxis werden bisher nur Methoden mit einem niedrigen analytischen Niveau, wie z. B. Datenzugriff, eingesetzt.

Im Folgenden wird aufgrund des fehlenden Einsatzes in Forschung und Praxis zum einen der Begriff KI und zum anderen der Begriff XAI erläutert. Hierbei wird der Themenbereich zu ML-Methoden als Synonym für prädiktive Methoden verwendet.

# 5 Grundlagen ausgewählter Methoden der Künstlichen Intelligenz

## 5.1 Vorbemerkung

Das Forschungsgebiet um das Thema „KI“ hat sich in den vergangenen Jahren stark weiterentwickelt und wird in den kommenden Jahren weiter an Dynamik gewinnen. Obwohl die Grundlagen bereits in der Mitte des letzten Jahrhunderts geschaffen wurden, sind viele praktische Anwendungen erst heute möglich. Dies folgt einerseits aus der kontinuierlich steigenden Leistungszunahme von Hardware und Software (bekannt als „Moore’s Law“), aber auch andererseits aus der Verfügbarkeit von wachsenden Datenmengen und deren Vernetzung im Zuge der Digitalisierung. Die Definitionen des Begriffs der KI sind durch die anhaltende Entwicklung oft unscharf, unterscheiden sich je nach Quelle und lassen sich darüber hinaus nicht klar voneinander abgrenzen. Teilgebiete der KI (Kapitel 5.2) sind ML (Kapitel 5.3), KNN (Kapitel 5.4) und Deep Learning (DL) (Kapitel 5.5) (siehe Abbildung 5.1), die in den folgenden Abschnitten in ihren Grundzügen erläutert werden. Um den Anforderungen der Prognosegenauigkeit und Erklärbarkeit zu genügen, werden ausgewählte Methoden der erklärbaren KI (eXplainable Artificial Intelligence, XAI) in Kapitel 5.6 vorgestellt und abschließend zusammengefasst (Kapitel 5.7).

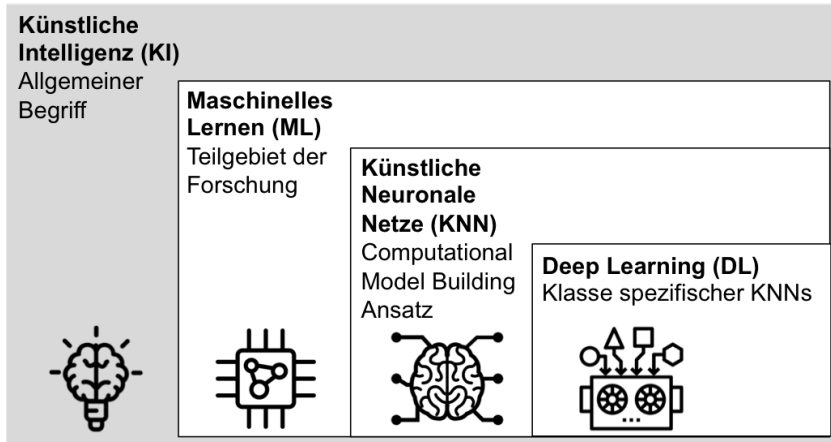


Abbildung 5.1: KI, ML, DL (vgl. Copeland 2016, Denk 2017)

## 5.2 Künstliche Intelligenz

### Definitionen

KI befasst sich mit dem Nachbau menschlicher Intelligenz durch Maschinen, d. h. meist Computer (vgl. Kroll 2013, S. 8). Alan Turing entwickelte 1950 dafür den nach ihm benannten Turing-Test. Bei diesem muss eine Person in mindestens 30 % der Testläufe die Ausgabe einer Maschine, für die eines Menschen halten, damit sie als „intelligent“ gilt (vgl. Ertel 2016, S. 4). John McCarthy, der den Begriff der KI als Erster definierte, schrieb 1955, dass das Ziel der KI sei, Maschinen zu entwickeln, die sich verhalten, als verfügten sie über Intelligenz (vgl. McCarthy et al. 1955).

Die Schwierigkeit dieser Definition besteht in der Deutung des Begriffs Intelligenz. Das Vorbild, die menschliche Intelligenz, zeichnet sich durch komplexe kognitive Fähigkeiten aus, ermöglicht situationsabhängige Entscheidungen (vgl. Lunze 2016, S. 2), bezieht ein Bewusstsein, Emotionalität und Intuition ein. Diese Entscheidungen können u. a. mathematischer, neurologischer, psychologischer, kybernetischer oder linguistischer Natur sein (vgl. Russell und Norvig



2012, S. 29 f.). Aus diesem Grund gestaltet sich eine einheitliche Definition von KI als schwierig. Kühl et al. (2022, S. 7) nennen weiter, dass unter Methoden der KI auch einfache regelbasierte Systeme und Formeln gehören können. Hingegen eignet sich der Begriff ML für ein eindeutiges Verständnis besser. ML gibt Computern die Möglichkeit zu lernen, ohne dabei explizit programmiert zu sein (vgl. Samuel 1959). Dieser Begriff wird in Abschnitt 5.3 weiter erläutert.

In diesem Zusammenhang wird daher auch zwischen „starker“ und „schwacher KI“ unterschieden. Nach Russell und Norvig (2012, S. 1176) zeichnet sich eine starke KI dadurch aus, dass sie wirklich „denken“ kann, während schwache KI ihre Intelligenz nur simuliert. In der Praxis eignen sich die schwachen Ansätze, um spezifische Anwendungsprobleme zu lösen (z. B. Bilderkennung beim autonomen Fahren oder Berechnung multidimensionaler Probleme), wohingegen starke Intelligenz sich vornehmlich vielschichtigen Aufgaben widmet (z. B. logisches Denken oder komplexe Kommunikation in natürlicher Sprache) und bisher keine Anwendung findet. Im Englischen werden hierfür auch die Begriffe „Artificial Narrow Intelligence“ (ANI) und „Artificial General Intelligence“ (AGI) verwendet, die den Unterschied besser verdeutlichen (vgl. Goertzel 2007, S. 1). Aktuell ist die schwache KI Stand der Forschung. Demnach werden auf Basis eines bestimmten Datensatzes konkrete und genau auf diesen Datensatz trainierte Prognosen erstellt. Diese sind nicht direkt übertragbar auf andere Datensätze bzw. generalisierbar. Daher sind die Ergebnisse dieser Arbeit eher als Fallstudien zum Einsatz von KI zu verstehen und dienen weniger dazu, eine allgemeingültige Aussage zum Einsatz von KI-Methoden zur Reduktion von Abweichungen geplanter Dauern in frühen Phasen eines Bauprojektes abzuleiten.

## **Intelligente Agenten**

Für eine schwache KI bilden intelligente Agenten die Grundlage. Als Agent (vom lat. „agere“ = tun, handeln, machen) wird im Allgemeinen ein System bezeichnet, das Informationen einer Eingabe verarbeitet und daraus eine Ausgabe produziert. Die Wahrnehmung der Daten erfolgt dabei über Sensoren und die daraus folgenden Aktionen über Aktuatoren. Dieses sehr allgemeine Konzept lässt sich auf verschiedene Weisen klassifizieren. Software-Agenten interagieren durch Ein-

und Ausgaben mit dem Benutzer. Hardware-Agenten hingegen agieren durch Wahrnehmung und Veränderung ihrer physischen Umgebung. (vgl. Ertel 2016, S. 18 f.)

Im Bereich der KI wird von einem intelligenten Agenten gesprochen. Die einfachste Form ist hierbei der einfache Reflexagent, der lediglich die aktuelle Wahrnehmung auswertet und vorherige Informationen nicht einbezieht. Die ausgeführte Aktion ermittelt sich dabei durch festgelegte Bedingungs-Aktions-Regeln (wenn-dann-Regeln). Der Vorteil einfacher Reflexagenten ist ihre Simplizität, die jedoch auch nur eine begrenzte Intelligenz ermöglicht. (vgl. Russell und Norvig 2012, S. 76 f.)

Lernende Agenten können hingegen selbstständig durch die ihnen zur Verfügung gestellten Informationen lernen. „Ein Agent heißt lernfähig, wenn sich seine Leistungsfähigkeit auf neuen, unbekanntem Daten im Laufe der Zeit (nachdem er viele Trainingsbeispiele gesehen hat) verbessert (gemessen auf einem geeigneten Maßstab).“ (Ertel 2016, S. 195) Dies ermöglicht, dass der Agent komplexe Aufgaben lösen kann, ohne dass diese explizit programmiert werden. So sind explizit programmierte wenn-dann-Regeln nicht mehr notwendig, da der Agent selbstständig die Regeln findet und diese lernt. Darüber hinaus kann ein solcher Agent in ihm unbekanntem Umgebungen funktionieren und dabei kompetenter werden, als er in seinem Ausgangszustand war. (vgl. Russell und Norvig 2012, S. 83)

Grundsätzlich enthalten lernende Agenten ein Leistungselement für Aktionen und ein Lernelement für die Verbesserung des Arbeitserfolges. Das sogenannte Kritikelement wiederum liefert dem Lernelement das Feedback abhängig von den Ergebnissen des Agenten. Diese Ergebnisse werden dann mittels des Leistungsstandards bewertet. Daraufhin kann das Lernelement das Leistungselement entsprechend ändern, um zukünftig bessere Ergebnisse zu erreichen. Der Problemgenerator schlägt Aktionen vor, die neue Erfahrungen des Agenten ermöglichen. Ohne den Problemgenerator würde der Agent immer die beste ihm bereits bekannte Aktion durchführen. (vgl. Russell und Norvig 2012, S. 83 f.)

Eine Anwendung lernender Agenten ist das Maschinelle Lernen, das sich mit der Auswertung großer Datenmengen beschäftigt und im Folgenden eingehender beschrieben wird.

## 5.3 Maschinelles Lernen

### 5.3.1 Ziele des Maschinellen Lernens

„[ML ist ein] Forschungsgebiet, das Computer in die Lage versetzen soll, zu lernen, ohne explizit darauf programmiert zu sein.“ (Samuel 1959) Das Ziel des ML ist demnach die Programmierung von Computern, sodass diese anhand von Daten lernen. Die Stärke des Verfahrens liegt insbesondere in der Fähigkeit, große Datenmengen verarbeiten zu können (vgl. Géron und Rother 2018, S. 4). Die klassische Programmierung zeichnet aus, dass die Regeln der Programme durch den Programmierer vorgegeben werden, und dadurch aus Daten (Input) Antworten (Output) generiert werden. Das ML nutzt hingegen sowohl Daten als auch dazugehörige Antworten, um dadurch Regeln abzuleiten. Dieses Programm mit Regeln kann anschließend auf weitere (neue) Daten angewendet werden, um dazugehörige Antworten zu generieren (vgl. Chollet 2018, S. 23). Mit dieser Definition des ML können die Methoden der bayesschen Statistik oder Fuzzy-Systeme zwar den Methoden der KI zugeordnet werden, sie werden jedoch aus dem Themenbereich des ML ausgeschlossen. Der Unterschied zwischen der klassischen Programmierung und dem ML ist in Abbildung 5.2 dargestellt.

ML zeichnet sich durch folgende Vorteile aus (vgl. Géron und Rother 2018, S. 4 ff.):

- **Automatisierung:** Arbeitsintensive Lernaufgaben, die viel Handarbeit oder viele Regeln benötigen, können mittels ML automatisiert werden.
- **Komplexität:** ML kann Lösungen für Probleme finden, die zu komplex für herkömmliche Methoden sind oder für die (noch) kein Programm existiert.

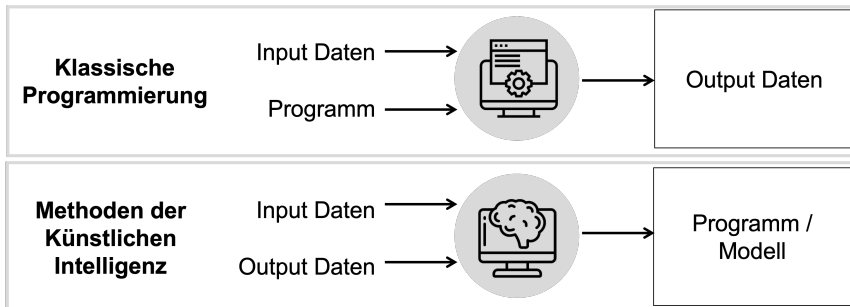


Abbildung 5.2: Das Programmierparadigma des ML im Vergleich zur klassischen Programmierung (vgl. Chollet 2018, S. 23)

- **Lernunterstützung:** ML kann Menschen beim Finden von Lösungen für Lernprobleme unterstützen. So können gelernte Regeln der ML-Modelle in vielen Fällen nachträglich untersucht werden, um Zusammenhänge und Beziehungen zu erkennen und zu verstehen.
- **Auswertung großer Datenmengen:** Durch die Digitalisierung stehen immer größere Datenmengen zur Verfügung, die von Hand kaum mehr bearbeitet werden können. ML kann es ermöglichen, nicht direkt sichtbare Muster zu erkennen.
- **Fluktuation:** ML-Algorithmen können sich sehr gut an neue Daten anpassen, die später zur Verfügung gestellt werden.

Das Vorgehen zum Erstellen eines ML-Modells wird in Unterabschnitt 6.3.1 beschrieben.

## 5.3.2 Lernen

Das Lernen aus zur Verfügung gestellten Daten ist das zentrale Konzept des ML. Beim Lernen geht es hierbei nicht um das „Auswendiglernen“ von Inhalten, sondern um das Erlernen eines Verfahrens, das auf beliebige andere Beispiele angewandt werden kann. Dies nennt man Generalisierung (vgl. Ertel 2016, S. 192)

(s. auch Gleichung 5.3.4). Konkreter ausgedrückt wird aus einer Sammlung von Ein- und Ausgabepaaren eine Funktion gelernt, durch die für weitere Eingaben (Input) eine Ausgabe (Output) vorhergesagt werden kann (vgl. Russell und Norvig 2012, S. 809).

In mathematischer Form ausgedrückt kann das Lernen folgendermaßen beschrieben werden: Es existiere eine Menge  $\mathcal{X}$  (gültige Eingaben) und eine unbekannte Funktion  $f$  zu  $\mathcal{X}$ , die jedem Element aus  $\mathcal{X}$  seine entsprechende Ausgabe zuweist. Das Lernen besteht nun darin, eine möglichst genaue Approximation  $\hat{f}$  zu  $f$  mithilfe der Daten zu finden. Wie  $f$  definiert ist, hängt von der Lernaufgabe ab. (vgl. Knox 2018, S. 3)

Dabei wird der Definitionsbereich  $\mathcal{X}$  als Merkmalsraum bezeichnet. Ein einzelnes Element  $X \in \mathcal{X}$  heißt Merkmalsvektor (oder Eingabe). Die einzelnen Koordinaten von  $X$  wiederum werden Merkmale (vom engl. „features“) genannt.

$f(\mathcal{X})$  wird als Wertebereich bezeichnet. Je nach Ausgabebetyp kann es sich dabei um finite ungeordnete Mengen handeln oder der Wertebereich in kontinuierlicher Form vorliegen. Ein einzelnes Element  $Y \in f(\mathcal{X})$  wird bei einer finiten ungeordneten Menge als Klasse und als Antwort bei kontinuierlicher Form bezeichnet. (vgl. Knox 2018, S. 4)

Werden nun konkrete Daten betrachtet, so können diese als zufällige Ziehungen  $(X, Y)$  aus einer Zufallsverteilung  $P$  auf  $\mathcal{X} \times f(\mathcal{X})$  betrachtet werden.

Es existieren verschiedene Kategorien, in die maschinelle Lernprozesse eingeteilt werden können, je nach Art der zur Verfügung stehenden Daten und der Steuerung des Lernprozesses (vgl. Géron und Rother 2018, S. 8).

## Überwachtes Lernen

Der Agent lernt beim überwachten Lernen (vom engl. „supervised learning“) einen Zusammenhang in jeweiligen Ein- und Ausgabepaaren und versucht so eine Funktionalität bzw. Regeln abzuleiten (vgl. Russell und Norvig 2012, S. 811). Mit bekannten Beispieldaten wird auf eine bestimmte Ausgabe und die damit einhergehende Interpretation trainiert. Anschließend können die gefundenen Regeln

verwendet werden, um neue Ausgaben zu erstellen. (vgl. Kirste und Schürholz 2019, S. 25)

Mathematisch gesehen bestehen die Daten in diesem Fall aus Definitionsbereich-Wertebereich-Paaren  $(X, Y)$  mit den Merkmalen  $(x_1, y_1), \dots, (x_n, y_n) \in \mathcal{X} \times f(\mathcal{X})$ . Jedes  $(x_i, y_i)$  wird hier aus einer gemeinsamen Wahrscheinlichkeitsverteilung  $P(X, Y)$  aus  $\mathcal{X} \times f(\mathcal{X})$  gezogen. Das Maschinelle Lernen hat meist die Aufgabe, einen Algorithmus zu approximieren, sodass gilt  $\hat{f}: \mathcal{X} \rightarrow f(\mathcal{X})$  (vgl. Knox 2018, S. 4 f.). Diese Art Daten werden als beschriftete Daten (engl. „labeled data“) und die Menge  $Y$  als Beschriftung (engl. „label“) bezeichnet, die die zugehörigen Klassen oder Antworten enthält. (vgl. Dasgupta 2010, S. 15)

Typische Anwendungsgebiete überwachten Lernens sind die Klassifizierung und die Regression (siehe Kapitel 5.3.3) (vgl. Knox 2018, S. 5).

## Unüberwachtes Lernen

Dahingegen lernt der Agent beim unüberwachten Lernen (von engl. „unsupervised learning“) von den Eingabedaten ohne explizites Feedback über das Ergebnis der Ausgabe (vgl. Russell und Norvig 2012, S. 811).

Mathematisch gesehen bestehen die Daten in diesem Fall nur aus Werten  $X$  aus dem Definitionsbereich (ohne  $Y$ ) mit den Merkmalen  $x_1, \dots, x_n \in \mathcal{X}$  (vgl. Knox 2018, S. 4 f.). Diese Art Daten werden als unbeschriftete Daten (engl. „unlabeled data“) bezeichnet (vgl. Dasgupta 2010, S. 15). Sie enthalten keine Labels wie beim überwachten Lernen und müssen deshalb vom Agenten selbst ermittelt werden. Das Erheben der Daten ist deshalb einfacher, der ML-Prozess jedoch aufwändiger und fehleranfälliger.

Typischerweise wird unüberwachtes Lernen für Clustering und Dichteschätzung (siehe Kapitel 5.3.3) genutzt (vgl. Knox 2018, S. 5).

## Semi-überwachtes Lernen

Zwischen den beiden zuvor vorgestellten Konzepten liegt das semi-überwachte Lernen, auch halbüberwachtes Lernen (von engl. „semi-supervised learning“)

genannt. Hier sind nur wenig beschriftete Daten verfügbar, dagegen viele unbeschriftete (vgl. Russell und Norvig 2012, S. 811).

Das Ziel ist hier bessere Vorhersagen zu treffen, als wenn ausschließlich die beschrifteten oder die unbeschrifteten Daten zum Lernen verwendet würden. Die Verwendung der beschrifteten Daten ermöglicht es dem Agenten, einen Startpunkt für die Bearbeitung der unbeschrifteten Daten zu finden. Bei der Klassifizierung (siehe Kapitel 5.3.3) ordnet der Agent zunächst die Daten mittels unüberwachtem Lernen in verschiedene Cluster. Anschließend kann er aus den wenigen gelabelten Daten in den jeweiligen Klassen schließen, um welche Klasse es sich bei welchem Cluster handelt. (vgl. Rebala et al. 2019, S. 22)

Auch hier sind Regression und Klassifizierung typische Anwendungsgebiete (siehe Kapitel 5.3.3) (vgl. Zhu 2010, S. 892 f.).

### **Bestärkendes Lernen**

Bestärkendes (auch verstärkendes) Lernen (engl. „Reinforcement Learning“) beschreibt eine Technik des Maschinellen Lernens, bei der der Lernerfolg durch Verstärkungen (d. h. Belohnungen und Bestrafungen) gesteuert wird (vgl. Russell und Norvig 2012, S. 811).

Dem Agenten wird hierbei nicht exakt mitgeteilt, wie groß der Unterschied der Ausgabe zum Zielwert ist (vgl. Lippe 2006, S. 56), sondern es wird lediglich ein Belohnungswert generiert, den der Agent versucht zu optimieren (vgl. Kirste und Schürholz 2019, S. 29). Das Feedback („bestärkendes Signal“) erhält der Agent dabei nicht nach jedem Beispiel, sondern nach Erreichen eines Zielzustands (vgl. Jones 2017). Dies ist insbesondere in Situationen relevant, wenn der Agent seinen Lernerfolg durch die Interaktion mit seiner Umgebung erfährt (vgl. Russell und Norvig 2012, S. 960). Eine beispielhafte Anwendung dieses Lerntyps ist ein Agent, der Schachspielen lernt. Die Belohnungsfunktion richtet sich hier nach dem Sieg oder der Niederlage des Spiels nach einer Abfolge von Zügen (vgl. Raschka und Mirjalili 2017, S. 30).

### 5.3.3 Aufgaben

Maschinelle Lernprozesse können auch nach dem Typ ihrer Aufgabe kategorisiert werden. Diese nicht abschließende Liste stellt verschiedene Konzepte vor.

#### Regression

Die Regressionsanalyse ist eine Anwendung des überwachten Lernens (s. Unterabschnitt 5.3.2), die zum Ziel hat, Zusammenhänge in Datensätzen zu finden (vgl. Rebala et al. 2019, S. 25). Die Regressionsanalyse bzw. multivariate Regression ist bereits ohne einen lernenden Agenten bzw. als einfacher Reflexagent in Abschnitt 3.3.2 beschrieben.

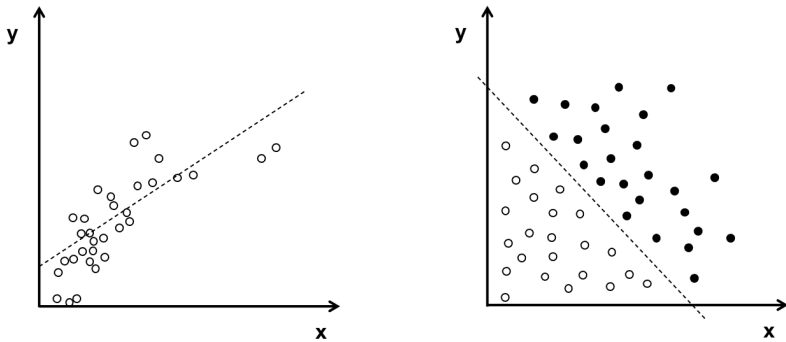
#### Klassifizierung

Auch die Klassifizierung wird dem überwachten Lernen zugeordnet, da sie beschriftete Daten verwendet. Hierbei werden Daten anhand von gemeinsamen Charakteristika Gruppen zugeordnet, den sogenannten Klassen (vgl. Drummond 2010, S. 168). Im Gegensatz zur Regression ist der Wertebereich der Zielvariablen  $f(\mathcal{X})$  eine finite ungeordnete Menge (vgl. Knox 2018, S. 4).

Existieren lediglich zwei Klassen, spricht man von Binärklassifizierung (vgl. Chollet 2018, S. 96). Sind dagegen mehr als zwei Klassen vorhanden, handelt es sich um Mehrfachklassifizierung. Noch genauer wird unterschieden zwischen der Single-Label-Mehrfachklassifizierung (Zuordnung eines Datensatzes zu ausschließlich einer Klasse möglich) und der Multi-Label-Mehrfachklassifizierung (Zuordnung eines Datensatzes zu mehreren Klassen möglich) (vgl. Chollet 2018, S. 109). Ein Beispiel einer einfachen Binärklassifizierung ist in Abbildung 5.3b abgebildet.

Regression und Klassifikation sind in Abbildung 5.3 visualisiert.





(a) Beispiel einer eindimensionalen Regression mit der gefundenen Approximation (gestrichelte Linie)      (b) Beispiel einer einfachen Klassifizierung mit der Trenngeraden

**Abbildung 5.3:** Regression und Klassifikation

## Clustering

Das Clustering ist die bekannteste Form des unüberwachten Lernens. Es geht hierbei darum, Häufungen („Cluster“) in den Trainingsdaten zu finden. Ein solches Cluster zeichnet aus, dass der Abstand zwischen den einzelnen Datenpunkten innerhalb eines Clusters kleiner ist, als zwischen unterschiedlichen Clustern (erkennbar in Abbildung 5.4) (vgl. Ertel 2016, S. 244 f.). In diesem Fall liegt ein finiter Wertebereich  $f(\mathcal{X})$  vor.

## Weitere Typen

Im Falle des unüberwachten Lernens mit kontinuierlichem Wertebereich  $f(\mathcal{X}) \in [0, \infty)$  wird von Dichteschätzung gesprochen (vgl. Knox 2018, S. 4).

Die Sequenzvorhersage (engl. „sequence prediction“) wird verwendet, um Reihenfolgen zu lernen. Es gilt:  $s_i, s_{i+1}, \dots, s_j \rightarrow s_{j+1}, \dots, s_{j+k}$  mit  $1 \leq i \leq j < \infty$  und  $k > 1$ . Dabei steht  $s$  für eine Teil der Sequenz. Das heißt, mit einem gegebenen  $s_i, s_{i+1}, \dots, s_j$  soll  $s_{j+1}, \dots, s_{j+k}$  vorhergesagt werden (vgl. Sun und Giles 2003, S. 4).

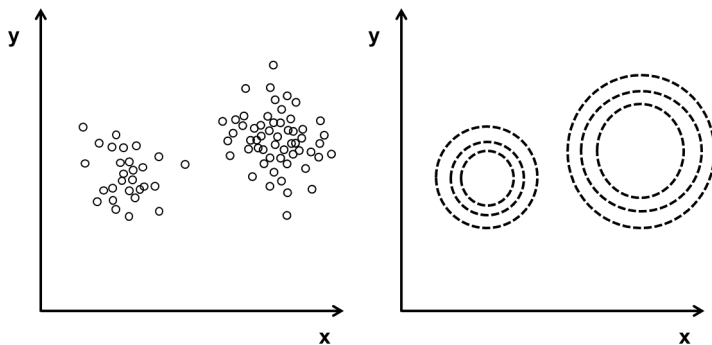


Abbildung 5.4: Beispiel eines Clustering (vgl. Russell und Norvig 2012, S. 945)

### 5.3.4 Verlustfunktion

Um den Lernerfolg während und nach dem Lernprozess bewerten zu können, bedarf es einer Verlustfunktion (vom engl. „loss function“). Mit dieser kann der Nutzen des Agenten optimiert werden (vgl. Russell und Norvig 2012, S. 828). Das Ziel dieser Funktion ist es, die Abweichung der gelernten Ausgabe von den Trainingsdaten zu messen. Der Lernerfolg ist umso schlechter, je größer der errechnete Verlust ist (vgl. Davidson-Pilon 2016, S. 127). Die Verlustfunktion ist allgemein definiert als  $L : f(\mathcal{X}) \times \hat{f}(\mathcal{X}) \rightarrow \mathbb{R}$ . Somit wird die Ausgabe der gelernten Funktion  $\hat{f}$  mit der bekannten vorgegebenen Ausgabe der Funktion  $f$  der Trainingsdaten verglichen. Dabei ist für jedes  $(x, y) \in \mathcal{X} \times f(\mathcal{X})$  der Verlust als  $L(y, \hat{f}(x))$  definiert, indem die Abweichung der Vorhersage  $\hat{f}(x)$  von dem Ergebnis  $y$  gemessen wird (vgl. Knox 2018, S. 6). Ziel des Lernens ist es,  $\hat{f}(x)$  an  $f$  anzunähern. Es existieren unterschiedliche Verlustfunktionen, je nachdem, um welches Problem es sich handelt. Im Folgenden werden zwei gängige Funktionen vorgestellt.

### Mittlerer absoluter Fehler

Der mittlere absolute Fehler (engl. „mean absolute error“ (MAE)) wird auch  $L_1$ -Verlust genannt. Für jede bekannte Ausgabe  $y$  und den zugehörigen vorhergesagten Wert  $\hat{y}$  berechnet sich der Verlust aus  $L_1(y, \hat{y}) = |y - \hat{y}|$  (vgl. Russell und Norvig 2012, S. 829, Hyndman und Koehler 2006). Diese Fehlerberechnung zeichnet sich durch eine stärkere Robustheit aus, d. h. Ausreißer beeinflussen das Ergebnis nicht so stark (vgl. Knox 2018, S. 17). Auf der anderen Seite kann eine Optimierung mit dem MAE zu Instabilitäten führen, da mehrere Optima existieren können. Über den ganzen Datensatz  $(x_1, \dots, x_n)$  betrachtet, berechnet sich der MAE nach Sammut und Webb (2010, S. 652) aus

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}|}{n}. \quad (5.1)$$

### Mittlerer absoluter prozentualer Fehler

Der mittlere absolute prozentuale Fehler (engl. „mean absolute percentage error“ (MAPE)) berechnet das Verhältnis der Differenz zwischen dem tatsächlichen Ausgabewert  $y$  und dem prognostizierten Wert  $\hat{y}$  zum tatsächlichen Ausgabewert  $y$  über alle Datenpunkte mit (vgl. Hyndman und Koehler 2006):

$$MAPE = \frac{100\%}{n} \sum \left| \frac{y - \hat{y}}{y} \right|. \quad (5.2)$$

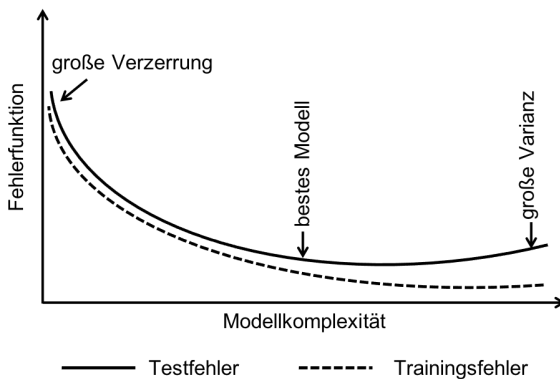
Relative Fehlerwerte haben in der Regel den Vorteil, dass sie skalunenabhängig sind (vgl. Hyndman 2006).

### Generalisierung, Überanpassung und Regularisierung

Kann ein Modell, das bereits mit gewissen Daten trainiert wurde, erfolgreich auf neue, ihm unbekannte Daten angewendet werden, spricht man von Generalisierung (vgl. Ertel 2016, S. 192). Zum Testen der Generalisierungsfähigkeit werden Testdaten benötigt, die das Modell noch nicht gesehen hat. Diese Daten müssen dem Agenten noch unbekannt sein – ansonsten würde er immer gut abschneiden,

da die Regeln, mit denen er entscheidet, auf den bereits bekannten Daten basieren (vgl. Ertel 2016, S. 195). Meist werden dazu die vorhandenen Daten in einen Trainings- und einen Testdatensatz aufgeteilt (siehe Kapitel 6.3.1).

Überanpassung (vom engl. „overfitting“) eines ML-Modells liegt vor, wenn nicht die zugrunde liegende Verteilung der genutzten Daten beschrieben wird, sondern ein vorhandenes Rauschen oder eine Varianz abgebildet wird. Als Rauschen wird eine zufällige und irrelevante Form der Daten verstanden, die nicht in das Modell einbezogen werden sollte. Die Folge ist ein Anstieg des Verlusts (Kapitel 5.3.4), wenn neue, nicht trainierte Daten auf das Modell angewandt werden (vgl. Webb 2010b, S. 744). Grob gesagt ist das Modell zu komplex für die verwendeten Daten (vgl. Raschka 2016, S. 65). So nimmt die Verlustfunktion der Trainingsdaten mit steigender Anpassung des Modells an die Daten im Normalfall konstant ab, während der Fehler bei den Testdaten zunächst abnimmt und bei weiterem Training im Anschluss jedoch erneut ansteigt. Das Phänomen der Überanpassung ist auch in Abbildung 5.5 dargestellt.



**Abbildung 5.5:** Idealisierter Test- und Trainingsverlauf. Optimales Modell am Punkt des geringsten Testfehlers (vgl. Prechelt 2012, S. 54, VanderPlas 2016, S. 366)

Auch der gegenteilige Effekt ist möglich: Bei der Unteranpassung (vom engl. „underfitting“) ist das Modell nicht komplex genug, um die Muster in den Daten abzubilden. Man sagt auch, das Modell hat eine große Verzerrung. Auch mit

neuen Daten wird das Modell nicht gut umgehen können (vgl. Raschka 2016, S. 65).

Eine der Schwierigkeiten des Maschinellen Lernens ist es, eine gute Abwägung zwischen einer hohen Verzerrung (Unteranpassung) und einer hohen Varianz (Überanpassung) zu finden (engl. „Bias-Variance-Tradeoff“). Dabei ist die Verzerrung (engl. „bias“) das Maß, wie sehr die Ausgabe des Modells bei anderen Datensätzen abweicht, die es nicht gelernt hat. Sie misst somit einen systematischen Fehler, der nicht durch Zufälligkeit verursacht ist. Andererseits gibt die Varianz (engl. „variance“) an, wie stark konsistent oder variabel das Modell ist, wenn es mehrfach mit einem unterschiedlichen Teilbereich eines Datensatzes trainiert wird (vgl. Raschka 2016, S. 66).

## 5.4 Künstliche Neuronale Netze

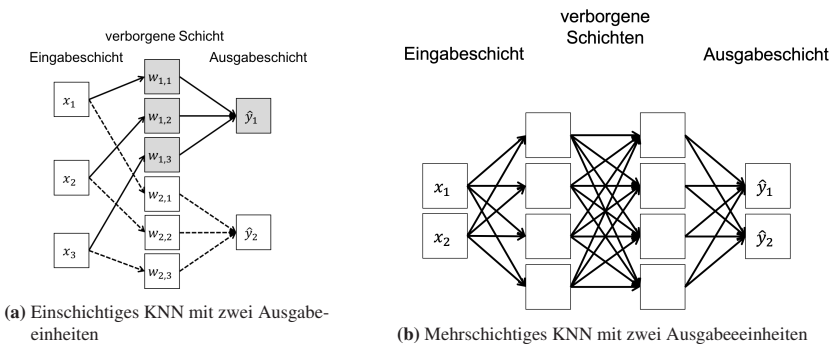
„Neuronale Netze in der künstlichen Intelligenz versuchen die Arbeitsweise des menschlichen Gehirns nachzubilden. [...] Man [möchte] Computerprogramme [...] befähigen, in analoger Weise [zum menschlichen Gehirn] aus Beispielen Gelerntes auf neue Situationen anzuwenden.“ (Lämmel und Cleve 2008, S. 195)

KNN (engl. „artificial neural networks“) sind ein ML-Algorithmus, dessen Aufbau und Funktion sich am Nervensystem und Gehirn von Mensch und Tier orientieren (vgl. Oksenberg und Kalton 1991, Braun 1997, S. 207 f.). Nachdem in der Biologie das Konzept der Neuronen als elektrochemische Transmitter von Informationen in Organismen entdeckt worden war, kam man in der Folge auf die Idee, dieses Konzept durch Maschinen nachzubilden. Das Konzept des biologischen Neurons wurde durch McCulloch und Pitts (1943) in die Mathematik und Datenverarbeitung übertragen. Es berücksichtigte ausschließlich die binäre Ausgabe, das heißt, wenn die ankommenden Signale einen gewissen Schwellenwert überschreiten, wird ein fixes Ausgabesignal generiert (vgl. Raschka 2016, S. 18).

Der Aufbau besteht aus mehreren parallel arbeitenden Neuronen. Diese Neuronen können über gerichtete Verbindungen Informationen zueinander übertragen (vgl. Kruse et al. 2015, S. 7).

Ein KNN folgt dem Konzept eines biologischen Neurons insoweit, als es aus mehreren Neuronen besteht, die in Schichten aufgeteilt sind. Ein einfaches KNN besteht aus mindestens drei Schichten (engl. „layer“). Die erste Schicht wird als Eingangsschicht (engl. „input layer“) bezeichnet, bei der den Neuronen die Eingangswerte des Modells übergeben werden. Darauf folgen eine oder mehrere verborgene Schichten (engl. „hidden layer“), die für das Verarbeiten der Eingangsdaten zuständig sind. Zum Schluss liefert die Ausgangsschicht (engl. „output layer“) die Werte, die das Verfahren ausgeben soll. Die Neuronen besitzen jeweils eine Aktivierungsfunktion und gewichtete Verbindungen zu Neuronen aus anderen Schichten. Die Aktivierungsfunktion eines Neurons bestimmt nun, in welcher Höhe ein Signal weitergegeben wird. Die Gewichtung der Verbindungen ist entscheidend, wie stark das Signal beim nächsten Neuron ankommt. Die Anpassung dieser beiden Parameter entspricht nun dem ML eines KNN. Da die Optimierung jedoch für jedes Neuron und jede Verbindung erfolgen muss, steigt der Lernaufwand exponentiell mit der Knoten- bzw. Verbindungsanzahl.

Eine grafische Darstellung eines KNN mit einer oder mehreren verborgenen Schichten findet sich in Abbildung 5.6.



**Abbildung 5.6:** Verschiedene KNN (nach Russell und Norvig (2012, S. 848))

## 5.5 Deep Learning

DL (engl. etwa „tiefes Lernen“) ist ein in der jüngeren Zeit häufig anzutreffender Begriff im Bereich der KI. Er wurde durch Dechter (1986) für ML allgemein und durch Aizenberg et al. (2000) für KNN geprägt. Anwendungsgebiet sind anpassungsfähige Systeme, die lange Ketten potenzieller kausaler Verknüpfungen zwischen Aktionen und deren Konsequenzen enthalten (vgl. Schmidhuber 2017, S. 338 ff.). So gehören mehrschichtige KNN ebenfalls in den Themenbereich DL.

Mit der Anforderung der Erklärbarkeit von KI-Modellen beschäftigt sich der Teilbereich „XAI“, der im Folgenden beschrieben wird.

## 5.6 Ausgewählte Methoden der erklärbaren Künstlichen Intelligenz

### 5.6.1 Ziele der erklärbaren Künstlichen Intelligenz

Modelle der KI (z. B. KNN) bauen meistens auf einem „Black Box“-Ansatz auf und sind im Allgemeinen nicht nachvollziehbar. Zielgruppe dieser Modelle sind daher vor allem die Entwickler der Modelle selbst (vgl. Miller et al. 2017, S. 1). Im Vergleich zu einfachen statistischen Methoden können andere Anwender neben dem Entscheidungsprozess hin zum Ergebnis auch den algorithmischen Aufbau des Modells schwer nachvollziehen. Folge können Misstrauen und eine Unzuverlässigkeit des entwickelten Modells sein.

Jedoch sind Erklärungen u. a. relevant, um Vertrauen in die Entscheidungen zu bilden, Kausalitäten ableiten zu können, die Ergebnisse zu transferieren, Informationen zu erhalten, faire Schlussfolgerungen ziehen zu können und mit dem Modell zu interagieren (siehe Kapitel 2.6) (vgl. Arrieta et al. 2020). Erklärungen sind eng mit dem Begriff der Interpretierbarkeit verbunden (vgl. Biran und Cotton 2017, Anjomshoae et al. 2019). Die Interpretierbarkeit setzt ein gewisses

Vorwissen über Kategorien, Beziehungen und Aktivitäten beim Anwender voraus (vgl. Langley et al. 2017, S. 1). Systeme sind allgemein interpretierbar, wenn die Prozesse von einem Menschen verstanden werden können (vgl. Biran und Cotton 2017).

Arrieta et al. (2020) definieren drei Arten von Interpretierbarkeit im Zusammenhang mit XAI. Dies sind die algorithmische Transparenz, die Zerlegbarkeit und Simulierbarkeit.

Die „Simulierbarkeit“ bezeichnet die Fähigkeit eines Modells von einem Menschen simuliert werden zu können. Damit muss es in sich geschlossen sein, so dass der Mensch es als Ganzes betrachten kann. Nach dem Aspekt der Simulierbarkeit sind einfache lineare Modelle besser interpretierbar als dichte Modelle (vgl. Tibshirani 1996).

Die „Zerlegbarkeit“ steht für die Möglichkeit, das Modell in die Bestandteile der Inputparameter, der Modellparameter und seine Kalkulation zu unterteilen. Jedes dieser Elemente muss daher für den Menschen ohne weitere Werkzeuge interpretierbar sein.

Unter der „algorithmischen Transparenz“ ist die Nachvollziehbarkeit des Prozesses zu verstehen, den das Modell durchführt. Damit muss das Modell durch mathematische Analysen und Methoden vollständig erklärbar sein.

Neben diesen drei Kriterien kann noch die „nachträgliche Erklärbarkeit“ von Black Box-Modellen gesehen werden. Dies kann durch die Erklärung bspw. durch Texte, Visualisierungen, Beispiele, einem abstrahierten Modell oder das Aufzeigen der Relevanz der Merkmale unterstützt werden.

Nach Arrieta et al. (2020) umfassen XAI-Methoden lineare Regressionsmodelle, Entscheidungsbäume, Nächste-Nachbarn-Klassifikation, regelbasierte Lernsysteme, allgemeine additive Modelle und Modelle der bayesschen Statistik. Bewertet man diese Modelle nach den drei Kriterien der Interpretierbarkeit, so ist bis auf die Entscheidungsbäume mindestens jeweils eins der Kriterien nicht gegeben. Hier müssen die Modelle zur Simulierbarkeit eingeschränkt werden. Die



Anzahl der Variablen und Regeln ist zu hoch, Beziehungen zwischen den Variablen sind zu komplex (Zerlegbarkeit). Oder es sind weitere mathematische und statistische Werkzeuge notwendig, um die Modellergebnisse zu verstehen. Entscheidungsbäume hingegen erfüllen alle drei Kriterien: Der Entscheidungsweg kann durch den Menschen verstanden werden (Simulierbarkeit); das Modell umfasst Regeln, die die Daten nicht verändern und die Lesbarkeit bewahren (Zerlegbarkeit) und das Ergebnis zeigt lesbare Regeln bzw. Wissen, das aus den Daten gelernt wurde (algorithmische Transparenz).

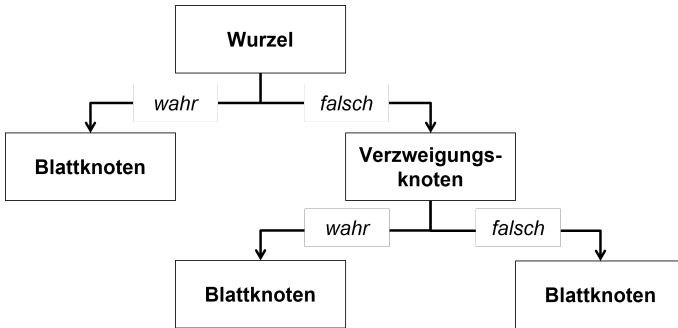
Zusammengefasst beschäftigt sich der Themenbereich XAI mit inhärent erklärbaren Modellen und Methoden zur nachträglichen Erklärung von komplexen Black Box-Modellen, wie bspw. KNN (vgl. Molnar 2019). Im Folgenden werden beide Varianten an beispielhaften Modellen vorgestellt. Auf Basis der Schlussfolgerungen von Arrieta et al. (2020) wird der Entscheidungsbaum als inhärent erklärbare Methode betrachtet. Hingegen wird das SHAP-Framework zur nachträglichen Erklärung von Black Box-Modellen angewendet.

## 5.6.2 Entscheidungsbäume

Entscheidungsbäume sind Teil der inhärent erklärbaren Modelle. Sie nehmen einen Vektor mit Eingabewerten entgegen und geben einen einzelnen Zielwert aus. Dabei können sowohl Ein- als auch Ausgabewerte diskret oder stetig sein (Regression und Klassifizierung, siehe Kapitel 5.3.3).

Entscheidungsbäume als Teil von KI-Modellen teilen den Datensatz an Attributen rekursiv, sodass mit der entstehenden Struktur eine möglichst genaue Prognose entsteht. Abbildung 5.7 zeigt eine vereinfachte Darstellung eines Entscheidungsbaums (vgl. Myles et al. 2004). Ausgehend von der Wurzel teilt sich der Entscheidungsbaum an den Ästen oder Kanten in Verzweigungsknoten oder Blattknoten. Die Blattknoten zeigen die Prognose an.

Es lassen sich aus der Darstellung des Entscheidungsbaumes weiter wenn-dann-Regeln erstellen. Damit ist der Entscheidungsbaum eine prädiktive Methode, die



**Abbildung 5.7:** Vereinfachte Darstellung eines Entscheidungsbaums (vgl. Myles et al. 2004)

Blattknoten können jedoch auch weiter verwendet werden, um Optimierungen durchzuführen.

Der große Vorteil der Verwendung von Entscheidungsbäumen liegt daher in ihrer Anschaulichkeit, sodass der Mensch sie leicht verstehen und interpretieren kann (vgl. Ertel 2016, S. 217).

Das Grundprinzip während des Trainings eines Entscheidungsbaumes ist wie folgt: Innerhalb des Modells wird an jeder Gabel des Entscheidungsbaums eine Aufteilung nach einem oder mehreren Attributen innerhalb von Verzweigungs-knoten vorgeschlagen. Auf dieser Basis wird der Vorteil jeder Aufteilung berechnet und jene mit dem höchsten Zugewinn gewählt. Der höchste Gewinn bezieht sich, intuitiv gesprochen, auf die Aufteilung, die die Daten in die kohärentesten Teilmengen trennt. Dies wird als eine Optimierung des Informationsgewinns (engl. „information gain“) bezeichnet (vgl. Rokach und Maimon 2005). Je kleiner dieser Zugewinn wird, desto besser ist die gewählte Struktur, da sie zu einer genaueren Prognose führt. Die Blattknoten geben die Prognose mit den zuvor angegebenen Attributen wieder.

Entscheidungsbäume des ML kombinieren eine Vielzahl an einzelnen Entscheidungsbäumen, um ein leistungsfähiges Modell zu erstellen und die Prognose zu verbessern. Damit gehören sie zur maschinellen Lerntechnik des „Ensembling“.

Innerhalb von Entscheidungsbäumen kann zwischen zwei Varianten unterschieden werden, dem „Bagging“ und dem „Boosting“. Beim Bagging werden parallel und unabhängig Entscheidungsbäume erstellt. Die Prognosegenauigkeit gibt den Durchschnitt aller Bäume wieder. Durch die Parallelisierung ist das Modell abhängig von der Datenmenge entweder sehr schnell oder beansprucht mehr Zeit. Beim Boosting hingegen werden sequentiell Bäume erstellt und auf den vorherigen Ergebnissen angepasst, um die Vorhersage schrittweise zu verbessern. Dabei werden die Ergebnisse der einzelnen Bäume entsprechend ihrer Prognosegenauigkeit gewichtet. Die Prognosegenauigkeit des Modells ist die Summe der gewichteten Ergebnisse. (vgl. Schapire et al. 1998, Breiman 1996)

Im Vergleich zu anderen KI-Methoden können Entscheidungsbäume sehr gut mit fehlenden Attributen umgehen (vgl. Sheh 2017, S. 631), nehmen Kategorien in der Prognose mit auf und zeigen bereits mit einer kleinen Datenmenge gute Ergebnisse. Fortgeschrittene Erweiterungen von Entscheidungsbäumen sind bspw. der Random Forest (vgl. Liaw et al. 2002), Gradient Boosting Regression (GBR), eXtreme Gradient Boosting (XGBoost) Tree (vgl. Chen und Guestrin 2016), Light Gradient Boosting Machine (LightGBM) (vgl. Ke et al. 2017) oder der CatBoost (vgl. Prokhorenkova et al. 2018), die darauf basieren, mehrere Entscheidungsbäume miteinander zu kombinieren. Der Random Forest basiert auf dem „Bagging“. XGBoost, LightGBM und CatBoost verwenden dasselbe zugrundeliegende Framework, „Boosting“, unterscheiden sich aber nur in Nuancen in ihrer Implementierung.

**Random Forest:** Mit dem Random Forest werden zufällig viele Entscheidungsbäume unkorreliert erstellt. Dies unterstützt eine geringe Anfälligkeit für eine Überanpassung. Die Bäume werden nach der Methode des Bagging hin zu einem „Wald“ aus zufälligen und leicht unterschiedlichen Entscheidungsbäumen erstellt. Jeder Baum trägt einen Teil zur Prognose bei. (vgl. Liaw et al. 2002)

**GBR:** Für den GBR (deutsch „Gradientenverstärkung“) wird das Boosting verwendet. Nach der Analyse des ersten Baumes wird das Gewicht für die

schwierige Regression bzw. Klassifikation erhöht und für eine einfache Regression bzw. Klassifikation verringert. Es werden demnach in jeder Iteration neue Prädiktoren erstellt, um schrittweise bessere Vorhersagen generieren zu können. Die Vorhersage des letzten Modells enthält die gewichteten Prognosen der vorherigen Modelle. Der GBR kann keine Kategorien einbinden und eine vorherige Transformation in numerische Werte ist notwendig. Anwendung kann hier das „one-hot encoding“ finden, das jede Kategorie in eine Spalte überführt und Zellen entsprechend mit „1“ oder „0“ befüllt (vgl. Dorogush et al. 2018). Durch das sequentielle Vorgehen nimmt mit Anzahl der Merkmale und Datenpunkte die Rechenkapazität zu.

**XGBoost:** Im Vergleich zum GBR existieren hier einige Erweiterungen. Dazu gehört eine intelligente Bestrafung von Bäumen (siehe auch Kapitel 5.3.4), eine zusätzliche Randomisierung der Parameter zur Verringerung der Korrelation zwischen Bäumen, ein proportionales Schrumpfen der Endknoten, um die Vergleichbarkeit dieser zu erhöhen, sowie das Newton-Boosting. Das Newton-Boosting bietet einen direkteren Weg zum minimalen Gewicht als der Gradientenabstieg. (vgl. Chen und Guestrin 2016)

**LightGBM:** Der LightGBM beinhaltet ein gradientenbasiertes einseitiges Sampling (engl. „Gradient one-sided sampling“ (GOSS)). Hierbei wird der Baum schrittweise einseitig entlang einzelner Knoten in die Tiefe gegangen. Instanzen mit größeren Gradienten zur Berechnung des Informationsgewinns werden behalten und kleine Gradienten ausgeschlossen. Damit wird eine höhere Prognosegenauigkeit erreicht. Dies ist insbesondere bei einer breiten Informationsgrundlage sinnvoll. Der Algorithmus bezieht weitere Merkmale mit ein, die als Kategorien beschrieben sind. (vgl. Ke et al. 2017)

**CatBoost:** Der CatBoost ist eine fortgeschrittene Methode des Boosting (vgl. Prokhorenkova et al. 2018), da es zwei wesentliche Erweiterungen zum XGBoost oder LightGBM gibt: erstens das „Ordered Target Encoding“ für die Behandlung von Kategorien und zweitens das „Ordered Boosting“ für

eine höhere Prognosegenauigkeit. Mit der Anwendung des „one-hot encoding“ besteht die Gefahr einer großen zusätzlichen Anzahl neuer Merkmale (vgl. Prokhorenkova et al. 2018, S. 2) und damit eines „Overfittings“. Beim Ordered Target Encoding werden die Daten zufällig sortiert und es wird eine „künstliche Zeit“ eingeführt. Diese Zeit wird verwendet, um eine Zielkodierung Schritt für Schritt pro Zeitstempel durchzuführen. (vgl. Prokhorenkova et al. 2018, S. 2ff.)

Beim Ordered Boosting werden weiter verschiedene Permutationen auf Basis des Datensatzes einbezogen: So werden bei der ersten Permutation die ersten Knotennachbarn berechnet, bei der zweiten Permutation die nächsten usw. (vgl. Prokhorenkova et al. 2018, S. 5ff.)

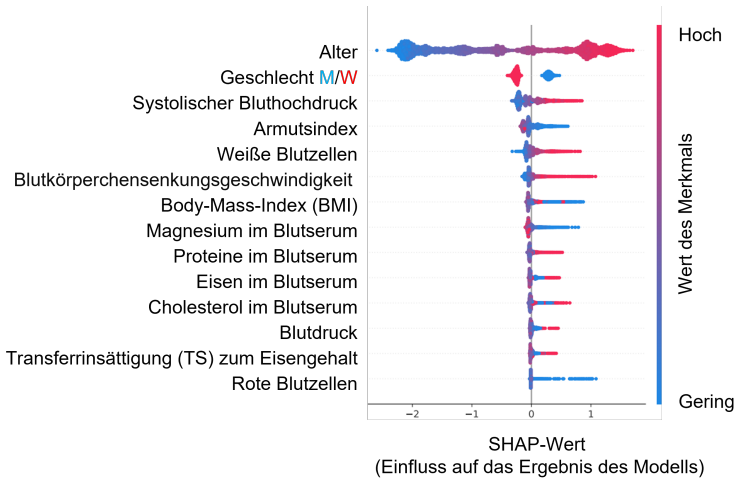
### 5.6.3 SHAP-Framework

Das SHAP-Framework unterstützt die nachträgliche Interpretierbarkeit von „Black Box“ KI-Modellen, indem relevante Merkmale identifiziert und bezüglich ihres Einflusses zur Prognose bewertet werden (vgl. Lundberg et al. 2018). Auch kann das SHAP-Framework dazu dienen, relevante Merkmale einheitlich zu bewerten. Werden bspw. verschiedene Entscheidungsbäume verwendet, so werden die Merkmale durch das Modell und die Wahl der Modellparameter unterschiedlich angeordnet und die Bewertung der Modelle ist nicht konsistent. SHAP dient demnach als einheitliches Bewertungssystem. Shapely-Regressionswerte stammen aus der Spieltheorie und weisen jedem Merkmal einen SHAP-Wert zu (vgl. Shapely 1953). Dieser Wert zeigt in vereinfachter Form den spezifischen Beitrag eines Merkmals zur Prognose in einem Datensatz. Dabei sind  $F$  und  $S$  eine Teilmenge aller Modellmerkmale  $N$  und  $i$  hieraus ein spezifisches Merkmal. Es wird davon ausgegangen, dass die Merkmale voneinander unabhängig sind. Mit  $\theta_i$  wird der Wert von  $i$  zur Prognose mit und ohne das Merkmal berechnet. Dabei werden alle möglichen Permutationen ermittelt. Der Wert von  $S$  wird Shapely-Wert  $\theta_i$  bezeichnet und wie folgt berechnet (Lundberg 2018, S. 3):

$$\theta_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(F - |S| - 1)!}{F!} [f_x(S \cup \{i\}) - f_x(S)] \quad (5.3)$$

Dabei ist  $f_x(S \cup \{i\})$  der Wert mit dem Merkmal und  $f_x(S)$  ohne das spezifische Merkmal. Die Differenz wird für alle Teilmengen  $S \subseteq N \setminus \{i\}$  berechnet.

Mit dem SHAP-Framework werden je Datensatz die Shapely-Werte zu den einzelnen Merkmalen berechnet und nach ihrem gemittelten Wert angeordnet. Durch die gemittelte Betrachtung aller Permutationen kann eine konsistente Vorhersage getroffen werden (vgl. Lundberg 2018).



**Abbildung 5.8:** Beispielhafte Darstellung eines SHAP-Frameworks (vgl. Lundberg 2018)

Abbildung 5.8 zeigt ein SHAP-Framework für einen Datensatz von Miller (1973) mit 18 Merkmalen für eine Mortalität innerhalb von 20 Jahren. Die Bewertung des globalen Einflusses wird absteigend angeordnet (engl. „Feature Importance“). Der Wert des Merkmals (engl. „Feature Value“) gibt hingegen den Originalwert aus der Datenbasis an. Dieser wird mit der entsprechenden Farbskala in der Grafik angezeigt. Der SHAP-Wert gibt den Wert des einzelnen Merkmals zur

Prognose wieder. Dabei wird jeder Datenpunkt bzw. jede Person als ein Punkt dargestellt. Die horizontale Achse zeigt demnach den positiven oder negativen Einfluss auf die Prognose. Damit stellt das SHAP-Framework die globale Relevanz der Merkmale zur Zielvariablen sowie gleichzeitig auch lokale Werte durch den SHAP-Wert dar.

## 5.7 Zusammenfassung

KI beschäftigt sich mit der Entwicklung von Maschinen, die Intelligenz imitieren (vgl. McCarthy et al. 1955). Hierbei liegt jedoch die Herausforderung in der Bedeutung des Wortes Intelligenz, da dieses u. a. ein Bewusstsein, Emotionalität und Intuition einbezieht. Da die Definition des ML enger zu fassen ist, wird diese im Folgenden als Synonym für KI verwendet. ML als Teilbereich der KI gibt Computern die Möglichkeit zu lernen, ohne dabei explizit programmiert zu werden (vgl. Samuel 1959). Im Vergleich zur klassischen Programmierung werden hier Daten und Antworten gegeben, aus denen der Computer eigenständig Regeln ableitet (vgl. Chollet 2018, S. 23). Das Modell passt sich demnach bei einer sich verändernden Datenmenge automatisch an diese Regeln an.

Durch die zunehmende Menge und Vielfalt von Daten nehmen auch die Anforderungen an die Modelle immer weiter zu. Viele der ML-Modelle sind sogenannte „Black Box“-Modelle, die für den Anwender nicht erklärbar bzw. interpretierbar sind. Mit der Erklärbarkeit von ML-Modellen beschäftigt sich der Bereich der erklärbaren KI (XAI). Es kann zwischen ML-Modellen unterschieden werden, die inhärent erklärbar sind, und Modellen, die zur nachträglichen Interpretation von Black Box-Modellen verwendet werden. Erklärbare inhärente ML-Modelle zeichnen sich durch ihre Simulierbarkeit, Zerlegbarkeit und algorithmische Transparenz aus (vgl. Arrieta et al. 2020). Diesen drei Anforderungen genügen Entscheidungsbäume ohne den Einsatz zusätzlicher Hilfsmittel.

Aus diesem Grund werden im Folgenden Entscheidungsbäume als Methode für inhärent erklärbare ML-Modelle verwendet. ML-Entscheidungsbäume sind u. a.

der Random Forest (vgl. Liaw et al. 2002), GBR, XGBoost Tree (vgl. Chen und Guestrin 2016), LightGBM (vgl. Ke et al. 2017) oder CatBoost (vgl. Prokhorenkova et al. 2018). ML-Entscheidungsbäume basieren darauf, mehrere Bäume intelligent miteinander zu kombinieren. Das SHAP-Framework hingegen eignet sich, um Black Box-Modelle nachträglich zu erklären und relevante Merkmale des Modells konsistent bewerten zu können.

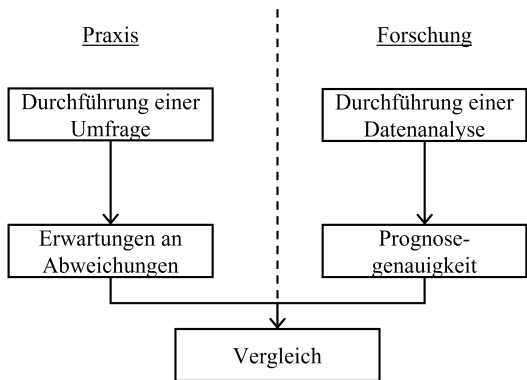
Methoden der XAI, wie ML-Entscheidungsbäume oder das SHAP-Framework, stellen analytische Planungsmethoden dar, die die Prognosegenauigkeit und Erklärbarkeit der Prognose von Dauern in frühen Phasen erhöhen bzw. verbessern könnten. Die vorgestellten XAI-Methoden sowie KNN als Stand der Forschung im Bereich des ML werden in dieser Arbeit weiter betrachtet. Im folgenden Kapitel wird eine Datenanalyse an beispielhaften Datensätzen durchgeführt, die die Planungsgenauigkeit verschiedener analytischer Planungsmethoden miteinander vergleicht.



# 6 Vergleich und Bewertung ausgewählter analytischer Planungsmethoden

## 6.1 Vorbemerkung

In diesem Kapitel werden verschiedene analytische Planungsmethoden hinsichtlich ihrer Prognosegenauigkeit miteinander verglichen. Insbesondere werden lineare Regressionsmodelle und KNN als State-of-the-Art sowie Entscheidungsbäume als Teilbereich prädiktiver Planungsmethoden untersucht. Im Hinblick auf die Planung von Dauern werden lineare Regressionsmodelle in der Literatur am häufigsten untersucht und KNN stellen den aktuellen Einsatz von ML-Methoden zur Prognose von Dauern einzelner Projektphasen dar (siehe auch Kapitel 3). Die Fragestellung der zweiten Forschungsfrage in dieser Arbeit lautet: Welche ausgewählten Modelle analytischer Planungsmethoden sind im Vergleich in frühen Planungsphasen zur Prognose von Dauern dienlich? Hierzu werden zuerst mit einer Umfrage Ziele für die Prognosegenauigkeit von Dauern definiert (Kapitel 6.2). Kapitel 6.3 beschäftigt sich mit einer Datenanalyse anhand beispielhafter Datensätze. Hierfür werden die verwendeten Datensätze vorgestellt und entsprechend den Ergebnissen der als relevant identifizierten Merkmale aus Kapitel 3.5 und Kapitel 4 bewertet. Die vorliegenden Datensätze enthalten Informationen über Projekte in der Phase Bauausführung. Die vorherzusagende Dauer wird demnach auf die Phase der Bauausführung beschränkt. Das Vorgehen ist in Abbildung 6.1 zusammengefasst. Die Ergebnisse der angewendeten Methoden



**Abbildung 6.1:** Zusammenfassung des Vorgehens in Kapitel 6

werden anschließend verglichen und mit den Anforderungen an die Prognosegenauigkeit auf ihre Anwendbarkeit bewertet. Weiter zeigt das SHAP-Framework relevante Merkmale auf, die in der Anwendung der prädiktiven Planungsmethoden als relevant einzuordnen sind. Kapitel 6.4 fasst die wesentlichen Punkte zusammen und definiert über eine kritische Betrachtung weitere Punkte zur Forschung.

## 6.2 Herleitung der Anforderungen an die Prognosegenauigkeit

### 6.2.1 Ziele der Umfrage

Zur Rahmenterminplanung eines Bauprojektes sind eine Vielzahl an Entscheidungen bei unvollständigen Informationen notwendig. Je nach Planungsstand sind durch die fehlenden Informationen Folgen von Entscheidungen und deren Eintrittswahrscheinlichkeiten häufig unbekannt. (vgl. Rogel und Osebold 2013, S. 211)

Mit welchem Streuungsbereich von Abweichungen Beteiligte der Baubranche in der DACH-Region (Deutschland, Österreich, Schweiz) zufrieden sind, wurde in bisherigen Studien nicht quantifiziert. Diese Quantifizierung stellt ein Akzeptanzlevel dar, nach der die Zeitreserven im Terminplan optimiert werden können. Dieser zufriedenstellende Streuungsbereich von Abweichungen ist als Momentaufnahme zu verstehen. Mit der kontinuierlichen Verbesserung von Strategien und Methoden zur Vorhersage von Dauern ist es das Ziel, die Planabweichungen und damit deren Akzeptanzlevel zu reduzieren. Die Quantifizierung des Streuungsbereichs kann demnach ebenfalls bei weiteren Aufnahmen zu späteren Zeitpunkten als Anhaltspunkt für die Weiterentwicklung von Strategien und Methoden der Rahmenterminplanung verstanden werden.

Diese explorative Studie in Form einer Umfrage untersucht die Zufriedenheit der an Bauprojekten beteiligten Personen mit zeitlichen Abweichungen der Bauausführung von der jeweiligen Terminplanungsstufe. Abweichungen werden im Sinne von Überschreitungen und Unterschreitungen der geplanten Dauer verstanden (siehe auch Kapitel 2.5). Überschreitungen bedeuten eine längere Ausführung. Unterschreitungen bedeuten eine schnellere Fertigstellung gegenüber der Planung. Zufriedenstellende Abweichungen werden prozentual und nicht absolut abgefragt, um eine Vergleichbarkeit bei unterschiedlichen Projektdauern gewährleisten zu können. Es werden im Hinblick auf die Bauausführung die Terminpläne der Rahmenterminplanung, Generalablaufplanung, Grobterminplanung und Bauausführungsplanung (vgl. HOAI 2018) betrachtet. Damit lässt sich der zufriedenstellende Streuungsbereich von Abweichungen zwischen der Rahmenterminplanung und ihrer Realisierung auch mit den weiteren Detaillierungsstufen des Terminplans vergleichen.

## 6.2.2 Methodisches Vorgehen der Umfrage

Der Aufbau der Umfrage beinhaltet zuerst vier Fragen zur demografischen Zusammensetzung der Teilnehmer. Hierzu zählen die Rolle des Unternehmens (auftraggeber- (bauherren-) oder auftragnehmerseitige Rolle, beratende Rolle),

dessen Spezialisierung (Hochbau, Industriebau, Schlüsselfertigbau, Ingenieurbau, Tiefbau, Verkehrsbau, Fertighäuser und Sonstige), die Unternehmensgröße (gemessen an Mitarbeitern und Umsatz) und die Berufserfahrung des Teilnehmers. Weiter wurde jeweils eine Matrix-Frage zu Über- und Unterschreitungen gestellt. Die Fragestellung hierzu lautete jeweils zu den vier Terminplantypen: „Ich bin immer noch mit dem Ergebnis des Bauprojektes zufrieden, wenn die Bauausführung um maximal ... % langsamer (bzw. schneller) gegenüber der entsprechenden Terminplanung ist.“ Dies ergibt in Summe sechs Fragen.

In der Einleitung der Umfrage werden deren Zweck, die erhebende Organisation, Kontaktdaten bei Rückfragen und die anonyme Behandlung der Daten beschrieben.

Mit zwei Personen wurde ein Pre-Test durchgeführt. Die erste Person stammt aus der Forschung und gab Feedback zu Fachbegriffen sowie zur Struktur der Umfrage. Die zweite Person, eine branchenfremde Person, überprüfte die Umfrage nochmals auf Verständlichkeit.

Die Umfrage wurde an Beteiligte der Baubranche in der DACH-Region verteilt. Hierzu zählen Auftraggeber (Bauherren), deren Auftragnehmer und Berater. Die Umfrage war im November 2021 für vier Wochen verfügbar.

Über LinkedIn-Gruppen, persönliche Kontakte und die Kontaktaufnahme mit Experten der Terminplanung wurde der Link zur Umfrage verteilt. Hier wurde die Person jeweils gebeten, den Link an eine weitere Person zu verteilen, um eine größere Reichweite der Umfrageteilnehmer zu erhalten.

Es nahmen insgesamt 87 Teilnehmer teil, davon haben 66 Teilnehmer die Umfrage vollständig beendet. Dies resultiert in einer Abbruchquote von 24,1 %. Nach Hill (1998, S. 7) sind in explorativen Studien bereits zehn bis 30 Teilnehmer ausreichend. Damit sollen statistische Auswertungen durchgeführt werden können und schwache Zusammenhänge in den Daten sichtbar sein. Roscoe (1975, S. 163) und Abranovic (1997, S. 307) definieren hingegen ein unteres Limit der Stichprobe von 30. Chassan (1979) fordert weiter eine untere Grenze der Stichprobe

zwischen 50 und 100. Nach diesen Aussagen ist die Stichprobengröße der durchgeführten Umfrage ausreichend. Die Teilnahme dauerte im Median drei Minuten. Abgebrochen wurde die Umfrage insbesondere nach den demografischen Fragen zur Rolle, Spezialisierung, Unternehmensgröße und Berufserfahrung.

### 6.2.3 Ergebnisse der Umfrage

Die vollständigen Antworten sind auf Mendeley veröffentlicht<sup>4)</sup>. Auf Basis der Ergebnisse werden univariate und bivariate Analysen durchgeführt. In den univariaten Analysen wird die Demografie der Teilnehmer betrachtet, aber auch der zufriedenstellende Streuungsbereich von Abweichungen wird analysiert.

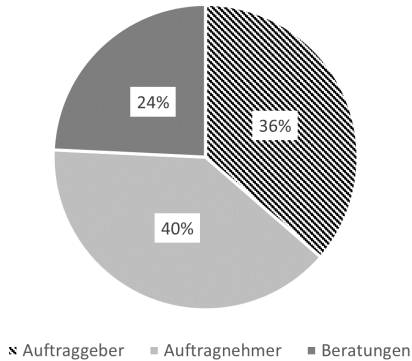
Die 66 Teilnehmer teilen sich auf in Auftraggeber, seine Auftragnehmer und Berater (siehe Abbildung 6.2). Die 36 % teilgenommenen Auftraggeber (entsprechen 24 Personen) bestehen aus Eigenbedarfsbauherren, Bauträgern, BOT-Bauherren, Investoren und der Projektsteuerung/Projektmanagement. Die 40 % Auftragnehmer setzen sich aus Bauunternehmen, Planern und Nachunternehmern zusammen. Zuletzt nahmen 24 % Berater teil, die Auftraggeber und Auftragnehmer vertreten können.

Bei der Spezialisierung des Unternehmens war eine Mehrfachauswahl möglich. Die Unternehmen sind vor allem auf den Hochbau (44 Angaben), den Schlüsselfertigbau (31 Angaben), den Industriebau (29 Angaben) und Ingenieurbau (21 Angaben) spezialisiert.

---

<sup>4)</sup> <https://data.mendeley.com/datasets/2275rhvkt7/1>

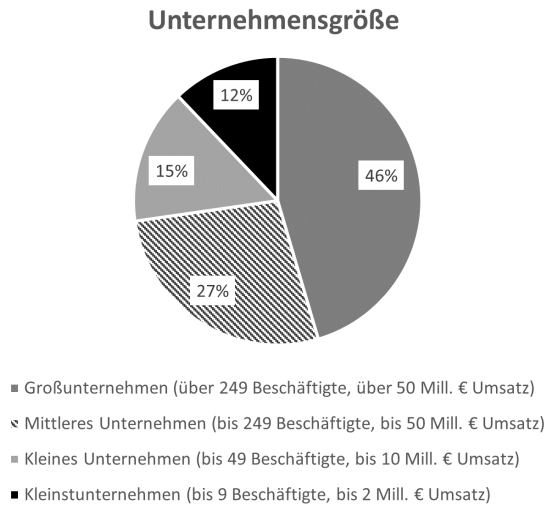
### Rolle des Unternehmens



**Abbildung 6.2:** Ergebnisse der Umfrage: Rolle des Unternehmens aufgeteilt in Auftraggeber, Auftragnehmer und Beratungen

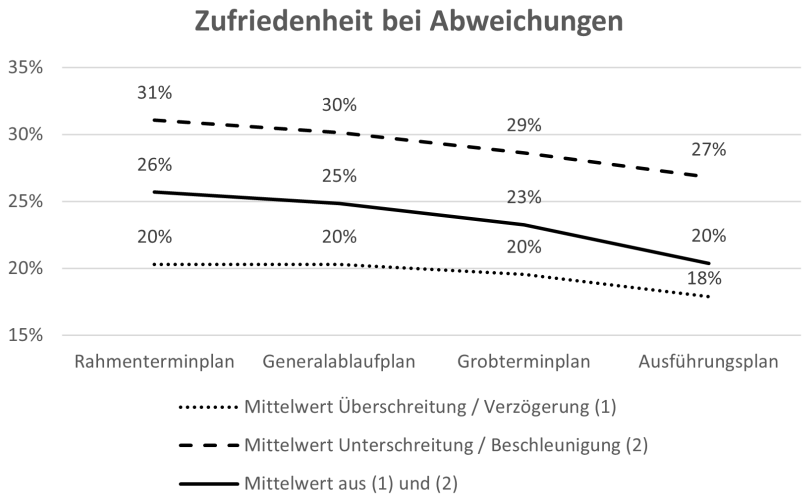
Weiter nahmen an der Umfrage überwiegend Großunternehmen und mittlere Unternehmen teil (siehe Abbildung 6.3). Zu beachten ist, dass dies nicht den Markt der Baubranche in der DACH-Region repräsentiert, der vor allem durch Kleinunternehmen und kleine Unternehmen geprägt ist.

Die Teilnehmer haben im Mittelwert 12,3 Jahre Berufserfahrung. Dabei streute die Berufserfahrung zwischen einem und 63 Jahren.



**Abbildung 6.3:** Ergebnisse der Umfrage: Unternehmensgröße anhand Mitarbeiteranzahl und Umsatz

Wie in Abbildung 6.4 zu sehen ist, nimmt erwartungsgemäß die Grenze der Zufriedenheit bei Abweichungen mit zunehmendem Detaillierungsgrad des Terminplans ab: Sind bei der Rahmenterminplanung die Teilnehmer durchschnittlich noch zufrieden, wenn sich die Realisierung um 20 % von der hier geplanten Fertigstellung verzögert, so sind dies in der Ausführungsplanung 18 %. Ist das Bauprojekt schneller als geplant, werden bei der Rahmenterminplanung durchschnittlich 31 % Abweichungen akzeptiert, bei der Ausführungsplanung 27 %. Nimmt man den Mittelwert je Teilnehmer aus Über- und Unterschreitungen der Plandauer, kann die gleiche Tendenz betrachtet werden. Hier kann von einer Zufriedenheit bei 26 % Abweichungen vom Rahmenterminplan ausgegangen werden, bei der Ausführungsplanung sind es 20 %.



**Abbildung 6.4:** Ergebnisse der Umfrage zur Zufriedenheit bei Abweichungen des jeweiligen Terminplans

Vergleicht man die Zufriedenheit bei Abweichungen von Auftragnehmern zu Auftraggebern anhand der Rahmenterminplanung (siehe Tabelle 6.1), so sind Auftragnehmer bei Verzögerungen von 27 % noch zufrieden. Hingegen sind Auftraggeber durchschnittlich nur mit 16 % Überschreitungen des Plans zufrieden. Ist die Bauausführung schneller als geplant, haben Auftraggeber und Auftragnehmer eine annähernd gleiche Erwartungshaltung. Auftraggeber sind im Mittel mit 25 % Abweichungen noch zufrieden. Auftragnehmer hingegen sind noch mit 28 % Abweichungen zufrieden. Berater sind noch bei Überschreitungen des Plans von bis zu 16 % zufrieden, akzeptieren jedoch eine weitaus größere Unterschreitung der Bauausführung von 35 % im Vergleich zu Auftraggebern und Auftragnehmern.

Auftraggeber haben demnach höhere Erwartungen als ihre Auftragnehmer. Berater werten Unterschreitungen vergleichsweise deutlich weniger negativ. Bei Überschreitungen haben sie eine ähnliche Einstellung wie die Auftraggeber. Es



**Tabelle 6.1:** Vergleich von Auftraggeber (AG), Auftragnehmer (AN) und Beratungen bei der Zufriedenheit von Abweichungen zwischen der Rahmenterminplanung und der Bauausführung

	Anzahl	Mittelwert Überschreitung (1)	Mittelwert Unterschreitung (2)	Mittelwert aus (1) und (2)
<b>AG</b>	24	16 %	25 %	20 %
<b>AN</b>	26	27 %	28 %	28 %
<b>Beratungen</b>	16	16 %	35 %	25 %

ist demnach immer eine Frage der Perspektive, was noch zufriedenstellende Abweichungen bedeuten. Dennoch wird auf Basis dieser Umfrage im Hinblick auf die Phase der Bauausführung von einer Spanne noch zufriedenstellender Abweichungen zwischen 16 und 35 % mit einem durchschnittlichen Wert von 26 % in der Rahmenterminplanung ausgegangen.

Weiter werden bivariate Analysen durchgeführt. Damit sollen lineare Zusammenhänge zwischen der Zufriedenheit bei Abweichungen und demografischen Kennzahlen überprüft werden. Es werden die demografischen Kennzahlen, der Unternehmensgröße und der Berufserfahrung untersucht. Bei einer Untersuchung der Rolle und Spezialisierung des Unternehmens sollte durch die hohe Anzahl der Gruppen in diesen die Stichprobengröße durch eine weitere Umfrage vergrößert werden. Die Unternehmensgröße zeigt in den Analysen keine signifikante Korrelation nach Pearson auf.

Wie in Tabelle 6.2 zu sehen ist, können bei der Untersuchung der Berufserfahrung und Akzeptanz von Abweichungen insbesondere bei den Über- und Unterschreitungen geringe negative Korrelation gefunden werden. Bei der Untersuchung des Mittelwerts wurde beim Rahmenterminplan, Generalablaufplan und Grobterminplan eine moderate negative Korrelation gefunden. Diese ist statistisch signifikant, da  $p$  als Signifikanzwert kleiner als 5 % ist ( $p < 0,05$ ). Dies bedeutet, dass je länger die Berufserfahrung ist, desto geringere Abweichungen werden akzeptiert bzw. Berufseinsteiger akzeptieren folglich tendenziell größere Planabweichungen. Dieser Zusammenhang sowie die zufriedenstellenden Abweichungen

**Tabelle 6.2:** Pearson Korrelationskoeffizient und p-Wert für den Zusammenhang zwischen der Berufserfahrung und der Zufriedenheit bei Abweichungen (in kursiv moderate Korrelation)

	<b>Rahmen-</b> <b>terminplan</b>	<b>General-</b> <b>ablaufplan</b>	<b>Grob-</b> <b>terminplan</b>	<b>Ausführ-</b> <b>ungsplanung</b>
<b>Korrelation</b>				
Verzögerung	-0,21	-0,26	-0,27	-0,11
Beschleunigung	-0,28	-0,28	-0,27	-0,24
Mittelwert	<i>-0,34</i>	<i>-0,35</i>	<i>-0,34</i>	<i>-0,25</i>
<b>p-Wert</b>				
Verzögerung	8,62E-08	6,86E-08	4,40E-08	1,60E-08
Beschleunigung	1,71E-06	1,31E-06	8,74E-07	5,76E-07
Mittelwert	3,40E-07	2,730E-07	1,80E-07	8,73E-08

von Terminplänen sollte in einem nächsten Schritt durch eine größere Stichprobe nochmals überprüft werden.

## 6.2.4 Zusammenfassung der Anforderungen an die Prognosegenauigkeit

Die Ergebnisse der Umfrage zeigen, dass je mehr sich das Bauprojekt der Realisierung nähert, desto weniger Abweichungen vom Plan akzeptiert werden bzw. noch zufriedenstellend sind. Abweichungen werden in Unter- und Überschreitungen der geplanten Dauern für eine differenzierte Betrachtung unterteilt. Auch wird der Mittelwert aus Über- und Unterschreitungen berechnet. Zum Zeitpunkt der Rahmenterminplanung sind die Teilnehmer der Umfrage durchschnittlich bei einer Differenz von Plan zu Ausführung mit 20 % Überschreitungen und 31 % Unterschreitungen noch zufrieden. Auftraggeber akzeptieren in der Rahmenterminplanung weniger Überschreitungen als Auftragnehmer (16 % im Vergleich zu 27 %). Berater sind noch zufrieden, wenn die Bauausführung 35 % schneller als in der Rahmenterminplanung geplant, durchgeführt wird. Bei Auftraggebern sind dies 25 % und bei Auftragnehmern 28 %. Aus den Ergebnissen folgen Zielwerte,

die durch entsprechende Strategien und Methoden mindestens eingehalten werden sollten, um ein für die am Bauprojekt beteiligten Personen zufriedenstellendes Ergebnis zu erreichen. Es ist jedoch anzumerken, dass die Ist-Abweichungen häufig höher als die hier definierten Zielabweichungen sind (vgl. Ballard 2000). So zeigen beispielhaft Studien beim Bau von Straßen 38 % Zeitüberschreitungen auf (vgl. Flyvbjerg et al. 2016). Die deutlich höhere Abweichung in der Praxis zeigt auch die Studie von Changali et al. (2015). Laut dieser Studie werden 77 % aller Megaprojekte mit mindestens 40 % Überschreitung zur geplanten Dauer übergeben.

Es ist anzumerken, dass die Umfrage explorativ ist und damit für eine konfirmative Analyse die Auswertungen nochmals mit einer größeren Stichprobe durchgeführt werden sollten. Diese konfirmative Analyse könnte im internationalen Raum durchgeführt werden, um eine größere Zielgruppe abdecken zu können. Auch sind in dieser Umfrage mittlere und große Unternehmen überrepräsentiert. Die Ergebnisse der Umfrage sind als Momentaufnahme zu verstehen. Ziel von Beteiligten des Bauprojektes sollte es sein, Abweichungen gegenüber dem Plan zu reduzieren und damit auch die Erwartungshaltung zu Abweichungen zu verändern. Mit immer besser werdenden Strategien und Methoden zur Prognose von Dauern in Bauprojekten kann das Akzeptanzniveau gegenüber Abweichungen reduziert werden.

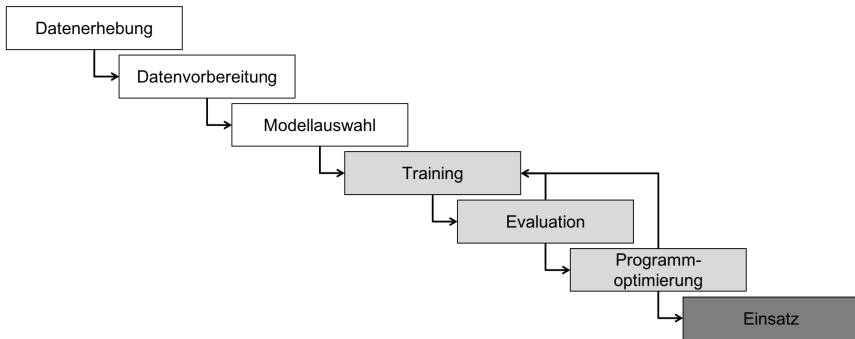
Im Weiteren wird das explorativ ermittelte Akzeptanzniveau mit der Prognosegenauigkeit der ausgewählten Planungsmethoden verglichen.

## **6.3 Bewertung der Prognosegenauigkeit**

### **6.3.1 Methodisches Vorgehen zur Modellanwendung**

Dieser Abschnitt beschreibt sieben Schritte zur Anwendung von ML-Methoden und ist an das Konzept von Guo (2017) angelehnt. Die einzelnen Schritte sind in Abbildung 6.5 dargestellt. Die ersten Schritte in den weißen Boxen

(Datenerhebung, Datenvorbereitung und Modellauswahl) sind dabei als Vorbereitung auf den darauffolgenden Lernprozess (grau) zu verstehen. Nachdem das Modell fertiggestellt ist, kann ein praktischer Einsatz erfolgen (dunkelgrau). Der letzte Schritt, der Einsatz, wird in dieser Analyse ausgelassen. Anzumerken ist, dass das Vorgehen oft kein linearer Prozess ist. Häufig müssen Entscheidungen überdacht und somit Schritte wiederholt werden.



**Abbildung 6.5:** Schritte bei der Anwendung eines Konzeptes des Maschinellen Lernens (Guo 2017)

**Schritt 1: Datenerhebung.** Der erste Schritt beeinflusst maßgebend das Ergebnis des Lernprozesses, je nach Qualität und Quantität der erhobenen Daten (vgl. Guo 2017). Formell wird hier eine Stichprobe aus einer Grundgesamtheit entnommen, die deshalb möglichst repräsentativ sein sollte. Hierbei ist insbesondere darauf zu achten, relevante Merkmale aufzunehmen (vgl. von Luxburg und Rätsch 2004, S. 2).

**Schritt 2: Datenvorbereitung.** Nachdem die gewünschten Daten erhoben sind, müssen sie je nach verwendetem Algorithmus auf verschiedene Weise vorbereitet und organisiert werden. Die Vorbereitung der Daten nimmt häufig die meiste Zeit im Prozess des Maschinellen Lernens ein (vgl. Webb 2010a, S. 259 f.). Um das Modell nach dem Lernen testen zu können, werden die vorhandenen Daten geteilt in Trainingsdaten und Testdaten (vgl. Ertel 2016, S. 195).

Sollte in den Daten ein Rauschen vorhanden sein, kann es hilfreich sein, Ausreißer und andere verdächtige Datenpunkte in den Daten zu finden und diese zu löschen (vgl. Webb 2010a, S. 259 f.).

Ein weiterer gängiger Vorgang ist bspw. die Normierung von Merkmalen, wenn diese in verschiedenen Größenordnungen (heterogene Daten) vorliegen (vgl. Chollet 2018, S. 154).

Fehlen Werte einzelner Merkmale bei Datensätzen, muss ein geeignetes Vorgehen gewählt werden, um mit diesen Datensätzen umzugehen. Oft werden für die fehlenden Werte Daten generiert („Imputation“). Dies kann bspw. durch den Mittelwert des jeweiligen Merkmals geschehen, durch den ansonsten am häufigsten verwendeten Wert dieses Merkmals oder durch weitergehende Verfahren (vgl. VanderPlas 2016, S. 381).

**Schritt 3: Modellauswahl.** Um mit den so vorbereiteten Daten lernen zu können, muss nun eine Wahl des Modells erfolgen. Diese wird teilweise durch die Rahmenbedingungen des Problems vorgegeben, zum Beispiel durch die Art des Lernens (siehe Kapitel 5.3.2), die Art der Ein- und Ausgabedaten (siehe Kapitel 5.3.3) oder die Anzahl der Merkmale.

**Schritt 4: Training.** Nach den Schritten der Vorbereitung beginnt nun der eigentliche sukzessive Lernprozess des Modells. Anhand der zur Verfügung gestellten Daten wird eine Approximation  $\hat{f}$  gesucht (siehe Kapitel 5.3.2), indem das Modell schrittweise mithilfe des Feedback-Signals an die Daten angepasst wird (vgl. Chollet 2018, S. 73).

Zum Training kann eine Kreuzvalidierung verwendet werden. Hier werden die Daten mehrfach in unterschiedliche Trainings- und Testdatensätze geteilt, um die Prognosegenauigkeit weiter zu verallgemeinern.

Je nach Modell werden unterschiedliche Gewichtungen und andere Parameter im Training angepasst. Ein einzelner, kompletter Durchlauf der Trainingsdaten wird „(Lern-/Trainings-) Epoche“ genannt. Dabei werden zunächst alle Trainingsdaten durchlaufen und erst im Anschluss die aggregierten Änderungen der Parameter ausgeführt (vgl. Kruse et al. 2015, S. 23 f.). Diesen Modus nennt man

„Batch-Lernen“ oder „Batch-Training“ (engl. „batch“: Stapel), wobei sämtliche Trainingsbeispiele bei Trainingsbeginn vorhanden sein müssen (vgl. Kruse et al. 2015, S. 24). Dem steht das „Online-Lernen“ oder „Online-Training“ (engl. „online“ : mitlaufend, schritthaltend) gegenüber. Hierbei wird ein Trainingsschritt ausgeführt, sobald ein Trainingsbeispiel verfügbar wird (vgl. Kruse et al. 2015, S. 23). Diese Lernmethode ist insbesondere interessant, wenn sich die Daten, die sukzessive verfügbar werden, über die Zeit verändern (vgl. Russell und Norvig 2012, S. 871 f.).

**Schritt 5: Evaluation.** Nach dem Training des Modells (oft auch schon währenddessen), muss der Lernerfolg anhand der Generalisierungsfähigkeit überprüft werden (siehe auch Überanpassung in Kapitel 5.3.4).

**Schritt 6: Programoptimierung.** Durch Training und Evaluation werden die Gewichtungen des Modells angepasst. Jedes Modell hat jedoch auch verschiedene Hyperparameter, die bspw. die Struktur des Modells oder das Verhalten im Lernprozess steuern. Diese zeichnen sich dadurch aus, dass sie einmal vor Beginn des Trainings gesetzt werden und nicht durch das Modell gelernt werden, d. h. bei Veränderung muss der Trainingsprozess neu gestartet werden (vgl. VanderPlas 2016, S. 349). Durch die Anpassung der Parameter kann gegebenenfalls der Lernerfolg substantiell verbessert werden. Die manuelle Konfiguration durch einen Experten oder über eine maschinelle Optimierung sind notwendig.

**Schritt 7: Einsatz.** Sind alle Anpassungen des Modells erfolgt, ist es bereit für den Einsatz. Für zukünftig verfügbare unbeschriftete Merkmalsvektoren  $X_{n+1}, \dots$  kann der Algorithmus jetzt Vorhersagen  $\hat{f}(X_{n+1}), \dots$  treffen. Dies wird voraussagendes Lernen (engl. „predictive learning“) genannt (vgl. Knox 2018, S. 13).

### 6.3.2 Beschreibung der verwendeten Datensätze

Zur Analyse werden zwei Datensätze mit Informationen zum Neubau von Wohngebäuden verwendet. Ein Datensatz stammt aus Teheran, Iran, und der andere

aus San Francisco, USA. Beide Datensätze werden um weitere Daten angereichert, die die wirtschaftliche sowie politische Lage beschreiben. Auf diese Datensätze werden unterschiedliche Entscheidungsbäume angewendet und mit den Ergebnissen von linearen Hochrechnungsmodellen und KNN verglichen. Hierdurch ergeben sich nicht nur Möglichkeiten für einen Vergleich von verschiedenen Prognosemodellen, sondern es kann auch ein Vergleich des Einflusses von Eigenschaften des Datensatzes auf die Prognosegenauigkeit vollzogen werden. Denn die beiden Datensätze sind unterschiedlich groß und beinhalten jeweils die Dauer der Bauausführung in einer unterschiedlichen Granularität sowie unterschiedliche Merkmale.

### **Datensatz 1: Teheran, Iran**

Der erste zur Verfügung stehende Datensatz beschreibt Daten zu Bauprojekten von 372 Wohngebäuden in Teheran, Iran. Die Bauzeit wird in Quartalen dokumentiert. Die Dauer zur Planung ist nicht dokumentiert. Als beschreibende Variablen sind Kosten, Projektvariablen und ökonomische Randdaten im Datensatz enthalten. Die ökonomischen Variablen enthalten Werte für jeweils fünf Quartale vor Baubeginn (vgl. Rafiei und Adeli 2016). Der Datensatz kann online im „Machine Learning Repository“ der University of California, Irvine abgerufen werden<sup>5)</sup>. Die Projekte wurden in einem Zeitraum von 16 Jahren aufgenommen (im persischen Kalender Jahr 72 bis 88). Eine vollständige Auflistung aller Merkmale ist im Anhang C.1 enthalten. Die durchschnittliche Dauer beträgt 6,3 Quartale mit einer Standardabweichung von 2,1 Quartalen.

### **Datensatz 2: San Francisco, USA**

Der zweite Datensatz beinhaltet 184 abgeschlossene Neubauprojekte zu Wohnungen in San Francisco, USA<sup>6)</sup>. Es werden dabei die Projekte nach der Baugenehmigung „Neubau“ und „Appartments“ als vorgeschlagene Verwendung selektiert. Damit ist Datensatz zwei inhaltlich mit Datensatz eins vergleichbar. Der erste Tag der Baurealisierung findet zwischen 1993 und 2018 statt. Damit sind

<sup>5)</sup> <https://archive.ics.uci.edu/ml/datasets/Residential+Building+Data+Set> (besucht am 06.05.2020)

<sup>6)</sup> <https://data.sfgov.org/Housing-and-Buildings/Building-Permits/198e-djp9> (besucht am 06.05.2020)

die Projekte über einen Zeitraum von 25 Jahren dokumentiert worden. Es wird die nach der Baugenehmigung zur Verfügung stehende Bauzeit in Tagen betrachtet. Auch hier ist die benötigte Zeit zur Planung nicht dokumentiert. Die Projekte sind durch die Geschosse, Wohneinheiten, genehmigten Kosten sowie durch die eingereichten Planunterlagen beschrieben. Weiter sind der Straßename und die Nachbarschaftsgegend dokumentiert (siehe Tabelle C.3 im Anhang). Die durchschnittliche Dauer beträgt in Datensatz zwei 1.179,4 Tage (entspricht ca. 13,1 Quartale bei durchschnittlich 90 Tagen je Quartal) und weist im Vergleich zu Datensatz eins eine deutlich höhere Standardabweichung mit 808,69 Tagen (etwa 9,0 Quartale) auf.

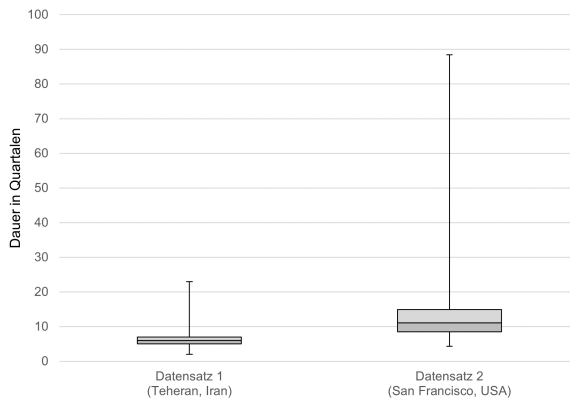
Die wesentlichen Eigenschaften beider Datensätze sind in Tabelle 6.3 beschrieben. Mit beiden Datensätzen kann der Einfluss des Detaillierungsgrads in der Dokumentation der Projekte, durch die Anzahl der dokumentierten Projekte, der Anzahl der Merkmale sowie der Standardabweichung der Dauer als Zielvariable analysiert werden.

Abbildung 6.6 zeigt ein Boxplot der beiden verwendeten Datensätze. Die Abbildung zeigt die vergleichsweise längeren Dauern und hohen Ausreißer in Datensatz zwei (San Francisco, USA).



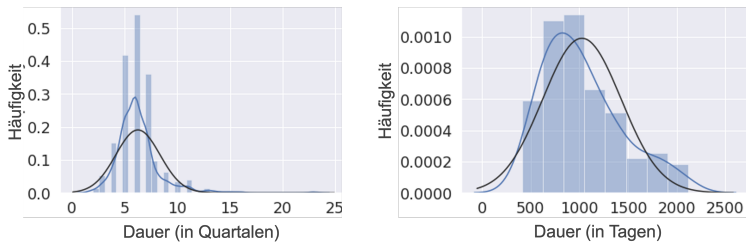
**Tabelle 6.3:** Beschreibung der verwendeten Datensätze (Zur besseren Vergleichbarkeit sind für Datensatz 2 die Dauern in Quartale umgerechnet. Es wird angenommen, dass 90 Tage einem Quartal entsprechen)

	<b>Datensatz 1 (Teheran, Iran)</b>	<b>Datensatz 2 (San Francisco, USA)</b>
Anzahl dokumentierter Projekte	372	184
Anzahl dokumentierter Features	32	10
Durchschnittliche Dauer	6,3 Quartale	1.179 Tage 13,1 Quartale
Standardabweichung der Dauer	2,1 Quartale	809 Tage 9,0 Quartale
Minimale Dauer	2,0 Quartale	392 Tage 4,4 Quartale
25% Quantil der Dauer	5,0 Quartale	763 Tage 8,5 Quartale
50% Quantil der Dauer	6,0 Quartale	999 Tage 11,1 Quartale
75% Quantil der Dauer	7,0 Quartale	1.341 Tage 14,9 Quartale
Maximale Dauer	23,0 Quartale	7.958 Tage 88,4 Quartale



**Abbildung 6.6:** Boxplot zum Vergleich des Datensatzes 1 (Teheran, Iran) und Datensatz 2 (San Francisco, USA) in Quartalen

Abbildung 6.7 zeigt jeweils die Verteilung der Bauzeit in den Projekten für die beiden vorgestellten Datensätze als Histogramm. Die Kurven repräsentieren die Normalverteilung vor (blau) und nach (schwarz) einer log-Transformation. Es wird im Rahmen der Datenvorbereitung eine log-Transformation verwendet, um die Datensätze so nah wie möglich an eine Normalverteilung anzunähern und die Schrägheit der Datenverteilung zu reduzieren. Zu erkennen ist, dass beide Datensätze annähernd normalverteilt sind.



**Abbildung 6.7:** Verteilung der Datensätze: Linke Abbildung Teheran, Iran in Quartalen; rechte Abbildung San Francisco, USA in Tagen (Normalverteilung vor (blau) und nach (schwarz) der log-Transformation)

### Externe Daten zur weiteren Anreicherung

Die zwei Datensätze werden weiter jeweils mit 65 externen Datenpunkten angereichert (siehe Tabelle C.4 und C.5 im Anhang), die globale Informationen über die Wirtschaft und Politik des jeweiligen Standorts ergänzen<sup>7)</sup>. Unter anderem enthalten diese Kennzahlen zur Inflation, dem Korruptions- und Innovationsindex sowie die Anzahl der Baugenehmigungen in dem jeweiligen Land. Diese Informationen werden beispielhaft für das erste, zweite und letzte Jahr des jeweiligen Bauprojektes integriert. Es wurde angenommen, dass die wirtschaftliche Lage in der nahen Zukunft in die Planung miteinbezogen wird. Es ergeben sich 195 weitere Merkmale pro Projekt.

<sup>7)</sup> <https://data.oecd.org/>; <https://www.theglobaleconomy.com/> (besucht am 06.05.2020)

### 6.3.3 Bewertung der ausgewählten Datensätze

Die Bewertung der ausgewählten Datensätze wird hinsichtlich der Güte und der Repräsentativität vorgenommen. Die Bewertung der Güte erfolgt anhand der Objektivität der Merkmale und ihrer Relevanz. Die Relevanz bezieht sich auf die häufig genannten Merkmale nach der Literaturrecherche in Kapitel 3.5. Die Repräsentativität wird auf Basis der Umfrageergebnisse aus Kapitel 4 geschlussfolgert.

#### Güte der ausgewählten Datensätze

Beide Datensätze enthalten numerische Merkmalsattribute, die zu großen Teilen auf einer Ratio-Skala zu verordnen sind (z. B. Gesamtgeschossfläche, Zinssatz für ein Darlehen, Bodenpreisindex). Damit sind ihre Ausprägungen zum Nullpunkt durch ihre Reihenfolge und Abstände eindeutig bewertbar. In Datensatz eins (Teheran, Iran) ist der Projektort anhand einer Intervallskala und in Datensatz zwei ist der Straßename sowie die Nachbarschaft als Text geschrieben. Durch eine Verortung des Bauprojektes in einer eindeutigen Postleitzahl, Straße und Nachbarschaftsgegend sind jedoch auch diese Merkmalsattribute eindeutig und objektiv. Datensatz zwei (San Francisco, USA) enthält eine eindeutige Kategorisierung von drei verschiedenen Bauweisen als Text. Die Datensätze enthalten keine Texte, die individuell durch die Person, die die Dokumentation der Projekte vorgenommen hat, beschrieben sind und subjektive Elemente beinhalten könnten.

Es werden zur Bewertung, ob die Datensätze zur Prognose der Dauer relevante Merkmale enthalten, die drei häufigsten Nennungen der Kategorien Prozesse, Personen, Produkt, Position bzw. Örtlichkeit und Preis vorgenommen (siehe Kapitel 3.5). Folgende Merkmale mit häufigen Nennungen nach der Literaturrecherche sind in den Datensätzen enthalten:

- Bei beiden Datensätzen ist der Produkttyp identisch. Es handelt sich um den Neubau von Wohnungen (Kategorie Produkt).

- Das Design des Gebäudes wird in Datensatz eins mit der Gesamtgeschossfläche und in Datensatz zwei mit der Anzahl der Stockwerke und der vorgeschlagenen Anzahl an Wohneinheiten beschrieben (Kategorie Produkt).
- Datensatz zwei enthält weiter eine Beschreibung des eingesetzten Materials für den Rohbau durch die Kategorisierung in drei Bauweisen (Kategorie Produkt).
- Die Baustelle wird in Datensatz eins mit der Grundstücksfläche beschrieben (Kategorie Position).
- Das Gebiet, in dem sich die Baustelle findet, wird in Datensatz eins mit der Postleitzahl und in Datensatz zwei mit Straße und Nachbarschaftsgegend beschrieben (Kategorie Position).
- Die Kosten sind in Datensatz eins detailliert mit der vorläufigen Schätzung der Baukosten und Gesamtbaukosten sowie den tatsächlichen Baukosten aufgeschlüsselt. In Datensatz zwei sind diese mit einer Position definiert (Kategorie Preis).
- Datensatz eins enthält weiter eine Vielzahl finanzieller Indikatoren, wie die kumulative Liquidität oder Zinssätze des Darlehens (Kategorie Preis).
- Datensatz eins beinhaltet auch Informationen zu Einnahmen mit den Preisen der Einheiten und den tatsächlichen Verkaufspreisen (Kategorie Preis).

Damit enthalten beide Datensätze relevante Kennzahlen zur Prognose der Dauer. Die dokumentierten Merkmale sind jedoch auf Basis der Liste aus der Literaturrecherche keinesfalls vollständig. So hätte u. a. der Bauherr mit seiner Expertise (z. B. Anzahl Projekt pro Jahr, Anzahl Mitarbeiter), seine Prozesse (z. B. Abwicklungsmodelle), das Wetter und die fehlenden Merkmale entsprechend der obigen Liste im direkten Vergleich beider Datensätze ebenfalls dokumentiert werden können. Wie die Ergebnisse der Umfrage in Kapitel 4 jedoch gezeigt haben, dokumentieren Organisationen der Baubranche die Projekte nicht vollständig, sondern im Durchschnitt sind 14 Datenpunkte dokumentiert. Beide Datensätze

werden im Folgenden anhand der in der Praxis verfügbaren Datensätze hinsichtlich ihrer Repräsentativität bewertet.

### **Repräsentativität der ausgewählten Datensätze**

Die Datensätze beinhalten 372 und 184 Projekte. Nach der Umfrage in Kapitel 4 verfügen 16 % der Unternehmen über mehr als 100 Projekte. Damit sind diese Datensätze zwar vorhanden, jedoch besitzt die Mehrheit der an der Umfrage teilgenommenen Organisationen keine Datensätze in diesem Umfang.

Die an der Umfrage teilgenommenen Organisationen dokumentieren laut der univariaten Ergebnisse der Umfrage 14 Datenpunkte bzw. Merkmale. Der Datensatz eins (Teheran, Iran) enthält 32 Merkmale und Datensatz zwei (San Francisco, USA) zehn Merkmale. Dabei sind in Datensatz eins jedoch lediglich zwölf Merkmale projektspezifisch, die 20 weiteren Merkmale beziehen sich auf wirtschaftliche Kennzahlen (z. B. Wechselkurse, Preisindex für Land, Anzahl Baugenehmigungen pro Jahr; siehe Beschreibung der Merkmale in Kapitel C.1 im Anhang). Damit liegen die Anzahl der dokumentierten Merkmale in beiden Datensätzen unter den in der Umfrage aufgenommenen 14 Merkmalen.

Vergleicht man die Inhalte der häufig dokumentierten Datenpunkte bzw. Merkmale mit denen der verwendeten Datensätze, so können folgende Merkmale wiedergefunden werden:

- Die Projektart (97 % dokumentieren dieses Merkmal) wird in beiden Datensätzen durch die Klassifizierung des Datensatzes in Wohnungen (Neubau) abgedeckt.
- Das Bauvolumen (77 % dokumentieren dieses Merkmal) als Beschreibung des Baukörpers wird in Datensatz zwei durch die Anzahl der Stockwerke und Appartements angegeben. Datensatz eins enthält hingegen die Gesamtgeschossfläche.

- Der Standort (67 % dokumentieren dieses Merkmal) ist in Datensatz eins durch die Postleitzahl, in Datensatz zwei durch den Straßennamen sowie die Nachbarschaftsgegend beschrieben.
- Die Bebauungsfläche (66 % dokumentieren dieses Merkmal) ist in Datensatz eins mit der Grundfläche des Gebäudes enthalten.
- Die Bausumme (60 % dokumentieren dieses Merkmal) wird in beiden Datensätzen genannt. Hierbei ist die Dokumentation in Datensatz eins deutlich detaillierter (z. B. mit den Kosten pro  $m^2$  zum Zeitpunkt der Planung, tatsächliche Kosten, Verkaufspreise der Wohnungen).
- Auch die Bauzeit (51 % dokumentieren dieses Merkmal) ist in beiden Datensätzen enthalten.

Nach der Umfrage werden weiter Nutzeranforderungen (62 %) und die Projektorganisation (53 %) häufig dokumentiert. Diese Merkmale sind in den ausgewählten Datensätzen nicht enthalten. Bei diesen beiden Merkmalen ist jedoch zu beachten, dass eine in Kategorien standardisierte und damit objektive Dokumentation vorliegen muss, um die Inhalte miteinander vergleichen zu können. Hingegen ist in Datensatz zwei die Anzahl der erstellten Plansätze enthalten, die laut Umfrage nicht zu häufig dokumentierten Merkmalen gehört und ein Anhaltspunkt für die Planungsgenauigkeit sein kann.

Demnach sind die in den Datensätzen enthaltenen Merkmale als repräsentativ anzusehen, die Größe des Datensatzes liegt jedoch nach der Umfrage lediglich bei weniger als 16 % der Unternehmen vor. Beide Datensätze sind daher als extreme Fälle zu betrachten. Aufgrund der Analyse lediglich zweier Datensätze lassen sich weiter keine statistischen Rückschlüsse ziehen. Die Datensätze repräsentieren Fallstudien, mit denen ein spezifischer Fall detailliert analysiert werden soll. Nach Denzin und Lincoln (2011) können mit einer Fallstudie Rückschlüsse auf eine größere Grundgesamtheit geschlossen werden. Es werden nach den Auswahlstrategien von Denzin und Lincoln (2011, S. 307) extreme Datensätze ausgewählt, die eine Vielzahl an dokumentierten Projekten im Vergleich zu einem durchschnittlichen Unternehmen besitzen. Damit kann überprüft werden, ob bei

entsprechend großen Datenmengen analytische Planungsmethoden sinnvoll sind. Bei einer Negierung ergeben sich ebenfalls Rückschlüsse für kleine Datensätze. Mit dieser Fallstudie können daher Rückschlüsse auf die Gesamtheit geschlossen werden, auch wenn diese nicht einem durchschnittlichen Datensatz gleichen.

### 6.3.4 Ergebnisse der Datenanalyse

In bestehenden Forschungen werden insbesondere lineare Regressionen sowie KNN als prädiktive Methode verwendet (siehe Kapitel 3).

Mit den zwei definierten Zielkriterien der Prognosegenauigkeit und Erklärbarkeit werden die Ergebnisse von linearen Regressionen und KNN mit denen von Entscheidungsbäumen verglichen. Hierfür werden verschiedene Entscheidungsbäume in der Modellauswahl zum Vergleich ausgewählt: der Random Forest (vgl. Liaw et al. 2002), Gradient Boosting Regression (GBR), XGBoost Tree (vgl. Chen und Guestrin 2016), LightGBM (vgl. Ke et al. 2017) und der CatBoost (vgl. Prokhorenkova et al. 2018).

Durch die unterschiedliche Granularität der Dauer in beiden Datensätzen (Dokumentation in Quartalen und Tagen) eignet sich der MAPE als einheitlicher Vergleichswert. Mit diesem Ergebnis kann ein Vergleich zwischen den Modellen, linearen Regressionsmodellen, einem KNN sowie Entscheidungsbäumen, vorgenommen werden. Dieser Vergleich gibt Aufschluss über die Anwendbarkeit von Entscheidungsbäumen zur Prognose der Dauern von Projektphasen. Ferner kann die Visualisierung des Entscheidungsbaums verwendet werden, um für den Anwender die Referenzklassen als Endknoten des Entscheidungsbaums darzustellen. Ein Performance-Indikator zur Abweichung innerhalb der Blatt- bzw. Endknoten bewertet die Prognosequalität. Die verwendeten Datensätze werden in einen Trainings- und Testdatensatz mit einem Verhältnis von 80/20 aufgeteilt (siehe Empfehlungen zur Aufteilung in Kapitel 6.3.1). Es wird ein Batch-Training angewendet und zur Kreuzvalidierung  $k$  wird  $k = 3$  gewählt. Zur technischen Umsetzung wird für jeden Entscheidungsbaum eine Tiefe von sechs als Hyperparameter gewählt, um vergleichbare Ergebnisse zu erzielen. Hierfür wird

Python zur Programmierung mit entsprechenden Hyperparametern verwendet. Der verwendete Code ist in C.4 im Anhang zu finden.

Die Programmierung wird mit Google Colab und folgenden Bibliotheken durchgeführt: TensorFlow (1.14.0), keras-applications (1.0.6), CatBoost (0.24.3), shap (0.37.0).

Um die Prognosegenauigkeit der ausgewählten analytischen Planungsmethoden zu bewerten und mit bestehenden Methoden zu vergleichen, wird der vorgestellte Code auf beide Datensätze jeweils mit und ohne die externen Daten angewendet. So kann zusätzlich der Mehrwert der externen Daten bewertet werden. Tabelle 6.4 zeigt für die Entscheidungsbäume die mittlere prozentuale Abweichung (MAPE) als Ergebnis auf.

**Tabelle 6.4:** Vergleich der Prognosegüte der Entscheidungsbäume anhand des Performance-Indikators MAPE (\* ohne externe Daten; \*\* mit externen Daten; in fett bestes Ergebnis pro Datensatz)

	Teheran, Iran*	Teheran, Iran**	San Francisco, USA*	San Francisco, USA**
Random Forest	19,58 %	17,79 %	<b>36,79 %</b>	26,78 %
GBR	17,34 %	15,47 %	38,71 %	27,50 %
XGBoost	18,11 %	15,34 %	53,57 %	29,13 %
LightGBM	18,85 %	16,83 %	48,42 %	34,85 %
CatBoost	<b>16,76 %</b>	<b>13,48 %</b>	39,67 %	<b>26,45 %</b>

Im Datensatz von Iran ist der geringste MAPE 13,48 % (entspricht MAE 0,84 Quartale) und in San Francisco 26,45 % (entspricht MAE 423 Tage). Mit und ohne die Integration der externen Daten zeigt der CatBoost die besten Prognosewerte auf. Die Hinzunahme der externen Daten steigert weiter die Genauigkeit der Prognose.

Beim Vergleich der beiden Datensätze zeigt sich, dass der Datensatz von San Francisco vergleichsweise schlechtere Ergebnisse aufweist. Dies kann erstens an



der höheren Standardabweichung, zweitens an der geringeren Größe des Datensatzes und drittens an fehlenden Merkmalen liegen. In Datensatz eins (Teheran, Iran) sind deutlich mehr Merkmale je Projekt in den Kategorien Preis, Produkt und Örtlichkeit des Projektes dokumentiert.

Das Ergebnis des CatBoosts wird nun für den Datensatz San Francisco mit linearen Regressionsmodellen und dem KNN verglichen. Die Formeln zur linearen Regression sind nicht auf den Datensatz aus Teheran anwendbar, da dort die Dokumentation in Quartalen angegeben ist und die entsprechenden linearen Regressionsmodelle der Autoren Ireland (1983), Chan (1999, 2001), Choudhury und Rajan (2003), Ogunsemi und Jagboro (2006), Le-Hoai und Lee (2009) eine Dokumentation in Tagen benötigen. Weiter wird zum Vergleich ein KNN angewendet, das den Stand der Forschung zur Prognose der Bauausführung darstellt. Die Ergebnisse des Vergleichs sind in Tabelle 6.5 mithilfe des MAPE als Performance-Indikator dargestellt.

**Tabelle 6.5:** Vergleich der Prognosegüte zwischen KNN, dem CatBoost und linearen Regressionsmodellen repräsentiert durch die jeweiligen Autoren mit dem Performance-Indikator MAPE (\* mit externen Daten; in fett beste Ergebnisse des Datensatzes zwei)

	<b>Teheran, Iran (in Quartalen)</b>	<b>San Francisco, USA (in Tagen)</b>
Ireland (1983)	-	32,12 %
Chan (1999)	-	44,68 %
Chan (2001)	-	30,80 %
Ogunsemi (2006)	-	32,06 %
Choudhoury (2003)	-	69,92 %
LeHoai (2009)	-	36,92 %
Own regression	21,50 %	27,70 %
KNN*	59,86 %	36,16 %
CatBoost*	<b>13,48 %</b>	<b>26,45 %</b>

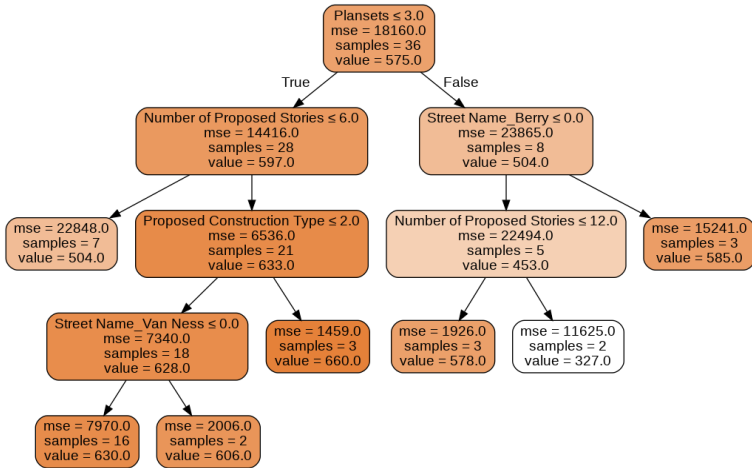
Im Vergleich zu bestehenden linearen Regressionsmodellen (vgl. Ireland 1983, Chan 1999, 2001, Choudhury und Rajan 2003, Ogunsemi und Jagboro 2006, Le-Hoai und Lee 2009) und auch im Vergleich zum KNN zeigt der CatBoost die deutlich besseren Ergebnisse auf. In Datensatz zwei beträgt der MAPE des KNN 36,16 % (entspricht 573 Tage MAE) und der des CatBoost 26,46 % (423 Tage MAE). In Datensatz eins ist die Prognosegenauigkeit beim CatBoost deutlich besser mit 13,48 %. Weiterhin zeigt Datensatz eins einen größeren Unterschied in der Vorhersagegenauigkeit von KNN und CatBoost auf. Der Unterschied beträgt 46,38 Prozentpunkte.

Die Funktionsweise des CatBoost begründet im Vergleich zu den anderen angewendeten Entscheidungsbäumen die vergleichsweise gute Prognosegenauigkeit (siehe Kapitel 5.6.2). Die geringere Prognosegenauigkeit von KNN ist möglicherweise auf die relativ geringe Trainingsdatenmenge zurückzuführen. Die Zunahme der externen Daten steigert die Prognosegenauigkeit allgemein weiter. Hier werden innerhalb der Entscheidungsbäume nur die besten externen Faktoren einbezogen, in KNN kann dies hingegen durch den Einbezug aller Merkmale zu einem übertrainierten Modell führen. Da eine Evaluation und Parameteroptimierung des KNN daher lediglich zu schlechteren Ergebnissen führen würde, werden diese beiden Schritte ausgelassen.

Somit kann durch die Anwendung von Entscheidungsbäumen nicht nur die Erklärbarkeit im Bereich maschineller Lernmodelle verbessert werden, sondern auch die Prognosegenauigkeit im Vergleich zu linearen Regressionsmodellen und KNN erhöht werden.

Abbildung 6.8 stellt die Repräsentation eines beispielhaften Random Forest mit einer maximalen Tiefe von fünf auf Basis historischer Projektdaten visuell dar. Der Anwender kann den Ästen durch eine Einordnung des aktuellen Projektes entlang der Entscheidungen in den Blättern folgen. Das letzte Blatt stellt die Referenzklasse dar. Ein Wert gibt die prognostizierte Dauer der Ausführungszeit wieder. Die quadrierte mittlere Abweichung dient in diesem beispielhaften Entscheidungsbaum als Maß der Prognosegenauigkeit. Damit werden zwei Inhalte visualisiert: Zum einen werden vergleichbare Projektbeispiele und deren dabei

beeinflussende Projektmerkmale aufgezeigt. Zum anderen wird ein Performance-Indikator als Maß der Prognosegenauigkeit aufgezeigt.



**Abbildung 6.8:** Entscheidungsbaum eines Random Forest (auf Basis von Datensatz zwei - San Francisco, USA)

Zusammenfassend lässt sich nach der Datenanalyse schlussfolgern, dass Entscheidungsbäume und insbesondere der CatBoost bessere Ergebnisse liefern als lineare Regressionen und derzeit angewendete KNN. Als Grund hierfür ist die vergleichsweise geringe Datenmenge zu nennen, mit der Entscheidungsbäume sehr gut umgehen können. KNN können zu einem späteren Zeitpunkt bei Verfügbarkeit von mehr Daten eventuell bessere Ergebnisse aufzeigen. Weiter wird festgestellt, dass externe Daten und eine Vorgruppierung der Daten die Prognosegüte verbessern können.

### 6.3.5 Evaluierung relevanter Merkmale in den Modellen

Entscheidungsbäume stellen mit ihrer einfachen Visualisierung menschliche Gedankengänge nach. Um besser analysieren zu können, welche Merkmale einen

großen Einfluss auf die Prognose besitzen, wird das SHAP-Framework verwendet. Ziel ist es, die Relevanz von Merkmalen im Sinne der Dokumentation und Unterschiede in beiden Datensätzen bewerten zu können. SHAP unterstützt die Interpretierbarkeit der KI-Modelle (vgl. Lundberg et al. 2018). Der SHAP-Wert zeigt, vereinfacht dargestellt, den spezifischen Beitrag eines Merkmals zur Prognose auf (siehe Kapitel 5.6.3).

Wendet man einen Random Forest Regressor auf den Datensatz eins (Teheran, Iran; mit externen Daten) an, können die als besonders relevant identifizierten Faktoren visualisiert werden.

Abbildung 6.9 und 6.10 zeigen die SHAP-Werte für den Datensatz eins (Teheran, Iran) und Datensatz zwei (San Francisco, USA). Die Bewertung des globalen Einflusses wird absteigend angeordnet („Feature Importance“). Der Wert des Merkmals („Feature Value“) gibt den Originalwert aus der Datenbasis an. Dieser wird mit der entsprechenden Farbskala in der Grafik angezeigt. Der SHAP-Wert

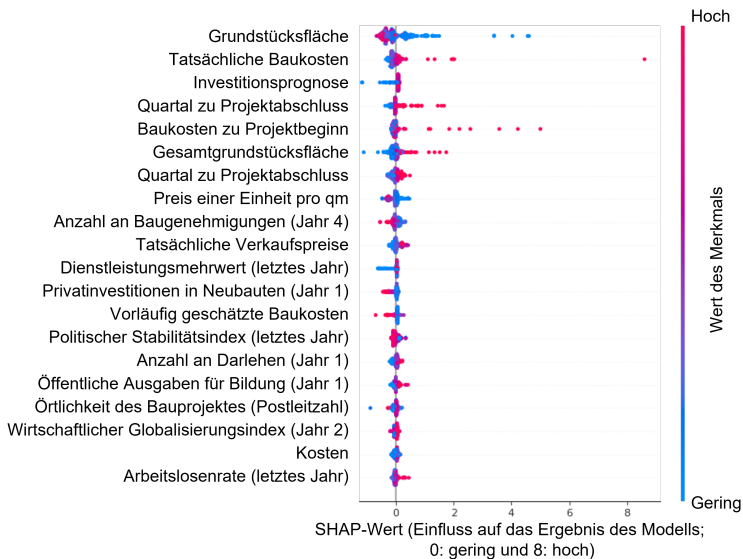
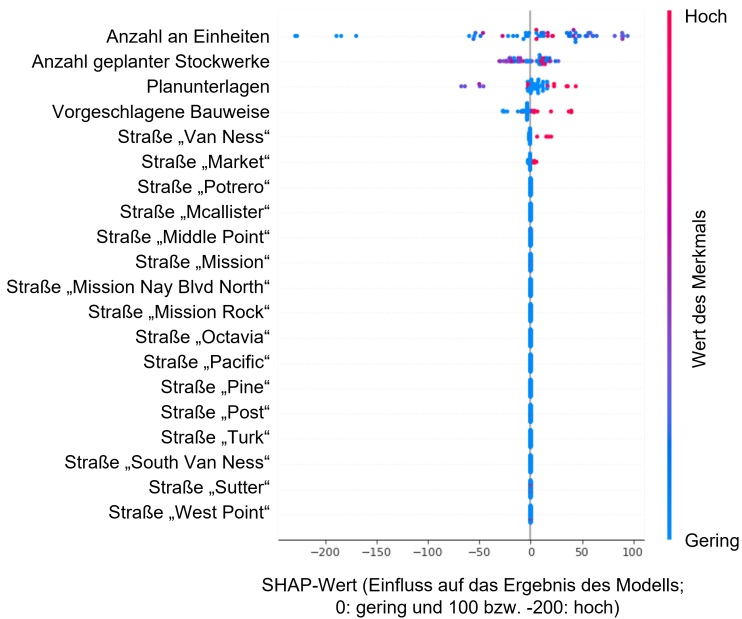


Abbildung 6.9: SHAP-Framwork zu Datensatz eins (Teheran, Iran)



**Abbildung 6.10:** SHAP-Framework zu Datensatz zwei (San Francisco, USA)

gibt den Beitrag des einzelnen Merkmals zur Prognose wieder. Dabei wird jedes Projekt als ein Punkt dargestellt. Die horizontale Achse zeigt demnach den positiven oder negativen Einfluss auf die Prognose.

Aus beiden Abbildungen lassen sich folgende Aussagen im Sinne der Dokumentation treffen:

1. Informationen zum Produkt (z. B. die Fläche, Anzahl an Geschossen) haben den größten Einfluss auf die Prognose der Dauer.
2. Ein weiteres wichtiges Merkmal sind die Kosten.
3. Weiterhin ist die Lage (z. B. beschrieben durch die Straße) relevant.

Die Relevanz der Kostenvariablen bestätigt das Prognosemodell von Bromilow (vgl. Bromilow 1969) und auch die Relevanz der Preiskategorie im 5P-Framework (siehe Kapitel 3.5) kann hervorgehoben werden.

## 6.4 Zusammenfassung und kritische Betrachtung

In diesem Kapitel wird zuerst die erwartete Abweichung bei der Planung von Dauern, die gerade noch zufriedenstellend ist, ermittelt. An der Umfrage in der DACH-Region nahmen 87 Teilnehmer aus Organisationen der Baubranche teil. Geplante Dauern in frühen Projektphasen, wie der Rahmenterminplanung, dürfen den Planwert in der Realisierung um 31 % unterschreiten und um 20 % überschreiten. Dies bedeutet ein gemitteltes Akzeptanzniveau von 26 % noch zufriedenstellender Abweichungen zum aufgestellten Rahmenterminplan. Weitere Differenzierungen nach Rollen wurden festgestellt.

In einem zweiten Teil werden Entscheidungsbäume als Teil erklärbarer KI-Methoden mit linearen Regressionsmodellen und KNN verglichen, um die Dauer der Bauausführungsphase vorherzusagen. Die beiden hierfür ausgewählten Datensätze stammen aus unterschiedlichen Ländern (Teheran, Iran, und San Francisco, USA), besitzen beide Informationen über Wohngebäude, sind unterschiedlich groß und weisen unterschiedliche Merkmale auf. Diese Variablen können als Moderatorvariablen gesehen werden, die die Effektstärke von analytischen Planungsmethoden auf die Prognosegenauigkeit bestimmen. Beide Datensätze enthalten Merkmale wie zur Örtlichkeit des Projektes, den Preisen und Produktinformationen. Die Datensätze werden mit ihrer Beschreibung als extreme Fallstudien bewertet, um zu überprüfen, ob eine als groß bewertete Datenbasis ausreichend zufriedenstellende Ergebnisse liefert. Die Größe beider Datensätze ist nach der durchgeführten Umfrage in Kapitel 4 lediglich bei weniger als 16 % der Unternehmen vorhanden.

Die Entscheidungsbäume zeigen neben ihrer Interpretierbarkeit in beiden Datensätzen auch Vorteile in der Prognosegenauigkeit auf: Es konnte gezeigt werden, dass Entscheidungsbäume bessere Ergebnisse liefern als linearen Regressionsmodelle und KNN. In den vorliegenden Datensätzen konnte die Vorhersage um bis zu 46,38 Prozentpunkte verbessert werden (von 59,86 % auf 13,48 % von einem KNN zu einem CatBoost für den Datensatz eins in Teheran, Iran). Für beide Datensätze zeigt der CatBoost besonders vielversprechende Ergebnisse. Die Prognosegenauigkeit liegt bei einer durchschnittlichen prozentualen Abweichung (MAPE) von 13,48 % (Teheran, Iran) bzw. 26,45 % (San Francisco, USA) für die beiden Datensätze.

Dabei spielen in der Prognose externe Daten, die die wirtschaftliche und politische Lage im jeweiligen Land beschreiben, eine Rolle. Auch Informationen zum Produkt bzw. dem Bauwerk, zu den Kosten und der Lage sind in ihrer Dokumentation relevant.

Mit den Ergebnissen kann die Abweichung des Rahmenterminplans (Mittelwert aus Über- und Unterschreitung), die mit durchschnittlich 26 % noch als zufriedenstellend bewertet wird, bei Datensatz eins mit 13,48 % deutlich eingehalten und bei Datensatz zwei mit 26,45 % fast eingehalten werden. Nach einer Studie von Changali et al. (2015) werden 77 % aller Megaprojekte mit mindestens 40 % Überschreitung zur geplanten Dauer übergeben. Dieser Wert übersteigt das hier ermittelte Akzeptanzniveau von 26 % und zeigt damit das weitere Potenzial der ausgewählten Methoden für die Praxis.

Betrachtet man die Ergebnisse von Datensatz zwei differenzierter und nimmt diese als Unter- oder Überschreitung an, so wird die mittlere prozentuale Abweichung als Unterschreitung mit 31 % noch akzeptiert, als Überschreitung mit 20 % jedoch nicht akzeptiert. Weiter wird die Erwartungshaltung der Auftraggeber in Unter- und Überschreitungen nicht erreicht. Bei Beratern liegen die maximal 16 % Verzögerungen des Rahmenterminplans über den mit dem CatBoost erreichten 26,45 %.

Demnach kann die Anforderung an die Prognosegenauigkeit größtenteils eingehalten werden. Datensatz eins zeigt durch die Vielzahl der dokumentierten Merkmale, die Größe des Datensatzes und die geringere Standardabweichung der Dauer seine Stärken. So entspricht die Granularität der Vorhersage eher Monaten. Eine tagesgenaue Vorhersage ist aber auch mit maschinellen Lernmodellen bisher nicht möglich. ML hat daher kaum das Potenzial in der Breite bei allen Bauunternehmen, als alleinige Methode eingesetzt zu werden. Die Modelle könnten mit dem Wissen von Experten und durch weitere Datenanalysen kombiniert werden. Zum Beispiel könnte eine Vorgruppierung der Daten durch Experten (z. B. in kurze, mittlere und lange Bauvorhaben) oder eine Anwendung in Kombination mit der Fuzzy-Methode die Prognosegenauigkeit weiter erhöhen.

Die Umfrage, die Datenanalyse sowie der Vergleich der ermittelten Anforderungen zu den Ergebnissen der Datenanalysen müssen kritisch betrachtet werden. In der Umfrage ist das zufriedenstellende Akzeptanzniveau für Abweichungen von der Bauausführung für die DACH-Region zu einem bestimmten Zeitpunkt ermittelt worden. Für eine höhere Repräsentativität der Ergebnisse sollte eine weitere Umfrage zu weiteren Zeitpunkten sowie im internationalen Raum durchgeführt werden. Wie auch in Kapitel 6.2.4 beschrieben, sind Kleinstunternehmen unterrepräsentiert. Bei der Beantwortung der Umfrage kann weiter das Verständnis des Begriffs „Zufriedenheit“ unterschiedlich aufgefasst werden und individuell unterschiedliche Leistungsziele enthalten. Das objektive und statistische Messbarmachen der inneren Ausgeglichenheit bei erreichten Zielen stellt eine große Herausforderung dar. Das Ergebnis der Umfrage zeigt eine differenzierte Betrachtung zufriedenstellender Über- und Unterschreitungen der geplanten Dauer sowie einen gemittelten Wert. Diese Erwartungen werden dem Performance-Indikator MAPE als Ergebnis der Datenanalyse gegenübergestellt. Der gemittelte Wert wurde aus Gründen der Vereinfachung gleich verteilt aus Über- und Unterschreitungen erstellt. Eine Gewichtung der Über- und Unterschreitungen ist jedoch ebenfalls möglich und verändert damit die Bewertung der Ergebnisse der Datenanalyse.

In der Datenanalyse ist, wie in Kapitel 6.3.3 angemerkt, die Repräsentativität der Daten zu hinterfragen. Die analysierten Datensätze sind aufgrund ihrer Größe



aktuell bei größeren Unternehmen (gemessen am Umsatz und an ihren Mitarbeitern) zu finden. Dennoch kann damit das zufriedenstellende Akzeptanzniveau von Abweichungen bei Dauern verglichen werden, das ebenfalls repräsentativ für größere Unternehmen ist. Nimmt die Anzahl an dokumentierten Daten in den kommenden Jahren weiter zu, so entstehen vergleichbare Datenbanken auch bei kleineren Unternehmen. Hier sei auf die Studie von Thomas und Bowman (2021) verwiesen, die auf Basis einer Umfrage mit 3.916 Beteiligten der Baubranche im internationalen Raum aufzeigen, dass sich die Anzahl der dokumentierten Daten in den vergangenen drei Jahren verdoppelte. Zuletzt ist kritisch anzumerken, dass die Datenanalyse zur Prognose der Dauer der Bauausführung durchgeführt wurde und andere Phasen eines Bauprojektes, wie einzelne Planungsphasen, nicht im Fokus der Betrachtung sind. Eine Übertragbarkeit der Ergebnisse auf diese Phasen ist in weiteren Studien zu ermitteln. Die Bauausführungsphase hat jedoch durch das Zusammenkommen vieler Organisationen und Sichtbarkeit für die Öffentlichkeit eine besondere Relevanz hinsichtlich bei der Einhaltung der Dauer.

In der Umfrage aus Kapitel 4 konnte ein direkter Zusammenhang zwischen der Relevanz der verfügbaren Datensätze und den Zeitabweichungen aufgezeigt werden, jedoch finden analytische Planungsmethoden mit einer hohen Stufe in der Praxis bisher eine geringe Anwendung. Daher stellt sich die Frage, ob dennoch bessere Ergebnisse durch die Intuition von Experten auf Basis eines Datenzugriffs erzielt werden als durch die in diesem Kapitel analysierten prädiktiven Planungsmethoden. Aus diesem Grund wird im nächsten Kapitel überprüft, inwieweit die analytische Planungsmethode mit der höchsten Prognosegenauigkeit, der CatBoost, im Vergleich zur Expertenintuition eingesetzt werden sollten.



# 7 Bewertung der Expertenintuition im Vergleich zu ausgewählten analytischen Planungsmethoden

## 7.1 Vorbemerkung

In den vorherigen Kapiteln konnte gezeigt werden, dass eine große Datengrundlage zeitliche Abweichungen reduzieren kann und die Prognosegenauigkeit maschineller Lernmodelle als ausreichend gut bewertet wird. Die Expertenintuition könnte jedoch aufgrund der hohen Unsicherheit durch Erfahrungswerte zu einer besseren Vorhersage als maschinelle Lernmodelle führen. In keiner der vorliegenden Veröffentlichungen wird der Vorteil eines maschinellen Lernmodells im Vergleich zur Expertenintuition für die Planung in frühen Phasen von Bauprojekten untersucht. Daher soll in diesem Kapitel die dritte Forschungsfrage beantwortet werden: Inwiefern lässt sich im Vergleich zur Planung durch Experten mit den ausgewählten Modellen analytischer Planungsmethoden die Prognosegenauigkeit von Dauern eines Bauprojektes in einer frühen Planungsphase unterstützen und verbessern? Zur Beantwortung wird eine Fallstudie mit einem Datensatz durchgeführt, der eine detaillierte Analyse der Fragestellung zulässt.

Hierfür wird ein Datensatz aus New York, USA, verwendet, der Informationen zur Dauer von Planungsbeginn bis Bauausführungsende enthält. Dieser Datensatz wird wie in Kapitel 6 mit externen marktbeschreibenden Faktoren angereichert. Der Datensatz enthält die zugehörige prognostizierte und die tatsächliche Dauer der Bauprojekte. Die prognostizierte Dauer spiegelt die Planung durch den

Experten wider und mit der tatsächlichen Dauer wird das maschinelle Lernmodell trainiert. Als Ergebnis kann eine vergleichende Analyse der Expertenintuition und der ausgewählten analytischen Planungsmethoden vorgenommen werden. Weiter können Projektmerkmale und ihre Ausprägungen identifiziert werden, die für eine Fallunterscheidung zwischen Experte und maschinellen Lernmodellen relevant sind. Diese Merkmalsausprägungen sind durch den Datensatz gegeben und damit durch diesen beschränkt.

Da in diesem Datensatz Informationen zu Projektänderungen enthalten sind, können die Ergebnisse in zwei Szenarien unterstützend wirken: Zum einen kann eine Prognose zu Beginn des Bauprojekts simuliert werden, für die noch keine Änderungen bekannt sind (Szenario eins). Zum anderen kann die Dauer während des Projektes auf Basis von Änderungen validiert werden (Szenario zwei). Hier können bisherige Änderungen zu Kosten und der Dauer in die Planung einbezogen werden. Die Betrachtung beider Szenarien wird durchgeführt, da frühere Studien (z. B. Akintoye und MacLeod 1997, Enshassi et al. 2009) ebenfalls den Einfluss von Änderungen auf Kosten und die Projektdauer bestätigt hatten.

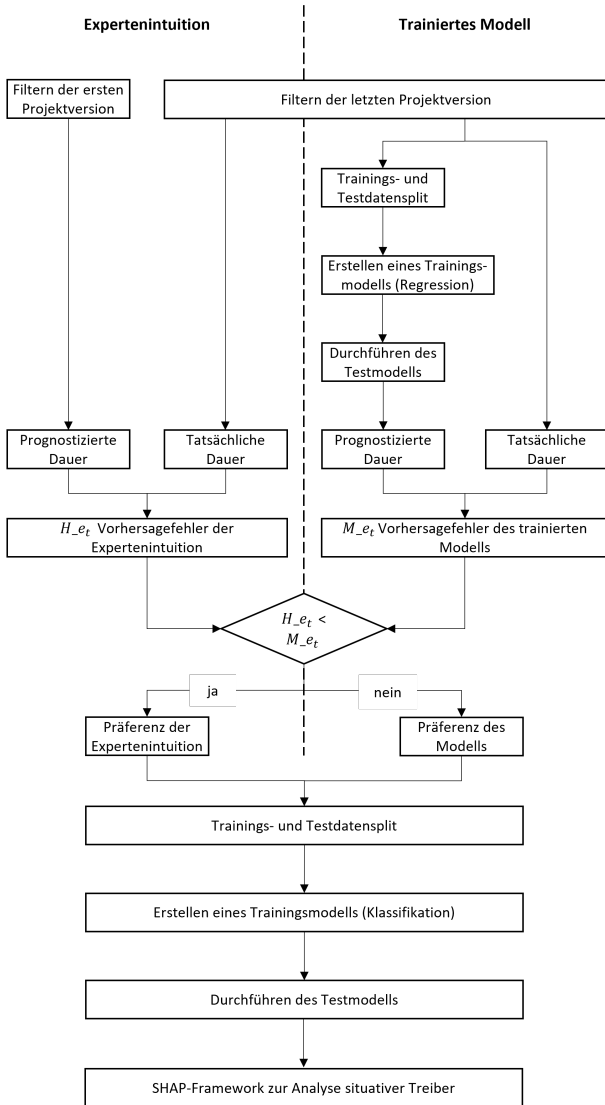
Das Kapitel ist wie folgt gegliedert: In Kapitel 7.2 wird das methodische Vorgehen vorgestellt. Der verwendete Datensatz ist in Kapitel 7.3 beschrieben und in Kapitel 7.4 bewertet. Die Ergebnisse der maschinellen Lernmodelle zur Vorhersage der Dauer sowie der Vergleich der Vorhersagefehler der Expertenintuition mit dem trainierten Modell sind in Kapitel 7.5 erläutert. Auf Basis des verwendeten Datensatzes können Einflussfaktoren für eine maschinen- bzw. expertengestützte Prognose aufgezeigt werden. Kapitel 7.6 enthält eine Zusammenfassung der Ergebnisse sowie eine kritische Betrachtung.

## 7.2 Methodisches Vorgehen zur Modellanwendung

Abbildung 7.1 zeigt schematisch das methodische Vorgehen zur Analyse von Prognosen der Gesamtprojektdauer durch Experten und das maschinelle Lernmodell: Die Expertenschätzung wird als die initiale Projektplanung angenommen. Als Experten werden demnach die Personen angesehen, die die Projekte des Datensatzes initial geplant haben. Für das maschinelle Lernmodell werden die Projektdaten in Trainings- und Testdaten aufgeteilt. Hierfür wird das Modell mit den realisierten Dauern trainiert. Mit dem maschinellen Lernmodell wird die Dauer auf der Grundlage des Testdatensatzes vorhergesagt. Daraus ergibt sich der maschinelle Vorhersagefehler. Durch den Vergleich der durch den Experten prognostizierten Dauer mit der realisierten Dauer kann der Vorhersagefehler des Experten ermittelt werden. Dazu wird der Performance-Indikator MAE zur Bewertung der Prognosegenauigkeit herangezogen. Außerdem können die absoluten Ergebnisse miteinander verglichen werden: In wie vielen Fällen lag die Vorhersage der Maschine und in wie vielen Fällen die des Menschen näher am beobachteten Wert? Wie groß ist die Differenz zwischen beiden?

Anhand dieser Ergebnisse können auf Basis des gegebenen Datensatzes relevante Einflussfaktoren und Treiber für den Einsatz der Expertenintuition und maschineller Lernmodelle analysiert werden. Zum Vergleich und der Bewertung der Expertenintuition und ausgewählter analytischer Planungsmethoden sowie ihrer Einflussfaktoren werden die Schritte aus Kapitel 6.3.1 durchlaufen:

**Schritt 1: Datenerhebung.** Es wird ein frei zugänglicher Datensatz mit öffentlichen Bauprojekten gewählt. Dieser umfasst die initiale Planung sowie die realisierte Dauer des Bauprojektes. Dabei wird die Dauer als Dauer des Gesamtprojektes, von Planung bis Übergabe des Bauwerks verstanden.



**Abbildung 7.1:** Methodisches Vorgehen zum Vergleich der Vorhersagefehler der Expertenintuition mit dem trainierten Modell sowie Analyse ihrer situativen Treiber

**Schritt 2: Datenvorbereitung.** Der Datensatz wird in insgesamt vier Datensätze untergliedert:

1. Datensatz eins: Der erste Datensatz repräsentiert das Projekt während seiner Realisierung. Aus diesem Grund werden Informationen zu Änderungen in die Prognose einbezogen. Informationen zu externen Marktfaktoren sind jedoch nicht enthalten.
2. Datensatz zwei: Im Vergleich zu Datensatz eins werden die externen Marktfaktoren hinzugefügt.
3. Datensatz drei: Aus Datensatz eins werden die Spalten „Letzte Budgetänderungen“, „Gesamtbudgetänderungen“, „letzte Terminplanänderung“, „Gesamte Terminplanänderung“ entfernt, sodass der Datensatz lediglich die Projektinformationen zu Projektbeginn enthält. Diese Merkmale stellen die Projektänderungen dar.
4. Datensatz vier: Zu Datensatz drei werden die externen Marktfaktoren hinzugefügt.

Damit beschreiben Datensatz drei und vier das Szenario eins, in dem eine Prognose zu Projektbeginn durchgeführt werden soll. Datensatz eins und zwei hingegen können verwendet werden, wenn während des Projektes bereits Änderungen vorliegen (Szenario zwei). Um das Modell nach dem Lernen testen zu können, werden die vorhandenen Daten in Trainings- und Testdaten aufgeteilt (Ertel 2016, S. 195). Es wird eine Aufteilung in einen Trainings- und Testdatensatz im Verhältnis 80/20 gewählt. Dies liegt im üblichen Bereich von 60/40 bis 80/20 (Raschka und Mirjalili 2017, S. 137).

**Schritt 3: Modellauswahl.** Die Programmierung wird mit Google Colab und den folgenden Bibliotheken durchgeführt: TensorFlow (1.14.0), keras-applications (1.0.6), CatBoost (0.24.3), shap (0.37.0).

Zur Modellauswahl wird der CatBoost (vgl. Prokhorenkova et al. 2018) mit dem Ziel der Regression verwendet. Dieser hat sich durch die Ergebnisse aus dem vorherigen Kapitel bewährt.

**Schritt 4: Training.** Es wird ein Batch-Training durchgeführt und eine Kreuzvalidierung mit  $k = 3$  ausgewählt.

**Schritt 5: Evaluation.** Die Ergebnisse des maschinellen Lernmodells werden mit dem MAE als Performance-Indikator bewertet. Der Vergleich der Modelle dient der Überprüfung des Einsatzes von Entscheidungsbäumen zur Prognose von Dauern in einem Bauprojekt.

**Schritt 6: Optimierung der Parameter.** Zur Optimierung des maschinellen Lernmodells werden die Hyperparameter analysiert und diese entsprechend im Modell verbessert. Hyperparameter von Entscheidungsbäumen sind die optimale Tiefe des Entscheidungsbaums und die Anzahl der maximal trainierten Bäume („n\_estimators“). Dies wird über eine „Grid Search“ durchgeführt, die eine der am häufigsten verwendeten Methoden zur Optimierung der Hyperparameter ist (vgl. Alibrahim und Ludwig 2021, S. 1552). Mit der Grid Search werden alle möglichen Kombinationen der Hyperparameter in das Trainingsmodell eingesetzt und die Leistung für jede Kombination wird mithilfe einer Kreuzvalidierung bewertet. Dabei haben alle Parameter die gleiche Wahrscheinlichkeit zur Beeinflussung des Prozesses. Als Ergebnis werden die besten Hyperparameter wiedergegeben. Nachteilig an der Grid Search ist, dass bei einer hohen Anzahl an Merkmalen und Ausprägungen die Grid Search eine große Rechenleistung sowie viel Zeit benötigen kann (vgl. Alibrahim und Ludwig 2021, S. 1552). Vorteil einer einfachen Grid Search ist hingegen die Interpretierbarkeit ihrer Ergebnisse und die Fähigkeit der Interaktion zwischen verschiedenen Hyperparametern (vgl. Hutter et al. 2015, S. 331).

Das resultierende maschinelle Lernmodell ist die Basis, um den Vorhersagefehler der menschlichen Expertenschätzung mit dem des maschinellen Lernmodells vergleichen zu können. Hierfür werden erneut die Schritte zwei bis fünf durchlaufen.

**Schritt 2: Datenvorbereitung.** Um die Vorhersage des maschinellen Lernmodells mit der Expertenschätzung zur Dauer vergleichen zu können, wird den Datensätzen eins bis vier jeweils die prognostizierte Dauer hinzugefügt und über



den identifizierenden Schlüssel, der „Projekt ID“, die Dauer der initialen Planung hinzugefügt. Der Vorhersagefehler der Expertenintuition und des maschinellen Lernmodells werden nun berechnet. Daraus kann gefolgert werden, ob der Experte oder das maschinelle Lernmodell pro Projekt die genauere Prognose erstellt hätte. Es wird eine Spalte im Datensatz hinzugefügt, die zwischen den beiden Fällen unterscheidet. Dabei beschreibt „0“ eine vorteilhafte Vorhersage durch den Experten und „1“ eine vorteilhafte Vorhersage durch das maschinelle Lernmodell.

**Schritt 3: Modellauswahl und Schritt 4: Training.** Mit dem zuvor ausgewählten maschinellen Lernmodell wird ein neues maschinelles Lernmodell zur Klassifikation auf diese Spalte hin trainiert. Dieses soll vorhersagen, ob die Vorhersage vorrangig durch den Experten oder das maschinelle Lernmodell durchgeführt werden sollte.

**Schritt 5: Evaluation.** Um beeinflussende Merkmale auf die Entscheidung zwischen einer Prognose durch den Experten und einem maschinellen Lernmodell zu identifizieren, wird ein SHAP-Framework verwendet. Dies stellt die relevanten Einflussfaktoren für eine Differenzierung der beiden Fälle dar.

Der letzte Schritt ist der Einsatz des maschinellen Lernmodells. Der gesamte Python Code ist im Anhang in Kapitel D.2 enthalten.

## 7.3 Beschreibung des ausgewählten Datensatzes

Die analysierten Daten sind verfügbar auf „NYC Open Data“<sup>8)</sup> (vgl. Office of Operations 2018) und umfassen öffentliche Bauprojekte in New York Stadt, USA. Insgesamt enthält der Datensatz 2.400 Einträge mit 15 Spalten, die die Projektmerkmale beschreiben. Die Spalte der Projektbeschreibung (Text) wird

---

<sup>8)</sup> Datensatz: [www.opendata.cityofnewyork.us](http://www.opendata.cityofnewyork.us) (besucht am 06.05.2020)

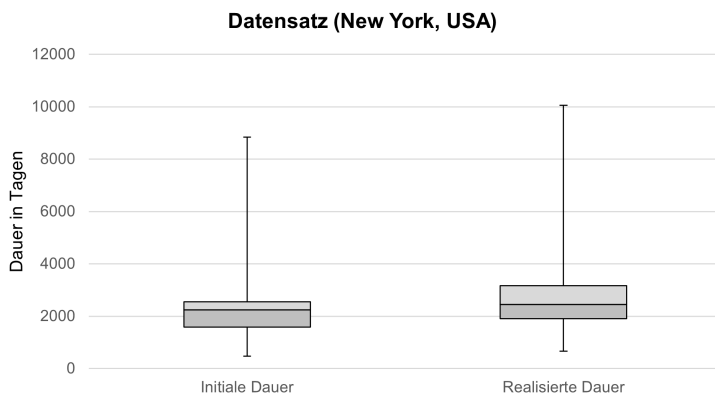
aufgrund der fehlenden Vergleichbarkeit in der folgenden Analyse nicht verwendet. Die Projekte sind zu verschiedenen Berichtszeitpunkten dokumentiert. Die erste und letzte dokumentierte Version eines jeden Projekts werden aus dem Datensatz extrahiert. Die erste Version stellt die Planung durch den Experten dar und auf Basis der letzten Version des Projektes wird das maschinelle Lernmodell zur Vorhersage trainiert. Nach der Extraktion sind in beiden Datensätzen 367 Projekte zu unterschiedlichen Projektkategorien (z. B. Straßen, Brücken, Schulen) enthalten. Die verwendeten Einzelmerkmale sind im Anhang in Tabelle D.1 aufgeführt. Die Berichtszeitpunkte der Projekte beginnen zur Planung und enden mit der Fertigstellung der Bauausführung. So kann die Dauer des Projektes als Differenz zwischen der prognostizierten Fertigstellung und dem Planungsstart berechnet werden. Die dokumentierte Dauer umfasst daher die Dauer der Planung und der Realisierung des Bauprojekts. Der Start der Planung ist zwischen dem Jahr 1993 und dem Jahr 2000 dokumentiert. Der Dokumentationszeitraum umfasst daher sieben Jahre. Der angedachte Projektabschluss ist zwischen dem Jahr 2014 und dem Jahr 2032 im Datensatz festgehalten.

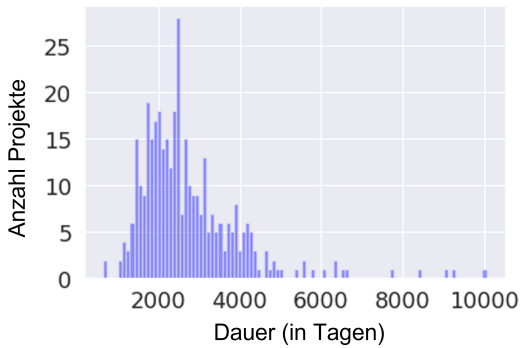
Eine Beschreibung des Datensatzes, der initial geplanten und der realisierten Projektdauer, sind in Tabelle 7.1 dargestellt. Erkennbar sind hier die Planungsfehl-schlüsse der Expertenintuition. In allen Kategorien zur Beschreibung der Dauer kann eine Unterschätzung der initialen Planung identifiziert werden. So weicht die durchschnittliche Projektdauer um 431 Tage von der initialen zur realisierten Dauer ab und die Standardabweichung um 97 Tage. Der menschliche Vorhersa-gefehler wird weiter in Tabelle 7.3 dargestellt.

Zusätzlich zeigt Abbildung 7.2 ein Boxplot zur initialen und realisierten Dauer sowie Abbildung 7.3 die Verteilung der Dauer (in Tagen) über die Bauprojekte im Datensatz.

**Tabelle 7.1:** Beschreibung des verwendeten Datensatzes New York, USA

	<b>Datensatz (New York, USA)</b>	
	Initiale Dauer	Realisierte Dauer
Anzahl dokumentierter Projekte		367
Anzahl dokumentierter Merkmale		15
Durchschnittliche Dauer	2.291 Tage	2.723 Tage
Standardabweichung der Dauer	1.153 Tage	1.250 Tage
Minimale Dauer	476 Tage	656 Tage
25 % Quantil der Dauer	1.584 Tage	1.902 Tage
50 % Quantil der Dauer	2.241 Tage	2.447 Tage
75 % Quantil der Dauer	2.558 Tage	3.172 Tage
Maximale Dauer	8.830 Tage	10.049 Tage

**Abbildung 7.2:** Boxplot zur Dauer des Datensatzes aus New York, USA



**Abbildung 7.3:** Histogramm über die Projektdauer (in Tagen) und Anzahl der Bauprojekte

Die Anzahl der Projekte in den 20 Kategorien ist in der folgenden Tabelle 7.2 dargestellt. Projekte zu Straßen und Verkehrswegen sind am häufigsten enthalten.

Die Projektdauer kann neben den rein projektbeschreibenden Merkmalen auch durch externe Marktfaktoren beeinflusst werden. Aus diesem Grund wird der Datensatz wie im vorherigen Kapitel 6 zusätzlich mit den 65 externen Datenpunkten pro Projekt angereichert, die globale Informationen über die Wirtschaft und Politik des jeweiligen Standorts ergänzen<sup>9)</sup>. Zu den Daten gehören u.a. Kennzahlen zur Inflation, zum Korruptions- und Innovationsindex und zur Anzahl der Baugenehmigungen im jeweiligen Land. Diese externen Daten sind im Anhang C.3 in Tabelle C.4 zusammengefasst. Sie werden für das Jahr, in dem die Planung begonnen hat, dem jeweiligen Bauprojekt zugeordnet, um die wirtschaftliche und politische Ausgangslage der Planung miteinzubeziehen.

<sup>9)</sup> <https://data.oecd.org/>; <https://www.theglobaleconomy.com/> (besucht am 06.05.2020)

**Tabelle 7.2:** Übersicht über die Anzahl der Projekte je Kategorie

<b>Kategorie</b>	<b>Anzahl Projekte</b>
Straßen und Verkehrswege	87
Abwasserkanäle	40
Industrielle Entwicklung	35
Wasserversorgung	33
Abwasseraufbereitung	30
Schulen	29
Parks	19
Brücken	19
Öffentliche Sicherheit und Strafjustiz	16
Brücken, Straßen und Verkehrswege	15
Sonstige öffentliche Einrichtungen	12
Abwasserentsorgung	9
Fähren	5
Gesundheit und Krankenhäuser	5
Bibliotheken	4
Kunst und Kultur	4
Parks, Straßen und Verkehrswege	2
Soziale Dienste	1
Industrielle Entwicklung, Parks	1
Industrielle Entwicklung, Straßen und Wege	1

## 7.4 Bewertung des ausgewählten Datensatzes

Wie im vorherigen Kapitel 6 wird der Datensatz anhand der Güte mit einer objektiven Beschreibung relevanter Merkmale anhand der Ergebnisse der Literaturrecherche aus Kapitel 3 und der Repräsentativität dieser anhand der Umfrage aus Kapitel 4 bewertet.

## Güte des ausgewählten Datensatzes

Der Datensatz enthält die Merkmale, Budgetprognose sowie letzte und gesamte Budgetänderungen, die auf einer Ratio-Skala verordnet sind. In diesem Datensatz sind eine Vielzahl an Merkmalen enthalten, die durch einen Text beschrieben sind (u. a. Kategorie, Gemeinde, Verwaltungsbehörde, Kundenagentur). Durch eine Kategorisierung sind jedoch auch diese Texte eindeutig und objektiv. Die Merkmale Projektname und die Projektbeschreibung werden aus Gründen der Objektivität nicht in die Prognose miteinbezogen.

Wie im vorherigen Kapitel werden die Merkmale des Datensatzes mit den Nennungen relevanter Merkmale in bestehender Fachliteratur zugeordnet. Folgende Merkmale mit hohen Nennungen sind im Datensatz enthalten:

- Die Änderungen mit den letzten und gesamten Budgetänderungen sowie Terminplanänderungen beschreiben die Stabilität des Prozesses (Kategorie Prozess).
- Der Bauherr wird mit der Verwaltungsbehörde und Kundenagentur genannt. Eine fachliche Beschreibung des Bauherrn (z. B. Anzahl Projekte pro Jahr) ist nicht enthalten (Kategorie Person).
- Der Produkttyp wird mit der Kategorie (z. B. Parks, Straßen, Schulen, öffentliche Sicherheit) definiert (Kategorie Produkt).
- Die Gemeinde, in der das Bauprojekt stattfindet, beschreibt die Örtlichkeit (Kategorie Position).
- Die Kosten mit der Budgetprognose dienen als finanzieller Indikator (Kategorie Preis).

Damit deckt dieser Datensatz alle fünf Kategorien ab. Jedoch sind auch hier einzelne häufige Nennungen nicht enthalten. So fehlen u. a. Informationen zum Entwurf der Gebäude mit der Bruttogeschossfläche oder Stockwerken, Materialien, das Wetter während der Durchführung, die verfügbare Baustellenfläche mit Möglichkeiten der Baustelleneinrichtung sowie weitere finanzielle Indikatoren.

## Repräsentativität der ausgewählten Datensätze

Der Datensatz enthält 367 Projekte mit 15 Merkmalen. Wie im vorherigen Kapitel entspricht der Datensatz damit einem extremen Fall, da nach der Umfrage in Kapitel 4 lediglich 16 % der Unternehmen Datensätze mit mehr als 100 Projekten besitzen. Die 15 Projektmerkmale des Datensatzes entsprechen hingegen annähernd den 14 durchschnittlich dokumentierten Merkmalen. Vergleicht man die Inhalte der häufig dokumentierten Datenpunkte bzw. Merkmale mit denen der verwendeten Datensätze, so können folgende Merkmale wiedergefunden werden:

- Die Projektart (97 % dokumentieren dieses Merkmal) wird durch die Kategorie abgedeckt.
- Der Standort (67 % dokumentieren dieses Merkmal) wird durch die Nennung der Gemeinde, in der das Projekt gebaut wird, beschrieben.
- Die Bausumme (60 % dokumentieren dieses Merkmal) ist als Budgetprognose dokumentiert.
- Auch die Bauzeit (51 % dokumentieren dieses Merkmal) ist im Datensatz auf Tagesebene enthalten.

Nach der Umfrage aus Kapitel 4 werden weiter das Bauvolumen (77 %), die Bebauungsfläche (66 %), Nutzeranforderungen (62 %) und die Projektorganisation (53 %) häufig dokumentiert. Diese Merkmale sind im ausgewählten Datensatz nicht enthalten. Demnach decken die dokumentierten Merkmale zwar alle fünf Kategorien (Prozess, Person, Produkt, Position, Preis) ab, einige durch die Praxis häufig dokumentierte Projektmerkmale (wie z.B. die Bebauungsfläche) sind jedoch nicht im Datensatz enthalten. Hingegen zeigt dieser Datensatz durch die Vielzahl möglicher Projektkategorien eine extreme Vielfalt auf.

Der Datensatz entspricht daher aufgrund der Anzahl der dokumentierten Projekte sowie der Vielfalt der Projektkategorien nicht einem typischen Datensatz eine Organisation in Deutschland. Auch dieser Datensatz wird nach der Auswahlstrategie von Denzin und Lincoln (2011, S. 307) als extremer Fall betrachtet. So wird mit dieser Fallstudie insbesondere das Ziel verfolgt, mit der Vielfalt der

dokumentierten Projekte relevante Merkmalsattribute zu identifizieren, die eine Fallunterscheidung zwischen Experte und maschinellen Lernmodell aufzeigen. Auch hier können durch die Definition eines extremen Falls, ohne statistische Auswertungen durchzuführen, Rückschlüsse auf die Grundgesamtheit gezogen werden.

## 7.5 Ergebnisse der Datenanalyse

Auf Basis der Ergebnisse des vorherigen Kapitels wird ein CatBoost trainiert. Ergebnis der Grid Search sind 100 maximal trainierte Bäume („n\_estimators“) und eine maximale Baumtiefe von drei. Diese Hyperparameter sind Grundlage des verwendeten maschinellen Lernmodells.

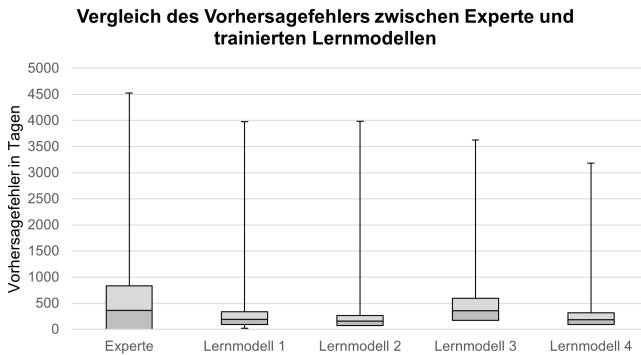
Auf diesen Ergebnissen aufbauend wird jeweils die Vorhersage des trainierten CatBoost-Modells mit der menschlichen Prognose (initiale Planung) verglichen. Tabelle 7.3 gibt wesentliche statistische Kennzahlen zur vorhergesagten Dauer durch den Experten und des maschinellen Lernmodells für die vier Datensätze wieder.

Zusätzlich visualisiert Abbildung 7.4 die Werte der Tabelle 7.3 als Boxplot. Die Abbildung zeigt insbesondere die hohen Ausreißer.



**Tabelle 7.3:** Vergleich des absoluten Vorhersagefehlers (MAE) in Tagen zwischen der Expertenintuition und dem trainierten Modell CatBoost (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren), in fett die Ergebnisse mit dem geringsten Vorhersagefehler je Zeile

	Vorhersagefehler der Expertenintuition ( $H_{e_t}$ )	Vorhersagefehler des trainierten Modells ( $M_{e_t}$ )			
		1	2	3	4
Durchschnitt	528	<b>290</b>	212	486	253
Standardabweichung	653	427	278	491	<b>274</b>
Minimum	<b>0</b>	1	0	2,19	2
25 % Quantil	<b>0</b>	94	73	171	94
50 % Quantil	365	191	<b>157</b>	356,91	184
75 % Quantil	833	336	<b>263</b>	595	319
Maximum	4.528	3.976	3.986	3.625	<b>3.182</b>



**Abbildung 7.4:** Boxplot zum Vergleich des absoluten Vorhersagefehlers in Tagen zwischen der Expertenintuition und dem trainierten Modell CatBoost

Es zeigt sich, dass die durchschnittliche Fehlschätzung des Experten zur Dauer höher ist. Der absolute durchschnittliche Fehler (MAE) des Experten beträgt 528 Tage. Betrachtet man nicht den absoluten Wert, also die Über- und Unterschätzungen der tatsächlichen Dauer, so beträgt der durchschnittliche Vorhersagefehler des Experten -431 Tage. Dies bedeutet, dass durchschnittlich die Dauer unterschätzt wird und bestätigt die systematischen Planungsfehlschlüsse nach Kapitel 2.5.5. Ebenfalls sind bei der Planung durch den Menschen ein deutlich größeres 75 % Quantil sowie ein größerer maximaler Vorhersagefehler zu erkennen. Ausreißer sind demnach beim Menschen höher als beim maschinellen Lernmodell. Vergleicht man die Vorhersagefehler der maschinellen Lernmodelle für die vier Datensätze, so zeigt sich weiter, dass in Datensatz zwei und vier, bei dem externe Daten einbezogen sind, der Vorhersagefehler durchschnittlich geringer ist als in Datensatz eins und drei. Damit ist die Relevanz externer Daten zur Prognose höher als Informationen zu Änderungen. Weiter ist anzumerken, dass jedes maschinelle Lernmodell einen Vorhersagefehler aufweist, wenn auch nur geringe (0,05 Tage Abweichungen beim Minimum in Datensatz zwei).

Aufbauend auf diesen Ergebnissen kann nun entschieden werden, ob jeweils die Differenz der Vorhersagefehler zwischen Experte und CatBoost größer oder kleiner Null ist. Tabelle 7.4 zeigt die Fälle, in denen die bessere Vorhersage entweder durch das maschinelle Lernmodell oder den Experten getroffen werden sollte. So hat z. B. bei Datensatz eins das trainierte Modell in 57,49 % der Fälle eine bessere Prognose getroffen.

**Tabelle 7.4:** Vergleich der Fälle, in denen das trainierte Modell oder der Experte die bessere Prognose durchgeführt hat (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren)

	1		2		3		4	
	Anzahl	Prozent	Anzahl	Prozent	Anzahl	Prozent	Anzahl	Prozent
Trainiertes Modell	211	57,49 %	217	59,12 %	165	44,96 %	147	40,05 %
Experte	156	42,50 %	150	40,87 %	202	55,05 %	220	59,95 %

Vergleicht man die Ergebnisse, so beträgt der Unterschied zwischen trainiertem Modell und Experte hinsichtlich der besseren Prognose maximal 19,90 Prozentpunkte. Insbesondere in Datensatz eins und zwei, bei dem Informationen zu Änderungen vorliegen, zeigt das maschinelle Lernmodell seine Vorteile. Demnach verbessern die externen Daten die Prognosegenauigkeit deutlich, bei Informationen zu Änderungen wird das maschinelle Lernmodell präferiert. Geht man nach diesen Fallzahlen, sollte die Initialprognose verstärkt durch den Experten getätigt werden. Im Laufe des Projektes eignet sich das Lernmodell jedoch besser. Bewertet man nun den Vergleich zwischen Experte und maschinellern Lernmodell, weist das Ausmaß der Fehlschätzung (siehe Tabelle 7.3) durch die hohen Ausreißer der Expertenintuition eine Bevorzugung des maschinellen Lernmodells auf. Nach den absoluten Fallzahlen sollte hingegen eine weitere Fallunterscheidung erfolgen bzw. Experte und Lernmodell mit annähernd gleich gewichteten Machtverhältnissen verwendet werden. Es kann daher für die Bauherren bzw. das Bauprojektmanagement von Interesse sein, die Fälle weiter zu unterscheiden, um relevante Einflussfaktoren zu finden.

Für die Untersuchung relevanter Merkmale wird ein weiteres Modell trainiert, das als Ergebnisvariable die Fälle („0“ für die vorteilhafte Vorhersage durch den Experten oder „1“ für die vorteilhafte Vorhersage durch das maschinelle Lernmodell) unterscheiden soll. Es wird eine CatBoost-Klassifizierung verwendet. Mit diesem Modell werden in einem ersten Schritt die Projektkategorien analysiert.

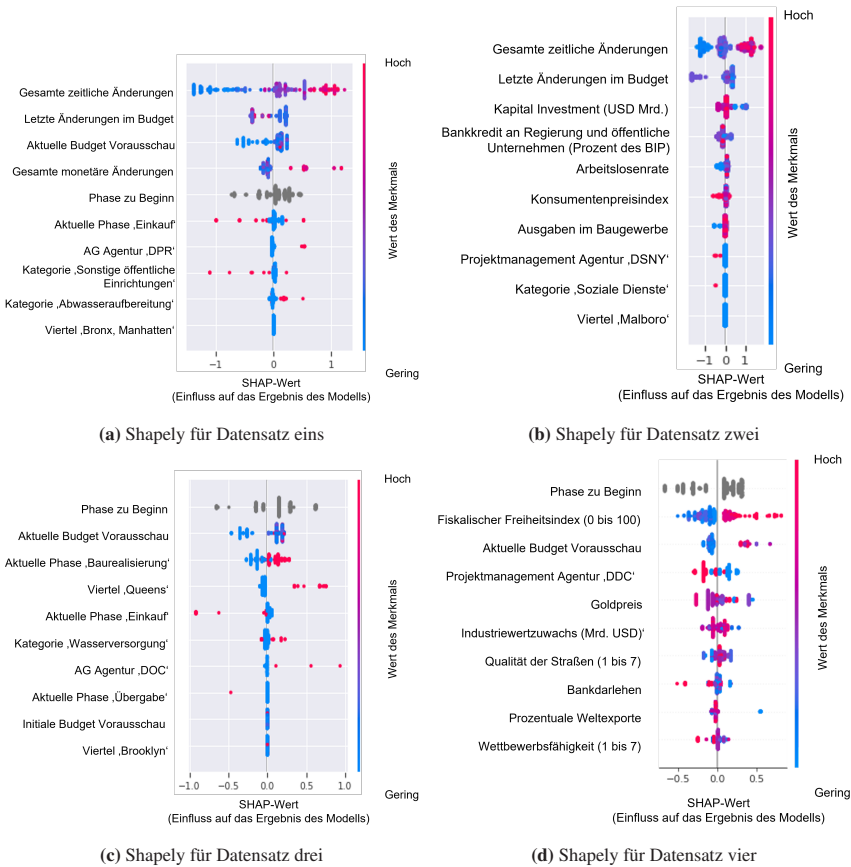
Tabelle 7.5 zeigt die Ergebnisse der prozentualen Fälle je Projektkategorie, in denen das maschinelle Lernmodell, der CatBoost, eine bessere Vorhersage getroffen hat als der Experte. Es werden dabei Kategorien dargestellt, die mehr als zehn Projekte beinhalten.

Betrachtet man die Fallunterscheidung in den einzelnen Kategorien nach den Datensätzen, so lässt sich hier kein klarer Trend je Kategorie beobachten. Vielmehr ist eine Betrachtung je Datensatz notwendig. So eignet sich bspw. das maschinelle Lernmodell insbesondere in der Kategorie „Brücken, Straßen und Verkehrswege“, wenn bereits Informationen zu Änderungen vorhanden sind. Die Fallunterscheidung kann daher nicht allein anhand eines Kriteriums vollzogen werden,

**Tabelle 7.5:** Vergleich der präferierten Fälle für die Expertenintuition oder der CatBoost je Projektkategorie (in Prozent) (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren)

Kategorie	Anzahl Projekte	1		2		3		4	
		Experte	CatBoost	Experte	CatBoost	Experte	CatBoost	Experte	CatBoost
Straßen und Verkehrswege	87	48,28	51,72	48,28	51,72	55,17	44,83	52,87	47,13
Abwasserkanäle	40	40,00	60,00	35,00	65,00	50,00	50,00	62,50	37,50
Industrielle Entwicklung	35	17,14	82,86	48,57	51,43	65,71	34,28	54,29	45,71
Wasserversorgung	33	33,33	66,67	30,30	69,70	54,55	45,45	69,70	30,30
Abwasseraufbereitung	30	30,00	70,00	33,33	66,67	56,67	43,33	76,67	23,33
Schulen	29	44,83	55,17	41,38	58,62	55,17	44,83	58,62	41,38
Brücken	19	42,11	57,89	10,53	89,47	68,42	31,58	68,42	31,58
Parks	19	26,32	73,68	47,37	52,63	63,16	36,84	63,16	36,84
Öffentliche Sicherheit und Strafjustiz	16	50,00	50,00	43,75	56,25	62,50	37,50	50,00	50,00
Brücken, Straßen und Verkehrswege	15	13,33	86,67	6,67	93,33	73,33	26,67	86,67	13,33
Sonstige öffentliche Einrichtungen	12	50,00	50,00	50,00	50,00	50,00	50,00	50,00	50,00

vielmehr scheinen neben den Informationen zu Änderungen mehrere Faktoren beeinflussend zu sein. Aus diesem Grund wird für jeden der Datensätze eine SHAP-Darstellung erstellt, um relevante Merkmalsausprägungen zu ermitteln, die die Bevorzugung von Experte und maschinellem Lernmodell unterscheiden. Abbildung 7.5 zeigt hierfür die zehn wichtigsten Merkmale auf der Grundlage des jeweils verwendeten Datensatzes. Dabei ist der globale Einfluss des Merkmals („Feature Importance“) vertikal absteigend abgebildet. Das Merkmal mit der höchsten Relevanz steht ganz oben, die anderen Merkmale werden in absteigender Reihenfolge angezeigt. Auf der vertikalen Achse ist der lokale Einfluss dargestellt. Jeder Punkt stellt ein Projekt dar. Der Wert des Merkmals („Feature Value“) gibt den Originalwert aus der Datenbank an. Er wird mit der entsprechenden Farbskala in der Grafik dargestellt. Der SHAP-Wert spiegelt den Beitrag des einzelnen Merkmals zur Prognose wider. Da eine präferierte Vorhersage durch das maschinelle Lernmodell mit „1“ gekennzeichnet ist, unterstützen Ausprägungen rechts der Horizontalachse diese Prognose. Umgekehrt zeigen Ausprägungen auf der linken Seite die Vorhersage durch den Experten, da dies mit „0“ in der Zielvariable markiert ist.



**Abbildung 7.5:** Shapely-Darstellung für Datensatz eins bis vier (1: mit Informationen zu Änderungen und ohne externe Marktfaktoren, 2: mit Informationen zu Änderungen und mit externen Marktfaktoren, 3: ohne Informationen zu Änderungen und ohne externen Marktfaktoren, 4: ohne Informationen zu Änderungen und mit externen Marktfaktoren)

Aus den Abbildungen können vier wesentliche Erkenntnisse gezogen werden. Bei diesen Erkenntnissen sind Auswertungen einzelner Merkmale ausgeschlossen, wenn lediglich sehr wenige Projekte zu einer bestimmten Schlussfolgerung führen.

**(1) Bei umfangreichen Projektänderungen sind maschinelle Lernmodelle zu bevorzugen.**

Bei Datensatz eins (mit Informationen zu Änderungen und ohne externe Marktfaktoren) und zwei (mit Informationen zu Änderungen und mit externen Marktfaktoren) zeigt sich, dass je höher die gesamten zeitlichen (und monetären) Änderungen sind, das maschinelle Lernmodell zu bevorzugen ist. Umgekehrt, ist bei geringen zeitlichen Änderungen, der Experte zu bevorzugen.

**(2) Das maschinelle Lernmodell ist in frühen Projektphasen zu bevorzugen.**

Für Datensatz drei (ohne Informationen zu Änderungen und ohne externe Marktfaktoren) und Datensatz vier (ohne Informationen zu Änderungen und mit externe Marktfaktoren) zeigt sich, dass die Phase zu Beginn der Dokumentation das global wichtigste Merkmal bei der Fallunterscheidung ist. Die Phase zu Beginn enthält im Datensatz die Phase „Planung“, „Einkauf“ oder „Baurealisierung“. Extrahiert man aus Datensatz drei die Spalten „Phase zu Beginn“ und die Spalte zur Fallunterscheidung, können die Fälle weiter analysiert werden: Es ist zu erkennen, dass für Projekte, die möglichst früh (in der Planung) im System dokumentiert werden, ein maschinelles Lernmodell zur Planung zu präferieren ist (siehe Tabelle 7.6). Eine erste Projektdokumentation während des Einkaufs oder der Baurealisierung zeigte eine vorteilhafte Prognose durch den Experten auf.

**Tabelle 7.6:** Vergleich der Präferenz des maschinellen Lernmodells und der Expertenintuition anhand der „Phase zu Beginn“ in Datensatz drei

Phase zu Beginn	Präferenz des maschinellen Lernmodells	Präferenz der Expertenintuition	Präferierte Fälle des maschinellen Lernmodells (in Prozent)
Planung	142	42	77,17 %
Einkauf	24	28	46,15 %
Baurealisierung	28	36	43,75 %

Weiter wird in Datensatz drei zusätzlich bei der aktuellen Phase „Baurealisierung“ das maschinelle Lernmodell bevorzugt. Dies kann ein Indikator dafür sein,

dass es zwischen der Planung und Realisierung zu Änderungen kam (siehe Punkt eins).

Die Präferenz einer frühen Dokumentation im System und einer aktuellen späten Phase können ebenfalls auf hohe gesamt Änderungen des Projektes hinweisen.

**(3) In einem instabilen Marktumfeld sind maschinelle Lernmodelle zu bevorzugen.**

Eine hohe Arbeitslosenrate (in Datensatz drei), geringe Kapital Investments (Datensatz drei) oder ein hoher fiskalischer Freiheitsindex (in Datensatz vier), können Hinweise für ein instabiles Marktumfeld sein. Der fiskalische Freiheitsindex gibt die maximale Steuerlast privater und unternehmerischer Einkommen sowie die prozentuale Steuerlast gemessen am BIP an. Damit misst er die steuerlichen Freiheiten eines Staates. Diese Erkenntnis bestätigt die Grundlagen in Kapitel 2.4.1. Hier wird u. a. genannt, dass Experten insbesondere bei stabilen Umweltbedingungen Entscheidungen treffen sollten, trainierte Lernmodelle hingegen in instabilen Umgebungen (vgl. Kahneman und Klein 2009, S. 515ff., Agor 1986, S. 49ff.). Die Punkte eins bis drei bestätigen eine vorteilhafte Prognose durch das maschinelle Lernmodell bei hohen Unsicherheiten.

**(4) Es können Einzelfälle zur Fallunterscheidung identifiziert werden.**

Einzelne Attribute der Merkmale geben weiter Aufschluss über die Planung und Fähigkeiten einzelner Organisationen. So ist die Projektkategorie (z. B. Abwasseraufbereitung in Datensatz eins oder Wasserversorgung in Datensatz drei), das Viertel des Bauprojektes (z. B. Queens in Datensatz drei) oder die zuständige Organisation (z. B. die Projektmanagement Agentur „DDC“ mit guten planerischen Fähigkeiten) in diesem Datensatz ausschlaggebend für eine Fallunterscheidung. Diese Merkmale und ihre Attribute können datensatzspezifisch bei der Fallunterscheidung hinzugezogen werden.

## 7.6 Zusammenfassung und kritische Betrachtung

In Kapitel 7 wird auf Basis eines beispielhaften Datensatzes der Vorhersagefehler maschineller Lernmodelle mit der Expertenintuition verglichen. Weiter werden relevante Einflussfaktoren analysiert, die eine maschinelle Prognose oder die Prognose durch einen Experten von Dauern in Bauprojekten beeinflussen. Diese Analyse soll Unternehmen der Baubranche in einem differenzierten Einsatz maschineller Lernmodelle unterstützen.

Auf Basis des Vergleichs der Vorhersagefehler können folgende Aussagen getroffen werden:

- Maschinelle Lernmodelle weisen zwar zufriedenstellende Ergebnisse auf. Eine eindeutige Präferenz gegenüber der Expertenintuition ist jedoch nicht gegeben (siehe Vergleich der präferierten Fälle in Tabelle 7.4).
- Der menschliche Vorhersagefehler zeigt eine häufige Unterschätzung der Dauer. Dies bestätigt die Theorie der Planungsfehlschlüsse (siehe Kapitel 2.5.5).
- Maschinelle Lernmodelle liefern konsistente Prognosen, der Mensch überschätzt sich in einzelnen Situationen erheblich (Ausreißer).
- Maschinelle Lernmodelle sind insbesondere in einem instabilen Umfeld mit hohen Änderungen im Projektumfeld und in einer frühen Planungsphase zu bevorzugen. Mit der Datensatzanalyse zeigt sich, dass maschinelle Lernmodelle besser mit Unsicherheiten umgehen können. Weiter können Einzelfälle (z. B. Projektkategorie, Viertel oder beteiligte Organisationen) bei der Fallunterscheidung zwischen Experte und maschinellem Lernmodell ausschlaggebend sein.

So können mit dem Einsatz des maschinellen Lernmodells in Situationen mit einer hohen Unsicherheit:



1. Zeit eingespart werden, indem erste Vorschläge generiert werden,
2. die Qualität erhöht werden, indem Ausreißer und damit Planungsfehlschlüsse identifiziert werden,
3. Kosten reduziert werden, indem Zeit eingespart und Planungsfehlschlüsse reduziert werden,
4. die Kreativität erhöht werden, indem geeignetere Vergleichsprojekte sowie relevante Merkmalsausprägungen transparent aufgezeigt werden,
5. neue Geschäftsmodelle generiert werden, indem durch den Einsatz analytischer Planungsmethoden erweiterte Dienstleistungen angeboten werden können.

Die Auswertung der Performance-Indikatoren in Kapitel 6 und Kapitel 7 zeigen jedoch, dass sowohl ein maschinelles Lernmodell als auch eine Prognose durch Experten allein nicht zu einem Optimum einer Entscheidung führen, sondern sich diese aktuell eher gegenseitig ergänzen können (vgl. Jakoby 2019, S. 180). Da bessere maschinelle Vorhersagen mit mehr Daten erzielt werden, kann davon ausgegangen werden, dass die Prognosegenauigkeit schrittweise verbessert wird und maschinelle Lernmodelle immer stärker zu bevorzugen sind.

Der Einsatz des Datensatzes und der durchgeführten Untersuchung ist insbesondere in einem Punkt kritisch zu hinterfragen. Die prognostizierte Dauer innerhalb des Datensatzes wird durch unbekannte Planer erstellt. Über diese liegen keine Daten zu ihrer Erfahrungstiefe und ihrer demografischen Zusammensetzung vor. Es stellt sich daher die Frage, ob diese als wirkliche Experten angesehen werden können. Ist diese Ausgangsbasis als nicht valide anzusehen, ist der Vergleich der Expertenintuition mit dem ausgewählten maschinellen Lernmodell als kritisch zu hinterfragen.

Es ergeben sich folgende anschließende Fragen für die weitere Forschung:

1. Da das maschinelle Lernmodell auf der Basis eines spezifischen Datensatzes trainiert ist, kann kein generisches Modell abgeleitet werden. Es stellt

sich die Frage, inwieweit die identifizierten relevanten Merkmale auf Projekte anderer Kategorie, Lage oder Größe übertragbar sind.

2. Wie beeinflussen weiche Faktoren, die schwer objektiv zu dokumentieren sind, das Training von KI-Modellen? Hier sind zum Beispiel Unternehmenswerte oder die Zusammenarbeit der Projektbeteiligten zu nennen.
3. Schließlich lassen sich auf Basis des Datensatzes keine Aussagen über die Disponenten als Experten ableiten. Erfahrungstiefe und Vorwissen könnten die Prognosefähigkeit des Menschen, aber auch die Interpretation der Ergebnisse der Maschine beeinflussen. Denn die Interpretierbarkeit setzt ein gewisses Vorwissen über Kategorien, Zusammenhänge und Tätigkeiten aufseiten des Nutzers voraus (vgl. Langley et al. 2017, S. 1). Systeme sind interpretierbar, wenn die Prozesse durch Selbsterkenntnis und Erklärungen von einem Menschen verstanden werden können (vgl. Biran und Cotton 2017). Welches Vorwissen der Experten liegt dem Datensatz zugrunde? Welches Vorwissen unterstützt weiter eine maschinelle Vorhersage? Welches Vorwissen ist für die Zusammenarbeit mit der Maschine relevant?

# 8 Handlungsempfehlungen für die Praxis

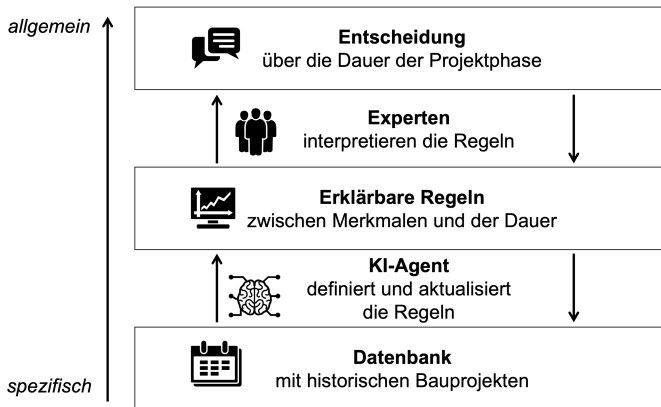
## 8.1 Empfehlungen zur Mensch-Maschinen-Kollaboration

Die vorherigen Kapitel konnten das Potenzial der strukturierten Projektdokumentation, der Auswertung dieser mit analytischen Planungsmethoden und insbesondere der Auswertung mit Entscheidungsbäumen als Teil maschineller Lernmodelle in frühen Phasen eines Bauprojektes aufzeigen. Ziel dieses Kapitels ist die Zusammenstellung der in dieser Arbeit ermittelten Ergebnisse und deren Aufbereitung in Form von Handlungsempfehlungen für Unternehmen in der Praxis. Unternehmen sind nach Kapitel 2.3.3 Bauherrenorganisationen, Unternehmen mit Geschäftsbereichen im Bauprojektmanagement, Generalübernehmer und Totalübernehmer sowie Unternehmen von Baubetreuern und Bauträgern.

Für diese Organisationen haben Fehlprognosen durch maschinelle Lernmodelle ebenso wie durch die Experten weitreichende Auswirkungen. Die Datenanalysen in Kapitel 6 haben gezeigt, dass die Prognosequalität maschineller Lernmodelle zwar weitestgehend zufriedenstellend ist. Die Anzahl der Fälle, in denen die Prognose durch den Menschen zu bevorzugen ist, ist jedoch etwa ausgeglichen mit den Vorhersagen des maschinellen Lernmodells (siehe Kapitel 7). Dies weist auf eine gleichberechtigte Prognose zwischen Experte und maschinellem Lernmodell hin. Eine geringe Prognosegenauigkeit des maschinellen Lernmodells kann zu einer Ablehnung dieses Modells führen. Eine hohe Prognosegenauigkeit hingegen kann zu einem blinden Vertrauen in das Modell führen. Es ist daher zu

empfehlen, dass der Mensch mit seiner Expertise und das maschinelle Lernmodell zusammenarbeiten. Man spricht hier auch von hybriden Systemen, mit denen durch ein gegenseitiges Lernen bessere Ergebnisse erreicht werden können, als sie jeder für sich allein erreicht hätte (vgl. Dellermann et al. 2019, S. 640). Die Stärken beider Systeme werden miteinander kombiniert. So geht auch Jakoby (2019, S. 180) davon aus, dass durch die Kombination von intuitiven und analytischen Methoden Fehler erkannt, korrigiert oder beseitigt werden können. Weiter zeigen sich minimale Unterschiede in der Bevorzugung zwischen Experte und maschinellen Lernmodell bei der Prognose von Dauern eines Bauprojektes in frühen Phasen. Vorhandene Unsicherheiten, wie eine Planung in einer frühen Projektphase, die Instabilität des Marktumfeldes und hohe Projektänderungen, unterstützen die Stärken des maschinellen Lernmodells, hingegen zeigt die Einbindung des Wissens aus nicht dokumentierten Projekteigenschaften Stärken des Experten auf. Als Ergebnis kann das hybride System einen größeren Vorteil als eine selektive Nutzung erbringen. Mit dem maschinellen Lernmodell könnten daher Ausreißer reduziert werden und die Prognose kann in unsicheren, instabilen Situationen unterstützt werden. Der Experte hingegen kann bspw. sein Wissen zu Merkmalen, die nicht dokumentiert sind oder subjektiv und nicht dokumentierbar sind, einbringen. Um hier Erfolge zu generieren, sind erklärbare, nachvollziehbare Ergebnisse des maschinellen Lernmodells notwendig (siehe 2.6). Eine Lösung hierfür können erklärbare KI-Methoden sein (siehe Kapitel 5.6). In Abbildung 8.1 wird ein beispielhaftes Konzept vorgestellt, in dem der Anwender durch KI-Methoden in der Prognose von Dauern eines Bauprojektes unterstützt wird (vgl. Bernard et al. 1991).

Den Modellen liegt eine Datengrundlage historischer Bauprojekte zugrunde. In diesen sind verschiedene Merkmale sowie je Projekt die Dauer der Projektphase, die zu bestimmende Zielgröße, dokumentiert. Auf dieser Datengrundlage werden Regeln durch einen lernenden Agenten (KI-Agent) erstellt. Diese Ebene stellt die algorithmische Umsetzung des Problems dar. Ergebnis sind erklärable Regeln, die auf Basis historischer Daten den Zusammenhang zwischen Projektmerkmalen und der Dauer aufzeigen. Diese Regeln können entweder als Performance-Indikatoren (siehe MAE oder MAPE in Kapitel 5.3.4), Feature Importance (siehe



**Abbildung 8.1:** Ein beispielhaftes Modell zur Zusammenarbeit zwischen Experte und maschinellem Lernmodell (vgl. Bernard et al. 1991)

Kapitel 5.6.3), Beispiele oder Regeln mit Wenn-Dann-Beziehungen (vgl. van der Waa et al. 2021) dargestellt werden. Anwender interpretieren diese Regeln und können abschließend das aktuelle Projekt in die entsprechende Referenzklasse einordnen. Diese Regeln werden automatisch aktualisiert, wenn die Datenmenge wächst. Darüber hinaus können durch den Einsatz von KI-Modellen auch nicht lineare Abhängigkeiten zwischen einzelnen Merkmalen identifiziert werden.

Hemmer et al. (2021) leitet folgende Empfehlungen zur Kollaboration zwischen Experten und maschinellen Lernmodellen ab:

1. Der Experte sollte vor dem Anwenden des maschinellen Lernmodells nach einer eigenen Prognose gefragt werden. Dies verbessere insgesamt die Prognosequalität des hybriden Systems.
2. Je geringer die Prognosequalität des maschinellen Lernmodells, desto mehr Erklärungen sind notwendig.
3. Die Visualisierung des maschinellen Lernmodells kann variieren. So ist zu entscheiden, ob das Ergebnis der Vorhersage oder eher beeinflussende Merkmale dem Experten gezeigt werden sollten.

In diesem Zusammenhang sollte überprüft werden, ob fehlerhafte Prognosen des maschinellen Lernmodells durch den Experten angenommen und sogar durch den Experten verstärkt werden. Das maschinelle Lernmodell sollte in diesem Zusammenhang eher als Erklärung für den Experten verstanden werden. Je nach Übereinstimmung beider Prognosen könnte sich das maschinelle Lernmodell anpassen und als Grundlage für ein kritisches Hinterfragen der Expertenmeinung verwendet werden.

## **8.2 Zusammenfassung der Empfehlungen in Handlungsfelder**

In dem hybriden System spielen ebenfalls die Fähigkeiten, das Wissen und die Persönlichkeit des Experten eine Rolle (vgl. Hemmer et al. 2021). Aus diesem Grund sind durch das Unternehmen weitere Maßnahmen neben der Einführung des technisch hybriden Systems zu beachten. Diese schaffen die entsprechenden Rahmenbedingungen, beginnen bei der Geschäftsstrategie und reichen bis zum Verhalten der Experten.

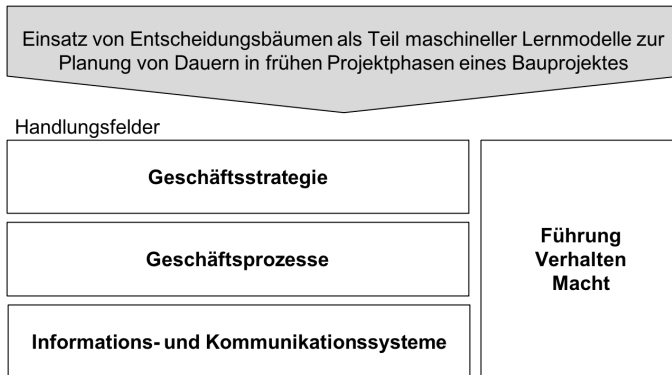
Die Handlungsempfehlungen werden nach den vier übergeordneten Handlungsfeldern zur digitalen Transformation eines Unternehmens nach:

1. Geschäftsstrategie,
2. Geschäftsprozesse,
3. Informations- und Kommunikationssysteme,
4. Führung, Verhalten und Macht

untergliedert (siehe Abbildung 8.2). Mit diesen soll nicht nur die fachlich-technologisch geprägte Perspektive abgedeckt werden, sondern auch politisch-kulturelle Maßnahmen zur Einführung sind maßgeblich. So sind die Motivation,

die Fähigkeiten und Kommunikation oder Machtverhältnisse des prognostizierenden Bauherren bzw. Projektleiters relevant. (vgl. Österle und Winter 2000, S. 12)

Diese Handlungsfelder werden im Folgenden erläutert.



**Abbildung 8.2:** Die vier Handlungsfelder zum Einsatz von Entscheidungsbäumen als Teil maschineller Lernmodelle zur Planung von Dauern in frühen Projektphasen eines Bauprojektes (angelehnt an Österle und Winter 2000, S. 12)

## Geschäftsstrategie

### (1) Erarbeiten eines gemeinsamen Verständnisses von Potenzialen und Grenzen beim Einsatz von Entscheidungsbäumen als Teil maschineller Lernmodelle zur Planung von Dauern eines Bauprojektes in frühen Projektphasen

In Kapitel 6 konnten die theoretischen Potenziale und Grenzen von Entscheidungsbäumen zur Prognose von Dauern aufgezeigt werden. Die durchschnittliche Abweichung (MAPE) zwischen den Ergebnissen des maschinellen Lernmodells und der tatsächlichen Dauer liegt zwischen 13,5 % und 36,8 % (siehe Tabelle 6.4). Dies ist zwar für die Baubeteiligten nach den Ergebnissen der Umfrage in Kapitel 6.2 noch zufriedenstellend, der Experte erzielt aber in annähernd gleich vielen Fällen bessere Ergebnisse (siehe Ergebnisse in Tabelle 7.4). Es ist

daher relevant, die Potenziale und Grenzen der maschinellen Lernmodelle sowie der Expertenintuition zu verstehen. Ziele und eine Vision für den Einsatz von Entscheidungsbäumen sollten gesetzt werden. Auch sind im Rahmen einer Datenstrategie und Verbesserung der Ergebnisse maschineller Lernmodelle Erwartungen an die Dokumentation von Daten (Frequenz und Anzahl und Art der Variablen) festzuhalten.

### **(2) Erarbeiten einer Transformationsstrategie zum Einsatz analytischer Planungsmethoden bzw. maschineller Lernmodelle**

Zur nachhaltigen Implementierung maschineller Lernmodelle im Unternehmen ist der Rahmen für die finanziellen Ressourcen, technologischen Möglichkeiten und die Managementtätigkeiten festzulegen.

## **Geschäftsprozesse**

### **(3) Einführung einer strukturierten digitalen Dokumentation von Bauprojekten entlang der Geschäftsprozesse**

Es sind hierbei die Geschäftsprozesse zu verstehen, in denen aktuell relevante Daten aufgenommen werden und zukünftig dokumentiert werden sollten. Wie Kapitel 6.3.5 aufzeigt, sind insbesondere eine hohe Anzahl an Informationen zu Kosten, dem Produkt und der Lage für die Prognose von Dauern in frühen Projektphasen notwendig. Geschäftsbereiche wie die Buchhaltung sind daher einzu beziehen und zu dokumentierende Merkmale sind gemeinsam zu definieren. Es ist demnach festzulegen, welche Informationen benötigt werden und wie diese Informationen für den Austausch und die Zusammenarbeit strukturiert sein müssen.

### **(4) Einführung der analytischen Planungsmethoden bzw. maschineller Lernmodelle in den Prozess der Rahmenterminplanung**

Es ist ein Zusammenarbeitsmodell entsprechend Abbildung 8.1 festzulegen. Dieses kann in die Planung zu Projektbeginn und während des Projektes unterschieden werden (siehe unterschiedene Szenarien in Kapitel 7). Hier ist das Verständnis nach Empfehlung (1) und auch die Frequenz der Rückkopplung neuer Daten nach Empfehlung (3) relevant.



## **Informations- und Kommunikationssysteme**

### **(5) Implementierung einer Datenplattform**

Die Ergebnisse der Umfrage auf Basis des Interviewleitfadens (siehe Tabelle 4.1) haben aufgezeigt, dass Projektkennzahlen aktuell teilweise sogar in Powerpoint-Präsentationen gespeichert werden. Auf diese kann nur mit hohem manuellen Aufwand zurückgegriffen werden und analytische Auswertungen sind nur erschwert möglich. Es sollte daher eine zentrale Datenplattform für Bauprojekte eingeführt werden. Auf dieser werden die Daten und Informationen gespeichert, ausgetauscht und gemeinsam genutzt werden. Damit werden einheitliche Projektinformationen und -prozesse gewährleistet. Weiterhin werden den Projektbeteiligten hierdurch rechtzeitig aktuelle, relevante und zuverlässige Projektinformationen in einer gemeinsamen und strukturierten Umgebung für die Durchführung eines Projekts zur Verfügung gestellt.

### **(6) Auswahl eines oder mehrerer maschineller Lernmodelle zur Rahmenterminplanung und entsprechende Performance-Metriken zur Bewertung dieser**

In Kapitel 6 werden verschiedene Modelle für Entscheidungsbäume hinsichtlich ihrer Prognosegenauigkeit analysiert. Für die hier verwendeten Datensätze hat sich das CatBoost-Modell als besonders geeignet erwiesen. Es sollte jedoch für den spezifischen Datensatz des Unternehmens eine erneute Überprüfung der Modelleignung (z.B. Random Forest, LightGBM, CatBoost) mit der Prognosegenauigkeit vorgenommen werden. Eine entsprechende Ergebnisvisualisierung im Sinne der Erklärbarkeit kann ebenfalls je Unternehmen unterschiedlich präferiert werden. Auch sind zu nutzende Performance-Metriken auszuwählen. In dieser Arbeit werden der MAPE und MAE als Metriken verwendet.

## **Führung, Verhalten und Macht**

### **(7) Definition der Machtverhältnisse zwischen Experte und maschinellem Lernmodell**

Die Ergebnisse der Fälle in Tabelle 7.4 zeigen, dass ein maschinelles Lernmodell nicht per se zu präferieren ist. Vielmehr eignet sich ein hybrides System, um

die Stärken des Experten und des maschinellen Lernmodells miteinander kombinieren zu können und diese entsprechend zu gewichten. Dieses Machtverhältnis sollte unternehmensintern überprüft und gemeinsam mit den Experten definiert werden.

**(8) (Entwicklung und) Angebote von Qualifikationen zu maschinellen Lernmodellen**

Zur Akzeptanzförderung der maschinellen Lernmodelle und dem Aneignen von entsprechendem Wissen sind Trainings und Qualifikationen zu befürworten. Mit diesen können Potenziale und Herausforderungen besser verstanden werden sowie maschinelle Lernmodelle implementiert und weiterentwickelt werden.

# 9 Schlussbetrachtung

## 9.1 Zusammenfassung

Die Rahmenterminplanung ist in der frühen Planungsphase aufgrund der Projektkomplexität durch eine Vielzahl von Unsicherheiten geprägt. Systematische Planungsfehlschlüsse führen zu weiteren Abweichungen von der tatsächlichen Dauer. Diese Abweichungen können Auswirkungen auf die Kosten und Qualität haben. Sie führen zu Unzufriedenheit und Demotivation. Weiter können sie Konflikte, Nacharbeiten, eine Gefahr der Überarbeitung und fehlende Entscheidungen für Folgeprozesse zur Folge haben. Eine möglichst genaue Planung dieser Dauern als erste Planungsgrundlage des Projektes ist daher essenziell. Mit der zunehmenden Menge dokumentierter Projektdaten und modernen Techniken kann der Einsatz analytischer Planungsmethoden in der Baubranche erwartet werden. In bisherigen Publikationen zur Prognose von Dauern in Bauprojekten werden zum einen sehr einfache nachvollziehbare statistische Methoden, wie lineare Regressionen, zum anderen komplexe Methoden der KI, die häufig einem „Black Box“-Ansatz zugrunde liegen, verwendet. Eine Kombination der Erklärbarkeit in komplexen Methoden ist auf Grundlage der durchgeführten Literaturrecherche kein Gegenstand der Forschung.

Der Einsatz analytischer Planungsmethoden in der Forschung bildet die Ausgangsbasis für die **erste Forschungsfrage**: Wie werden derzeit Rahmenterminpläne für Bauprojekte im Untersuchungsraum Deutschland durch das Projektmanagement erstellt und welche Ansätze zur Reduktion von Abweichungen geplanter Dauern sind statistisch signifikant? In einer Umfrage mit 131 Teilnehmern in Deutschland werden nach dem Vorgehen von Hinkin (1998) Informationen

zum Datenmanagement, der Verwendung dieser Daten und ihrem Einfluss auf die Projektqualität aufgenommen. Die Projektqualität wird zusammen aus den Zeitabweichungen, den Zeitpuffern, Maßnahmen und der Zufriedenheit der Projektbeteiligten definiert. Es wird festgestellt, dass im Mittel 14 Projektmerkmale dokumentiert werden und nur 16 % der Befragten über mehr als 100 dokumentierte Projekte verfügen. Es werden in den meisten Fällen (83 %) analytische Methoden mit der Expertenintuition kombiniert. Bei den analytischen Methoden wird jedoch insbesondere der Datenzugriff für Projektvergleiche (80 %) verwendet. In der bivariaten Untersuchung konnte mithilfe einer EFA ein statistischer Zusammenhang zwischen der Anzahl der dokumentierten Projekte inklusive Projektmerkmale zu den Zeitabweichungen aufgedeckt werden. Um Abweichungen zu reduzieren, sollte daher das Projektmanagement Bauprojekte dokumentieren und auf dieses Wissen zugreifen. Mit der Beantwortung der ersten Forschungsfrage zeigt sich, dass mit der Zunahme dokumentierter Projekte und ihrer Merkmale Abweichungen der geplanten Dauern reduziert werden können. Diese Dokumentation bildet einen Ausgangspunkt für den Einsatz analytischer Planungsmethoden. In der Praxis werden jedoch bisher insbesondere analytische Planungsmethoden mit einem geringen Grad an Intelligenz genutzt.

Darauf aufbauend stellt sich die **zweite Forschungsfrage**: Welche ausgewählten Modelle analytischer Planungsmethoden sind zur Prognose von Dauern in der frühen Bauprojektphase dienlich? Auf Basis der Grundlagen werden die gestalterischen Anforderungen der Erklärbarkeit und Prognosegenauigkeit definiert. Zur weiteren Definition der Anforderungen an die Prognosegenauigkeit aus der Praxis wird mithilfe einer Umfrage mit 87 Teilnehmern die Schwelle definiert, bei der Projektbeteiligte gerade noch zufrieden mit den Abweichungen der geplanten Dauern zum Zeitpunkt der Rahmenterminplanung sind. Der Zielwert analytischer Planungsmethoden wird auf Basis der Umfrageergebnisse mit durchschnittlich 26 % festgelegt. In einem zweiten Schritt werden in zwei Fallstudien ausgewählte analytische Planungsmethoden miteinander verglichen. Hier dienen lineare Regressionsmodelle als Vergleichsmethode, da diese in Publikationen am

häufigsten eingesetzt werden. Weiter werden KNN verwendet, da diese den aktuellen Stand der Forschung mit modernen Techniken darstellen und in Publikationen vielversprechende Ergebnisse aufzeigen. Lineare Regressionsmodelle haben den Vorteil, dass sie einfach nachvollziehbar sind. KNN hingegen können auch schwache Zusammenhänge in Daten identifizieren und diese in Lernmodelle einbeziehen. Diese Methode ist im Vergleich zu einer linearen Regression ohne weitere Methoden nicht nachvollziehbar. Daher werden für die folgende Arbeit Methoden gewählt, die nachvollziehbar und zugleich schwache Zusammenhänge in maschinelle Lernmodelle einbeziehen. Diese beiden Anforderungen sind in Entscheidungsbäumen als Teil von Methoden des maschinellen Lernens zu finden. In beiden Fallstudien werden die Abweichungen der Modellergebnisse zu den tatsächlich geplanten Dauern verglichen. Es zeigt sich, dass Entscheidungsbäume und insbesondere das Modell des CatBoost vergleichsweise gut abschneiden und den Zielwert von 26 % Abweichungen weitestgehend einhalten. Zur Beantwortung der zweiten Forschungsfrage werden zwei Datensätze untersucht, die eine höhere Anzahl an Projekten dokumentiert haben als der Durchschnitt der befragten Organisationen. Schließt man darauf aufbauend auf aktuell durchschnittlich verfügbare Datensätze, zeigt sich, dass davon ausgegangen werden kann, dass hier die Prognosegenauigkeit analytischer Planungsmethoden nicht zufriedenstellend ist. Mit der Zunahme der Projektdokumentation können jedoch auch hier zukünftig analytische Planungsmethoden zu einer zufriedenstellenden Prognosegenauigkeit führen.

Die **Forschungsfrage drei** beschäftigt sich abschließend mit einem Vergleich der Prognosegenauigkeit maschineller Lernmodelle und der Vorhersage durch Experten. Die Forschungsfrage lautet: Inwiefern lässt sich im Vergleich zur Planung durch Experten mit den ausgewählten Modellen analytischer Planungsmethoden die Prognosegenauigkeit von Dauern eines Bauprojektes in einer frühen Planungsphase unterstützen und verbessern? Auch hier dient eine Fallstudie als Grundlage, in der der Vorhersagefehler der Expertenintuition zu dem des maschinellen Lernmodells verglichen wird. Es kann gezeigt werden, dass die Planung durch den Experten im Vergleich zum maschinellen Lernmodell höhere Ausreißer zur tatsächlichen Dauer aufzeigt. Ebenfalls sind Dauern durch den Experten

häufig zu kurz prognostiziert. Dennoch ist eine generelle Präferenz des maschinellen Lernmodells nicht zu bestätigen. Trotz durchschnittlich höherer Fehlschätzungen des Experten sind die Anzahl der Fälle, in denen die Expertenintuition zu besseren Ergebnissen führen, annähernd gleich verteilt mit der bevorzugten Prognose durch maschinelle Lernmodelle. Die Analyse relevanter Merkmale für eine Fallunterscheidung zwischen Experte und maschinellem Lernmodell zeigt weiter, dass sich in einem unsicheren, instabilen und sich stark ändernden Umfeld maschinelle Lernmodelle besser eignen.

Um die durchschnittlich hohen Fehlschätzungen und Ausreißer der Experten zu reduzieren, kann ein maschinelles Lernmodell hinzugezogen werden. Es wird daher eine Kombination zwischen Experte und maschinellem Lernmodell angestrebt und auf Basis der Ergebnisse der drei Forschungsfragen der Einsatz eines KI-Assistenzsystems empfohlen.

Abbildung 9.1 fasst die wesentlichen Ergebnisse auf Basis der drei Forschungsfragen zusammen.

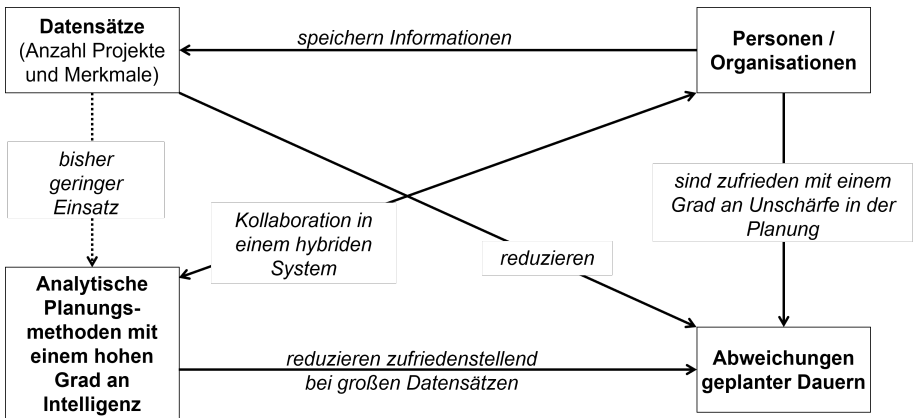


Abbildung 9.1: Zusammenfassung der Ergebnisse

Personen bzw. Organisation speichern Informationen zu Projekten in frühen Planungsphasen ab. Ergebnisse sind Datensätze, bestehend aus einzelnen Projekten

mit Projektmerkmalen. Die Größe der Datenbasis hat einen statistisch signifikanten Einfluss zur Reduktion von Abweichungen geplanter Dauern. Auf Basis der Datensätze werden bisher kaum analytische Planungsmethoden mit einem hohen Grad an Intelligenz eingesetzt. Insbesondere sind Datenzugriffe und Reportings in der Praxis im Untersuchungsraum Deutschland zu finden. Analytische Planungsmethoden mit einem hohen Grad an Intelligenz, wie prädiktive Methoden, erzielen bei einem entsprechend großen Datensatz zufriedenstellende Ergebnisse. Dennoch sollte der Experte gleichberechtigt in die Prognose einbezogen werden, da die Anzahl der Fälle, in denen die Prognosegenauigkeit des Experten höher ist, vergleichbar mit der Prognose des maschinellen Lernmodells ist. Jedoch sollten die Machtverhältnisse zwischen Experte und maschinellem Lernmodell in einzelnen Situationen unterschieden werden. So spielt der Umfang der Änderungen am Projekt und die Stabilität des Marktumfeldes eine Rolle bei der Unterscheidung.

Demnach konnte abschließend aufgezeigt werden, dass analytische Planungsmethoden und insbesondere erklärbare KI-Methoden mit dem CatBoost zur Reduktion von Abweichungen geplanter Dauern in frühen Phasen eines Bauprojekts beitragen. Die Beantwortung der drei Forschungsfragen zeigt die Grenzen der analytischen Planungsmethoden auf: Die aktuelle Verfügbarkeit ausreichend großer Datensätze, Prognosen mit Abweichungen zur tatsächlichen Dauer und eine explizite Bevorzugung in einem unsicheren und instabilen Umfeld.

## 9.2 Kritische Würdigung und Ausblick

In Kapitel 6 und 7 sind unterschiedliche Datensätze als Fallbeispiele analysiert worden. Hier stellt sich die kritische Frage, ob die Ergebnisse dieser Datensätze formal verallgemeinert werden können und damit als induktiver Ansatz ausreichend sind. Es existieren über die Möglichkeiten zu einer Verallgemeinerung auf Basis von Fallstudien verschiedene Auffassungen (vgl. Ruddin 2006). Aus diesem Grund ist anzumerken, dass vielmehr das Ziel dieser Arbeit ist, durch eine gezielte Auswahl an Fallbeispielen Informationen zu aktuellen Herausforderungen und Lösungsmöglichkeiten bei der Rahmenterminplanung zu identifizieren.

Es werden daher nach Flyvbjerg (2021, siehe Tabelle A) extreme Fallstudien betrachtet. Zum einen wird betrachtet, ob selbst mit großen Datensätzen zufriedenstellende Ergebnisse geliefert werden. Zum anderen sind in der Fallstudie aus Kapitel 7 eine maximale Variation an Projekttypen ausgewählt. Durch die Wahl extremer Fälle kann nach Denzin und Lincoln (2011) ein Rückschluss auf die Grundgesamtheit geschlossen werden. Hier kann geschlussfolgert werden, dass die Prognose mit analytischen Planungsmethoden, insbesondere dem CatBoost, und bei einer großen Datenmenge im jeweiligen Unternehmen größtenteils zufriedenstellend ist. Dennoch ist die Prognosegenauigkeit sehr ungenau und zeigt hohe Abweichungen zur tatsächlichen Dauer auf.

Die Datensätze der Fallstudien sind aufgrund der Wahl extremer Fälle keine Grundlage für statistische Analysen. Auf die Berechnung von Effektstärken für die Durchführung einer Metaanalyse wird daher bewusst verzichtet. Es sollten dennoch im Anschluss Analysen mit weiteren Datensätzen durchgeführt werden, um statistische Aussagen treffen können. Zur Verbesserung der Prognosegüte sind daher folgende Forschungsthemen für Folgearbeiten genannt:

- Die hier vorgestellten analytischen Planungsmethoden beziehen keine unvorhergesehenen externen Ereignisse mit ein, bspw. im Bereich Naturschutz, eine unerwartete Bodenqualität oder Kriege. Diese Ursachen könnten weiter zu Abweichungen in den Datenanalysen führen. Ihr Einbezug ist in weiteren Datenanalysen zu untersuchen.
- Es sollten weitere maschinelle Lernmodelle getestet werden. Entscheidungsbäume sind ein möglicher Ansatz zur Verbesserung bestehender Planungsmethoden. Sie versprechen einen Vorteil in der Erklärbarkeit und dem Umgang mit Daten aus Bauprojekten (Merkmale, Fehlwerte, Anzahl dokumentierter Projekte).
- Auch die Kombination von verschiedenen Methoden stellt einen Lösungsansatz dar (KI-Methoden im Ensemble), um die Prognosegenauigkeit weiter zu verbessern. Hier sollten verschiedene Kombinationen zusammen mit Entscheidungsbäumen erprobt werden.



- Zu dokumentierende Projektmerkmale sollten erweitert und anschließend auf deren Einfluss im Prognosemodell untersucht werden.
- Weitere Untersuchungen mit der Zunahme an Datensätzen zu Projekten sind durchzuführen. So ist der Zuwachs an der Prognosegenauigkeit festzuhalten.
- Die Zusammenarbeit zwischen der Expertenintuition und dem maschinellen Lernmodell sollte weiter ausgearbeitet und getestet werden. Der Experte könnte die Prognose des maschinellen Lernmodells durch seine Erfahrungen weiter verbessern und damit die Ausreißer in den Prognosen reduzieren.

Für eine vollständige Analyse der Mensch-Maschinen-Kollaboration müssen immer die drei Dimensionen, die menschliche Prognose, die Prognose des maschinellen Lernmodells und die Realität analysiert werden. Zur Abbildung der Realität ist eine statistisch ausreichend große Anzahl an Projekten notwendig. Diese Analyse wird aufgrund der langen Projektlaufzeiten in dieser Arbeit ausgeschlossen. Es sei angemerkt, dass in den Fallstudien latente Merkmale (z. B. Intelligenz der Experten, Unternehmenskultur) keine Betrachtung fanden. Auch ist der regionale Fokus in Kapitel 4 zu Kapitel 6 und Kapitel 7 zu bemerken. Hier werden die Ergebnisse von Umfragen im Untersuchungsraum Deutschland und der DACH-Region zu Datensätzen aus Amerika und Iran verglichen. Durch die nicht standardisierte Definition der Rahmenterminplanung wird der Fokus der Umfragen regional eingegrenzt und so eine einheitliche Ausgangsbasis als Verständnis geschaffen. Die verwendeten Datensätze haben jedoch einen regional anderen Bezug, sind aber öffentlich verfügbar und entsprechen dem Datensatz einer Organisation im Untersuchungsraum Deutschland, die überdurchschnittlich viele Projekte dokumentiert hat. Dennoch sind weitere Analysen mit Datensätzen aus der DACH-Region zu empfehlen bzw. wäre eine weitere Umfrage im amerikanischen und iranischen Raum durchzuführen, um eine direkte Vergleichbarkeit herstellen zu können. Weiter sei auf die kritischen Würdigungen innerhalb der einzelnen Kapitel verwiesen.

Mithilfe der Datenerhebung und -analysen konnten Ansätze zur Reduktion von Abweichungen geplanter Dauern überprüft werden. Dabei stellen analytische Planungsmethoden mit einem hohen Grad an Intelligenz und insbesondere Modelle der Entscheidungsbäume erfolgversprechende Ansätze dar. Mit der wachsenden Dokumentation an Daten in Bauprojekten eignet sich der Einsatz dieser Methoden immer mehr und eine entsprechende Expertise bei Organisationen der Baubranche sollte aufgebaut werden.

# Literaturverzeichnis

- M.S.B.A. Abd El-Karim, O.A. Mosa El Nawawy, und A.M. Abdel-Alim. Identification and assessment of risk factors affecting construction projects. *HBRC journal*, 13(2):202–216, 2017.
- Mohammad Abdolmohammadi und Arnold Wright. An examination of the effects of experience and task complexity on audit judgments. *Accounting Review*, Seiten 1–13, 1987.
- W.A. Abranovic. *Statistical Thinking and Data Analysis for Managers*. MA: Addison-Wesley, Reading, 1997.
- Andrew Agapiou, David Notman, Roger Flanagan, und George Norman. A time-series analysis of UK annual and quarterly construction output data (1955–1995). *Construction Management and Economics*, 16(4):409–416, 1998.
- Weston H. Agor. How top executives use their intuition to make important decisions. *Business horizons*, 29(1):49–53, 1986.
- Dominic D. Ahiaga-Dagbui und Simon D. Smith. Rethinking construction cost overruns: cognition, learning and estimation. *Journal of financial management of property and construction*, 2014.
- AHO. *Projektmanagementleistungen in der Bau- und Immobilienwirtschaft. Unter Mitarbeit von Norbert Preuß*. Ausschuss der Verbände und Kammern der Ingenieure und Architekten für die Honorarordnung, Bundesanzeiger-Verlagsgesellschaft (Leistungsbild und Honorierung, Nr. 9), Berlin, Köln, 4., vollständig überarbeitete Auflage, Stand: Mai 2014 (korrigierter Nachdruck) edition, 2014.

- Hannsjörg Ahrens, Klemens Bastian, und Lucian Muchowski. *Handbuch Projektsteuerung-Baumanagement - Ein praxisorientierter Leitfaden mit zahlreichen Hilfsmitteln und Arbeitsunterlagen auf CD-ROM*. Stuttgart: Fraunhofer-IRB-Verlag, 2004.
- Igor N. Aizenberg, Naum N. Aizenberg, und Joos Vandewalle. *Multi-Valued and Universal Binary Neurons*. Springer US, 2000.
- Akintola S. Akintoye und Malcolm J. MacLeod. Risk analysis and management in construction. *International journal of project management*, 15(1):31–38, 1997.
- Abdullah Albogamy, Nashwan Dawood, und Darren Scott. A risk management approach to address construction delays from the client aspect. *Computing in Civil and Building Engineering*, Seiten 1497–1505, 2014.
- Hussain Alibrahim und Simone A. Ludwig. Hyperparameter optimization: Comparing genetic algorithm against grid search and bayesian optimization. In *2021 IEEE Congress on Evolutionary Computation (CEC)*, Seiten 1551–1559. IEEE, 2021.
- Yvan Allaire und Mihaela E. Firsirotu. Coping with strategic uncertainty. *MIT Sloan Management Review*, 30(3):7, 1989.
- Stuart D. Anderson, Keith Robert Molenaar, und Cliff J. Schexnayder. *Guidance for cost estimation and management for highway projects during planning, programming, and preconstruction*, volume 574. Transportation Research Board, 2007.
- S. Anjomshoae, A. Najjar, D. Calvaresi, und K. Främling. Explainable agents and robots: Results from a systematic literature review. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, Seiten 1078–1088, Richland, SC, 2019. International Foundation for Autonomous Agents and Multiagent Systems.

- Ghassan Aouad, Michail Kagioglou, Rachel Cooper, John Hinks, und Martin Sexton. Technology management of IT in construction: a driver or an enabler? *Logistics Information Management*, 1999.
- Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bénézet, Siham Tabik, Alberto Barbado, Salvador García, Sergio Gil-López, Daniel Molina, Richard Benjamins, et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58:82–115, 2020.
- Sadi Assaf und Sadiq Al-Hejji. Causes of delay in large construction projects. *International Journal of Project Management*, 24:349–357, 2006.
- David Baccarini. The concept of project complexity - a review. *International journal of project management*, 14(4):201–204, 1996.
- Herman Glenn Ballard. *The last planner system of production control*. PhD thesis, University of Birmingham, 2000.
- Anasse Bari, Mohamed Chaouchi, und Tommy Jung. *Predictive Analytics for Dummies*. John Wiley & Sons, Incorporated, Somerset, United States, 2014.
- Nina Baur und Jörg Blasius. Handbuch Methoden der empirischen Sozialforschung. *Springer VS*, 2., vollständig überarbeitete und erweiterte Auflage, 2019.
- Jean Bedard. Expertise and its relation to audit decision quality. *Contemporary Accounting Research*, 8(1):198–222, 1991.
- D. Bernard, M. Borillo, und B. Gaume. From event calculus to the scheduling problem. semantics of action and temporal reasoning in aircraft maintenance. *Applied Intelligence*, 1:195–221, 01 1991.
- Fritz Berner, Vitali Kochkine, Ilka Habenicht, Sven Spieckermann, und Cornelius Vöth. Simulation in manufacturing planning of buildings. In *2013 Winter Simulations Conference (WSC)*, Seiten 3306–3317. IEEE Press, 2013.

- Önder Halis Bettemir und Rifat Sonmez. Hybrid genetic algorithm with simulated annealing for resource-constrained project scheduling. *Journal of Management in Engineering*, 31:1–8, 2014.
- S. Bhokha und S. Ogunlana. Application of artificial neural network to forecast construction duration of buildings at the predesign stage. *Engineering, Construction & Architectural Management*, 6:133–144, 12 1999.
- S Biggs. Improving auditor judgment through research: a problem and some potential solutions. *Advances in Accounting*, 2:169–84, 1985.
- Muhammad Bilal, Lukumon O. Oyedele, Junaid Qadir, Kamran Munir, Saheed O. Ajayi, Olugbenga O. Akinade, Hakeem A. Owolabi, Hafiz A. Alaka, und Maruf Pasha. Big data in the construction industry: A review of present status, opportunities, and future trends. *Advanced Engineering Informatics*, 30(3): 500–521, 7 2016.
- O. Biran und C. Cotton. Explanation and justification in machine learning: A survey. In *IJCAI-17 workshop on explainable AI (XAI)*, volume 8.1, 2017.
- A. Bockmayr und F. J. Radermacher. Künstliche Intelligenz und Operations Research. In Otthein Herzog, Thomas Christaller, und Dieter Schütt, editors, *Grundlagen und Anwendungen der Künstlichen Intelligenz*, Seiten 249–254, Berlin, Heidelberg, 1993. Springer.
- Eric Bonabeau. Don't Trust Your Gut. *Harvard Business Review*, 81(5):116–23, May 2003. ISSN 0017-8012.
- P. Bonnal, Didier Gourc, und Germain Lacoste. Where do we stand with fuzzy project scheduling? *Journal of Construction Engineering and Management, American Society of Civil Engineers*, 130(1):114–123, 2014.
- Dino Bouchlaghem, Andrew G. Kimmance, und Chimay J. Anumba. Integrating product and process information in the construction sector. *Industrial Management and Data Systems*, 104(3):218–233, 2004.

- Nuhu Braimah und Issaka Ndekugri. Factors influencing the selection of delay analysis methodologies. *International Journal of Project Management*, 26(8): 789–799, 2008.
- Jürg Brandenberger und Ernst Ruosch. Projektmanagement im Bauwesen. *Die-tikon: Baufachverlag AG Zürich*, 1991.
- Dale E. Brashers. Communication and uncertainty management. *Journal of communication*, 51(3):477–497, 2001.
- Frank Braun und Patrick Benz. *Genese natürlicher Entscheidungsprozesse und Determinanten kluger Entscheidungen*. Springer Gabler, Wiesbaden, 2015.
- Heinrich Braun. *Neuronale Netze: Optimierung durch Lernen und Evolution*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.
- Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.
- F.J. Bromilow. Contract time performance expectations and the reality. In *Building forum*, volume 1.3, Seiten 70–80, 1969.
- P. Brézillon. Joint cognitive systems, cooperative systems and decision support systems: A cooperation in context. *Proceedings of the European Conference on Cognitive Science*, 01 1997.
- Roger Buehler, Dale Griffin, und Michael Ross. Exploring the planning fallacy: Why people underestimate their task completion times. *Journal of personality and social psychology*, 67(3):366, 1994.
- Hans-Jörg Bullinger, Kai Wörner, und Juan Prieto. Wissensmanagement - Modelle und Strategien für die Praxis. In *Wissensmanagement*, Seiten 21–39. Springer, 1998.
- Statistisches Bundesamt. Produzierendes Gewerbe: Tätige Personen und Umsatz der Betriebe im Baugewerbe (Zugriff am: 01.10.2022), 2018a. URL [https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Bauen/Publicationen/Downloads-Baugewerbe-Struktur/personen-umsatz-baugewerbe-2040510187004.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Bauen/Publicationen/Downloads-Baugewerbe-Struktur/personen-umsatz-baugewerbe-2040510187004.pdf?__blob=publicationFile).

- Statistisches Bundesamt. Statistisches Jahrbuch 2018: Kapitel 23 Bauen (Zugriff am: 01.10.2022), 2018b. URL <https://www.destatis.de/DE/Themen/Querschnitt/Jahrbuch/jb-bauen.html>.
- JS Busby und EJ Hughes. Projects, pathogens and incubation periods. *International Journal of Project Management*, 22(5):425–434, 2004.
- Stephanie J Byram. Cognitive and motivational factors influencing time prediction. *Journal of Experimental Psychology: Applied*, 3(3):216–239, 1997.
- Danilo Bzdok, Naomi Altman, und Martin Krzywinski. Statistics versus machine learning. *Nature Methods*, 15(4):233–234, 2018.
- Fritz Böhle, Sabine Pfeiffer, und Nese Sevsay-Tegethoff. *Die Bewältigung des Unplanbaren*. Springer-Verlag, 2015.
- Guangming Cao, Yanqing Duan, und Trevor Cadden. The link between information processing capability and competitive advantage mediated through decision-making effectiveness. *International Journal of Information Management*, 44:121–131, 2019.
- Peter Carey. *Data protection: a practical guide to UK and EU law*. Oxford University Press, Inc., 2018.
- A.P.C. Chan. Modelling building durations in hong kong. *Construction Management and Economics*, 17(2):189–196, 1999.
- A.P.C. Chan. Time–cost relationship of public sector projects in malaysia. *International Journal of Project Management*, 19(4):223–229, 2001.
- Daniel W.M. Chan und Mohan M. Kumaraswamy. A comparative study of causes of time overruns in hong kong construction projects. *International Journal of Project Management*, 15(1):55–63, 1997.
- Daniel W.M. Chan und Mohan M. Kumaraswamy. Compressing construction durations: lessons learned from hong kong building projects. *International journal of project management*, 20(1):23–35, 2002.



- Stefan Chanas und Jerzy Kamburowski. The use of fuzzy variables in pert. *Fuzzy Sets and Systems*, 5(1):11–19, 1981.
- Sriram Changali, Azam Mohammad, und Mark van Nieuwland. The construction productivity imperative. *McKinsey*, 2015.
- Jacob Bernard Chassan. *Research design in clinical psychology and psychiatry*. Irvington Pub, 1979.
- T. Chen und C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, Seiten 785–794, 2016.
- W. Tong Chen und Y.H. Huang. Approximately predicting the cost and duration of school reconstruction projects in taiwan. *Construction Management and Economics*, 24(12):1231–1239, 2006.
- Nicola J Chevallier und Alan D Russell. Developing a draft schedule using templates and rules. *Journal of construction engineering and management*, 127(5):391–398, 2001.
- François Chollet. *Deep Learning mit Python und Keras : Das Praxis-Handbuch vom Entwickler der Keras-Bibliothek*, volume 2018 of *mitp Professional*. MITP, 2018.
- I. Choudhury und S. Rajan. Time-cost relationship for residential construction in texas. *CIB Report*, 01 2003.
- David Cleden. *Managing project uncertainty*. Routledge, 2017.
- European Commission. Empfehlung der Kommission vom 6. Mai 2003 betreffend die Definition der Kleinstunternehmen sowie der kleinen und mittleren Unternehmen. Technical Report 32003H0361, Europäische Kommission, May 2003. Code Number: 124 Code: OJ L Legislative Body: COM.
- Michael Copeland. *What’s the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning?* NVIDIA (Hrsg.), 2016.

- Jose M Cortina. What is coefficient alpha? an examination of theory and applications. *Journal of applied psychology*, 78(1):98–104, 1993.
- Robert F Cox, Raja RA Issa, und Dar Ahrens. Management’s perception of key performance indicators for construction. *Journal of construction engineering and management*, 129(2):142–151, 2003.
- Erhard Cramer und Udo Kamps. *Grundlagen der Wahrscheinlichkeitsrechnung und Statistik: Eine Einführung für Studierende der Informatik, der Ingenieur- und Wirtschaftswissenschaften*. Springer Spektrum, Berlin, 2020.
- Adnan Darwiche, Raymond E Levitt, und Barbara Hayes-Roth. Oarplan: generating project plans by reasoning about objects, actions and resources. *AI EDAM*, 2(3):169–181, 1988.
- Sanjoy Dasgupta. Active Learning Theory. In Claude Sammut und Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning*, Seiten 14–19. Springer US, 2010.
- Thomas H. Davenport und Jeanne G. Harris. *Competing on analytics: the new science of winning / Thomas H. Davenport and Jeanne G. Harris*. Harvard Business School Press, Boston, Mass, 2007.
- Cameron Davidson-Pilon. *Bayesian Methods for Hackers: Probabilistic Programming and Bayesian Inference*. Addison-Wesley, 2016.
- Rina Dechter. Learning while searching in constraint-satisfaction problems, 1986.
- Dominik Dellermann, Philipp Ebel, Matthias Söllner, und Jan Marco Leimeister. Hybrid intelligence. *Business & Information Systems Engineering*, 61(5):637–643, 2019.
- W. Denk. Artificial intelligence, machine learning and data science: Same same but different?! (Zugriff am 01.07.2022), 2017. URL <https://www.capgemi.com/de-de/2017/09/artificial-intelligence-machine-learning-und-data-science-same-same-but-different/#>.

- Norman K Denzin und Yvonna S Lincoln. *The Sage handbook of qualitative research*, chapter 17, Seiten 301–316. Thousand Oaks, CA: Sage, 4 edition, 2011.
- Claus Jürgen Diederichs. *Immobilienmanagement im Lebenszyklus. Projektentwicklung, Projektmanagement, Facility Management, Immobilienbewertung*. Springer-Verlag, Berlin Heidelberg New York, 2. erw. und aktualisierte Aufl. edition, 2006.
- Claus Jürgen Diederichs, Norbert Preuß, und Alexander Malkwitz. *Projektmanagement*, Seiten 281–314. Springer Fachmedien, Wiesbaden, 2020.
- DIN69901-5. *2009-01 Projektmanagement - Projektmanagementsysteme - Teil 5: Begriffe*. Beuth Verlag, 2009.
- S. Dissanayaka und M. Kumaraswamy. Evaluation of factors affecting time and cost performance in hong kong building projects. *Construction and Architectural Management*, 6(3):287–298, 1999.
- Vijaya Dixit, Atanu Chaudhuri, und Rajiv K. Srivastava. Procurement scheduling in engineer procure construct projects: a comparison of three fuzzy modelling approaches. *International Journal of Construction Management*, 18(3):189–206, 2018.
- Hemanta Doloi, Anil Sawhney, K. Iyer, und Sameer Rentala. Analysing factors affecting delays in indian construction projects. *International Journal of Project Management*, 30, 05 2012.
- Marco Dorigo und Luca Maria Gambardella. Ant colonies for the travelling salesman problem. *biosystems*, 43(2):73–81, 1997.
- Anna Veronika Dorogush, Vasily Ershov, und Andrey Gulin. Catboost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*, 2018.
- E. Douglas. Schedule levels of detail — as applied in engineering, procurement and construction. *AACE Int. Recommended Practice No. 37R*, 6, 2010.

- Chris Drummond. Classification. In Claude Sammut und Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning*. Springer US, 2010.
- Melvin Dubnick. Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review*, 28(3):376–417, 2005.
- Duden. Stichwort: Methode (Zugriff am: 2022-06-15), 2022a. URL <https://www.duden.de/rechtschreibung/Methode>.
- Duden. Stichwort: Modell (Zugriff am: 2022-06-15), 2022b. URL <https://www.duden.de/rechtschreibung/Modell>.
- Duden. Stichwort: Prognose (Zugriff am: 15.06.2022), 2022c. URL <https://www.duden.de/rechtschreibung/Modell>.
- Bruce Edmonds. *Syntactic measures of complexity*. The University of Manchester (United Kingdom), 1999.
- Mohamed A. El-Gafy. Construction time-cost trade-off analysis using ant system. *International Journal of Construction Education and Research*, 3(2): 99–108, 2007.
- M. El-Razek, H. Bassioni, und A.M. Mobarak. Causes of delay in building construction projects in egypt. *Journal of Construction Engineering and Management*, 134:831–841, 11 2008.
- Augustine Elinwa und Joshua Mangyvat. Time-overrun factors in nigerian construction industry. *Journal of Construction Engineering and Management*, 127:419–425, 10 2001.
- Adnan Enshassi, Sherif Mohamed, und Saleh Abushaban. Factors affecting the performance of construction projects in the gaza strip. *Journal of Civil engineering and Management*, 15(3):269–280, 2009.
- Ercan Erdis. The effect of current public procurement law on duration and cost of construction projects in turkey. *Journal of Civil Engineering and Management*, 19:121–135, 01 2013.

- Wolfgang Ertel. *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*. Springer Vieweg, Wiesbaden, 4., überarbeitete Auflage edition, 2016.
- Klaus Eschenbruch und C. Rintelen. Bauablaufstörung und Terminfortschreibung nach der VOB/B. *NZBau*, 7:401–464, 2010.
- Martin A. Fischer und Florian Aalami. Scheduling with computer-interpretable construction method models. *Journal of Construction Engineering & Management*, 122(4):337–347, 1996.
- Bent Flyvbjerg. From nobel prize to project management: Getting risks right. *Project management journal*, 37(3):5–15, 2006.
- Bent Flyvbjerg. Policy and planning for large-infrastructure projects: problems, causes, cures. *Environment and Planning B: planning and design*, 34(4):578–597, 2007.
- Bent Flyvbjerg. Curbing optimism bias and strategic misrepresentation in planning: Reference class forecasting in practice. *European planning studies*, 16(1):3–21, 2008.
- Bent Flyvbjerg. Yes, you can generalize from a case study - despite what your supervisor says (Zugriff am: 13.08.2022), 2021. URL <https://medium.com/geekculture/yes-you-can-generalize-from-a-case-study-f8000bd647b9>.
- Bent Flyvbjerg, Mette Skamris Holm, und Soren Buhl. Underestimating costs in public works projects: Error or lie? *Journal of the American planning association*, 68(3):279–295, 2002.
- Bent Flyvbjerg, Chi-keung Hon, und Wing Huen Fok. Reference class forecasting for hong kong’s major roadworks projects. In *Proceedings of the Institution of Civil Engineers-Civil Engineering*, volume 169.6, Seiten 17–24. Thomas Telford Ltd, 2016.
- Nuria Forcada, Alba Fuertes, Marta Gangolells, Miquel Casals, und Marcel Marcella. Knowledge management perceptions in construction and design companies. *Automation in construction*, 29:83–91, 2013.

- J. Kevin Ford, Robert C. MacCallum, und Marianne Tait. The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel psychology*, 39(2):291–314, 1986.
- David M. Frederick und Robert Libby. Expertise and auditors' judgments of conjunctive events. *Journal of Accounting Research*, Seiten 270–290, 1986.
- Yaw Frimpong, Jacob Oluwoye, und Lynn Crawford. Causes of delay and cost overruns in construction of groundwater projects in a developing countries; ghana as a case study. *International Journal of Project Management*, 21(5): 321–326, 2003.
- Mark Funkhouser. Accountability, performance and performance auditing: reconciling the views of scholars and auditors. *Performance auditing: Contributing to accountability in democratic government*, 209, 2011.
- Jay Galbraith. Designing complex organizations. *Reading, Mass*, 1973.
- Michael P. Gallaher, Michael P. Gallaher, John L. Dettbarn, und Linda T. Gilday. *Cost analysis of inadequate interoperability in the US capitol facilities industry*. US Department of Commerce, National Institute of Standards and Technology, 2004.
- José María Rodríguez García. Scientia potestas est—knowledge is power: Francis bacon to michel foucault. *ANGLIA*, 2001.
- Joana Geraldi, Harvey Maylor, und Terry Williams. Now, let's make it really complex (complicated): A systematic review of the complexities of projects. *International journal of operations & production management*, Seiten 966–990, 2011.
- Joana G. Geraldi und Gerald Adlbrecht. On faith, fact, and interaction in projects. *Project Management Journal*, 38(1):32–43, 2007.
- Gerd Gigerenzer. *Bauchentscheidungen: die Intelligenz des Unbewussten und die Macht der Intuition*. C. Bertelsmann Verlag, 2015.

- Gerhard Girmscheid. *Risikomanagement in Bauprojekten und Bauunternehmen*, Seiten 645–754. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.
- Ben und Cassio Pennachin Goertzel. *Artificial General Intelligence. Cognitive Technologies*. Springer-Verlag, Berlin Heidelberg New York, 2007.
- Eliyahu M Goldratt. *Die Kritische Kette: Das neue Konzept im Projektmanagement*. Campus Verlag, Frankfurt/Main, 2002.
- Hamed Golizadeh, Aidin Sadeghifam, Hamid Aadal, und Muhd-Zaimi Abd Majid. Automated tool for predicting duration of construction activities in tropical countries. *KSCE Journal of Civil Engineering*, 20:12–22, 03 2015.
- Mike Gralla. *Baubetriebslehre-Bauprozessmanagement*. Werner, 2011.
- C. Gray. ‘Intelligent’ construction time and cost analysis. *Construction Management and Economics*, 4(2):135–150, 1986.
- Peter Greiner, Peter E Mayer, und Karlhans Stark. *Baubetriebslehre-Projektmanagement*. Springer-Verlag, 2005.
- Peter Greiner, Peter E. Mayer, und Karlhans Stark. *Baubetriebslehre - Projektmanagement. Erfolgreiche Steuerung von Bauprojekten*. Vieweg + Teubner (Studium), Wiesbaden, 4. akt. Aufl. 2009 edition, 2009.
- Heinz Grote. Hohe Produktivität in selbstorganisierten Organisationsstrukturen–Komplexitätsbeherrschung mit Managementkybernetik. *Entscheiden in komplexen Systemen*, 20, 2000.
- Yufeng Guo. *The 7 Steps of Machine Learning. Towards Data Science*, 2017.
- Aurélien Géron und Kristian Rother. *Praxiseinstieg Machine Learning mit Scikit-Learn und TensorFlow : Konzepte, Tools und Techniken für intelligente Systeme*. O’Reilly, 2018.
- S Haghsheno, C Baier, MR-D Budau, A Schilling Miguel, P Talmon, und L Frantz. Structuring approach for integrated project delivery [strukturierungsansatz für das modell der integrierten projektentwicklung (ipa)]. *BAUINGENIEUR*, 97(3):63–76, 2022.

- Shervin Haghsheno. *Analyse der Chancen und Risiken des GMP-Vertrags bei der Abwicklung von Bauprojekten*. Mensch & Buch Verlag, 2004.
- Arthur D. Hall. Fagen, re, 1956: Definition of system. *A: General Systems: Yearbook for*, 1956.
- Peter Hall. *Great Planning Disasters: With a New Introduction*, volume 1. Univ of California Press, 1982.
- M. Hapke, Andrzej Jaszkiwicz, und Slowinski Roman. Fuzzy project scheduling with multiple criteria. In *Proceedings of 6th International Fuzzy Systems Conference*, Seiten 1277 – 1282 vol.3, 08 1997.
- Robert J. Harvey, Robert S. Billings, und Kevin J. Nilan. Confirmatory factor analysis of the job diagnostic survey: Good news and bad news. *Journal of applied psychology*, 70(3):461–468, 1985.
- Patrick Hemmer, Max Schemmer, Michael Vössing, und Niklas Kühl. Human-ai complementarity in hybrid intelligence systems: A structured literature review. *PACIS*, Seiten 1–14, 2021.
- Chris Hendrickson, Carlos Zozaya-Gorostiza, Daniel Rehak, Eduardo Baracco-Miller, und Peter Lim. Expert system for construction planning. *Journal of Computing in Civil Engineering*, 1(4):253–269, 1987.
- Z. Herbsman und R. Ellis. Research of factors influencing construction productivity. *Construction management and economics*, 8(1):49–61, 1990.
- Robin Hill. What sample size is “enough” in internet survey research. *Interpersonal Computing and Technology: An electronic journal for the 21st century*, 6(3-4):1–12, 1998.
- Timothy R. Hinkin. A Brief Tutorial on the Development of Measures for Use in Survey Questionnaires. *Organizational Research Methods*, 1(1):104–121, January 1998. Publisher: SAGE Publications Inc.



- Jimmie Hinze, Gregory Selstead, und Joe P Mahoney. Cost overruns on state of washington construction contracts. *Transportation Research Record*, 1351: 87–93, 1992.
- HOAI. *HOAI 2013-Textausgabe/HOAI 2013-Text Edition - Honorarordnung für Architekten und Ingenieure vom 10. Juli 2013*. Springer-Verlag, Berlin Heidelberg New York, 6. aufl. edition, 2018.
- Robert Hoffman, G. Klein, und S. Mueller. Explaining explanation for “explainable ai”. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 62:197–201, 09 2018.
- Wilfried Johannes Hoffmann. *Zum Umgang mit der Komplexität von Bauvorhaben (Indikatorbezogenes Modell zur Bewertung von Komplexität in Bauprojekten)*. PhD thesis, Technische Universität Kaiserslautern, 2017.
- White House. Preparing for the future of artificial intelligence. executive office of the president national science and technology council. committee on technology, 2016.
- Laura Huang. When It’s OK to Trust Your Gut on a Big Decision. *Harvard Business Review*, October 2019. ISSN 0017-8012. URL <https://hbr.org/2019/10/when-its-ok-to-trust-your-gut-on-a-big-decision>.
- Frank Hutter, Jörg Lücke, und Lars Schmidt-Thieme. Beyond manual tuning of hyperparameters. *KI-Künstliche Intelligenz*, 29(4):329–337, 2015.
- R.J. Hyndman und A.B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006.
- Rob Hyndman. Another look at forecast accuracy metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting*, 4:43–46, 01 2006.
- Michael Häder. *Empirische Sozialforschung: eine Einführung*. Springer VS, Wiesbaden, 3. Auflage edition, 2015.

- V. Ireland. The role of managerial actions in the cost, time and quality performance of high-rise commercial building projects. *Construction Management and Economics*, 3:59–87, 01 1983.
- Muhammad Irfan, Muhammad Bilal Khurshid, Panagiotis Anastasopoulos, Samuel Labi, und Fred Moavenzadeh. Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type, and contract type. *International Journal of Project Management*, 29(1):78–92, 2011.
- Walter Jakoby. *Projektmanagement für Ingenieure*. Springer Vieweg, Wiesbaden, 4. Auflage edition, 2019.
- Lawrence R Jauch und Kenneth L Kraft. Strategic management of uncertainty. *Academy of management review*, 11(4):777–790, 1986.
- M. Tim Jones. Models for machine learning (Zugriff am: 01.07.2019). *IBM Developer*, 2017. URL <https://developer.ibm.com/articles/cc-models-machine-learning/>.
- Minhyuk Jung, Moonseo Park, Hyun-Soo Lee, und Hyunsoo Kim. Weather-delay simulation model based on vertical weather profile for high-rise building construction. *Journal of Construction Engineering and Management*, 142, 01 2016.
- Daniel Kahneman. Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, 93(5):1449–1475, 2003.
- Daniel Kahneman und Gary Klein. Conditions for intuitive expertise: a failure to disagree. *American psychologist*, 64(6):515–526, 2009.
- Daniel Kahneman und Amos Tversky. Intuitive prediction: Biases and corrective procedures. Technical report, Decisions and Designs Inc Mclean Va, 1977.
- E. Kalhor, M. Khanzadi, E. Eshtehardian, und A. Afshar. Stochastic time–cost optimization using non-dominated archiving ant colony approach. *Automation in Construction*, 20(8):1193–1203, 2011.

- Chabota Kaliba, Mundia Muya, und Kanyuka Mumba. Cost escalation and schedule delays in road construction projects in zambia. *International Journal of Project Management*, 27(5):522–531, 2009.
- Konrad Wolfgang Kallus. *Erstellung von Fragebogen*. Facultas, Wien, 2016.
- Peter F. Kaming, Paul O. Olomolaiye, Gary D. Holt, und Frank C. Harris. Factors influencing construction time and cost overruns on high-rise projects in indonesia. *Construction Management & Economics*, 15(1):83–94, 1997.
- Alaattin Kanoglu. An integrated system for duration estimation in design/build projects and organizations. *Engineering, Construction and Architectural Management*, 10:272–282, 08 2003.
- Andreas Kaplan und Michael Haenlein. Siri, siri, in my hand: Who’s the fairest in the land? on the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1):15–25, 2019.
- Sadahisa Kato und Jack Ahern. ‘learning by doing’: adaptive planning as a strategy to address uncertainty in planning. *Journal of environmental planning and management*, 51(4):543–559, 2008.
- Erdal Kayacan und Mojtaba Ahmadi Khamesar. *Fuzzy neural networks for real time control applications: Concepts, modeling and algorithms for fast learning*. Elsevier, Amsterdam, 2016.
- G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, und T.Y. Liu. Lightgbm: A highly efficient gradient boosting decision tree. In *Advances in neural information processing systems*, Seiten 3146–3154, 2017.
- T. Kessel, M. Gawlitta, C. Hilbig, und M. Walther. *Aspekte der Baubetriebslehre in Forschung und Praxis - Festschrift zum 60. Geburtstag von Univ.-Prof. Dr.-Ing. Dipl.-Kfm. Dieter Jacob*. Springer-Verlag, Berlin Heidelberg New York, 2015.
- Scott Kirkpatrick, D. Jr, und Mario Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 01 1983.

- Moritz Kirste und Markus Schürholz. Einleitung: Entwicklungswege zur KI. In *Künstliche Intelligenz: Technologie | Anwendung | Gesellschaft*, Seiten 21–35. Springer Berlin Heidelberg, 2019.
- Steven W. Knox. *Machine Learning: A Concise Introduction*. Wiley Series in Probability and Statistics. Wiley, 1 edition, 2018.
- Bernd Kochendörfer, Jens Liebchen, und Markus Viering. *Bau-Projekt-Management - Grundlagen und Vorgehensweisen*. Springer-Verlag, Berlin Heidelberg New York, 4. Aufl. edition, 2010.
- Sander Koole und Mascha van't Spijker. Overcoming the planning fallacy through willpower: Effects of implementation intentions on actual and predicted task-completion times. *European Journal of Social Psychology*, 30(6): 873–888, 2000.
- P.A. Koushki, K. Al-Rashid, und N. Kartam. Delays and cost increases in the construction of private residential projects in kuwait. *Construction Management and Economics*, 23:285–294, 03 2005.
- Thomas Krickhahn. *Statistik für Naturwissenschaftler für Dummies / Thomas Krickhahn*. Wiley, Weinheim, 2017.
- Andreas Kroll. *Computational Intelligence: eine Einführung in Probleme, Methoden und technische Anwendungen*. Oldenbourg, 2013.
- Justin Kruger und Matt Evans. If you don't want to be late, enumerate: Unpacking reduces the planning fallacy. *Journal of Experimental Social Psychology*, 40(5):586–598, 2004.
- Rudolf Kruse, Christian Borgelt, Christian Braune, Frank Klawonn, Christian Moewes, und Matthias Steinbrecher. *Computational intelligence: eine methodische Einführung in künstliche neuronale Netze, evolutionäre Algorithmen, Fuzzy-Systeme und Bayes-Netze*. Computational intelligence. Springer Vieweg, 2., überarbeitete und erweiterte Auflage edition, 2015.

- Niklas Kühl, Marc Goutier, Robin Hirt, und Gerhard Satzger. Machine learning in artificial intelligence. *Electron Markets*, 8(3):159–164, 2022.
- Vellanki Kumar und Wael Abdullah. Optimization of uncertain construction time-cost trade off problem using simulated annealing algorithm. *2011 World Congress on Information and Communication Technologies - WICT 2011*, Seiten 489–494, 2011.
- Vellanki Kumar und G. Reddy. Fuzzy logic approach to forecast project duration in construction projects. In *Construction Research Congress 2005: Broadening Perspectives*, Seiten 1–10, 08 2005.
- Mohan M. Kumaraswamy und Daniel W.M. Chan. Determinants of construction duration. *Construction Management and Economics*, 13(3):209–217, 1995.
- Ziva Kunda. The case for motivated reasoning. *Psychological bulletin*, 108(3): 480–498, 1990.
- Carmen Lacave und Francisco Díez. A review of explanation methods for heuristic expert systems. *The Knowledge Engineering Review*, 19:133–146, 06 2004.
- K. Lam und O. Olalekan. Forecasting construction output: A comparison of artificial neural network and box-jenkins model. *Engineering, Construction & Architectural Management*, 23:302–322, 05 2016.
- Werner Langen. Die bauzeit im rahmen der vertragsgestaltung. *NZBau-Neue Zeitschrift für Baurecht und Vergaberecht*. Frankfurt a. M.: Verlag CH Beck, 3:145–208, 2009.
- P. Langley, B. Meadows, M. Sridharan, und D. Choi. Explainable agency for intelligent autonomous systems. In *AAAI*, Seiten 4762–4763, 2017.
- Georg Langs und Rania Wazir. Statistik Grundlagen. In *Handbuch Data Science: mit Datenanalyse und Machine Learning Wert aus Daten generieren / Stefan Papp, Wolfgang Weidinger, Mario Meir-Huber, Bernhard Ortner, Georg Langs, Rania Wazir*. Hanser, München, 2019.

- Sir Michael Latham. Constructing the team. *Joint Review of Procurement and Contractual Arrangements in the United Kingdom Construction Industry*, 1994.
- Svenja Lauble und Shervin Haghsheno. Predicting the construction duration in the predesign phase with decision trees. In *eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2022*. CRC Press, September, 2022 2023.
- Long Le-Hoai und Young Lee. Time-cost relationships of building construction project in korea. *Facilities*, 27:549–559, 07 2009.
- Long Le-Hoai, Young Lee, und Jun-Yong Lee. Delay and cost overruns in vietnam large construction projects: A comparison with other selected countries. *KSCCE Journal of Civil Engineering*, 12:367–377, 11 2008.
- A. Liaw, M. Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.
- C.S. Lim und M. Zain Mohamed. Criteria of project success: an exploratory re-examination. *International journal of project management*, 17(4):243–248, 1999.
- W.-M. Lippe. *Soft-Computing: Mit Neuronalen Netzen, Fuzzy-Logic Und Evolutiven Algorithmen*. eXamen.Press. Springer, 2006.
- Min Liu, William Rasdorf, Donna A. Hollar, Joseph E. Hummer, und Shalin C. Parikh. *Predicting the Preliminary Engineering Duration of Bridge Projects*, Seiten 505–514. Construction Research Congress 2012, 2012.
- Daniel Lovallo, Dan; Kahneman. Delusions of success. *Harvard business review*, 81(7):56–63, 2003.
- Peter E.D. Love und Dominic Ahiaga-Dagbui. Debunking fake news in a post-truth era: the plausible untruths of cost underestimation in transport infrastructure projects. *Transportation research part A: policy and practice*, 113:357–368, 2018.

- Peter E.D. Love, David J Edwards, und Zahir Irani. Moving beyond optimism bias and strategic misrepresentation: An explanation for social infrastructure project cost overruns. *IEEE transactions on engineering management*, 59(4): 560–571, 2011.
- Peter E.D. Love, Chun-Pong Sing, Brad Carey, und Jeong Tai Kim. Estimating construction contingency: Accommodating the potential for cost overruns in road construction projects. *Journal of Infrastructure Systems*, 21(2):1–10, 2015a.
- Peter E.D. Love, Jim Smith, Ian Simpson, Michael Regan, und Oluwole Olatunji. Understanding the landscape of overruns in transport infrastructure projects. *Environment and Planning B: Planning and Design*, 42(3):490–509, 2015b.
- Y. Lu und S.M. AbouRizk. Automated box–jenkins forecasting modelling. *Automation in Construction*, 18(5):547–558, 2009.
- Scott M. Lundberg. Interpretable machine learning with xgboost. *Towards Data Science*, 2018.
- Scott M. Lundberg, Gabriel G. Erion, und Su-In Lee. Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*, 2018.
- Jan Lunze. *Künstliche Intelligenz für Ingenieure: Methoden zur Lösung ingenieurtechnischer Probleme mit Hilfe von Regeln, logischen Formeln und Bayesnetzen*. De Gruyter Studium. De Gruyter Oldenbourg, 3., überarbeitete Auflage edition, 2016.
- Uwe Lämmel und Jürgen Cleve. *Künstliche Intelligenz: Lehr- und Übungsbuch*. Fachbuchverlag Carl Hanser, München, 3. Aufl. edition, 2008.
- Ole M. Magnussen und Nils O.E. Olsson. Comparative analysis of cost estimates of major public investment projects. *International Journal of Project Management*, 24(4):281–288, 2006.
- James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Angela Hung Byers, et al. *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute, 2011.

- Harvey R. Maylor, Neil W. Turner, und Ruth Murray-Webster. How hard can it be?: Actively managing complexity in technology projects. *Research-Technology Management*, 56(4):45–51, 2013.
- John McCarthy, Marvin L. Minsky, Nathaniel Rochester, und Claude E. Shannon. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (Zugriff am 08.05.2019), 1955. URL <http://www-formal.stanford.edu/jmc/history/dartmouth.pdf>.
- Warren S. McCulloch und Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.
- Xianhai Meng. The effect of relationship management on project performance in construction. *International Journal of Project Management*, 30, 02 2012.
- Isaac Mensah, Theophilus Adjei-Kumi, und Gabriel Nani. Duration determination for rural roads using the principal component analysis and artificial neural network. *Engineering, Construction and Architectural Management*, 23:638–656, 09 2016.
- Giovanni C. Migliaccio und Pramen P. Shrestha. *Analysis of Design-Build Procurement Activities Durations for Highway Projects*, Seiten 229–238. Construction Research Congress 2009, 2009.
- Henry W. Miller. Plan and operation of the health and nutrition examination survey, united states, 1971-1973. *DHEW publication no.(PHS)-Dept. of Health, Education, and Welfare (USA)*, 1973.
- T. Miller, P. Howe, und L. Sonenberg. Explainable ai: Beware of inmates running the asylum or: How i learnt to stop worrying and love the social and behavioural sciences. *arXiv preprint arXiv:1712.00547*, Seiten 1–7, 12 2017.
- Hans-Joachim Mittag und Katharina Schüller. *Grundbegriffe der Statistik*, Seiten 21–30. Springer Berlin Heidelberg, Berlin, Heidelberg, 2020.
- C. Molnar. *Interpretable Machine Learning*. Lulu.com, 2019.



- Yvonne Moret und Herbert Einstein. Construction cost and duration uncertainty model: Application to high-speed rail line project. *Journal of Construction Engineering and Management*, 142:1–13, 04 2016.
- Katharina Morik. Schlüsseltechnologie maschinelles lernen. *Digitale Welt*, 2(4): 22–27, 2018.
- Erhard Motzel und Thor Möller. *Projektmanagement Lexikon. Referenzwerk zu den aktuellen nationalen und internationalen PM-Standards*. John Wiley and Sons, New York, 3. Aufl. edition, 2017.
- A.J. Myles, R.N. Feudale, Y. Liu, N.A. Woody, und S.D. Brown. An introduction to decision tree modeling. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 18(6):275–285, 2004.
- Gopal M. Naik und M. Kumar. Project cost and duration optimization using soft computing techniques. *Journal of Advanced Management Science*, 1(3): 299–303, 2013.
- F. Nasirzadeh und P. Nojedehi. Dynamic modeling of labor productivity in construction projects. *International journal of project management*, 31(6):903–911, 2013.
- D. Navinchandra, D. Sriram, und R.D. Logcher. Ghost: project network generator. *Journal of Computing in Civil Engineering*, 2(3):239–254, 1988.
- Stefan Nickel, Oliver Stein, und Karl-Heinz Waldmann. *Operations Research*. Springer, Berlin, 2 edition, 2014.
- Paul Niebler und Dominic Lindner. *Datenbasiert entscheiden: Ein Leitfaden für Unternehmer und Entscheider*. Springer Fachmedien, Wiesbaden, 2019.
- Klaus North. *Wissensorientierte Unternehmensführung: Wissensmanagement gestalten*. Springer-Verlag, 2016.
- Abdalla M Odeh und Hussien T Battaineh. Causes of construction delay: traditional contracts. *International Journal of Project Management*, 20(1):67–73, 2002.

- H. Odeyinka, K. Larkin, R. Weatherup, G. Cunningham, M. McKane, und G. Bogle. Modelling risk impacts on the variability between contract sum and final account. *Royal Institution of Chartered Surveyors, London*, Seiten 1–19, 2012.
- Mayor Office of Operations. Capital projects (Zugriff am: 01.05.2020), 2018. URL <https://data.cityofnewyork.us/City-Government/Capital-Projects/n7gv-k5yt>.
- Stephen O Ogunlana, Krit Promkuntong, und Vithool Jearkijrm. Construction delays in a fast-growing economy: Comparing thailand with other economies. *International Journal of Project Management*, 14(1):37–45, 1996.
- D.R. Ogunsemi und G.O. Jagboro. Time-cost model for building projects in nigeria. *Construction management and economics*, 24.3:253–258, 2006.
- Lois Oksenberg und Graham Kalton. New strategies for pretesting survey questions. *Journal of official statistics*, 7(3):349, 1991.
- Svenja Oprach und Maximilian Budau. Zeitliche Einflussfaktoren auf die Fertigstellung von Bauprojekten - Ein Status Quo aktueller Forschungsschwerpunkte. 30. BBB-Assistententreffen in Karlsruhe-Fachkongress der wissenschaftlichen Mitarbeiter Bauwirtschaft\ Baubetriebe\ Bauverfahrenstechnik: 10.-12. Juli 2019, Institut für Technologie und Management im Baubetrieb (TMB), Karlsruher Institut für Technologie (KIT). Hrsg.: S. Haghsheno, K. Lennerts, S. Gentes, 2019.
- Svenja Oprach, Tobias Bolduan, Dominik Steuer, Michael Vössing, und Shervin Haghsheno. Building the future of the construction industry through artificial intelligence. *Digitale Welt, Sonderausgabe Artificial Intelligence*, 2019a.
- Svenja Oprach, Dominik Steuer, Viktoria Krichbaum, und Shervin Haghsheno. Smart data - dealing with task complexity in construction scheduling. In *Proc. 27th Annual Conference of the International Group for Lean Construction (IGLC)*, Seiten 347–358, Dublin, Ireland, 2019b.
- Hubert Österle und Robert Winter. Business engineering. In *Business engineering*, Seiten 3–20. Springer, 2000.

- Cyril Northcote Parkinson und Osbert Lancaster. *Parkinson's Law or the Pursuit of Progress*. Murray London, 1958.
- S. Petrusseva, V. Zujo, und V. Zileska Pancovska. Neural network prediction model for construction project duration. *International Journal of Engineering Research and Technology*, 1, 01 2012.
- W. Pewdum, T. Rujiranyong, und V. Sooksatra. Forecasting final budget and duration of highway construction projects. *Engineering, Construction and Architectural Management*, 16:544–557, 11 2009.
- Jeffrey Pfeffer und Robert I. Sutton. Evidence-based management. *Harvard Business Review*, 84(1):62–74, January 2006.
- Jeffrey K Pinto und Peter WG Morris. *The Wiley guide to managing projects*. John Wiley & Sons New Jersey, 2004.
- PMBOK. *A Guide to the Project Management Body of Knowledge (PMBOK(R) Guide-Sixth Edition / Agile Practice Guide Bundle (GERMAN) -*. Project Management Institute, Newton Square, 2018.
- K Potts. The new scottish parliament building a critical examination of the lessons to be learned. In *Third International Conference on Construction in the 21st Century (CITC-III) Advancing Engineering, Management and Technology*, Seiten 15–17, 2005.
- Lutz Prechelt. Early Stopping — But When? In Grégoire Montavon, Geneviève B. Orr, und Klaus-Robert Müller, editors, *Neural Networks: Tricks of the Trade: Second Edition*, Seiten 53–67. Springer Berlin Heidelberg, 2012.
- Gilbert Probst, Steffen Raub, und Kai Romhardt. *Wissen managen: Wie Unternehmen ihre wertvollste Ressource optimal nutzen*, volume 6., überarbeitete und erweiterte Auflage. Gabler-Verlag / GWV Fachverlage GmbH Wiesbaden, 2010.
- L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush, und A. Gulin. Catboost: unbiased boosting with categorical features. In *Advances in neural information processing systems*, Seiten 6638–6648, 2018.

- Novi Quadrianto und Wray L. Buntine. Regression. In Claude Sammut und Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning*, Seiten 838–842. Springer US, 2010.
- Mohammad Hossein Rafiei und Hojjat Adeli. A Novel Machine Learning Model for Estimation of Sale Prices of Real Estate Units. *Journal of Construction Engineering and Management*, 142(2):1–10, 2016.
- Sebastian Raschka. *Python Machine Learning: Unlock Deeper Insights into Machine Learning with This Vital Guide to Cutting-Edge Predictive Analytics*. Community Experience Distilled. Packt Publishing open source, 2016.
- Sebastian Raschka und Vahid Mirjalili. *Machine Learning mit Python und Scikit-Learn und TensorFlow: Das umfassende Praxis-Handbuch für Data Science, Predictive Analytics und Deep Learning*. MITP-Verlags GmbH & Co. KG, 2017.
- Gopinath Rebala, Ajay Ravi, und Sanjay Churiwala. *An Introduction to Machine Learning*. Springer International Publishing, 2019.
- Nicholas Rescher. *Wissenschaftlicher Fortschritt: Eine Studie über die Ökonomie der Forschung*. Walter de Gruyter GmbH & Co KG, 1982.
- RIBA. *RIBA Plan of Work (Zugriff am: 03.07.2019)*. RIBA Publications, 2013. URL <https://www.ribaplanofwork.com/Download.aspx>.
- Pieter M.A. Ribbers und Klaus-Clemens Schoo. Program management and complexity of erp implementations. *Engineering management journal*, 14(2):45–52, 2002.
- Francisco Loforte Ribeiro und Vanessa Leitão Tomásio Ferreira. Using knowledge to improve preparation of construction projects. *Business Process Management Journal*, Seiten 361–376, 2010.
- M.T. Ribeiro, S. Singh, und C. Guestrin. “why should i trust you?”: Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD

- '16, Seite 1135–1144. Association for Computing Machinery, New York, NY, USA, 2016.
- Dirk Rogel und Rainard Osebold. *Skizzierung fachspezifischer Unsicherheiten im Bauwesen als Entwicklungsansatz zur Steigerung der Zeiteffizienz in der baubetrieblichen Terminplanung*, Seiten 209–244. Springer Fachmedien Wiesbaden, Wiesbaden, 2013.
- Lior Rokach und Oded Maimon. *Decision Trees*. Springer US, Boston, MA, 2005.
- John T Roscoe. *Fundamental research statistics for the behavioral sciences*. Holt Rinehart & Winston, New York, 1975.
- Michael M. Roy, Nicholas JS Christenfeld, und Craig RM. MCKENZIE. Underestimating the duration of future events: Memory incorrectly used or memory bias? *Psychological bulletin*, 131(5):738–756, 2005.
- Lee Peter Ruddin. You can generalize stupid! social scientists, bent flyvbjerg, and case study methodology. *Qualitative Inquiry*, 12(4):797–812, 2006.
- R.J. Rummel. *Applied factor analysis*, Evanston, IL: Northwestern University, 1970.
- Stuart J. Russell und Peter Norvig. *Künstliche Intelligenz: Ein moderner Ansatz*. Pearson, München, 3., aktualisierte Aufl. edition, 2012.
- P.D. Rwelamila und K.A. Hall. Total systems intervention: an integrated approach to time, cost and quality management. *Construction Management and Economics*, 13(3):235–241, 1995.
- Murali Sambasivan und Yau Wen Soon. Causes and effects of delays in Malaysian construction industry. *International Journal of Project Management*, 25(5): 517–526, 2007.
- Claude Sammut und Geoffrey I. Webb. Mean Absolute Error. In *Encyclopedia of Machine Learning*, Seiten 652–652. Springer US, 2010.

- A. L. Samuel. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3):210–229, 1959.
- Robert E Schapire, Yoav Freund, Peter Bartlett, Wee Sun Lee, et al. Boosting the margin: A new explanation for the effectiveness of voting methods. *The annals of statistics*, 26(5):1651–1686, 1998.
- Ewald Scherm, Christian Julmi, und Florian Lindner. Intuitive versus analytische entscheidungen Überlegungen zur situativen stimmigkeit. In *Nachhaltiges entscheiden*, Seiten 299–318. Springer, 2016.
- Hartmut Schiefer und Felix Schiefer. *Statistik für Ingenieure: eine Einführung mit Beispielen aus der Praxis*. Springer Vieweg, Wiesbaden, 2018.
- Jürgen Schmidhuber. Deep Learning. In Claude Sammut und Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning and Data Mining*, Seiten 338–348. Springer US, 2017.
- Manfred Schulte-Zurhausen. *Organisation*. Vahlen, 1995.
- Stefan Seyffert. *Optimierungspotenziale im Lebenszyklus eines Gebäudes: Entwicklung und Nachweis eines Modells zur Anwendung der Radio-Frequenz-Identifikation im Bauwesen*. Vieweg+Teubner Verlag, 2011. Aufl. edition, 2011.
- L.S.A. Shapely. Value for n—pemon game, 1953.
- R. Sheh. Why Did You Do That? Explainable Intelligent Robots. In *Workshops at the Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- Weiming Shen, Qi Hao, Helium Mak, Joseph Neelamkavil, Helen Xie, John Dickinson, Russ Thomas, Ajit Pardasani, und Henry Xue. Systems integration and collaboration in architecture, engineering, construction, and facilities management: A review. *Advanced engineering informatics*, 24(2):196–207, 2010.
- James A. Shepperd. Taking stock of unrealistic optimism. *Perspectives on Psychological Science*, 8(4):395–411, 2013.

- Vyas Shubham. Causes of delay in project construction in developing countries. *Indian Journal of Commerce & Management Studies* ISSN, 24-29:0310, 2013.
- SIA. *Ordnung SIA 103 Ordnung für Leistungen und Honorare der Bauingenieurinnen und Bauingenieure*. Schweizer Ingenieur- und Architektenverein (SIA), 2014.
- Matti Siemiatycki. What's the secret? confidentiality in planning infrastructure using public/private partnerships. *Journal of the American Planning Association*, 73(4):388–403, 2007.
- Matti Siemiatycki. Academics and auditors: Comparing perspectives on transportation project cost overruns. *Journal of Planning Education and Research*, 29(2):142–156, 2009.
- Karlhans Stark. *Baubetriebslehre*. Springer, 2006.
- Auditor State North Carolina. Performance audit: Department of transportation highway project schedules and costs, 2008.
- Eugene F Stone. *Research methods in organizational behavior*. Goodyear Publishing Company, 1978.
- Ron Sun und C Lee Giles. *Sequence learning: Paradigms, algorithms, and applications*. Springer, 2003.
- Ghaleb Sweis, Rateb Sweis, Ayman Abu Hammad, und A. Shboul. Delays in construction projects: The case of Jordan. *International Journal of Project Management*, 26:665–674, 08 2008.
- Eric Thomas und Jay Bowman. Harnessing the data advantage in construction (Zugriff am: 07.09.2022), 2021. URL [https://a.storyblok.com/f/64835/x/137c281eda/harnessing\\_the\\_data\\_advantage\\_in\\_construction.pdf](https://a.storyblok.com/f/64835/x/137c281eda/harnessing_the_data_advantage_in_construction.pdf).
- Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

- Jasper van der Waa, Elisabeth Nieuwburg, Anita Cremers, und Mark Neerinx. Evaluating xai: A comparison of rule-based and example-based explanations. *Artificial Intelligence*, 291:1–52, 2021.
- Jacob T. VanderPlas. *Python Data Science Handbook: Essential Tools for Working with Data*. O’Reilly Media, Inc, 1 edition, 2016.
- Verein Deutscher Ingenieure e.V. 2870:Ganzheitliche Produktionssysteme; Blatt 1: Grundlagen, Einführung und Bewertung, 2012.
- Verein Deutscher Ingenieure e.V. 3633:Simulation von Logistik-, Materialfluss- und Produktionssystemen - Grundlagen, 2014.
- Ludovic-Alexandre Vidal und Franck Marle. Understanding project complexity: implications on project management. *Kybernetes*, 2008.
- P. Voigt und A. Von dem Bussche. The EU general data protection regulation (GDPR). *A Practical Guide, 1st Ed.*, Cham: Springer International Publishing, 2017.
- Olivier Bousquet Ulrike von Luxburg und Gunnar Rätsch. *Advanced lectures on machine learning*. Lecture, 2004.
- Derek H. T. Walker. An investigation into construction time performance. *Construction Management and Economics*, 13(3):263–274, 1995.
- C.Y. Wang, Y.R. Yu und H.H. Chan. Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models. *International Journal of Project Management*, 30(4): 470–478, 2012.
- Geoffrey I. Webb. Data preparation. In Claude Sammut und Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning*, Seiten 259–260. Springer US, 2010a.
- Geoffrey I. Webb. Overfitting. In Claude Sammut und Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning*, Seiten 744–744. Springer US, 2010b.



- Neil D. Weinstein. Unrealistic optimism about future life events. *Journal of personality and social psychology*, 39(5):806–820, 1980.
- Terry Williams. Assessing and moving on from the dominant project management discourse in the light of project overruns. *IEEE Transactions on engineering management*, 52(4):497–508, 2005.
- Terry M Williams. The need for new paradigms for complex projects. *International journal of project management*, 17(5):269–273, 1999.
- Ricky Wai-kin Wu und Fabian C Hadipriono. Fuzzy modus ponens deduction technique for construction scheduling. *Journal of construction engineering and management*, 120(1):162–179, 1994.
- Falk Würfele, Bert Bielefeld, und Mike Gralla. *Terminplanung*, Seiten 25–65. Springer Fachmedien Wiesbaden, Wiesbaden, 2017.
- Weidong Xia und Gwanhoo Lee. Complexity of information systems development projects: conceptualization and measurement development. *Journal of management information systems*, 22(1):45–83, 2005.
- Jyh-Bin Yang und Pei-Rei Wei. Causes of delay in the planning and design phases for construction projects. *Journal of Architectural Engineering*, 16(2): 80–83, 2010.
- Lotfi A. Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.
- Yanshuai Zhang und S. Ng. An ant colony system based decision support system for construction time-cost optimization. *Journal of Civil Engineering and Management*, 18:580–589, 08 2012.
- Yongli Zhang und Sergio Koreisha. Adaptive order determination for constructing time series forecasting models. *Communications in Statistics - Theory and Methods*, 44(22):4826–4847, 2015.
- Daisy Zheng, S. Ng, und Mohan Kumaraswamy. Applying a genetic algorithm-based multiobjective approach for time-cost optimization. *Journal of Construction Engineering and Management*, 130:168–176, 04 2004.

Xiaojin Zhu. Semi-Supervised Learning. In Claude Sammut und Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning*, Seiten 892–897. Springer US, 2010.

Tarila Zuofa, Edward Ochieng, und Alan Burns. Appraising knowledge management perceptions among construction practitioners. *Proceedings of the Institution of Civil Engineers-Management, Procurement and Law*, 168:89–98, 2015.

# **A Anhang zu Kapitel 3**

**Tabelle A.1:** Literaturübersicht der gefilterten Veröffentlichungen (1/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

<b>Titel der Veröffentlichung</b>	<b>Journal</b>	<b>Jahr</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
'Intelligent' construction time and cost analysis	Construction Management and Economics	1986	0	1	0	0
"Important Causes of Delay in Public Utility Projects in Saudi Arabia	Construction Management and Economics	1999	1	0	0	0
20. Study of the Relationship between Procurement Duration and Project Performance in Design-Build Projects: Comparison between Water/Wastewater and Transportation Sectors	Journal of Management in Engineering	2013	1	0	0	0
A Benchmark Model for Construction Duration in Public Housing Developments	Construction Management and Economics	2003	0	1	0	0
A comparative study of causes of time overruns in Hong Kong construction project	International Journal of Project Management	1997	1	0	0	0
A computational intell. fuzzy model approach for excavator cycle time simulation	Automation in Construction	2003	0	1	0	0
A computer-based cost prediction model for institutional building projects in Nigeria: An artificial neural network approach	Journal of Engineering Design and Technology	2014	1	0	1	0
A construction time information system for the building industry	Ph.D. thesis	1991	0	1	0	0
A decision support framework for estimating project duration under the impact of weather	Automation in Construction	2018	0	1	0	0
A decision support model for construction cash flow management	Computer-Aided Civil and Infrastructure Engineering	2007	0	1	0	0
A formwork method selection model based on boosted decision trees in tall building construction	Automation in Construction	2012	0	0	1	0
A framework for estimating negotiation amounts in construction projects	Construction Innovation	2009	0	1	0	0
A Method for Highway Agency to Estimate Highway Construction Durations and Set Contract Times	Construction Management and Economics	2007	0	1	0	0
A novel multiple linear regression model for forecasting S-curves	Engineering, Construction and Architectural Management	2006	0	1	0	0
A Petri net simulation model for virtual construction of earthmoving operations	Automation in Construction	2011	1	0	0	0
A Present Value Approach to Determine Optimum Project Duration	The Engineering Economist	1992	0	0	0	1
A Process for the Estimation of the Duration of Activities in Fuzzy Project Scheduling	Vulnerability, Uncertainty, and Risk: Analysis, Modeling, and Management	2011	0	1	0	0
A risk management approach to address construction delays from the client aspect	Civil Building Engineering	2014	1	1	0	0
A simplified approach to the planning and control of cost and project duration	Construction Management and Economics	1985	0	0	0	1
A Study of Contribution Factors to Building Failures and Defects in Construction Industry	Procedia Engineering	2011	1	0	0	0
A study of the factors affecting construction durations in Hong Kong	Construction Management and Economics	1995	0	1	0	0
A Survey of Factors Influencing the Productivity of Construction Operatives in the State of Qatar	Construction Management and Economics	2012	1	0	0	0
A survey of site management practice on some selected sites in Nigeria	Building and Environment	1989	1	0	0	0
A survey of time-cost relationships in Hong Kong construction projects	Building Technology and Management Journal	1993	1	0	0	0
A time-series analysis of UK annual and quarterly construction output data (1955-95)	Construction Management and Economics	1998	0	1	0	0
Achieving safety and productivity in construction projects	Journal of Civil Engineering and Management	2015	1	0	0	0
Adaptive Forecasting in Earthmoving Operation Using DES and Site Captured Data	Procedia Engineering	2014	0	1	0	0
Adaptive Order Determination for Constructing Time Series Forecasting Models	Communications in Statistics - Theory and Methods	2015	0	1	0	0
Addressing Construction Delays in the Kingdom of Saudi Arabia	Centre for Construction Industry Studies			1	0	0
Aggregation of factors causing cost overruns and time delays in large public construction projects Trends and implications	ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT	2013	1	0	0	0
An Investigation into the Delays in Road Projects in Bahrain	Int. Journal of Research in Engineering & Science	2014	1	0	0	0
An Investigation into Causes and Effects of Construction Delays in UAE	Proc 26th Annual ARCOM Conference, 6-8 September 2010. Association of Researchers in Construction Management, Leeds, UK.	2012	1	0	0	0
An ant colony system based decision support system for construction timecost optimization	Journal of Civil Engineering and Management	2012	0	0	0	1
An efficient approximation for project composite learning curves	Project Management Journal	1995	0	1	0	0
An Empirical Study of Schedule Delay Causes Based on Taiwan's Litigation Cases	PROJECT MANAGEMENT JOURNAL	2010	1	0	0	0
An evaluation of construction time performance in the building industry	Building and Environment	1996	1	0	0	0
An integrated regression analysis and time series model for construction tender price index forecasting	Construction Management and Economics	2004	0	1	0	0

**Tabelle A.2:** Literaturübersicht der gefilterten Veröffentlichungen (2/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

<b>Titel der Veröffentlichung</b>	<b>Journal</b>	<b>Jahr</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
An integrated simulation model for construction	Simulation Practice and Theory	1996	0	1	0	0
An intelligent simulation approach to cost-time forecasting for housing modernization works	Computing Systems in Engineering	1995	0	1	0	0
An interactive risk visualisation tool for large-scale and complex engineering and construction projects under uncertainty and interdependence	International Journal of Production Research	2018	0	1	0	0
An investigation into construction time performance	Construction Management and Economics	1995	1	0	0	0
An investigation into factors that determine building construction time performance	Ph.D. thesis	1994	1	0	0	0
An investigation into the development of a benchmark for U.K. construction time performance	Master thesis	1995	1	0	0	0
Analysing delay causes and effects in Ghanaian state housing construction projects	INTERNATIONAL JOURNAL OF MANAGING PROJECTS IN BUSINESS	2018	1	0	0	0
Analysing factors affecting delays in Indian construction projects	International Journal of Project Management	2012	1	0	0	0
Analysis of causes of delay and time performance in construction projects	Journal Construction Engineering Management	2014	1	0	0	0
Analysis of characteristics affecting completion time for Malaysian construction projects	Built Environment Project and Asset Management	2015	1	0	0	0
Analysis of Design-Build Procurement Activities Durations for Highway Projects	Building a Sustainable Future	2009	1	0	0	0
Analysis of factors critical to construction project success in Malaysia	Engineering, Construction and Architectural Management	2012	1	0	0	0
ANALYSIS OF PROJECT SUCCESS FACTORS IN CONSTRUCTION INDUSTRY	TECHNOLOGICAL AND ECONOMIC DEVELOPMENT OF ECONOMY	2014	1	0	0	0
Analysis of schedule deviations in road construction projects and the effects of project physical characteristics	Journal of Financial Management of Property and Construction	2016	1	0	0	0
Analyzing delay causes in Egyptian construction projects	Journal Advanced Research	2014	1	0	0	0
Analyzing Delays of Road Construction Projects in Cambodia: Causes and Effects	JOURNAL OF MANAGEMENT IN ENGINEERING	2014	1	0	0	0
Application of artificial neural network to forecast construction duration of buildings at the predesign stage	Engineering Construction and Architectural Management	1999	0	1	1	0
Application of artificial neural network to forecast construction duration of buildings at the predesign stage	ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT	2015	0	0	1	0
Application of discrete-event simulation to study the paving operation of asphalt concrete	Construction Innovation	2008	1	0	0	0
APPLICATION OF IMPORTANCE-SATISFACTION ANALYSIS AND INFLUENCE-RELATIONS MAP TO EVALUATE DESIGN DELAY FACTORS	JOURNAL OF CIVIL ENGINEERING AND MANAGEMENT	2013	1	0	0	0
Application of Time Buffers to Construction Project Task Durations	Journal of Construction Engineering and Management	2013	1	0	0	0
Applying a genetic algorithm-based multiobjective approach for time-cost optimization	Journal Construction Management	2004	0	0	0	1
Approximately predicting the cost and duration of school reconstruction projects in Taiwan	Construction Management and Economics	2008	0	1	1	0
Assessing Understanding of Planning and Scheduling Theory and Practice on Construction Projects	Engineering Management Journal	2015	1	0	0	0
Attributes of UK construction clients influencing project consultants' performance	Construction Management and Economics	1994	1	0	0	0
Australian project time-cost analysis: Statistical analysis of intertemporal trends	Construction Management and Economics	2011	0	1	0	0
Automated estimation of completion time for dam projects	International Journal of Construction Management	2017	0	0	1	0
Automated tool for predicting duration of construction activities in tropical countries	KSCE JOURNAL OF CIVIL ENGINEERING	2018	0	0	1	0
Benchmarking of project planning and success in selected industries	Benchmarking: An International Journal	2006	0	1	0	0
Bestimmung der Bauzeit von Bauprojekten zum Zeitpunkt der Realisierungsentscheidung	Ph.D. thesis	2017	0	1	0	0
BIM's Impact on the Success Measures of Construction Projects	Computing in Civil Engineering	2009	1	0	0	0
Cash flow forecasting and updating for building projects	Project Management Journal	1996	0	1	0	0
Cash flow modeling for construction projects	Engineering Construction and Architectural Management	2014	0	1	0	0

**Tabelle A.3:** Literaturübersicht der gefilterten Veröffentlichungen (3/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
Cause and effect 3-D model for measuring performance in construction acceleration: A decision support system: Preliminary structure of decision support system to help managers decide whether and how to accelerate progress in construction projects when delays occur	Building Research & Information	1996	0	0	0	1
Cause Of Delay In Construction Industry In Libya	The International Conference on Administration and Business	2009	1	0	0	0
Cause of delay in construction project in Kuwait	Engineering Journal of the University of Qatar	2002	1	0	0	0
Causes for Delays in Construction Projects in Kuwait	Engineering Journal of the University of Qatar	2001	1	0	0	0
Causes of Delay in Completion of Construction Projects in Oman	International Conference on Innovations in Engineering and Technology, Bangkok, Thailand	2019	1	0	0	0
Causes and Effects of Delay in Iranian Construction Projects	International Journal of Engineering and Technology	2012	1	0	0	0
Causes and effects of delays in Malaysian construction industry	International Journal of Project Management	2007	1	0	0	0
Causes of construction delay: traditional contracts	International Journal of Project Management	2002	1	0	0	0
Causes of construction delays for engineering projects: an Egyptian perspective	AEI 2013	2013	1	0	0	0
Causes of Construction Delays in Countries with High Geopolitical Risks	JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT	2010	1	0	0	0
Causes of contractors failures in Saudi Arabia	Master thesis	1993	1	0	0	0
CAUSES OF CONTRACTORS' CLAIMS IN INTERNATIONAL ENGINEERING-PROCUREMENT-CONSTRUCTION PROJECTS	JOURNAL OF CIVIL ENGINEERING AND MANAGEMENT	2014	1	0	0	0
Causes of delay and cost overruns in construction of groundwater projects in a developing countries	International Journal of Project Management	2003	1	0	0	0
Causes of delay and cost overruns in construction of groundwater projects in a developing countries: Ghana as a case study	International Journal of Project Management	2003	1	0	0	0
Causes of delay and cost overruns in Nigeria construction projects	International Journal of Project Management	1994	1	0	0	0
Causes of delay in building construction projects in Egypt	Journal Construction Engineering Management	2008	1	0	0	0
Causes of delay in construction industry in Libya	Proceedings of International Conference on Administration and Business	2009	1	0	0	0
Causes of delay in construction projects of Punjab-Pakistan: an empirical study	Journal Basic Application Science Research	2013	1	0	0	0
Causes of delay in large building construction projects	International Journal of Project Management	2006	1	0	0	0
Causes of Delay in Large Construction Projects	International Journal of Project Management	2016	1	0	0	0
Causes of Delay in MARA Management Procurement Construction Projects	Journal of Surveying, Construction & Property	2010	1	0	0	0
Causes of delay in residential construction projects in Cambodia	Cogent Engineering	2017	1	0	0	0
Causes of delay in road construction projects	Journal Management Engineering	2012	1	0	0	0
Causes of delay in road construction projects in Malawi	JOURNAL OF THE SOUTH AFRICAN INSTITUTION OF CIVIL ENGINEERING	2018	1	0	0	0
Causes of delay in the construction industry	Journal Construction Division	1971	1	0	0	0
Causes of delay in the planning and design phases for construction projects	Journal Architecture Engineering	2010	1	0	0	0
Causes of delays and overruns of construction projects in developing countries	Proceedings of the CIB	2003	1	0	0	0
Causes of delays and overruns of construction projects in developing countries	Proceedings of the Fourth International Symposium on Organization and Management of Construction: Developing Countries	1984	1	0	0	0
Causes of Delays in Construction Projects in Turkey	Journal of Civil Engineering and Management	2012	1	0	0	0
Causes of Delays in Iranian Construction Projects	International Journal of Construction Management	2010	1	0	0	0
Causes of delays in large building construction projects	Journal Management Engineering	1995	1	0	0	0
Causes of delays in Saudi Arabian public sector construction projects	Construction Management and Economics	2009	1	0	0	0
Causes of delays in the construction industry in Lebanon	Engineering, Construction and Architectural Management	1998	1	0	0	0

**Tabelle A.4:** Literaturübersicht der gefilterten Veröffentlichungen (4/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
Causes of delays in the construction phase of Chinese building projects	ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT	2018	1	0	0	0
Causes of high costs of construction in Nigeria	Journal of Construction Engineering and Management	1988	1	0	0	0
Causes of project delay in the construction industry in Afghanistan	ECPPM2012	2012	1	0	0	0
Causes of schedule delays in construction projects in Algeria	International Journal of Construction Management	2018	1	0	0	0
Causes of Time Buffer and Duration Variation in Construction Project Tasks: Comparison of Perception to Reality	Journal of Construction Engineering and Management	2014	1	0	0	0
Causes of Time Waste in Construction Projects in Palestine	ARAB GULF JOURNAL OF SCIENTIFIC RESEARCH	2017	1	0	0	0
Causes of Variation in Construction Project Task Starting Times and Duration	Journal of Construction Engineering and Management	2011	1	0	0	0
Chance-constrained time-cost tradeoff analysis considering funding variability	Journal Construction Management	2005	0	1	0	0
Chaotic harmony search algorithms	Applied Mathematics and Computation	2010	0	0	0	1
Developing a draft schedule using templates and rules	Journal of Construction Engineering and Management	2001	0	1	0	0
Comparison of schedule delay and causal factors between traditional and green construction projects	Technological and Economic Development of Economy	2013	1	0	0	0
Compressing construction durations: lessons learned from Hong Kong building projects	International Journal of Project Management	2002	1	0	0	0
Construction claims and disputes: causes and cost/time overruns	Journal Construction Engineering Management	1994	1	0	0	0
Construction Cost and Duration Uncertainty Model: Application to High-Speed Rail Line Project	Journal of Construction Engineering and Management	2016	0	1	0	0
Construction delay: a quantitative analysis	International Journal of Project Management	2000	1	0	0	0
Construction delay: quantitative analysis	International Journal of Project Management	2000	1	0	0	0
Construction delays and their causative factors in Nigeria	Journal Construction Engineering Management	2006	1	0	0	0
Construction delays in a fast-growing economy: comparing Thailand with other economies	International Journal of Project Management	1996	1	0	0	0
Construction delays in clients opinion – multivariate statistical analysis	The Proceedings of Selected Papers from Creative Construction Conference 2015, Procedia Engineering	2015	1	0	0	0
Construction delays in Florida	www.sweeneylawpa.com	2002	1	0	0	0
Construction delays in Florida: an empirical study	Planning Consultant Department of Community Affairs	2001	1	0	0	0
Construction delays in Hong Kong civil engineering projects	Journal Construction Engineering Management	2006	1	0	0	0
Construction delays: a quantitative analysis	International Journal of Project Management	2000	0	1	0	0
Construction duration of residential building projects in Germany	Engineering Construction and Architectural Management	2007	0	1	0	0
Construction of reinforced concrete multi-storey office buildings: A simulation model for time/cost calculations: First stage computer simulation model for the duration and cost estimation of the construction of a typical multi-storey concrete framed office building providing rapid response to changes in input values and their effect on the duration and cost of the project	Building Research & Information	1995	0	1	0	0
Construction Project Cost and Duration Optimization Using Artificial Neural Network	AEI 2015	2015	0	0	1	0
Construction resource/method factors influencing productivity for high rise concrete construction	Construction Management and Economics	1999	1	0	0	0
CONSTRUCTION SCHEDULE DELAY RISK ASSESSMENT BY USING COMBINED AHP-RII METHODOLOGY FOR AN INTERNATIONAL NPP PROJECT	NUCLEAR ENGINEERING AND TECHNOLOGY	2017	1	0	0	0
Construction Time-cost Trade-off Analysis Using Ant System	International Journal of Construction Education and Research	2007	0	0	0	1
Construction time-influencing factors: the contractor's perspective	Construction Management and Economics	1995	1	0	0	0
Contributors to Schedule Delays in Public Construction Projects in Saudi Arabia: Owners' Perspective	Journal of Construction Project Management	2018	1	0	0	0
Contributors to construction delay	Construction Management and Economics	1998	1	0	0	0
Contributors to construction delays	Construction Management and Economics	1998	1	0	0	0

**Tabelle A.5:** Literaturübersicht der gefilterten Veröffentlichungen (5/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
COPRAS based Comparative Analysis of the European Country Management Capabilities within the Construction Sector in the Time of Crisis	Journal of Business Economics and Management	2011	1	0	0	0
Cost and Schedule Risk Analysis of Bridge Construction in Pakistan: Establishing Risk Guidelines	JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT	2016	1	1	0	0
Cost and time control of construction projects: inhibiting factors and mitigating measures in practice	Construction Management and Economics	2010	1	0	0	0
Cost and time implications of abandoned project resuscitation: A case study of educational institutional buildings in Nigeria	Journal of Financial Management of Property and Construction	2018	1	0	0	0
Cost escalation and schedule delays in road construction projects in Zambia	International Journal of Project Management	2009	1	0	0	0
Critical success factors for different project objectives	Journal of Construction Engineering and Management	1999	1	0	0	0
Generating project plans by reasoning about objects, actions and resources.	Artificial intelligence for engineering design, analysis and manufacturing	1988	0	1	0	0
Decision support for tower crane selection with building information models and genetic algorithms	Automation in Construction	2016	0	0	0	1
Delay and Cost Overruns in Vietnam Large Construction Projects: A Comparison with Other Selected Countries	Journal of Civil Engineering	2008	1	0	0	0
Delay causes analysis in complex construction projects: a Semantic Network Analysis approach	Production Planning & Control	2018	1	0	0	0
Delay causes in Iran gas pipeline projects	International Journal of Project Management	2013	1	0	0	0
Delay causes in Kuwait construction projects	AICSGE7, Proceedings of Seventh Alexandria International Conference on Structural and Geotechnical Engineering	2010	1	0	0	0
Delay Factor Analysis for Hospital Projects in Vietnam	KSCE JOURNAL OF CIVIL ENGINEERING	2016	1	0	0	0
Delays and cost increases in construction of private residential projects in Kuwait	Journal Construction Management	2005	1	0	0	0
Delays in building construction projects in Ghana	Australian Journal Construction Economics	2010	1	0	0	0
Delays in construction projects – the case of Jordan	International Journal of Project Management	2008	1	0	0	0
Delays in Construction: A Brief Study of the Florida Construction Industry	ASC Proceedings of the 39th Annual Conference	2003	1	0	0	0
Delays in project completion: a comparative study of construction delay factors in Saudi Arabia and the United Kingdom	Master thesis	2004	1	0	0	0
Delays in the construction of public utility projects in Saudi Arabia	Master thesis	1995	1	0	0	0
Delays on Large Construction Projects	International Journal of Operations & Production Management	1986	1	0	0	0
Demotivating factors influencing the productivity of civil engineering projects	International Journal of Project Management	2004	1	0	0	0
Determinants of construction duration	Journal of Construction Management and Economics	1995	1	0	0	0
Determinants of construction duration for building projects in Germany	Engineering Construction and Architectural Management	2012	1	0	0	0
Developing a benchmark model for project construction time performance in Hong Kong	Building and Environment	2004	0	1	0	0
Developing a Construction-Duration Model Based on a Historical Dataset for Building Project	Journal of Civil Engineering and Management	2011	0	1	0	0
Development of a client-based risk management methodology for the early design stage of construction processes: Applied to the KSA	Engineering Construction and Architectural Management	2015	0	1	0	0
Discrete time–cost–environment trade-off problem for large-scale construction systems with multiple modes under fuzzy uncertainty and its application to Jinping-II Hydroelectric Project	International Journal of Project Management	2012	0	1	0	0
Duration determination for rural roads using the principal component analysis and artificial neural network	Engineering, Construction and Architectural Management	2015	0	0	1	0



**Tabelle A.6:** Literaturübersicht der gefilterten Veröffentlichungen (6/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

<b>Titel der Veröffentlichung</b>	<b>Journal</b>	<b>Jahr</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
Duration estimation model for bridge construction projects in Ghana	Journal of Engineering Design and Technology	2017	0	1	0	0
Dynamic modeling of labor productivity in construction projects	International Journal of Project Management	2013	0	1	0	0
Using Product Models to Plan Construction	Proceedings 5th International Conference on Computing in Civil and Building Engineering	1993	0	1	0	0
Early Stage Cost Estimation of Buildings Construction Projects using Artificial Neural Networks	Journal of Artificial Intelligence	2011	0	0	1	0
Effect of organizational culture on delay in construction	International Journal of Project Management	2017	1	0	0	0
Effects of resource management regimes on project schedule	International Journal of Project Management	2002	0	1	0	0
Empirical Assessment of the Likelihood and Duration of Highway Project Time Delays	Journal of Construction Engineering and Management	2011	1	0	0	0
Empirical Perspective on Activity Durations for Project-Management Simulation Studies	Journal of Construction Engineering and Management	2015	0	1	0	0
Empirical Study of Factors Influencing Schedule Delays of Public Construction Projects in Burkina Faso	JOURNAL OF MANAGEMENT IN ENGINEERING	2004	1	0	0	0
Empirical Study on the Key Drivers Affecting Durations for Right-of-Way Acquisition on Highway Projects	Journal of Management in Engineering	2013	1	0	0	0
Entry mode selection for international construction markets: the influence of host country related factors	Construction Management and Economics	2008	1	0	0	0
Entwicklung einer agentenbasierten Terminplanoptimierung im Bauwesen	Ph.D. thesis	2019	0	1	0	0
Entwicklung eines Expertensystems fuer die Planung kerntechnischer Rueckbauprojekte	Forschungsbericht KIT	2018	0	1	0	0
Environmental factors and work performance of project managers in the construction industry	International Journal of Project Management	2006	1	0	0	0
Estimating performance time for construction projects	Journal of Management in Engineering	2007	0	1	0	0
Estimating project and activity duration: a risk management approach using network analysis	Construction Management and Economics	1998	0	1	0	0
Estimating project S-curves using polynomial function and neural networks	Journal Construction Management	2009	0	0	1	0
Estimation of Highway Project Duration at the Planning Stage and Analysis of Risk Factors Leading To Time Overrun	T&DI Congress 2014: Planes, Trains, and Automobiles	2014	0	1	0	0
Estimation of project total cost and duration for housing projects in the UK	Building and Environment	1996	0	1	0	0
Evaluating schedule delay causes for private participating public construction works under the Build-Operate-Transfer model	International Journal of Project Management	2013	1	0	0	0
Evaluation of Delay Causes for BOT Projects Based on Perceptions of Different Stakeholders in Turkey	JOURNAL OF MANAGEMENT IN ENGINEERING	2012	1	0	0	0
Evaluation of delays in Indian construction contracts	Journal of the Institution of Engineers (India)	1996	1	0	0	0
Evaluation of factors affecting time and cost performance in Hong Kong building projects	Engineering Construction and Architectural Management	2002	1	0	1	0
Evaluation of the modular method for industrial plant construction projects	International Journal of Construction Management	2014	1	0	0	0
Evaluation von Schedulingproblemen fuer die Projektplanung von Großprojekten am Beispiel des kerntechnischen Rueckbaus	EconPapers	2017	0	1	0	0
Examination of Project Duration, Project Intensity, and Timing of Cost Certainty in Highway Project Delivery Methods	Journal of Management in Engineering	2018	0	1	0	0
Examining project learning, project management competencies, and project efficiency in project-based firms (PBFs)	INTERNATIONAL JOURNAL OF MANAGING PROJECTS IN BUSINESS	2019	1	0	0	0
Exploratory factor analysis of skills requirement for PPP contract governance	Built Environment Project and Asset Management	2018	1	0	0	0

**Tabelle A.7:** Literaturübersicht der gefilterten Veröffentlichungen (7/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
Exploring delay causes of road construction projects in Egypt	Alexandria Engineering Journal	2016	1	0	0	0
Exploring delays in Victoria-based Australian pipeline projects	The Proceedings of the Twelfth East Asia-Pacific Conference on Structural Engineering and Construction – EASEC12, Procedia Engineering	2011	1	0	0	0
Exploring Project Complexity through Project Failure Factors: Analysis of Cluster Patterns Using Self-Organizing Maps	COMPLEXITY	2011	1	0	0	0
Exposure envelopes: an assessment of the exposure to time and cost overruns during construction projects	International Journal of Project Management	1993	1	0	0	0
Factor Analysis for Development of Construction Period Calculation Model in Apartment House Remodeling	Procedia Engineering	2017	1	0	0	0
Factor selection for delay analysis using knowledge discovery in databases	Automation in Construction	2008	1	0	0	0
Factors Affecting Construction Labor Productivity: Iran Case Study	IRANIAN JOURNAL OF SCIENCE AND TECHNOLOGY-TRANSACTIONS OF CIVIL ENGINEERING	2015	1	0	0	0
Factors affecting construction labour productivity for Malaysian residential projects	Structural Survey	2005	1	0	0	0
Factors Affecting Construction Project Outcomes: Case Study of Vietnam	JOURNAL OF PROFESSIONAL ISSUES IN ENGINEERING EDUCATION AND PRACTICE	2013	1	0	0	0
Factors affecting contractor performance	Master thesis	1991	1	0	0	0
Factors Affecting Cost Overruns in Micro-scaled Construction Companies	Procedia Engineering	2014	1	0	0	0
Factors affecting schedule delay, cost overrun, and quality level in public construction projects	Journal Management Engineering	2016	1	0	0	0
Factors Causing Design Schedule Delays in Turkey Projects in Taiwan: An Empirical Study of Power Distribution Substation Projects	PROJECT MANAGEMENT JOURNAL	2017	1	0	0	0
Factors effecting construction contractors' performance	Building Research & Information	1996	1	0	0	0
Factors impacting construction project duration: a comparison between France, Germany and the u.k.	Building and Environment	1998	1	0	0	0
Factors influencing construction time and cost overruns on high-rise projects in Indonesia	Construction Management and Economics	1997	1	0	0	0
Factors influencing development of construction enterprises in Singapore	Construction Management and Economics	2001	1	0	0	0
Factors influencing labour productivity in Bahrain's construction industry	International Journal of Construction Management	2015	1	0	0	0
Factors influencing sub-contractors selection in construction projects	HBRC Journal	2013	1	0	0	0
Factors influencing the construction time of civil engineering projects in Malaysia	Engineering Construction and Architectural Management	2006	1	0	0	0
Factors influencing the effectiveness of construction site managers	International Journal of Project Management	1998	1	0	0	0
Factors influencing the selection of delay analysis methods in construction projects in UAE	International Journal of Construction Management	2018	1	0	0	0
Fischer, M. und Aalami, F. (1996): Scheduling with Computer-Interpretable Construction Method Models.	Journal of Construction Engineering and Management	1996	0	1	0	0
Flexible model for time–cost trade-off problem	Journal Construction Management	2005	0	1	0	0
Forecasting construction durations for public housing projects: a Hong Kong perspective	Building and Environment	1999	0	1	0	0
Forecasting construction output: a comparison of artificial neural network and Box-Jenkins model	Engineering, Construction and Architectural Management	2015	0	1	1	0
Forecasting final budget and duration of highway construction projects	Engineering Construction and Architectural Management	2009	0	0	1	0
Forecasting labor productivity changes in construction using the PMV index	International Journal of Industrial Ergonomics	2005	0	1	0	0
Forecasting project duration as a function of scope changes and delay	Project management journal	2002	0	1	0	0
ISIS: a constraint-directed reasoning approach to job shop scheduling.	Ph.D. thesis	1983	0	1	0	0
Framework for minimising time overruns of Malaysian construction projects	Master thesis	2008	1	0	0	0
Fuzzy Logic Approach to Forecast Project Duration in Construction Projects	Construction Research Congress 2005: Broadening Perspectives	2005	0	1	0	0

**Tabelle A.8:** Literaturübersicht der gefilterten Veröffentlichungen (8/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
Fuzzy Randomness, Uncertainty in Civil Engineering and Computational Mechanics	Springer	2005	0	1	0	0
Green building projects: schedule performance, influential factors and solutions	ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT	2017	1	0	0	0
Expert System for Construction Planning	Journal of Computing in Civil Engineering	1987	0	1	0	0
Housing market dynamics with delays in the construction sector	Journal Housing Economics	2014	0	1	0	0
Identification of coordination factors affecting building projects performance	Alexandria Engineering Journal	2018	1	0	0	0
Identification of the causes of construction delay in Malaysia	International Journal Civil Science Engineering	2012	1	0	0	0
Identifying the critical factors affecting schedule performance of public housing projects	HABITAT INTERNATIONAL	2012	1	0	0	0
Identifying the important causes of delays in building construction projects	Proceedings Ninth East Asia-Pacific Conference on Structural Engineering and Construction	2003	1	0	0	0
Identifying the key process factors affecting project performance	ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT	2013	1	0	0	0
Importance and ranking evaluation of delay factors for development construction projects in Benin	KSCE JOURNAL OF CIVIL ENGINEERING	2018	1	0	0	0
Important Causes of Delay in Public Utility Projects in Saudi Arabia	Construction Management and Economics	1999	1	0	0	0
Important causes of delay in public utility projects in Saudi Arabia	Construction Management and Economics	1999	1	0	0	0
Improving Project Performance through Partnering and Emotional Intelligence	Practice Periodical on Structural Design and Construction	2013	1	0	0	0
Integrated duration-cost influence network for modelling risk impacts on construction tasks	Construction Management and Economics	2006	0	1	0	0
Internal causes of delays in highway construction projects in Thailand	American Journal of Civil Engineering and Architecture	1999	1	0	0	0
Interrelationships of factors causing delays in the relocation of utilities A cognitive mapping approach	ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT	2015	1	0	0	0
Investigating the causes of delay in grain bin construction projects: the case of China	Construction Management and Economics	2019	1	0	0	0
Kostensimulation aus Sicht des Auftraggebers	Springer	2014	0	1	0	0
Large construction projects in developing countries: a case study from Vietnam	International Journal of Project Management	2004	1	0	0	0
Learning curve modelling of gas pipeline construction in Egypt	International Journal of Construction Management	2015	0	1	0	0
Major causes of construction time and cost overruns: A case of selected educational sector projects in Ghana	Journal of Engineering Design and Technology	2017	1	0	0	0
Managing Constructibility Reviews to Reduce Highway Project Durations	Journal of Construction Engineering and Management	2004	1	0	0	0
Measuring the impact of lean construction practices on project duration and variability: A simulation-based study on residential buildings	Journal of Civil Engineering and Management	2017	0	1	0	0
Micro and macro level of dispute causes in residential building projects: Studies of Saudi Arabia	Journal of King Saud University - Engineering Sciences	2016	1	0	0	0
Modeling construction time in Spanish building projects	International Journal of Project Management	2014	0	1	0	0
Modelling and predicting construction durations in Hong Kong public housing	Construction Management and Economics	1999	0	1	0	0
Modelling building construction durations	Journal of Financial Management of Property and Construction	2015	0	1	0	0
Modelling building durations in Hong Kong	Construction Management and Economics	1999	0	1	0	0
Multi-criteria analysis of Projects' performance in construction	Archives of Civil and Mechanical Engineering	2014	1	0	0	0
Multimodelllaufbau unter Einsatz von Bauteilbibliotheken	Springer	2014	0	1	0	0
GHOST (1988): Project Network Generator.	Journal of Computing in Civil Engineering	1988	0	1	0	0
Neural network model for predicting building projects	Conference: proceedings of association of researchers in construction management	1996	0	0	1	0
Neural network prediction model for construction project duration	International Journal of Engineering Research & Technology (IJERT)	2013	0	1	1	0

**Tabelle A.9:** Literaturübersicht der gefilterten Veröffentlichungen (9/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
Neurofuzzy modelling of construction projects' duration I: principles	Engineering, Construction and Architectural Management	2001	0	1	0	0
Neurofuzzy modelling of construction projects' duration II: application	Engineering Construction and Architectural Management	2001	0	1	0	0
Nonlinear time/cost trade-off models in project management	Computing Industry Engineering	1994	0	1	0	0
OPTIMIZATION MODEL FOR CONSTRUCTION PROJECT DURATIONS USING A MULTISTAGE DECISION PROCESS	Engineering Optimization	1998	0	0	0	1
Optimizing Transparency and Disclosure to Reduce Right-of-Way Acquisition Duration for Construction Projects in Mississippi	Journal of Legal Affairs and Dispute Resolution in Engineering and Construction	2013	1	0	0	0
Parametric time overrun estimation of building projects	Journal of Financial Management of Property and Construction	2016	0	1	0	0
Planning the reconstruction of a historical building by using a fuzzy stochastic network	Automation in Construction	2017	0	1	0	0
Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type, and contract type	International Journal of Project Management	2011	0	1	0	0
Predicting contract duration for building construction: is Bromilow's time-cost model a panacea?	Journal of Management in Engineering	2015	0	1	0	0
Predicting Military Construction Project Time Outcomes Using Data Analytics	Engineering Management Journal	2018	0	1	0	0
Predicting project progress via estimation of S-curve's key geometric feature values	Automation in Construction	2015	0	1	0	0
Predicting the Preliminary Engineering Duration of Bridge Projects	Construction Research Congress 2012: Construction Challenges in a Flat World	2012	0	1	0	0
Primary factors influencing bid mark-up size decisions of general contractors in Kuwait	Journal of Financial Management of Property and Construction	2013	1	0	0	0
Proactive-reactive, robust scheduling and capacity planning of deconstruction projects under uncertainty	Ph.D. thesis	2017	0	1	0	0
Probabilistic Forecasting of Project Duration Using Bayesian Inference and the Beta Distribution	Journal of Construction Engineering and Management	2009	0	1	0	0
Problems causing delays in major construction projects in Thailand	Construction Management and Economics	2008	1	0	0	0
Problems of projects and effects of delays in the construction industry of Pakistan	Australian Journal Business Management	2011	1	0	0	0
Procurement and project performance D a comparison of management contracting and traditional contracting	Occasional paper	1991	1	0	0	0
Procurement scheduling in engineer procure construct projects: a comparison of three fuzzy modelling approaches	International Journal of Construction Management	2017	0	1	0	0
Profiling causative factors leading to construction project delays in the United Arab Emirates	Engineering Construction and Architectural Management	2017	1	0	0	0
PROJECT DELAY: KEY ELECTRICAL CONSTRUCTION FACTORS IN HONG KONG	JOURNAL OF CIVIL ENGINEERING AND MANAGEMENT	2018	1	0	0	0
Project delays and cost: stakeholder perceptions of traditional v. PPP procurement	Journal of Financial Management of Property and Construction	2012	1	0	0	0
Project managers' perception of production problems — An Indonesian case study: Lack of material was considered of most importance to productivity problems whilst re-work was equally a problem under kendall concordance analysis	Building Research & Information	2007	1	0	0	0
Project Schedule Forecasting for Skyscrapers	JOURNAL OF MANAGEMENT IN ENGINEERING	2016	0	1	0	0
Project schedule influenced by financial issues: Evidence in construction industry	SCIENTIFIC RESEARCH AND ESSAYS	2017	1	0	0	0

**Tabelle A.10:** Literaturübersicht der gefilterten Veröffentlichungen (10/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
Quantification of delay factors using Relative Importance Index (RII) method for construction projects in Turkey	Journal Management Engineering	2013	1	0	0	0
Quantifying the causes and cost of rework in construction	Construction Management and Economics	2000	1	0	0	0
Ranking of delay factors in construction projects after Egyptian revolution	Alexandria Engineering Journal	2013	1	0	0	0
Reasons for delay in civil engineering projects—the case of Hong Kong	The Hong Kong Institution of Engineers Transactions	1996	1	0	0	0
Reasons for delay in selected hydro-power projects in Khyber Pakhtunkhwa (KPK), Pakistan	RENEWABLE & SUSTAINABLE ENERGY REVIEWS	2014	1	0	0	0
Reasons for delays in public projects in Turkey	Construction Management and Economics	2006	1	0	0	0
Reasons for overrun in public sector construction projects in Nigeria	International Journal of Project Management	1990	1	0	0	0
Reasons for overrun in public sector construction projects in Nigeria.	International Journal of Project Management	1990	1	0	0	0
Recurring themes in value management practice	Engineering Construction and Architectural Management	2002	1	0	0	0
Reducing Variability of Workforce as a Tool to Improve Plan Reliability	Procedia Engineering	2017	0	1	0	0
Regression modelling of risk impacts on construction cost flow forecast	Journal of Financial Management of Property and Construction	2012	1	0	0	0
Relationship between duration and cost of Malaysian construction projects	International Conference in the Built Environment in the 21st Century (ICIBE 2006)	2006	0	1	0	0
Relationship between Organizational Sizes and Contractors' Risk Pricing Behaviors for Weather Risk under Different Project Values and Durations	Journal of Construction Engineering and Management	2008	1	0	0	0
Relationship between value and duration of construction projects	Construction Management and Economics	1991	0	1	0	0
Research of factors influencing construction productivity	Construction Management and Economics	1990	1	0	0	0
Revisiting the construction project failure factors in Vietnam	Built Environment Project and Asset Management	2015	2	0	0	0
Right-of-Way Acquisition Duration Prediction Model for Highway Construction Projects	Journal of Construction Engineering and Management	2011	0	1	0	0
Risk management and contingency sum of construction projects	Journal of Financial Management of Property and Construction	2017	1	0	0	0
Robust multi-mode resource constrained project scheduling of building deconstruction under uncertainty	Ph.D. thesis	2015	0	1	0	0
Role of Communication and Coordination in Project Success: Case Study	JOURNAL OF PERFORMANCE OF CONSTRUCTED FACILITIES	2009	1	0	0	0
Root Causes of Construction Project Delays in Dubai	Association of Researchers in Construction Management	2015	1	0	0	0
Root Causes of Recurring Contractual Conflicts in International Construction Projects: Five Case Studies from Vietnam	JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT	2014	1	0	0	0
Root-Cause Analysis of Construction-Cost Overruns	JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT	2016	1	0	0	0
Safety and Productivity in the Construction Industry	ARAB GULF JOURNAL OF SCIENTIFIC RESEARCH	2017	1	0	0	0
Schedule delays in engineering, procurement, and construction petrochemical projects in Vietnam: A qualitative research study	International Journal of Energy Sector Management	2014	1	0	0	0
Significance ranking of parameters impacting construction labour productivity	Construction Innovation	2012	1	1	1	0
Significant Factors Causing Delay in the UAE Construction Industry	Construction Management and Economics	2006	1	0	0	0
Significant factors causing delay in the UAE construction industry	Construction Management and Economics	2006	1	0	0	0
Simulation modeling and sensitivity analysis of a tunneling construction project's supply chain	Engineering Construction and Architectural Management	2011	1	0	0	0

**Tabelle A.11:** Literaturübersicht der gefilterten Veröffentlichungen (11/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

<b>Titel der Veröffentlichung</b>	<b>Journal</b>	<b>Jahr</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
Simulation von Bauausführungsvarianten mit einem prozessbasierten Simulationsbaukasten	Springer	2014	0	1	0	0
Simulationsgestützte Bauablaufplanung	Springer	2011	0	1	0	0
Statistical Analysis on the Cost and Duration of Public Building Projects	JOURNAL OF MANAGEMENT IN ENGINEERING	2012	1	0	0	0
Stochastic time–cost optimization using non-dominated archiving ant colony approach	Automation in Construction	2011	0	1	0	1
Study on the influence factors of urban subway project delay in the initial construction stage	Journal of Discrete Mathematical Sciences and Cryptography	2018	1	0	0	0
Survey research on quality costs and problems in the construction environment	TOTAL QUALITY MANAGEMENT & BUSINESS EXCELLENCE	2017	1	0	0	0
The accuracy of housing forecasting in Australia	Engineering, Construction and Architectural Management	2003	0	1	0	0
The causes and effects of construction delays on completion cost of housing project in Nigeria	Financial Management Property Construction	1997	1	0	0	0
The causes and effects of construction delays on completion cost of housing projects in Nigeria	Journal of Financial Management of Property and Construction	1997	1	0	0	0
The development of a benchmark model that uses historical data for monitoring the progress of current construction projects	Engineering Construction and Architectural Management	2008	0	1	0	0
The effect of critical success factors on project success in Public-Private Partnership projects: A case study of highway projects in Iran	TRANSPORT POLICY	2019	2	0	0	0
The effect of current public procurement law on duration and cost of construction projects in Turkey	Journal of Civil Engineering and Management	2013	0	0	2	0
The effect of relationship management on project performance in construction	International Journal of Project Management	2017	1	0	0	0
The impact of contractors' management capability on cost and time performance of construction projects in Nigeria	Journal of Financial Management of Property and Construction	2009	1	0	0	0
The impact of interpersonal conflict on construction project performance A moderated mediation study from China	INTERNATIONAL JOURNAL OF CONFLICT MANAGEMENT	2016	1	0	0	0
The Importance of Project Management in the Planning Process of Transport Infrastructure Projects in Germany	Procedia - Social and Behavioral Sciences	2014	1	0	0	0
The influence of a benchmarking system for construction time performance of contracts in South Wales	Master thesis	1996	0	1	0	0
The influence of personal characteristics on effectiveness of construction site managers	Construction Management and Economics	2000	1	0	0	0
The predictive ability of Bromilow's timecost model	Construction Management and Economics	2001	0	1	0	0
The professionals' perspective on the causes of project delay in the construction industry	ENGINEERING CONSTRUCTION AND ARCHITECTURAL MANAGEMENT	2017	1	0	0	0
The relationship between project performance of emerging contractors in government infrastructure projects and their experience and technical qualifications	JOURNAL OF THE SOUTH AFRICAN INSTITUTION OF CIVIL ENGINEERING	2015	1	0	0	0
The role of human error in construction defects	Structural Survey	1999	1	0	0	0
The role of managerial actions in the cost, time and quality performance of high-rise commercial buildings	Construction Management and Economics	1985	1	0	0	0
The root causes of delays in highway construction	Submitted for Presentation at the 82nd Annual Meeting of the Transportation Research Board	2003	1	0	0	0
The significant factors causing delay of building construction projects in Malaysia	Engineering, Construction and Architectural Management	2007	1	0	0	0
The time-cost trade-off analysis in construction project using computer simulation and interactive procedure	Technological and Economic Development of Economy	2009	0	1	0	0
The top 10 universal delay factors in construction projects	INTERNATIONAL JOURNAL OF MANAGING PROJECTS IN BUSINESS	2018	1	0	0	0

**Tabelle A.12:** Literaturübersicht der gefilterten Veröffentlichungen (12/12) (A: Statistische Analysen, B: Hochrechnung, C: Prädiktive Modelle, D: Optimierungen)

Titel der Veröffentlichung	Journal	Jahr	A	B	C	D
The use of exploratory tunnels as a tool for scheduling and cost estimation	Technological and Economic Development of Economy	2010	0	1	0	0
Time – cost relationship for building projects in Korea	Facilities	2009	0	1	0	0
Time – cost relationship for residential construction in Texas	College Station	2003	0	1	0	0
Time and cost overruns in the UAE construction industry: a critical analysis	Construction Management and Economics	2018	1	0	0	0
Time and Cost Performance of Design-Build Projects	JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT	2018	0	1	0	0
Time-cost relationship for predicting construction duration	Archives of Civil and Mechanical Engineering	2013	0	1	0	0
Time-cost relationship of public sector projects in Malaysia	International Journal of Project Management	2001	1	0	0	0
Time-cost relationships of building construction project in Korea	Facilities	2009	0	1	0	0
Time-overruns factors in Nigerian construction industry	Journal of Construction Engineering and Management	2001	1	0	0	0
Time-cost model for building projects in Nigeria	Construction Management and Economics	2006	0	1	0	0
Time–cost optimization of construction projects with generalized activity constraints	Journal Construction Management	2005	0	1	0	0
Time–cost relationship for predicting construction duration	Archives of Civil and Mechanical Engineering	2013	0	1	0	0
Time–cost relationship of public sector projects in Malaysia	International Journal of Project Management	2001	0	1	0	0
UCI Machine Learning Repository: Residential Building Data Set Data Set	online dataset	2018	0	0	1	0
Uncertainty Forecasting in Civil Engineering	Springer	2015	0	2	0	0
UNDERSTANDING IMPACTS OF TIME AND COST RELATED CONSTRUCTION RISKS ON OPERATIONAL PERFORMANCE OF PPP PROJECTS	INTERNATIONAL JOURNAL OF STRATEGIC PROPERTY MANAGEMENT	2012	1	0	0	0
Understanding the outcomes of megaprojects: a quantitative analysis of very large civilian projects	The Rand Corporation	1988	0	1	0	0
Using neural network to predict performance of design-build projects in Singapore	Building and Environment	2017	0	0	1	0
Using soft precedence relations for reduction of the construction project duration	Technological and Economic Development of Economy	2012	0	1	0	0
Using Weibull analysis for evaluation of cost and schedule performance	Journal Construction Management	2005	0	1	0	0
Validating factors influencing monitoring and evaluation in the Ghanaian construction industry: a Delphi study approach	Construction Management and Economics	2018	0	1	0	0
Vom Bieten zum Bauen – Transparenz im Spannungsfeld von Markt und Machbarkeit	Ernst & Sohn Verlag fuer Architektur und technische Wissenschaften GmbH & Co. KG	2016	1	0	0	0
Weather-delay simulation model based on vertical weather profile for high-rise building construction	Journal Construction Engineering Management	2016	0	0	0	1
Weather-wise: A weather-aware planning tool for improving construction productivity and dealing with claims	Automation in Construction	2017	0	1	0	0





# B Anhang zu Kapitel 4

## B.1 Übersicht der Steuerungsfragen im Interview

**Tabelle B.1:** Steuerungsfragen im Interviewleitfaden

Thema	Memospalte / Steuerungsfragen
1. Einleitung	Frage: Aufzeichnung des Interviews
	Vorstellung des Interviewers, Ziel des Interviews, Forschungsziel
	Definition der Rahmenterminplanung
	Rolle in der Rahmenterminplanung und gesammelte Erfahrungen
	Erfahrung des Produkttyps (Spezialisierung)
	Anzahl Bauprojekte pro Jahr (bzw. Rahmenterminplanung)
	Anzahl Mitarbeiter
	Durchschnittsalter der Mitarbeiter
	Umsatz (2019)
2. Ablauf der Planung	Welche Prozesse, bspw. Meetings, QS, Einzelarbeit, durchläuft die Rahmen- terminplanung? Wann wird welche Personengruppe in die Rahmenterminplanung eingebunden?
3. Angaben (Input)	Wie viele Meilensteine/Phasen verwenden Sie?
	Welche Phasen verwenden Sie? (Aufzählung)
	Wie viele Planungsebenen werden in Ihrem Unternehmen unterschieden?
4. Benötigte Zeit	Wie lange dauert die Rahmenterminplanung insgesamt?
	Können Sie dies in Stunden / Tagen angeben?
	Wie lange dauern die einzelnen Bestandteile im Erstellungsprozess?
5. Qualität	Wie zufrieden sind Sie insgesamt?
	Wie ist die Qualität des Gesamtprojektes: Wird der Zeitplan des Gesamtprojektes eingehalten?
	Wie werden einzelne Projektphasen eingehalten?
	Können eventuelle Abweichungen später aufgeholt werden?

**Tabelle B.2:** Steuerungsfragen im Interviewleitfaden

Thema	Memospalte / Steuerungsfragen
6. Hürden und Probleme	Welche Faktoren beeinflussen die Schätzung der Dauer einer Phase (rechtlich, Bauherrenanforderungen, ausschreibungstechnisch, intern, ...)? Verwenden Sie Software bei der Rahmenterminplanung? Wenn ja, welche?
7. Wissensdatenbank	Für welche Aufgaben verwenden sie die Software genau bei der Rahmenterminplanung? Welche Daten werden bei der Rahmenterminplanung verwendet? Nach welchen Kriterien/Umweltbedingungen werden Projekte dabei unterschieden? Welche Personen (-gruppen) sind an der Rahmenterminplanung beteiligt?
8. Fähigkeiten und Erfahrungen	Welche Aufgaben haben die beschriebenen Personengruppen? Welche Entscheidungskompetenz haben die einzelnen Personengruppen? Inwieweit sind die einzelnen Personengruppen verantwortlich für den Projekterfolg/-misserfolg?
9. Anreize und Motivation	Gibt es Anreizsysteme, die Rahmenterminplanung richtig zu gestalten (z.B. monetär)? Inwieweit werden die Personengruppen verantwortlich gemacht für die Richtigkeit der Rahmenterminplanung (geplante vs. tatsächliche Dauer)?
10. Steuerungsprozess	Wie weit vor dem Projektstart wird mit der Rahmenterminplanung begonnen? Welche Prozesse, bspw. Meetings, QS, Einzelarbeit, durchläuft die Steuerung der Rahmenterminplanung? Wann wird welche Personengruppe in die Steuerung der Rahmenterminplanung eingebunden?
11. Sonstiges	Werden im Projektverlauf oder/und nach Projektabschluss die tatsächlichen Phasendauern/Projektdauern dokumentiert? Wenn ja, in welcher Form und wo? Werden die Ergebnisse vorheriger Projekte in die Rahmenterminplanung neuer Projekte einbezogen? Wenn ja, in welcher Form und wann? Weitere Anmerkungen? Haben Sie wahrheitsgemäß geantwortet?

## B.2 Online-Umfrage

## Befragung zur Rahmenterminplanung von Bauprojekten

### Herzlich willkommen

Sehr geehrte Damen und Herren,

das Institut für Technologie und Management im Baubetrieb des Karlsruher Instituts für Technologie (KIT) möchte Sie herzlich zur Teilnahme an dieser Umfrage einladen. Inhalt der Umfrage ist die Thematik **Rahmenterminplanung von Bauprojekten**.

Die Rahmenterminplanung umfasst das gesamte Bauprojekt und setzt zu einem sehr frühen Zeitpunkt wesentliche Projektziele. Als früheste Terminplanung müssen trotz hoher Ungewissheit bezüglich der zukünftigen Entwicklungen viele Faktoren berücksichtigt werden, um einen zuverlässigen Rahmenterminplan zu erstellen.

Wie deutsche Unternehmen diese Herausforderung meistern, soll mit dieser Umfrage untersucht werden.

Bei Interesse an den Ergebnissen der Umfrage senden Sie bitte eine Email an: [sekretariat@tmb.kit.edu](mailto:sekretariat@tmb.kit.edu)

Die Ergebnisse werden Ihnen nach Auswertung zugesendet.

Die Bearbeitung der Umfrage nimmt in etwa **15 Minuten** in Anspruch und ist bis einschließlich **22.05.2020** möglich.

### Vielen Dank für Ihre Unterstützung!

Mit freundlichen Grüßen,

Ihr Forschungsteam  
des Instituts für Technologie und Management im Baubetrieb

Hinweise:

Ihre Angaben werden anonym erhoben.

Aus Gründen der besseren Lesbarkeit wird in der Befragung ausschließlich die männliche Form verwendet. Sämtliche Personenbezeichnungen gelten für beide Geschlechter.

Bei Rückfragen wenden Sie sich bitte an [sekretariat@tmb.kit.edu](mailto:sekretariat@tmb.kit.edu)

### Datenschutzerklärung

Auf den folgenden Seiten werden Ihnen Fragen zu Ihrem Vorgehen und Ihren Ergebnissen bei der Rahmenterminplanung von Bauprojekten gestellt.

Am Ende der Umfrage werden zudem Informationen zu Ihrem Unternehmen und Ihrer Person abgefragt. Die Bearbeitung der Befragung ist **ohne die Nennung von Ihrem Namen und Unternehmen** möglich.

Die Teilnahme an der Umfrage und die Auswertung der Angaben erfolgt anonym. Es können keine Rückschlüsse auf Ihre Person oder Ihr Unternehmen gemacht werden.

1. **Sind Sie mit der oben beschriebenen Datenschutzerklärung einverstanden? \***

- Ja, ich bin einverstanden und möchte teilnehmen.
- Nein, ich bin nicht einverstanden und möchte nicht teilnehmen.

## Allgemeine Fragen

Der Begriff Rahmenterminplanung wird in der Literatur sowie in der Praxis unterschiedlich verwendet. In dieser Umfrage wird von folgender Definition ausgegangen:

**"Die Rahmenterminplanung bezeichnet die früheste Terminplanung und umfasst die wesentlichen Phasen und Meilensteine des gesamten Bauprojektes."**

Bitte berücksichtigen Sie dies bei der Bearbeitung der Fragen.  
Beantworten Sie die Fragen bitte passend zu der Durchführung in Ihrem Unternehmen.

2. **Führen Sie selbst die Rahmenterminplanung von Bauprojekten durch? \***

Bitte wählen Sie aus.

- Ja, ich führe diese selbst durch.
- Nein, aber ich bin daran beteiligt.
- Nein, und ich bin auch nicht daran beteiligt.

3. **Ist das Vorgehen bei der Rahmenterminplanung in Ihrem Unternehmen bei jedem Projekt gleich? \***

Bitte bewerten Sie wie häufig das Vorgehen gleich ist.

- nie    selten    weder noch    häufig    immer    ich weiß es nicht

4. **Bitte bewerten Sie die Bedeutung der Rahmenterminplanung für die Projektdurchführung? \***

Bitte wählen Sie aus.

- gering    eher gering    weder noch    eher groß    groß    ich weiß es nicht

5. **Wie zufrieden sind Sie mit der Rahmenterminplanung in Ihrem Unternehmen? \***

Bitte wählen... ▼

## Fragen zum Prozess

Erinnern Sie sich bitte an das **letzte abgeschlossene Projekt** und beantworten Sie die kommenden Fragen passend zu diesem Projekt.

6. **Bitte wählen Sie alle Disziplinen aus, die bei der Rahmenterminplanung in Ihrem letzten abgeschlossenen Projekt beteiligt waren. \***

Mehrfachauswahl möglich.

- |   |  |
|---|--|
| <input type="checkbox"/> Bauherr            | <input type="checkbox"/> Fachplanung         |
| <input type="checkbox"/> Projektleitung     | <input type="checkbox"/> Objektplanung       |
| <input type="checkbox"/> Projektsteuerung   | <input type="checkbox"/> Genehmigungsbehörde |
| <input type="checkbox"/> Bauleitung         | <input type="checkbox"/> Tragwerksplanung    |
| <input type="checkbox"/> Ausführende Firmen | <input type="checkbox"/> Ich weiß es nicht   |
| <input type="checkbox"/> Andere und zwar    | <input type="text"/>                         |

7. **Wie wurde die Rahmenterminplanung in Ihrem letzten abgeschlossenen Projekt vorwiegend erarbeitet? \***

Bitte wählen Sie aus.

- in Einzelarbeit  
 im Team  
 ich weiß es nicht

8. **Wie viele Phasen hatte der Rahmenterminplan in Ihrem letzten abgeschlossenen Projekt? \***

Bitte wählen Sie die passende Anzahl an Phasen aus. Wenn Sie den genauen Wert nicht wissen, schätzen Sie bitte ab.

- < 3    4 - 9    10 - 19    29 - 29    39 - 40    > 40    ich weiß es nicht

9. **Wie wurde bei Ihrem letzten abgeschlossenen Projekt bestimmt, welche Phasen im Rahmenterminplan aufgezeigt werden? \***

Mehrfachauswahl möglich.

- |  |   |
|--|---|
| <input type="checkbox"/> Orientierung an HOAI Phasen   | <input type="checkbox"/> Automatisiert durch ein Planungstool       |
| <input type="checkbox"/> Auf Basis von einem Projektvergleich mit bereits durchgeführten Projekten | <input type="checkbox"/> Die Phasen sind vom Unternehmen vorgegeben |
| <input type="checkbox"/> Auf Basis von Literaturwerken   | <input type="checkbox"/> Es sind bei jedem Projekt dieselben Phasen |
| <input type="checkbox"/> Auf Basis der persönlichen Erfahrung                                      | <input type="checkbox"/> Ich weiß es nicht                          |
| <input type="checkbox"/> Aus Basis der Erfahrung von zusätzlichen Experten                         |   |
| <input type="checkbox"/> Anders und zwar <input type="text"/>                                      |   |

10. **Wie wurden die Dauern der Phasen bei Ihrem letzten abgeschlossenen Projekt bestimmt? \***

Mehrfachauswahl möglich.

- |  |   |
|--|---|
| <input type="checkbox"/> Berechnung auf Basis von mathematischen Formeln (z. B. für Rohbauzeitwerte) | <input type="checkbox"/> Schätzung auf Basis der persönlichen Erfahrung               |
| <input type="checkbox"/> Berechnung auf Basis von Benchmarks   | <input type="checkbox"/> Schätzung auf Basis der Erfahrung von zusätzlichen Experten  |
| <input type="checkbox"/> Berechnung mit Methoden der künstlichen Intelligenz                         | <input type="checkbox"/> Mit Berücksichtigung der möglichen Risiken im Projektverlauf |
| <input type="checkbox"/> Schätzung auf Basis von Vergleichen mit bereits durchgeführten Projekten    | <input type="checkbox"/> Ich weiß es nicht  |
| <input type="checkbox"/> Schätzung auf Basis von Literaturwerken                                     |   |
| <input type="checkbox"/> Anders und zwar <input type="text"/>  |   |

11. **Bitte wählen Sie alle Aussagen aus, die bei Ihrem letzten abgeschlossenen Projekt zutreffend waren. \***

- Das Ergebnis der Rahmenterminplanung war ein Plan mit fixierten Zeitpunkten.
- Das Ergebnis der Rahmenterminplanung war ein Plan mit flexiblen Zeitspannen.
- Das Ergebnis der Rahmenterminplanung waren mehrere Pläne mit verschiedenen Planungsvarianten.
- Zum Ergebnis der Rahmenterminplanung gehörte eine Analyse verschiedener Risiken.
- Die Rahmenterminplanung wurde mit Erläuterungen dokumentiert.
- Der Rahmenterminplan wurde während der Projektumsetzung fortgeschrieben.

## Fragen zu verfügbaren Informationen

Erinnern Sie sich bitte an das **letzte abgeschlossene Projekt** und beantworten Sie die kommende Frage passend zu diesem Projekt.

### 12. Welche Informationen zum planenden Projekt waren beim Zeitpunkt der Rahmenterminplanung in Ihrem letzten abgeschlossenen Projekt definiert? \*

Bitte wählen Sie alle Informationen aus, die Ihnen bei Ihrem letzten abgeschlossenen Projekt zum Zeitpunkt der Rahmenterminplanung definiert waren.

Wählen Sie keine Daten aus, für welche Sie selbst Annahmen treffen mussten.

Mehrfachauswahl möglich.

- |  |   |  |
|--|---|--|
| <input type="checkbox"/> Projektart (z. B. Bürogebäude, Wohngebäude) | <input type="checkbox"/> Qualitätsanforderungen       | <input type="checkbox"/> Projektbeteiligte       |
| <input type="checkbox"/> Nutzeranforderungen                         | <input type="checkbox"/> Technischer Ausstattungsgrad | <input type="checkbox"/> Logistikkonzept         |
| <input type="checkbox"/> Bauvolumen                                  | <input type="checkbox"/> Besondere Anforderungen      | <input type="checkbox"/> Vorhandene Erschließung |
| <input type="checkbox"/> Bebauungsfläche                             | <input type="checkbox"/> Bauliche Besonderheiten      | <input type="checkbox"/> Standort                |
| <input type="checkbox"/> Gestalt des Baukörpers                      | <input type="checkbox"/> Bausumme                     | <input type="checkbox"/> Baugrundbeschaffenheit  |
| <input type="checkbox"/> Objektplanung                               | <input type="checkbox"/> Bauzeit                      | <input type="checkbox"/> Risikoanalyse           |
| <input type="checkbox"/> Ausbaustandard                              | <input type="checkbox"/> Vergabemodell                | <input type="checkbox"/> Ich weiß es nicht       |
| <input type="checkbox"/> Gebäudedatenmodell                          | <input type="checkbox"/> Projektorganisation          |  |
| <input type="checkbox"/> Andere und zwar                             | <input type="text"/>                                  |  |

### 13. Besitzen Sie dokumentierte Projektdaten von bereits durchgeführten Projekten? \*

Bitte wählen Sie aus.

- ja
- nein
- ich weiß es nicht

## Fragen zu dokumentierten Projektdaten

Auf dieser Seite geht es um dokumentierte Daten von bereits durchgeführten Projekten.

Beantworten Sie die Fragen auf dieser Seite bitte passend zum **Durchschnitt aller dokumentierten Projekte** Ihres Unternehmens.

14. Von wie vielen Projekten besitzen Sie dokumentierte Daten? \*

Bitte wählen Sie aus. Wenn Sie die genaue Anzahl nicht wissen, schätzen Sie bitte ab.

- < 5    5 - 9    10 - 30    30 - 99    > 100    ich weiß es nicht

15. Wie werden die Projektdaten vorgehalten? \*

Beispiele für Datenbanken sind MS Access, MS Excel oder ein BIM Model.  
Bitte wählen Sie aus. Mehrfachauswahl möglich.

- in Papierform in Ordnern  
 digital in persönlichen Projektordnern verteilt  
 digital in unternehmensweiten Projektordnern verteilt  
 digital in einer persönlichen Datenbank  
 digital in einer unternehmensweiten Datenbank  
 ich weiß es nicht

16. Welche Daten von durchgeführten Projekten sind dokumentiert? \*

Bitte wählen Sie aus. Mehrfachauswahl möglich.

- |  |   |   |
|--|---|---|
| <input type="checkbox"/> Projektart (z. B. Büro, Schule, etc.) | <input type="checkbox"/> Technischer Ausstattungsgrad | <input type="checkbox"/> Logistikkonzept                    |
| <input type="checkbox"/> Nutzeranforderungen                   | <input type="checkbox"/> Besondere Anforderungen      | <input type="checkbox"/> Standort                           |
| <input type="checkbox"/> Bauvolumen                            | <input type="checkbox"/> Bauliche Besonderheiten      | <input type="checkbox"/> Baugrundbeschaffenheit             |
| <input type="checkbox"/> Bebauungsfläche                       | <input type="checkbox"/> Bausumme                     | <input type="checkbox"/> Risikoanalyse                      |
| <input type="checkbox"/> Ausbaustandard                        | <input type="checkbox"/> Fortgeschriebene Terminpläne | <input type="checkbox"/> Auftraggeber                       |
| <input type="checkbox"/> Objektplanung                         | <input type="checkbox"/> Vergabemodell                | <input type="checkbox"/> Gewonnene Erkenntnisse             |
| <input type="checkbox"/> Gebäudedatenmodell                    | <input type="checkbox"/> Projektorganisation          | <input type="checkbox"/> Ursachen für Plan-Ist Abweichungen |
| <input type="checkbox"/> Qualitätsanforderungen                | <input type="checkbox"/> Projektbeteiligte            | <input type="checkbox"/> Ich weiß es nicht                  |
| <input type="checkbox"/> Andere und zwar                       | <input type="text"/>                                  |   |



## Fragen zu Ihrer Datenbank

### 17. Können Sie innerhalb der Datenbank die Projekte nach Variablen filtern? \*

Variablen können zum Beispiel die Bausumme oder der Auftraggeber sein.  
Bitte wählen Sie aus.

- ja
- nein
- ich weiß es nicht

## Fragen zur Qualität

Erinnern Sie sich bitte an das **letzte abgeschlossene Projekt** und beantworten Sie die Fragen auf dieser Seite passend dazu.

### 18. Ist bei Ihrem letzten abgeschlossenen Projekt eine Plan-Ist Abweichung aufgetreten? \*

Bitte wählen Sie die Größe der Abweichung aus. Wenn Sie es nicht genau wissen, schätzen Sie bitte ab.  
Die Größe der Abweichung ist in % zu der Dauer nach Plan angegeben.

- keine    1 - 9 %    10 - 24 %    25 - 49 %    > 50 %    ich weiß es nicht

### 19. Wurden bei Ihrem letzten abgeschlossenen Projekt besondere Maßnahmen zur Planeinhaltung eingesetzt? \*

Besondere Maßnahmen können zum Beispiel eine Erhöhung des Budgets oder der Kolonnenstärke sein.  
Bitte wählen Sie aus.

- ja
- nein
- ich weiß es nicht

### 20. Wie groß war der im Rahmenterminplan eingeplante Zeitpuffer bei Ihrem letzten abgeschlossenen Projekt? \*

Die Größe des Puffers ist hier in % von der tatsächlichen Dauer ohne Puffer angegeben.  
Bitte wählen Sie aus.

- 0 %    1 - 9 %    10 - 24 %    25 - 50 %    > 50 %    ich weiß es nicht

## Weitere Fragen zur Qualität

21. Bitte wählen Sie alle Gründe aus, die für die Plan-Ist Abweichung bei Ihrem letzten abgeschlossenen Projekt mitverantwortlich waren. \*

- Fehlende Informationen zum Zeitpunkt der Planung
- Änderungen vom Auftraggeber
- Unrealistische zeitliche Zielanforderungen des Auftraggebers
- Externe unerwartete Einflüsse
- Ungenügende Risikobetrachtung
- Ungenaue Planung

## Weitere Fragen zu Ihrem letzten abgeschlossenen Bauprojekt

Erinnern Sie sich bitte an das **letzte abgeschlossene Projekt** und beantworten Sie die Fragen auf dieser Seite passend dazu.

22. Bitte bewerten Sie, wie neuartig bzw. routiniert das letzte abgeschlossene Projekt war? \*

Neuartige Projekte sind von einem hohen Unikatcharakter der Projektziele geprägt und ähneln sich zu Ihren sonstigen Projekten wenig bis nicht.

Routinierte Projekte sind in ihren Projektzielen ähnlich bis gleich zu anderen Projekten von Ihrem Unternehmen. Bewerten Sie bitte.

- neuartig    eher neuartig    weder noch    eher routiniert    routiniert    ich weiß es nicht

23. Bitte wählen Sie die Projektsumme Ihres letzten abgeschlossenen Projekts aus. \*

24. Wie repräsentativ war das Vorgehen bei der Rahmenterminplanung in Ihrem letzten abgeschlossenen Projekt bezüglich Ihrer sonstigen Vorgehensweise? \*

## Fragen zu Ihrer Person

### 25. Bitte wählen Sie Ihre Funktion aus. \*

Mehrfachauswahl möglich.

- |  |  |
|--|--|
| <input type="checkbox"/> Bauherr         | <input type="checkbox"/> Berater             |
| <input type="checkbox"/> Bauingenieur    | <input type="checkbox"/> Projektentwickler   |
| <input type="checkbox"/> Projektmanager  | <input type="checkbox"/> Generalplaner       |
| <input type="checkbox"/> Projektleiter   | <input type="checkbox"/> Generalunternehmer  |
| <input type="checkbox"/> Projektsteuerer | <input type="checkbox"/> Generalübernehmer   |
| <input type="checkbox"/> Architekt       | <input type="checkbox"/> Real Estate Manager |
| <input type="checkbox"/> Andere und zwar | <input type="text"/>                         |

### 26. Wie lange arbeiten Sie bereits in dieser Funktion? \*

Bitte tippen Sie die Anzahl in Jahren ein.

Jahre in dieser Funktion

### 27. Wie lange arbeiten Sie bereits in der Baubranche? \*

Bitte tippen Sie die Anzahl in Jahren ein.

Jahre in der Baubranche

### 28. Bei wie vielen Projekten waren Sie bereits bei der Rahmenterminplanung beteiligt? \*

Wenn Sie die genaue Anzahl nicht wissen, schätzen Sie bitte ab.

Bitte wählen...

## Fragen zu Ihrem Unternehmen

### 29. Wie viele Bauprojekte hat Ihr Unternehmen insgesamt im letzten Jahr bearbeitet? \*

Bitte wählen Sie aus. Wenn Sie die genaue Anzahl nicht wissen, schätzen Sie bitte ab.

- < 3    3 - 9    10 - 30    30 - 99    > 100    ich weiß es nicht

### 30. In welcher Art von Unternehmen sind Sie tätig? \*

Bitte wählen Sie aus. Mehrfachauswahl möglich.

- |   |  |   |
|---|--|---|
| <input type="checkbox"/> Projektmanagement    | <input type="checkbox"/> Öffentliche Behörde     | <input type="checkbox"/> Projektentwicklung |
| <input type="checkbox"/> Bauunternehmen       | <input type="checkbox"/> Professioneller Bauherr | <input type="checkbox"/> Systemanbieter     |
| <input type="checkbox"/> Architektur          | <input type="checkbox"/> Baumanagement           | <input type="checkbox"/> Ingenieurbüro      |
| <input type="checkbox"/> Beratungsunternehmen | <input type="checkbox"/> Gelegenheitsbauherr     | <input type="checkbox"/> Fachplanung        |
| <input type="checkbox"/> Andere und zwar      | <input type="text"/>                             |   |

### 31. Bitte wählen Sie die Spezialisierung Ihres Unternehmens aus. \*

Mehrfachauswahl möglich.

- |  |   |                                       |
|--|---|---------------------------------------|
| <input type="checkbox"/> Keine Spezialisierung | <input type="checkbox"/> Schlüsselfertigbau | <input type="checkbox"/> Verkehrsbau  |
| <input type="checkbox"/> Tiefbau               | <input type="checkbox"/> Industriebau       | <input type="checkbox"/> Ingenieurbau |
| <input type="checkbox"/> Hochbau               | <input type="checkbox"/> Wohnungsbau        | <input type="checkbox"/> Fertighaus   |
| <input type="checkbox"/> Andere und zwar       | <input type="text"/>                        |                                       |

### 32. Bitte wählen Sie die Anzahl der Angestellten in Ihrem Unternehmen aus. \*

Wenn Sie den genauen Wert nicht wissen, schätzen Sie bitte ab.

Bitte wählen...

**33. Wie hoch ist der Jahresumsatz von Ihrem Unternehmen ? \***

Wenn Sie den genauen Wert nicht wissen, schätzen Sie bitte ab.

**34. Befindet sich der Standort Ihres Unternehmens in Deutschland? \***

Bitte wählen Sie aus.

ja

nein

**Ende der Umfrage**

**35. Hier können Sie gerne Ihre Anmerkungen und Verbesserungsvorschläge zur Rahmenterminplanung von Bauprojekten festhalten.**

**Vielen Dank für Ihre Teilnahme an der Erhebung. Bitte klicken Sie auf "Fertig" um die Umfrage zu beenden.**

Bei Interesse an den Umfrageergebnissen, schreiben Sie bitte an folgende Emailadresse: [sekretariat@tmb.kit.edu](mailto:sekretariat@tmb.kit.edu)

Datenschutzversprechen: Die Auswertung der Angaben in der Umfrage erfolgt anonym. Es können keine Rückschlüsse auf Ihre Person oder Ihr Unternehmen gemacht werden.

Mit freundlichen Grüßen,

Ihr Forschungsteam  
des Instituts für Technologie und Management im Baubetrieb

» **Umleitung auf Schlussseite von Umfrage Online**



## **C Anhang zu Kapitel 6**

### **C.1 Umfrage zu Anforderungen an die Prognosegenauigkeit**

# Erwartungshaltung zur Güte von Bauterminplänen (Umfrage)

## Seite 1

6 Fragen - max. 6 Minuten

Sehr geehrte Damen und Herren,  
im Rahmen meiner Dissertation am Institut für Technologie und Management im Baubetrieb des Karlsruher Instituts für Technologie (KIT) möchte ich Sie herzlich zur Teilnahme an dieser Umfrage einladen. Inhalt der Umfrage ist die Erwartungshaltung zur Güte der Terminplanung von Bauprojekten mit Fokus auf der Bauausführung. Mit Ihrem Feedback tragen Sie dazu bei, die Erwartungshaltung von Auftraggebern und -nehmern zu quantifizieren und zu vergleichen.  
Falls Sie Interesse an den Ergebnissen, weitere Kommentare oder Anregungen haben, senden Sie bitte eine Email an: svenja.oprach@kit.edu  
Die Bearbeitung der Umfrage nimmt in etwa 6 Minuten in Anspruch. Ihre Angaben werden anonym erhoben. Es können keine Rückschlüsse auf Ihre Person oder Ihr Unternehmen gemacht werden.  
Vielen Dank für Ihre Unterstützung!

## Angaben zur Organisation und Ihrer Person

**Bitte wählen Sie die primäre Rolle Ihrer Organisation. \***

- |  |                                       |
|--|---------------------------------------|
| <input type="radio"/> Eigenbedarfsbauherr                      | <input type="radio"/> Bauunternehmen  |
| <input type="radio"/> BOT-Bauherr                              | <input type="radio"/> Planer          |
| <input type="radio"/> Projektträger                            | <input type="radio"/> Berater         |
| <input type="radio"/> Investor                                 | <input type="radio"/> Nachunternehmer |
| <input type="radio"/> Bauträger                                | <input type="radio"/> Behörde         |
| <input type="radio"/> Anderes und zwar... <input type="text"/> |                                       |

**Bitte wählen Sie Spezialisierung Ihrer Organisation (Mehrfachauswahl möglich). \***


- |   |   |
|---|---|
| <input type="checkbox"/> Fertighäuser                             | <input type="checkbox"/> Hochbau            |
| <input type="checkbox"/> Ingenieurbau                             | <input type="checkbox"/> Industriebau       |
| <input type="checkbox"/> Tiefbau                                  | <input type="checkbox"/> Schlüsselfertigbau |
| <input type="checkbox"/> Verkehrsbau                              |   |
| <input type="checkbox"/> Anderes und zwar... <input type="text"/> |   |





Ich bin immer noch mit dem Ergebnis des Bauprojektes zufrieden, wenn die Bauausführung um maximal ... % schneller ist gegenüber der entsprechenden Terminplanung.

0% schneller gegenüber dem Terminplan (bzw. die Bauausführung ist identisch wie im entsprechenden Terminplan geplant)

100% schneller gegenüber dem Terminplan (bzw. die Bauausführung ist doppelt so schnell wie im entsprechenden Terminplan geplant) \* 

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Ablaufplan der Bauausführung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Grobterminplan Bauausführung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Generalablaufplan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rahmenterminplan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

» [Umleitung auf Schlussseite von Umfrage Online \(ändern\)](#)

## C.2 Datensätze aus Teheran, Iran, und San Francisco, USA

Die Tabellen C.1 und C.2 beschreiben mit 29 Features die Projekte in Teheran, Iran. Der Datensatz ist verfügbar unter <https://archive.ics.uci.edu/ml/datasets/Residential+Building+Data+Set>.

**Tabelle C.1:** Datensatz 1: Teheran, Iran (1/2)

<b>Projekt-physische und finanzielle Variablen</b>		
V-1	Projektort anhand von Postleitzahlen	N/A
V-2	Gesamtgeschossfläche des Gebäudes	m <sup>2</sup>
V-3	Grundstücksfläche	m <sup>2</sup>
V-4	Vorläufig geschätzte Gesamtbaukosten auf der Grundlage der Preise zu Beginn des Projekts	10000000 IRR (Währung in Iran (Rial))
V-5	Vorläufige Schätzung der Baukosten auf der Grundlage der Preise, die zu Beginn des Projekts galten	10000 IRR
V-6	Äquivalente vorläufige geschätzte Baukosten auf der Grundlage der Preise zu Beginn des Projekts in einem ausgewählten Basisjahr (z.B. 1383 (iranischer Kalender))	10000 IRR
V-7	Dauer der Bauarbeiten	Zahl in Zeiteinheiten
V-8	Preis der Einheit zu Beginn des Projekts pro m <sup>2</sup>	10000 IRR
V-9	Tatsächliche Verkaufspreise (Output)	10000 IRR
V-10	Tatsächliche Baukosten (Leistung)	10000 IRR
V-11	Die Anzahl der erteilten Baugenehmigungen	N/A
V-12	Index der Gebäudedienstleistungen (BSI) (stellt die Gesamtsumme der Aufträge für Subunternehmer dar, z. B. Löhne für Arbeiter und Rohrinstallationen)	N/A
V-13	Großhandelspreisindex (WPI) (oder Erzeugerpreisindex stellt die Kosten eines Warenkorbs von Lebensmitteln/Dienstleistungen dar, um Inflation/Deflation zu erkennen (U.S. Dept. of Labor 2015)) für Baumaterialien für das Basisjahr	N/A
V-14	Gesamtgeschossflächen der von der Stadt/Gemeinde erteilten Baugenehmigungen	m <sup>2</sup>
V-15	Kumulative Liquidität (zeigt, wie schnell verschiedene Arten von Vermögenswerten in Bargeld umgewandelt werden können)	10000000 IRR
V-16	Investitionen des Privatsektors in neue Gebäude	10000000 IRR
V-17	Bodenpreisindex für das Basisjahr	10000000 IRR
V-18	Die Anzahl der von den Banken in einer Zeitauflösung (d.h. einem Quartal (persischer Kalender)) vergebenen Kredite	N/A

**Tabelle C.2:** Datensatz 1: Teheran, Iran (2/2)

<b>Wirtschaftliche Variablen und Indizes</b>		
V-19	Die Höhe der von den Banken in einer Zeitspanne gewährten Kredite	10000000 IRR
V-20	Der Zinssatz für ein Darlehen in einer zeitlichen Auflösung	%
V-21	Die durchschnittlichen Baukosten von Gebäuden des privaten Sektors zum Zeitpunkt der Fertigstellung des Baus	10000 IRR
V-22	Der Durchschnitt der Baukosten von Gebäuden im privaten Sektor zu Beginn der Bauarbeiten	10000 IRR/m <sup>2</sup>
V-23	Offizieller Wechselkurs gegenüber dem Dollar	IRR
V-24	Inoffizieller (Straßenmarkt-) Wechselkurs gegenüber dem Dollar (wird nur in Ländern mit kontrollierten Währungen verwendet)	IRR
V-25	Verbraucherpreisindex (VPI) im Basisjahr (repräsentiert die Veränderung der Preise eines Warenkorbs von Waren und Dienstleistungen, die von städtischen Haushalten gekauft werden (U.S. Dept. of Labor 2015))	N/A
V-26	VPI für Wohnung, Wasser, Brennstoff	Energie im Basisjahr
V-27	Aktienmarktindex (stellt die Rückzahlungsbedingungen von Investitionen im Aktienmarkt dar)	N/A
V-28	Einwohnerzahl der Stadt	N/A
V-29	Goldpreis pro Unze	IRR

Tabelle C.3 beschreibt mit zehn Features die Projekte in San Francisco, USA. Der Datensatz ist verfügbar unter: <https://data.sfgov.org/Housing-and-Buildings/Building-Permits/i98e-djp9>.

**Tabelle C.3:** Datensatz 2: San Francisco, USA

<b>Projekt-physische und finanzielle Variablen</b>		
V-1	Erstellungsdatum der Genehmigung	Datum & Zeit
V-2	Datum des ersten Baudokuments	Datum & Zeit
V-3	Datum der Fertigstellung	Datum & Zeit
V-4	Straßenname	Text
V-5	Überarbeitete Kosten	Numerisch in USD
V-6	Vorgeschlagene Einheiten	Numerisch
V-7	Vorgeschlagene Stockwerke	Numerisch
V-8	Beschreibung der vorgeschlagenen Bauweise	Text
V-9	Plansätze	Numerisch
V-10	Nachbarschaften	Text

## C.3 Externe Marktbeschreibende Daten

In Tabelle C.4 und C.5 sind 65 Variablen enthalten, mit denen die wirtschaftliche und politische Lage eines Landes beschrieben werden kann. Die Daten sind für das jeweilige Land unter folgenden Links verfügbar: <https://data.oecd.org/> and <https://www.theglobaleconomy.com/>. Es wurden für das entsprechende Land die Variablen für die Dauer der Bauausführung extrahiert.

**Tabelle C.4:** Externe Daten zur Beschreibung der wirtschaftlichen und politischen Lage (1)

Quelle: <a href="https://data.oecd.org/">https://data.oecd.org/</a>		
V-1	Investitionen in Anlagegüter	% der Bruttoanlageinvestitionen (BAI)
V-2	Reale BIP-Prognose	jährliche Wachstumsrate (%)
V-3	Beschäftigung im Baugewerbe	tausend Personen
V-4	Gearbeitete Stunden	Stunden/Arbeiter
V-5	Inflation (CPI)	jährliche Wachstumsrate (%)
V-6	Preise für Häuser	langfristiger Durchschnitt = 100
V-7	Bebautes Gebiet	Quadratmeter pro Kopf
V-8	Verbrauchervertrauensindex (CCI)	langfristiger Durchschnitt = 100 (0-200)
V-9	Geschäftsvertrauensindex (BCI)	langfristiger Durchschnitt = 100 (0-200)
V-10	Bevölkerung	Millionen Menschen
V-11	Erwerbstätige Bevölkerung	Millionen Menschen
V-12	Preisniveauindex	OECD = 100
Quelle: <a href="https://www.theglobaleconomy.com/">https://www.theglobaleconomy.com/</a>		
V-13	Kapitalanlage	Milliarde USD
V-14	Wechselkurse	Einheiten der Landeswährung pro U.S. Dollar
V-15	Arbeitslosenzahl	Prozent
V-16	Erwerbstätigenquote	Prozent
V-17	Ausgaben der Regierung	Milliarde USD
V-18	Prognose der Investitionen	Verhältnis zwischen Gesamtinvestitionen und BIP
V-19	Wettbewerbsfähigkeit - Index des Weltwirtschaftsforums (1)	1-7
V-20	Wettbewerbsfähigkeit - Index des Weltwirtschaftsforums (2)	0-100
V-21	Schattenwirtschaft	Prozent des BIP
V-22	Kontrolle der Korruption	-2.5 schwach; 2.5 stark
V-23	Index der politischen Stabilität	-2.5 schwach; 2.5 stark
V-24	Kurzfristiges politisches Risiko	1=gering, 7=hoch
V-25	Mittel-/langfristiges politisches Risiko	1=gering, 7=hoch
V-26	Internetnutzer, Prozent der Bevölkerung	Prozent
V-27	Qualität der Straßen	1=gering, 7=hoch
V-28	Innovationsindex	0-100
V-29	Export von Informationstechnologie	Prozent der gesamten Warenexporte
V-30	Bankkredite an Staat und öffentliche Unternehmen	Prozent des BIP
V-31	Anzahl der börsennotierten Unternehmen	Anzahl
V-32	Innovationsindex	0-100
V-33	Index der Eigentumsrechte	0-100
V-34	Index der Freiheit von Korruption	0-100

**Tabelle C.5:** Externe Daten zur Beschreibung der wirtschaftlichen und politischen Lage (2)

<b>Quelle: <a href="https://www.theglobaleconomy.com/">https://www.theglobaleconomy.com/</a></b>		
V-35	Index der fiskalischen Freiheit	0-100
V-36	Index der unternehmerischen Freiheit	0-100
V-37	Index der Arbeitsfreiheit	0-100
V-38	Index der monetären Freiheit	0-100
V-39	Index der Handelsfreiheit	0-100
V-40	Index der Investmentfreiheit	0-100
V-41	Index der finanziellen Freiheit	0-100
V-42	Wirtschaftliche Freiheit, Gesamtindex	0-100
V-43	Gesundheitsausgaben pro Kopf	USD pro Einwohner
V-44	Todesrate	pro 1000 Personen
V-45	Armut	Prozent der Bevölkerung
V-46	Öffentliche Ausgaben für Bildung	Prozent des BIP
V-47	Globalisierungsindex	0-100
V-48	Index der wirtschaftlichen Globalisierung	0-100
V-49	Index der politischen Globalisierung	0-100
V-50	Index der sozialen Globalisierung	0-100
V-51	Prozentsatz der Weltbevölkerung	Prozent
V-52	Prozentsatz des weltweiten BIP	Prozent
V-53	Prozent der weltweiten Exporte	Prozent
V-54	Prozent der weltweiten Importe	Prozent
V-55	Wertschöpfung der Industrie	Milliarde USD
V-56	Mehrwert der Dienstleistungen	Milliarde USD
V-57	Mengenindex der lokalen Anbieter	1=gering, 7=hoch
V-58	Glücksindex	0=unglücklich, 10=glücklich
V-59	Index der menschlichen Entwicklung	0 - 1
V-60	Landfläche	Quadratkilometer
V-61	Hauspreisindex (HPI)	prozentuale Veränderung; Basisjahr = 100
V-62	Baugenehmigungen	Anzahl
V-63	Reale Preise für Wohnimmobilien im Vergleich zum Vorjahr	prozentuale Veränderung
V-64	Verbraucherpreisindex (CPI)	prozentuale Veränderung; Basisjahr = 10
V-65	Staatliche Ausgaben	Milliarden Währungseinheiten



## C.4 Python Code

```
# Auswertung: Prognose Dauer von Bauprojekten

## Import der benötigten Bibliotheken

!pip install tensorflow==1.14.0
%tensorflow_version 1.x

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# Bibliotheken für System I/O
import os
import io

# Sklearn: universelle Bibliothek im ML-Bereich
from sklearn import preprocessing

# Keras: Bibliothek für KNN
from keras.models import Sequential
from keras.layers import Dense, Activation, Dropout
from keras.optimizers import rmsprop

# Bibliotheken, um Daten aus Google Drive zu laden
from pydrive.auth import GoogleAuth
from pydrive.drive import GoogleDrive
from google.colab import auth
from oauth2client.client import GoogleCredentials
```

```
# XGBoost is an optimized distributed gradient boosting library  
import xgboost as xgb
```

```
# Talos zur Optimierung von Keras KNN
```

```
!pip install talos  
import talos as ta  
from talos.utils.best_model import best_model, activate_model  
from talos.model.normalizers import lr_normalizer
```

```
# Shap Bibliothek: A game theoretic approach to  
explain the output of any machine learning model.
```

```
!pip install shap  
import shap
```

```
# gradient boosting on decision trees (Similar to XGBoost)
```

```
!pip install catboost
```

```
## Wahl des Modus & Setzen des Seeds
```

Wahl des entsprechenden Modus durch setzen der Konstante MODE.

Mögliche Optionen: "1a", "1b", "2a", "2b"

Setzen des Seeds zur Reproduzierbarkeit des Zufallszustands

```
MODE = "2a"  
SEED = 314
```

```
## Auswahl des Datensatzes
```

Wahl der gewünschten Datenquelle, d.h. ausführen

des entsprechenden Blocks

Der Datensatz muss `in` einer csv-Datei vorliegen. Spaltenüberschriften müssen vorhanden sein. Die Zielvariablen müssen `'Duration'` bzw. `'Cost'` benannt sein.

```
# A: Irandatensatz ohne externen Daten - nicht  
bereinigt (Schritt a)  
datasource = "https://drive.google.com/open?id=1z4c0lxV7HktVoqPpnn7j_xM0yOZ2Y5yY"  
datagoogleid = "1z4c0lxV7HktVoqPpnn7j_xM0yOZ2Y5yY"  
decimalseparator = ","
```

```
# B: Irandatensatz mit externen Daten - nicht  
bereinigt (Schritt b)  
datasource = "https://drive.google.com/open?id=1X8IY9fsKxl_2KQhcGDQUuOn-e9Mq2WFu"  
datagoogleid = "1X8IY9fsKxl_2KQhcGDQUuOn-e9Mq2WFu"  
decimalseparator = ","
```

```
# C: Irandatensatz mit externen Daten - bereinigt  
(Schritt c)  
datasource = "https://drive.google.com/open?id=13LNqNopyZuKAC1baqKWYPbdBMAcsqGQL"  
datagoogleid = "13LNqNopyZuKAC1baqKWYPbdBMAcsqGQL"  
decimalseparator = ","
```

```
# B: NY Datensatz ohne externen Daten - nicht  
bereinigt - Apartments (Schritt b)  
datasource = "https://drive.google.com/open?id=1hcmJPidl7QRELQvYR4RbogyDG59eBYGa"  
datagoogleid = "1hcmJPidl7QRELQvYR4RbogyDG59eBYGa"
```

```
decimalseparator = ","
#Fold=3 -> Catboost: 55,98 Tage (ohne externe
Daten!)

# B: NY Datensatz mit externen Daten - nicht
bereinigt - Schulen (Schritt b)
datasource = "https://drive.google.com/open?id=1pCQ
jwyKdBSImnKi2BP-vUFKMH-sH7x4x"
datagoogleid = "1pCQjwyKdBSImnKi2BP-vUFKMH-sH7x4x"
decimalseparator = ","

## Laden und einlesen der Datenbank

# Laden und einlesen der zuvor ausgewählten
Datenbank
auth.authenticate_user()
gauth = GoogleAuth()
gauth.credentials =
GoogleCredentials.get_application_default()
drive = GoogleDrive(gauth)

file = drive.CreateFile({'id': datagoogleid})
data = pd.read_csv(io.StringIO(file.GetContentStrin
g()), sep=';', decimal=decimalseparator, header=0)

## Datenvorbereitung

Funktion, die bei den verschiedenen Algorithmen
aufgerufen wird und somit jeweils die Daten passend
vorbereitet

from pandas import read_csv
```

```
from numpy import nan

data.fillna(data.mean(), inplace=True)

#outlier detection

df = pd.DataFrame(data)
Q1 = df.quantile(0.25)
Q3 = df.quantile(0.75)
IQR = Q3 - Q1

print(IQR)

data = df[~((df < (Q1 - 1.5 * IQR)) |(df > (Q3 + 1.5
* IQR))).any(axis=1)]

def prepare_data(data, target, delete_nans=True,
convert_categorical=True, normalize=True,
splitratio=0.8, seed=SEED):

    if(delete_nans == True):
        #Ausgabe der Daten vor Entfernung der
        Leerdaten
        print('Mit Leerdaten: ' + str(data.shape))

        #Löschen von Projekten, in denen die Dauer
        fehlt
        #data = data.dropna(axis=0,
        subset=['Duration'])

        #Löschen von Projekten, in denen die Dauer 0
        ist
        data = data[~data['Duration'].isin([0])]
```

```
#Ausgabe der Daten nach Entfernung der
Leerdaten
print('Ohne Leerdaten: ' + str(data.shape))

if(convert_categorical == True):
    #Ausgabe der Daten vor One Hot Codierung
    print('Vor One Hot Codierung: ' +
          str(data.shape))

    # One Hot Codierung
    data = pd.get_dummies(data)

    #Ausgabe der Daten nach One Hot Codierung
    print('Nach One Hot Codierung: ' +
          str(data.shape))

if(normalize == True):
    # Normalisieren der Daten [Werte zwischen 0
und 1], außer Target
    tmp = data
    data_min = data.min()
    data_max = data.max()
    data = (data - data_min) / (data_max -
                               data_min)
    data[target] = tmp[target]

# Trennung Trainings-/ Testdaten
train = data.sample(frac=splitratio,
                    random_state=seed)
test = data.drop(train.index)

return train, test
```

```
## Datenverteilung

### Schiefe Lage und Kurtosis

import seaborn as sns
from scipy.stats import norm, skew

# Verteilungsplot
sns.distplot(data['Duration'], kde=True, fit=norm)
(mu,sigma)=norm.fit(data['Duration'])
plt.legend(['Normal dist. ( $\mu$ = ${:.2f}$  and
 $\sigma$ = ${:.2f}$ )'.format(mu,sigma)], loc='best')
plt.xlabel('Duration')
plt.ylabel('Frequency')
plt.title('Duration Verteilung')

# Werte der Skewness und Kurtosis
print("Skewness: %f"% data['Duration'].skew())
print("Kurtosis: %f"% data['Duration'].kurt())

### Standardisierung mit log1p

y_log1p = np.log1p(data['Duration'])
sns.distplot(y_log1p, fit=norm)
(mu,sigma) = norm.fit(y_log1p)

plt.legend(['Normal dist. ( $\mu$ = ${:.2f}$  and
 $\sigma$ = ${:.2f}$ )'.format(mu,sigma)], loc='best')
plt.xlabel('Duration')
plt.ylabel('Frequency')
plt.title('Duration Verteilung nach
```

```
log-Transformation')
```

```
#Schieflage und Kurtosis
```

```
print("Skewness: %f"% y_log1p.skew())  
print ("Kurtosis: %f"% y_log1p.kurt())
```

```
### Q-Q-Plot nach der log-Transformation
```

```
from scipy import stats  
res = stats.probplot(np.log1p(data['Duration']),  
plot=plt)  
plt.title('Q-Q-Plot der Dauer nach der  
log-Transformation')  
plt.show()
```

```
#Skewness aller numerischen Features überprüfen  
skewed_feats = data.apply(lambda x:  
skew(x.dropna())).sort_values(ascending=False)  
skewness = pd.DataFrame({'Skew' :skewed_feats})  
skewness.head(250)
```

```
## Auswahl der Merkmale und Trennung von  
Zielvariable
```

Definiert die Funktionen, die bei den jeweiligen Algorithmen nach der Vorbereitung der Daten aufgerufen werden, um die relevanten Features je nach Modus aufzurufen.

```
def select_data(train, test, target,  
unwanted_features=[]):  
    if target not in unwanted_features:  
        unwanted_features.append(target)
```



```
train_target = train.loc[:, target]
test_target = test.loc[:, target]
for feature in unwanted_features:
    if feature in train:
        train_data = train.drop(columns=[feature])
        test_data = test.drop(columns=[feature])
return train_data, train_target, test_data,
test_target

def select_mode(mode):
    duration = "Duration"
    cost = "Costs"
    target_1a = target_2a = duration
    target_1b = target_2b = cost
    cost_dependent_features = ["Costs", cost] # V-8:
    Preis/m2

    # 1a) Vorhersage der Dauer exklusive Kosten in
    Trainingsdaten
    if mode == "1a":
        return target_1a, cost_dependent_features

    # 1b) Vorhersage der Kosten exklusive Dauer in
    Trainingsdaten
    elif mode == "1b":
        return target_1b, ([duration] +
            cost_dependent_features)

    # 2a) Vorhersage der Dauer inklusive Kosten in
    Trainingsdaten
    elif mode == "2a":
        return target_2a, []
```

```
# 2b) Vorhersage der Kosten exklusive Dauer in
Trainingsdaten
elif mode == "2b":
    return target_2b, cost_dependent_features

else:
    raise ValueError(str(mode) + " is not a valid
mode.")

# Regression

### Feature Auswahl

target, unwanted_features = select_mode(MODE)
train, test = prepare_data(data,
                           target,
                           delete_nans=False,
                           convert_categorical=True,
                           normalize=False)
train_data, train_target, test_data, test_target =
select_data(train, test, target, unwanted_features)
#definition features and target variable
# x = train.drop(['Duration'], axis=1, inplace = False)
# y = train['Duration']
x = train_data
y = train_target

#Feature Selection using Random Forest
from sklearn.ensemble import RandomForestRegressor
#do not use RandomForestClassifier!

#parameter setting
# rf = RandomForestRegressor(max_features=100)
```

```
rf = RandomForestRegressor()
rf.fit(x,y)
importances = pd.DataFrame({'feature':x.columns,
'importance':np.round(rf.feature_importances_,3)})
importances = importances.sort_values('importance',
ascending=False).set_index('feature')

#display 25 most important features
import matplotlib.pyplot as plt
%matplotlib inline
importances.head(25).plot(kind='bar',figsize=(30,20
),alpha=1,
width = 0.8,facecolor = 'yellowgreen',
edgecolor = 'white',lw=1,fontsize=40)
plt.tick_params(labelsize=18)
font1 = {'family': 'Times New Roman','weight':
'normal','size': 50,}
legend = plt.legend(prop=font1)
font2 = {'family': 'Times New Roman','weight':
'normal','size': 50,}
plt.xlabel('Features', font2)
font3 = {'family': 'Times New Roman','weight':
'normal','size': 50,}
plt.ylabel('Correlation', font3)

### Ranking der Features
Heatmap

import seaborn as sns

#Top 25 in der Heatmap
k = 26 #include target 'Duration'
corrmat = data.corr()
```

```
cols = corrmatrix.nlargest(k,
    'Duration')['Duration'].index
cm = np.corrcoef(data[cols].values.T)
sns.set(font_scale = 1.6)
plt.figure(figsize = (30,20))
hm = sns.heatmap(cm, cbar=True,
    annot=True,square=True,fmt='%.2f',cmap='YlGnBu',
    linewidths=.5,
    annot_kws={'size':
16},yticklabels=cols.values,xticklabels=cols.values)
plt.show()
```

```
##Regressionsanalyse
```

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# cols_ratio = ['Cost', 'V-2', 'V-18.4', 'V-28.4',
'V-28', 'V-27', 'V-27.1', 'V-18', 'V-24', 'V-27.3']
cols_ratio = cols.tolist()[1:5]
cols_target = ['Duration']

#pd.plotting.scatter
grr = pd.plotting.scatter_matrix(train[cols_target +
cols_ratio]
```

```
,c = train_target
,figsize=(20, 20)
,marker = 'o'
,hist_kws={'bins'
: 20}
,s = 15
,alpha = 0.8)
```

```
for ax in grr.ravel():
    ax.set_xlabel(ax.get_xlabel(), fontsize = 15,
rotation = 90)
    ax.set_ylabel(ax.get_ylabel(), fontsize = 15,
rotation = 0)

plt.show()

##Lineare Regression

X = data[['Costs']]
y = data['Duration']

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test =
train_test_split(X, y, test_size = 0.3,
random_state = SEED)

from sklearn.linear_model import LinearRegression
lr = LinearRegression()
lr.fit(X_train, y_train)

print('----- Lineare Regression -----')
print('Funktion via sklearn: y = %.3f * x + %.3f' %
(lr.coef_[0], lr.intercept_))
print("Alpha: {}".format(lr.intercept_))
print("Beta: {}".format(lr.coef_[0]))
print("Training Set R2 Score:
{:.2f}".format(lr.score(X_train, y_train)))
print("Test Set R2 Score:
```

```
{:.2f}").format(lr.score(X_test, y_test)))
print("\n")

plt.figure(figsize=(10,10))
plt.scatter(X_train, y_train, color = 'blue')
# Blaue Punkte sind Trainingsdaten
plt.scatter(X_test, y_test, color = 'green')
# Grüne Punkte sind Testdaten
plt.plot(X_train, lr.predict(X_train), color =
'red')
plt.xlabel(X_train.columns[0])
plt.ylabel(cols_target[0])
plt.show()

##Polynomiale Regression

from sklearn.preprocessing import PolynomialFeatures

quadratic = PolynomialFeatures(degree = 3)

X_train_quad = quadratic.fit_transform(X_train)
X_test_quad = quadratic.fit_transform(X_test)

pr = LinearRegression()

pr.fit(X_train_quad, y_train)

X_fit = np.arange(X_train.min().values,
X_train.max().values)
X_fit = X_fit[:, np.newaxis]

X_fit_quad = quadratic.fit_transform(X_fit)
```

```
## Decision Tree Regressor

from sklearn.tree import DecisionTreeRegressor

tree = DecisionTreeRegressor(max_depth=5)

tree.fit(X_train, y_train)

plt.figure(figsize=(10,10))
plt.scatter(X_train, y_train, color = 'blue')
plt.scatter(X_test, y_test, color = 'green')
plt.plot(X_train, lr.predict(X_train), color =
'red', linewidth = 2.0)
plt.plot(X_fit, pr.predict(X_fit_quad), color =
'purple', linewidth = 2.0)
plt.plot(X_fit, tree.predict(X_fit), color =
'orange', linewidth = 2.0)
plt.xlabel(X_train.columns[0])
plt.ylabel(cols_target[0])

plt.show()

#Classification

### Random Forest Modellierung

from sklearn import preprocessing
from sklearn.ensemble import RandomForestRegressor

target, unwanted_features = select_mode(MODE)
train, test = prepare_data(data,
                           target,
                           delete_nans=False,
```

```
        convert_categorical=True,
        normalize=False)
train_data, train_target, test_data, test_target =
select_data(train, test, target, unwanted_features)

RF = RandomForestRegressor( n_estimators=500,
                           max_depth=6,
                           min_samples_split=2,
                           min_samples_leaf=2,
                           max_features=10,
                           Min_weight_fraction_leaf=0)
RF.fit(train_data, train_target)
y_pred = RF.predict(test_data)

#average error of Duration
MD=sum(abs(y_pred - test_target)/len(y_pred))

#Normalized RMSE
RMSE=(sum(((y_pred) - (test_target))*((test_target)
- (test_target))/len(y_pred)))*0.5
print('average error of Duration=',MD)
print('root-mean-square error of Duration=',RMSE)

#Random Forest Visualization
from sklearn.tree import export_graphviz
estimator = RF.estimators_[5]
export_graphviz(estimator, out_file='tree.dot',
                class_names=['Duration'],
                rounded = True, proportion = False,
                precision = 0, filled = True)
!dot -Tpng tree.dot -o tree.png
from IPython.display import Image
Image(filename = 'tree.png')
```



```
#plot one tree of XGboost
import xgboost as xgb
model = xgb.XGBRegressor(n_estimators=166,
learning_rate=0.05,max_depth=6)
model.fit(train_data, train_target,
early_stopping_rounds=5, eval_set=[( test_data,
test_target)], verbose=True)
xgb.to_graphviz(model)
#print

#### Model Tuning Random Forest

from sklearn.model_selection import GridSearchCV
from sklearn.model_selection import StratifiedKFold
from sklearn.metrics import roc_auc_score

#search the best parameter using param_grid
param_grid = { 'max_features': ['auto', 'sqrt',
'log2'],
'max_depth': [5,6,10,13],
}
rfc = RandomForestRegressor()
CV_rfc = GridSearchCV(estimator=rfc,
param_grid=param_grid, cv=10)
CV_rfc.fit(train_data, train_target)
CV_rfc.best_params_

## Gradient Boosting Regression Modelling

import numpy as np
import matplotlib.pyplot as plt
```

```
from sklearn import ensemble
from sklearn import datasets
from sklearn.utils import shuffle
from sklearn.metrics import mean_squared_error

# Fit regression model
params = {'n_estimators': 500, 'max_depth': 4,
         'min_samples_split': 2,
         'learning_rate': 0.01, 'loss': 'ls'}
clf = ensemble.GradientBoostingRegressor(**params)

clf.fit(train_data, train_target)
mse = mean_squared_error(test_target,
                          clf.predict(test_data))
print("MSE: %.4f" % mse)

# Plot training deviance

# compute test set deviance
test_score = np.zeros((params['n_estimators'],),
                      dtype=np.float64)

for i, pred_target in
    enumerate(clf.staged_predict(test_data)):
    test_score[i] = clf.loss_(test_target,
                              pred_target)

plt.figure(figsize=(12, 30))
plt.subplot(1, 2, 1)
plt.title('Deviance')
plt.plot(np.arange(params['n_estimators']) + 1,
         clf.train_score_, 'b-',
         label='Training Set Deviance')
```

```
plt.plot(np.arange(params['n_estimators']) + 1,
test_score, 'r-',
        label='Test Set Deviance')
plt.legend(loc='upper right')
plt.xlabel('Boosting Iterations')
plt.ylabel('Deviance')

# Plot feature importance
feature_importance = clf.feature_importances_

# make importances relative to max importance
feature_importance = 100.0 * (feature_importance /
feature_importance.max())
sorted_idx = np.argsort(feature_importance)
pos = np.arange(sorted_idx.shape[0]) + .5
plt.subplot(1, 2, 2)
plt.barh(pos, feature_importance[sorted_idx],
align='center')
# plt.yticks(pos, data.feature_names[sorted_idx])
plt.yticks(pos, data.columns[sorted_idx].to_list())
plt.xlabel('Relative Importance')
plt.title('Variable Importance')
plt.show()

##XGB Rregression Modelling

# plot decision tree
from numpy import loadtxt
from xgboost import XGBClassifier
from xgboost import plot_tree
import matplotlib.pyplot as plt

# fit model no training data
```

```
model = XGBClassifier()
model.fit(train_data, train_target)

# plot
plot_tree(model, num_trees=16)
# wenn Ansicht von Links nach Rechts
# plot_tree(model, num_trees=6, rankdir='LR')
plt.show()

## LightGBM Modelling

import lightgbm as lgb
import pandas as pd

if lgb.compat.MATPLOTLIB_INSTALLED:
    import matplotlib.pyplot as plt
else:
    raise ImportError('You need to install matplotlib
    for plot_example.py.')

print('Loading data...')

# create dataset for lightgbm
lgb_train = lgb.Dataset(train_data, train_target)
lgb_test = lgb.Dataset(test_data, test_target,
reference=lgb_train)

# specify your configurations as a dict
params = {'num_leaves': 5,
          'metric': ('l1', 'l2'),
          'verbose': 0}

evals_result = {} # to record eval results for
```

plotting

```
print('Starting training...')

# train
gbm = lgb.train(params,
                lgb_train,
                num_boost_round=100,
                valid_sets=[lgb_train, lgb_test],
                #feature_name=['f' + str(i + 1) for
                i in range(train_data.shape[-1])],
                categorical_feature=[21],
                evals_result=evals_result,
                verbose_eval=10)

print('Plotting metrics recorded during
training...')
ax = lgb.plot_metric(evals_result, metric='l1')
plt.show()

print('Plotting feature importances...')
ax = lgb.plot_importance(gbm, max_num_features=10)
plt.show()

print('Plotting 54th tree...') # one tree use
categorical feature to split
ax = lgb.plot_tree(gbm, tree_index=53, figsize=(15,
15), show_info=['split_gain'])
plt.show()

print('Plotting 54th tree with graphviz...')
graph = lgb.create_tree_digraph(gbm, tree_index=53,
name='Tree54')
```

```
graph.render(view=True)

##Catboost Modellierung

import numpy as np
import catboost
from catboost import CatBoostRegressor, Pool

# find non-numerical
is_cat = (train_data.dtypes != float)

# fill in NaNs for non-numericals
for feature, feat_is_cat in
is_cat.to_dict().items():
    if feat_is_cat:
        train_data[feature].fillna("NaN",
            inplace=True)

cat_features_index = np.where(is_cat)[0]
pool = Pool(train_data, train_target,
cat_features=cat_features_index,
            feature_names=list(train_data.columns))

#model = CatBoostClassifier(
#max_depth=6, verbose=False, max_ctr_complexity=1,
iterations=2).fit(pool)

model = CatBoostRegressor(
    max_depth=2, verbose=False,
    max_ctr_complexity=1, random_seed=42,
    iterations=2).fit(pool)

model.plot_tree(
```

```
    tree_idx=0,
    pool=None
)

# Prediction accuracy of classification ohne KNN
(incl. Stacking Ensemble)

import numpy as np
import pandas as pd
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns
color = sns.color_palette()
sns.set_style('darkgrid')
import warnings
def ignore_warn(*args, **kwargs):
    pass
warnings.warn = ignore_warn
from scipy import stats
from scipy.stats import norm, skew
from sklearn.pipeline import make_pipeline #using a
pipeline we can chain together modeling and scaling
where it is necessary
from sklearn.preprocessing import RobustScaler
#scaling in pipeline (robust to outliers!)
from sklearn.base import BaseEstimator,
TransformerMixin, RegressorMixin, clone
from sklearn.model_selection import KFold,
cross_val_score, train_test_split
from sklearn.metrics import mean_absolute_error
from sklearn.ensemble import RandomForestRegressor,
GradientBoostingRegressor
import xgboost as xgb
```

```
import lightgbm as lgb
import catboost
from catboost import CatBoostRegressor

import keras
from keras.layers import Dense, Dropout
from keras import metrics
from keras import backend as K
from keras.wrappers.scikit_learn import
KerasRegressor
from keras import optimizers

#K-folds validation to evaluate machine learning
models
n_folds = 10

def scores(model):
    kf = KFold(n_folds, shuffle=True,
              random_state=SEED).get_n_splits(train_data.values)
    scores= -cross_val_score(model, train_data.values,
                              train_target.values,
                              scoring='neg_mean_absolute_error', cv=kf)
    return(scores)

#parameters without Model Tunning
#Random Forest
RF = RandomForestRegressor(max_depth=10,
                           random_state=SEED,
                           max_features='auto')

#Gradient Boosting Regression
GBoost = GradientBoostingRegressor(learning_rate=0.
1, max_depth=3, random_state=SEED,
max_features='auto')
```



```
#XGBoost
model_xgb = xgb.XGBRegressor(learning_rate=0.08,
                             max_depth=3,
                             min_child_weight=4,
                             max_features='auto')

#LightGBM
model_lgb = lgb.LGBMRegressor(num_leaves=200,
                              max_depth=10)

#Catboost:
model_catboost = CatBoostRegressor(max_depth=6,
                                   verbose=0)

#Base models scores
score = scores(RF)
print("Random Forest score:",score.mean())
score = scores(GBoost)
print("GBRT score:",score.mean())
score = scores(model_xgb)
print("XGBoost score:",score.mean())
score = scores(model_lgb)
print("LightGBM score:",score.mean())
score = scores(model_catboost)
print("Catboost score:",score.mean())

models_score = pd.DataFrame({"random forest":
                             scores(RF),
                             "gboost": scores(GBoost),
                             "xgboost": scores(model_xgb),
                             "lightgbm": scores(model_lgb),
                             "catboost": scores(model_catboost),
                             "stacked models":
                             scores(model_catboost)})
```

```
models_score.plot.bar()
plt.legend(loc='center left', bbox_to_anchor=(1.0,
0.5))
plt.show()
```

```
#Stackingmodels with meta-model
#definite StackingAveragedModels
class StackingAveragedModels(BaseEstimator,
RegressorMixin, TransformerMixin):

    def __init__(self, base_models, meta_model,
n_folds=10):
        self.base_models = base_models
        self.meta_model = meta_model
        self.n_folds = n_folds

    def fit(self, data, target):
        self.base_models_ = [list() for target in
self.base_models]
        self.meta_model_ = clone(self.meta_model)
        kfold = KFold(n_splits=self.n_folds,
shuffle=True, random_state=SEED)
        # Train cloned base models then create
out-of-fold predictions
        # that are needed to train the cloned meta-model
        out_of_fold_predictions =
np.zeros((data.shape[0], len(self.base_models)))
        for i, model in enumerate(self.base_models):
            for train_index, holdout_index in
kfold.split(target, data):
                instance = clone(model)
                self.base_models_[i].append(instance)
                instance.fit(data[train_index],
```

```

        target[train_index])
        y_pred =
        instance.predict(data[holdout_index])
        out_of_fold_predictions[holdout_index, i] =
        y_pred
        # Now train the cloned meta-model using the
        out-of-fold predictions as new feature
        self.meta_model_.fit(out_of_fold_predictions,
        target)
        return self
        #Do the predictions of all base models on the test
        data and use the averaged predictions as
        #meta-features for the final prediction which is
        done by the meta-model

    def predict(self, target):
        meta_features = np.column_stack([
            np.column_stack([model.predict(target) for
            model in base_models]).mean(axis=1)
            for base_models in self.base_models_ ])
        return self.meta_model_.predict(meta_features)
        #build stacked_averaged_models, base_models von
        RF,GBoost,XGBoost,Catboost, SVM, meta_models von
        GBoost
        stacked_averaged_models =
        StackingAveragedModels(base_models =
        (RF,GBoost,model_xgb,model_catboost),
        meta_model = GBoost)
        #score of stacked_averaged_models
        score = scores(stacked_averaged_models)
        print("stacked_averaged_models score:",score.mean())

        #define AveragingModels

```

```
class AveragingModels(BaseEstimator, RegressorMixin,
TransformerMixin):
    def __init__(self, models):
        self.models = models

    # define clones of the original models to fit the
    data in
    def fit(self, data, target):
        self.models_ = [clone(model) for model in
self.models]

        # Train cloned base models
        for model in self.models_:
            model.fit(data, target)
        return self

    # do the predictions for cloned models and average
    them
    def predict(self, data):
        predictions = np.column_stack([
            model.predict(data) for model in self.models_
        ])
        return np.mean(predictions, axis=1)

#average all of models with AveragingModels function
averaged_models = AveragingModels(models =
(RF,GBoost,model_xgb,model_catboost))# model_lgb is
weak, not be used

#score of averaged_models
score = scores(averaged_models)

RF.fit(train_data.values, train_target.values)
```

```
RF_test_pred = RF.predict(test_data.values)
GBoost.fit(train_data.values, train_target.values)
GBoost_test_pred = GBoost.predict(test_data.values)
model_xgb.fit(train_data.values,
train_target.values)
xgb_test_pred = model_xgb.predict(test_data.values)
model_lgb.fit(train_data.values,
train_target.values)
lgb_test_pred = model_lgb.predict(test_data.values)
model_catboost.fit(train_data.values,
train_target.values)
catboost_test_pred
=model_catboost.predict(test_data.values)

#averaged_models
stacked_averaged_models.fit(train_data.values,
train_target.values)
stacked_train_pred =
stacked_averaged_models.predict(train_data.values)
#stacked_train_pred =
stacked_averaged_models.predict(train_data.values)

#stacked_averaged_models
stacked_averaged_models.fit(train_data.values,
train_target.values)
stacked_pred =
stacked_averaged_models.predict(test_data.values)
#stacked_pred =
stacked_averaged_models.predict(test_data.values)

random_forest_mae = mean_absolute_error(test_target,
RF_test_pred)
gboost_mae = mean_absolute_error(test_target,
```

```
GBoost_test_pred)
xgboost_mae = mean_absolute_error(test_target,
xgb_test_pred)
lightgbm_mae = mean_absolute_error(test_target,
lgb_test_pred)
catboost_mae = mean_absolute_error(test_target,
catboost_test_pred)
#knn_mae = mean_absolute_error(test_target, )
#stacked_mae = mean_absolute_error(test_target,
stacked_test_pred)
#stacked_models_mae =
mean_absolute_error(test_target, knn_test_pred)
stacked_mae = mean_absolute_error(test_target,
stacked_pred)

print('Random Forest mean_absolute_error = ',
random_forest_mae)
print('GBoost mean_absolute_error = ', gboost_mae)
print('XGBoost mean_absolute_error = ', xgboost_mae)
print('LightGBM mean_absolute_error = ',
lightgbm_mae)
print('CatBoost mean_absolute_error = ',
catboost_mae)
print('stacked_averaged_models MAE = ', stacked_mae)

models_mae = pd.DataFrame({"randam forest":
[random_forest_mae],
"gboost": [gboost_mae],
"xgboost": [xgboost_mae],
"lightgbm": [lightgbm_mae],
"catboost": [catboost_mae],
"stacked models": [stacked_mae]})
```

```
models_mae.plot.bar()
plt.legend(loc='center left', bbox_to_anchor=(1.0,
0.5))
plt.show()

random_forest_mape = np.mean(np.abs((test_target -
RF_test_pred) / test_target)) * 100
gboost_mape = np.mean(np.abs((test_target -
GBoost_test_pred) / test_target)) * 100
xgboost_mape = np.mean(np.abs((test_target -
xgb_test_pred) / test_target)) * 100
lightgbm_mape = np.mean(np.abs((test_target -
lgb_test_pred) / test_target)) * 100
catboost_mape = np.mean(np.abs((test_target -
catboost_test_pred) / test_target)) * 100
stacked_models_mape = np.mean(np.abs((test_target -
stacked_pred) / test_target)) * 100

print('MAPE Random Forest = ', random_forest_mape)
print('MAPE XBoost = ', gboost_mape)
print('MAPE XGBoost = ', xgboost_mape)
print('MAPE Lightgbm = ', lightgbm_mape)
print('MAPE Catboost = ', catboost_mape)
print('MAPE Stacked Models = ', stacked_models_mape)

models_mape = pd.DataFrame({"random forest":
[random_forest_mape],
"gboost": [gboost_mape],
"xgboost": [xgboost_mape],
"lightgbm": [lightgbm_mape],
"catboost": [catboost_mape],
"stacked models":
[stacked_models_mape]})
```

```
models_mape.plot.bar()

plt.legend(loc='center left', bbox_to_anchor=(1.0,
0.5))
plt.show()

# Prediction accuracy of classification mit KNN
(incl. Stacking Ensemble)

import numpy as np
import pandas as pd
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns
color = sns.color_palette()
sns.set_style('darkgrid')
import warnings
def ignore_warn(*args, **kwargs):
    pass
warnings.warn = ignore_warn
from scipy import stats
from scipy.stats import norm, skew
from sklearn.pipeline import make_pipeline #using a
pipeline we can chain together modeling and scaling
where it is necessary
from sklearn.preprocessing import RobustScaler
#scaling in pipeline (robust to outliers!)
from sklearn.base import BaseEstimator,
TransformerMixin, RegressorMixin, clone
from sklearn.model_selection import KFold,
cross_val_score, train_test_split
from sklearn.metrics import mean_absolute_error
```



```
from sklearn.ensemble import RandomForestRegressor,
GradientBoostingRegressor
import xgboost as xgb
import lightgbm as lgb
import catboost
from catboost import CatBoostRegressor

import keras
#from keras.models import Sequential
from keras.layers import Dense, Dropout
from keras import metrics
from keras import backend as K
from keras.wrappers.scikit_learn import
KerasRegressor
from keras import optimizers

#K-folds validation to evaluate machine learning
models
n_folds = 4

def scores(model):
    kf = KFold(n_folds, shuffle=True,
random_state=SEED).get_n_splits(train_data.values)
    scores= -cross_val_score(model, train_data.values,
train_target.values,
scoring='neg_mean_absolute_error', cv=kf)
    return(scores)

#parameters without Model Tunning
#Random Forest
RF = RandomForestRegressor(max_depth=10,
random_state=SEED,
max_features='auto')
```

```
#Gradient Boosting Regression
GBoost = GradientBoostingRegressor(learning_rate=0.
1, max_depth=3, random_state=SEED,
max_features='auto')
#XGBoost
model_xgb = xgb.XGBRegressor(learning_rate=0.08,
                             max_depth=3,
                             min_child_weight=4,
                             max_features='auto')

#LightGBM
model_lgb = lgb.LGBMRegressor(num_leaves=200,
                              max_depth=10)

#Catboost:
model_catboost = CatBoostRegressor(max_depth=6,
                                   verbose=0)

#Base models scores
score = scores(RF)
print("Random Forest score:",score.mean())
score = scores(GBoost)
print("GBRT score:",score.mean())
score = scores(model_xgb)
print("XGBoost score:",score.mean())
score = scores(model_lgb)
print("LightGBM score:",score.mean())
score = scores(model_catboost)
print("Catboost score:",score.mean())

from sklearn.neural_network import MLPClassifier
from sklearn.datasets import make_classification
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
```

```
model_knn = MLPClassifier(hidden_layer_sizes=(100,1
00,100), max_iter=500, alpha=0.0001,
                           solver='sgd', verbose=10,
                           random_state=21,
                           tol=0.000000001)

score = scores(model_knn)
print("KNN:", score.mean())

#Stackingmodels with meta-model
#definite StackingAveragedModels
class StackingAveragedModels(BaseEstimator,
RegressorMixin, TransformerMixin):

    def __init__(self, base_models, meta_model,
n_folds=10):
        self.base_models = base_models
        self.meta_model = meta_model
        self.n_folds = n_folds

    def fit(self, data, target):
        self.base_models_ = [list() for target in
self.base_models]
        self.meta_model_ = clone(self.meta_model)
        kfold = KFold(n_splits=self.n_folds,
shuffle=True, random_state=SEED)
        # Train cloned base models then create
        out-of-fold predictions
        # that are needed to train the cloned meta-model
        out_of_fold_predictions =
np.zeros((data.shape[0], len(self.base_models)))
        for i, model in enumerate(self.base_models):
            for train_index, holdout_index in
```

```
kfold.split(target, data):
    instance = clone(model)
    self.base_models_[i].append(instance)
    instance.fit(data[train_index],
                target[train_index])
    y_pred =
    instance.predict(data[holdout_index])
    out_of_fold_predictions[holdout_index, i] =
    y_pred
    # Now train the cloned meta-model using the
    out-of-fold predictions as new feature
    self.meta_model_.fit(out_of_fold_predictions,
                        target)
    return self
#Do the predictions of all base models on the test
data and use the averaged predictions as
#meta-features for the final prediction which is
done by the meta-model

def predict(self, target):
    meta_features = np.column_stack([
        np.column_stack([model.predict(target) for
            model in base_models]).mean(axis=1)
        for base_models in self.base_models_ ])
    return self.meta_model_.predict(meta_features)
#build stacked_averaged_models, base_models von
RF,GBoost,XGBoost,Catboost, SVM, meta_models von
GBoost
stacked_averaged_models =
StackingAveragedModels(base_models =
(RF,GBoost,model_xgb,model_catboost,model_knn),
meta_model = GBoost)
#score of stacked_averaged_models
```

```
score = scores(stacked_averaged_models)
print("stacked_averaged_models score:",score.mean())

#define AveragingModels
class AveragingModels(BaseEstimator, RegressorMixin,
TransformerMixin):
    def __init__(self, models):
        self.models = models

    # define clones of the original models to fit the
data in
    def fit(self, data, target):
        self.models_ = [clone(model) for model in
self.models]

        # Train cloned base models
        for model in self.models_:
            model.fit(data, target)
        return self

    # do the predictions for cloned models and average
them
    def predict(self, data):
        predictions = np.column_stack([
            model.predict(data) for model in self.models_
        ])
        return np.mean(predictions, axis=1)

#average all of models with AveragingModels function
averaged_models = AveragingModels(models =
(RF,GBoost,model_xgb,model_catboost,model_knn))#
model_lgb is weak, not be used
```

```
#score of averaged_models
```

```
score = scores(averaged_models)
```

```
RF.fit(train_data.values, train_target.values)
```

```
RF_test_pred = RF.predict(test_data.values)
```

```
GBoost.fit(train_data.values, train_target.values)
```

```
GBoost_test_pred = GBoost.predict(test_data.values)
```

```
model_xgb.fit(train_data.values,  
train_target.values)
```

```
xgb_test_pred = model_xgb.predict(test_data.values)
```

```
model_lgb.fit(train_data.values,  
train_target.values)
```

```
lgb_test_pred = model_lgb.predict(test_data.values)
```

```
model_catboost.fit(train_data.values,  
train_target.values)
```

```
catboost_test_pred
```

```
=model_catboost.predict(test_data.values)
```

```
model_knn.fit(train_data.values,  
train_target.values)
```

```
knn_test_pred = model_knn.predict(test_data.values)
```

```
knn_test_pred = model_knn.predict(test_data.values)
```

```
#averaged_models
```

```
stacked_averaged_models.fit(train_data.values,  
train_target.values)
```

```
stacked_train_pred =
```

```
stacked_averaged_models.predict(train_data.values)
```

```
#stacked_averaged_models
```

```
stacked_averaged_models.fit(train_data.values,  
train_target.values)
```

```
stacked_pred =
```

```
stacked_averaged_models.predict(test_data.values)

random_forest_mae = mean_absolute_error(test_target,
RF_test_pred)
gboost_mae = mean_absolute_error(test_target,
GBoost_test_pred)
xgboost_mae = mean_absolute_error(test_target,
xgb_test_pred)
lightgbm_mae = mean_absolute_error(test_target,
lgb_test_pred)
catboost_mae = mean_absolute_error(test_target,
catboost_test_pred)
knn_mae = mean_absolute_error(test_target,
knn_test_pred)
stacked_mae = mean_absolute_error(test_target,
stacked_pred)

print('Random Forest mean_absolute_error = ',
random_forest_mae)
print('GBoost mean_absolute_error = ', gboost_mae)
print('XGBoost mean_absolute_error = ', xgboost_mae)
print('LightGBM mean_absolute_error = ',
lightgbm_mae)
print('CatBoost mean_absolute_error = ',
catboost_mae)
print('KNN mean_absolute_error = ', knn_mae)
print('stacked_averaged_models MAE = ', stacked_mae)

models_mae = pd.DataFrame({"random forest":
[random_forest_mae],
"gboost": [gboost_mae],
"xgboost": [xgboost_mae],
"lightgbm": [lightgbm_mae],
```

```
    "catboost": [catboost_mae],  
    "ann": [knn_mae],  
    "stacked models": [stacked_mae]})
```

```
models_mae.plot.bar()  
plt.legend(loc='center left', bbox_to_anchor=(1.0,  
0.5))  
plt.show()
```

### Mean Absolute Percentage Error

```
random_forest_mape = np.mean(np.abs((test_target -  
RF_test_pred) / test_target)) * 100  
gboost_mape = np.mean(np.abs((test_target -  
GBoost_test_pred) / test_target)) * 100  
xgboost_mape = np.mean(np.abs((test_target -  
xgb_test_pred) / test_target)) * 100  
lightgbm_mape = np.mean(np.abs((test_target -  
lgb_test_pred) / test_target)) * 100  
catboost_mape = np.mean(np.abs((test_target -  
catboost_test_pred) / test_target)) * 100  
knn_mape = np.mean(np.abs((test_target -  
knn_test_pred) / test_target)) * 100  
stacked_models_mape = np.mean(np.abs((test_target -  
stacked_pred) / test_target)) * 100  
  
print('MAPE Random Forest = ', random_forest_mape)  
print('MAPE XBoost = ', gboost_mape)  
print('MAPE XGBoost = ', xgboost_mape)  
print('MAPE Lightgbm = ', lightgbm_mape)  
print('MAPE Catboost = ', catboost_mape)  
print('MAPE KNN = ', knn_mape)  
print('MAPE Stacked Models = ', stacked_models_mape)
```



```
models_mape = pd.DataFrame({"random forest":
[random_forest_mape],
    "gboost": [gboost_mape],
    "xgboost": [xgboost_mape],
    "lightgbm": [lightgbm_mape],
    "catboost": [catboost_mape],
    "knn": [knn_mape],
    "stacked models":
    [stacked_models_mape]})

models_mape.plot.bar()

plt.legend(loc='center left', bbox_to_anchor=(1.0,
0.5))
plt.show()

#XG Boost Algorithm

target, unwanted_features = select_mode(MODE)
train, test = prepare_data(data,
    target,
    delete_nans=True,
    convert_categorical=True,
    normalize=False)
train_data, train_target, test_data, test_target =
select_data(train, test, target, unwanted_features)

dtrain = xgb.DMatrix(train_data, label=train_target)
dtest = xgb.DMatrix(test_data, label=test_target)

## Hyperparameteroptimierung für XGBoost
```

```
# Fixe Parameter
params = {"max_depth": 6,
          "min_child_weight": 1,
          "eta": 0.3,
          "subsample": 1,
          "colsample_bytree": 1,
          "eval_metric": "mae",
          "objective": "reg:squarederror"
        }

# Zu optimierende Parameter
gridsearch_params = {
    "max_depth": range(3, 12),
    "min_child_weight": range(3, 8),
    "subsample": [i/10. for i in range(7, 11)],
    "colsample_bytree": [i/10. for i in range(7,
    11)],
    "eta": [.3, .2, .1, .05, .01, .005]
}

num_boost_round = 999
nfold = 5

# Anzahl der Runden, nach denen ohne Verbesserung
abgebrochen wird
early_stopping_rounds = 10

def tune_parameter_set(gridsearch_parameter, params,
data,
                        param_name_1,
                        param_name_2=None):
    if param_name_2 is None:
        gridsearch_params =
```

```
        gridsearch_parameter[param_name_1]
else:
    gridsearch_params = [
        (param_1, param_2)
        for param_1 in
            gridsearch_parameter[param_name_1]
        for param_2 in
            gridsearch_parameter[param_name_2]
    ]
min_loss = float("Inf")
best_params = None
for param_set in gridsearch_params:
    params[param_name_1] = param_set
    if param_name_2 is not None:
        params[param_name_1] = param_set[0]
        params[param_name_2] = param_set[1]
    cv_results = xgb.cv(
        params,
        data,
        num_boost_round = num_boost_round,
        seed = SEED,
        nfold = nfold,
        metrics = {params["eval_metric"]},
        early_stopping_rounds =
            early_stopping_rounds
    )
    mean_loss = cv_results["test-" +
        params["eval_metric"] + "-mean"].min()
    boost_rounds = np.argmin(np.array(cv_results["test-" +
        params["eval_metric"] +
        "-mean"]))
    if mean_loss < min_loss:
        min_loss = mean_loss
```

```
        best_params = param_set

print("Best params: {}, {}:
{}.format(best_params, params["eval_metric"],
min_loss))
params[param_name_1] = best_params
if param_name_2 is not None:
    params[param_name_1] = best_params[0]
    params[param_name_2] = best_params[1]
return params

def tune_parameter(gridsearch_params, params,
dtrain):
    params = tune_parameter_set(gridsearch_params,
params, dtrain,
                                "max_depth",
                                "min_child_weight")
    params = tune_parameter_set(gridsearch_params,
params, dtrain,
                                "subsample",
                                "colsample_bytree")
    params = tune_parameter_set(gridsearch_params,
params, dtrain, "eta")
    return params

### Tuning Parameter

params = tune_parameter(gridsearch_params, params,
dtrain)

## Training des Modells

def train_xgb(params, train, test, num_boost_round,
```

```
early_stopping_rounds):
    model = xgb.train(
        params,
        train,
        num_boost_round = num_boost_round,
        evals = [(test, "Test")],
        early_stopping_rounds =
            early_stopping_rounds
    )
    best_boost_round = model.best_iteration + 1
    best_model = xgb.train(
        params,
        train,
        num_boost_round = best_boost_round,
        evals=[(test, "Test")]
    )
    return best_model
```

*# Eigentliches Model Training*

```
model = train_xgb(params, dtrain, dtest,
                 num_boost_round,
                 early_stopping_rounds)
```

*# Berechnung des MAE mit den Testdaten*

```
from sklearn.metrics import mean_absolute_error
```

```
xgboost_tuned_mae =
mean_absolute_error(model.predict(dtest),
dtest.get_label())
print("MAE Modell {}: {}".format(MODE,
xgboost_tuned_mae))
```

```
# models_mae =
models_mae.join(pd.DataFrame({"xgboost_tuned":
[xgboost_tuned_mae]}))
# print(models_mae)

### Speichern und Laden von Modellen

# Model speichern
model.save_model("model_{}.model".format(MODE))

# Model laden
loaded_model = xgb.Booster()
loaded_model.load_model("model_{}.model".format
(MODE))
#loaded_model.predict(dtest)

model = loaded_model

#Erklärbarkeit der Modelle

import shap

shap.initjs()
#rf_shap_values =
shap.KernelExplainer(model.predict, dtest)

explainer = shap.TreeExplainer(model)
shap_values = explainer.shap_values(train_data)
shap.summary_plot(shap_values, train_data)

### Dependence Plot
```

```
shap.dependence_plot("Lotarea", shap_values,
train_data, interaction_index=None, show=False)

plt.title("Lot Area")
plt.ylabel("SHAP value for the 'Merkmal X' feature")
# plt.savefig("my_dependence_plot.pdf")
plt.show()

shap.dependence_plot("Lotarea", shap_values,
train_data, interaction_index="ActualsalesPricesOutput", show=False)

plt.title("Lot Area and Actual Sales Prices Output")
plt.ylabel("SHAP value for the 'Merkmal X' feature")
# plt.savefig("my_dependence_plot.pdf") # we can
save a PDF of the figure if we want
plt.show()

inds = shap.approximate_interactions("Totalfloorareaofthebuilding", shap_values, train_data)

for i in range(3):
    shap.dependence_plot("Total floor area of the
    building", shap_values, train_data,
    interaction_index=inds[i])
```





## **D Anhang zu Kapitel 7**

### **D.1 Datensatz aus New York, USA**

**Tabelle D.1:** Datensatz von Investitionsprojekten in New York City Office of Operations (2018)

Spaltenname	Spaltenbeschreibung
Berichtszeitpunkt	Berichtszeitpunkt (mm/jjjj)
Projektname	Projektbezeichnung
Beschreibung	Beschreibung der Projektdetails Der Zweck des Projekts: Kunst und Kultur; Brücken; Brücken, Straßen und Wege; Fähren; Industrielle Entwicklung; Industrielle Entwicklung, öffentliche Sicherheit und Strafjustiz; Industrielle Entwicklung, Straßen und Wege; Bibliotheken;
Kategorie	Andere staatliche Einrichtungen; Parks; Parks, Straßen und Wege; Öffentliche Sicherheit und Strafjustiz; Abwasserentsorgung; Schulen; Kanalisation; Kanalisation, Straßen und Wege, Wasserversorgung; Straßen und Wege; Straßen und Wege, Wasserversorgung; Abwasserbehandlung, Wasserversorgung
Gemeinde	Gemeinde, in der das Projekt gebaut wird, falls zutreffend
Verwaltungsbehörde	Behörde, die die Projektentwicklung verwaltet
Kundenagentur	Die Agentur, die der Geschäftsinhaber/Sponsor für das Projekt ist
Aktuelle Phase	Die Projektphase zum Zeitpunkt des Berichtszeitraums
Entwurfsbeginn	Das Datum des Projektbeginns
Budgetprognose	Die zum Zeitpunkt des Berichtszeitraums geschätzten Gesamtkosten des Projekts
Letzte Budgetänderung	Anzahl der Dollar, um die das Projekt seit der letzten Berichterstattung über oder unter dem Budget liegt
Gesamtbudgetänderungen	Anzahl der Dollar, um die das Projekt seit dem Datum des Planungs- beginns das Budget über- oder unterschritten hat
Prognostizierte Fertigstellung	Das voraussichtliche Datum für den Projektabschluss
Letzte Terminplanänderung	Anzahl der Tage, um die das Projekt dem Zeitplan voraus oder hinterher ist seit der letzten Meldung der Daten
Gesamte Terminplanänderungen	Anzahl der Tage, die das Projekt seit dem Datum des Entwurfsbeginns vor oder hinter dem Zeitplan liegt

## D.2 Python Code

```
## Import Daten
```

```
from google.colab import drive
drive.mount('/content/drive')
```

```
import pandas as pd
import numpy as np

### Original Data
# https://data.cityofnewyork.us/City-Government/
Capital-Projects/n7gv-k5yt
data_raw = pd.read_csv("drive/MyDrive/Daten/
New York Raw Data/Capital_Projects.csv", sep=",")

# Data

## Preprocessing

data1 = data_raw.copy()

data1['Design Start'].replace('', np.nan, inplace=True)
data1['Forecast Completion'].replace('', np.nan, inplace=True)

data1.dropna(subset=['Design Start'], inplace=True)
data1.dropna(subset=['Forecast Completion'], inplace=True)

data1 = data1.drop(columns=['Project Name', 'Description'])

def stats(df):
    print(f"Shape: {df.shape}")
    print(f"Anzahl Projekte: {len(df)}")
    print(f"features: {list(df.columns)}")

stats(data1)

# Date Format
data1['Forecast Completion'] = pd.to_datetime(data1
['Forecast Completion'], format="%m/%d/%Y")
```

```
data1['Design Start'] = pd.to_datetime(data1['Design Start'],
format="%m/%d/%Y")
```

```
data1['Date Reported As Of'] = pd.to_datetime(data1
['Date Reported As Of'], format="%m/%d/%Y", exact=False)
```

```
data1['Duration1'] = data1['Forecast Completion'] -
data1['Design Start']
data1['Duration'] = data1['Duration1'].dt.days
```

```
data1 = data1.drop(columns=['Duration1'])
data1.head()
```

```
## Exploration
```

```
print(f"Number of Projects: {len(data1['PID'].unique())}")
```

```
# Group Projects by labels
```

```
categories = data1.groupby(by='PID').first()['Category'].
value_counts().index.to_list()
print(f"Categories: {len(categories)}")
```

```
categories_water_labels = ['Sewers', 'Wastewater Treatment',
'Water Supply', 'Sanitation']
```

```
categories_streets_labels = ['Bridges', 'Streets and Roadways',
'Bridges, Streets and Roadways', 'Parks, Streets and
Roadways', 'Industrial Development, Streets and Roadways']
```

```
categories_hochbau_labels = [ 'Libraries', 'Schools', 'Other
Government Facilities'] + ['Arts and Culture', 'Industrial
Development', 'Health and Hospitals',
'Public Safety and Criminal Justice']
```

```
categories_others_labels = [ 'Ferries', 'IT', 'Health and Hospitals,
IT', 'Parks', 'Social Services', 'IT, Public Safety and Criminal
```

```
Justice', 'Industrial Development, Parks', 'Social Services', 'IT,  
Public Safety and Criminal Justice']
```

```
#Projects per Category
```

```
for category_name, category_labels in zip(['Water', 'Streets',  
'Hochbau', 'Others'], [categories_water_labels,  
categories_streets_labels, categories_hochbau_labels,  
categories_others_labels]):  
    count = len(data1[data1['Category'].isin(category_labels)].  
groupby (by="PID").first())  
    print(f"{category_name}: {count}")
```

```
## Clean Data
```

```
data1.head(5)
```

```
data_final = data1.copy()
```

```
# sort by PID and Date Reported as of and take last
```

```
data_final = data_final.sort_values(by=['PID', 'Date Reported As  
Of'], ascending= [True, True])
```

```
#
```

```
data_final = data_final.drop_duplicates(subset=['PID'],  
keep='last')
```

```
data_final.head(5)
```

```
print("The average of 'Duration' is:")
```

```
print(data_final['Duration'].mean())
```

```
print("The standard variation of 'Duration' is:")
```

```
print(data_final['Duration'].std())
```

```
data_initial = data1.copy()

# sort by PID and Date Reported as of and take last
data_initial = data_initial.sort_values(by=['PID', 'Date
Reported As Of'], ascending= [True, True])
#
data_initial = data_initial.drop_duplicates(subset=['PID'],
keep='first')

data_initial = data_initial.drop(columns=['Date Reported As Of',
'Category', 'Borough', 'Managing Agency', 'Client Agency',
'Design Start', 'Latest Budget Changes',
'Total Budget Changes', 'Forecast Completion',
'Latest Schedule Changes', 'Total Schedule Changes'])

data_initial.head(5)

## Data sets creation

data2 = data_final.drop(columns=['Latest Budget Changes',
'Design Start', 'Total Budget Changes', 'Latest Schedule
Changes', 'Total Schedule Changes'])

print('Vor One Hot Codierung: ' + str(data2.shape))

# One Hot Codierung
data2 = pd.get_dummies(data2)

#Ausgabe der Daten nach One Hot Codierung
print('Nach One Hot Codierung: ' + str(data2.shape))

data_external = pd.read_csv("drive/MyDrive/Daten/New
York Raw Data/Externe Daten_USA.csv", sep=";")
```

```
data_external.head(5)

data_final['year'] = data_final['Design Start'].dt.year

from pandas import read_csv
from numpy import nan

print('Vor One Hot Codierung: ' + str(data_final.shape))

# One Hot Codierung
data_final = pd.get_dummies(data_final)

#Ausgabe der Daten nach One Hot Codierung
print('Nach One Hot Codierung: ' + str(data_final.shape))

data_final.head()

import pandas as pd
data3 = pd.merge(data_final, data_external, on="year",
how="inner")
#data3 = data3.drop(columns=['Design Start'])
data3.head()

from pandas import read_csv
from numpy import nan

print('Vor One Hot Codierung: ' + str(data3.shape))

# One Hot Codierung
data3 = pd.get_dummies(data3)

#Ausgabe der Daten nach One Hot Codierung
```

```
print('Nach One Hot Codierung: ' + str(data3.shape))

data4 = data3.drop(columns=['Latest Budget Changes',
'Total Budget Changes', 'Latest Schedule Changes',
'Total Schedule Changes'])

print('Vor One Hot Codierung: ' + str(data4.shape))

# One Hot Codierung
data4 = pd.get_dummies(data4)

#Ausgabe der Daten nach One Hot Codierung
print('Nach One Hot Codierung: ' + str(data4.shape))

#Übersicht Datensätze
## data_final: **ohne** externe Daten und **mit**
#Informationen zu Änderungen
## data2: **ohne** externe Daten und **ohne**
#Informationen zu Änderungen
## data3: **mit** externe Daten und **mit**
#Informationen zu Änderungen
## data4: **mit** externe Daten und **ohne**
#Informationen zu Änderungen

#Features:
##- **Data Reported As Of:** Reporting period (mm/yyyy)
##- **Project Name:** Project Name
##- **Description:** Description of the project details
##- **Category:** The purpose of type of project
##- **Borough:** Borough where the project is being built
##- **Managing Agency:** Agency that is managing the project
#development
##- **Client Agency:** The agency that is the business
```



```
#owner/sponsor for the project
#- **Current Phase:** The project phase at the point
#in time of the reporting period
#- **Design Start Date:** The date the project started
#- **Budget Forecast:** The total cost of the project
#estimated at the pointin time of the reporting
#- **Latest Budget Change:** Number of dollars the project
#is over or under budgetsince the last time the data was reported
#- Total Budget Changes: Number of dollars the project is
# over or under budgetsince the design start date
#- **Forecast Completion:** The date projected for
#project completion
#- **Latest Schedule Change:** Number of days the project
#is ahead or behind schedule since the last time the data
#was reported
#- **TotalSchedule Changes:** Number of days the project
#is ahead or behind schedulesince the design start date

#Comparison

data = data3.copy()

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline

hplot = sns.distplot(data['Duration'], kde=False, color='blue',
bins=100)
plt.title('Duration', fontsize = 18)
plt.xlabel('Duration', fontsize = 18)
plt.ylabel('Duration', fontsize = 18)
```

```
import seaborn as sns

#Top 25 in der Heatmap
k = 26 #include target 'Duration'
corrmat = data.corr()
cols = corrmat.nlargest(k, 'Duration')['Duration'].index
cm = np.corrcoef(data[cols].values.T)
sns.set(font_scale = 1.6)
plt.figure(figsize = (30,20))
hm = sns.heatmap(cm, cbar=True,
annot=True,square=True,fmt='.2f',cmap='YlGnBu',linewidths=.5,
annot_kws={'size':
16},yticklabels=cols.values,xticklabels=cols.values)
plt.show()

data.head()

data = data.drop(columns=['PID', 'Date Reported As Of', 'Forecast
Completion', 'Design Start'])
#'Design Start', 'Forecast Completion' 'Date Reported As Of',
'Design Start', 'Forecast Completion'

X = data.drop(columns=['Duration'])
y = data['Duration']

X.isnull().sum()
X.fillna(-999, inplace = True)
y.fillna(-999, inplace = True)

# Splitting the dataset into the Training set and Test set
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split
```

```
(X, y, test_size = 0.2, random_state = 1)

##Random Forest

# Fitting Random Forest Regression to the dataset
# import the regressor
from sklearn.ensemble import RandomForestRegressor

# create regressor object
regressor = RandomForestRegressor(n_estimators = 100,
random_state = 1)

# Train the model on training data
regressor.fit(X_train, y_train)

# Predicting the Test set results
y_pred = regressor.predict(X_test)

# Evaluating the Algorithm
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error
(y_test, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error
(y_test, y_pred))
print('Root Mean Squared Error:',
np.sqrt(metrics.mean_squared_error(y_test, y_pred)))

mape = np.mean(np.abs((y_test - y_pred)/y_test))*100
print("MAPE in y units:", np.mean(mape))

import numpy as np
#from sklearn.metrics import
symmetrical_mean_absolute_percentage_error as smape
```

```
smape = 1/len(y_test) * np.sum(2 * np.abs(y_pred-y_test) /
(np.abs(y_test) + np.abs(y_pred))*100)
print("SMAPE in y units:", np.mean(smape))

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer

scores = cross_validate(regressor, X, y, cv=3,
                        scoring=('r2', 'neg_mean_squared_error'),
                        return_train_score=True)
print('MSE:', scores['test_neg_mean_squared_error'].mean())

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer

scores = cross_validate(regressor, X, y, cv=3,
                        scoring=('r2', 'neg_mean_absolute_error'),
                        return_train_score=True)
print('MAE:', scores['test_neg_mean_absolute_error'].mean())

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer

scores = cross_validate(regressor, X, y, cv=3,
                        scoring=('r2',
                                'neg_mean_absolute_percentage_error'),
                        return_train_score=True)
print('MAPE:', scores['test_neg_mean_absolute_
percentage_error'].mean())

##Gradient Boosting Regression

from sklearn.ensemble import GradientBoostingRegressor
```

```
# with new parameters
gbr = GradientBoostingRegressor(n_estimators=100,
                                max_depth=3,
                                learning_rate=0.01,
                                min_samples_split=3)

gbr.fit(X_train, y_train)

y_pred = gbr.predict(X_test) # Predictions

# Evaluating the Algorithm
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error
      (y_test, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error
      (y_test, y_pred))
print('Root Mean Squared Error:',
      np.sqrt(metrics.mean_squared_error(y_test, y_pred)))

mape = np.mean(np.abs((y_test - y_pred)/y_test))*100
print("MAPE in y units:", np.mean(mape))

import numpy as np
#from sklearn.metrics import
symmetrical_mean_absolute_percentage_error as smape
smape = 1/len(y_test) * np.sum(2 * np.abs(y_pred-y_test) /
                              (np.abs(y_test) + np.abs(y_pred))*100)
print("SMAPE in y units:", np.mean(smape))

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer
```

```
scores = cross_validate(gbr, X, y, cv=3,
                        scoring=('r2', 'neg_mean_squared_error'),
                        return_train_score=True)
print('MSE:', scores['test_neg_mean_squared_error'].mean())

scores = cross_validate(gbr, X, y, cv=3,
                        scoring=('r2', 'neg_mean_absolute_error'),
                        return_train_score=True)
print('MAE:', scores['test_neg_mean_absolute_error'].mean())

scores = cross_validate(gbr, X, y, cv=3,
                        scoring=('r2',
                                'neg_mean_absolute_percentage_error'),
                        return_train_score=True)
print('MAPE:', scores['test_neg_mean_absolute_
percentage_error'].mean())

## XGBoost

import xgboost as xgb

# Instantiate an XGBoost object with hyperparameters
xgb_reg = xgb.XGBRegressor(max_depth=3, n_estimators=100, n_jobs=2,
                           objective='reg:squarederror',
                           booster='gbtree',
                           random_state=1, learning_rate=0.05)

# Train the model with train data sets
xgb_reg.fit(X_train, y_train)

y_pred = xgb_reg.predict(X_test) # Predictions

# Evaluating the Algorithm
```

```
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error
      (y_test, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error
      (y_test, y_pred))
print('Root Mean Squared Error:',
      np.sqrt(metrics.mean_squared_error(y_test, y_pred)))

mape = np.mean(np.abs((y_test - y_pred)/y_test))*100
print("MAPE in y units:", np.mean(mape))

import numpy as np
#from sklearn.metrics import
symmetrical_mean_absolute_percentage_error as smape
smape = 1/len(y_test) * np.sum(2 * np.abs(y_pred-y_test) /
                               (np.abs(y_test) + np.abs(y_pred))*100)
print("SMAPE in y units:", np.mean(smape))

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer

scores = cross_validate(xgb_reg, X, y, cv=3,
                       scoring=('r2', 'neg_mean_squared_error'),
                       return_train_score=True)
print('MSE:', scores['test_neg_mean_squared_error'].mean())

scores = cross_validate(xgb_reg, X, y, cv=3,
                       scoring=('r2', 'neg_mean_absolute_error'),
                       return_train_score=True)
print('MAE:', scores['test_neg_mean_absolute_error'].mean())

scores = cross_validate(xgb_reg, X, y, cv=3,
                       scoring=('r2',
```

```
        'neg_mean_absolute_percentage_error'),
        return_train_score=True)
print('MAPE:', scores['test_neg_mean_absolute_
percentage_error'].mean())

##LightGBM

import lightgbm as lgb

model = lgb.LGBMRegressor(learning_rate=0.09,max_depth=3,
random_state=1)
model.fit(X_train,y_train,eval_set=[(X_test,y_test),
(X_train,y_train)],
        verbose=20, eval_metric='logloss')
# Predicting the Test set results
y_pred = model.predict(X_test)

# Evaluating the Algorithm
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error
(y_test, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error
(y_test, y_pred))
print('Root Mean Squared Error:',
np.sqrt(metrics.mean_squared_error(y_test, y_pred)))

mape = np.mean(np.abs((y_test - y_pred)/y_test))*100
print("MAPE in y units:", np.mean(mape))

import numpy as np
#from sklearn.metrics import
symmetrical_mean_absolute_percentage_error as smape
smape = 1/len(y_test) * np.sum(2 * np.abs(y_pred-y_test) /
```



```
(np.abs(y_test) + np.abs(y_pred))*100
print("SMAPE in y units:", np.mean(smape))

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer

scores = cross_validate(model, X, y, cv=3,
                        scoring=('r2', 'neg_mean_squared_error'),
                        return_train_score=True)
print('MSE:', scores['test_neg_mean_squared_error'].mean())

scores = cross_validate(model, X, y, cv=3,
                        scoring=('r2', 'neg_mean_absolute_error'),
                        return_train_score=True)
print('MAE:', scores['test_neg_mean_absolute_error'].mean())

scores = cross_validate(model, X, y, cv=3,
                        scoring=('r2',
                                'neg_mean_absolute_percentage_error'),
                        return_train_score=True)
print('MAPE:', scores['test_neg_mean_absolute_
percentage_error'].mean())

lgb.plot_tree(model,figsize=(30,40))

##Regression

from sklearn.linear_model import LinearRegression
# importieren der Klasse

lr = LinearRegression()
# instanzieren der Klasse
```

```
lr.fit(X_train, y_train)
# trainieren

y_pred = lr.predict(X_test)

# Evaluating the Algorithm
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error
      (y_test, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error
      (y_test, y_pred))
print('Root Mean Squared Error:',
      np.sqrt(metrics.mean_squared_error(y_test, y_pred)))

mape = np.mean(np.abs((y_test - y_pred)/y_test))*100
print("MAPE in y units:", np.mean(mape))

import numpy as np
#from sklearn.metrics import
symmetrical_mean_absolute_percentage_error as smape
smape = 1/len(y_test) * np.sum(2 * np.abs(y_pred-y_test) /
                               (np.abs(y_test) + np.abs(y_pred))*100)
print("SMAPE in y units:", np.mean(smape))

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer

scores = cross_validate(lr, X, y, cv=3,
                       scoring=('r2', 'neg_mean_squared_error'),
                       return_train_score=True)
print('MSE:', scores['test_neg_mean_squared_error'].mean())

scores = cross_validate(lr, X, y, cv=3,
```

```
        scoring=('r2', 'neg_mean_absolute_error'),
        return_train_score=True)
print('MAE:', scores['test_neg_mean_absolute_error'].mean())

scores = cross_validate(lr, X, y, cv=3,
                        scoring=('r2',
                                'neg_mean_absolute_percentage_error'),
                        return_train_score=True)
print('MAPE:', scores['test_neg_mean_absolute_
percentage_error'].mean())

##ANN

from sklearn.neural_network import MLPClassifier

model_ann = MLPClassifier(hidden_layer_sizes=(100,100,100),
                           max_iter=500, alpha=0.0001,
                           solver='sgd', verbose=10, random_state=1,
                           tol=0.000000001)

# Fitting the ANN to the Training set
model_ann.fit(X_train, y_train)

# Generating Predictions on testing data
y_pred = model_ann.predict(X_test)

# Evaluating the Algorithm
from sklearn import metrics
print('Mean Absolute Error:', metrics.mean_absolute_error
(y_test, y_pred))
print('Mean Squared Error:', metrics.mean_squared_error
(y_test, y_pred))
print('Root Mean Squared Error:',
```

```
np.sqrt(metrics.mean_squared_error(y_test, y_pred)))

mape = np.mean(np.abs((y_test - y_pred)/y_test))*100
print("MAPE in y units:", np.mean(mape))

import numpy as np
#from sklearn.metrics import
symmetrical_mean_absolute_percentage_error as smape
smape = 1/len(y_test) * np.sum(2 * np.abs(y_pred-y_test) /
(np.abs(y_test) + np.abs(y_pred))*100)
print("SMAPE in y units:", np.mean(smape))

from sklearn.model_selection import cross_validate
from sklearn.metrics import make_scorer

scores = cross_validate(model_ann, X, y, cv=3,
                        scoring=('r2', 'neg_mean_squared_error'),
                        return_train_score=True)
print('MSE:', scores['test_neg_mean_squared_error'].mean())

scores = cross_validate(model_ann, X, y, cv=3,
                        scoring=('r2', 'neg_mean_absolute_error'),
                        return_train_score=True)
print('MAE:', scores['test_neg_mean_absolute_error'].mean())

scores = cross_validate(model_ann, X, y, cv=3,
                        scoring=('r2',
                                'neg_mean_absolute_percentage_error'),
                        return_train_score=True)
print('MAPE:', scores['test_neg_mean_absolute_
percentage_error'].mean())

##Tree Explainer SHAP
```

```
!pip install shap
import shap
import shap

shap.initjs()
#rf_shap_values = shap.KernelExplainer(model.predict, dtest)

explainer = shap.TreeExplainer(model)
shap_values = explainer.shap_values(X_train)
shap.summary_plot(shap_values, X_train)

# Catboost

## Model

data = data3.copy()
data.head()

data = data.drop(columns=['Date Reported As Of',
'Forecast Completion', 'Design Start'])

X = data.drop(columns=['Duration'])
y= data['Duration']

X.isnull().sum()
X.fillna(-999, inplace = True)
y.fillna(-999, inplace = True)

!pip install catboost
import catboost
from catboost import CatBoostRegressor, Pool
```

```
cb = CatBoostRegressor(n_estimators=100,
                       loss_function='RMSE',
                       learning_rate=0.4,
                       depth=3, task_type='CPU',
                       random_state=1,
                       verbose=False)

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split
(X, y, test_size=0.20, random_state=1)

#data2["Category"] = data2["Category"].astype(np.int)
#train["Current Phase"] = train["Current Phase"].astype('int64')

X.dtypes

categorical_features_indices = np.where
(X.dtypes != np.float)[0] #X_train

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split
(X, y, test_size=0.20, random_state=1)

#pool_train = Pool(X_train, y_train, cat_features = ['Category',
# 'Borough', 'Managing Agency', 'Client Agency', 'Current Phase'])

pool_train = Pool(X_train, y_train,
cat_features=categorical_features_indices)
pool_test = Pool(X_test, cat_features=categorical_features_indices)

#pool_test = Pool(X_test, cat_features = ['Category', 'Borough',
# 'Managing Agency', 'Client Agency', 'Current Phase'])
```

```
cb.fit(pool_train)
y_pred = cb.predict(pool_test)

## Prediction Results CatBoost

import numpy as np
from sklearn.metrics import mean_squared_error as mse

cb_rmse = np.sqrt(mse(y_test, y_pred))
print("RMSE in y units:", np.mean(cb_rmse))

import numpy as np
from sklearn.metrics import mean_absolute_error as mae

mae = np.absolute(np.subtract(y_test, y_pred)).mean()
print("MAE in y units:", np.mean(mae))

import numpy as np
from sklearn.metrics import mean_absolute_percentage_error
as mape

mape = np.mean(np.abs((y_test - y_pred)/y_test))*100
print("MAPE in y units:", np.mean(mape))

import numpy as np
#from sklearn.metrics import
symmetrical_mean_absolute_percentage_error as smape

smape = 1/len(y_test) * np.sum(2 * np.abs(y_pred-y_test) /
(np.abs(y_test) + np.abs(y_pred))*100)
print("SMAPE in y units:", np.mean(smape))

from sklearn.model_selection import cross_validate
```

```
from sklearn.metrics import make_scorer

scores = cross_validate(cb, X, y, cv=10,
                       scoring=('r2', 'neg_mean_squared_error'),
                       return_train_score=True)
print('MSE:', scores['test_neg_mean_squared_error'].mean())

scores = cross_validate(cb, X, y, cv=10,
                       scoring=('r2', 'neg_mean_absolute_error'),
                       return_train_score=True)
print('MAE:', scores['test_neg_mean_absolute_error'].mean())

scores = cross_validate(cb, X, y, cv=10,
                       scoring=('r2',
                                'neg_mean_absolute_percentage_error'),
                       return_train_score=True)
print('MAPE:', scores['test_neg_mean_absolute_
percentage_error'].mean())

y.describe()

## Hyperparameter Optimization

data_ = data2.copy()

from sklearn import datasets
from catboost import CatBoostRegressor
from sklearn.model_selection import GridSearchCV

#create the grid
grid = {'max_depth': [3,4,5], 'n_estimators':[100, 200, 300]}

X_ = data_.drop(columns=['Duration'])
```



```
y_ = data_['Duration']

X_.isnull().sum()
X_.fillna(-999, inplace = True)
y_.fillna(-999, inplace = True)
categorical_features_indices = np.where
(X_.dtypes != np.float)[0] #X_train

cbr = CatBoostRegressor()

GridSearchCV

#Instantiate GridSearchCV
gscv = GridSearchCV (estimator = cbr, param_grid = grid,
scoring = 'accuracy', cv = 5)

#fit the model
gscv.fit(X_,y_)

#returns the estimator with the best performance
print(gscv.best_estimator_)

#returns the best score
print(gscv.best_score_)

#returns the best parameters
print(gscv.best_params_)

# Comparison of Machine and Human Prediction

##Preparation

prediction = cb.predict(X)
```

```
data['Duration_by_Machine'] = prediction

## import dependencies
from sklearn import tree #For our Decision Tree
import pandas as pd # For our DataFrame
import pydotplus # To create our Decision Tree Graph
from IPython.display import Image
# To Display a image of our graph

data6 = data.copy()
print(data6)
print(len(data6.columns))

data_initial.rename(columns={'Duration':'Duration_by_Human'},
inplace=True)
data_initial.rename(columns={'Budget Forecast':'Budget
Forecast_initial'}, inplace=True)
data_initial.rename(columns={'Current Phase':'Current
Phase_initial'}, inplace=True)

data6.rename(columns={'Duration':'Duration_actual'},
inplace=True)
data6.rename(columns={'Budget Forecast':
'Budget Forecast_current'}, inplace=True)
data6.rename(columns={'Current Phase':'Current Phase_current'},
inplace=True)
data6.head()

import pandas as pd
data5 = pd.merge(data6, data_initial, on="PID", how="inner")
data5.head()

data5['Duration_by_Human'].describe()
```

```
data5['Duration_actual'].describe()

data5['Human Error'] = abs(data5['Duration_by_Human'] - data5
['Duration_actual'])
data5['Human Error'].describe()

data5['Machine Error'] = abs(data5['Duration_by_Machine'] -
data5 ['Duration_actual'])
data5['Machine Error'] .describe()

data5 = data5.drop(columns = ['Machine Error',
'Human Error', 'PID'])

data5['Machine_Human_Difference'] = data5['Machine Error'] -
data5['Human Error']

data5['Machine_Human_Difference'] .describe()

data5.loc[data5['Machine_Human_Difference'] <= 0,
'Machine_Human'] = '0'
data5.loc[data5['Machine_Human_Difference'] > 0,
'Machine_Human'] = '1'
#1= Machine, 0=Human
print(data5)

data5.head(5)

data5['Machine_Human'].value_counts()

data5['Machine_Human'].value_counts(normalize=True)

(data5[['Category_Streets and
```

```
Roadways'])*data5['Machine_Human']).sum(axis=1)
```

```
data5['sumproduct'] = data5['Category_Streets and  
Roadways']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] =  
data5['Category_Sewers']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] = data5['Category_Industrial  
Development']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] = data5['Category_Water  
Supply']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] = data5['Category_Wastewater  
Treatment']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] =  
data5['Category_Schools']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] =  
data5['Category_Bridges']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] = data5['Category_Parks']*  
data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] = data5['Category_Public Safety and  
Criminal Justice']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] = data5['Category_Bridges, Streets  
and Roadways']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5['sumproduct'] = data5['Category_Other Government  
Facilities']*data5['Machine_Human']  
data5['sumproduct'].value_counts()
```

```
data5 = data5.drop(columns=['sumproduct'])
```

```
##Residual Plots
```

```
#Plotten von Predicted (by AI und by Humans) vs. Actual Values,  
#um zu schauen was die Modelle genau machen. Haben  
#die Ansätzen die Tendenz über bzw. unter zuschätzen: z.B.  
#https://towardsdatascience.com/how-to-use-residual-plots-  
#for-regression-model-validation-c3c70e8ab378
```

```
#import necessary libraries
```

```
import matplotlib.pyplot as plt  
import statsmodels.api as sm  
from statsmodels.formula.api import ols
```

```
#fit simple linear regression model
```

```
model = ols('Duration_by_Machine ~ Duration_actual',  
data=data5).fit()
```

```
#view model summary
```

```
print(model.summary())

#define figure size
fig = plt.figure(figsize=(12,8))

#produce regression plots
fig = sm.graphics.plot_regress_exog(model, 'Duration_actual',
fig=fig)

#import necessary libraries
import matplotlib.pyplot as plt
import statsmodels.api as sm
from statsmodels.formula.api import ols

#fit simple linear regression model
model = ols('Duration_by_Human ~ Duration_actual',
data=data5).fit()

#view model summary
print(model.summary())

#define figure size
fig = plt.figure(figsize=(12,8))

#produce regression plots
fig = sm.graphics.plot_regress_exog(model, 'Duration_actual',
fig=fig)

## Influencing factors

import pandas as pd

data12= data5.copy()
```

```
del data12['Machine_Human_Difference']
del data12['Duration_by_Machine']
del data12['Duration_by_Human']
#data12.drop(columns=['Duration_by_Machine', 'Duration_by_Human'])
# 'Duration_by_Machine', 'Duration_by_Human'

data12.fillna(data.mean(), inplace=True)
categorical_features_indices = np.where(X.dtypes
!= np.float)[0] #X_train

data12.head()

def select_data(train, test, target, unwanted_features=[]):
    if target not in unwanted_features:
        unwanted_features.append(target)
    train_target = train.loc[:, target]
    test_target = test.loc[:, target]
    for feature in unwanted_features:
        if feature in train:
            train_data = train.drop(columns=[feature])
            test_data = test.drop(columns=[feature])
    return train_data, train_target, test_data, test_target

def select_mode(mode):
    duration = "Machine_Human"
    #duration = "Change"
    cost = "BudgetForecast"
    target_1a = target_2a = duration
    target_1b = target_2b = cost
    cost_dependent_features = ["BudgetForecast", cost]

    # 1a) Vorhersage der Dauer exklusive Kosten in Trainingsdaten
    if mode == "1a":
```

```
    return target_1a, cost_dependent_features

# 1b) Vorhersage der Kosten exklusive Dauer in Trainingsdaten
elif mode == "1b":
    return target_1b, ([duration] + cost_dependent_features)

# 2a) Vorhersage der Dauer inklusive Kosten in Trainingsdaten
elif mode == "2a":
    return target_2a, []

# 2b) Vorhersage der Kosten exklusive Dauer in Trainingsdaten
elif mode == "2b":
    return target_2b, cost_dependent_features

else:
    raise ValueError(str(mode) + " is not a valid mode.")

# Trennung Trainings-/ Testdaten
train = data.sample(frac=splitratio, random_state=seed)
test = data.drop(train.index)

return train, test
```

```
SEED = 314
```

```
def prepare_data(data, target, delete_nans=True,
convert_categorical=True, normalize=True, splitratio=0.8,
seed=SEED):
```

```
    if(delete_nans == True):
        #Ausgabe der Daten vor Entfernung der Leerdaten
        print('Mit Leerdaten: ' + str(data.shape))
```



```
#Löschen von Projekten, in denen die Dauer fehlt
data = data.dropna(axis=0, subset=['Duration in
days (real)'])

#Löschen von Projekten, in denen die Dauer 0 ist
data = data[~data['Duration in days (real)'].isin([0])]

#Ausgabe der Daten nach Entfernung der Leerdaten
print('Ohne Leerdaten: ' + str(data.shape))

if(convert_categorical == True):
    #Ausgabe der Daten vor One Hot Codierung
    print('Vor One Hot Codierung: ' + str(data.shape))

    # One Hot Codierung
    data = pd.get_dummies(data)

    #Ausgabe der Daten nach One Hot Codierung
    print('Nach One Hot Codierung: ' + str(data.shape))

if(normalize == True):
    # Normalisieren der Daten [Werte zwischen 0 und 1], außer
    Target
    tmp = data
    data_min = data.min()
    data_max = data.max()
    data = (data - data_min) / (data_max - data_min)
    data[target] = tmp[target]

# Trennung Trainings-/ Testdaten
train = data.sample(frac=splitratio, random_state=seed)
test = data.drop(train.index)
```

```
    return train, test

data12.head()

from sklearn.preprocessing import LabelEncoder
labelencoder= LabelEncoder()

#data12['Current Phase_initial'] = pd.get_dummies(data12['Current
Phase_initial'])

MODE = "2a"

target, unwanted_features = select_mode(MODE)
train, test = prepare_data(data12,
                           target,
                           delete_nans=False,
                           convert_categorical=False,
                           normalize=False)
train_data, train_target, test_data, test_target =
select_data(train, test, target, unwanted_features)

x = train_data
y = train_target

train_data.head()

import numpy as np
import catboost
from catboost import CatBoostClassifier, Pool

# find non-numerical
is_cat = (train_data.dtypes != float)
```

```
# fill in NaNs for non-numericals
for feature, feat_is_cat in is_cat.to_dict().items():
    if feat_is_cat:
        train_data[feature].fillna("NaN", inplace=True)

cat_features_index = np.where(is_cat)[0]
pool = Pool(train_data, train_target,
            cat_features=cat_features_index,
            feature_names=list(train_data.columns))

model = CatBoostClassifier(
    max_depth=6, verbose=False, max_ctr_complexity=1,
    iterations=2).fit(pool)

#model = CatBoostRegressor(
#    # max_depth=2, verbose=False, max_ctr_complexity=1,
#    random_seed=42, iterations=2).fit(pool)

!pip install shap
import shap

#data.drop(columns=['DeviationHuman', 'DeviationMachine'])
shap.initjs()
#rf_shap_values = shap.KernelExplainer(model.predict, dtest)

explainer = shap.TreeExplainer(model)
shap_values = explainer.shap_values(train_data)

shap.summary_plot(shap_values, train_data, max_display = 10)

#binary_shap_values = shap_values
```