

Forschungszentrum Karlsruhe

Technik und Umwelt

Wissenschaftliche Berichte

FZKA 6613

Diagnoseunterstützung für die Instrumentelle Ganganalyse (Projekt GANDI)

Ralf Mikut, Hagen Malberg, Norbert Peter,
Jens Jäkel, Lutz Gröll, Georg Bretthauer,
Rainer Abel*, Leonhard Döderlein*,
Rüdiger Rupp*, Matthias Schablowski*,
Andrea Siebel*, Hans-Jürgen Gerner*

Institut für Angewandte Informatik
Programm Medizintechnik

*Orthopädische Universitätsklinik Heidelberg

Impressum der Print-Ausgabe:

**Als Manuskript gedruckt
Für diesen Bericht behalten wir uns alle Rechte vor**

**Forschungszentrum Karlsruhe GmbH
Postfach 3640, 76021 Karlsruhe**

**Mitglied der Hermann von Helmholtz-Gemeinschaft
Deutscher Forschungszentren (HGF)**

ISSN 0947-8620

Arbeitsbericht zum DFG-Fördervorhaben
BR-1303/6 und GE-1139/1:
Anwendung von Methoden der Computational Intelligence zur
automatisierten Diagnoseerstellung in der Ganganalyse

Zusammenfassung

Die Instrumentelle Ganganalyse ist ein Verfahren zur quantitativen Analyse von Bewegungen. Dabei werden Videodaten von Bewegungsabläufen, gemessene Bodenreaktionskräfte sowie Muskelaktivitäten simultan aufgezeichnet. Die Methodik erlaubt die Erfassung von dynamischen Bewegungsvorgängen, die in ihrer Komplexität durch einfache optische Kontrolle nicht in ausreichendem Umfang ermittelt werden können. Die so gewonnenen Daten werden unter anderem zur Diagnose und Therapieplanung bei neurologisch bedingten Bewegungsstörungen (Querschnittslähmungen, frühkindliche Hirnschäden – Zerebralpareesen) eingesetzt. Gegenwärtig existieren zwar ausgereifte Messsysteme, aber die Auswertung und klinische Interpretation der Daten erfolgt noch manuell, was ein hohes Maß an jeweiligem Expertenwissen erfordert. Im hier beschriebenen Projekt wird am Beispiel der Schrittphasenerkennung versucht, mit Hilfe von statistischen und Fuzzy-Ansätzen diese Auswertung zumindest teilweise zu automatisieren.

Abstract

Diagnosis Support in Clinical Gait Analysis (Project GANDI):

Instrumented gait analysis is a method for quantitative analysis of movements. It is based on the simultaneous measurement of video data of gait, ground reaction forces and electromyographic data. In contrast to purely visual observation, this method allows for a more detailed analysis of complex patterns in the dynamic behaviour during movements. The resulting data is often used in diagnosis and treatment planning of neurologic diseases (Paraplegia, Cerebral Palsy). Currently, highly sophisticated commercial measurement systems are being used for data acquisition. On the other hand, clinical interpretation is solely being done manually and requires long-term experience of highly specialized clinical experts. In this project, a statistical and a fuzzy logic approach for a semi-automatic analysis and interpretation will be investigated for the example of gait phases recognition to overcome this limitation.

Inhaltsverzeichnis

1	Motivation	1
2	Instrumentelle Ganganalyse	3
2.1	Neurologisch bedingte Bewegungsstörungen	3
2.2	Historische Entwicklung	4
2.3	Entwicklungsstand	5
2.4	Messverfahren	6
2.5	Subjektive Bewertung zur Diagnose und Therapieplanung	10
2.5.1	Allgemeines Vorgehen	10
2.5.2	Subjektive Bewertung mit Videodaten	10
2.5.3	Subjektive Bewertung mit Messdaten der Ganganalyse	11
2.6	Ansätze zur automatisierten Bewertung	11
2.7	Schrittphasen	13
3	Datenbasis und Vorverarbeitung	15
3.1	Datenbasis	15
3.2	Signalvorverarbeitung und abgeleitete Merkmale	16
4	Statistische Verfahren	19
4.1	Ziele	19
4.2	Merkmalsselektion	19
4.3	Dimensionsreduzierende Merkmalstransformation	22
4.4	Klassifikation mit Bayes-Klassifikatoren	26
4.5	Interpretation der transformierten Merkmale	28
4.6	Weitere Anwendungen	29
5	Fuzzy-Verfahren	33
5.1	Übersicht	33
5.1.1	Struktur	33
5.1.2	Fuzzifizierung	34
5.1.3	Inferenz	35
5.1.4	Defuzzifizierung	38
5.2	Entwicklungsstand beim Entwurf von Fuzzy-Systemen	39
5.3	Verwendetes Entwurfsverfahren für Fuzzy-Systeme	41
5.4	Entwurf von Zugehörigkeitsfunktionen	42
5.4.1	Ziele	42
5.4.2	Informationstheoretische Maße für Fuzzy-Partitionen	43
5.4.3	Entwurfsverfahren	48

5.5	Merkmalsselektion	50
5.6	Regelgenerierung	54
5.6.1	Ziele	54
5.6.2	Hypothesengenerierung mit Entscheidungsbäumen	55
5.6.3	Regelpruning und Bewertung von Einzelregeln	57
5.7	Suche nach Regelbasen	68
6	Vergleich	72
7	Implementierung	74
8	Zusammenfassung und Ausblick	76
A	Merkmale, Merkmalsrelevanzen und Regeln	78
A.1	Merkmalsbezeichnungen	79
A.2	Univariate statistische Merkmalsrelevanzen	80
A.3	Univariate ID3-Merkmalsrelevanzen	81
A.4	Multivariate ID3-Merkmalsrelevanzen	82
A.5	Zugehörigkeitsfunktionen der ausgewählten Merkmale	83
A.6	Vertrauensintervalle bei Binomialverteilungen	84
A.7	Regelbasis nach Pruning	85
B	Implementierung	90
B.1	Funktion <code>compute_features.m</code>	90
B.2	Wichtige Symbole und Bezeichnungen	92

Kapitel 1

Motivation

Die Instrumentelle Ganganalyse stellt ein Verfahren zur quantitativen Untersuchung und Dokumentation des menschlichen oder tierischen Ganges dar [88]. Dabei werden Videodaten von Bewegungsabläufen, über Kraftmessplatten gemessene Bodenreaktionskräfte sowie unter Umständen Muskelaktivitäten (Elektromyogramme – EMG) und der Sauerstoffverbrauch simultan aufgezeichnet. Die Methodik erlaubt die Erfassung und anschließende Beurteilung von dynamischen Bewegungsvorgängen, die in ihrer Komplexität durch einfache optische Kontrolle nicht in ausreichendem Maße erfasst werden können. Die so gewonnenen Daten werden zur Diagnose und Therapieplanung bei Bewegungsstörungen [6, 67, 88, 96, 99, 101], zur Generierung von Bewegungsmustern bei der funktionellen Elektrostimulation [92], zur Regelung von Prothesen [117] und zur Übertragung der Ergebnisse auf zwei- bzw. mehrbeinige Roboter [22] verwendet.

Bei der Diagnose und Therapieplanung wird die Ganganalyse hauptsächlich bei frühkindlichen Hirnschädigungen (*Infantile Zerebralparese – ICP*, 20.000 Fälle/Jahr in der Bundesrepublik Deutschland), bei Schlaganfallpatienten (150.000 Fälle/Jahr) und bei inkompletten Querschnittlähmungen (1200 Fälle/Jahr) eingesetzt.

Für die gesamte Diagnostik und Therapieplanung werden aus der klinischen Erfahrung stammende, unscharf definierte Entscheidungsregeln von Seiten der Untersucher verwendet. Der unterschiedliche Kenntnisstand dieser Untersucher und die Schwierigkeit, eine Vielzahl von Parametern gleichzeitig auswerten zu müssen, verhindern zum heutigen Stand eine standardisierte Auswertung.

Erste Versuche, diese Entscheidungen mittels einer automatisierten Klassifikation zu objektivieren, wurden bereits in der Vergangenheit unternommen. Dazu wurden für eng umrissene Fragestellungen mit Hilfe von datenbasierten Methoden Klassifikatoren entworfen (z. B. unter Verwendung Künstlicher Neuronaler Netze [4] oder statistischer Methoden [43]) oder das vorhandene Wissen in Expertensysteme implementiert (z. B. [11, 25, 53, 68]). Obwohl dabei interessante Ergebnisse erzielt wurden, verhindern bisher schwerwiegende Nachteile der genannten Verfahren den routinemäßigen klinischen Einsatz: Datenbasierte Verfahren führen meist zu schwer interpretierbaren, nicht nachvollziehbaren Lösungen und Lösungswegen. Der Grund für die eingeschränkte Verwendbarkeit von Expertensystemen in der klinischen Entscheidungsfindung liegt hauptsächlich an dem bekannten Problem, Expertenwissen in vollem Umfang zu erfassen und zu formalisieren. Eine Übertragbarkeit der Verfahren zur Klassifikation von Ganganalysedaten auf andere Institutionen und Fragestellungen wurde somit nur in begrenztem Maße erreicht.

Um diese Nachteile zu umgehen, wird im Unterschied zu bisher bekannten Verfahren angestrebt, expertenunabhängige datenbasierte Methoden und auf Expertenwissen basie-

rende Methoden zu kombinieren und die Interpretierbarkeit der datenbasiert gewonnenen Ergebnisse für den Experten zu garantieren. Damit sollen transparente Entscheidungen und eine quantitative Bewertung ermöglicht werden.

Die mittelfristigen Teilziele zum Erreichen dieser Vorhaben sind:

- Definieren und Extrahieren von geeigneten Einzelmerkmalen auf der Grundlage von medizinischem Expertenwissen aus den erhobenen Datensätzen,
- Formulierung geeigneter Regeln, die zum Ausdruck bringen, in welcher Art und Weise diese Merkmale in die Diagnoseerstellung eingehen,
- eine offene, interpretierbare und jederzeit erweiterbare Wissensbasis,
- Suche nach weitgehend redundanten Entscheidungsregeln, die auch bei unterschiedlicher gerätetechnischer Ausstattung (z. B. fehlende EMG-Erfassung, fehlende Kraftmessung) noch eine valide Entscheidung ermöglichen,
- rechen-technische Implementierung der regelbasierten Auswerteverfahren,
- automatisierte Auswertung von Datensätzen,
- klinische Validierung der automatisch erstellten Diagnose sowie
- Weiterentwicklung vorhandener Verfahren zur Regelgenerierung und Klassifikation zur Anpassung an die speziellen Besonderheiten ganganalytischer Daten.

In der hier beschriebenen ersten Projektphase werden die Vor- und Nachteile von zwei methodischen Ansätzen, einem statistischen Verfahren und einem wissensbasierten Verfahren mit Fuzzy-Regeln, anhand der Schrittphasenerkennung in der Instrumentellen Ganganalyse diskutiert. Die Schrittphasenerkennung dient dabei als Beispiel für eine typische Klassifikationsaufgabe in der Instrumentellen Ganganalyse [1, 82]. Sie bildet eine Voraussetzung für die Anwendung nachfolgender Auswerteverfahren, mit denen langfristig eine Unterstützung und Objektivierung von Diagnose- und Therapieentscheidungen angestrebt wird. Bei beiden Verfahren steht nicht nur das Ergebnis, sondern insbesondere die Transparenz und Interpretierbarkeit des Lösungsweges im Mittelpunkt.

Die Ziele dieses Berichtes sind

- die Einführung in die Problemstellungen bei der Instrumentellen Ganganalyse,
- die Vorstellung der verwendeten Klassifikationsverfahren und
- die Darstellung der Vor- und Nachteile dieser Verfahren anhand der Schrittphasenerkennung.

Das Projekt 'Anwendung von Methoden der Computational Intelligence zur autonomen Diagnosestellung in der Ganganalyse' (GE 1139/1, BR 1303/6) wird seit dem 1. Januar 2001 von der Deutschen Forschungsgemeinschaft (DFG) gefördert. Der vorliegende Bericht fasst die Vorarbeiten und erste Ergebnisse zusammen.

Kapitel 2

Instrumentelle Ganganalyse

2.1 Neurologisch bedingte Bewegungsstörungen

Die Ausprägungen neurogener Bewegungsstörungen hängen stark von der Art und der Lokalisation der beeinträchtigten neuronalen Strukturen ab. Die zu Grunde liegenden Ursachen können sowohl angeboren als auch erworben sein. Eine Übersicht über typische Erkrankungsbilder aus Sicht der Ganganalyse findet sich in [8].

Querschnittslähmungen als Folge einer Schädigung des Rückenmarks (Spina) ziehen abhängig von der Höhe der Schädigung dramatische Funktionsstörungen sowohl des somatischen als auch des vegetativen Nervensystems nach sich. Je nach Umfang der Beeinträchtigung ist zwischen kompletten und inkompletten Querschnittslähmungen zu unterscheiden. Die resultierenden Bewegungsstörungen erstrecken sich entweder auf die beiden unteren Extremitäten (Paraplegie) oder auf alle vier Extremitäten (Tetraplegie). Die betreffenden Muskeln sind nicht mehr willkürlich ansteuerbar und bzw. oder die sensomotorische Rückkopplung (Bewegungsgefühl) für bestimmte Körperregionen fehlt vollständig oder teilweise. Bei partiellen Querschnittslähmungen gibt es starke individuelle Unterschiede bezüglich der betroffenen Regionen. Mit Hilfe einer individuell auszuwählenden geeigneten Therapie (z. B. Laufbandtraining) kann in einigen Fällen die Stand- und Gehfähigkeit zumindest teilweise wiederhergestellt werden. Zudem trägt eine geeignete Therapie entscheidend zur Kreislaufstabilisierung und somit zur Erhöhung der Lebenserwartung bei.

Infantile Zerebralpareesen (ICP) bzw. frühkindliche Hirnschädigungen sind häufig die Folge eines Geburtstraumas, wobei bestimmte Hirnregionen geschädigt werden. Auch hier kommt es zu individuell variierenden motorischen Beeinträchtigungen. Sie können beidseitig (Diplegie) oder auf eine Körperseite beschränkt (Hemiplegie) sein. In der Regel treten Muskelspastiken (nicht willkürlich beeinflussbare Muskelspannungen - Erhöhungen des Muskeltonus) auf. Damit entstehen komplexe pathologische Veränderungen des Gangbildes, die beispielsweise zu Wachstumsstörungen, zur Überlastung von Gelenken und mit zunehmendem Alter zum vollständigen Verlust der Gehfähigkeit führen können. Als operative Therapiemethoden stehen orthopädisch-chirurgische Eingriffe zur Verbesserung der biomechanischen Voraussetzungen beim Gehen wie Muskelverlängerung, die Verlagerung oder der Transfer eines Muskels bzw. seiner Sehne sowie Eingriffe zur Korrektur am knöchernen Bewegungsapparat zur Verfügung. Da gerade Operationen dieser Art irreversibel sind, sollen sie einen wesentlichen funktionellen Zugewinn versprechen und müssen exakt geplant werden. Außerdem können durch Behandlungen mit Botulinum Toxin [105] spastische Muskeln gezielt ruhig gestellt werden, was das Gangbild ebenfalls verbessern kann.

In den genannten Fällen ist die Instrumentelle Ganganalyse ein wichtiges Diagnoseverfahren und beeinflusst damit die nachfolgende Therapieentscheidung.

2.2 Historische Entwicklung

Die Versuche, den menschlichen Gang zu analysieren und objektiv zu beschreiben, reichen bis in die Zeit von Aristoteles (384 v. Chr. bis 322 v. Chr.) und Galenus (129 bis 199). Als einer der Meilensteine in der Entwicklung der Ganganalyse gilt das Werk *de motu animalium* von Borelli (1682). Hier wird zum ersten Mal der Körperschwerpunkt (center of gravity) bestimmt. Die moderne wissenschaftliche Messung und Ausarbeitung von Prinzipien des Ganges geht auf Werke der Gebrüder Weber *Die Mechanik der Gehwerkzeuge* (1836) zurück. Ihnen ist z. B. die Erkenntnis zu verdanken, dass der Anteil der Zeit eines Gangzyklus, bei dem sich beide Füße auf dem Boden befinden, mit zunehmender Geschwindigkeit immer kleiner wird. Die Möglichkeiten der Fotografie eröffneten auch im Bereich der Ganganalyse neue Möglichkeiten. Der Fotograf Muybridge aus San Francisco nutzte sie zur Darstellung der Bewegungsphasen eines Pferdes (*The horse in motion, as shown by instantaneous photography, 1882*), die von William veröffentlicht wurden. Braune und Fischer stellten 1889 erstmals eine Methode zur Bestimmung der dreidimensionalen Ganganalysedaten aus fotografischem Material vor. Die von den hier genannten Autoren und Arbeitsgruppen vorgestellten Messergebnisse erforderten einen enormen Messaufwand und Fleiß bei der Auswertung, decken sich aber sehr gut mit denjenigen, die heute in Ganglabors mit moderner Rechner- und Geräteausstattung bestimmt werden. Scherb entwickelte in der ersten Hälfte des letzten Jahrhunderts ein Laufband, welches ihm erlaubte, den Probanden ohne Ortsänderung beim Gehen zu untersuchen. Er hat durch Palpation die Muskelaktivierung in bestimmten Gangzyklen bestimmt und festgehalten.

Ein wesentlicher Durchbruch erfolgte nach 1980 durch die Zusammenführung von moderner Computertechnologie mit den Analysemethoden der fotografischen Technik. Erst hierdurch kam es zur klinischen, in der Routinediagnostik einsetzbaren Ganganalyse. Sie wurde in der Folge z. B. von Winter, Gage, Perry und Sutherland formalisiert und ihre Effektivität in der besseren Behandlungsplanung bei Krankheiten mit Gangstörungen nachgewiesen. Querschnittgelähmte zählten lange Zeit nicht zu den besonderen Zielgruppen für die Ganganalyse. Aufgrund der vielfältigen Komplikationen der Erkrankung war die Lebenserwartung kurz, die Schwere der Verletzung führte in den allermeisten Fällen zu kompletten Lähmungen ohne Restgefunktionen.

Durch die grundlegenden Arbeiten von Sir Ludwig Guttmann ist es gelungen, die Komplikationen einer Querschnittlähmung zu beherrschen und ein sicheres Langzeitüberleben zu garantieren. Außerdem hat die Entwicklung der Technologie gerade im Kfz-Bereich und die Verbesserung der Rettungskette dazu geführt, dass bei immer mehr Querschnittgelähmten ein inkomplettes Lähmungsbild vorliegt. Hier ist das Ausnutzen der noch verbliebenen Restgefunktionen sowie das Wiedererlernen des Gehens ein wesentliches Rehabilitationsziel. Vor etwa 15 Jahren haben Wernig und Barbeau mit Gehtraining auf dem Laufband unter Zuhilfenahme eines Entlastungssystems, ähnlich einem Fallschirmspringergurt, begonnen [3, 112]. Diese Therapiemethode ist mittlerweile in der Praxis etabliert. Sie erlaubt es, einen Patienten gehen zu lassen, der noch nicht sein volles Körpergewicht selbst tragen kann (Bild 2.1). Für diagnostische Zwecke erlaubt diese Anordnung, den Patienten genau dosiert zu belasten und zu beobachten, welche Folgen dies auf die einzelnen Bewegungsabschnitte der unteren Extremitäten hat.

Mit dieser Ausstattung kann in Verbindung mit den Möglichkeiten der Ganganalyse



Bild 2.1: Partiiell querschnittgelähmter Patient bei einer Ganganalyse auf dem Laufband

der Einsatz von Hilfsmitteln, wie Orthesen, Schuhzurichtungen und eventuell zukünftig zur Verfügung stehender Hilfsmittel wie Neuroprothesen auf seine Notwendigkeit überprüft und optimiert werden.

2.3 Entwicklungsstand

Heute wird die Instrumentelle Ganganalyse im Rahmen vielfältiger Zielsetzungen angewandt:

- zur Erweiterung des grundlegenden Verständnisses und der Nachbildung des Gangbildes von gesunden Menschen oder Tieren, insbesondere zur Übertragung natürlicher Fortbewegungsweisen auf zwei oder mehrbeinige Roboter (Laufmaschinen) [22, 23, 26, 64, 97, 103, 104],
- zur Modellbildung für die Entwicklung von Neuroprothesen, bei der sich die Forschung in erster Linie auf die Bestimmung von Aktivierungsmustern von Muskelgruppen und von notwendigen Sensorinformationen zur Generierung eines stabilen Gang- oder Standmusters konzentriert, [91–93, 111, 117],
- für die Diagnose, die Therapieplanung und –validierung bei Gang- und Bewegungsstörungen [6, 7, 13, 14, 18, 20, 27, 31–33, 37, 43, 54, 56, 66, 67, 70, 86–88, 96, 98, 99, 101, 109, 113, 114].

Je nach Zielsetzung kommen unterschiedliche Abstraktionsniveaus einer muskulo-skeletalen und biomechanischen Modellierung zum Einsatz. Bei der Bearbeitung der ersten beiden Ziele dominieren detaillierte, parametrisierte und rechenaufwendige Modelle, die aufgrund ihrer komplexen Parameterbestimmung nur in seltenen Fällen individuell an Probanden

anpassbar sind. Die so gewonnenen Modelle dienen hauptsächlich dem Entwurf und der Simulation von Regelalgorithmen, die unter definiert vorgegebenen Randbedingungen in Form von Sensorinformationen und Stellmöglichkeiten (Muskelaktivierung bei Neuroprothesen oder Stellmotoren bzw. hydraulische Stellglieder bei Laufmaschinen) für spezialisierte Aufgaben (z. B. Stabilisierung beim Stehen, Laufen bei konstanten Geschwindigkeiten) entwickelt werden. Beim dritten klinischen Schwerpunkt werden stark abstrahierte Modelle zur übersichtlichen Auswertung von kinematischen Bewegungsdaten und zur Berechnung von kinetischen Größen verwendet.

In der Orthopädischen Universitätsklinik Heidelberg besteht das Ziel der Instrumentellen Ganganalyse hauptsächlich darin, die Diagnosestellung und Therapieplanung bei neurogenen Gangstörungen (z. B. frühkindliche Hirnschädigung, Querschnittlähmung) zu objektivieren.

2.4 Messverfahren

Alle im Folgenden beschriebenen Messungen basieren auf Untersuchungen, die mit einem zum größten Teil in Heidelberg entwickelten Diagnostiklaufband [96] mit der Möglichkeit zur dreidimensionalen Bodenreaktionskraftmessung vorgenommen werden (Bild 2.2). Durch die Anwendung der Ganganalyse in Kombination mit einem Laufband und einer Vorrichtung zur Gewichtsentlastung ist eine Erweiterung der Methodik der Ganganalyse auf Patienten möglich, die nur mittels Gewichtsentlastung gehfähig sind, wie z. B. querschnittgelähmte Patienten.

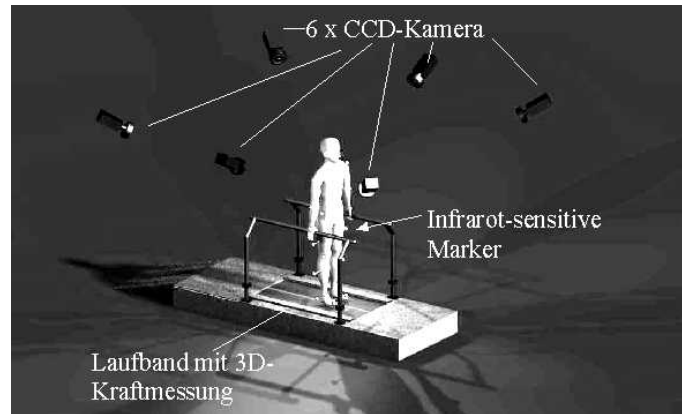


Bild 2.2: Schematische Darstellung der Datenaufnahme mit Laufband

Vor der Messung werden zunächst patientenspezifische Daten ermittelt. Diese Daten umfassen Basisdaten wie

- Geschlecht,
 - Alter,
 - Diagnose,
 - Anamnese am Muskelskelettsystem (evtl. Voroperationen und Therapie) und
 - Begleiterkrankungen (besonders Herz/Lunge mit Einfluss auf Leistungsfähigkeit),
- und anthropometrische Daten (Maßverhältnisse am menschlichen Körper):
- Körpergröße [cm],

- Körpergewicht [kg],
- Fußlänge (Ferse-Großzehenspitze [cm]),
- Beinlänge (Spina iliaca anterior-superior (vorderer oberer Darmbeinstachel) [cm]),
- maximale knöcherne Kniegelenksbreite (mittleres und seitliches Oberschenkelknochenhöckerchen (Epicondylen) [cm]),
- Sprunggelenksbreite (oberer Teil des oberen Sprunggelenkes (Maleolengabel) [cm]),
- Tibial Torsion (Schienbeindrehung, Fußöffnungswinkel gemessen in Bauchlage, Unterschenkel 90° gebeugt, Fuß neutral),
- BMI: Body mass index BMI (BMI=Gewicht [kg]/(Größe [m])²) und
- BSA: Body surface area [cm²] nach Mosteller: $\sqrt{\frac{\text{Größe[cm]} \cdot \text{Gewicht [kg]}}{3600}}$ (zugeschnittene Größengleichung).

Für die Erhebung der ganganalytischen Daten kommt meist ein markergestütztes Infrarot-Bewegungsanalysesystem mit sechs Kameras zum Einsatz (hier z. B. EVaHiRes der Firma Motion Analysis). Hierbei nehmen Videokameras die Positionen der reflektierenden Marker in einem festen Raumkoordinatensystem auf. Die Marker werden dabei an festgelegten Punkten auf der Haut befestigt. Solche Bewegungsanalysesysteme enthalten Komponenten zur

- Ansteuerung der Aufnahmehardware für die Erfassung der Videodaten (CCD-Kameras und zugehörige Framegrabberkarten für die Digitalisierung der Videodaten) und der Analogmesswerte (AD-Wandlerkarte),
- Bereitstellung der mathematischen Methoden zur statischen und dynamischen Kalibration der Aufnahmehardware,
- Bereitstellung der mathematischen Methoden zur Verrechnung der aus sechs Kameraperspektiven aufgenommenen zweidimensionalen Markertrajektorien zu Bewegungsbahnen im dreidimensionalen Raum (Discrete Linear Transform – DLT) und
- Möglichkeiten zur Bearbeitung der berechneten Trajektorien.

Im Ergebnis entstehen Trajektorien aller Marker im dreidimensionalen Raum.

Die weitere Auswertung der Daten (hier z. B. mit dem Programmpaket OrthoTrak 4.1 der Firma Motion-Analysis [85]) beruht auf einem stark vereinfachten Körpermodell. Es gestattet die Zuordnung der einzelnen Marker zu Segmenten auf der Basis eines Starrkörpermodells des menschlichen Körpers und die Berechnung der Gelenkwinkelverläufe durch Bestimmung der Orientierung der definierten Segmente zueinander bzw. im Raum. Dieses Körpermodell enthält Annahmen über die Bewegungsmöglichkeiten zwischen verschiedenen, in sich starren Körpersegmenten sowie deren Abmessungen, Massen und Trägheitsmomente, die aus den oben genannten patientenspezifischen Daten geschätzt werden.

Damit werden

- Kinematikdaten (Gelenkwinkelverläufe, Raum-Zeit-Parameter) und
- Kinetikdaten

bestimmt (Bild 2.3). Die Datensätze enthalten somit für verschiedene Gelenke (Fußgelenk, Knie, Hüfte, Becken, Hals, Schulter usw.) jeweils Schätzungen für die Gelenkwinkel, die Kräfte, die Momente und die mechanische Arbeit in drei verschiedenen Projektionsebenen (frontal – von vorn, sagittal – von der Seite, transversal – von oben). Diese Daten liegen

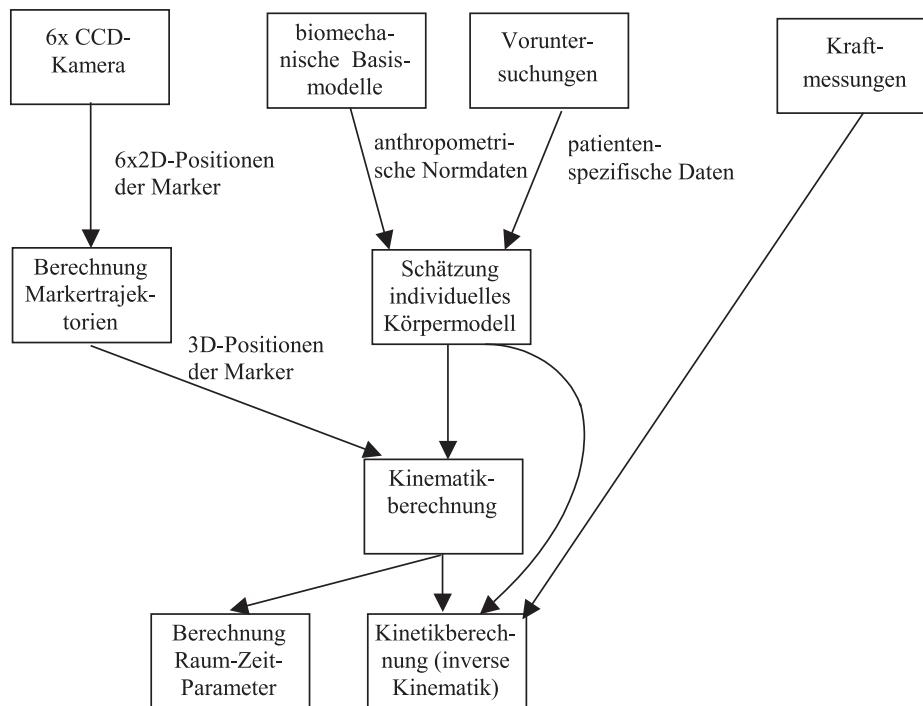


Bild 2.3: Ablauf bei der Berechnung der Ganganalyseparametern

als Standard für gesunde Probanden weltweit in den Ganglabors vor und werden teilweise durch EMG- oder Sauerstoffverbrauchsdaten ergänzt.

Bei den *Kinematikdaten* sind beim Becken die Kippung (sagittal), der Schiefstand (frontal) und die Rotation (transversal) relevant. Die Kippung (sagittal) folgt aus der relativen Bewegung einer vertikalen Achse bezüglich einer horizontalen Achse von links nach rechts (l/r) durch das Becken. Ein Winkel von 0° bedeutet einen gerade aufgerichteten Oberkörper, positive Werte einen nach vorn geneigten Oberkörper. Beim Schiefstand wird die Bewegung einer horizontalen Achse (l/r) bezüglich einer horizontalen Achse von vorn nach hinten (v/h) durch das Becken bestimmt. Ein Winkel von 0° bedeutet einen geraden Stand. Bei der Rotation wird der Winkel einer horizontalen Achse (l/r) bezüglich der vertikalen Längsachse gemessen.

Die Hüftwinkel bringen die Bewegung des Oberschenkels relativ zum Becken zum Ausdruck. Bei der Bestimmung der Flexion (Beugung) bzw. Extension (Streckung) in der sagittalen Ebene wird die Bewegung der Längsachse durch den Oberschenkel aus Sicht der horizontalen Achse (l/r) durch das Becken bestimmt. Die Abduktion (Bewegung vom Körper weg) bzw. Adduktion (Bewegung zum Körper hin) bestimmt die Bewegung der Längsachse durch die Oberschenkel bezüglich zur horizontalen Achse (v/h) durch das Becken. Die innere bzw. äußere Rotation (transversal) beschreibt die relative Bewegung zwischen der horizontalen Achse (l/r) durch den Oberschenkel und der horizontalen Achse (l/r) durch das Becken aus Sicht des Beckens entlang der Längsachse des Oberschenkels (gegenseitige Verdrehung). Diese Messung liefert die Beziehung zwischen dem Becken und dem entferntesten Oberschenkelknochen.

Die Kniewinkel bringen die Bewegung des Unterschenkels relativ zum Oberschenkel zum Ausdruck. Klinisch relevant ist nur die Beugung (Flexion) bzw. Streckung (Extension)

in der sagittalen Ebene, wobei die Bewegung der Längsachse des Unterschenkels aus Sicht der horizontalen Achse (l/r) des Knies gemessen wird. Ein durchgestrecktes Knie bedeutet 0° . Die Knierotation ist ebenso wie die Knieabduktion sehr gering und kann im Vergleich zu den Messfehlern nicht valide bestimmt werden. Folglich sind solche Merkmale wegen schlechter Stör-Nutz-Signalverhältnisse wenig aussagekräftig.

Beim Sprunggelenk ist die sagittale und die transversale Ebene von Interesse. Der sagittale Gelenkwinkel (Dorsiflexion in Richtung des Fußrückens, Plantarflexion in Richtung der Fußsohle) beschreibt den Winkel zwischen Unterschenkel und Fußrücken. Bei der Rotation wird die Stellung des Fußes bezüglich zur Körperlängsachse betrachtet.

Die *Raum-Zeit-Parameter* sind ein Oberbegriff für komprimierte Merkmale, die aus den Gelenkwinkelverläufen berechnet werden. Dazu gehören die Schrittlänge (links/rechts, Doppelschritt) [cm], die Gehgeschwindigkeit [cm/s], die Anzahl von Schritten pro Minute bzw. Kadenz [min^{-1}], die Zeitdauer des Doppelschrittes [s], die Spurbreite [cm] und die prozentualen Anteile verschiedener Phasen am Gangzyklus (Standphase, Double Support – Phase mit beiden Füßen am Boden usw.).

Die *Kinetikdaten* berechnen sich aus den Gelenkwinkeln und den gemessenen Bodenreaktionskräften. Sie umfassen Gelenkkräfte [N], auf das Körpergewicht normierte Gelenkmomente [Nm/kg] und auf das Körpergewicht normierte mechanische Gelenkleistungen [W/kg] an den einzelnen Gelenken. Die Gelenkleistung ist das Produkt aus Gelenkmoment und Winkelgeschwindigkeit. Klinisch besonders wichtig sind dabei die sagittalen Kinetikdaten der Hüfte, des Knies und des Sprunggelenks sowie die frontalen Werte der Hüfte.

Aufgrund der Ungenauigkeiten in der Positionierung der Marker am Körper, der Messfehler bei der Positionsbestimmung der Marker und der Kraftmessung sowie der vereinfachten Modelle kommt es zu Abweichungen der gemessenen Merkmale von den realen Größen.

Für die strukturierte Verwaltung der verschiedenen Messmodalitäten der Instrumentellen Ganganalyse (Video-, Kraftmess-, Druckverteilungs-, EMG- und Sauerstoffverbrauchsdaten) wurde an der Orthopädischen Universitätsklinik Heidelberg eine hierarchisch strukturierte Datenbank entworfen. In der Datenbank wurden ein universelles, standardisiertes Interface zur Erstellung von Auswertelgorithmen (MATLAB) und ein ebenfalls standardisiertes Interface zu allgemeinen Visualisierungsfunktionen (MATLAB) im Sinne von Bibliotheksfunktionen implementiert. Mit Hilfe dieser Software sind standardisierte Ausdrücke der Kinematik-, Kinetik- und EMG-Verläufe möglich, die für die klinische Entscheidungsfindung verwendet werden.

Die in Heidelberg aufgebauten Datenbanken enthalten neben den instrumentell erhobenen Messergebnissen auch ärztliche Diagnosen – hauptsächlich von Querschnittgelähmten und Patienten mit Infantiler Zerebralparese (ICP). Diese liegen als nicht standardisierter Erklärungstext und damit in nicht unmittelbar rechentechnisch auswertbarer Form (z. B. als Zahlencode) vor. Die standardisierte Klassifikation von Diagnosen ist aber eine unabdingbare Voraussetzung für die Anwendung der Verfahren der Computational Intelligence, die auf überwachten Lernverfahren aufbauen. Als weitere Voraussetzung für die Bearbeitung des Projektes ist daher ein Konzept zu entwickeln, um die bestehenden Diagnosen in einer standardisierten Form zu dokumentieren. Dazu soll geprüft werden, inwieweit das ICD-System (International Statistical Classification of Diseases and Related Health Problems) als Basis verwendet werden kann und wie es gegebenenfalls erweitert werden muss.

2.5 Subjektive Bewertung zur Diagnose und Therapieplanung

2.5.1 Allgemeines Vorgehen

Auf der Basis der subjektiven Analyse der verarbeiteten Daten erfolgt in der klinischen Praxis das Stellen von Arbeitsdiagnosen. Die Bedeutung dieser Diagnosen liegt derzeit vor allem in der Evaluierung konservativer Therapiemethoden, die eingesetzt werden, um z. B. Operationen bei Kindern aufzuschieben, sowie bei der Überprüfung des Operationsergebnisses von orthopädisch-chirurgischen Eingriffen. Für die gesamte Diagnostik und Therapieplanung werden aus der klinischen Erfahrung stammende, nur unscharf definierte Entscheidungsregeln von Seiten der Untersucher verwendet.

Im folgenden Unterabschnitt werden am Beispiel der Spitzfußproblematik bei der Infantilen Zerebralparese (ICP) wesentliche Merkmale erläutert, die in der klinischen Praxis eingesetzt werden.

2.5.2 Subjektive Bewertung mit Videodaten

Bei der Ganganalyse selbst werden subjektive und objektive Gangparameter unterschieden. Gage [29] und Koman kommen auf unterschiedliche Weise zu ersten subjektiven Diagnosen. Gage legt beim Betrachten des Patientenvideos besonderes Augenmerk auf folgende seitendifferente Gangparameter, ohne diese jedoch quantitativ zu bewerten:

- Standstabilität (rechts/links (r/l), vorhanden/nicht vorhanden),
- Bodenfreiheit in der Schwungphase (r/l, vorhanden/nicht vorhanden),
- IC (Initial Contact) (r/l, Ferse, Sohle, Vorfuß),
- Adäquate Schrittlänge (r/l, adäquat/nicht adäquat) sowie
- Gangökonomie (mühsam/flüssig).

Koman legt seinen Beobachtungen eine ärztliche Schätzskala zu Grunde, die speziell für die sogenannte Spitzfußsymptomatik verwendet wird. Dabei betrachtet er jedes Bein zu unterschiedlichen Zeitpunkten und teilt den aufgeführten Gangparametern Werte gemäß seiner Schätzskala zu:

- Kauergang (sagittal): Der Grad der Beugung wird unterteilt in: stark ausgeprägt ($> 20^\circ$ Hüfte, Knie, Sprunggelenk, Skalenwert 0), mäßig ausgeprägt (5° - 20° Hüfte, Knie, Sprunggelenk, Skalenwert 1), schwach ausgeprägt ($< 5^\circ$ Hüfte, Knie, Sprunggelenk, Skalenwert 2) oder keine Beugung (Skalenwert 3).
- Spitzfuß: Bei spitzem Winkel der Sprunggelenkposition ist die Wadenmuskulatur häufig verkürzt. Eingeteilt wird in fest (Gelenk vollständig steif, wird im Liegen untersucht, Skalenwert 0), dynamisch (passive Bewegung möglich, wird im Liegen untersucht, Skalenwert 1), gelegentlicher Fersenkontakt (nach Auftreten mit dem Vorderfuß, Skalenwert 2), Auftritt mit Ferse und Abrollen bis Zehen (Skalenwert 3).
- Die Rückfußposition (transversal) gibt an, ob der Fuß nach innen oder nach außen gedreht ist: varus (bei Fußauftritt nach außen gedreht, Skalenwert 0), valgus (bei Fußauftritt nach innen gedreht, Skalenwert 1), gelegentlich neutral bei Fußauftritt (Skalenwert 2) oder neutral bei Fußauftritt (Skalenwert 3).
- Die Überstreckung des Knies in der Standphase wird unterteilt in: Überstreckung $> 5^\circ$ (Skalenwert 0), Überstreckung $0^\circ - 5^\circ$ (Skalenwert 1), keine Überstreckung (Skalenwert 2).

- Ganggeschwindigkeit: nur langsam (Skalenwert 0), variabel langsam – schnell (Skalenwert 1).
- Gangart: nur auf den Zehen (Skalenwert 0), gelegentlicher Fersenkontakt (nach Auftreten mit dem Vorderfuß (Skalenwert 1), Auftritt mit Ferse und Abrollen bis Zehen (Skalenwert 2).

Die zugeordneten Skalenwerte geben Auskunft über die einzelnen Anomalien. Veränderungen des Gangbildes, insbesondere vor und nach einer Therapie, können durch Aufaddieren der Skalenwerte quantifiziert werden. Allerdings sind diese Daten aufwendig zu erheben, untersucherabhängig und z. T. schwer reproduzierbar.

2.5.3 Subjektive Bewertung mit Messdaten der Ganganalyse

Im Vergleich zu diesen relativ einfachen qualitativen Abschätzungen liefern die in Abschnitt 2.4 vorgestellten Messverfahren wesentlich detailliertere Informationen. Die heute übliche Vorgehensweise besteht darin, die gemessenen klinisch relevanten Kinematikverläufe und teilweise auch die Kinetikverläufe mit Normdaten von gesunden Probanden zu vergleichen.

Ein Beispiel für die daraus entstehenden Charts gibt Bild 2.4. Hier sind die Kinematikdaten für Becken, Hüfte, Knie- und Sprunggelenk (von oben nach unten) dargestellt. Von links nach rechts sind die Gelenkwinkelverläufe in der Sagittal-, Frontal- und Transversalebene angegeben (y-Achse). Auf der x-Achse der jeweiligen Bilder ist die normierte Zeit zwischen 0 und 100% eines Gangzyklus aufgetragen. Grau hinterlegt sind Normbereiche, die bei gesunden Probanden und mittleren Ganggeschwindigkeiten auftreten. Die durchgezogenen Linien geben die Messwerte für die rechte, die gestrichelten Werte für die linke Körperseite an. Klinisch irrelevante Verläufe (z. B. transversale Kniegelenkwinkel), die wegen schlechter Stör-Nutzsignal-Verhältnisse (Bewegungen gegen Messfehler) nicht mit ausreichender Genauigkeit bestimmt werden können, sind nicht mit angegeben.

Die Diagnosestellung und Therapievalidierung basiert auf dem Vergleich der Messwerte mit den Normkurven. Eine Diagnosestellung ist deshalb schwierig, weil im gesamten Gangbild Ursachen und Kompensationsmechanismen schwer zu unterscheiden sind. Solche Kompensationsmechanismen sind erforderlich, um Schmerzen zu vermeiden, das Gleichgewicht zu halten und ein Vorwärtskommen zu ermöglichen. Deshalb hat häufig eine Ursache (z. B. eine dynamische Bewegungseinschränkung im Sprunggelenk) eine Vielzahl von Wirkungen im Gangbild (z. B. eine stärkere sagittale Knieflexion, stärkere Rotationsbewegungen im Becken usw.).

Das Ziel einer Therapie besteht in der Regel darin, eine Annäherung an das Normverhalten zu erreichen. Da diese 'Annäherung' nicht ohne weiteres zu definieren ist, bleibt diese Definition bisher subjektiv.

Das wichtigste Problem liegt darin, anhand des gemessenen Gangbildes die Auswirkungen einer möglichen Therapie abzuschätzen, was eine detaillierte Diagnose voraussetzt und langjährige klinische Erfahrung erfordert.

2.6 Ansätze zur automatisierten Bewertung

Für einige bei der Diagnose auftretende Fragestellungen wurden deshalb mit Hilfe verschiedener Ansätze Auswerteverfahren entwickelt, um hier zu einer objektiveren Vorgehensweise zu kommen (Übersicht siehe [15, 16]).

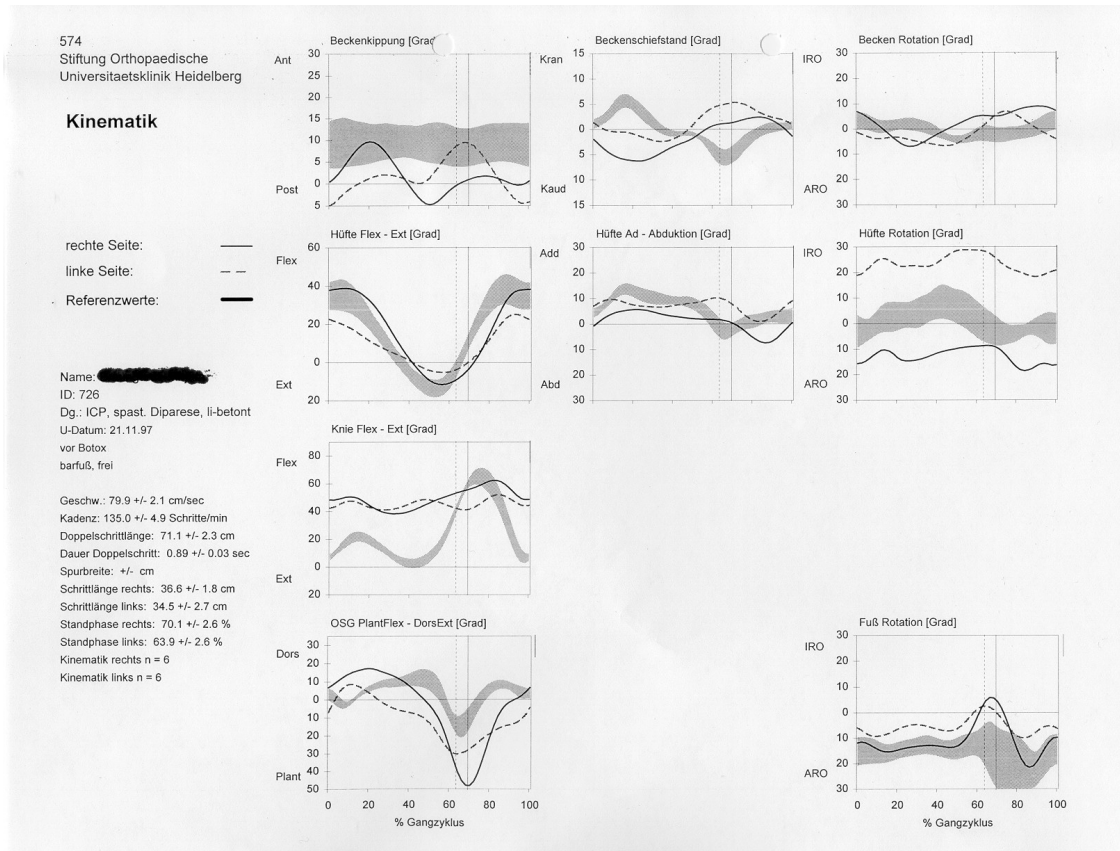


Bild 2.4: Kinematikdaten als Ergebnis der Ganganalyse für einen ICP-Patienten

Eine Vielzahl von Arbeiten beschäftigt sich mit der Analyse von Bewegungsmustern mit Hilfe einfacher statistischer Verfahren, wie Mittelwertvergleichen und Berechnungen von Kovarianzmatrizen. So untersucht beispielsweise [116] Merkmale wie Hüftflexion und deren Änderungsgeschwindigkeit in verschiedenen Schrittphasen.

Eine Reihe von Anwendungen nutzt die Hauptkomponenten- (HKA) und Faktoranalyse. In [19] wird anhand transformierter Merkmale ein Vergleich pre- und postoperativer Daten vorgenommen. Die HKA zeigt an, in welchem Bereich des Gangzyklus signifikante Unterschiede im Datenmaterial auftreten. In [71] wird mit Hilfe der Faktoranalyse nach Funktionsmechanismen und Synergien beim Zusammenwirken unterschiedlicher Muskelgruppen gesucht.

Die Diskriminanzanalyse wird beispielsweise dazu verwendet, aus Schwarz-Weiß-Bildern in einer sagittalen Projektion räumliche und zeitliche Muster zur Unterscheidung verschiedener Personen abzuleiten [43].

Fuzzy-Regeln werden beispielsweise zur Identifikation von vereinfachten Sequenzen von Schrittphasen (5 Phasen) unter Nutzung von Winkelgebern als Eingangssignal zum Entwurf von Neuroprothesen für die funktionelle Elektrostimulation verwendet [84]. Nachteilig sind hier allerdings extrem große Regelbasen von 100-2000 Regeln und manuell entworfene Zugehörigkeitsfunktionen.

Künstliche Neuronale Netze werden dazu verwendet, Ganggeschwindigkeiten und modifizierte Gangmuster (fixierte Gelenke, Zusatzgewichte) aus den Daten zu erkennen [5, 30].

Ähnliche Arbeiten existieren zur Unterscheidung normaler und pathologischer Gangmuster [41] und zu Zusammenhängen mit EMG-Daten [100].

Ein markerloser Ansatz über die direkte Auswertung von Videodaten wird in [72–76] vorgeschlagen. Nach einem Clusterverfahren werden dann Hidden-Markov-Modelle zur Erkennung von verschiedenen Gangarten (langsames Gehen, Springen usw.) verwendet.

Der Grund für die eingeschränkte Verwendbarkeit all dieser Verfahren liegt in dem zugrundeliegenden Klassifikationsprinzip, das eine starre Formalisierung des in der jeweiligen Institution verwendeten Erfahrungswissens darstellt. Insbesondere neuronale Ansätze führen zu schwer interpretierbaren, nicht nachvollziehbaren Ergebnissen. Eine Übertragbarkeit des Verfahrens auf andere Institutionen und Fragestellungen ist nur in begrenztem Maße zu erreichen und konnte sich deshalb bisher auch nicht in der breiten klinischen Anwendung durchsetzen. Ansätze, die diese Klassifikationsaufgabe auf der Basis von detaillierten muskulo-skeletalen und biomechanischen Modellen vornehmen, sind bisher nicht bekannt und scheinen auch erst dann vielversprechend, wenn das Problem der individuellen Adaption dieser aufwendigen Modelle befriedigend gelöst ist. Ein erster Schritt in diese Richtung ist die Verwendung individueller Modelle für die Bestimmung der Kinetik [66]. Zur Unterstützung der im Moment stark untersucherabhängigen Therapieentscheidung ist aber dringend ein verlässliches, objektives Verfahren der Diagnosestellung notwendig, jedoch beim derzeitigen Stand der Forschung nicht existent.

2.7 Schrittphasen

Eine wichtige Voraussetzung aus diagnostischer Sicht ist es, Schrittphasen im Bewegungsmuster eines Patienten zu erkennen und eventuell individuelle Besonderheiten innerhalb der einzelnen Schrittphasen zu formulieren. Diese Information dient einer späteren Bildung neuer Merkmale, die aus einer zusammenfassenden Beschreibung der Bewegung innerhalb einer Schrittphase entstehen.

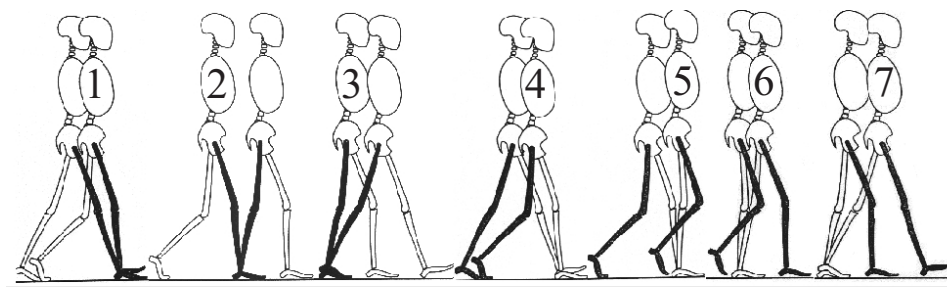


Bild 2.5: Schrittphasen – rechte (R) Körperseite (1: Initial Contact und Loading Response (LRE), 2: Mid Stance (MST), 3: Terminal Stance (TST), 4: Pre Swing (PSW), 5: Initial Swing (ISW), 6: Mid Swing (MSW), 7: Terminal Swing (TSW); nach [88])

Innerhalb eines Schrittzklus werden nach [88] sieben Schrittphasen unterschieden (Bild 2.5). Die erste Phase (*Loading Response*) beginnt mit dem Aufsetzen des betrachteten Fußes auf dem Boden (*Initial Contact*). Diese Phase umfasst beim normalen Gehen mit mittlerer Geschwindigkeit etwa 0-10 % des Schrittzklus (SZ) und endet, wenn die Zehe des anderen Fußes vom Boden abhebt. Die zweite Phase (*Midstance*) enthält den Zeitraum, bis sich der Schwerpunkt des Körpers über dem betrachteten Vorderfuß befindet (10-30% SZ). In der dritten Phase (*Terminal Stance*) ist nach wie vor nur der betrachtete Fuß am Boden (30-50% des SZ). Das Aufsetzen des anderen Fußes leitet zur vierten Phase über, wobei der

betrachtete Fuß mit der Ferse abhebt (*Pre Swing*, 50-60% SZ). Der andere Fuß befindet sich zu dieser Zeit in der ersten Phase. Wenn die Zehe des betrachteten Fußes abhebt, beginnt die fünfte Phase (*Initial Swing*). Diese endet, wenn sich das Schwungbein auf Höhe des Standbeins befindet (60-73% SZ). Die sechste Phase (*Mid Swing*, 73-87% SZ) umfasst den Zeitraum, bis das Schienbein senkrecht zum Boden steht. In der abschließenden siebenten Phase (*Terminal Swing*, 87-100% SZ) wird die Schwungphase bis hin zum Aufsetzen der Ferse beendet.

Bei anderen Geschwindigkeiten kann die prozentuale Phasenaufteilung deutlich von den angegebenen Werten abweichen. Beispielsweise sind die Phasen 1 und 4, bei denen sich beide Füße am Boden befinden, bei langsamem Gehen tendenziell länger. Beim schnellen Laufen werden diese Phasen hingegen kürzer. Zudem ist die angegebene prozentuale Phasenaufteilung nur ein Durchschnittswert und weist individuelle Unterschiede auf.

Die datenbasierte Klassifikation von Schrittzyklen wird im folgenden Abschnitt vorgestellt. Weil eine manuelle Klassifikation der Ausgangsklassen für den Datensatz nicht zur Verfügung stand, wurde für den Klassifikatorentwurf (Lernvorgang) die in [88] angegebene prozentuale Aufteilung angesetzt.

Kapitel 3

Datenbasis und Vorverarbeitung

3.1 Datenbasis

Das Ziel der hier beschriebenen ersten Projektphase besteht darin, anhand einiger gut überschaubarer und illustrativer Beispiele die Leistungsfähigkeit von datenbasierten Algorithmen zu evaluieren. Dazu stand ein Beispieldatensatz mit Normdaten zur Bewegung auf dem Laufband zur Verfügung. Er beinhaltet Messungen mit verschiedenen Laufgeschwindigkeiten (mittel und schnell) sowie mit unterschiedlichen Gewichtsentlastungen zwischen 0 und 40% (Tabelle 3.1), die für verschiedene Probanden (Tabelle 3.2) aufgenommen wurden. Die ermittelten Datensätze bestehen jeweils aus Messungen mit 87 Merkmalen mit Kinematik- und Kinetikdaten. Tabelle 3.3 enthält die Bezeichnungsbausteine. In der Zusammensetzung bezeichnet beispielsweise R-KNEE Flex Ang den Flexionswinkel des rechten Knies.

Klassen-Nr.	Messung Nr. (alt)	Geschwindigkeit	Barren	Entlastung [%]	Anzahl Datensätze
-	1	mittel	nein	0	-
1	2	mittel	ja	0	2577
-	3	langsam	nein	0	-
2	4	langsam	ja	0	568
-	5	schnell	nein	0	-
3	6	schnell	ja	0	199
4	7	mittel	ja	10	262
-	8	langsam	ja	10	-
-	9	schnell	ja	10	-
-	10	mittel	ja	25	-
5	11	langsam	ja	25	579
-	12	schnell	ja	25	-
6	13	mittel	ja	40	273
-	14	langsam	ja	40	-
7	15	schnell	ja	40	186

Tabelle 3.1: Kodierung der Versuchsreihen bei der Erkennung von Laufgeschwindigkeiten und Gewichtsentlastungen

Klassen-Nr.	Personen-Nr.	Anzahl Datensätze
1	10001	325
2	10023	272
3	10024	280
4	10025	2613
5	10026	287
6	10027	275
7	10028	273
8	10029	319

Tabelle 3.2: Kodierung der Versuchsreihen bei der Erkennung von Personen

Codebestandteil	Bedeutung
L-	Links
R-	Rechts
ANK	Fußgelenk
KNEE	Kniegelenk
HIP	Hüfte
Pelvis	Becken
ELBOW	Ellenbogengelenk
Shl	Schultergelenk
Trunk	oberer Rumpf (Wirbelsäule als 'Gelenk')
Head	Position zwischen Kopf und Körper (Hals als 'Gelenk')
Null	Merkmal nicht belegt
GRF	Kraft auf Kraftmessplatte [N]
ANG	Winkel [°]
FRC	Kraft [N]
MOM	Moment [Nm]
PWR	Arbeit [Watt]
Flex	Flexion (positiver Winkel in Sagittalebene), Beugung
Rot	Rotation (Winkel in Transversalebene)
Abd	Abduktion (positiver Winkel in Frontalebene), Bewegung, die sich von Mittellinie entfernt
Tilt	Kippen, Neigung
Lat	lateral (seitwärts)
Fwd	forward (vorwärts)
Vrt	vertikal (senkrecht)
V	Änderungsgeschwindigkeit

Tabelle 3.3: Textbausteine zur Bezeichnung der Merkmale

3.2 Signalvorverarbeitung und abgeleitete Merkmale

Aus den vorliegenden 87 Merkmalen können zusätzliche Merkmale berechnet werden. Eine wichtige Information ist dabei die zeitliche Änderung eines Merkmals x_l zwischen den

Zeitpunkten $i - 1$ und $i + 1$, die mittels

$$\Delta x_{li} = \frac{1}{2} ((x_{l,i+1} - x_{li}) + (x_{li} - x_{l,i-1})) \quad (3.1)$$

$$= \frac{1}{2} (x_{l,i+1} - x_{l,i-1}) \quad (3.2)$$

abgeschätzt wird¹. Dieser Wert ist bei äquidistanter Abtastung proportional der Änderungsgeschwindigkeit des jeweiligen Merkmals und wird in der Bezeichnung durch ein nachgestelltes V charakterisiert. Wenn diese Änderung für alle Merkmale berechnet wird, ergeben sich nun 174 Merkmale. Eine vollständige Liste aller Merkmalsnummern und -bezeichnungen ist in Anhang A.1 zu finden.

Der Datensatz umfasst jeweils ca. zehn Schritte von acht gesunden Probanden bei unterschiedlichen Geschwindigkeiten (selbst gewählte Normalgeschwindigkeit, langsames Gehen, schnelles Gehen). Der Datensatz wird durch zufällige Auswahl in einen Lerndatensatz und einen Testdatensatz zur Validierung aufgespalten.

Die nachfolgenden Klassifikationsaufgaben erfordern eine Erkennung der Schrittlängen und des Anfangs neuer Schrittzyklen. Diese Erkennung beruht auf der Auswertung des Flexionswinkels der linken Hüfte (Merkmal x_{15}) sowie des Flexionswinkels des rechten Knies (Merkmal x_6) und ist in Bild 3.1 verdeutlicht. In einem ersten Schritt werden lokale Minima des Merkmals x_{15} gesucht. Dabei werden Minima nicht berücksichtigt, die innerhalb von fünf Abtastwerten vor oder nach einem lokalen Minimum mit kleinerem Wert auftreten. Die so gefundenen Minima sind im Bild durch * gekennzeichnet. Die Aufgabe wird dadurch erleichtert, dass Merkmal x_{15} meist nur ein ausgeprägtes Minimum innerhalb eines Schrittzyklusses aufweist.

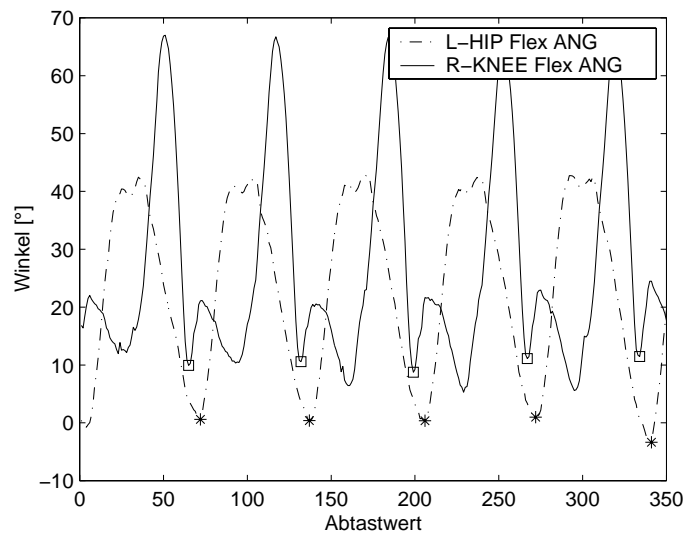


Bild 3.1: Erkennung von Schrittzyklen durch Auswertung der Merkmale x_6 und x_{15}

Damit resultiert eine erste grobe Schätzung der Schrittzyklen und ihrer (zeitlichen) Längen, die nun mit Hilfe des Merkmals x_6 verfeinert wird.

In der Nachbarschaft von einem Viertel der geschätzten Länge des Schrittzyklusses vor dem gefundenen Minimum sucht der Algorithmus nun nach einem lokalen Minimum des

¹Dabei wird vorausgesetzt, dass die Merkmale im Lerndatensatz zeitlich sortiert sind und dass die Werte gestrichen werden, bei denen die Berechnung wegen $i = 0$ und $i = N + 1$ innerhalb eines Versuchs (gleiche Person, gleiche Geschwindigkeit) scheitert.

Flexionswinkels des rechten Knies. Unter lokalem Minimum wird dabei zur Unterdrückung von Rauscheffekten ein Wert verstanden, der innerhalb von fünf Abtastwerten (zwei vor und zwei nach dem betreffenden Wert) minimal ist.

Dieser Punkt ist im jeweiligen Kurvenverlauf von x_6 mit \square markiert. Er wird in allen nachfolgenden Berechnungen als Initial Contact des rechten Fußes innerhalb eines Schrittzylusses angenommen.

Der beschriebene Algorithmus ist besonders bei mittleren und schnellen Geschwindigkeiten leistungsfähig. Probleme ergeben sich eventuell bei geringen Geschwindigkeiten, weil hier das lokale Minimum von x_6 weniger stark ausgeprägt ist (Bild 3.2). Für pathologische Gangbilder sind modifizierte Algorithmen sinnvoll, die z. B. auf verschiedenen Algorithmen mit einem nachfolgenden Voting-System in Abhängigkeit der erreichten Ergebnisse beruhen. Ein solches Voting-System kann auf Maßen wie der Minimierung der Standardabweichung erkannter Schrittlängen aufbauen, wenn von der Hypothese vergleichbarer Schrittlängen der Testpersonen während eines Versuches ausgegangen wird.

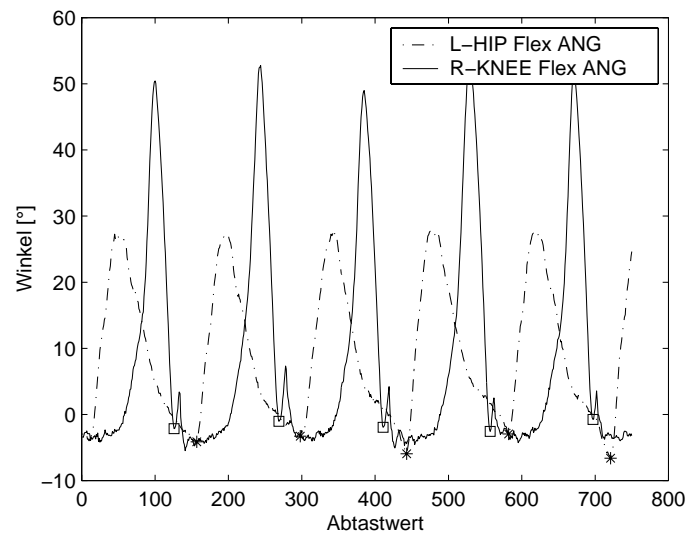


Bild 3.2: Probleme bei der Erkennung von Schrittzyklen durch Auswertung der Merkmale x_6 und x_{15} bei langsamer Gehgeschwindigkeit

Kapitel 4

Statistische Verfahren

4.1 Ziele

Mit der Anwendung klassischer statistischer Verfahren können zwei Ziele verfolgt werden. Zum Einen können statistische Verfahren Merkmale und Regeln liefern, die in ein Fuzzy-Modell einfließen und somit mit Hilfe der Computational Intelligence eine Entscheidungsfindung ermöglichen. Zum Anderen können diese Verfahren auch direkt das vorhandene Datenmaterial analysieren.

Im Folgenden steht die direkte Auswertung des Datenmaterials im Vordergrund. Dabei wird in einem vierstufigen Verfahren vorgegangen:

- Merkmalsselektion (Auswahl der wichtigsten Merkmale),
- dimensionsreduzierende Merkmalstransformation (Berechnung einer geringeren Anzahl von neuen, transformierten Merkmalen aus den ausgewählten Merkmalen),
- Klassifikation im Raum dieser neuen Merkmale und
- Interpretation der neuen Merkmale.

Die einzelnen Schritte werden in den folgenden Abschnitten näher erläutert.

4.2 Merkmalsselektion

In einem ersten Arbeitsschritt sind zunächst wichtige Merkmale für die Klassifikation herauszufinden. Das Ziel besteht dabei darin, eine kleine Anzahl s_m (z. B. $s_m = 6$) aus den potentiellen $s = 174$ Merkmalen auszuwählen und dabei möglichst wenig Information zu verlieren. Die Nummern der so ausgewählten Merkmale sollen in eine Indexmenge \mathcal{I} geschrieben werden. Dazu wird mit dem MANOVA-Verfahren (Multivariate ANalysis Of VAriance [2]) nach einer Gruppe von s_m Merkmalen gesucht, die *in ihrem Zusammenwirken* besonders wichtig sind. Dieses Vorgehen liefert bessere Ergebnisse als das Heraussuchen der wichtigsten Merkmale ohne die Analyse ihres Zusammenwirkens, weil Redundanzen (korrelierte Merkmale) berücksichtigt werden. Ein Experte kann zwar verbale Hinweise geben (Im Beispiel: Betrachtung von der Seite wichtiger als von vorn, Gelenkwinkel am Knie wichtiger als Gelenkwinkel am Fußgelenk usw.), eine quantitative Betrachtung oder gar eine Analyse von Redundanzen überfordert jedoch den medizinischen Experten.

Das Datenmaterial umfasst je N Messwerte des s -dimensionalen Merkmalsvektors \mathbf{x} und der Ausgangsgröße y (hier: Schrittphase), die in einer Matrix der Eingangsgrößen \mathbf{X}

der Dimension (N, s)

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1s} \\ x_{21} & x_{22} & \cdots & x_{2s} \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Ns} \end{pmatrix} = \begin{pmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_N^T \end{pmatrix} \quad (4.1)$$

und einem Vektor der Ausgangsgrößen

$$\mathbf{y} = (y_1 \quad y_2 \quad \cdots \quad y_N)^T \quad (4.2)$$

der Dimension $(N, 1)$ angeordnet werden. Die Matrix \mathbf{X} enthält reellwertige Elemente und der Vektor \mathbf{y} enthält diskrete Werte B_j , $j = 1, \dots, m_y$.

Das MANOVA-Verfahren beruht auf der Annahme, dass jede Klasse der Ausgangsgröße (Schrittphase) näherungsweise durch eine s -dimensionale Normalverteilung der reellwertigen Merkmale \mathbf{x} beschrieben werden kann. Die Bezeichnung B_j kodiert hier die Nummer der Schrittphase und kann folglich die Werte Eins bis Sieben annehmen. Die s -dimensionale Wahrscheinlichkeitsdichte für die j -te Ausgangsklasse im Eingangsraum der Merkmale wird somit durch

$$p(\mathbf{x}|y = B_j) = \frac{1}{(2\pi)^{\frac{s}{2}} \sqrt{\det(\boldsymbol{\Sigma}_j)}} \cdot e^{(-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}_j^{-1}(\mathbf{x}-\boldsymbol{\mu}_j))} \quad (4.3)$$

angegeben. Dabei sind $\boldsymbol{\Sigma}_j$ die Kovarianzmatrix und $\boldsymbol{\mu}_j$ der Vektor der Erwartungswerte der Merkmale für die j -te Ausgangsklasse.

Eine Gruppe von ausgewählten Merkmalen \mathcal{I} ist umso aussagekräftiger, je dichter die Messwerte der gleichen Klasse zusammen liegen und je weiter Messwerte unterschiedlicher Klassen auseinander liegen. Äquivalent zur zweiten Forderung ist die Aussage, dass die Mittelwerte der Messwerte unterschiedlicher Klassen möglichst weit auseinander liegen.

Um ein Maß für diese Forderungen zu gewinnen, werden aus der Matrix der Eingangsgrößen \mathbf{X} diejenigen s_m Merkmale ausgesucht, die in der Indexmenge \mathcal{I} enthalten sind. Daraus ergibt sich eine neue Matrix

$$\tilde{\mathbf{X}} = \begin{pmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \cdots & \tilde{x}_{1s_m} \\ \tilde{x}_{21} & \tilde{x}_{22} & \cdots & \tilde{x}_{2s_m} \\ \vdots & \vdots & & \vdots \\ \tilde{x}_{N1} & \tilde{x}_{N2} & \cdots & \tilde{x}_{Ns_m} \end{pmatrix} = \begin{pmatrix} \tilde{\mathbf{x}}_1^T \\ \tilde{\mathbf{x}}_2^T \\ \vdots \\ \tilde{\mathbf{x}}_N^T \end{pmatrix} \quad (4.4)$$

mit N Messwerten und s_m Merkmalen. Das Auswählen der s_m Merkmale aus der Indexmenge \mathcal{I} lässt sich auch durch Multiplizieren mit einer Transformationsmatrix $\tilde{\mathbf{V}}$ der Dimension (s, s_m) darstellen, die in jeder Spalte genau eine Eins für das zugehörige Element der Indexmenge \mathcal{I} enthält und sonst nur Nullelemente hat:

$$\tilde{\mathbf{X}} = \mathbf{X} \cdot \tilde{\mathbf{V}}, \quad \tilde{v}_{ij} = \begin{cases} 1 & \text{wenn } \mathcal{I}_j = i \\ 0 & \text{sonst.} \end{cases} \quad (4.5)$$

Die Schätzung der Mittelwerte der Merkmale lautet

$$\bar{\mathbf{x}}^T = \frac{1}{N} \mathbf{1}_{1,N} \cdot \tilde{\mathbf{X}} \quad (4.6)$$

und die der s_m -dimensionalen Kovarianzmatrix der Merkmale $\tilde{\mathbf{S}}$ (Maximum-Likelihood-Schätzung) berechnet sich mit

$$\tilde{\mathbf{S}} = \frac{1}{N} \cdot (\tilde{\mathbf{X}} - \mathbf{1}_{N,1} \cdot \tilde{\mathbf{x}}^T)^T \cdot (\tilde{\mathbf{X}} - \mathbf{1}_{N,1} \cdot \tilde{\mathbf{x}}^T) \quad (4.7a)$$

$$= \frac{1}{N} \cdot (\tilde{\mathbf{X}} - \frac{1}{N} \mathbf{1}_{N,N} \cdot \tilde{\mathbf{X}})^T \cdot (\tilde{\mathbf{X}} - \frac{1}{N} \mathbf{1}_{N,N} \cdot \tilde{\mathbf{X}}) \quad (4.7b)$$

$$= \frac{1}{N} \cdot \tilde{\mathbf{X}}^T \cdot \mathbf{Z}_{N,N} \cdot \tilde{\mathbf{X}} \quad \text{mit} \quad \mathbf{Z}_{N,N} = \mathbf{I}_{N,N} - \frac{1}{N} \mathbf{1}_{N,N}. \quad (4.7c)$$

Dabei bezeichnet \mathbf{I} die Einheitsmatrix und $\mathbf{1}$ eine Matrix, die aus Eins-Elementen besteht.

Für jede der m_y Klassen der Ausgangsgröße (im Beispiel $m_y = 7$) werden jetzt alle N_j Datensätze herausgesucht ($N = \sum_{j=1}^{m_y} N_j$), die zur j -ten Klasse gehören, und in der Matrix $\tilde{\mathbf{X}}_j$ zusammengefasst. Deren Mittelwertvektor berechnet sich aus

$$\tilde{\mathbf{x}}_j^T = \frac{1}{N_j} \mathbf{1}_{1,N_j} \cdot \tilde{\mathbf{X}}_j \quad (4.8)$$

und die Klassenkovarianzmatrix aus

$$\tilde{\mathbf{S}}_j = \frac{1}{N_j} \cdot \tilde{\mathbf{X}}_j^T \cdot \mathbf{Z}_{N_j,N_j} \cdot \tilde{\mathbf{X}}_j. \quad (4.9)$$

Aus diesen Schätzungen lassen sich die Gesamtvariationsmatrix $\tilde{\mathbf{T}}$ (Total Variance) als Maß für die Streuung des gesamten Datenmaterials, die Innerklassenvariationsmatrix $\tilde{\mathbf{W}}$ (Within-Groups-Variance) für die Streuung innerhalb der einzelnen Klassen und die Zwischenklassenvariationsmatrix $\tilde{\mathbf{B}}$ (Between-Groups-Variance) für die Streuung zwischen den Klassen gewinnen:

$$\tilde{\mathbf{T}} = N \cdot \tilde{\mathbf{S}} \quad \tilde{\mathbf{W}} = \sum_{j=1}^{m_y} N_j \cdot \tilde{\mathbf{S}}_j \quad (4.10)$$

$$\tilde{\mathbf{B}} = \sum_{j=1}^{m_y} N_j \cdot \left(\frac{1}{N_j} \mathbf{1}_{1,N_j} \cdot \tilde{\mathbf{X}}_j - \frac{1}{N} \mathbf{1}_{1,N} \cdot \tilde{\mathbf{X}} \right)^T \cdot \left(\frac{1}{N_j} \mathbf{1}_{1,N_j} \cdot \tilde{\mathbf{X}}_j - \frac{1}{N} \mathbf{1}_{1,N} \cdot \tilde{\mathbf{X}} \right). \quad (4.11)$$

Dabei gilt der Streuungszzerlegungssatz $\tilde{\mathbf{T}} = \tilde{\mathbf{B}} + \tilde{\mathbf{W}}$. Je größer $\tilde{\mathbf{B}}$ im Verhältnis zu $\tilde{\mathbf{W}}$ ist, desto besser eignen sich die Merkmale und erfüllen die oben genannte Forderung zur Lage von Messungen gleicher und unterschiedlicher Klassen. Die Lösung des Problems führt auf ein verallgemeinertes Eigenwertproblem, das sich bei Invertierbarkeit von $\tilde{\mathbf{W}}$ in ein klassisches Eigenwertproblem umwandeln lässt:

$$(\tilde{\mathbf{B}} - \lambda_i \tilde{\mathbf{W}}) \mathbf{v}_i = \mathbf{0} \quad (\tilde{\mathbf{W}}^{-1} \tilde{\mathbf{B}} - \lambda_i \mathbf{I}) \mathbf{v}_i = \mathbf{0}. \quad (4.12)$$

Auf den geordneten Eigenwerten ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{s_m} \geq 0$) von (4.12) bauen nun verschiedene Maße $M_{\mathcal{I}}$ zur Beurteilung der in \mathcal{I} enthaltenen Merkmale auf. Das Ziel dieser Maße ist es, große Eigenwerte zu favorisieren, wobei es dabei insbesondere auf den oder die ersten Eigenwerte ankommt. Im Sonderfall $s_m = 1$ entsprechen alle Tests dem ANOVA-Verfahren und dem t -Test. Häufig verwendet wird das Likelihood-Quotienten-Maß:

$$M_{\mathcal{I}} = 1 - \prod_{i=1}^{s_m} \frac{1}{1 + \lambda_i}. \quad (4.13)$$

Merkmal	Bezeichnung	Güte	Merkmal	Bezeichnung	Güte
x_3	R-HIP Flex ANG	0.726	x_3	R-HIP Flex ANG	0.726
x_6	R-KNEE Flex ANG	0.721	$+x_6$	R-KNEE Flex ANG	0.939
x_{90}	R-HIP Flex ANG V	0.718	$+x_{102}$	L-HIP Flex ANG V	0.977
x_{102}	L-HIP Flex ANG V	0.700	$+x_{105}$	L-KNEE Flex ANG V	0.991
x_{15}	L-HIP Flex ANG	0.680	$+x_{93}$	R-KNEE Flex ANG V	0.995
x_{93}	R-KNEE Flex ANG V	0.639	$+x_{15}$	L-HIP Flex ANG	0.997

Tabelle 4.1: Beste Merkmalsrelevanzen der Einzelmerkmale (links) und von Merkmalsgruppen (rechts) mit dem Likelihood-Quotienten-Maß – Bezeichnungen: Knie (KNEE), Hüfte (HIP), links (L), rechts (R), Flexion in der Sagittalebene (Flex), Winkel (ANG) und Änderungsgeschwindigkeit (V)

Mit dem Kriterium können nun verschiedene Hypothesen sinnvoller Merkmalskombinationen \mathcal{I} geprüft werden. Ein suboptimales Verfahren mit geringem Rechenaufwand sucht zunächst das beste Einzelmerkmal ($s_m = 1$) heraus. Anschließend wird iterativ zur bisher gefundenen besten Merkmalskombination immer das Merkmal gesucht, das den Wert von (4.13) maximiert.

Im Beispiel führt diese Auswahl dazu, dass unter den sechs wichtigsten Einzelmerkmalen und Merkmalskombinationen nur Gelenkwinkel und ihre Änderungsgeschwindigkeiten bei einer Betrachtung von der Seite (sagittal) im Knie- und Hüftbereich zu finden sind (Tabelle 4.1). Das relevanteste Einzelmerkmal ist der Beugungswinkel des Hüftgelenks auf der rechten Seite (x_3 , Winkel zwischen rechtem Oberschenkel und Oberkörper). Das dazu am besten passende Merkmal ist der Beugungswinkel des rechten Kniegelenks (x_6). Das dritte Merkmal ist nicht das drittbeste Einzelmerkmal x_{90} (Änderungsgeschwindigkeit des rechten Hüftwinkels), weil dieses wiederum sehr stark mit Merkmal x_6 korreliert ist und somit kaum neue Informationen liefert. Stattdessen werden nacheinander die Merkmale x_{102} (Winkelgeschwindigkeit der linken Hüftseite, 4. Einzelmerkmal), x_{105} (Winkelgeschwindigkeit des linken Knies, 7. Einzelmerkmal), x_{93} (Winkelgeschwindigkeit des rechten Knies, 6. Einzelmerkmal) und x_{15} (Winkel der linken Hüftseite, 5. Einzelmerkmal) hinzugefügt. Die erreichten Werte der Merkmalsrelevanzen auf einer Skala zwischen 0 und 1 zeigen, dass diese Kombination aus sechs Merkmalen relevante Ergebnisse liefert.

Andere Einzelmerkmale (Betrachtung von vorn und oben, Kräfte usw.) werden drastisch schlechter bewertet. Damit liefert diese Merkmalsselektion auch einen Beitrag zur Interpretierbarkeit der Lösungen, indem sie über eine sortierte Merkmalsliste Auskunft über wesentliche, redundante und irrelevante Merkmale gibt.

4.3 Dimensionsreduzierende Merkmalstransformation

Bei dimensionsreduzierenden Verfahren zur Merkmalstransformation werden s_d neue Merkmale aus Linearkombinationen der ursprünglichen Merkmale gebildet. Die Erfolgchancen einer manuellen Suche nach solchen Merkmalen durch Experten sind ebenfalls sehr gering. Wichtig ist hierbei, dass durch die geringere Anzahl neuer Merkmale eine Visualisierung der Ergebnisse möglich und die nachfolgende Klassifikation erleichtert wird.

Diese dimensionsreduzierende Transformation kann auf der Basis der s ursprünglichen Merkmale (einstufige Transformation) oder der s_m ausgewählten Merkmale (zweistufige Transformation) erfolgen.

Die Verfahren werden im Folgenden beispielhaft anhand der zweistufigen Transformation vorgestellt, können aber anhand einer formalen Ersetzung der entsprechenden Matrizen

(\mathbf{X} statt $\tilde{\mathbf{X}}$, \mathbf{S} statt $\tilde{\mathbf{S}}$, \mathbf{B} statt $\tilde{\mathbf{B}}$, \mathbf{W} statt $\tilde{\mathbf{W}}$ usw.) auf den einstufigen Fall übertragen werden.

Die Matrix der Messwerte $\tilde{\mathbf{X}}$ wird mit einer Transformationsmatrix \mathbf{V} der Dimension (s_m, s_d) multipliziert:

$$\tilde{\mathbf{X}}_D = \tilde{\mathbf{X}} \cdot \mathbf{V}. \quad (4.14)$$

Damit entsteht aus Linearkombinationen der ursprünglichen Merkmale ein niederdimensionaler ($s_d \leq s_m$) Raum neuer Merkmale. Auf diese Art werden zunächst durch die im vorhergehenden Abschnitt beschriebene Merkmalsselektion aus den ursprünglichen s Merkmalen s_m Merkmale ausgewählt, aus denen mittels einer Linearkombination s_d neue Merkmale resultieren.

Als Verfahren werden die Hauptkomponentenanalyse und die Diskriminanzanalyse untersucht.

Bei der *Diskriminanzanalyse* ist die Transformationsmatrix

$$\mathbf{V} = (\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_{s_d}) \quad (4.15)$$

die Matrix der ersten $i = 1, \dots, s_d$ Eigenvektoren \mathbf{v}_i des Problems (4.12), die zu den s_d größten Eigenwerten gehören.

Bei der *Hauptkomponentenanalyse* (Karhunen-Loeve-Transformation) wird anstelle der Matrix $\tilde{\mathbf{W}}^{-1} \tilde{\mathbf{B}}$ in (4.12) nur die Gesamtvariationsmatrix $\tilde{\mathbf{T}}$ betrachtet. Das ist besonders bei unbekanntem Ausgangsklassen sinnvoll oder wenn zunächst die Menge potentieller Merkmale reduziert werden soll. Hier wird die Matrix der ersten $i = 1, \dots, s_d$ Eigenvektoren

$$\mathbf{V} = (\tilde{\mathbf{v}}_1 \quad \cdots \quad \tilde{\mathbf{v}}_{s_d}) \quad (4.16)$$

eines Problems

$$(\tilde{\mathbf{T}} - \lambda_i \mathbf{I}) \tilde{\mathbf{v}}_i = \mathbf{0} \quad (4.17)$$

verwendet, die zu den s_d größten Eigenwerten dieses Eigenwertproblems gehören.

Aus Sicht der Klassifikation ist die Diskriminanzanalyse vorzuziehen, wenn der Lern Datensatz bekannte Ausgangsklassen enthält. Im Unterschied zu anderen Verfahren (z. B. Hauptkomponentenanalyse) gehen sowohl die Eingangsgrößen als auch die bekannten Ausgangsgrößen (Klassen – hier: Schrittphasen) des Lern Datensatzes ein.

Verfahren	Eigenwertproblem	Ausgangsklassen	notwendige Beispielanzahl
Diskriminanzanalyse	$(\tilde{\mathbf{W}}^{-1} \tilde{\mathbf{B}} - \lambda \mathbf{I}) \mathbf{v} = \mathbf{0}$	bekannt	groß
Hauptkomponentenanalyse	$(\tilde{\mathbf{T}} - \lambda \mathbf{I}) \tilde{\mathbf{v}} = \mathbf{0}$	unbekannt	mittel

Tabelle 4.2: Vergleich Hauptkomponenten- und Diskriminanzanalyse

Einen Vergleich zwischen Diskriminanz- und Hauptkomponentenanalyse nimmt Tabelle 4.2 vor. Die Diskriminanzanalyse stellt wegen der Berechnung von Innerklassenkovarianzmatrizen hohe Anforderungen an die verfügbaren Datenmengen, insbesondere bei

Problemen mit vielen Ausgangsklassen. Bei zu geringen Datenmengen fällt die Klassifikationsgüte über Testdaten deutlich ab. Deshalb ist der in der Literatur häufig angegebene Weg, eine Diskriminanzanalyse ohne vorherige Merkmalsselektion durchzuführen ($s_m = s$) nur bei Probleme mit großen Datenmengen einsetzbar. Auswege bieten das hier verwendete zweistufige Verfahren mit vorgeschalteter Merkmalsselektion ($s_m \ll s$) oder ein zweistufiges Verfahren mit einer vorgeschalteten Hauptkomponentenanalyse zur Merkmalsreduktion und einer nachgeschalteten Diskriminanzanalyse.

Das zweistufige Vorgehen mit vorgeschalteter Merkmalsselektion lässt sich auch als eine Einschnittstransformation

$$\tilde{\mathbf{X}}_D = \mathbf{X} \cdot \mathbf{V}_{ges} = \mathbf{X} \cdot \tilde{\mathbf{V}} \cdot \mathbf{V} \quad (4.18)$$

darstellen, bei dem sich die Gesamttransformationsmatrix \mathbf{V}_{ges} aus der Transformationsmatrix zur Merkmalsselektion $\tilde{\mathbf{V}}$ (Dimension $s \times s_m$ mit je einer Eins pro Spalte in den Zeilen der ausgewählten Merkmale und sonst Nullen) und der Transformationsmatrix der Diskriminanzanalyse \mathbf{V} ergibt.

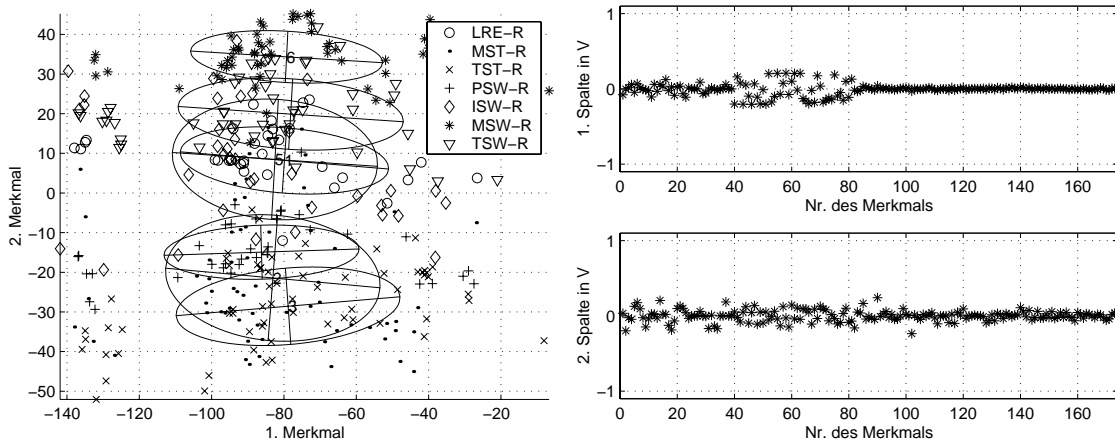


Bild 4.1: Klassifikation der Schrittphasen in einem zweidimensionalen Merkmalsraum ($\tilde{\mathbf{x}}_{H,1}, \tilde{\mathbf{x}}_{H,2}$) nach einer Hauptkomponentenanalyse auf zwei Merkmale ohne vorgeschaltete Merkmalsselektion (links) und zugehörige Transformationsmatrix \mathbf{V}_{ges} (rechts)

Zu unbefriedigenden Ergebnissen führt eine alleinige Hauptkomponentenanalyse mit $s_d = 2$ ohne vorherige Merkmalsselektion (Bild 4.1, links). Zwar sind noch Unterschiede in den Mittelwerten der Schrittphasen zu sehen, aber der Schrittzyklus ist nicht erkennbar. Zudem sind die Phasen nicht voneinander trennbar. Im rechten Teilbild sind die Elemente der Transformationsmatrix \mathbf{V} über der Nummer der Merkmale dargestellt (rechts oben: 1. Spalte, rechts unten: 2. Spalte), die beim Lösen des Problems $(\mathbf{T} - \lambda_i \mathbf{I})\tilde{\mathbf{v}}_i = \mathbf{0}$ entsteht. Hier zeigt sich, dass nahezu alle Merkmale in die Entscheidung eingehen, was die Interpretierbarkeit reduziert.

Etwas günstigere Ergebnisse als eine alleinige Hauptkomponentenanalyse sichert eine Hauptkomponentenanalyse mit $s_d = 2$ nach vorheriger Merkmalsselektion mit dem MANOVA-Verfahren (Bild 4.2, links). Ein Schrittzyklus ist in Ansätzen zu erkennen, wobei die Werte bei unterschiedlichen Geschwindigkeiten deutlich voneinander abweichen. Deswegen sind die Phasen nur mit erheblichen Fehlern voneinander trennbar. Im rechten Teilbild sind die Elemente der Transformationsmatrix \mathbf{V} über der Nummer der Merkmale dargestellt (rechts oben: 1. Spalte, rechts unten: 2. Spalte), die aus (4.18) mit (4.5) und (4.16) folgt. Nur die sechs ausgewählten Merkmale gehen in die Entscheidung ein.

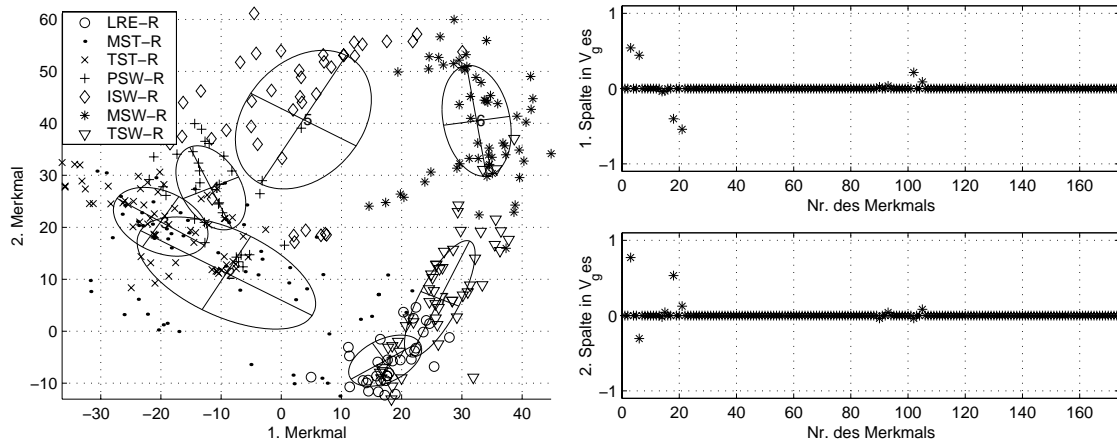


Bild 4.2: Klassifikation der Schrittphasen in einem zweidimensionalen Merkmalsraum $(\tilde{\mathbf{x}}_{H,1}, \tilde{\mathbf{x}}_{H,2})$ nach einer Hauptkomponentenanalyse auf zwei Merkmale mit vorgeschalteter Merkmalsselektion (links) und zugehörige Gesamttransformationsmatrix \mathbf{V}_{ges} (rechts)

Eine Diskriminanzanalyse ohne vorherige Maßnahmen ist nicht durchführbar, weil die Daten nicht zur Berechnung der klassenspezifischen Kovarianzmatrizen ausreichen.

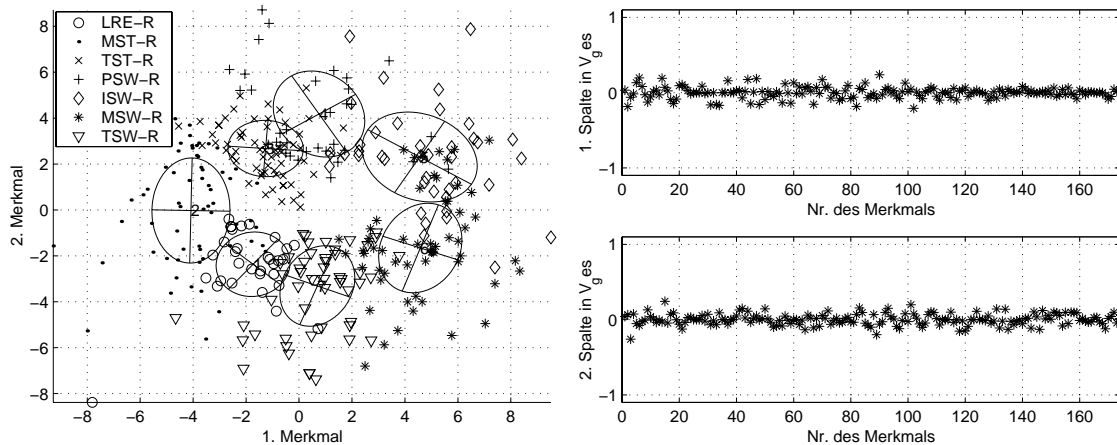


Bild 4.3: Klassifikation der Schrittphasen in einem zweidimensionalen Merkmalsraum $(\tilde{\mathbf{x}}_{D,1}, \tilde{\mathbf{x}}_{D,2})$ nach einer Hauptkomponentenanalyse auf sechs Merkmale mit nachgeschalteter Diskriminanzanalyse (links) und zugehörige Gesamttransformationsmatrix \mathbf{V}_{ges} (rechts)

Relativ gute Ergebnisse liefert eine Hauptkomponentenanalyse ($s_d = 6$) mit nachfolgender Diskriminanzanalyse mit $s_d = 2$ (Bild 4.3, links). Ein Schrittzyklus ist zu erkennen, wobei die Werte bei unterschiedlichen Geschwindigkeiten immer noch voneinander abweichen. Deswegen sind die Phasen nur mit Fehlern voneinander trennbar. Im rechten Teilbild sind die Elemente der Transformationsmatrix \mathbf{V} über der Nummer der Merkmale dargestellt (rechts oben: 1. Spalte, rechts unten: 2. Spalte). Weil alle Merkmale bei der Hauptkomponentenanalyse eingehen, sind auch in der Gesamttransformationsmatrix alle Merkmale beteiligt – mit den genannten Nachteilen bei der Interpretation der Ergebnisse.

Die mit Abstand besten Ergebnisse für das Beispiel ergeben sich für das vorgestellte zweistufige Verfahren mit Merkmalsselektion und Dimensionsreduktion durch Diskriminanzanalyse mit $s_d = 2$ (Bild 4.4, links). Der Schrittzyklus ist deutlich durch die Kreis-

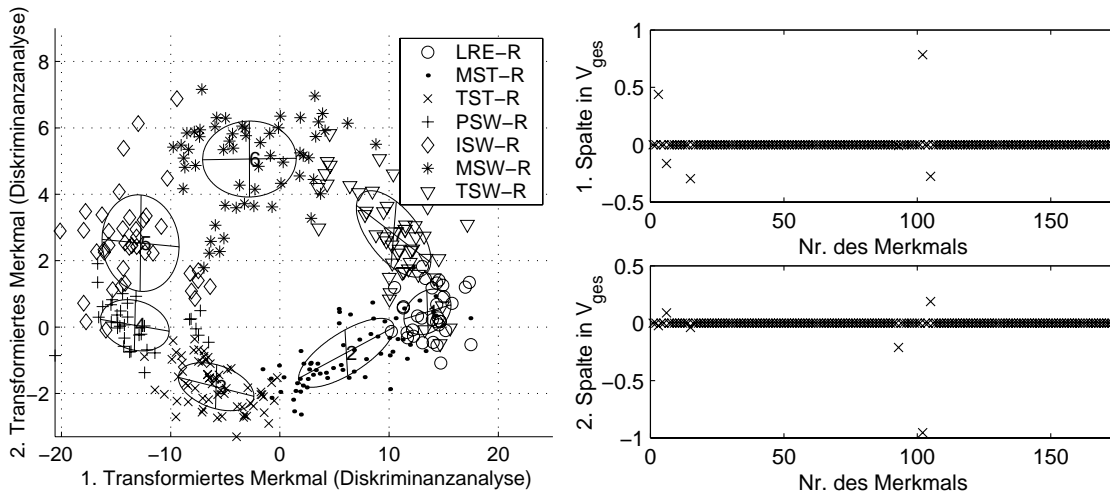


Bild 4.4: Klassifikation der Schrittphasen in einem zweidimensionalen Merkmalsraum ($\tilde{\mathbf{x}}_{D,1}, \tilde{\mathbf{x}}_{D,2}$) nach einer Diskriminanzanalyse mit vorgeschalteter Merkmalsreduktion auf sechs Merkmale (links) und zugehörige Gesamttransformationsmatrix \mathbf{V}_{ges} gemäß (4.18)

form der Datenprojektion erkennbar, beginnt unten rechts mit dem Zustand LRE-R und wird im Uhrzeigersinn durchlaufen. Die benachbarten Phasen sind qualitativ voneinander trennbar. Im rechten Teilbild sind die Elemente der Transformationsmatrix \mathbf{V}_{ges} über der Nummer der Merkmale dargestellt (rechts oben: 1. Spalte, rechts unten: 2. Spalte). Dabei ist festzustellen, dass nur in \mathcal{I} ausgewählte Merkmale Werte ungleich Null aufweisen.

4.4 Klassifikation mit Bayes-Klassifikatoren

Mit Hilfe eines Klassifikators muss nun für die aus den Messdaten berechneten transformierten Merkmale eine Entscheidung (hier: Welche Schrittphase?) getroffen werden. Die Idee eines Bayes-Klassifikators besteht darin, die anhand der vorliegenden Informationen wahrscheinlichste Entscheidung zu treffen. Diese Entscheidung kann prinzipiell anhand der Matrizen der Messwerte der ursprünglichen Merkmale \mathbf{X} , der ausgewählten Merkmale $\tilde{\mathbf{X}}$ oder der transformierten Merkmale $\tilde{\mathbf{X}}_D$ getroffen werden. Aufgrund der Einfachheit der Schreibweise werden alle nachfolgenden Untersuchungen anhand von \mathbf{X} demonstriert, können aber durch formales Ersetzen der jeweiligen Mittelwerte und geschätzten Kovarianzmatrizen auf andere Fälle übertragen werden.

Ein Bayes-Klassifikator schätzt die A-posteriori-Wahrscheinlichkeiten für eine Klassifikation

$$\hat{p}(y = B_j | \mathbf{x}) = \frac{\hat{p}(y = B_j) \cdot \hat{p}(\mathbf{x} | y = B_j)}{\sum_{k=1}^{m_y} \hat{p}(y = B_k) \cdot \hat{p}(\mathbf{x} | y = B_k)} \quad (4.19)$$

(Wie wahrscheinlich ist die Ausgangsklasse $y = B_j$ bei einem gegebenen Messwert \mathbf{x} ?) und wählt dann die Ausgangsklasse B_j mit der maximalen Wahrscheinlichkeit aus:

$$\hat{y} = \operatorname{argmax}_j (\hat{p}(y = B_j | \mathbf{x})). \quad (4.20)$$

Dabei gehen die A-priori-Wahrscheinlichkeiten der Ausgangsklassen $\hat{p}(y = B_j)$ (hier: Wie häufig kommt eine bestimmte Ausgangsklasse bzw. Schrittphase überhaupt vor?) ebenso in die Entscheidung ein wie die geschätzte Verteilung der Messwerte für jede Ausgangsklasse.

Werden mehrdimensionale Normalverteilungen entsprechend (4.3) mit den Schätzungen der Kovarianzmatrizen \mathbf{S}_j und der Mittelwerte $\bar{\mathbf{x}}_j$ angenommen, kann nun eine Entscheidung berechnet werden. Nach Einsetzen in (4.19) und Kürzen aller konstanten Faktoren ergibt sich

$$\hat{p}(y = B_j | \mathbf{x}) = \frac{\hat{p}(y = B_j) \det(\mathbf{S}_j)^{-1/2} \cdot e^{(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_j)^T \mathbf{S}_j^{-1} (\mathbf{x} - \bar{\mathbf{x}}_j))}}{\sum_{k=1}^{m_y} \hat{p}(y = B_k) \det(\mathbf{S}_k)^{-1/2} \cdot e^{(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_k)^T \mathbf{S}_k^{-1} (\mathbf{x} - \bar{\mathbf{x}}_k))}}. \quad (4.21)$$

Wenn nur die Entscheidung gesucht wird, muss wegen des konstanten Nennerterms nur der Zählerterm maximiert werden. Nach Logarithmieren, Durchmultiplizieren mit 2 und Vorzeichenumkehr wird ein Maß minimiert, das Parallelen zu einem Abstandsmaß aufweist:

$$\hat{y} = \operatorname{argmin}_j (\ln(\det(\mathbf{S}_j)) - 2 \ln(\hat{p}(y = B_j)) + (\mathbf{x} - \bar{\mathbf{x}}_j)^T \mathbf{S}_j^{-1} (\mathbf{x} - \bar{\mathbf{x}}_j)). \quad (4.22)$$

Allerdings kann dieses Maß wegen des Determinantenterms auch negativ werden.

Unter den folgenden Annahmen, deren Gültigkeit jeweils problemspezifisch zu prüfen ist, ergeben sich Vereinfachungen von (4.22):

- Die A-priori-Wahrscheinlichkeiten aller Ausgangsklassen sind gleich:

$$\hat{p}(y = B_j) = \frac{1}{m_y}. \quad (4.23)$$

Damit kann auf den Term mit den A-priori-Wahrscheinlichkeiten verzichtet werden, womit sich (4.22) vereinfacht:

$$\hat{y} = \operatorname{argmin}_j (\ln(\det(\mathbf{S}_j)) + (\mathbf{x} - \bar{\mathbf{x}}_j)^T \mathbf{S}_j^{-1} (\mathbf{x} - \bar{\mathbf{x}}_j)). \quad (4.24)$$

- Zusätzlich sind die Kovarianzmatrizen aller Ausgangsklassen gleich, aber nicht unbedingt Einheitsmatrizen:

$$\mathbf{S}_j = \mathbf{S}, \quad j = 1, \dots, m_y. \quad (4.25)$$

Damit beeinflusst der Determinantenterm das Ergebnis nicht mehr (Abstandsklassifikator mit Mahalanobis-Distanz):

$$\hat{y} = \operatorname{argmin}_j ((\mathbf{x} - \bar{\mathbf{x}}_j)^T \mathbf{S}^{-1} (\mathbf{x} - \bar{\mathbf{x}}_j)). \quad (4.26)$$

- Die Kovarianzmatrizen aller Ausgangsklassen haben Diagonalstruktur

$$\mathbf{S}_j = \sigma^2 \cdot \mathbf{I}, \quad j = 1, \dots, m_y, \quad (4.27)$$

woraus

$$\hat{y} = \operatorname{argmin}_j ((\mathbf{x} - \bar{\mathbf{x}}_j)^T (\mathbf{x} - \bar{\mathbf{x}}_j)) \quad (4.28)$$

resultiert (euklidischer Abstandsklassifikator).

Mit jeder Vereinfachung bleibt die mögliche Klassifikationsgüte über den Lerndaten bestenfalls gleich und fällt in der Regel sogar ab. Über den Testdaten ist die geeignete Vereinfachung erst zu bestimmen, weil jede Vereinfachung die Zahl der zu schätzenden Parameter reduziert und somit auch das Risiko einer Überanpassung an zu kleine Lerndatensätze sinkt.

Eine mögliche Erweiterung des Bayes-Konzeptes besteht hingegen in der Idee, nicht die Wahrscheinlichkeit, sondern die Kosten von Fehlentscheidungen zu minimieren.

Bei Verwendung von (4.19) können auch die Wahrscheinlichkeiten der Entscheidung und ihrer Alternativen geschätzt werden, was Möglichkeiten zur Beurteilung der Sicherheit einer Entscheidung und ihrer eventuellen Korrektur in nachfolgenden Verarbeitungsschritten eröffnet. Damit werden die Ergebnisse zudem zu den unscharfen Ergebnissen von Fuzzy-Klassifikatoren kompatibel.

Prinzipiell kann dieses Klassifikationskonzept bereits auf die Originaldaten mit s Merkmalen angewendet werden. Allerdings erfordert die Klassifikation dann die Bestimmung extrem vieler Parameter in den geschätzten Mittelwerten und Kovarianzmatrizen, so dass die erforderliche Menge an Lerndaten für eine brauchbare Schätzung jedes realisierbare Maß übersteigt.

Im Beispiel wird als einzige Vereinfachung von nahezu gleichen A-priori-Wahrscheinlichkeiten ausgegangen und somit mit klassenspezifischen Kovarianzmatrizen \mathbf{S}_j entsprechend (4.24) gearbeitet. Dabei ergeben sich für die verschiedenen Transformationen folgende Ergebnisse:

- Hauptkomponentenanalyse ohne vorhergehende Merkmalsselektion (Fehler Lerndatensatz: 49.8%, Fehler Testdatensatz 52.9 %, Bild 4.1, links),
- Hauptkomponentenanalyse mit vorhergehender Merkmalsselektion (Fehler Lerndatensatz: 33.6%, Fehler Testdatensatz 35.5%, Bild 4.2, links),
- Diskriminanzanalyse mit vorhergehender Hauptkomponentenanalyse (Fehler Lerndatensatz: 26.4%, Fehler Testdatensatz 29.8%, Bild 4.3, links) sowie
- Diskriminanzanalyse mit vorhergehender Merkmalsselektion (Fehler Lerndatensatz: 16.3%, Fehler Testdatensatz 18.2%, Bild 4.4, links).

Die mit Abstand besten Ergebnisse ergeben sich somit für eine Diskriminanzanalyse mit vorhergehender Merkmalsselektion. In Bild 4.4 (links) sind die Kovarianzmatrizen der einzelnen Schrittphasen zur Kennzeichnung der Metrik des Bayes-Klassifikators ebenfalls eingezeichnet. Die auftretenden Abweichungen zwischen Lerndatensatz und Klassifikatorergebnissen sind zum Teil durch die in Abschnitt 2.7 diskutierten Probleme erklärbar (individuelle Unterschiede, Abhängigkeiten von Laufgeschwindigkeiten) und betreffen bis auf 0.5% Fehler nur Abweichungen zur benachbarten Phase. Dennoch ergibt sich eine gute Klassifikation der Schrittphasen (Bild 6.1 in Kapitel 6). Bei der Verwendung der gleichen Transformation, aber eines Abstandsklassifikators gemäß (4.28), steigen die Fehler auf 18.6% für den Lerndatensatz und 21.0% über den Testdaten an.

Die Fehler der anderen Ansätze bestätigen den subjektiven visuellen Eindruck, der beim Vergleich der Bilder 4.1-4.4 gewonnen wurde.

4.5 Interpretation der transformierten Merkmale

Im rechten Teilbild von Bild 4.4 sind die Werte der Transformationsmatrix \mathbf{V}_{ges} dargestellt (Transformationsvorschrift für das erste Merkmal $\tilde{x}_{D,1}$ – oben – und das zweite Merkmal $\tilde{x}_{D,2}$ – unten). Sie hat nur für die ausgewählten Merkmale Werte ungleich Null und erlaubt noch Interpretationsversuche. So ist beispielsweise die Schwungphase (Schrittphasen 4, 5 und 6) durch kleine Werte des transformierten Merkmals $\tilde{x}_{D,1}$ charakterisiert, in das mit positivem Vorzeichen die Merkmale x_3 und x_{102} bzw. mit negativem Vorzeichen die Merkmale x_6 , x_{15} und x_{105} eingehen. Die Stellung der Gelenke ist in diesen Phasen tendenziell durch einen angewinkelten rechten Oberschenkel (rechtes Hüftgelenk, x_3), einen nach

Merkmal SZR	Bezeichnung	Merkmal SZL	Bezeichnung
x_3	R-HIP Flex ANG	x_{15}	L-HIP Flex ANG
x_6	R-KNEE Flex ANG	x_{18}	L-KNEE Flex ANG
x_{102}	L-HIP Flex ANG V	x_{90}	R-HIP Flex ANG
x_{105}	L-KNEE Flex ANG V	x_{93}	R-KNEE Flex ANG
x_{93}	R-KNEE Flex ANG V	x_{105}	L-KNEE Flex ANG
x_{15}	L-HIP Flex ANG	x_3	R-HIP Flex ANG

Tabelle 4.3: Äquivalente Merkmale für die Berechnung zwischen dem Schrittzzyklus der rechten Körperseite (SZR) und dem Schrittzzyklus der linken Körperseite (SZL)

hinten gestreckten linken Oberschenkel (linkes Hüftgelenk, x_{15}) und ein gebeugtes rechtes Knie (x_6) gekennzeichnet. Die Dynamik der Bewegung kommt durch eine Streckung im linken Hüftgelenk (x_{102}) und im linken Knie (x_{105}) zum Ausdruck.

Die so gefundene abstrakte Darstellung ist als Vorverarbeitungsstufe zur Informationsreduktion für den Vergleich zwischen verschiedenen Patientengruppen ebenso einsetzbar wie zum Vergleich der Auswirkungen unterschiedlicher Geschwindigkeiten im Gangbild (engerer Kreis in Bild 4.4 entspricht geringerer Geschwindigkeit).

4.6 Weitere Anwendungen

Mit Hilfe der gefundenen Merkmalsselektion und der transformierten Merkmale kann auf einer abstrahierten Ebene nun eine Reihe detaillierter Untersuchungen vorgenommen werden.

Eine nähere Betrachtung der ausgewählten Merkmale in Tabelle 4.1 zeigt, dass sie eine Koordination zwischen Knie- und Hüftgelenk beider Körperseiten bewerten. Da bisher alle Untersuchungen anhand der Schrittzzyklen der *rechten* Körperseite geführt wurden, kann mit Hilfe einer einfachen Inversion der Merkmale der Schrittzzyklus der *linken* Körperseite bestimmt werden. Dabei sind einfach die Seitenangaben in den Merkmalen auszutauschen und die Transformationsmatrizen entsprechend zu korrigieren (Tabelle 4.3).

Damit können nun u. a. Verläufe beider Körperseiten einer Person bei unterschiedlichen Geschwindigkeiten verglichen werden. Dazu wird wieder die Merkmalsselektion mit nachfolgender Diskriminanzanalyse verwendet, mit der die Messwerte für eine Person (hier: 10025) in einer zweidimensionalen Darstellung visualisiert werden können. Bild 4.5 zeigt einen Vergleich zwischen einer langsamen (Klasse 2 in Tabelle 3.1) und einer schnellen Geschwindigkeit (Klasse 3 in Tabelle 3.1) ohne Gewichtsentlastung. Durch die Werte für die linke und rechte Körperseite entstehen jetzt jeweils zwei Trajektorien für beide Geschwindigkeiten. Mit der Zuordnung der Loading Response zum unteren rechten Bereich der Verläufe und der Schwungphasen zum linken und oberen Bereich (vergleiche Bild 4.4, links) zeigt sich, dass die wesentlichen geschwindigkeitsabhängigen Unterschiede hauptsächlich in der Schwungphase auftreten. Mit größeren Geschwindigkeiten wird der Kreisdurchmesser im Vergleich zu kleineren Geschwindigkeiten deutlich größer.

Mit Hilfe dieser Transformation ist es ebenfalls möglich, für eine Person Unterschiede zwischen beiden Körperseiten zu bewerten. Drei Beispiele für unterschiedliche Personen zeigen die Bilder 4.6 (Person 10023), 4.7 (Person 10026) und 4.8 (Person 10028).

Während bei Person 10028 die Verläufe für beide Körperseiten nahezu deckungsgleich sind, treten bei den anderen Personen deutliche Unterschiede auf. Bei Person 10023 entsprechen im oberen rechten Bereich (Schwungphase) die Bewegungen der rechten Körper-

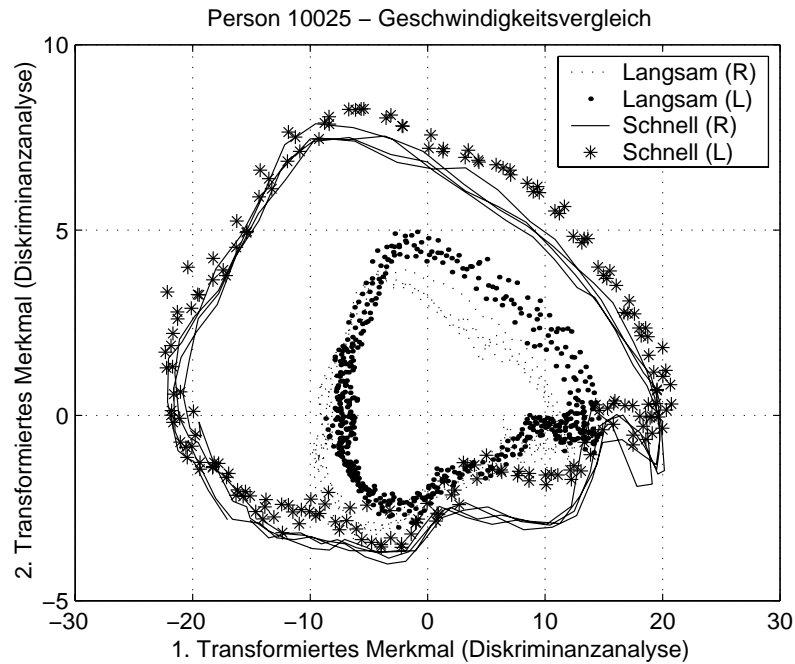


Bild 4.5: Vergleich unterschiedlicher Geschwindigkeiten

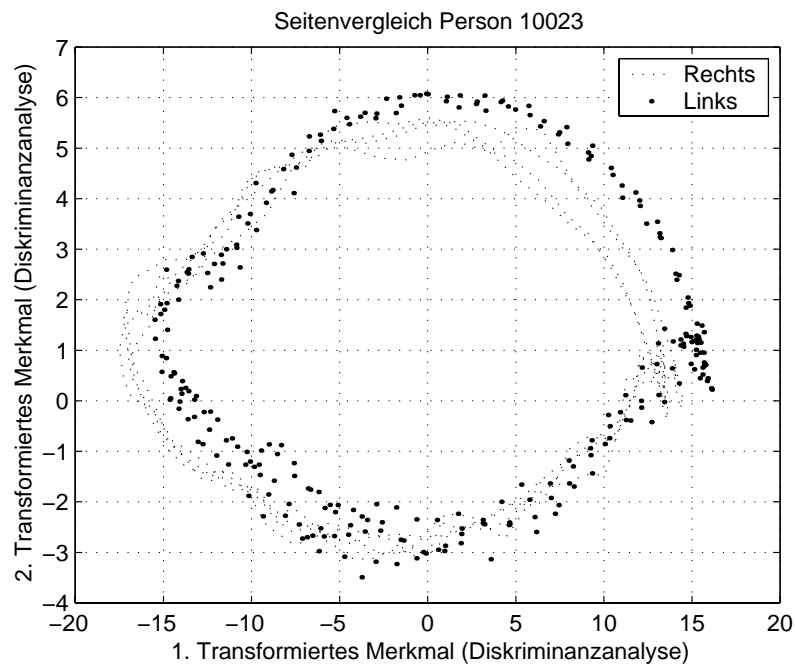


Bild 4.6: Links-Rechts-Seitenvergleich bei der Schrittphasenerkennung für Person 10023

seite einer geringeren Geschwindigkeit im Vergleich zur linken Körperseite, weil sie dort einen kleineren Kreisdurchmesser aufweisen. Bei Person 10026 weichen beide Verläufe hauptsächlich in der Standphase voneinander ab.

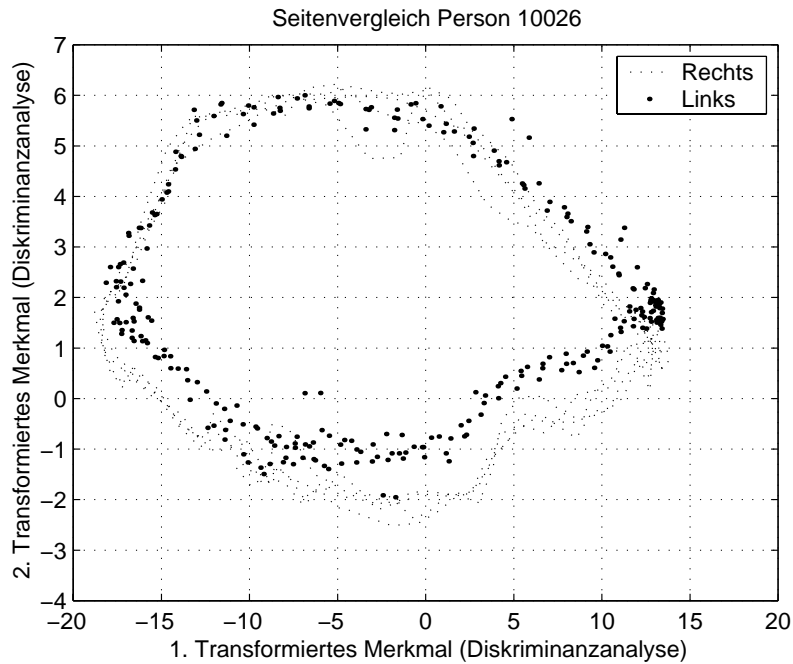


Bild 4.7: Links-Rechts-Seitenvergleich bei der Schrittphasenerkennung für Person 10026

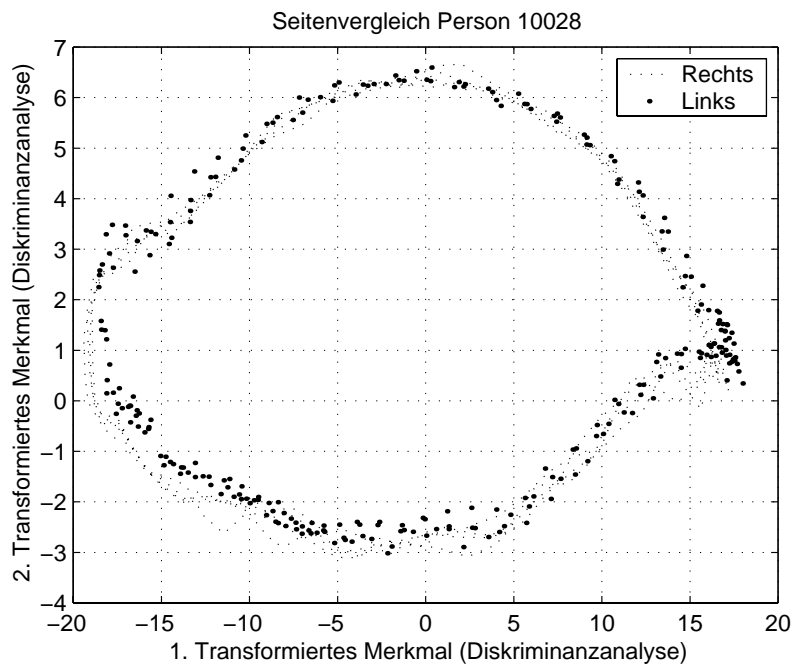


Bild 4.8: Links-Rechts-Seitenvergleich bei der Schrittphasenerkennung für Person 10028

Zudem unterscheiden sich bestimmte Kurvenbereiche in ihrer Form, was besonders bei der charakteristischen Spitze beim Aufsetzen des Fußes auffällt (Initial Contact als Beginn der Schrittphase Loading Response, rechts Mitte). Diese Spitze ist bei Person 10026

weniger ausgeprägt als bei den beiden anderen Personen.

Vergleiche dieser Art eröffnen Potenziale bei der Diagnose, weil hier mit Hilfe der transformierten Merkmale ein zusammengefasstes Gangbild einer Person mit unbekannter Diagnose gegenüber verschiedenen Personengruppen mit bekannten Diagnosen verglichen werden kann. Die daraufhin zu treffende Entscheidung ist zur Zeit manuell durchzuführen, allerdings lässt sich auch dieser Schritt prinzipiell automatisieren. Die Implementierung dieser automatischen Entscheidung ist nur dann sinnvoll, wenn eine manuelle Diagnose auf der Basis der transformierten Merkmale zu guten Ergebnissen führt.

Ein weiterer interessanter Schritt ist die Verlagerung der Transformationsmatrizen auf andere Körperbereiche. Während im Moment der Schwerpunkt auf dem Bereich Knie-Hüfte liegt, sind auch Kombinationen wie Fußgelenk-Knie oder Hüfte-Becken zu untersuchen.

Kapitel 5

Fuzzy-Verfahren

5.1 Übersicht

5.1.1 Struktur

Die Fuzzy-Logik ist ein Teilgebiet der Computational Intelligence, die außerdem noch die Teilgebiete Künstliche Neuronale Netze und Evolutionäre Algorithmen umfasst. Mit dem Einsatz von Fuzzy-Logik [115] werden menschliche Problemlösungsstrategien soweit wie möglich zur Beurteilung von Situationen herangezogen.

Das Grundprinzip von Fuzzy-Systemen besteht in der Nutzung natürlichsprachlicher (linguistischer) Regeln, die qualitative Zusammenhänge zwischen Eingangsgrößen (Merkmale) und Ausgangsgrößen beschreiben. Während klassische regelbasierte Systeme nur Ja-Nein-Aussagen über die Merkmale und Ausgangsgrößen (Voraussetzung erfüllt oder nicht erfüllt) verarbeiten, können Fuzzy-Systeme auch unscharfe Aussagen auswerten (Voraussetzung zu einem bestimmten Grad erfüllt). Die Verbindung zwischen quantitativen Werten und der qualitativen Beschreibung mittels linguistischer Terme wird durch Zugehörigkeitsfunktionen hergestellt.

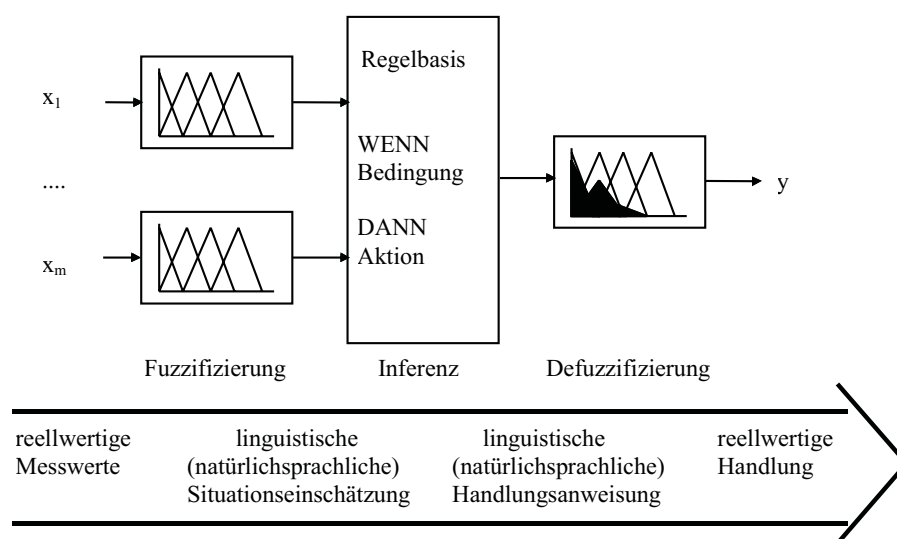


Bild 5.1: Prinzip der Fuzzy-Wissensverarbeitung

Beim Einsatz von Fuzzy-Systemen werden Messergebnisse mittels Zugehörigkeitsfunktionen von linguistischen Termen fuzzifiziert und damit in eine natürlichsprachliche Situationseinschätzung überführt. Durch eine Regelbasis und eine Inferenzmaschine werden aus diesen unscharfen Informationen ebenso natürlichsprachliche Schlussfolgerungen gezogen, aus denen bei Bedarf wieder eine konkrete Handlungsanweisung gewonnen wird (Defuzzifizierung). Die Struktur ist in Bild 5.1 schematisch dargestellt.

Vorteilhaft ist insbesondere die gute Interpretierbarkeit von Fuzzy-Regeln und Zugehörigkeitsfunktionen. Die einzelnen Schritte werden nun detaillierter erläutert.

5.1.2 Fuzzifizierung

Für jedes Merkmal (Eingangsgröße des Fuzzy-Systems) x_l werden so genannte linguistische Terme $A_{l,i}$, $i = 1, \dots, m_l$, eingeführt, die für natürlichsprachliche Situationseinschätzungen wie *groß*, *klein* usw. stehen. Zu jedem dieser linguistischen Terme gehört eine Zugehörigkeitsfunktion, die Werte zwischen Null und Eins annimmt.

Die Aufgabe bei der Fuzzifizierung besteht darin, jedem Messwert eine graduelle Zuordnung zu diesen linguistischen Termen zuzuweisen. So kann beispielsweise ein Messwert *Winkel ist 2.30°* den Aussagen *Winkel ist Null (Z)* mit einem Zugehörigkeitsgrad von 0.77 und *Winkel ist Positiv Klein (PK)* mit einem Zugehörigkeitsgrad von 0.23 zugeordnet werden. Dabei werden dreieckförmige Zugehörigkeitsfunktionen für *Null* mit einem Funktionswert von Eins bei $x_l = 0$ und für *Positiv Klein* mit einem Funktionswert von Eins bei $x_l = 10$ angenommen (Bild 5.2). Alle anderen linguistischen Terme erhalten einen Zugehörigkeitsgrad von Null.

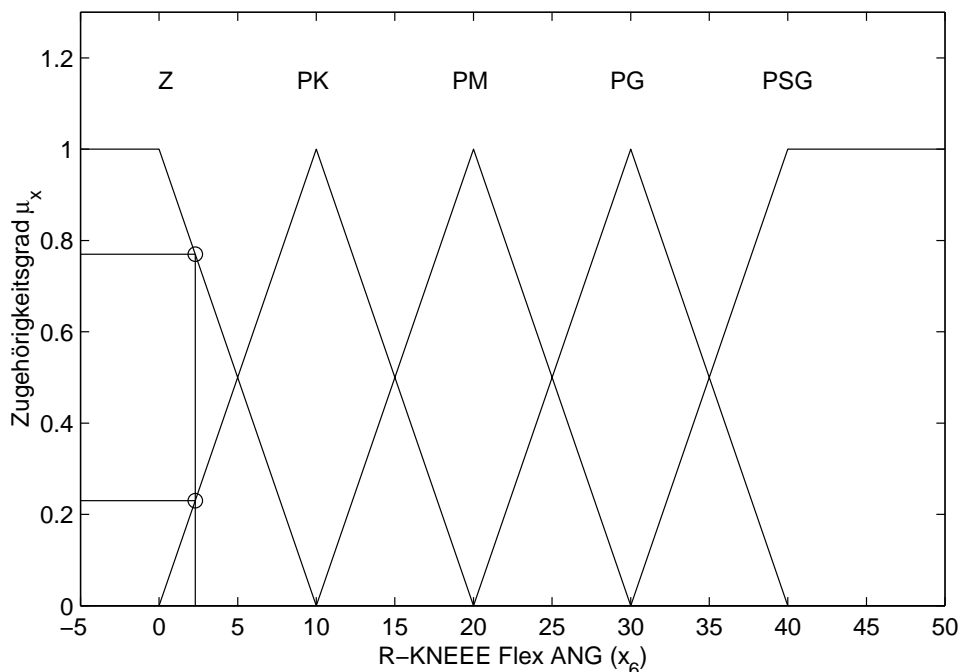


Bild 5.2: Prinzip der Fuzzifizierung: Merkmal x_6 am Beispiel eines Messwerts $x_6 = 2.30$ mit $m_6 = 5$ linguistischen Termen (Null (Z), Positiv Klein (PK), Positiv Mittel (PM), Positiv Groß (PG), Positiv Sehr Groß (PG)) und ihren Zugehörigkeitsfunktionen

Werden nun alle Messwerte (N Datensätze mit s Merkmalen) fuzzifiziert (den m_i linguistischen Termen zugeordnet), folgt

$$\boldsymbol{\mu}_X = \begin{pmatrix} \boldsymbol{\mu}_{x_1} \\ \boldsymbol{\mu}_{x_2} \\ \vdots \\ \boldsymbol{\mu}_{x_s} \end{pmatrix} \quad (5.1)$$

$$= \begin{pmatrix} \boldsymbol{\mu}_{A_1}(x_{11}) & \boldsymbol{\mu}_{A_1}(x_{21}) & \cdots & \boldsymbol{\mu}_{A_1}(x_{N1}) \\ \boldsymbol{\mu}_{A_2}(x_{12}) & \boldsymbol{\mu}_{A_2}(x_{22}) & \cdots & \boldsymbol{\mu}_{A_2}(x_{N2}) \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\mu}_{A_s}(x_{1s}) & \boldsymbol{\mu}_{A_s}(x_{2s}) & \cdots & \boldsymbol{\mu}_{A_s}(x_{Ns}) \end{pmatrix} \quad (5.2)$$

$$= \begin{pmatrix} \mu_{A_{11}}(x_{11}) & \mu_{A_{11}}(x_{21}) & \cdots & \mu_{A_{11}}(x_{N1}) \\ \mu_{A_{12}}(x_{11}) & \mu_{A_{12}}(x_{21}) & \cdots & \mu_{A_{12}}(x_{N1}) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{A_{1,m_1}}(x_{11}) & \mu_{A_{1,m_1}}(x_{21}) & \cdots & \mu_{A_{1,m_1}}(x_{N1}) \\ \mu_{A_{21}}(x_{12}) & \mu_{A_{21}}(x_{22}) & \cdots & \mu_{A_{21}}(x_{N2}) \\ \mu_{A_{22}}(x_{12}) & \mu_{A_{22}}(x_{22}) & \cdots & \mu_{A_{22}}(x_{N2}) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{A_{2,m_2}}(x_{12}) & \mu_{A_{2,m_2}}(x_{22}) & \cdots & \mu_{A_{2,m_2}}(x_{N2}) \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{A_{s1}}(x_{1s}) & \mu_{A_{s1}}(x_{2s}) & \cdots & \mu_{A_{s1}}(x_{Ns}) \\ \mu_{A_{s2}}(x_{1s}) & \mu_{A_{s2}}(x_{2s}) & \cdots & \mu_{A_{s2}}(x_{Ns}) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{A_{s,m_s}}(x_{1s}) & \mu_{A_{s,m_s}}(x_{2s}) & \cdots & \mu_{A_{s,m_s}}(x_{Ns}) \end{pmatrix}. \quad (5.3)$$

Die Matrix $\boldsymbol{\mu}_X$ hat die Dimension $(m = \sum_{l=1}^s m_l, N)$. Sie enthält in der Darstellung von (5.1) s Blockmatrizen $\boldsymbol{\mu}_{x_l}$ mit der Dimension (m_l, N) , in denen alle fuzzifizierten Werte des Merkmals x_l für alle N Beispiele des Datensatzes stehen. Diese Blockmatrizen setzen sich in (5.2) wiederum aus N Vektoren $\boldsymbol{\mu}_{A_l}(x_{li})$ der Dimension $(m_l, 1)$ zusammen, die die Zugehörigkeitsgrade aller linguistischen Terme von x_l für das i -te Beispiel beinhalten. Deren skalare Elemente $\mu_{A_{lk}}(x_{li})$ in (5.3) stehen für die Zugehörigkeit des k -ten Terms des l -ten Merkmals für das i -te Beispiel.

5.1.3 Inferenz

Die Regelbasis ist durch Regeln vom Typ WENN Bedingung (Prämisse) DANN Aktion (Konklusion) mit

$$\begin{array}{ll} R_1 : & \text{WENN } (x_1 = A_{1,R1}) \cap \cdots \cap (x_s = A_{s,R1}) & \text{DANN } y = B_{R1} \\ \cdots & & \\ R_r : & \text{WENN } (x_1 = A_{1,Rr}) \cap \cdots \cap (x_s = A_{s,Rr}) & \text{DANN } y = B_{Rr} \end{array}$$

gegeben. Die Prämisse der k -ten Regel lautet somit

$$P_k = (x_1 = A_{1,Rk}) \cap \cdots \cap (x_s = A_{s,Rk}).$$

Für die Ausgangsgröße y in der Konklusion werden linguistische Terme B_j , $j = 1, \dots, m_y$, eingeführt, die Ausgangsklassen (z. B. Schrittphase Loading Response) oder Handlungsanweisungen (z. B. Ausgangsgröße klein) kennzeichnen. Zusätzlich kann für jede Regel eine

Regelplausibilität zwischen Null und Eins angegeben werden, die die Glaubwürdigkeit einer Regel zwischen vollständig gültig (Eins) und absolut falsch (Null) bewertet.

Die Prämisse P_k ist eine UND-Verknüpfung (\cap) von Teilprämissen $x_l = A_{l,Rk}$. $A_{l,Rk}$ bezeichnet dabei eine ODER-Verknüpfung (\cup) von linguistischen Termen des l -ten Merkmals, die für die k -te Regel gilt.¹

$A_{l,Rk}$ kann entweder

- nur aus einem linguistischen Term $A_{l,k}$ bestehen,
- mehrere linguistische Terme $A_{l,k} \cup A_{l,p} \cdots$ oder
- alle linguistische Terme des Merkmals $A_{l,1} \cup A_{l,2} \cdots \cup A_{l,m_l}$

umfassen. Im letztgenannten Fall wird bei geeigneter Wahl der Zugehörigkeitsfunktionen die Prämisse nicht mehr durch die Teilprämisse beeinflusst, weil sie eine Zugehörigkeit von Eins und somit eine UND-Verknüpfung mit Eins in der Regelprämisse liefert. Damit kann die Teilprämisse aus der Prämisse der Regel entfernt werden.

Wesentliche Rechengrundlage der Inferenz sind geeignete Verknüpfungsoperatoren für die UND- (T-Norm [58]) bzw. ODER-Verknüpfung (T-Konorm [58]). Die gebräuchlichsten Operatoren sind dabei die zusammengehörigen Operatorenpaare

- UND: Minimum

$$\mu_3 = \mu_1 \cap \mu_2 = \min(\mu_1, \mu_2), \quad (5.4)$$

ODER: Maximum

$$\mu_3 = \mu_1 \cup \mu_2 = \max(\mu_1, \mu_2), \quad (5.5)$$

- UND: Produkt

$$\mu_3 = \mu_1 \cap \mu_2 = \mu_1 \cdot \mu_2, \quad (5.6)$$

ODER: algebraische Summe

$$\mu_3 = \mu_1 \cup \mu_2 = \mu_1 + \mu_2 - \mu_1 \cdot \mu_2, \quad (5.7)$$

- UND: beschränkte Differenz

$$\mu_3 = \mu_1 \cap \mu_2 = \max(\mu_1 + \mu_2 - 1, 0), \quad (5.8)$$

ODER: beschränkte Summe

$$\mu_3 = \mu_1 \cup \mu_2 = \min(\mu_1 + \mu_2, 1), \quad (5.9)$$

wobei sich der Zugehörigkeitsgrad eines Verknüpfungsergebnis μ_3 aus zwei UND- bzw. ODER-verknüpften linguistischen Termen mit den Zugehörigkeitsgraden μ_1 sowie μ_2 ergibt. Für die ODER-Verknüpfung wird teilweise auch die (gewöhnliche) Summe $\mu_3 = \mu_1 + \mu_2$ eingesetzt, die aber keine t -Konorm ist.

Für die Negation wird

$$\mu_2 = \overline{\mu_1} = 1 - \mu_1 \quad (5.10)$$

verwendet.

¹Die Operatoren \cap (UND, T-Norm, Schnitt von Mengen) und \cup (ODER, T-Konorm, Vereinigung von Mengen) werden im Folgenden sowohl für die Verknüpfung von logischen Aussagen, linguistischen Termen, Mengen, Fuzzy-Mengen und Zugehörigkeitsgraden verwendet, wobei jeweils unterschiedliche Berechnungsvorschriften zugrunde liegen.

Bei der Inferenz sind die Teilaufgaben

- Prämissenauswertung,
- Aktivierung und
- Akkumulation

zu bearbeiten. In den folgenden Ausführungen wird zunächst nur der Fall vollständiger (Für jede Kombination von Eingangsgrößen wird mindestens eine Regel aktiviert) und widerspruchsfreier (Vollständig aktivierte Regeln haben stets gleiche Konklusionen) Regelbasen diskutiert. Nähere Ausführungen zur Behandlung widersprüchlicher oder redundanter Regeln finden sich in Abschnitt 5.7.

Die *Prämissenauswertung* umfasst die Berechnung der Zugehörigkeitsgrade der Prämissen der r Regeln. Beim üblichen Rechenschema werden zunächst eventuelle ODER-Verknüpfungen zur Berechnung von $A_{l,Rk}$ bestimmt und dann die UND-Verknüpfungen über die linguistischen Variablen durchgeführt. Im Ergebnis entsteht für jede Regel und jeden Messwert ein Zugehörigkeitsgrad der Prämisse $\mu_{P_{k,i}}$, also eine Matrix der Dimension (r, N) . Im Folgenden wird für die ODER-Verknüpfung verschiedener linguistischer Terme die beschränkte Summe und für die UND-Verknüpfung der so entstehenden Teilprämissen das Produkt eingesetzt. Diese Kombination ist konsistent im Sinne der semantischen Restriktionen aus [45] und des Inferenz-Verfahrens in [77].

Die *Aktivierung* einer Regel erfolgt mit einer UND-Verknüpfung aus Regelplausibilität und Zugehörigkeitsgrad ihrer Prämisse. Wenn die Regelplausibilität Eins ist, ist der Zugehörigkeitsgrad der Konklusion identisch mit dem der Prämisse. Im Ergebnis entsteht für jede Regel und jeden Messwert ein Zugehörigkeitsgrad der Konklusion $\mu_{K_{k,i}}$, also eine Matrix der Dimension (r, N) . Für diese UND-Verknüpfung wird ebenfalls das Produkt verwendet.

Die *Akkumulation* wird in die Akkumulation über Regeln mit gleicher Konklusion (Akkumulation I, vertikale Akkumulation), die Konstruktion der entsprechenden Fuzzy-Mengen der linguistischen Terme der Ausgangsgröße (Akkumulation II) und deren Akkumulation (horizontale Akkumulation, Akkumulation III) zerlegt.

Die Akkumulation I ist eine ODER-Verknüpfung zwischen Regelaktivierungen von Regeln gleicher Konklusion. Als Operator wird im Folgenden der Summen-Operator verwendet, wobei durch eine spezielle Redundanzmatrix sich überlappende Prämissen berücksichtigt werden (siehe auch Abschnitt 5.7). Im Ergebnis entsteht für jeden linguistischen Term der Ausgangsgröße und jeden Messwert ein Zugehörigkeitsgrad der Prämisse $\mu_{B_j}(y_i)$, also eine Matrix der Dimension (m_y, N) .

Die Akkumulation II berechnet eine UND-Verknüpfung aus dem Ergebnis der Akkumulation I sowie den jeweiligen Zugehörigkeitsfunktionen der linguistischen Terme der Ausgangsgröße. Die Verknüpfung zur Konstruktion der Fuzzy-Mengen erfolgt durch punktweises Anwenden des UND-Operators. Hier wird der Produkt-Operator angewendet. Somit entsteht für jeden linguistischen Term der Ausgangsgröße und jeden Messwert eine Zugehörigkeitsfunktion zu einer Fuzzy-Menge $\mu_{B_{j,i}}(y)$, also eine Matrix der Dimension (m_y, N) . Deren Elemente sind Fuzzy-Mengen (also Funktionen), was die rechentechnische Behandlung deutlich erschwert. Bei der Verwendung der Schwerpunktmethode für Singletons² oder der Maximum-Defuzzifizierung ergeben sich allerdings wesentliche Vereinfachungen, weil die Akkumulation II nicht explizit ausgeführt werden muss (siehe Ab-

²Singletons sind Zugehörigkeitsfunktionen, die nur an einer Stelle (Stützpunkt) einen Zugehörigkeitsgrad von Eins und sonst einen Zugehörigkeitsgrad von Null aufweisen.

schnitt 5.1.4). Zudem ergibt bei der Konstruktion der Fuzzy-Mengen für Singletons jeder beliebige UND-Operator das gleiche Ergebnis.

Die Akkumulation III nimmt eine ODER-Verknüpfung der Ergebnisse der Akkumulation II vor. Für die Akkumulation der Fuzzy-Mengen wird punktweise ein ODER-Operator (meist der beschränkte Summen-Operator) verwendet. Im Ergebnis entsteht für jeden Messwert eine Fuzzy-Menge der Ausgangsgröße $\mu_{y_i}(y)$, also ein Vektor der Dimension $(1, N)$. Dessen Elemente sind aber wiederum Funktionen. Auch hier vereinfacht sich die Berechnung bei Verwendung der Schwerpunktmethod für Singletons oder der Maximum-Defuzzifizierung, weil die Akkumulation III nicht explizit ausgeführt werden muss (siehe Abschnitt 5.1.4).

Nach den verwendeten Operatoren für die UND- bzw. ODER-Verknüpfung werden die einzelnen Inferenzverfahren häufig unter den Begriffen MIN-MAX-Inferenz oder SUM-PROD-Inferenz zusammengefasst, wobei in den einzelnen Verarbeitungsschritten teilweise auch abweichende Operatoren verwendet werden.

Abweichend von der hier demonstrierten Vorgehensweise, möglichst lange (also in der Aktivierung und Akkumulation I) mit aussagenlogischen Ausdrücken (Zugehörigkeitsgraden) und nicht mit Fuzzy-Mengen (Funktionen) zu rechnen, wird häufig bereits bei der Aktivierung eine Fuzzy-Menge der Regelkonklusion ermittelt (Zusammenfassen von Aktivierung und Akkumulation II). Damit entsteht nach der Aktivierung eine Matrix von Fuzzy-Mengen der Dimension (r, N) . Die dann folgende Akkumulation entspricht dann nur noch der Akkumulation III. Weil bereits bei der Aktivierung mit Funktionen gerechnet werden muss, steigt der Implementierungsaufwand an, wenn es mehr Regeln als linguistische Terme der Ausgangsgröße gibt.

5.1.4 Defuzzifizierung

Die sich ergebende Fuzzy-Menge nach der Akkumulation wird bei Bedarf durch die Defuzzifizierung auf eine reelle Zahl abgebildet [58]. Verlangt das Problem eine kontinuierliche Ausgangsgröße, wird die Schwerpunktmethod (COG)

$$y_i = \frac{\int y \cdot \mu_{y_i}(y) \cdot dy}{\int \mu_{y_i}(y) \cdot dy} \quad (5.11)$$

bzw. die Schwerpunktmethod für Singletons (COGS) gewählt, für Klassifikationsaufgaben ist die Maximum-Defuzzifizierung (MOM) adäquat.

Die Schwerpunktmethod für Singletons (COGS) kann sowohl bei Singletons als Ausgangszugehörigkeitsfunktionen als auch bei dreieckförmigen Zugehörigkeitsfunktionen angewendet werden. Dabei werden alle Parameter (Stützpunkte der Singletons bzw. Maxima der dreieckförmigen Zugehörigkeitsfunktionen) der Ausgangszugehörigkeitsfunktionen b_j in den Vektor \mathbf{b} mittels

$$\mathbf{b} = (b_1 \ \cdots \ b_{m_y})^T \quad (5.12)$$

eingetragen. Zusammen mit dem Ergebnis der Akkumulation I kann dann der Ausgangswert berechnet werden:

$$y_i = \frac{\sum_{j=1}^{m_y} b_j \cdot \mu_{B_j}(y_i)}{\sum_{j=1}^{m_y} \mu_{B_j}(y_i)} \quad (5.13)$$

$$= \frac{\mathbf{1}_{1, m_y} \cdot (\mathbf{b} \odot \boldsymbol{\mu}_B(y_i))}{\mathbf{1}_{1, m_y} \cdot \boldsymbol{\mu}_B(y_i)}. \quad (5.14)$$

Der Operator \odot bezeichnet dabei eine elementweise Multiplikation. Akkumulation II und III sind in dieser Operation implizit enthalten.

Bei der Maximum-Defuzzifizierung wird für den i -ten Wert der Ausgangsgröße y_i der Wert b_j ausgewählt, für den der maximale Wert in $\mu_B(y_i)$ in der Akkumulation I auftritt. Gibt es mehrere Maxima, wird entweder das b_j mit dem kleineren j (Methode Linkes Maximum – LM), das mit dem größeren j (Methode Rechtes Maximum – RM) oder deren Mittelwert (Methode Mittleres Maximum – MOM) ausgewählt. Auch hier müssen Akkumulation II und III nicht explizit ausgeführt werden.

Ist für eine Ausgangsgröße y_i der Vektor der Zugehörigkeitsgrade $\mu_B(y_i)$ der Nullvektor, muss für die betreffende Größe ein Defaultwert für y_i ausgegeben werden. Anderenfalls treten bei der COGS-Methode numerische Probleme auf (Division durch Null).

5.2 Entwicklungsstand beim Entwurf von Fuzzy-Systemen

In Ergänzung zum klassischen Weg der direkten Umsetzung bekannter Expertenregeln gibt es in den letzten Jahren vielfältige Bemühungen, Regeln aus Daten automatisch zu generieren. Ausgangspunkt für datenbasierte Regelsuchverfahren bildet eine Menge von N Beispielen (Lerndatensatz) für $l = 1, \dots, s$ Merkmale x_l und eine Ausgangsgröße y , die in einer Matrix \mathbf{X} (4.1) und einem Vektor \mathbf{y} (4.2) angeordnet sind. Die Merkmale der zu klassifizierenden Objekte können kategorische, ordinale oder numerische Variable sein.³

Die Ausgangsgröße ist im Allgemeinen kategorisch und gibt die Zuordnung zu einer Klasse an. Liegt hingegen eine unscharfe Zuordnung zu verschiedenen Klassen vor, wird die Ausgangsgröße in Form einer Matrix

$$\mu_Y = \begin{pmatrix} \mu_{B_1}(y_1) & \mu_{B_1}(y_2) & \cdots & \mu_{B_1}(y_N) \\ \mu_{B_2}(y_1) & \mu_{B_2}(y_2) & \cdots & \mu_{B_2}(y_N) \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{B_{m_y}}(y_1) & \mu_{B_{m_y}}(y_2) & \cdots & \mu_{B_{m_y}}(y_N) \end{pmatrix} \quad (5.15)$$

von Zugehörigkeitsgraden angegeben. Dabei muss sie u. U. ebenso wie die Eingangsgröße fuzzifiziert werden.

Die Aufgabe besteht darin, eine regelbasierte Beschreibung der Klassen im Merkmalsraum zu finden. Anforderungen an die Regeln sind dabei:

- eine hohe Klassifikationsgüte,
- die Verständlichkeit der Einzelregeln,
- die Transparenz ihres Zusammenwirkens,
- eine vollständige Abdeckung des Eingangsraums (für jede Situation gibt es mindestens eine aktivierte Regel),
- ein niedriger Aufwand zur Speicherung und Auswertung der Regeln und
- ein niedriger Messaufwand für die in den Regelprämissen auftretenden Merkmale.

³Kategorische und ordinale Merkmale sind wertediskret, wobei nominale Merkmale keiner Ordnungsrelation (größer als, kleiner als) unterliegen. Ein Beispiel für nominale Merkmale sind Farben wie rot, grün und braun; ein Beispiel für ordinale Merkmale Größen wie sehr klein, klein, mittel und groß.

Die Regelsuche umfasst das Aufstellen und Bewerten von Hypothesen. Beim *direkten Zugang* sind die Hypothesen komplette Regelbasen (Pittsburgh-Ansatz [102]). Im Unterschied dazu werden beim *indirekten Zugang* (Michigan-Ansatz [39]) Einzelregeln als Hypothesen aufgestellt und erst später zu Regelbasen zusammengefasst. Übersichten über unterschiedliche Verfahren finden sich in [17, 81].

Da bei der Hypothesengenerierung aufgrund der Komplexität ein vollständiges Durchsuchen des Regelraums praktisch nicht durchführbar ist, müssen Methoden genutzt werden, die auch bei unvollständiger Suche annähernd alle relevanten Regeln bzw. möglichst gute Regelbasen finden. Dabei werden heuristische und induktive bzw. probabilistische Verfahren unterschieden.

Tabelle 5.1 zeigt eine Einteilung ausgewählter Verfahren zur Regelgenerierung für Fuzzy-Systeme entsprechend der beiden Zugänge sowie der Suchstrategien.

	direkter Zugang	indirekter Zugang
heuristisch/ induktiv	<ul style="list-style-type: none"> • ASMOD [65] • Fuzzy CART [50] • LOLIMOT [83] • Fuzzy-Entscheidungsbäume [12, 51, 94, 95] 	<ul style="list-style-type: none"> • Fuzzy-ROSA (explorative Standardstrategie) [28] • baumbasierte Regelgenerierung [24, 28, 35, 47, 57] • Fuzzy Version Space Learning [42, 107] • induktives Lernen modularer Fuzzy-Regeln [60, 108] • Induktion hierarchischer Fuzzy-Systeme [40]
probabilistisch	GA (Genetische Algorithmen)-basierter Entwurf von <ul style="list-style-type: none"> • Fuzzy-Systemen [38, 44, 55] • Neuro-Fuzzy-Systemen mit SMOG [9] 	<ul style="list-style-type: none"> • evolutionäre Regelgenerierung [10] • Fuzzy-ROSA (evolutionäre Suchstrategie) [62, 63] • GA-basierte Regelgenerierung [36]

Tabelle 5.1: Einteilung von Verfahren zur Regelgenerierung für Fuzzy-Systeme.

Beim direkten Zugang, bei dem im Allgemeinen globale Gütemaße angewendet werden, besteht das Problem, die Qualität einzelner Regeln nicht bewerten zu können. Das kann dazu führen, dass sich fehlerhafte Einzelregeln in ihren Auswirkungen kompensieren und so trotzdem ein gutes Gesamtverhalten ergeben, was die Interpretierbarkeit der Regeln reduziert. Außerdem ist der Suchraum wegen der großen Anzahl möglicher Hypothesen sehr groß.

Der Suchraum beim indirekten Zugang ist deutlich kleiner. Da sich die Bewertung bei der Regelsuche zunächst auf die einzelnen Regeln bezieht, besteht die Notwendigkeit, ihr Zusammenwirken in der Regelbasis ähnlich wie beim direkten Zugang zu bewerten. Zudem

ist meist nur ein Teil des Eingangsraums durch die Regeln abgedeckt.

Im Suchraum können zwei prinzipielle Suchrichtungen für neue Regelhypothesen unterschieden werden, und zwar Spezialisierungen und Verallgemeinerungen (Generalisierungen). Spezialisierung bedeutet, dass die Suche vom Allgemeinen zum Speziellen erfolgt, bei Verallgemeinerungen wird die umgekehrte Suchrichtung verfolgt.

Ausgehend von einer Menge von Anfangshypothesen werden sukzessive neue Hypothesen generiert, bewertet und entsprechend akzeptiert oder verworfen. Anfangshypothesen sind z. B. eine initiale Regelbasis (direkter Zugang), die allgemeinste Regel mit „leerer“ Prämisse oder die speziellsten mittels der Beispiele aufstellbaren Regeln (indirekter Zugang). Für das Aufstellen neuer Hypothesen kommen verschiedene Strategien in Frage.

Die bekannten Algorithmen müssen allerdings in erheblichem Maß manuell an die jeweilige Problemstellung angepasst werden (gewünschtes Abstraktionsniveau, Kompromiss zwischen Klarheit der Regeln und Interpretierbarkeit einerseits gegenüber einer hohen Klassifikationsgüte andererseits, Redundanz, statistische Absicherung).

Bei den meisten Verfahren erfordert der Entwurf aber noch die manuelle Festlegung der Zugehörigkeitsfunktionen und die manuelle Auswahl von Merkmalen. In vielen praktischen Anwendungen stellen aber gerade diese Entwurfsschritte den Anwender vor erhebliche Schwierigkeiten und begrenzen die Güte der gefundenen Regeln.

5.3 Verwendetes Entwurfsverfahren für Fuzzy-Systeme

Nachfolgend wird ein indirektes, baumorientiertes Verfahren zum Entwurf von Fuzzy-Systemen näher betrachtet, das am Forschungszentrum Karlsruhe entwickelt wurde (siehe Bild 5.3).

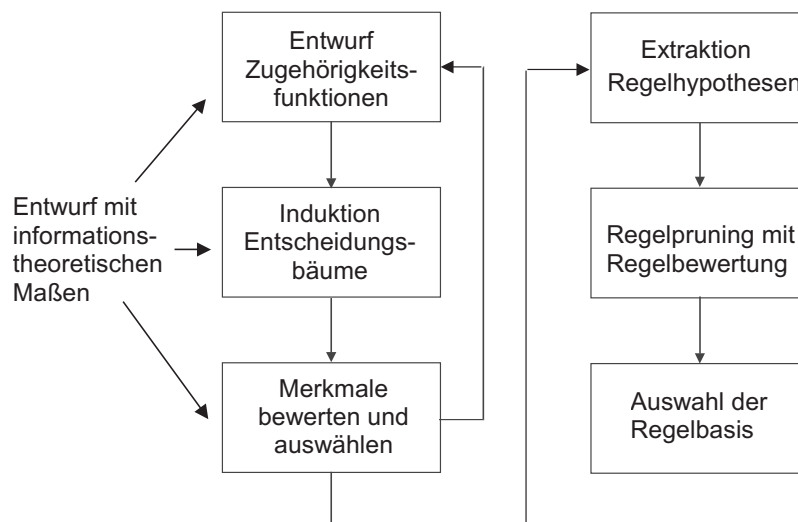


Bild 5.3: Überblick über das am Forschungszentrum Karlsruhe entwickelte Entwurfsverfahren für Fuzzy-Systeme

Es integriert alle Entwurfsschritte, die für einen automatischen Entwurf von Fuzzy-Systemen erforderlich sind. In den nächsten Abschnitten werden der

- Entwurf von Zugehörigkeitsfunktionen (Abschnitt 5.4),

- die Merkmalsselektion (Abschnitt 5.5),
- die Regelgenerierung (Abschnitt 5.6) und
- die Suche nach einer Regelbasis (Abschnitt 5.7)

beschrieben. Teilweise ausführlichere methodische Darstellungen finden sich in [45–49, 77, 79–81]).

5.4 Entwurf von Zugehörigkeitsfunktionen

5.4.1 Ziele

Zugehörigkeitsfunktionen übernehmen die graduelle Zuordnung von Messwerten zu linguistischen Termen $A_{l,i}$. Die Funktionswerte der Zugehörigkeitsgrade liegen meist zwischen Null (völlig falsche Aussage) und Eins (absolut richtige Aussage). Als Zugehörigkeitsfunktionen werden häufig dreieck- oder trapezförmige Funktionen angenommen, die über Parameter (Stützpunkte) an den Knickstellen beschrieben werden. Weitere sinnvolle Annahmen sind, dass sich alle Zugehörigkeitsfunktionen zu Eins ergänzen und immer nur höchstens zwei benachbarte Zugehörigkeitsfunktionen Funktionswerte größer Null aufweisen können. Eine theoretische Begründung für diese Festlegungen, die von semantischen Restriktionen ausgeht, wird in [45] gegeben.

Im Folgenden werden deshalb stets dreieckförmige Zugehörigkeitsfunktionen in der Mitte und trapezförmige Zugehörigkeitsfunktionen am Rand eingesetzt. Wenn die Zugehörigkeitsfunktionen jeweils Zugehörigkeitsgrade zwischen Null und Eins aufweisen, sich zu Eins ergänzen und nur einfach überlappen, reicht pro linguistischem Term ein skalarer Parameter $a_{l,i}$ zur Beschreibung der jeweiligen Zugehörigkeitsfunktion aus (Punkt mit Funktionswert Eins der Zugehörigkeitsfunktion). Im Beispiel aus Bild 5.4 (links) sind das folglich die Parameter $-10, -3, 0, 1$ und 10 , die in einem Vektor angeordnet werden:

$$\mathbf{a}_l = (-10 \quad -3 \quad 0 \quad 1 \quad 10)^T. \quad (5.16)$$

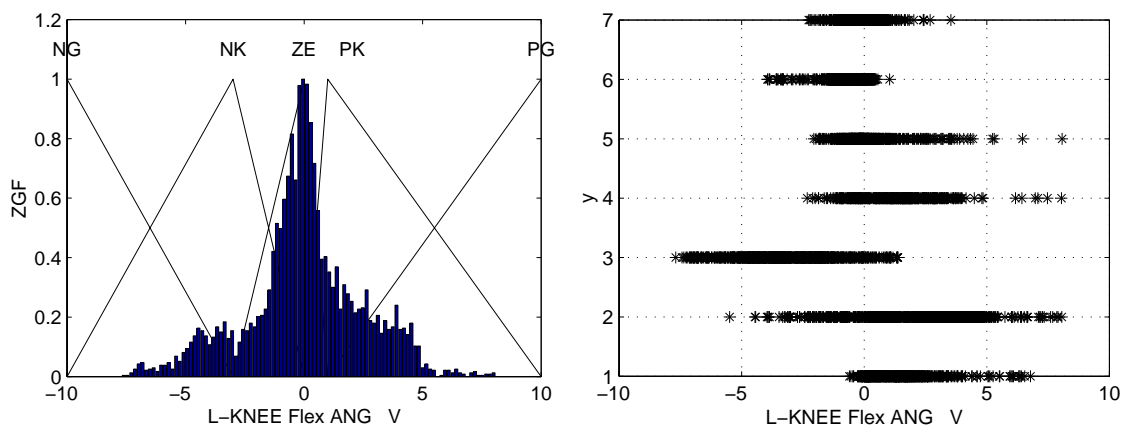


Bild 5.4: Automatisch entworfene Zugehörigkeitsfunktionen für das Merkmal x_{105} mit Histogramm (links) und Werte des Merkmals x_{105} für die sieben Schrittphasen (rechts), Abkürzungen der Namen der linguistischen Terme $A_{105,i}$: NG (Negativ Groß), NK (Negativ Klein), ZE (Null), Positiv Klein (PK) und Positiv Groß (PG)

Beim hier vorgestellten Verfahren werden Hypothesen für die Lage der Parameter von Zugehörigkeitsfunktionen (Stützpunkte) aufgestellt und bewertet. Die auszuwählenden Stützpunkte sollen

- eine gute Separierung der Ausgangsklassen ermöglichen,
- eine gleichmäßige Beispiellanzahl für alle Terme bewirken,
- nicht zu enge Abstände der Stützpunkte aufweisen,
- seltene Ausgangsklassen nicht unterdrücken und
- gut interpretierbare Werte annehmen (z. B. 0 und nicht 0.02 für den Term *Null*).

Die Problematik des zu findenden Kompromisses verdeutlicht Bild 5.4 anhand der automatischen Festlegung der Änderungsgeschwindigkeit der linken Kniebeugung (Merkmal x_{105}). Zu berücksichtigen ist dabei insbesondere, dass

- im Bereich um Null eine große Datendichte auftritt,
- die Platzierung der Parameter von ZE bei 0 und von PK bei 1 dazu führt, die Schrittphase $y = B_6$ zu separieren,
- die Platzierung der Parameter von ZE bei 0 und von NK bei -3 dazu führt, die Schrittphasen $y = B_4, y = B_5$ und $y = B_7$ gut von $y = B_6, y = B_2$ und $y = B_3$ zu separieren usw.

Zur Beurteilung dieser Forderungen werden informationstheoretische Maße eingesetzt, die auf einem hohen Abstraktionsniveau die Güte von disjunkten Partitionierungen⁴ bewerten, ohne bereits nach konkreten Fuzzy-Regeln suchen zu müssen. Die Idee besteht darin, auf der Basis von relativen Häufigkeiten einzelner linguistischer Terme der Merkmale und Ausgangsgrößen Entscheidungen über den Entwurf der Zugehörigkeitsfunktionen zu treffen. Dabei wird die Abweichung vom idealen Zustand untersucht, dass aus dem Vorliegen eines linguistischen Terms eines Merkmals auf die Ausgangsklasse geschlossen werden kann. Die notwendigen Zusammenhänge und ihre Anpassungen an Fuzzy-Systeme werden im folgenden Abschnitt erklärt.

5.4.2 Informationstheoretische Maße für Fuzzy-Partitionen

Informationstheoretische Maße [110] basieren auf Wahrscheinlichkeiten (bzw. relativen Häufigkeiten als Schätzungen für Wahrscheinlichkeiten) und stellen Maße für Zusammenhänge zwischen Größen dar. In Bild 5.5 sind verschiedene Maße am Beispiel des Zusammenhangs zwischen einem Merkmal x_l (Messgröße) und der Ausgangsgröße y (Klassifikationsentscheidung) dargestellt. Die Eingangsentropie $H(x_l)$, die Ausgangsentropie $H(y)$ und die Gesamtentropie $H(x_l, y)$ ergeben sich mit

$$H(x_l) = - \sum_{i=1}^{m_l} p(x_l = A_{l,i}) \cdot \log_2 p(x_l = A_{l,i}), \quad (5.17)$$

$$H(y) = - \sum_{j=1}^{m_y} p(y = B_j) \cdot \log_2 p(y = B_j), \quad (5.18)$$

$$H(x_l, y) = - \sum_{i=1}^{m_l} \sum_{j=1}^{m_y} p(x_l = A_{l,i} \cap y = B_j) \cdot \log_2 p(x_l = A_{l,i} \cap y = B_j), \quad (5.19)$$

⁴Bei disjunkten Partitionierungen haben die Mengen keine gemeinsamen Elemente.

wobei die Wahrscheinlichkeiten in der Regel über relative Häufigkeiten geschätzt werden, indem die Anzahl des Auftretens der jeweiligen Ereignisse n durch die Anzahl aller Datensätze N geteilt wird:

$$\hat{p}(x_l = A_{l,i}) = \frac{n(x_l = A_{l,i})}{N}, \quad (5.20)$$

$$\hat{p}(y = B_j) = \frac{n(y = B_j)}{N}, \quad (5.21)$$

$$\hat{p}(x_l = A_{l,i} \cap y = B_j) = \frac{n(x_l = A_{l,i} \cap y = B_j)}{N}. \quad (5.22)$$

Den Informationsfluss zwischen Ein- und Ausgangsgröße beschreibt die Transinformation $H(x_l; y)$. Die Irrelevanz $H(y|x_l)$ und die Äquivokation $H(x_l|y)$ sind Maße für die Anteile der Aus- bzw. Eingangsgröße, die nicht aus der jeweils anderen Größe resultieren.

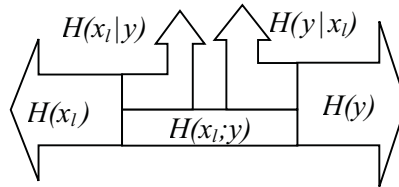


Bild 5.5: Entropiebewertung

Es gelten folgende Rechenregeln, die auf (5.17-5.19) basieren:

$$H(x_l; y) = H(x_l) + H(y) - H(x_l, y), \quad (5.23)$$

$$H(y|x_l) = H(x_l, y) - H(x_l), \quad (5.24)$$

$$H(x_l|y) = H(x_l, y) - H(y). \quad (5.25)$$

Bei Fuzzy-Systemen liegen wegen der graduellen Zuordnung keine wertediskreten Verteilungen vor, bei denen alle Merkmale und die Ausgangsgrößen nur bestimmte Werte (linguistische Terme) annehmen. Damit sind die Schätzungen für die Wahrscheinlichkeiten in (5.20)-(5.22) nicht ohne weiteres anwendbar.

Deshalb sind bei Fuzzy-Systemen spezielle Modifikationen zur Bestimmung der Häufigkeiten diskretisierter Klassen (linguistische Terme des Fuzzy-Systems) und ihrer Verbundverteilungen erforderlich. Dabei gibt es drei unterschiedliche Ansätze:

- Diskretisierung durch Verwendung der linguistischen Terme mit den größten Zugehörigkeitsgraden,
- Verwendung der Zugehörigkeitsgrade, des Produkt-Operators für UND-Verknüpfungen und des Summen-Operators für ODER-Verknüpfungen bzw.
- Schätzen der Verbundverteilung durch das Lösen eines restringierten Optimierungsproblems [81].

Bei der *Diskretisierung* können aus der unscharfen Partitionierung zwei scharfe Partitionierungen mit rechteckförmigen Zugehörigkeitsfunktionen ermittelt werden⁵:

⁵Die Transformation der hier verwendeten unscharfen Partitionierung in scharfe Partitionierungen mit rechteckförmigen Zugehörigkeitsfunktionen ist immer möglich und stets eindeutig. Die inverse Transformation von scharfen in unscharfe Partitionierungen mit den genannten Eigenschaften ist hingegen nur in Spezialfällen möglich und nicht immer eindeutig. Die stets mögliche Transformation von rechteck- in trapezförmige Zugehörigkeitsfunktionen ist hingegen nie eindeutig und führt tendenziell zu Interpretationsproblemen durch breite trapezförmige Zugehörigkeitsfunktionen.

1. Diskretisierung nach Termen: Der Einzugsbereich der m_l Terme erstreckt sich zwischen dem linken und rechten Schnittpunkt der jeweiligen Zugehörigkeitsfunktion mit der Linie $\mu = 0.5$ (α -Cut der Zugehörigkeitsfunktion mit $\alpha = 0.5$).
2. Diskretisierung nach Intervallen: Der Einzugsbereich der $m_l + 1$ Intervalle liegt zwischen den Stützpunkten mit $\mu = 1$ der Zugehörigkeitsfunktionen bzw. dem Rand des Definitionsbereiches (u. U. $+\infty$ bzw. $-\infty$).

Die Unterschiede zwischen einer Diskretisierung in Terme und Intervalle verdeutlicht Bild 5.6.

Die so diskretisierten Werte für die Terme werden dann auch bei der Berechnung der Verbundverteilung eingesetzt. Nach der Diskretisierung können nun wieder die Schätzungen (5.20)-(5.22) verwendet werden.

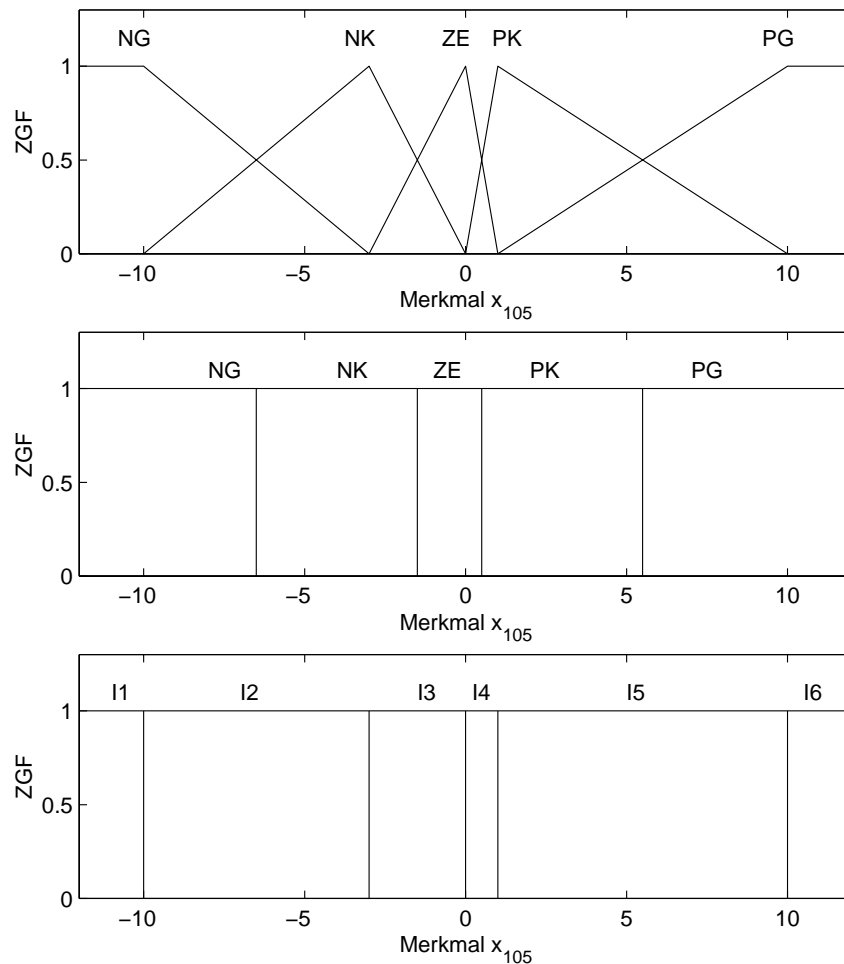


Bild 5.6: Diskretisierungen der Zugehörigkeitsfunktionen (oben) in Terme (Mitte) und Intervalle (unten) am Beispiel des Merkmals x_{105}

Nach einer Diskretisierung nach Intervallen lässt sich eine Eingangsentropie der Intervalle angeben:

$$H_I(x_l) = - \sum_{i=1}^{m_l+1} p(x_l = I_{l,i}) \cdot \log_2 p(x_l = I_{l,i}). \quad (5.26)$$

Bei Verwendung der *Zugehörigkeitsgrade*, des *Produkt-Operators* für *UND-Verknüpfungen* und des *Summen-Operators* für *ODER-Verknüpfungen* ergeben sich

$$\hat{p}(x_l = A_{l,i}) = \frac{1}{N} \sum_{k=1}^N \mu_{A_{l,i}}(x_{kl}), \quad (5.27)$$

$$\hat{p}(y = B_j) = \frac{1}{N} \sum_{k=1}^N \mu_{B_j}(y_k), \quad (5.28)$$

$$\hat{p}(x_l = A_{l,i} \cap y = B_j) = \frac{1}{N} \sum_{k=1}^N \mu_{A_{l,i}}(x_{kl}) \cdot \mu_{B_j}(y_k). \quad (5.29)$$

Allerdings produziert dieser naheliegende Ansatz insbesondere für die Elemente von $\hat{p}(x_l = A_{l,i} \cap y = B_j)$ fragwürdige Ergebnisse [78].

Wesentlich günstiger ist es deshalb, die Matrix der Verbundverteilungen durch das *Lösen eines restringierten Optimierungsproblems* zu schätzen. Der Lösungsweg dazu führt über den Zusammenhang

$$\hat{p}(x_l = A_{l,i} \cap y = B_j) = \hat{p}_{j,i}(y = B_j | x_l = A_{l,i}) \cdot \hat{p}(x_l = A_{l,i}) \quad (5.30)$$

und das Schätzen von $\hat{\mathbf{p}}(y|x_l)$ der bedingten relativen Häufigkeiten der Ausgangsklassen für gegebene linguistische Terme in x_l . $\hat{p}_{j,i}(y|x_l)$ ist dabei das Element (j, i) in einer Matrix $\hat{\mathbf{p}}(y|x_l)$, die die bedingten relativen Häufigkeiten der Ausgangsklassen für gegebene linguistische Terme in x_l enthält. Es gibt somit die geschätzte Häufigkeit dafür an, dass $y = B_j$ auftritt, wenn $x_l = A_{l,i}$ vorausgesetzt wird.

Mit den Matrizen der fuzzifizierten Werte für alle N Beispiele von x_l in $\boldsymbol{\mu}_{x_l}$ (5.1) mit der Dimension (m_l, N) und y in (5.15) mit der Dimension (m_y, N) folgt

$$Q(\hat{\mathbf{p}}(y|x_l)) = \frac{1}{2} \|\hat{\mathbf{p}}(y|x_l) \boldsymbol{\mu}_{x_l} - \boldsymbol{\mu}_y\|_F^2 \rightarrow \underset{\hat{\mathbf{p}}(y|x_l)}{\text{Min}}. \quad (5.31)$$

Die Restriktionen fordern positive Matrixelemente (5.32) mit Spaltensummen von Eins (5.33):

$$\hat{\mathbf{p}}(y|x_l) \geq \mathbf{0}_{m \times n}, \quad (5.32)$$

$$\mathbf{1}_m^T \hat{\mathbf{p}}(y|x_l) = \mathbf{1}_n^T. \quad (5.33)$$

In Analogie zu Wahrscheinlichkeiten können damit keine negativen Häufigkeiten auftreten und es wird garantiert, dass die Summe aller (relativen) Häufigkeiten Eins ist – es tritt garantiert eine der möglichen Ausgangsklassen auf. Die Ergebnisse von (5.27) und (5.28) können hingegen in diesem Ansatz verwendet werden.

In allen Fällen ist zu beachten, dass die informationstheoretischen Maße von den Parametern der Zugehörigkeitsfunktionen abhängen. Über den Ansatz in (5.31-5.33) können folglich nun für jeden Parametersatz \mathbf{a}_l der Zugehörigkeitsfunktionen eines Merkmals informationstheoretische Maße berechnet werden.

Jedes der informationstheoretischen Maße kann nun als Gütemaß für die Erfüllung einer oder mehrerer der oben genannten Forderungen für Zugehörigkeitsfunktionen dienen (siehe Tabelle 5.2).

Je größer die Transinformation ist, desto stärker ist der (durchschnittliche) Zusammenhang zwischen den linguistischen Termen des Merkmals und der Ausgangsgröße, was

	$H(x_l)$	$H_I(x_l)$	$H(x_l; y)$
Separierung der Ausgangsklassen	--	--	++
Beispielanzahl (Term)	++	+	--
Beispielanzahl (Intervall)	-	++	--
Berücksichtigung seltener Beispiele	--	--	+
Teilbarkeit durch zusätzliche Merkmale	+	++	-
gute Interpretation	--	-	--

Tabelle 5.2: Bewertung informationstheoretischer Maße nach verschiedenen Forderungen

eine günstige Ausgangsposition für die spätere Regelsuche darstellt. Ein Problem bei der Transinformation besteht aber darin, dass sie bei gestörten Daten zufällige, kleine Intervalle mit einer homogenen Zusammensetzung der Ausgangsklassen bevorzugt. Durch diese kleinen Intervalle werden tendenziell linguistische Terme mit einem kleinen Einzugsgebiet erzeugt, die für eine spätere Teilung durch andere Merkmale nicht geeignet sind. Das reduziert aber letztlich die Klassifikationsgüte und beeinträchtigt zudem die Interpretierbarkeit der Zugehörigkeitsfunktion.

Je größer die Eingangsentropie ist, desto gleichmäßiger sind die Daten zwischen den linguistischen Termen bzw. Intervallen verteilt, was auch eine spätere Teilbarkeit durch die UND-Verknüpfung mit anderen Merkmalen begünstigt. Zur besseren Vergleichbarkeit werden die Transinformation und die Eingangsentropien der Intervalle und Terme durch die jeweils maximal möglichen Werte $H(y)$ (Ausgangsentropie), $H_{max}(x_l)$ und $H_{I,max}(x_l)$ geteilt, um so Ergebnisse zwischen Null und Eins zu erhalten.

Ein Nachteil besteht aber darin, dass immer nur paarweise Zusammenhänge zwischen *einem* Merkmal und der Ausgangsgröße untersucht werden. Außerdem sind die Anforderungen widersprüchlich, so dass nach einem Kompromiss zu suchen ist.

Zudem leistet kein informationstheoretisches Maß einen Beitrag zur Verbesserung der Interpretierbarkeit der Ergebnisse. Besonders gut interpretierbar sind bei vielen Anwendungen ganzzahlige Werte innerhalb der jeweiligen Zehnerpotenz. Innerhalb dieser Werte ist häufig noch die Null besonders hervorzuheben, sofern sie im Bereich enthalten ist. Je stärker die Werte verfeinert sind, desto geringer ist die Akzeptanz.

Deswegen wird ein mehrstufiges Verfahren vorgeschlagen, das die Interpretierbarkeit von Parametern für Zugehörigkeitsfunktionen bewertet. Das Verfahren soll anhand des Merkmals x_{105} des Beispiels erläutert werden, das Werte zwischen -7.69 und 8.06 aufweist.

In einem ersten Schritt wird die Differenz zwischen dem maximalen und minimalen Wert des Merkmals berechnet, der Zehnerlogarithmus gebildet und der Wert gerundet (Operation rd). Dadurch entsteht eine Zahl p , die die Ausdehnung des Wertebereichs charakterisiert

$$p = \text{rd}(\log_{10}(\max(\mathbf{x}_l) - \min(\mathbf{x}_l))), \quad (5.34)$$

was im Beispiel auf $p = \text{rd}(\log_{10}(15.75)) = \text{rd}(1.20) = 1$ führt. In einem darauf folgenden Schritt werden Interpretierbarkeitsklassen $\mathfrak{A}_{l,q}$ aufgestellt, die auf zunehmend verfeinerten Rundungsoperationen im Bereich der reellen Zahlen \mathbb{R} aufbauen:

$$\mathfrak{A}_{l,q} = \left\{ \alpha \in \mathbb{R} : \alpha = \text{rd}\left(\frac{x_l}{10^{p+1-q}}\right) \cdot 10^{p+1-q} \text{ und } \alpha \notin \bigcup_{i=0}^{q-1} \mathfrak{A}_{l,i}; \right. \\ \left. q = 0, 1, \dots, q_{max}; \min(\mathbf{x}_l) \leq x_l \leq \max(\mathbf{x}_l) \right\}. \quad (5.35)$$

Die Werte sind äquidistant über den Wertebereich des jeweiligen Merkmals verteilt, wobei mit zunehmenden q die Feinheit der Teilung um eine Zehnerpotenz zunimmt. Werte von stärker abstrahierten Interpretierbarkeitsklassen gehören allerdings nicht mehr zu feineren Interpretierbarkeitsklassen. Dabei entstehen die in Tabelle 5.3 aufgeführten Hypothesen für Parameter durch Runden von Werten innerhalb des betrachteten Zahlenbereiches. Je abstrakter die Interpretierbarkeitsstufe, desto höher werden im Folgenden die Werte einer Interpretierbarkeit I gewählt.

Wenn außerdem anstelle des gesamten Wertebereichs $\min(\mathbf{x}_l) \leq x_l \leq \max(\mathbf{x}_l)$ nur die im Datensatz vorkommenden Werte \mathbf{x}_l entsprechend (5.35) gerundet werden, entstehen hauptsächlich Werte in jenen Bereichen, die auch durch den Datensatz abgedeckt sind.

Interpretierbarkeits- klasse $\mathfrak{A}_{l,q}$	Merkmal $x_6(p=2)$ $-5.50 \leq x_6 \leq 69.24$	Merkmal $x_{105}(p=1)$ $(-7.69 \leq x_{105} \leq 8.06)$	Interpretierbarkeit I
$\mathfrak{A}_{l,0}$	0	0	sehr groß (0.15)
$\mathfrak{A}_{l,1}$	100	-10; 10	groß (0.10)
$\mathfrak{A}_{l,2}$	-10;10;...;70	-8; -7; -6;...; 8	mittel (0.05)
$\mathfrak{A}_{l,3}$	-6;-5;...;69	-7.7; 7.6;...; 8.1	gering (0.00)
$\mathfrak{A}_{l,4}$	-5.5;-5.4;...;69.2	-7.69; 7.68;...; 8.06	sehr gering (0.00)

Tabelle 5.3: Interpretierbarkeit mit Werten I für verschiedene Hypothesen von ZGF-Parametern am Beispiel zweier Wertebereiche für die Merkmale x_6 und x_{105}

5.4.3 Entwurfsverfahren

Die Aufgabe besteht nun in der Festlegung der Parameter von dreieck- bzw. trapezförmigen Zugehörigkeitsfunktionen (Stützpunkte $a_{l,i}$) für die Merkmale x_l . Die Zahl der Zugehörigkeitsfunktionen wird dabei a priori festgesetzt, wobei im Folgenden $m_l = 5$ gewählt wird.

Der Algorithmus geht von einer Starthypothese aus, bei der als Parameter der minimale und der maximale Funktionswert des Merkmals x_l und die Werte verwendet werden, die eine möglichst gleiche Beispiellanzahl für die Terme bewirken. Damit ergibt sich der Vektor der Parameter \mathbf{a}_l . Für das bereits diskutierte Merkmal x_{105} lautet diese Starthypothese

$$\mathbf{a}_{105} = (-7.69 \quad -1.16 \quad 0.02 \quad 1.30 \quad 8.06)^T. \quad (5.36)$$

Ausgehend von diesen Werten werden nacheinander alle Parameter gestrichen und durch den neuen Stützpunkt ersetzt, für den ein gewichtetes Maß (Faktoren α und β) aus Transformation $H(x_l; y)$, Eingangsentropie der Intervalle $H_I(x_l)$ und Interpretierbarkeit I maximal wird:

$$Q = \alpha \frac{H(x_l; y)}{H(y)} + (1 - \alpha) \frac{H_I(x_l)}{H_{I,max}(x_l)} + \beta I \rightarrow \max_{\mathbf{a}_l}. \quad (5.37)$$

Wie bereits im vorherigen Abschnitt diskutiert, werden die Transformation und die Eingangsentropie der Intervalle durch die jeweils maximal möglichen Werte $H(y)$ (Ausgangsentropie) bzw. $H_{I,max}(x_l)$ geteilt, um so Ergebnisse zwischen Null und Eins zu erhalten.

Im Ergebnis wählt der Algorithmus jeweils gut interpretierbare Werte

$$\mathbf{a}_{105} = (-10 \quad -3 \quad 0 \quad 1 \quad 10)^T, \quad (5.38)$$

die in der Nähe optimaler, aber schwer interpretierbarer Werte liegen.

Einen Entwurfsschritt, die Festlegung des Parameters $a_{105,3}$ für das oben genannte Beispiel, zeigt Bild 5.7. Im Bild sind von oben nach unten die gewichteten Eingangsentropien für Terme $\frac{H(x_i)}{\bar{H}_{I,max}(x_i)}$ und für Intervalle $\frac{H_I(x_i)}{\bar{H}_{I,max}(x_i)}$, das Kompromiss-Maß Q , die auf die Ausgangsentropie bezogene Transinformation $\frac{H(x_i;y)}{H(y)}$ und die Interpretierbarkeit I aufgetragen. Das Bild verdeutlicht nochmals, dass jedes der Maße allein eine andere Entscheidung trifft. Die Kompromissentscheidung führt auf den Wert Null, was in erster Linie durch die Interpretierbarkeit entschieden wird.

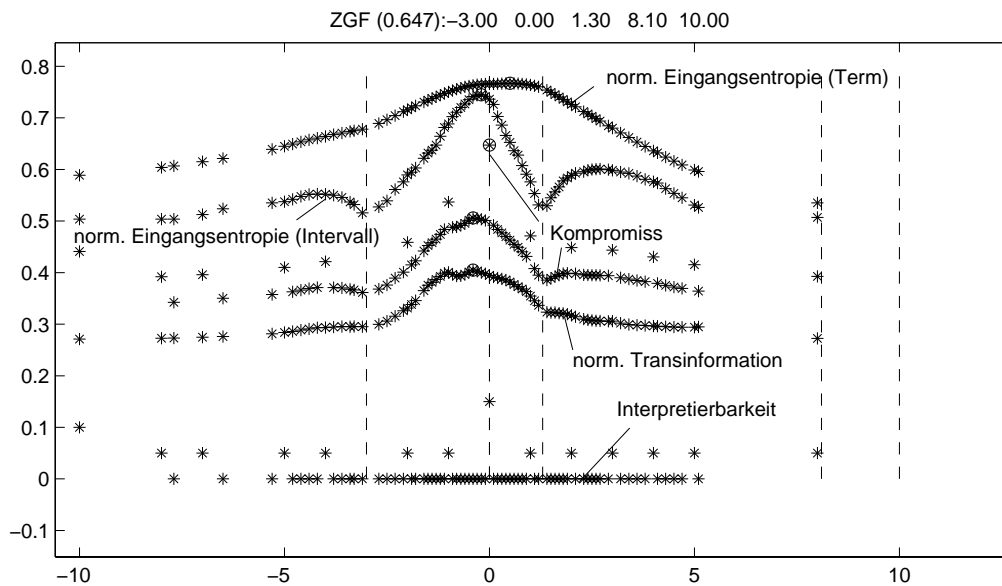


Bild 5.7: Entwurfsschritt für den Parameter $a_{105,3}$

Die so gefundene Entscheidung ist nicht optimal, weil durch das schrittweise Ersetzen aller Parameter nicht der gesamte Suchraum abgesucht wird. Aufgrund der aufwendigen Berechnungen für alle Bestandteile von Q ist eine vollständige Suche aber praktisch inakzeptabel.

Die linguistischen Terme werden dabei automatisch mit Namen wie Null (ZE), Positiv Klein (PK), Positiv Mittel (PM), Positiv Groß (PG), Positiv Sehr Groß (PSG), Negativ (N) usw. versehen.

Weitere Rechenvereinfachungen bei meist vertretbaren Güteverringern sind möglich, wenn die Werte der Startschätzung solange in aufsteigende Interpretierbarkeitsklassen gerundet werden, bis benachbarte Werte nicht mehr gleich sind.

Wenn die Parameter \mathbf{b} der Zugehörigkeitsfunktionen der Ausgangsgröße nicht explizit gegeben sind, empfiehlt sich eine separate Optimierung *vor* der Ermittlung der Parameter der Zugehörigkeitsfunktionen der Merkmale \mathbf{a}_l . Dabei sollten nur die Entropie der Intervalle der Ausgangsgröße und die Interpretierbarkeit berücksichtigt werden. Anderenfalls steigt der Berechnungsaufwand wegen der gegenseitigen Beeinflussung von Parametern der Zugehörigkeitsfunktionen von Merkmalen und Ausgangsgrößen bei der Ermittlung der Transinformation stark an.

Als Ergebnisse des Entwurfs von Zugehörigkeitsfunktionen entstehen folglich Parametervektoren von Zugehörigkeitsfunktionen \mathbf{a}_l für die Merkmale und evtl. \mathbf{b} für die Aus-

gangsgröße. Mit diesen Informationen können nun Messwerte und bekannte Ausgangsgrößen fuzzifiziert werden.

5.5 Merkmalsselektion

Bei einer Merkmalsselektion werden wie bei den statistischen Verfahren Merkmale ausgewählt, die später in Fuzzy-Regeln eingehen sollen. Die Ziele einer Merkmalsselektion sind eine bessere Klassifikation über den Testdaten durch Unterdrückung schlechter Merkmale, die Bewertung der Merkmale bei einer interaktiven Regelgenerierung, die automatisierte Suche nach zusätzlichen Merkmalen sowie reduzierte Rechenzeiten.

Sowohl die Merkmalsselektion als auch die nachfolgende Aufstellung von Regelhypothesen erfolgen über Entscheidungs bäume. Ein Entscheidungsbaum ist ein Modell für einen mehrstufigen Entscheidungsprozess zur Klassifikation eines Objekts anhand der Werte seiner Merkmale. Er besteht aus Knoten und Zweigen und kann als gerichteter Graph interpretiert werden (vgl. Beispiel in Bild 5.8). Details über die einzelnen Knoten, wie Name, Vorgängerknoten in der höheren Hierarchieebene, Spezialisierung des Merkmals der höheren Hierarchieebene, Entscheidung y im Knoten, Merkmal zur Auftrennung (Entscheidungsvariable), Fehler und Beispiele können Tabelle 5.4 entnommen werden. Die Buchstaben W und E hinter den Knotennamen kennzeichnen Wurzel- bzw. Endknoten.

In der obersten Hierarchieebene befindet sich der Wurzelknoten. Jeder Knoten legt für die Ausgangsgröße y einen linguistischen Term (Klasse) fest, der für die zugehörige Beispielmenge die beste Entscheidung darstellt. Im Allgemeinen ist das die häufigste Ausgangsklasse. Die gewählte Ausgangsklasse wird im Bild 5.8 als oberer Wert in jedem Knoten eingetragen, z. B. 3 für $y = B_3$ im Wurzelknoten).

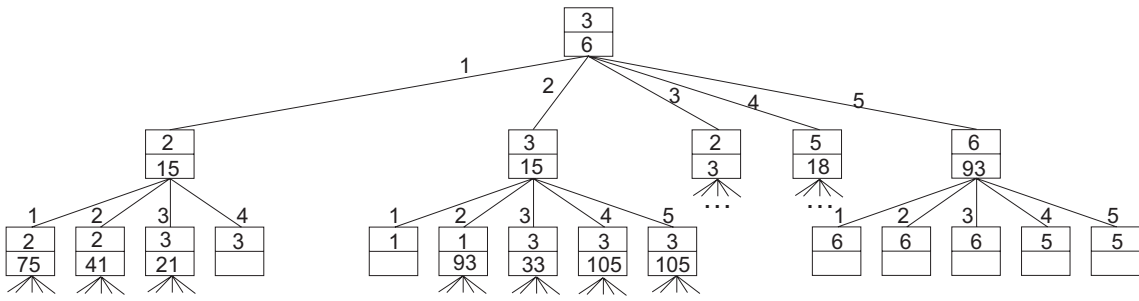


Bild 5.8: Entscheidungsbaum

Wenn die Beispielmenge in einem Knoten homogen bezüglich der Klassenzuordnung ist, d. h. korrekt klassifiziert wird, ist ihre weitere Auftrennung nicht notwendig und der Knoten wird somit zu einem Endknoten. Wenn die Beispielmenge noch Klassifikationsfehler aufweist, wird sie beim Entwurf des Entscheidungsbaums in Abhängigkeit vom Wert eines Merkmals x_l (Entscheidungsvariable) aufgespalten (im Bild 5.8 unterer Eintrag im Knoten, z. B. 6 für x_6 im Wurzelknoten). Das dafür optimale Merkmal x_l wird über ein Kriterium bestimmt.

Damit ergibt sich eine Verzweigung in die Nachfolgeknoten, deren Anzahl gleich der Anzahl der auftretenden Werte des Merkmals x_l ist (im Bild 5.8 Merkmalswerte am Verbindungsweig, z. B. 1 bis 5 für die linguistischen Terme $A_{6,1}$ bis $A_{6,5}$ im Wurzelknoten)⁶.

⁶Wenn bestimmte Spezifikationen im Lerndatensatz nicht vorkommen (z. B. $A_{15,5}$ im Knoten v_2), wird

Knoten	Elter	Verzweigung	y	x_i	Fehler	Beispiele
$v_1(W)$	–	–	B_3	x_6	912	1144
v_2	v_1	$A_{6,1}$	B_2	x_{15}	81	132
v_3	v_1	$A_{6,2}$	B_3	x_{15}	147	297
v_4	v_1	$A_{6,3}$	B_2	x_3	222	351
v_5	v_1	$A_{6,4}$	B_5	x_{18}	151	229
v_6	v_1	$A_{6,5}$	B_6	x_{93}	36	135
v_7	v_2	$A_{15,1}$	B_2	x_{75}	1	7
v_8	v_2	$A_{15,2}$	B_2	x_{41}	32	68
v_9	v_2	$A_{15,3}$	B_3	x_{21}	8	26
$v_{10}(E)$	v_2	$A_{15,4}$	B_3	–	3	31
$v_{11}(E)$	v_3	$A_{15,1}$	B_1	–	8	16
v_{12}	v_3	$A_{15,2}$	B_1	x_{93}	25	72
v_{13}	v_3	$A_{15,3}$	B_3	x_{33}	21	35
v_{14}	v_3	$A_{15,4}$	B_3	x_{105}	24	110
v_{15}	v_3	$A_{15,5}$	B_3	x_{105}	14	64
...						
$v_{26}(E)$	v_6	$A_{93,1}$	B_6	–	1	4
$v_{27}(E)$	v_6	$A_{93,2}$	B_6	–	1	53
$v_{28}(E)$	v_6	$A_{93,3}$	B_6	–	1	18
$v_{29}(E)$	v_6	$A_{93,4}$	B_5	–	27	56
$v_{30}(E)$	v_6	$A_{93,5}$	B_5	–	0	4

Tabelle 5.4: Ausschnitt aus einem Entscheidungsbaum

Dieses Vorgehen wird solange wiederholt, bis in der jeweils untersten Hierarchieebene nur noch Endknoten entstehen. In den tieferen Hierarchieebenen nehmen somit die Zahl der Beispiele und (zumindest tendenziell) auch die Zahl der Fehler ab.

Im Idealfall eines ungestörten und eindeutigen Datenmaterials garantiert der Entscheidungsbaum eine fehlerfreie Klassifikation des gesamten Lerndatensatzes. Bei gestörten und nicht eindeutigen Daten wird ein an den Lerndatensatz optimal angepasster Entscheidungsbaum generiert, der auch die Störungen modelliert. Deswegen ist er nicht in der Lage, befriedigend zu generalisieren, d. h. nicht im Lerndatensatz enthaltene Beispiele richtig zu klassifizieren.

Die Auswertung eines gegebenen Entscheidungsbaums für ein Beispiel mit unbekannter Ausgangsklasse beginnt im Wurzelknoten. Dort wird derjenige Zweig ausgewählt, der zum Wert der Entscheidungsvariable im Wurzelknoten gehört. Im Nachfolgeknoten, den der Zweig mit dem Wurzelknoten verbindet, wird wiederum die dortige Entscheidungsvariable untersucht und ein Zweig sowie der nächste Knoten ausgewählt. Das wird iterativ solange wiederholt, bis ein Endknoten erreicht wird. Dieser Endknoten liefert dann die Entscheidung für die Ausgangsklasse des Beispiels (oberer Eintrag im Knoten). Die Entscheidungen über die Ausgangsklasse in den vorher durchlaufenen Knoten werden nicht verwendet.

Die Algorithmen für den Entwurf von Entscheidungsbäumen unterscheiden sich nach

die Entscheidung des letzten durchlaufenen Knotens übernommen, der somit partiell zum Endknoten werden kann.

dem Kriterium, mit dem das Merkmal x_l für die Aufspaltung ausgewählt wird. Im Folgenden wird kurz auf das ID3- und das C4.5-Verfahren eingegangen.

Der ID3-Algorithmus [89] setzt nominale Merkmale voraus. Die Beispielmengende eines Knotens wird bezüglich der linguistischen Terme desjenigen Merkmals aufgespalten, das die Transinformation (5.23) maximiert. Dies ist gleichbedeutend mit einer Minimierung der Irrelevanz oder auch der Maximierung des Verhältnisses von Transinformation zu Ausgangsentropie

$$\frac{H(x_l; y)}{H(y)} \rightarrow \text{Max}_{l=1, \dots, m_l} . \quad (5.39)$$

Beim C4.5-Algorithmus [90] wird ein modifiziertes Entscheidungskriterium, und zwar das Verhältnis von Transinformation und Eingangsentropie

$$\frac{H(x_l; y)}{H(x_l)} \rightarrow \text{Max}_{l=1, \dots, m_l} \quad (5.40)$$

verwendet. Damit werden im Gegensatz zum ID3-Verfahren Aufteilungen nach Merkmalen unterdrückt, die eine sehr hohe Eingangsentropie (z. B. durch viele linguistische Terme) aufweisen. Bei gleicher Anzahl linguistischer Terme in den Merkmalen und ähnlichen Verteilungen sind hingegen die Ergebnisse beider Verfahren nahezu gleich. Im Folgenden wird das ID3-Verfahren eingesetzt.

Die Berechnung der relativen Häufigkeiten für die Transinformation bezieht sich dabei immer auf die verbliebene Beispieldatenmenge in einem Knoten und nicht auf den gesamten Lerndatensatz. Im diskutierten Entscheidungsbaum existieren im Wurzelknoten v_1 1144 Beispiele, von denen mit der Entscheidung $y = B_3$ 912 fehlerhaft klassifiziert werden (Tabelle 5.4). Nach Auftrennung mittels x_6 werden im Knoten v_2 noch die 132 Beispiele des Lerndatensatzes untersucht, die für x_6 den Wert $A_{6,1}$ aufweisen. Eine weitere Auftrennung dieser 132 Beispiele führt im Knoten v_7 noch zu 7 Beispielen nach der Auftrennung mit $x_{15} = A_{15,1}$ im Knoten v_2 usw.

Deswegen müssen alle Ergebnisse von (5.23) für jedes noch in Frage kommende Merkmal x_l in jedem Knoten jeweils neu berechnet werden. Damit ist keine vollständige Prämissenstruktur auszuwerten, sondern immer nur das aktuelle Merkmal zu untersuchen. Wenn kein weiteres Merkmal existiert oder der Zuwachs der Transinformation zu gering ist, wird der Knoten zu einem Endknoten.

Beim Aufstellen des Entscheidungsbaums steht für jeden der K einbezogenen Knoten für alle Merkmale x_l und die zum Knoten gehörenden N_k Beispiele das Verhältnis der Transinformation $H_k(x_l; y)$ zur Ausgangsentropie $H_k(y)$ zur Verfügung. Dieses Verhältnis ist stets ein Maß dafür, wie gut sich ein Merkmal für die jeweilige Klassifikationsaufgabe eignet.

Während bei Entscheidungsbäumen immer nur das beste Merkmal verwendet wird, liefern die bei der Entscheidung anfallenden Werte – Transinformation pro Ausgangsentropie (5.39) bzw. Transinformation pro Eingangsentropie (5.40) – zusätzlich quantitative Informationen über die Relevanz von Merkmalen, Informationen über nächstbeste Merkmale und, bei einer Mittelung über verschiedene Knoten, einen Durchschnittswert für die Merkmalsrelevanz.

Die Formalisierung dieser Idee liefert die Merkmalsrelevanz für das l -te Merkmal:

$$M_l = \frac{\sum_{k=1}^K N_k \frac{H_k(x_l; y)}{H_k(y)}}{\sum_{k=1}^K N_k} . \quad (5.41)$$

Merkmal	Bezeichnung	Güte
x_6	R-KNEE Flex ANG	0.276
x_{90}	R-HIP Flex ANG V	0.273
x_3	R-HIP Flex ANG	0.261
x_{102}	L-HIP Flex ANG V	0.260
x_{93}	R-KNEE Flex ANG V	0.246
x_{15}	L-HIP Flex ANG	0.245
x_{21}	L-ANK Flex ANG	0.239
x_{105}	L-KNEE Flex ANG V	0.233
x_{18}	L-KNEE Flex ANG	0.212
x_{14}	L-HIP Abd ANG	0.211

Merkmal	Bezeichnung	Güte
x_6	R-KNEE Flex ANG	0.175
x_3	R-HIP Flex ANG	0.161
x_{102}	L-HIP Flex ANG V	0.158
x_{93}	R-KNEE Flex ANG V	0.156
x_{15}	L-HIP Flex ANG	0.150
x_{18}	L-KNEE Flex ANG	0.150
x_{105}	L-KNEE Flex ANG V	0.147
x_{90}	R-HIP Flex ANG V	0.140
x_{21}	L-ANK Flex ANG	0.135
x_{14}	L-HIP Abd ANG	0.122

Tabelle 5.5: Beste Merkmalsrelevanzen der Einzelmerkmale (links, Wurzelknoten ID3) und durchschnittliche Merkmalsrelevanz unter Berücksichtigung von Redundanzen (rechts, klassenspezifische ID3) – Bezeichnungen: wie Tabelle 4.1, zusätzlich ANK (Fußgelenk), Abd (Abduktion – Gelenkwinkel in der Frontalebene)

Jeder Knoten k geht dabei um so stärker ein, je größer N_k ist. Die Merkmalsrelevanz im Wurzelknoten ($K = 1$) weist Parallelen zur Auswahl eines Einzelmerkmals mit univariaten statistischen Verfahren auf. Diese Relevanz berücksichtigt die spezielle Eignung des Merkmals für den Entwurf von Fuzzy-Systemen, weil sie Unterschiede in der Verteilung von Ausgangsklassen bei linguistischen Termen der Eingangsklasse bewertet.

Der linke Teil der Tabelle 5.5 zeigt die Merkmalsrelevanzen im Wurzelknoten des Entscheidungsbaums. Auch diese Maße nehmen wieder Werte zwischen Null (irrelevant) und Eins (relevant) an. Das Verfahren kommt zu ähnlichen Ergebnissen wie der t -Test in Tabelle 4.1 (links), wobei die konkreten Zahlenwerte keinen direkten Vergleich zwischen beiden Verfahren zulassen.

Für die Erkennung der Schrittphasen sind auch hier Merkmale wichtig, die in der sagittalen Ebene Knie- und Hüftgelenkwinkel und ihre Änderungsgeschwindigkeiten bewerten, wobei sich die Reihenfolge der Merkmale gegenüber dem t -Test geringfügig unterscheidet. Dabei werden im Folgenden 10 Merkmale verwendet, weil Fuzzy-Verfahren für eine gute Klassifikationsgüte erfahrungsgemäß eine größere Anzahl von Merkmalen im Vergleich zu statistischen Verfahren benötigen. Eine vollständige Liste der Merkmalsrelevanzen aller Merkmale kann Anhang A.3 entnommen werden.

Die Merkmalsrelevanzen in Knoten tieferer Hierarchieebenen berücksichtigen hingegen Redundanzen zwischen mehreren Merkmalen. Die Ursache liegt darin, dass durch die bereits getroffenen Entscheidungen in höheren Hierarchieebenen nur die Merkmale in tieferen Hierarchieebenen noch Vorteile für die Separierung widersprüchlicher Ausgangsklassen bringen, die nicht zu stark mit den bereits ausgewählten Merkmalen zusammenhängen.

Eine gewichtete Summe von Merkmalsrelevanzen aller Knoten eines Entscheidungsbaums ist somit ein Maß für die multivariaten Relevanzen verschiedener Merkmale.

Das Verfahren ist auch dann anwendbar, wenn mehrere Entscheidungsbäume generiert und bezüglich Merkmalsrelevanzen ausgewertet werden. Das können einerseits Entscheidungsbäume mit den nächstbesten Merkmalen im Wurzelknoten oder klassenspezifische Entscheidungsbäume sein, die immer die Unterschiede zwischen einer Ausgangsklasse und deren Negation auswerten. Klassenspezifische Entscheidungsbäume finden insbesondere Zusammenhänge, die eine spezielle Ausgangsklasse B_j von allen anderen unterscheiden, was genau der Philosophie einer Regel mit der Konklusion B_j entspricht. Zudem bieten sie geeignete Ansatzpunkte zur Suche nach negativen Regeln (\dots DANN NICHT B_j) entsprechend [58], was in dieser Arbeit aber nicht gesondert verfolgt wird.

Die mit sieben zusätzlichen klassenspezifischen Entscheidungsbäumen entstehenden Merkmalsrelevanzen sind in der rechten Teiltabelle von Tabelle 5.5 angegeben. Sie unterscheiden sich weder signifikant von den Relevanzen der Einzelmerkmale noch von den Ergebnissen der statistischen multivariaten Auswahl mit dem MANOVA-Verfahren. Allerdings geht die Wertigkeit des Merkmals x_{90} im Vergleich zur Einzelrelevanz zurück. Die vollständige Liste aller Merkmalsrelevanzen, die mit diesem Verfahren ermittelt wurde, steht in Anhang A.4.

Im Unterschied zu MANOVA beziehen sich die Relevanzen nicht auf eine Gruppe von Merkmalen, sondern auf eine durchschnittliche Relevanz des Merkmals unter Berücksichtigung der Auswahlentscheidungen in höheren Hierarchieebenen des Entscheidungsbaums. Deswegen ist auch hier weder ein Zahlenvergleich der Merkmalsrelevanzen zum MANOVA-Verfahren noch zu den Einzelrelevanzen im Wurzelknoten aussagekräftig.

Die Merkmalsselektion hat den Vorteil, dass das aufwendige Optimierungsverfahren für die Zugehörigkeitsfunktionen nur für die ausgewählten Merkmale durchgeführt werden muss, wenn für die Merkmalsselektion zunächst einfach entworfene Zugehörigkeitsfunktionen (hier: etwa gleiche Häufigkeiten für alle linguistischen Terme) akzeptiert werden. Der erhebliche Vorteil im Rechenaufwand bei dieser Variante hat allerdings Ungenauigkeiten in der Merkmalsselektion zur Folge. Zudem ändern sich dadurch die Zahlenwerte der Merkmalsrelevanzen. Da die Vorteile die Nachteile überwiegen, wird dieses Konzept hier verwendet. Die Parameter der so optimierten Zugehörigkeitsfunktionen für die ausgewählten Merkmale sind in Anhang A.5 aufgeführt.

Ergebnis der Merkmalsselektion ist somit eine Liste von 10 Merkmalen (Tabelle 5.5, rechts), die in der nachfolgenden Regelgenerierung Verwendung finden.

5.6 Regelgenerierung

5.6.1 Ziele

Zur Regelgenerierung werden Hypothesen über Regeln aufgestellt und bewertet. Bei der Bewertung von Fuzzy-Regeln ist zu unterscheiden,

- ob Einzelregeln, Gruppen von Regeln oder komplette Regelbasen zu bewerten sind,
- ob die Konklusionen bekannt sind oder nicht und
- welches Abstraktionsniveau gefordert ist.

Ein hohes Abstraktionsniveau verlangt einfache, klare Regeln, die die wesentlichen Zusammenhänge für komplexe Untersuchungen beschreiben. Sonderfälle (seltene Situationen usw.) bleiben unberücksichtigt. Bei einem geringen Abstraktionsniveau liegt der Schwerpunkt auf einer genauen Nachbildung des Systemverhaltens um den Preis detaillierterer Regeln. Für die Regelgenerierung sind aus den Bewertungsmaßen ableitbare Aussagen über

- den Einfluss einer Regel bzw. Gruppe von Regeln auf die Klassifikationsgüte,
- die Klarheit der Aussage einer Regel bzw. Gruppe von Regeln,
- den Unterschied zwischen Regeln einer Regelbasis (globale Bedeutung von Regeln),
- die Komplexität von Regeln und Regelbasen,
- die Redundanz einer Regelbasis sowie
- die (statistische) Absicherung einer Regel bzw. Regelbasis

von Interesse. Je nach Aufgabenstellung stehen verschiedene Forderungen im Mittelpunkt.

Im ersten Schritt der Hypothesengenerierung für Einzelregeln und Regelbasen werden erneut Entscheidungsbäume verwendet, um spezielle Regeln zu erzeugen. Diese speziellen Regeln werden unter Verwendung von modularen Maßen dann wieder verallgemeinert (Bild 5.9) und zu Regelbasen zusammengestellt.

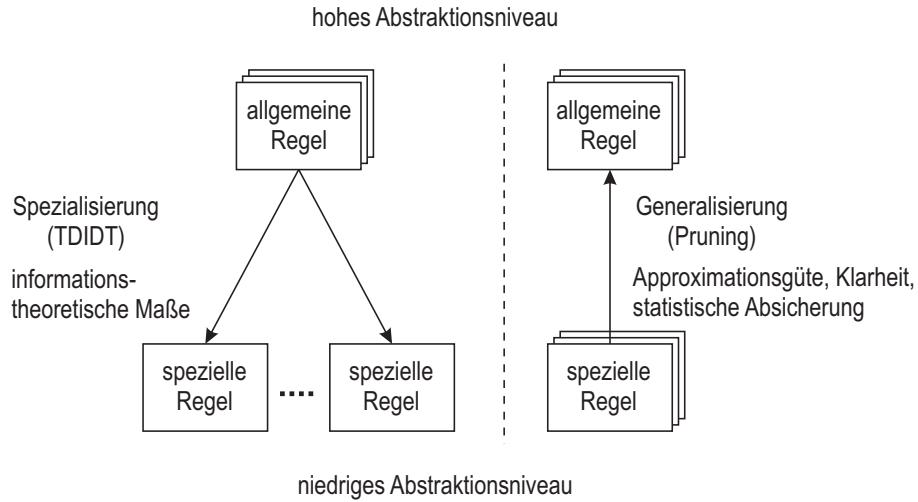


Bild 5.9: Prinzip der Regelgenerierung

5.6.2 Hypothesengenerierung mit Entscheidungsbäumen

Wegen der alleinigen Verwendung der ausgewählten Merkmale zur Aufspaltung und den optimierten Zugehörigkeitsfunktionen sind neue Entscheidungsbäume zu generieren (siehe Bild 5.10). Die Optimierung der Zugehörigkeitsfunktionen erklärt den Effekt, dass im Wurzelknoten des Entscheidungsbaums nun Merkmal x_{90} anstatt Merkmal x_6 Verwendung findet.

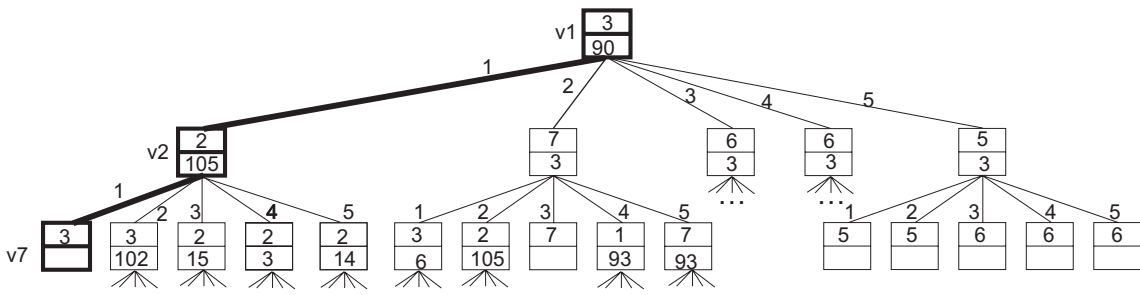


Bild 5.10: Entscheidungsbaum mit reduzierten Merkmalen

Aus Entscheidungsbäumen können dann Regeln generiert werden. Jeder Endknoten (Knoten ohne weitere Verzweigung) liefert eine Regel, deren Konklusion die Entscheidung über die Ausgangsgröße im Knoten ist. Die Prämisse folgt aus der UND-Verknüpfung (Bezeichnung: \cap) aller spezifizierten Merkmale x_l auf dem Rückweg zum Wurzelknoten.

Knoten	Elter	Verzweigung	y	x_i	Fehler	Beispiele
$v_1(W)$	–	–	B_3	x_{90}	912	1144
v_2	v_1	$A_{90,1}$	B_2	x_{105}	272	455
v_3	v_1	$A_{90,2}$	B_7	x_3	241	333
v_4	v_1	$A_{90,3}$	B_6	x_3	99	175
v_5	v_1	$A_{90,4}$	B_6	x_3	41	104
v_6	v_1	$A_{90,5}$	B_5	x_3	25	77
$v_7(E)$	v_2	$A_{105,1}$	B_3	–	0	12
v_8	v_2	$A_{105,2}$	B_3	x_{102}	12	159
v_9	v_2	$A_{105,3}$	B_2	x_{15}	46	72
v_{10}	v_2	$A_{105,4}$	B_2	x_3	64	206
v_{11}	v_2	$A_{105,5}$	B_2	x_{14}	2	6
v_{12}	v_3	$A_{3,1}$	B_3	x_6	85	140
v_{13}	v_3	$A_{3,2}$	B_2	x_{105}	7	20
$v_{14}(E)$	v_3	$A_{3,3}$	B_7	–	13	28
v_{15}	v_3	$A_{3,4}$	B_1	x_{93}	31	66
v_{16}	v_3	$A_{3,5}$	B_7	x_{93}	29	79
...						
$v_{27}(E)$	v_6	$A_{3,1}$	B_5	–	0	12
$v_{28}(E)$	v_6	$A_{3,2}$	B_5	–	3	34
$v_{29}(E)$	v_6	$A_{3,3}$	B_6	–	9	20
$v_{30}(E)$	v_6	$A_{3,4}$	B_6	–	0	10
$v_{31}(E)$	v_6	$A_{3,5}$	B_6	–	0	1

Tabelle 5.6: Ausschnitt aus einem Entscheidungsbaum in Bild 5.10

Das Prinzip ist anhand einer Regel in Bild 5.10 mit den fett gekennzeichneten Linien skizziert.

So entsteht die Regel WENN $x_{90} = A_{90,1} \cap x_{105} = A_{105,1}$ DANN $y = B_3$, wobei der Ausdruck $x_{90}=A_{90,1}$ aus der Verbindung zwischen Wurzelknoten und zweiter Hierarchieebene (Knoten v_1 und v_2 in Tabelle 5.6) und $x_{105}=A_{105,1}$ aus der Verbindung zwischen zweiter und dritter Hierarchieebene resultiert (Knoten v_2 und v_7 in Tabelle 5.6). Die Ausgangsklasse $y=B_3$ ist dem Knoten v_7 zugeordnet. Die entstehende Regel ist fehlerlos und durch 12 Beispiele abgesichert.

Im Beispiel werden acht verschiedene Entscheidungsbäume generiert: Ein Entscheidungsbaum spezifiziert dabei alle sieben Ausgangsklassen gegeneinander und die anderen sieben klassenspezifischen Entscheidungsbäume unterscheiden jeweils eine Ausgangsklasse gegen ihre Negation (ODER-Verknüpfung aller anderer Ausgangsklassen). Aus diesen Entscheidungsbäumen werden 220 Regeln extrahiert, wobei gleichlautende Regeln, die aus verschiedenen Entscheidungsbäumen entstehen, entfernt werden.

Der Effizienz des Baum-Induktionsverfahrens stehen allerdings verschiedene Nachteile gegenüber:

- Die Reihenfolge der Spezialisierungen, die von den Lerndaten abhängt, entscheidet darüber, welcher Entscheidungsbaum entsteht. Zwei unterschiedliche Datensätze desselben Prozesses können so zu unterschiedlichen Entscheidungsbäumen führen.

- Bei gestörten und nicht eindeutigen Daten lernt der Entscheidungsbaum auch die Störung auswendig und ist nicht in der Lage, befriedigend zu generalisieren.
- Abgeleitete Regeln sind im Allgemeinen zu speziell, u. a. wegen der vergleichsweise geringeren Repräsentationsfähigkeit eines Entscheidungsbaums.

Um die Generalisierungsfähigkeit zu gewährleisten, wird bei vielen Verfahren die Entwicklung des Baumes vorzeitig gestoppt oder Spezialisierungen am vollständig entwickelten Baum werden zurückgenommen (geprunt). Allerdings kann dieser Vorgang eine nichtoptimale Entwicklung des Baums (Auswahl eines bestimmten Merkmals in einem Knoten) im Allgemeinen nicht beheben. Zudem können bestimmte Regelbasen nicht durch einen Baum repräsentiert werden, weil u. a. das im Wurzelknoten ausgewählte Merkmal in allen Regeln vorkommt. Aus diesem Grund bringt es Vorteile, nicht den Baum, sondern die aus ihm abgeleiteten Regeln zu generalisieren, was im folgenden Unterabschnitt gezeigt wird.

5.6.3 Regelpruning und Bewertung von Einzelregeln

Zur Generalisierung stehen zwei Optionen zur Verfügung

- Streichen eines Merkmals (einer linguistischen Variablen mit allen linguistischen Termen in der Prämisse) oder
- eine ODER-Verknüpfung (\cup) mit zusätzlichen, benachbarten Termen.

Beide Ansätze gehen über das Pruning des Baumes hinaus, das nur das Streichen des letzten, hinzugefügten Merkmals gestattet. Jede so entstehende Regel ist stets eine Verallgemeinerung der ursprünglichen Regel. Damit entstehen in jedem Pruningschritt neue Regeln $R_k^{j,i}$, wobei k die Nummer der ursprünglichen Regel und j die Nummer der Modifikation im i -ten Durchlauf des Prunings bezeichnen.

Das Prinzip des Regelprunings wird anhand der oben genannten Beispielregel WENN $x_{90} = A_{90,1} \cap x_{105} = A_{105,1}$ DANN $y = B_3$ demonstriert. Mögliche Pruningschritte sind folglich das Streichen des Ausdrucks mit x_{90} ($x_{90}=A_{90,1}$, neue Regelhypothese $R_1^{2,1}$), das Streichen von x_{105} ($R_1^{3,1}$), das Ergänzen der benachbarten Terme $x_{90} = A_{90,2}$ ($R_1^{4,1}$) bzw. $x_{105} = A_{105,2}$ ($R_1^{5,1}$) (Tabelle 5.7).

$R_1^{1,1}$:	WENN $x_{90}=A_{90,1} \cap x_{105}=A_{105,1}$	DANN $y=B_3$
$R_1^{2,1}$:	WENN $x_{105}=A_{105,1}$	DANN $y=B_3$
$R_1^{3,1}$:	WENN $x_{90}=A_{90,1}$	DANN $y=B_3$
$R_1^{4,1}$:	WENN $(x_{90}=A_{90,1} \cup A_{90,2}) \cap x_{105}=A_{105,1}$	DANN $y=B_3$
$R_1^{5,1}$:	WENN $x_{90}=A_{90,1} \cap (x_{105}=A_{105,1} \cup A_{105,2})$	DANN $y=B_3$

Tabelle 5.7: Hypothesen für Regelpruning – 1. Durchlauf

Zur Entscheidung, welche der genannten Regelhypotesen die beste Regel darstellt, wird ein Bewertungsmaß benötigt. Dessen modulare Bestandteile sind

- Klassifikationsgüte,
- Klarheit sowie
- statistische Absicherungsgüte

und werden im Folgenden eingeführt.

Die *Klassifikationsgüte*

$$Q_P = \left(1 - \frac{F_P}{F_P^0}\right) \quad (5.42)$$

basiert auf dem Vergleich des Klassifikationsfehlers einer Regel

$$F_P = \|\hat{\boldsymbol{\mu}}_Y - \boldsymbol{\mu}_Y\|_F = \|\mathbf{R}\mathbf{P} - \boldsymbol{\mu}_Y\|_F \rightarrow \underset{\mathbf{R}}{\text{Min}} \quad (5.43)$$

mit einer Trivialschätzung, die aus einer Regel mit immer erfüllter Prämisse (Zugehörigkeitsgrad der Prämisse ist immer Eins)

$$F_P^0 = \|\hat{\boldsymbol{\mu}}_Y - \boldsymbol{\mu}_Y\|_F = \|\mathbf{r}_0 \cdot \mathbf{1}_{1,N} - \boldsymbol{\mu}_Y\|_F \rightarrow \underset{\mathbf{r}_0}{\text{Min}} \quad (5.44)$$

hervorgeht. $\hat{\boldsymbol{\mu}}_Y$ ist dabei die Schätzung der Zugehörigkeiten der Ausgangsklassen, die aus der Regelmatrix \mathbf{R} und der Prämisenmatrix \mathbf{P} resultiert, welche mit den Zugehörigkeiten der Ausgangsklassen im Lerndatensatz $\boldsymbol{\mu}_Y$ verglichen wird. Die Klassifikationsgüte liegt somit für alle richtigen Regeln zwischen Null (keine Verbesserung im Vergleich zur Trivialschätzung) und Eins (erklärt jeweilige Ausgangsklasse vollständig). Für falsche Regeln hat die Klassifikationsgüte negative Werte. Damit werden wie beim ROSA-Verfahren [58] nur Besonderheiten gesucht, die von der durchschnittlichen Verteilung der Ausgangsklassen abweichen. Mit der Klassifikationsgüte wird indirekt die statistische Absicherung bewertet, da nur Regeln, die hinreichend durch Beispiele abgedeckt sind, eine wesentliche Verbesserung der Klassifikationsgüte bewirken können.

Das Hauptproblem besteht darin, dass die Klassifikationsgüte prinzipiell nur für vollständige Regelbasen geschätzt werden kann. Deswegen werden bei der Bewertung von Einzelregeln immer Gruppen von vier Regeln untersucht:

- WENN Prämisse P_k DANN Konklusion B_j
- WENN negierte Prämisse \bar{P}_k DANN Konklusion B_j
- WENN Prämisse P_k DANN negierte Konklusion \bar{B}_j
- WENN negierte Prämisse \bar{P}_k DANN negierte Konklusion \bar{B}_j

Die Regelplausibilitäten für diese vier Regeln stehen in den Elementen r_{11} , r_{12} , r_{21} und r_{22} der Regelmatrix und können als Maß für die Gültigkeit der jeweiligen Regel zwischen Null und Eins interpretiert werden. Die Regelmatrix \mathbf{R} stellt somit den Zusammenhang zwischen den Prämisen (Spalten von \mathbf{R}) und den verschiedenen Klassen in den Konklusionen der Regeln (Zeilen von \mathbf{R}) her⁷. Die Elemente dieser Matrix liegen im Intervall $[0, 1]$, die Spaltensumme ist stets Eins. Die letzte Regel mit negierter Prämisse und Konklusion wird in den weiteren Ausführungen als Defaultregel bezeichnet.

\mathbf{P} ist die Prämisenmatrix mit der Dimension $(2, N)$:

$$\mathbf{P} = \begin{pmatrix} \mu_{P_{11}} & \mu_{P_{12}} & \cdots & \mu_{P_{1,N}} \\ 1 - \mu_{P_{11}} & 1 - \mu_{P_{12}} & \cdots & 1 - \mu_{P_{1,N}} \end{pmatrix}. \quad (5.45)$$

Jedes Element $\mu_{P_{1,i}}$ erfordert die Prämisenbewertung der betreffenden Regel für das i -te Beispiel im Lerndatensatz, wobei die fuzzifizierten Werte der Merkmale in (5.3) auszuwerten sind. Für die Prämisse einer einzelnen Regel geht es in die ODER-Verknüpfung aller

⁷Deswegen kann die Regelmatrix auch als eine Matrix der bedingten Häufigkeiten $(\hat{p}(y|P_k) \quad \hat{p}(y|\bar{P}_k))$ für die Prämisse der k -ten Regel P_k interpretiert werden

Zugehörigkeitsgrade der linguistischen Terme einer linguistischen Variablen (Teilprämissen) mit dem Summen-Operator⁸ und die UND-Verknüpfung der Zugehörigkeitsgrade aller Teilprämissen mit dem Produkt-Operator über. Im Beispiel gilt also für die Regel $R_1^{1,1}$

$$\mu_{P_{1,i}} = \mu_{A_{90,1}}(x_{i,90}) \cdot \mu_{A_{105,1}}(x_{i,105}) \quad (5.46)$$

und für die Regel $R_1^{4,1}$

$$\mu_{P_{1,i}} = (\mu_{A_{90,1}}(x_{i,90}) + \mu_{A_{90,2}}(x_{i,90})) \cdot \mu_{A_{105,1}}(x_{i,105}). \quad (5.47)$$

Die Zugehörigkeitsgrade für die Prämisse der Defaultregel ergeben sich aus der Berechnungsvorschrift für die Fuzzy-Negation der Prämisse (Ergänzung zu Eins $1 - \mu_{P_{1,i}}$, vgl. (5.10)).

Die Berechnungsvorschrift umfasst somit die Prämissenauswertung bei der Berechnung von \mathbf{P} sowie Aktivierung und Akkumulation I bei der Berechnung von \mathbf{RP} .

Das zu lösende Optimierungsproblem für \mathbf{R} (5.43) entspricht strukturell dem Problem in (5.31) und erfordert Restriktionen entsprechend (5.32) und (5.33)⁹:

$$\mathbf{R} \geq \mathbf{0}_{2 \times 2} \quad (5.48)$$

$$\mathbf{1}_2^T \mathbf{R} = \mathbf{1}_2^T. \quad (5.49)$$

Damit gibt es keine negativen Regelplausibilitäten und alle Regelplausibilitäten (Spaltensummen) ergänzen sich zu Eins¹⁰. Eine ideale, eindeutige Regel führt auf eine Regelmatrix

$$\mathbf{R} = \begin{pmatrix} 1.00 & 0.00 \\ 0.00 & 1.00 \end{pmatrix} \quad (5.50)$$

und eine Klassifikationsgüte von Eins. Mit dem Gegenüberstellen von Konklusion und negierter Konklusion in (5.43) ist μ_Y somit die Matrix der fuzzifizierten Ausgangsgröße aller Lernbeispiele mit der Dimension $(N, 2)$ (2 Klassen, N Beispiele). Im Unterschied zu (5.15) existieren bei der Untersuchung von Einzelregeln also nur $m_y = 2$ Ausgangsklassen, die Konklusion der Regel (im Beispiel $y = B_3$) und die ODER-Verknüpfung aller anderen Konklusionen ($y = B_1 \cup B_2 \cup B_4 \cup B_5 \cup B_6 \cup B_7$). Da im Beispiel alle Schrittphasen scharf zugeordnet werden, enthält diese Matrix pro Spalte nur genau eine Eins und eine Null.

Bei der Trivialschätzung (5.44) reduziert sich die Matrix \mathbf{P} zu einem Einsvektor der Dimension $(1, n)$, der für eine stets erfüllte Prämisse steht. Die negierte Prämisse der Trivialschätzung hat somit immer die Zugehörigkeit Null und wird nicht benötigt. Für die Berechnung des Regelvektors der Trivialschätzung \mathbf{r}_0 in gelten entsprechend die Restriktionen

$$\mathbf{r}_0 \geq \mathbf{0}_{2 \times 1} \quad (5.51)$$

$$\mathbf{1}_2^T \mathbf{r}_0 = 1. \quad (5.52)$$

⁸Wegen der ausschließlichen ODER-Verknüpfung von linguistischen Termen $A_{l,i}$, deren Zugehörigkeitsfunktionen sich maximal zu Eins ergänzen können, entspricht die Summe hier der beschränkten Summe.

⁹Bei Verwendung einer scharfen (disjunkten) Klasseneinteilung für die Merkmale und die Ausgangsgröße geht das hier vorgestellte Konzept in die Schätzung der absoluten bzw. bedingten Klassenwahrscheinlichkeiten der Ausgangsgröße über.

¹⁰Durch die Restriktionen für die Spaltensumme hat \mathbf{R} nur zwei Freiheitsgrade.

Das erste Element des Vektors \mathbf{r}_0 steht dabei für die relative Häufigkeit der Konklusion B_j der Regel:

$$\mathbf{r}_0 = \begin{pmatrix} \hat{p}(y = B_j) \\ 1 - \hat{p}(y = B_j) \end{pmatrix}. \quad (5.53)$$

Die *Klarheit* Q_K einer Regel ist mit

$$Q_K = r_{11} \quad (5.54)$$

definiert¹¹. Die Klarheit ist um so größer, je geringer der Anteil von Gegenbeispielen in der Menge der abgedeckten Beispiele ist.

Die Klarheit der Defaultregel ist ein Maß für die Homogenität des übrig bleibenden Gebietes, das nicht durch die Regel abgedeckt wird:

$$Q_D = r_{22}. \quad (5.55)$$

Ideal ist eine Klarheit der Defaultregel von Eins. Dieser Fall deutet darauf hin, dass die untersuchte Regel im Wesentlichen alle Beispiele der Ausgangsklasse abdeckt, die in der Konklusion der untersuchten Regel steht.

Die bisher eingeführten Maße enthalten nur indirekte Informationen über die statistische Absicherung einer Regel. Zwar ist es naheliegend, dass klare Regeln mit einer hohen Klassifikationsgüte vermutlich auch statistisch relevant sind. Allerdings existieren dafür Gegenbeispiele, z. B. kann die Erklärung seltener Klassen mit wenigen Beispielen zu sehr guten Gütewerten von Q_P , aber auch zu einer geringen statistischen Sicherheit führen.

Deswegen ist es wichtig, mit der *statistischen Absicherungsgüte* Q_S ein Maß für die statistische Absicherung von Regeln bereitzustellen. Die Schwierigkeit besteht darin, dass für die hier untersuchten Aufgabenstellungen keine genau passenden statistischen Verfahren zur Verfügung stehen. Dennoch gibt es Parallelen zur Untersuchung von Binomialverteilungen, die für eine gegebene A-priori-Wahrscheinlichkeit $p(y = B_j | \mathbf{x} = P_k)$ eines Ereignisses angeben, wie wahrscheinlich verschiedene Werte der Anzahl erfolgreicher Versuche $n(\mathbf{x} = P_k \cap y = B_j)$ bei $n(\mathbf{x} = P_k)$ Versuchen sind. Ebenso kann die Frage untersucht werden, welche geschätzten Wahrscheinlichkeiten $\hat{p}(y = B_j | \mathbf{x} = P_k)$ bei einer bekannten Anzahl erfolgreicher Versuche $n(\mathbf{x} = P_k \cap y = B_j)$ bei $n(\mathbf{x} = P_k)$ Versuchen zu einem bestimmten statistischen Signifikanzniveau α in Frage kommen.

Der Schlüssel zur Untersuchung des Problems ist die Frage, ob sich die bedingte Wahrscheinlichkeit der Regelkonklusion bei erfüllter Prämisse $\hat{p}(y = B_j | \mathbf{x} = P_k)$ signifikant von der bedingten Wahrscheinlichkeit der Regelkonklusion bei nicht erfüllter Prämisse $\hat{p}(y = B_j | \mathbf{x} = \bar{P}_k)$ unterscheidet.

Wenn davon ausgegangen wird, dass diese statistische Schlussweise auch bei Fuzzy-Zugehörigkeiten noch brauchbare Ergebnisse liefert, steht nun mit den Elementen der \mathbf{R} -Matrizen ein Äquivalent zu $\hat{p}(y = B_j | \mathbf{x} = P_k)$ zur Verfügung. Das Element r_{11} weist Parallelen zu einer bedingten Wahrscheinlichkeit $\hat{p}(y = B_j | \mathbf{x} = P_k)$ auf. Die Komplementärprämisse liefert das entsprechende Element r_{12} , das zu $\hat{p}(y = B_j | \mathbf{x} = \bar{P}_k)$ gehört. Wenn nun – wie oben diskutiert – unterstellt wird, dass beide Größen binomialverteilt sind, können untere und obere Abschätzungen für beide Größen eingeführt werden, die

¹¹Diese Definition ist ein Sonderfall der in [81] definierten Variante $\|\mathbf{R}_{\bullet j}\|_\infty = \max_{i=1, \dots, n} r_{ij}$, wobei $\mathbf{R}_{\bullet j}$ den j -ten Spaltenvektor von \mathbf{R} darstellt, die Regelmatrizen mit mehreren Regeln und m_y möglichen Konklusionen zulässt.

ein Vertrauensintervall kennzeichnen. Damit gilt für die wahren Parameter

$$r_{11}^l \leq r_{11} \leq r_{11}^u \quad (5.56)$$

$$r_{12}^l \leq r_{12} \leq r_{12}^u, \quad (5.57)$$

wobei l (lower) die untere und u (upper) die obere Grenze des Vertrauensintervalls bezeichnen. Die Schätzungen für

$$r_{11} = \hat{p}(y = B_j | \mathbf{x} = P_k) \quad (5.58)$$

$$r_{12} = \hat{p}(y = B_j | \mathbf{x} = \bar{P}_k) \quad (5.59)$$

mit (5.28) und für die Anzahl der Versuche

$$n(\mathbf{x} = P_k) = \sum_{i=1}^N \mu_{P_{1i}} = \sum_{i=1}^N \mu_P(\mathbf{x}_i = P_k) \quad (5.60)$$

$$n(\mathbf{x} = \bar{P}_k) = \sum_{i=1}^N \mu_{P_{2i}} = \sum_{i=1}^N 1 - \mu_P(\mathbf{x}_i = P_k) \quad (5.61)$$

gestatten eine Anwendung der Abschätzungen in (A.1) und (A.2) im Anhang A.6 [34]. Entsprechend der Schlussweise beim Relevanzmaß nach [58, 61] ist eine Regel statistisch abgesichert, wenn $r_{12}^u < r_{11}^l$ gilt (Bild 5.11).

Wenn hingegen $r_{11}^u < r_{12}^l$ erfüllt ist, handelt es sich um eine sogenannte negative Regel, die für die Prämisse P_k auf die negierte Konklusion führt. Negative Regeln [58] erlauben für einige Aufgabenstellungen eine kompaktere Formulierung der Regelbasis, erfordern allerdings eine aufwendigere Inferenz und Defuzzifizierung (Hyperinferenz und Hyperdefuzzifizierung). Deswegen wird dieser Ansatz hier nicht weiter verfolgt.

In [58, 61] findet allerdings ein Vergleich zwischen $\hat{p}(y = B_j)$ und $\hat{p}(y = B_j | \mathbf{x} = P_k)$ statt, während in dieser Arbeit $\hat{p}(y = B_j | \mathbf{x} = P_k)$ und $\hat{p}(y = B_j | \mathbf{x} = \bar{P}_k)$ gegenüber gestellt werden. Eine ausführliche Darstellung und Diskussion dieser Maße findet sich in [61]. Weiterentwicklungen dieser Maße enthält [52]. Die Modifikationen bei der Schätzung relativer Häufigkeiten für den Fall, dass den linguistischen Termen Fuzzy-Mengen zugeordnet sind, können zur Berechnung von $\hat{p}(y = B_j | \mathbf{x} = P_k)$ und $\hat{p}(y = B_j | \mathbf{x} = \bar{P}_k)$ verwendet werden, allerdings geht dabei die unmittelbare statistische Bedeutung verloren.

Neben der Ja-Nein-Frage nach der statistischen Absicherung interessiert auch die Frage nach der quantitativen Sicherheit dieser Entscheidung. Eine statistische Entscheidung ist um so sicherer, je schmaler die beiden Vertrauensintervalle im Vergleich zu ihrem Abstand untereinander sind. Die Idee besteht folglich darin, mit einem Faktor K beide Vertrauensintervalle multiplikativ so zu vergrößern, dass sie sich berühren. Für eine relevante Einzelregel (Fall 1) bedeutet das

$$r_{11} - K(r_{11} - r_{11}^l) = r_{12} + K(r_{12}^u - r_{12}) \quad (5.62)$$

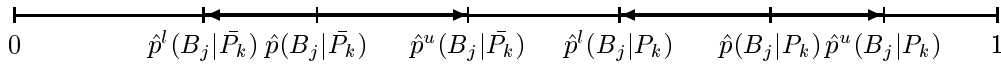
und für eine relevante negative Einzelregel (Fall 2)

$$r_{12} - K(r_{12} - r_{12}^l) = r_{11} + K(r_{11}^u - r_{11}). \quad (5.63)$$

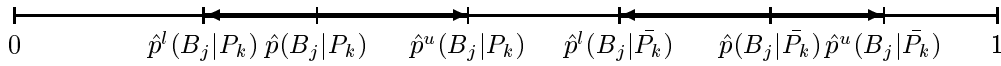
Damit gilt

$$K = \begin{cases} \frac{r_{11} - r_{12}}{r_{12}^u - r_{12} + r_{11} - r_{11}^l} & \text{wenn } r_{11} > r_{12} \\ \frac{r_{12} - r_{11}}{r_{11}^u - r_{11} + r_{12} - r_{12}^l} & \text{wenn } r_{11} \leq r_{12}. \end{cases} \quad (5.64)$$

Fall 1: relevante Einzelregel



Fall 2: relevante negative Einzelregel



Fall 3: irrelevante Regel

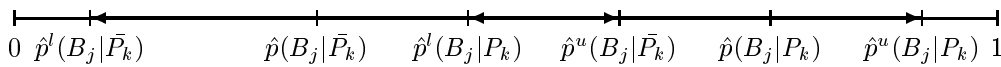


Bild 5.11: Vertrauensintervalle der geschätzten Wahrscheinlichkeiten zur Berechnung des Relevanzindex; abkürzende Schreibweise $\hat{p}(B_j|\bar{P}_k)$ statt $\hat{p}(y = B_j|\mathbf{x} = \bar{P}_k)$, $\hat{p}(B_j|P_k)$ statt $\hat{p}(y = B_j|\mathbf{x} = P_k)$ usw.

Bei einem Wert $K < 1$ ist die Regel statistisch nicht abgesichert. Je größer K wird, desto größer ist die statistische Sicherheit. Eine Metrik zur Formulierung dieses Problems führt auf die statistische Absicherungsgüte

$$Q_S = \begin{cases} 1 - \frac{1}{\max(1, K)} & \text{wenn } r_{11} > r_{12} \\ -1 + \frac{1}{\max(1, K)} & \text{wenn } r_{11} \leq r_{12}. \end{cases} \quad (5.65)$$

Diese statistische Absicherungsgüte darf aber nicht mit einem Signifikanzniveau α verwechselt werden. Bei einer Einzelregel bedeutet jeder Wert $Q_S > 0$, dass die Regel über dem vorgegebenen Signifikanzniveau liegt. Je näher dieser Wert in Richtung Eins geht, desto sicherer ist die Regel.

Mit diesen modularen Maßen Klassifikationsgüte, Klarheit und statistische Absicherungsgüte können nun Einzelregeln auf ihre Relevanz geprüft werden. Dazu dient hier das Kriterium [81]

$$Q = Q_P \cdot Q_K^\beta \quad (\beta > 0) \quad (5.66)$$

mit der Klassifikationsgüte Q_P und der Klarheit Q_K .¹² Der Exponent β erlaubt eine Gewichtung zwischen Klassifikationsgüte und Klarheit (hier: $\beta = 10$). Größere Werten von β führen zu einer Bevorzugung speziellerer Regeln mit wenigen Fehlern, weniger Beispielen und einer tendenziell geringeren statistischen Absicherung (niedrigeres Abstraktionsniveau) gegenüber allgemeineren Regeln mit einigen Fehlern, vielen Beispielen und einer höheren statistischen Absicherung (höheres Abstraktionsniveau).

Bei der Regelbewertung der Regel $R_1^{1,1}$ im Beispiel ergibt sich eine Regelbasis

¹²Während des Regelprunings ist es vorteilhaft, die statistische Absicherungsgüte erst am Ende des Pruningprozesses einzubeziehen, weil sonst viel versprechende spezielle Regeln nicht weiterverfolgt werden. Bei einer abschließenden Bewertung sind dann auch modifizierte Maße, z. B. $Q = Q_P \cdot Q_K^\beta \cdot Q_S$ sinnvoll.

WENN $x_{90}=A_{90,1} \cap x_{105}=A_{105,1}$	DANN $y=B_3$
WENN $x_{90}=\text{NICHT } A_{90,1} \cup x_{105}=\text{NICHT } A_{105,1}$	DANN $y=B_3$
WENN $x_{90}=A_{90,1} \cap x_{105}=A_{105,1}$	DANN $y=\text{NICHT } B_3$
WENN $x_{90}=\text{NICHT } A_{90,1} \cup x_{105}=\text{NICHT } A_{105,1}$	DANN $y=\text{NICHT } B_3$.

Die Trivialschätzung für r_0 spiegelt die relative Häufigkeit von $y = B_3$ im Lerndatensatz wider und liefert

$$r_0 = \begin{pmatrix} 0.20 \\ 0.80 \end{pmatrix} \quad (5.67)$$

mit einem Klassifikationsfehler von $F_P^0 = 19.23$. Die Regelmatrix \mathbf{R} für Regel $R_1^{1,1}$ lautet

$$\mathbf{R} = \begin{pmatrix} 1.00 & 0.17 \\ 0.00 & 0.83 \end{pmatrix}. \quad (5.68)$$

Damit hat die untersuchte Regel eindeutig die Konklusion B_3 Terminal Stance, was eine Klarheit von $Q_K = 1.00$ zeigt. Der Klassifikationsfehler reduziert sich auf $F_P=17.63$, was eine Klassifikationsgüte von $Q_P = 0.08$ zur Folge hat. Die Interpretation der Klassifikationsgüte als relative Reduzierung des Fehlers zwischen Null und Eins (hier in etwa: 'Die Regel erklärt 8% der Beispiele der Ausgangsklasse.') ist dabei anschaulicher als der Wertevergleich der Klassifikationsfehler und trägt folglich zur verbesserten Interpretierbarkeit der Ergebnisse bei.

Die optimale Defaultregel führt auf eine Entscheidung zu 0.17 für Terminal Stance und zu 0.83 für eine andere Schrittphase (Nicht Terminal Stance!), wobei die Klarheit der Defaultregel $Q_D = 0.83$ ist. Der Wert von 0.17 bedeutet, dass noch ein beträchtlicher Anteil von Beispielen mit Terminal Stance außerhalb des Einzugsbereichs der Regelprämisse liegt. Die statistische Abschätzung ergibt Intervallgrenzen von

$$0.74 \leq r_{11} \leq 1.00 \quad (5.69)$$

$$0.15 \leq r_{12} \leq 0.19, \quad (5.70)$$

was auf $K = 2.77$ und eine statistische Absicherungsgüte von $Q_S = 0.64$ führt.

Regel-Nr.	Q	Q_P	Q_K	Q_S	Fehler	Beispiele
$R_1^{1,1}$	0.08	0.08	1.00	0.64	0	11
$R_1^{2,1}$	0.09	0.09	1.00	0.67	0	12
$R_1^{3,1}$	0.00	0.08	0.43	0.86	285	455
$R_1^{4,1}$	0.09	0.09	1.00	0.67	0	12
$R_1^{5,1}$	0.38	0.38	1.00	0.97	11	166

Tabelle 5.8: Bewertung Regelpruning – 1. Durchlauf (Hypothesen siehe Tabelle 5.7)

Mit den vorgestellten Maßen können nun alle Regelhypothesen (Tabelle 5.7) für den 1. Pruningdurchlauf bewertet werden (Tabelle 5.8).¹³ Die beste Regel, die nach dem Kriterium (5.42) gefunden wurde, ist die Regel $R_1^{5,1}$, die durch Ergänzen eines Terms bei x_{105} hervorgegangen ist (Bild 5.12). Deren Klarheit ist ebenfalls Eins, aber die Klassifikationsgüte (0.38) und die Klarheit der Defaultregel (0.98) erhöhen sich deutlich im Vergleich

¹³Geringfügig kleinere Beispiellanzahlen bei Regeln im Vergleich zu Entscheidungsbäumen (11 statt 12 Beispiele) sind eine Folge unterschiedlicher Berechnungsmethoden. Während bei Entscheidungsbäumen auf jeder Hierarchieebene scharfe Entscheidungen getroffen werden, erfolgt bei Regeln die Prämissenauswertung mit den Operatoren Summe und Produkt. Ein Beispiel wird nur gewertet, wenn die Prämisse einen Zugehörigkeitsgrad größer gleich 0.5 aufweist.

zu $R_1^{1,1}$. Die Regel produziert zwar 11 Fehler, ist aber insgesamt durch 166 Beispiele abgesichert. Auffällig ist, dass alle Regeln statistisch abgesichert sind ($Q_S > 0$) und dass die ausgewählte Regel mit $Q_S = 0.97$ die beste statistische Absicherung aufweist.

WENN R-HIP Flex ANG V=NEG UND L-KNEE Flex ANG V=NG DANN y=TST-R

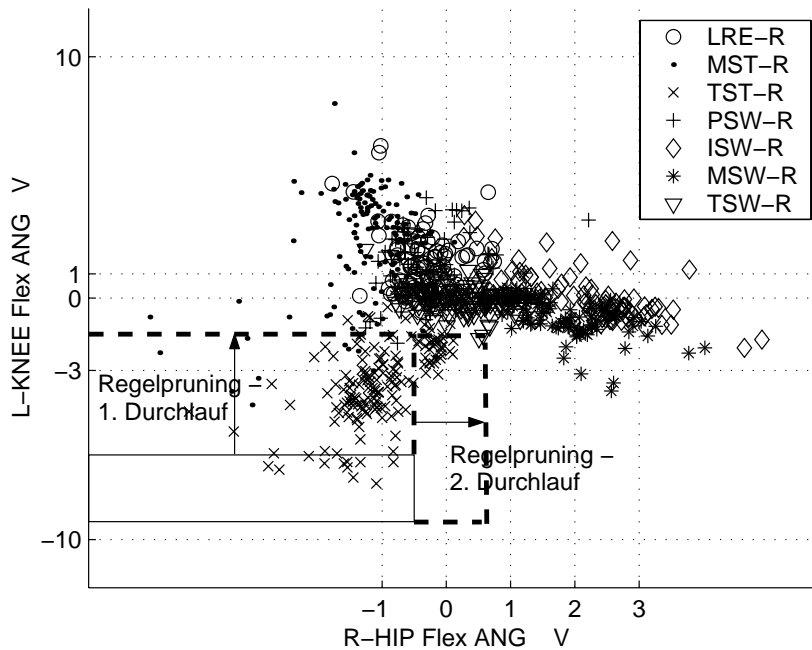


Bild 5.12: Visualisierung der Fuzzy-Regeln $R_1^{1,1}$ (durchgezogene Linie), $R_1^{5,1}$ (Bereich der Regel $R_1^{1,1}$ und Bereich der Regel $R_1^{5,1}$ nach dem 1. Durchlauf des Regelprunings) und $R_1^{4,2}$ (Bereich der Regel $R_1^{5,1}$ und Bereich fette, gepunktete Linie nach dem 2. Durchlauf des Regelprunings) zur Bestimmung der Schrittphase Terminal Stance mit den Merkmalen x_{90} (R-HIP Flex Ang V) und x_{105} (L-KNEE Flex Ang V)

Diese Regel bildet nun als neue Regel $R_1^{1,2}$ den Ausgangspunkt für den zweiten Durchlauf des Prunings, der in Tabelle 5.9 angegeben ist. Mögliche Pruningschritte sind wiederum das Streichen der Teilprämissen mit den Merkmalen x_{90} bzw. x_{105} oder das Ergänzen benachbarter linguistischer Terme. Mit der Regel $R_1^{4,2}$ wird erneut eine verbesserte Regel mit erhöhter Klassifikationsgüte (0.52), gleich bleibender Klarheit (1.00), verbesserter statistischer Absicherung (0.98) und verbesserter Klarheit der Defaultregel (0.99) gefunden (Bild 5.12). Durch das Ergänzen von $A_{90,2}$ wird das Einzugsgebiet auf 209 Beispiele erhöht, wobei sich die Fehleranzahl unwesentlich auf 13 vergrößert.

Im 3. Durchlauf (Tabelle 5.10) gelingt keine weitere Verbesserung, weil hier bei allen Generalisierungen zu viele Fehler entstehen und sich so Klassifikationsgüte und Klarheit verringern. Damit ist das Pruning der untersuchten Regel abgeschlossen.

Am Ende des Prunings wird überprüft, ob die jeweilige Regel bestimmte Mindestanforderungen bezüglich der erreichten Klarheit erfüllt und statistisch abgesichert ist ($Q_S > 0$). Wenn das nicht der Fall ist, wird die beste Regel übernommen, die während des Prunings der Regel die Mindestanforderungen erreicht hat. Wenn keine solche Regel existiert, wird die Regel gelöscht.

Außerdem kann für die beste Regel optional getestet werden, ob eine weitere Generalisierung zwar zu einer leichten Verschlechterung der Güte führt, aber der Generalisie-

$R_1^{1,2}$: WENN $x_{90}=A_{90,1} \cap (x_{105}=A_{105,1} \cup A_{105,2})$	DANN $y=B_3$
$R_1^{2,2}$: WENN $(x_{105}=A_{105,1} \cup A_{105,2})$	DANN $y=B_3$
$R_1^{3,2}$: WENN $x_{90}=A_{90,1}$	DANN $y=B_3$
$R_1^{4,2}$: WENN $(x_{90}=A_{90,1} \cup A_{90,2}) \cap (x_{105}=A_{105,1} \cup A_{105,2})$	DANN $y=B_3$
$R_1^{5,2}$: WENN $x_{90}=A_{90,1} \cap (x_{105}=A_{105,1} \cup A_{105,2} \cup A_{105,3})$	DANN $y=B_3$

Regel-Nr.	Q	Q_P	Q_K	Q_S	Fehler	Beispiele
$R_1^{1,2}$	0.38	0.38	1.00	0.97	11	166
$R_1^{2,2}$	0.13	0.40	0.89	0.96	33	230
$R_1^{3,2}$	0.00	0.08	0.43	0.86	285	455
$R_1^{4,2}$	0.52	0.52	1.00	0.98	13	209
$R_1^{5,2}$	0.03	0.29	0.81	0.93	62	232

Tabelle 5.9: Regelpruning – 2. Durchlauf

$R_1^{1,3}$: WENN $(x_{90}=A_{90,1} \cup A_{90,2}) \cap (x_{105}=A_{105,1} \cup A_{105,2})$	DANN $y=B_3$
$R_1^{2,3}$: WENN $(x_{105}=A_{105,1} \cup A_{105,2})$	DANN $y=B_3$
$R_1^{3,3}$: WENN $(x_{90}=A_{90,1} \cup A_{90,2})$	DANN $y=B_3$
$R_1^{4,3}$: WENN $(x_{90}=A_{90,1} \cup A_{90,2} \cup A_{90,3}) \cap (x_{105}=A_{105,1} \cup A_{105,2})$	DANN $y=B_3$
$R_1^{5,3}$: WENN $(x_{90}=A_{90,1} \cup A_{90,2}) \cap (x_{105}=A_{105,1} \cup A_{105,2} \cup A_{105,3})$	DANN $y=B_3$

Regel-Nr.	Q	Q_P	Q_K	Q_S	Fehler	Beispiele
$R_1^{1,3}$	0.52	0.52	1.00	0.98	13	209
$R_1^{2,3}$	0.13	0.40	0.89	0.96	33	230
$R_1^{3,3}$	0.00	0.07	0.31	0.86	556	788
$R_1^{4,3}$	0.47	0.50	0.99	0.98	15	212
$R_1^{5,3}$	0.00	0.26	0.58	0.92	207	434

Tabelle 5.10: Regelpruning – 3. Durchlauf

rungsschritt durch das Streichen eines Merkmals statistisch nicht ausgeschlossen werden kann. Ein typischer Fall dafür ist, dass nur eine sehr geringe Anzahl von Gegenbeispielen (z. B. ein oder zwei Fehler) eine Generalisierung verhindert, aber diese Gegenbeispiele keine Ablehnung der Generalisierung rechtfertigen. Eine Möglichkeit, diesen Fall statistisch zu überprüfen, ist ein Test der drei Gebiete

1. Regelprämisse der nicht generalisierten Regel,
2. neu hinzukommendes Gebiet durch Generalisierung und
3. komplementäre Regelprämisse der generalisierten Regel

auf statisch relevante Differenzen zwischen dem ersten und zweiten Gebiet analog zu (5.57-5.64).

Der vollständige Ablauf dieser Tests ist in Bild 5.13 dargestellt.

Der gleiche Pruningablauf wird mit allen 220 Regeln absolviert, die von den Entscheidungsbaum generiert werden. Da abschließend redundante Regeln entfernt werden (gleiche Regeln oder Spezialisierungen allgemeinerer Regeln mit niedrigerer Güte Q), bleiben nach dem Generalisieren noch 107 Regeln übrig (siehe Anhang A.7).

Eine weitere relativ einfache Regel, die Regel R_{10}

$$\text{WENN } (x_3=\text{PM} \cup \text{PG}) \cap x_6=\text{PSG}$$

$$\text{DANN } y=\text{MSW-R}$$

zur Beschreibung der mittleren Schwungphase, ist in Bild 5.14 dargestellt. Sie baut auf den

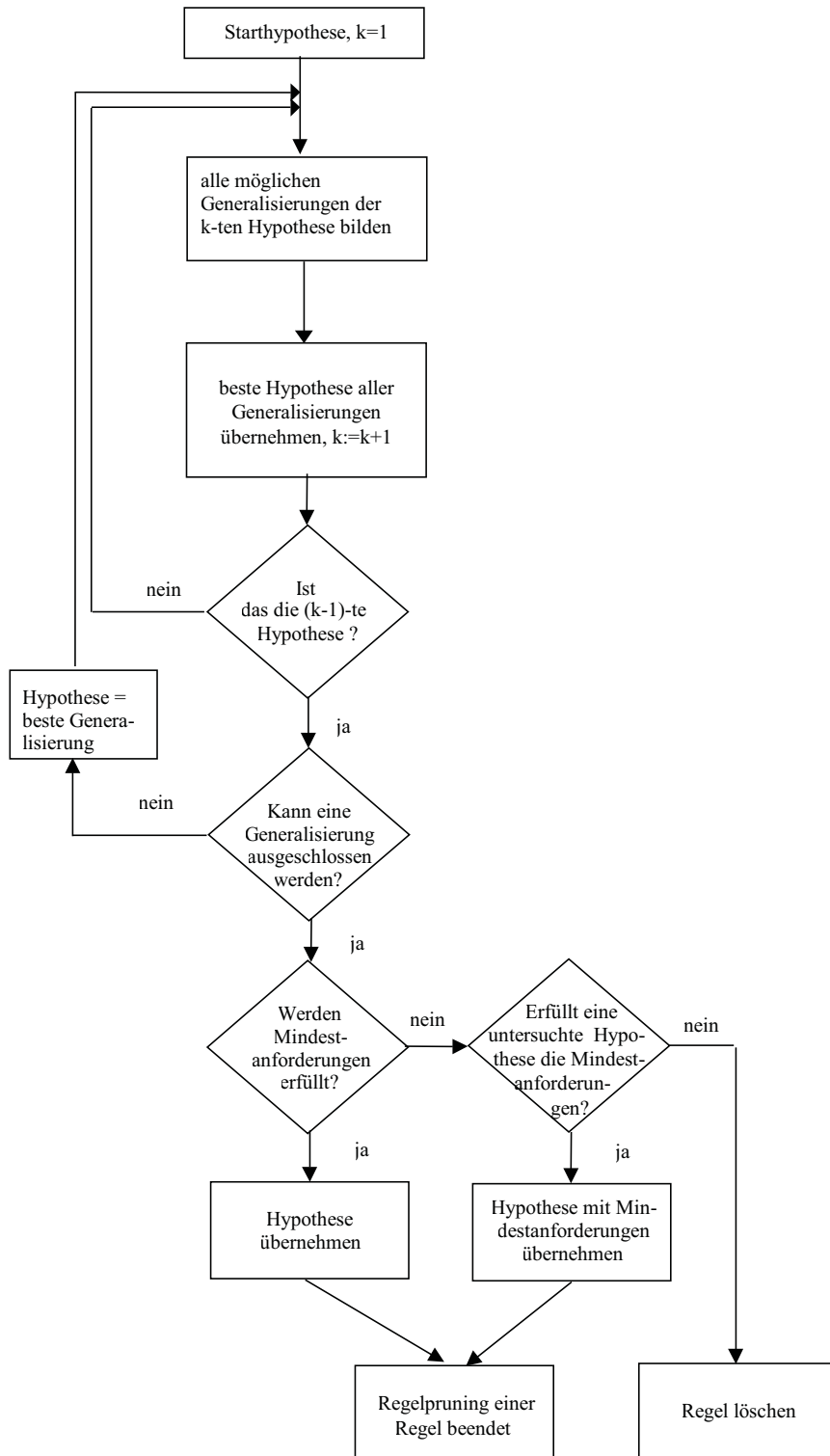


Bild 5.13: Ablauf des Prunings für eine Regel

Merkmalen x_3 (R-HIP Flex ANG, Hüftbeugung – Winkel zwischen rechtem Oberschenkel und Becken) und x_6 (R-KNEE Flex ANG, Beugung rechtes Knie) auf. In der mittleren Schwungphase ist dabei das Knie maximal angewinkelt (Term PSG) und die Hüftbeugung ist mittel (PM) bis groß (PG). Die genannte Phase ist in Bild 5.14 durch * markiert, das Gebiet der aktivierten Regel (Zugehörigkeitsgrad der Prämisse > 0.5 unter Vernachlässigung der Effekte durch den Summen- und den Produkt-Operator) ist eingerahmt. Die Regel erfasst einen Großteil der Messungen, die zur genannten Phase gehören. Die Parameter der Zugehörigkeitsfunktionen sind durch gepunktete Linien dargestellt. Während einige Parameter (10 und 20 für x_3) aus Sicht der Regel gut gewählt sind, lassen andere noch Reserven offen: Eine Verschiebung des Parameters 40 auf 50 für x_3 eröffnet die Möglichkeit, bisher nicht erfasste Datensätze mit in die Regel einzubeziehen. Eine Verschiebung des Parameters 40 auf ca. 45 für x_6 reduziert hingegen die Fehleranzahl zur Klasse TSW-R. Bei beiden Parametern ist aber zu berücksichtigen, dass sie einen Kompromiss für die Abgrenzung aller möglicher Klassen auf der Grundlage des Merkmals, nicht der Regel darstellen.

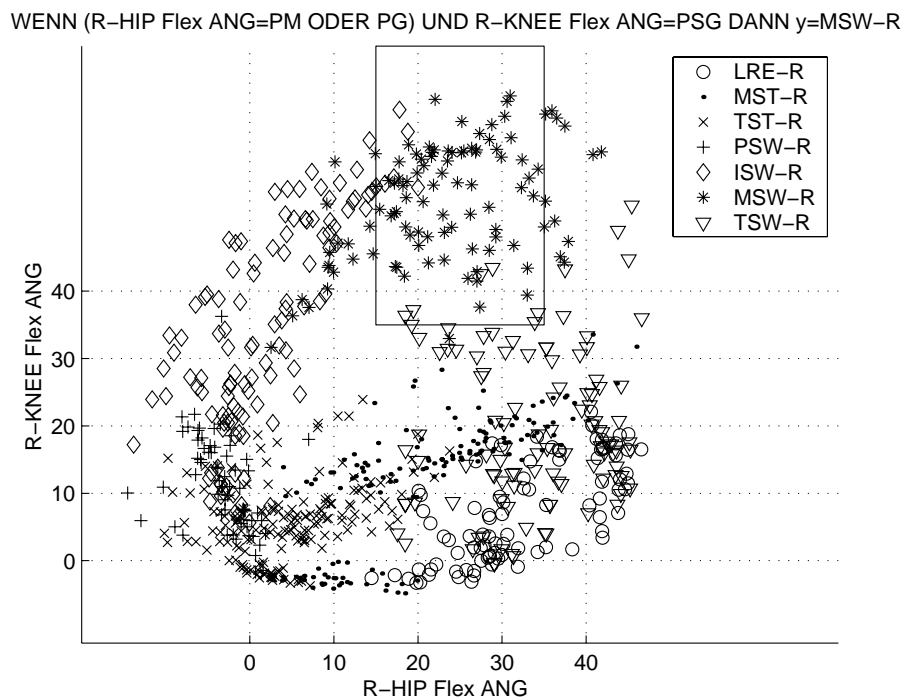


Bild 5.14: Visualisierung der Fuzzy-Regel R_{10} zur Bestimmung der Schrittphase Mid Swing mit den Merkmalen x_3 (R-HIP Flex Ang) und x_6 (R-KNEE Flex Ang)

Diese 107 Regeln können prinzipiell direkt zur Klassifikation der Schrittphasen verwendet werden, indem jeder Regel eine bestimmte Relevanz zur Auflösung von Konflikten konkurrierender Regeln zugewiesen wird. Allerdings ist diese große Regelbasis noch sehr unübersichtlich. Eine bessere Variante ist es deshalb, eine möglichst kleine Regelbasis gut zusammen passender Regeln auszuwählen, was im nächsten Abschnitt demonstriert wird.

5.7 Suche nach Regelbasen

Die Problematik der Auswahl von Regeln hängt wiederum eng mit den hier verwendeten Methoden zur Generalisierung zusammen. Generalisierte Regeln, die aus unterschiedlichen speziellen Regeln des gleichen oder ähnlicher Entscheidungsbäume hervorgehen, unterscheiden sich häufig nur in den linguistischen Termen der gleichen linguistischen Variablen (des gleichen Merkmals) oder in jeweils einer anderen linguistischen Variablen.

Ein typisches Beispiel für den zweiten Fall sind die Regeln

$$\begin{aligned}
 R_{22}: & \quad \text{WENN R-HIP Flex ANG=ZE} \cap (\text{R-HIP Flex ANG V=NEG} \cup \text{ZE}) \quad \text{DANN } y=\text{PSW-R} \\
 & \quad \cap \text{L-KNEE Flex ANG V=PK} \\
 R_{23}: & \quad \text{WENN R-HIP Flex ANG=ZE} \cap (\text{L-KNEE Flex ANG=ZE} \cup \text{PK}) \cap \quad \text{DANN } y=\text{PSW-R}, \\
 & \quad \text{L-KNEE Flex ANG V=PK}
 \end{aligned}$$

die zum Großteil die gleichen Beispiele abdecken.

Die wichtigsten Aufgaben der Regelauswahl bestehen darin, das Zusammenwirken der Regelbasis zu betrachten und partiell redundante Regeln zu reduzieren. Deswegen werden in einer Regelbasis gute Einzelregeln gesucht, die miteinander kooperieren und sich ergänzen. Dazu werden auf heuristischem Weg Hypothesen für aufzunehmende Regeln aufgestellt und ausgewertet.

Beginnend mit der besten Regel wird die Regelbasis sukzessive um eine weitere Regel ergänzt, die in ihrem Zusammenwirken die beste Güte bewirken.

Die Regelbasis ist dabei im Allgemeinen unvollständig und wird durch eine Defaultregel mit komplementärer Prämisse ergänzt (SONST $y = B_j$). Als deren Konklusion B_j wird entweder

- der Term B_j (bzw. die Klasse j) verwendet, der im Lerndatensatz im Einzugsbereich der komplementären Prämisse P_r am häufigsten vorkommt,
- ein fest vorgegebener Wert eingesetzt oder
- eine Rückweisungsklasse $m_y + 1$ eingeführt.

Welche Variante erfolgreicher ist, hängt sehr stark von der Topologie der Lerndaten ab. Besonders kompakte Regelbasen entstehen bei der ersten Variante dann, wenn nur Spezialfälle zu bestimmten Klassen abgebildet werden und große, nicht zusammenhängende Gebiete durch die Defaultregel abgedeckt werden. Die dritte Variante wird hier bevorzugt, weil so alle Klassen mit Regeln und nicht über eine Defaultregel erklärt werden.

Die Klassifikationsgüte (5.42) beruht nun jeweils auf allen Regeln der Regelbasis und der Defaultregel, wobei bei der Berechnung von F_P die Matrix \mathbf{R} durch die Konklusionen der Regeln in der Regelbasis B_{Rk} jetzt fest vorgegeben ist:

$$r_{ik} = \begin{cases} 1 & \text{für } B_{Rk} = B_i \\ 0 & \text{sonst.} \end{cases} \quad (5.71)$$

μ_Y ist dabei die Matrix der fuzzifizierten Ausgangsgröße aller Lernbeispiele mit der Dimension (m_y, N) (m_y Klassen, N Beispiele) oder $(m_y + 1, N)$ (m_y Klassen + Rückweisungsklasse, N Beispiele) – im Unterschied zu (5.43) mit nur zwei Ausgangsklassen. Deswegen sind die Zahlenwerte der Klassifikationsgüte für Regelbasen auch nicht mit den Zahlenwerten bei Einzelregeln zu vergleichen.

Wesentlich problematischer als bei der Untersuchung von Einzelregeln ist die Aufstellung der Prämissenmatrix \mathbf{P} , weil neben der Prämissenbewertung der einzelnen Regeln auch deren Überdeckungen berücksichtigt werden müssen. Diese Überdeckungen können sowohl zu widerspruchsfreien (sich überdeckende Regeln mit gleichen Konklusionen) oder

lokal widersprüchlichen (sich überdeckende Regeln mit verschiedenen Konklusionen) Regelbasen führen. Wenn diese Überdeckungen vernachlässigt werden, führt das u. U. zu irreführenden Ergebnissen. Deshalb werden zwei Matrizen aufgestellt, eine erweiterte Prämissenmatrix \mathbf{P}' und eine Redundanzmatrix \mathbf{V} .

Überdeckungen von Regeln können daran erkannt werden, dass die UND-Verknüpfungen der Prämissen mehrerer Regeln keine leeren Mengen sind. Ist das der Fall, müssen alle möglichen UND-Verknüpfungen zwischen den Regelprämissen mit berücksichtigt werden:

$$\mathbf{P}' = [\boldsymbol{\mu}_{P_1} \cdots \boldsymbol{\mu}_{P_r} \boldsymbol{\mu}_{P_1 \cap P_2} \cdots \boldsymbol{\mu}_{P_1 \cap P_r} \cdots \boldsymbol{\mu}_{P_{r-1} \cap P_r} \cdots \boldsymbol{\mu}_{P_1 \cap P_2 \cap P_3} \cdots]^T. \quad (5.72)$$

Die Vektoren $\boldsymbol{\mu}_{P_k}$ enthalten dabei für alle N Beispiele die Prämissenauswertung der k -ten Regel. Außerdem werden Prämissen von virtuellen Regeln eingeführt, die aus einer UND-Verknüpfung von Regelprämissen ($P_1 \cap P_2$ usw.) resultieren.¹⁴

Die Funktion der Redundanzmatrix \mathbf{V} besteht darin, die Aktivierungen der Regeln mit (partiell) redundanten Prämissen zu modifizieren. Die ersten r Spalten von \mathbf{V} bilden eine Einheitsmatrix. Die folgenden Spalten werden mit den Zugehörigkeitsgraden der Schnittmengen verknüpft. Das Vorzeichen berechnet sich zu $(-1)^{l-1}$, wobei l die Anzahl geschnittener Prämissen ist. Alle Spaltenelemente, die zu den am Schnitt beteiligten Prämissen gehören, erhalten den mit dem entsprechenden Vorzeichen versehenen Wert $1/l$:

$$v_{ij} = \begin{cases} (-1)^{(l-1)} \frac{1}{l} & \text{für } i \in \mathcal{I}_j \\ 0 & \text{sonst.} \end{cases} \quad (5.73)$$

Die Indexmenge \mathcal{I}_j enthält dabei die Indizes aller l Regeln, die im zugehörigen j -ten Element von $\boldsymbol{\mu}_{P'}$ UND-verknüpft sind.

Bei drei Regeln folgt so

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 & -1/2 & -1/2 & 0 & 1/3 \\ 0 & 1 & 0 & -1/2 & 0 & -1/2 & 1/3 \\ 0 & 0 & 1 & 0 & -1/2 & -1/2 & 1/3 \end{pmatrix} \quad (5.74)$$

und

$$\mathbf{P}' = \begin{pmatrix} \boldsymbol{\mu}_{P_{11}} & \boldsymbol{\mu}_{P_{12}} & \cdots & \boldsymbol{\mu}_{P_{1N}} \\ \boldsymbol{\mu}_{P_{21}} & \boldsymbol{\mu}_{P_{22}} & \cdots & \boldsymbol{\mu}_{P_{2N}} \\ \boldsymbol{\mu}_{P_{31}} & \boldsymbol{\mu}_{P_{32}} & \cdots & \boldsymbol{\mu}_{P_{3N}} \\ \boldsymbol{\mu}_{P_{11} \cap P_{21}} & \boldsymbol{\mu}_{P_{12} \cap P_{22}} & \cdots & \boldsymbol{\mu}_{P_{1N} \cap P_{2N}} \\ \boldsymbol{\mu}_{P_{11} \cap P_{31}} & \boldsymbol{\mu}_{P_{12} \cap P_{32}} & \cdots & \boldsymbol{\mu}_{P_{1N} \cap P_{3N}} \\ \boldsymbol{\mu}_{P_{21} \cap P_{31}} & \boldsymbol{\mu}_{P_{22} \cap P_{32}} & \cdots & \boldsymbol{\mu}_{P_{2N} \cap P_{3N}} \\ \boldsymbol{\mu}_{P_{11} \cap P_{21} \cap P_{31}} & \boldsymbol{\mu}_{P_{12} \cap P_{22} \cap P_{32}} & \cdots & \boldsymbol{\mu}_{P_{1N} \cap P_{2N} \cap P_{3N}} \end{pmatrix}^T. \quad (5.75)$$

Nach Ergänzung der Defaultregel kann nun \mathbf{P} berechnet werden:

$$\mathbf{P} = \begin{pmatrix} \mathbf{V} \cdot \mathbf{P}' \\ \mathbf{1}_N^T - \mathbf{1}_{m_y}^T \mathbf{V} \cdot \mathbf{P}' \end{pmatrix}. \quad (5.76)$$

Die so gefundene Hypothese für eine Regelbasis wird mit dem Klassifikationsfehler F_P^0 einer Regelbasis verglichen, die nur eine Defaultregel mit immer gültiger Prämisse (WENN 1

¹⁴Diese UND-Verknüpfung ist aussagenlogisch mit den linguistischen Termen unter Berücksichtigung ihrer Zusammenhänge und *nicht* aus einer UND-Verknüpfung von Zugehörigkeitsgraden mit feststehenden t -Normen zu berechnen, eine ausführliche Diskussion kann [79] entnommen werden.

DANN ...) enthält. Diese Regelbasis stellt eine Trivialschätzung dar und spiegelt je nach Verfahren die häufigste Ausgangsklasse B_j oder die Rückweisungsklasse B_{m_y+1} wider.

Alle Regelbasen gestatten nun eine Aussage über die relative Verbesserung zu diesem Klassifikationsfehler. Die Suche beginnt bei der Einzelregel, die den Klassifikationsfehler am stärksten reduziert. In jedem Suchschritt werden iterativ nacheinander alle Einzelregeln zur Regelbasis des vorhergehenden Suchschrittes probeweise hinzugefügt und die beste übernommen. Die Suche wird beendet, wenn eine festgelegte Anzahl von Regeln erreicht wird oder keine signifikante Reduzierung des Klassifikationsfehlers mehr auftritt.

Die so generierten Regelbasen decken wegen der Defaultregel wiederum den gesamten Eingangsraum ab, wobei durch die V -Matrix die partiellen Redundanzen korrigiert werden.

Im Beispiel wird im ersten Schritt zur besten Regel R_1 eine weitere Regel gesucht. Die beste Regel dafür ist die Regel R_5 , womit sich jetzt eine Regelbasis mit zwei Regeln durch

R_1 :	WENN $(x_{90}=\text{NEG} \cup \text{ZE}) \cap (x_{105}=\text{NG} \cup \text{NK})$	DANN $y=\text{TST-R}$
R_5 :	WENN $x_{90}=\text{NEG} \cap (x_{102}=\text{PK} \cup \text{PG}) \cap (x_{105}=\text{ZE} \cup \text{PK} \cup \text{PG})$	DANN $y=\text{MST-R}$
	SONST	Rückweisung

ergibt. Die Regeln R_1 und R_5 überlappen sich wegen der unterschiedlichen Terme bei x_{105} nicht, es gilt also für alle $i = 1, \dots, N$ Beispiele $\mu_{P_{1i} \cap P_{5i}} = 0$. Die Klassifikationsgüte steigt durch die Ergänzung der Regel R_5 von 0.10 auf 0.18 an.

Im nächsten Schritt wird Regel R_4 hinzugefügt, was eine Klassifikationsgüte der Regelbasis mit drei Regeln und einer Defaultregel von 0.27 zur Folge hat. Regel R_4 überlappt sich weder mit Regel R_1 noch mit Regel R_5 , also gilt $\mu_{P_{1i} \cap P_{4i}} = 0$, $\mu_{P_{4i} \cap P_{5i}} = 0$ und $\mu_{P_{1i} \cap P_{4i} \cap P_{5i}} = 0$.

Im vierten Schritt treten durch das Hinzufügen von Regel R_{10} (Klassifikationsgüte der Regelbasis 0.34)

R_{10} :	WENN $(x_3=\text{PM} \cup \text{PG}) \cap x_6=\text{PSG}$	DANN $y=\text{MSW-R}$
------------	---	-----------------------

erstmalig Überlappungen zu allen bisher eingefügten Regeln auf. Beispielsweise gilt

$$P_1 \cap P_{10} = (x_3=\text{PM} \cup \text{PG}) \cap x_6=\text{PSG} \cap (x_{90}=\text{NEG} \cup \text{ZE}) \cap (x_{105}=\text{NG} \cup \text{NK})$$

usw.

Nach den Hinzufügen der Regeln R_9 , R_{13} , R_{29} , R_{16} , R_6 und R_{44} entsteht eine Regelbasis von 10 Regeln und einer Defaultregel mit einer Klassifikationsgüte von 0.61, die wegen Erreichen der vorgegebenen Höchstanzahl von Regeln nicht weiter ergänzt wird. Dabei wurden die Regeln neu nummeriert, um sie nach den Konklusionen der Schrittphasen zu ordnen:

R_1 :	WENN $x_{15}=\text{ZE} \cap x_{21}=\text{PG} \cap (x_{93}=\text{ZE} \cup \text{PK} \cup \text{PG})$	DANN $y=\text{LRE-R}$
R_2 :	WENN $(x_3=\text{PM} \cup \text{PG} \cup \text{PSG}) \cap x_{90}=\text{ZE} \cap (x_{93}=\text{ZE} \cup \text{PK} \cup \text{PG})$	DANN $y=\text{LRE-R}$
R_3 :	WENN $(x_3=\text{PK} \cup \text{PM} \cup \text{PG}) \cap (x_{14}=\text{NG} \cup \text{NK}) \cap x_{105}=\text{PK}$	DANN $y=\text{MST-R}$
R_4 :	WENN $x_{90}=\text{NEG} \cap (x_{102}=\text{PK} \cup \text{PG}) \cap (x_{105}=\text{ZE} \cup \text{PK} \cup \text{PG})$	DANN $y=\text{MST-R}$
R_5 :	WENN $(x_{90}=\text{NEG} \cup \text{ZE}) \cap (x_{105}=\text{NG} \cup \text{NK})$	DANN $y=\text{TST-R}$
R_6 :	WENN $(x_{15}=\text{PM} \cup \text{PG} \cup \text{PSG}) \cap (x_{102}=\text{ZE} \cup \text{PK}) \cap (x_{105}=\text{NG} \cup \text{NK} \cup \text{ZE})$	DANN $y=\text{TST-R}$
R_7 :	WENN $x_3=\text{ZE} \cap (x_{18}=\text{ZE} \cup \text{PK} \cup \text{PM}) \cap (x_{90}=\text{NEG} \cup \text{ZE}) \cap (x_{105}=\text{ZE} \cup \text{PK})$	DANN $y=\text{PSW-R}$
R_8 :	WENN $(x_{15}=\text{NICHT ZE}) \cap (x_{90}=\text{PK} \cup \text{PM} \cup \text{PG}) \cap x_{93}=\text{PK}$	DANN $y=\text{ISW-R}$
R_9 :	WENN $(x_3=\text{PM} \cup \text{PG}) \cap x_6=\text{PSG}$	DANN $y=\text{MSW-R}$
R_{10} :	WENN $(x_{90}=\text{NEG} \cup \text{ZE} \cup \text{PK}) \cap x_{93}=\text{NK} \cap (x_{102}=\text{NG} \cup \text{NK} \cup \text{ZE})$	DANN $y=\text{TSW-R}$

Damit wird jede Schrittphase durch ein bis zwei Regeln linguistisch beschrieben. Auffällig

ist dabei, dass jeweils unterschiedliche Merkmale verwendet werden.

Die auftretenden Abweichungen zwischen Lerndatensatz und Klassifikatorergebnissen sind geringfügig besser im Vergleich zum statistischen Klassifikator (Fehler Lerndatensatz 16.3%, Testdatensatz 17.5%) und betreffen bis auf 1.0% Fehler nur Abweichungen zur benachbarten Phase (Bild 6.1 in Kapitel 6). Hervorzuheben ist dabei, dass mit einer relativ übersichtlichen Regelbasis mit 10 Regeln ein gutes Ergebnis erzielt wird.

Kapitel 6

Vergleich

Beide Ansätze, das statistische Verfahren auf der Basis des MANOVA-Verfahrens, der Diskriminanzanalyse und von Bayes-Klassifikatoren sowie das Verfahren zur Generierung von Fuzzy-Systemen, führen im hier genannten Problem auf eine gute Klassifikationsgüte unter Berücksichtigung der in Kapitel 2 ausgeführten Umstände der Entstehung des Lerndatensatzes. Bild 6.1 vergleicht die Ergebnisse des Testdatensatzes und beider Verfahren anhand von fünf Schritten mit zwei unterschiedlichen Geschwindigkeiten (links – mittel, rechts – langsam), wobei jeweils noch eine logische Nachkorrektur zur Beseitigung nichtplausibler Übergänge erfolgte. Es fällt auf, dass die Ergebnisse bei der mittleren Geschwindigkeit we-

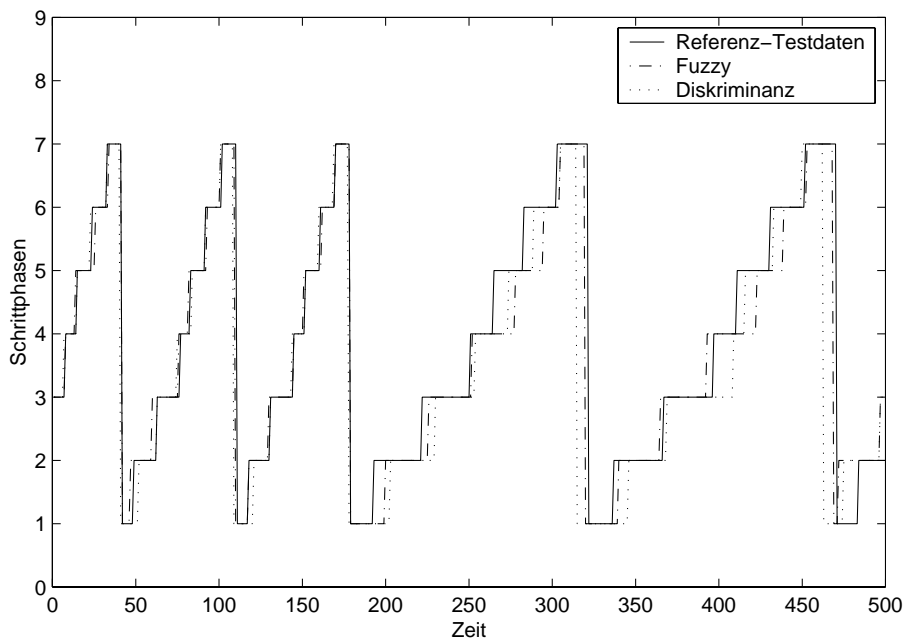


Bild 6.1: Vergleich zwischen den Schrittphasen im Testdatensatz (Referenzwerte aus Codierung entsprechend Kapitel 3), den Ergebnissen des Bayes-Klassifikators nach der Diskriminanzanalyse (Kapitel 4) und dem Fuzzy-Klassifikator (Kapitel 5)

niger voneinander abweichen. Dafür gibt es drei Ursachen, erstens sind die Unterschiede zwischen den einzelnen Schrittphasen bei mittleren Geschwindigkeiten stärker ausgeprägt, zweitens sind die für die Klassifikation der Lerndaten verwendeten Prozentangaben auf mittlere Gehgeschwindigkeiten bezogen und drittens ist die mittlere Geschwindigkeit auch

im Lerndatensatz stärker vertreten und folglich detaillierter angelernt. Dennoch gelingt es beiden Ansätzen, die qualitativen Abläufe richtig zu erfassen.

Die betrachteten statistischen Verfahren beruhen auf der Annahme einer mehrdimensionalen Normalverteilung, die kompakte Gebiete impliziert. Damit eignen sie sich in erster Linie zur Klassifikation von Problemen, die linear trennbar sind und für die Klassen im Eingangsraum näherungsweise normalverteilt sind. Selbst bei deutlichen Abweichungen von der Normalverteilung gibt es immer noch gute Ergebnisse, solange zumindest die Kompaktheit der Gebiete eingehalten wird und die Klassen relativ weit voneinander entfernt liegen. Ein großer Nachteil ist die Tatsache, dass die zugrunde liegenden transformierten Merkmale im niederdimensionalen Merkmalsraum nicht oder nur schwer interpretierbar sind. Hervorzuheben ist die zwar stark abstrahierte, aber bezüglich der Klassifikationsergebnisse, der Ähnlichkeiten zwischen verschiedenen Ausgangsklassen und der möglichen Klassifikationsfehler sehr informative grafische Darstellung der Datensätze.

Der Vorteil der statistischen Verfahren besteht darin, meist eine geschlossene optimale Lösung zu bieten. Das Vorgehen über die Kovarianzmatrizen findet die markantesten Merkmale heraus und vermeidet starke Korrelationen zwischen diesen Merkmalen. Damit berücksichtigt es Abhängigkeiten und Redundanzen von Merkmalen. Nachteilig sind hier allerdings parameterlineare Ansätze, die die Freiheitsgrade einschränken, und die ebenfalls reduzierte Interpretierbarkeit der Ergebnisse.

Beim vorgestellten datenbasierten Entwurf von Fuzzy-Systemen sind die Ergebnisse relativ gut interpretierbar, weil alle Regeln mit einer überschaubaren Anzahl von Merkmalen und linguistischen Termen arbeiten. Die geringe Anzahl der Regeln garantiert auch noch die Überschaubarkeit des Gesamtsystems. Da keine Rücksicht auf den logischen Zusammenhang der Merkmale genommen wird, muss sich der Experte u. U. um die Interpretation von Regeln mit heterogenen Bestandteilen (R_5 : linkes Knie und rechte Hüfte) bemühen. Allerdings findet sich für jedes dieser Merkmale auch ein korreliertes Ersatzmerkmal mit logischem Zusammenhang, das ohne wesentliche Verluste zu einer ähnlichen Regelgüte führt. Aufgrund der nichtlinearen Struktur ist der Klassifikator an viele Probleme besser anpassbar. Allerdings entstehen bei Erhalt der ursprünglichen Merkmale nur achsenparallele Klassifikatoren, was u.U. die Klassifikationsgüte begrenzt, und der Entwurfsaufwand steigt an.

Entsprechend dieser Unterschiede sind in vielen praktischen Anwendungen auch die Merkmale verschieden, die für die jeweiligen Verfahren ausgewählt werden. Bei entsprechenden Datensätzen ist es durchaus möglich, mit einer Gruppe von Merkmalen eine gute lineare, nicht achsenparallele Trennung zu bewirken, und mit einer anderen Gruppe gute achsenparallele, diskrete Partitionierungen für Fuzzy-Klassifikatoren zu erreichen.

Aus der Ähnlichkeit der Ergebnisse bei der Merkmalsselektion und der Klassifikationsgüte darf aber nicht geschlossen werden, dass beide Verfahren immer zu ähnlichen Ergebnissen führen. Bei anderen Aufgabenstellungen fallen die Vor- oder Nachteile der jeweiligen Verfahren stärker ins Gewicht, weswegen sich deutliche Unterschiede ergeben können. Insbesondere bei der Merkmalsselektion existieren Fälle, die zu vollkommen verschiedenen Merkmalskombinationen führen, was unterschiedlichen Lösungswegen entspricht.

Kapitel 7

Implementierung

Die beschriebenen Operationen in der Signalvorverarbeitung, die statistischen Verfahren und die Fuzzy-Regelgenerierung wurden in einem Prototyp unter MATLAB 5.3 [106] implementiert.

Die Schritte in der Signalvorverarbeitung umfassen das Einlesen von Messdaten aus Excel-Dateien (Funktion *open_gang.m*), die von OrthoTrack zur Verfügung gestellt werden, und die Erkennung von Schrittzyklen (Funktion *steprec2.m*). Die Funktion *gangchoice.m* stellt dann die interessierenden Daten zusammen und berechnet die Änderungsgeschwindigkeiten. In dieser Funktion werden auch die Referenzwerte für die Schrittphasen entsprechend der prozentualen Aufteilung in Abschnitt 3.2 festgelegt. Die so gewonnenen Daten werden als Projektdatei **.prj* im MATLAB-Format abgespeichert.

Die weitere Verarbeitung übernimmt ein Entwicklungsprototyp einer MATLAB-Toolbox, die erweiterte Funktionalitäten auf dem Gebiet Statistik und Fuzzy-Modellbildung bereitstellt, die nicht Ganganalyse-spezifisch sind. Dieser Prototyp setzt keine weiteren MATLAB-Toolboxen voraus. Er enthält Algorithmen

- zur uni- oder multivariate Merkmalsselektion mit dem MANOVA-Verfahren,
- für die Hauptkomponenten- und Diskriminanzanalyse,
- zum Klassifikatorentwurf von ausgewählten Bayes-Klassifikatoren (Euklidischer Abstandsklassifikator, Klassifikator mit klassenspezifischen Kovarianzmatrizen),
- zum Anwenden von Klassifikatoren auf neue Datensätze,
- zur Merkmalsselektion mit ID3- oder C4.5-Entscheidungsbäumen,
- zum Entwurf von Fuzzy-Zugehörigkeitsfunktionen (gleichverteilte Parameter, gleiche Häufigkeit, Optimierung mit informationstheoretischen Maßen),
- zur Regelgenerierung mit Entscheidungsbäumen für verschiedene Berechnungsvorschriften (mit und ohne statistische Abschätzungen),
- zum Regelpruning nach mehreren parametrierbaren Verfahren (Regel und Komplementärregel mit negierter Prämisse bei offenen Konklusionen, offener Konklusion und Komplementärkonklusion oder nur mit klaren Regeln),
- zur Auswahl von Fuzzy-Regelbasen mit verschiedenen Defaultregeln (fest vorgegeben, automatisch gewählt, Rückweisung),
- zum Anwenden von Fuzzy-Regelbasen auf neue Datensätze mit zwei verschiedenen Inferenzverfahren (Sum-Prod-Inferenz oder modifizierte Inferenz für überlappende und widersprüchliche Regelbasen [77]),

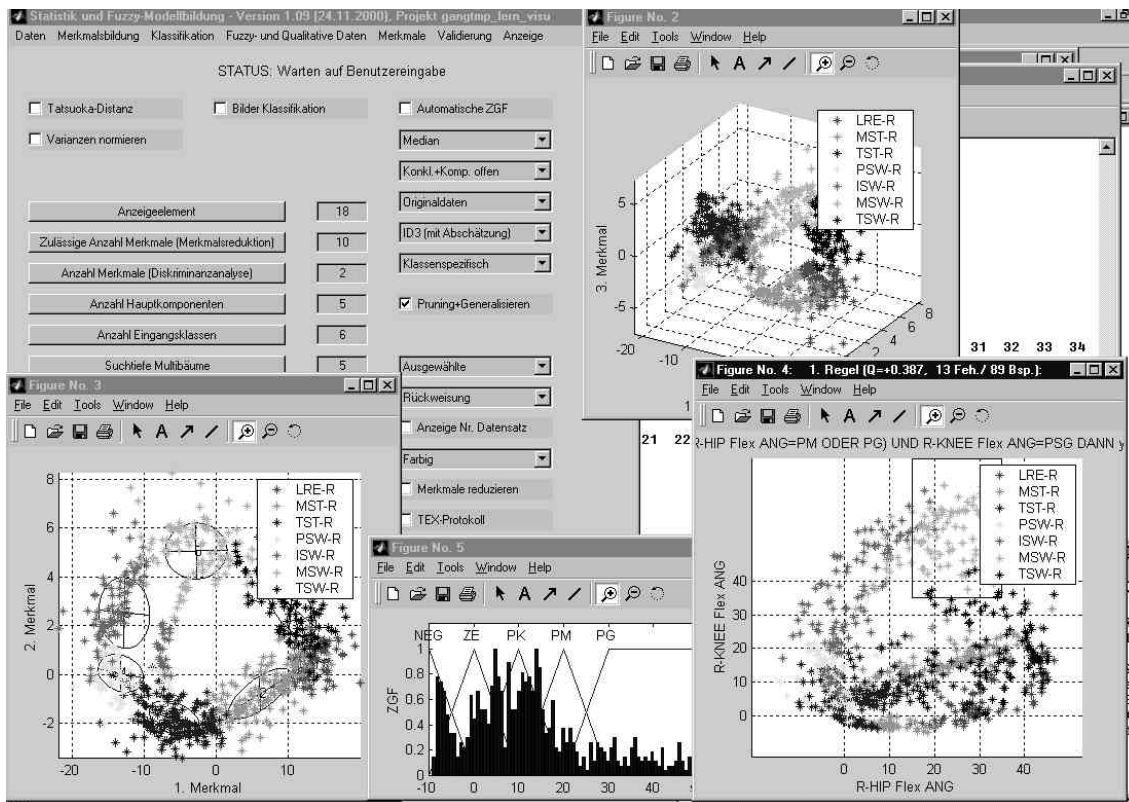


Bild 7.1: Entwicklungsprototyp einer MATLAB-Toolbox für Statistik und Fuzzy-Modellbildung

- zum Speichern von Klassifikatoren und Regelbasen,
- zur Crossvalidierung von linearen und Fuzzy-Klassifikatoren,
- zum Erzeugen von Lern- und Testdatensätzen,
- für verschiedene Visualisierungen und Protokollierungen in Text- oder TEX-Dateien sowie
- zum Export von Fuzzy-Systemen entsprechend der IEC 61131-7 bzw. der entsprechenden EN DIN 61131-7 [21].

Die Ergebnisse der Schrittphasenerkennung aus den Klassifikatoren werden mit der Funktion *rep_gang_sequenz.m* nochmals einer Optimierung unterzogen, indem nicht plausible Schrittphasenübergänge (z. B. Schrittphasen 1-2-1-2-3 oder 1-3-2-3) korrigiert werden. Zur Entscheidung, an welcher Stelle die Übergänge erfolgen sollen, werden die zweitbesten Wahrscheinlichkeitswerte der Bayes- oder Fuzzy-Klassifikatoren herangezogen.

Eine Teilkomponente, die Darstellung der Verläufe für linke und rechte Körperseiten, die aus einer Merkmalsselektion und einer nachfolgenden Diskriminanzanalyse resultiert, wird zur Zeit als Ganganalyse-spezifische Visualisierungskomponente in die in Heidelberg verwendete MS-Access-Datenbank integriert. Diese Datenbank ruft ihrerseits MATLAB-Funktionen zur Auswertung und Visualisierung auf, so dass hier kompatible Lösungen vorliegen. Der Programmcode dieser Komponente (Funktion *compute_features.m*) ist im Anhang angegeben.

Kapitel 8

Zusammenfassung und Ausblick

Der vorliegende Bericht stellt die Ergebnisse der ersten Phase eines Gemeinschaftsprojektes des Forschungszentrums Karlsruhe und der Stiftung Orthopädische Universitätsklinik Heidelberg zusammen, das sich mit der Diagnoseunterstützung in der Instrumentellen Ganganalyse beschäftigt.

Das mittelfristige Ziel des Projektes besteht darin, durch die Systematisierung des vorhandenen subjektiven Expertenwissens und die Analyse klinischer Daten mit datenbasierten Methoden einen Beitrag zur Objektivierung der Diagnose und Therapieplanung in der Instrumentellen Ganganalyse zu leisten. Wesentlich ist dabei nicht nur eine Entscheidungsunterstützung, sondern die Interpretierbarkeit und Nachvollziehbarkeit des Lösungsweges, der zu dieser Entscheidung geführt hat.

Nach einer Einführung (Kapitel 1) stellt das zweite Kapitel neben einer kurzen Übersicht über Methoden und Zielstellungen der Instrumentellen Ganganalyse wesentliche subjektive Vorgehensweisen in der klinischen Praxis zusammen. Die mangelnde Interpretierbarkeit herkömmlicher datenbasierter Verfahren wird als wesentliches Problem bei der klinischen Akzeptanz dieser Verfahren herausgearbeitet. Aus der Vielzahl möglicher Aufgabenstellungen für datenbasierte Verfahren wird mit der Erkennung von Schrittphasen eine typische, aber noch relativ überschaubare Aufgabe ausgewählt, um die Leistungsfähigkeit datenbasierter Methoden für die Instrumentelle Ganganalyse zu zeigen.

Für diese Erkennung von Schrittphasen steht ein Datensatz zur Verfügung, der Kinematik- und Kinetikdaten (87 Merkmale) für ein Normdatenkollektiv mit 8 Personen bei unterschiedlichen Geschwindigkeiten auf einem Diagnostiklaufband enthält (Kapitel 3). Diese Daten werden um Änderungsgeschwindigkeiten der Merkmale ergänzt und mit einer datenbasierten Erkennung von Schrittzyklen strukturiert.

Im Folgenden kommen zwei unterschiedliche Gruppen von datenbasierten Verfahren zum Einsatz, statistische Verfahren (Kapitel 4) und Fuzzy-Verfahren (Kapitel 5). Die Vorgehensweise beim Entwurf und die mathematischen Grundlagen werden ausführlich diskutiert. Beide Verfahrensgruppen gestatten einen vollautomatischen Entwurf, wenn mit Standardeinstellungen für einige Parameter gearbeitet wird.

Bei beiden Verfahrensgruppen steht neben der datenbasierten, möglichst fehlerfreien Erkennung der Schrittphasen die Interpretation des Lösungsweges im Vordergrund. Dabei werden jeweils wichtige Merkmale aus dem Messdatensatz ausgewählt, die sich besonders gut für die vorliegende Klassifikationsaufgabe eignen. Die so entstehenden Listen der Merkmalsrelevanzen sind das erste, wichtige Projektergebnis.

Bei den statistischen Verfahren werden MANOVA-Untersuchungen für die Merkmalsselektion, Hauptkomponenten- oder Diskriminanzanalyse zur Dimensionsreduktion und

Bayes-Klassifikatoren eingesetzt. Die entstehenden Transformationsmatrizen der Dimensionsreduktion erlauben noch eine Interpretation der entstehenden, abstrakten Merkmale.

Der Entwurf der Fuzzy-Systeme umfasst die Teilschritte Merkmalsselektion mit Entscheidungsbäumen, Entwurf der Zugehörigkeitsfunktionen mit informationstheoretischen Maßen, Suche nach Einzelregeln mit Entscheidungsbäumen, Regelpruning und modularen Maßen sowie die Suche nach Regelbasen. Die Interpretierbarkeit resultiert aus speziellen Vorgaben beim Entwurf der Zugehörigkeitsfunktionen, relativ einfachen Fuzzy-Regeln und überschaubaren Regelbasen.

Alle vorgestellten Verfahren erzielen ähnlich gute Ergebnisse (Kapitel 6) und wurden unter MATLAB implementiert (Kapitel 7). Der Anhang umfasst detaillierte Ergebnisse der datenbasierten Verfahren und mathematische Grundlagen.

Mit der erfolgreichen Analyse der Schrittphasen wurde die prinzipielle Anwendbarkeit der datenbasierten Verfahren für die Ganganalyse demonstriert. Die zukünftigen Arbeiten, die die DFG im Rahmen eines Projektes fördert, werden sich hauptsächlich auf die Diagnose pathologischer Fälle konzentrieren.

Die hier vorgestellten Algorithmen sind ohne wesentliche strukturelle Änderungen auf eine Vielzahl medizinischer und nichtmedizinischer Klassifikationsprobleme übertragbar. Besonders vorteilhaft ist neben der bereits erwähnten Interpretierbarkeit der Ergebnisse die Fähigkeit, mit großen Mengen redundanter Merkmale arbeiten zu können. Die Leistungsfähigkeit wurde bereits anhand der Analyse der spontanen Baroreflexsensitivität [69] bei Patienten mit dilatativer Kardiomyopathie gezeigt [81].

Anhang A

Merkmale, Merkmalsrelevanzen und Regeln

A.1 Merkmalsbezeichnungen

Nr.	Name	Nr.	Name	Nr.	Name
x_1	R-HIP Rot ANG	x_{59}	L-ANK Abd FRC	x_{117}	Trunk-Rotation V
x_2	R-HIP Abd ANG	x_{60}	L-ANK Flex FRC	x_{118}	Pelvis-Lat-Tilt V
x_3	R-HIP Flex ANG	x_{61}	L-GRF-FWD	x_{119}	Pelvis-Fwd-Tilt V
x_4	R-KNEE Rot ANG	x_{62}	L-GRF-LAT	x_{120}	Pelvis-Rotation V
x_5	R-KNEE Abd ANG	x_{63}	L-GRF-VRT	x_{121}	R-Foot-OrientationV
x_6	R-KNEE Flex ANG	x_{64}	R-HIP Rot MOM	x_{122}	R-Foot-Null V
x_7	R-ANK Rot ANG	x_{65}	R-HIP Abd MOM	x_{123}	R-Foot-Null V
x_8	R-ANK Abd ANG	x_{66}	R-HIP Flex MOM	x_{124}	L-Foot-OrientationV
x_9	R-ANK Flex ANG	x_{67}	R-KNEE Rot MOM	x_{125}	L-Foot-Null V
x_{10}	R-Elbow-Ang	x_{68}	R-KNEE Abd MOM	x_{126}	L-Foot-Null V
x_{11}	R-Shl-Add-Ang	x_{69}	R-KNEE Flex MOM	x_{127}	R-HIP Rot FRC V
x_{12}	R-Shl-Flex-Ang	x_{70}	R-ANK Rot MOM	x_{128}	R-HIP Abd FRC V
x_{13}	L-HIP Rot ANG	x_{71}	R-ANK Abd MOM	x_{129}	R-HIP Flex FRC V
x_{14}	L-HIP Abd ANG	x_{72}	R-ANK Flex MOM	x_{130}	R-KNEE Rot FRC V
x_{15}	L-HIP Flex ANG	x_{73}	L-HIP Rot MOM	x_{131}	R-KNEE Abd FRC V
x_{16}	L-KNEE Rot ANG	x_{74}	L-HIP Abd MOM	x_{132}	R-KNEE Flex FRC V
x_{17}	L-KNEE Abd ANG	x_{75}	L-HIP Flex MOM	x_{133}	R-ANK Rot FRC V
x_{18}	L-KNEE Flex ANG	x_{76}	L-KNEE Rot MOM	x_{134}	R-ANK Abd FRC V
x_{19}	L-ANK Rot ANG	x_{77}	L-KNEE Abd MOM	x_{135}	R-ANK Flex FRC V
x_{20}	L-ANK Abd ANG	x_{78}	L-KNEE Flex MOM	x_{136}	R-GRF-FWD V
x_{21}	L-ANK Flex ANG	x_{79}	L-ANK Rot MOM	x_{137}	R-GRF-LAT V
x_{22}	L-Elbow-Ang	x_{80}	L-ANK Abd MOM	x_{138}	R-GRF-VRT V
x_{23}	L-Shl-Add-Ang	x_{81}	L-ANK Flex MOM	x_{139}	L-HIP Rot FRC V
x_{24}	L-Shl-Flex-Ang	x_{82}	R-HIP PWR	x_{140}	L-HIP Abd FRC V
x_{25}	Head-Lat-Tilt	x_{83}	R-KNEE PWR	x_{141}	L-HIP Flex FRC V
x_{26}	Head-Fwd-Tilt	x_{84}	R-ANK PWR	x_{142}	L-KNEE Rot FRC V
x_{27}	Head-Rotation	x_{85}	L-HIP PWR	x_{143}	L-KNEE Abd FRC V
x_{28}	Trunk-Lat-Tilt	x_{86}	L-KNEE PWR	x_{144}	L-KNEE Flex FRC V
x_{29}	Trunk-Fwd-Tilt	x_{87}	L-ANK PWR	x_{145}	L-ANK Rot FRC V
x_{30}	Trunk-Rotation	x_{88}	R-HIP Rot ANG V	x_{146}	L-ANK Abd FRC V
x_{31}	Pelvis-Lat-Tilt	x_{89}	R-HIP Abd ANG V	x_{147}	L-ANK Flex FRC V
x_{32}	Pelvis-Fwd-Tilt	x_{90}	R-HIP Flex ANG V	x_{148}	L-GRF-FWD V
x_{33}	Pelvis-Rotation	x_{91}	R-KNEE Rot ANG V	x_{149}	L-GRF-LAT V
x_{34}	R-Foot-Orientation	x_{92}	R-KNEE Abd ANG V	x_{150}	L-GRF-VRT V
x_{35}	R-Foot-Null	x_{93}	R-KNEE Flex ANG V	x_{151}	R-HIP Rot MOM V
x_{36}	R-Foot-Null	x_{94}	R-ANK Rot ANG V	x_{152}	R-HIP Abd MOM V
x_{37}	L-Foot-Orientation	x_{95}	R-ANK Abd ANG V	x_{153}	R-HIP Flex MOM V
x_{38}	L-Foot-Null	x_{96}	R-ANK Flex ANG V	x_{154}	R-KNEE Rot MOM V
x_{39}	L-Foot-Null	x_{97}	R-Elbow-Ang V	x_{155}	R-KNEE Abd MOM V
x_{40}	R-HIP Rot FRC	x_{98}	R-Shl-Add-Ang V	x_{156}	R-KNEE Flex MOM V
x_{41}	R-HIP Abd FRC	x_{99}	R-Shl-Flex-Ang V	x_{157}	R-ANK Rot MOM V
x_{42}	R-HIP Flex FRC	x_{100}	L-HIP Rot ANG V	x_{158}	R-ANK Abd MOM V
x_{43}	R-KNEE Rot FRC	x_{101}	L-HIP Abd ANG V	x_{159}	R-ANK Flex MOM V
x_{44}	R-KNEE Abd FRC	x_{102}	L-HIP Flex ANG V	x_{160}	L-HIP Rot MOM V
x_{45}	R-KNEE Flex FRC	x_{103}	L-KNEE Rot ANG V	x_{161}	L-HIP Abd MOM V
x_{46}	R-ANK Rot FRC	x_{104}	L-KNEE Abd ANG V	x_{162}	L-HIP Flex MOM V
x_{47}	R-ANK Abd FRC	x_{105}	L-KNEE Flex ANG V	x_{163}	L-KNEE Rot MOM V
x_{48}	R-ANK Flex FRC	x_{106}	L-ANK Rot ANG V	x_{164}	L-KNEE Abd MOM V
x_{49}	R-GRF-FWD	x_{107}	L-ANK Abd ANG V	x_{165}	L-KNEE Flex MOM V
x_{50}	R-GRF-LAT	x_{108}	L-ANK Flex ANG V	x_{166}	L-ANK Rot MOM V
x_{51}	R-GRF-VRT	x_{109}	L-Elbow-Ang V	x_{167}	L-ANK Abd MOM V
x_{52}	L-HIP Rot FRC	x_{110}	L-Shl-Add-Ang V	x_{168}	L-ANK Flex MOM V
x_{53}	L-HIP Abd FRC	x_{111}	L-Shl-Flex-Ang V	x_{169}	R-HIP PWR V
x_{54}	L-HIP Flex FRC	x_{112}	Head-Lat-Tilt V	x_{170}	R-KNEE PWR V
x_{55}	L-KNEE Rot FRC	x_{113}	Head-Fwd-Tilt V	x_{171}	R-ANK PWR V
x_{56}	L-KNEE Abd FRC	x_{114}	Head-Rotation V	x_{172}	L-HIP PWR V
x_{57}	L-KNEE Flex FRC	x_{115}	Trunk-Lat-Tilt V	x_{173}	L-KNEE PWR V
x_{58}	L-ANK Rot FRC	x_{116}	Trunk-Fwd-Tilt V	x_{174}	L-ANK PWR V

A.2 Univariate statistische Merkmalsrelevanzen

Pos.	Nr.	Bezeichnung	Güte	Pos.	Nr.	Bezeichnung	Güte
1	x_3	R-HIP Flex ANG	0.726	61	x_{146}	L-ANK Abd FRC V	0.143
2	x_6	R-KNEE Flex ANG	0.721	62	x_{79}	L-ANK Rot MOM	0.135
3	x_{90}	R-HIP Flex ANG V	0.718	63	x_{86}	L-KNEE PWR	0.133
4	x_{102}	L-HIP Flex ANG V	0.700	64	x_{121}	R-Foot-OrientationV	0.130
5	x_{15}	L-HIP Flex ANG	0.680	65	x_{53}	L-HIP Abd FRC	0.128
6	x_{93}	R-KNEE Flex ANG V	0.639	66	x_{115}	Trunk-Lat-Tilt V	0.126
7	x_{105}	L-KNEE Flex ANG V	0.619	67	x_{19}	L-ANK Rot ANG	0.124
8	x_{21}	L-ANK Flex ANG	0.565	68	x_{84}	R-ANK PWR	0.119
9	x_{14}	L-HIP Abd ANG	0.558	69	x_{16}	L-KNEE Rot ANG	0.117
10	x_{97}	R-Elbow-Ang V	0.511	70	x_{160}	L-HIP Rot MOM V	0.116
11	x_{18}	L-KNEE Flex ANG	0.494	71	x_{108}	L-ANK Flex ANG V	0.114
12	x_{109}	L-Elbow-Ang V	0.491	72	x_{128}	R-HIP Abd FRC V	0.113
13	x_9	R-ANK Flex ANG	0.480	73	x_{163}	L-KNEE Rot MOM V	0.112
14	x_{44}	R-KNEE Abd FRC	0.473	74	x_{140}	L-HIP Abd FRC V	0.110
15	x_2	R-HIP Abd ANG	0.431	75	x_{151}	R-HIP Rot MOM V	0.110
16	x_{47}	R-ANK Abd FRC	0.416	76	x_{147}	L-ANK Flex FRC V	0.108
17	x_{67}	R-KNEE Rot MOM	0.402	77	x_{144}	L-KNEE Flex FRC V	0.108
18	x_{37}	L-Foot-Orientation	0.390	78	x_{22}	L-Elbow-Ang	0.106
19	x_{89}	R-HIP Abd ANG V	0.370	79	x_{170}	R-KNEE PWR V	0.105
20	x_{33}	Pelvis-Rotation	0.366	80	x_{73}	L-HIP Rot MOM	0.101
21	x_{34}	R-Foot-Orientation	0.363	81	x_{80}	L-ANK Abd MOM	0.101
22	x_{120}	Pelvis-Rotation V	0.356	82	x_{28}	Trunk-Lat-Tilt	0.100
23	x_{70}	R-ANK Rot MOM	0.354	83	x_{168}	L-ANK Flex MOM V	0.097
24	x_{45}	R-KNEE Flex FRC	0.354	84	x_{10}	R-Elbow-Ang	0.097
25	x_{59}	L-ANK Abd FRC	0.341	85	x_4	R-KNEE Rot ANG	0.096
26	x_{31}	Pelvis-Lat-Tilt	0.337	86	x_{159}	R-ANK Flex MOM V	0.093
27	x_{82}	R-HIP PWR	0.322	87	x_{171}	R-ANK PWR V	0.092
28	x_{56}	L-KNEE Abd FRC	0.321	88	x_{167}	L-ANK Abd MOM V	0.090
29	x_{50}	R-GRF-LAT	0.311	89	x_{54}	L-HIP Flex FRC	0.087
30	x_{48}	R-ANK Flex FRC	0.310	90	x_{158}	R-ANK Abd MOM V	0.083
31	x_{111}	L-Shl-Flex-Ang V	0.302	91	x_{165}	L-KNEE Flex MOM V	0.083
32	x_{17}	L-KNEE Abd ANG	0.300	92	x_{157}	R-ANK Rot MOM V	0.082
33	x_{20}	L-ANK Abd ANG	0.295	93	x_{173}	L-KNEE PWR V	0.082
34	x_{99}	R-Shl-Flex-Ang V	0.283	94	x_{166}	L-ANK Rot MOM V	0.081
35	x_{104}	L-KNEE Abd ANG V	0.280	95	x_{169}	R-HIP PWR V	0.080
36	x_{60}	L-ANK Flex FRC	0.280	96	x_{162}	L-HIP Flex MOM V	0.080
37	x_{101}	L-HIP Abd ANG V	0.279	97	x_{65}	R-HIP Abd MOM	0.077
38	x_{96}	R-ANK Flex ANG V	0.261	98	x_{154}	R-KNEE Rot MOM V	0.077
39	x_{131}	R-KNEE Abd FRC V	0.260	99	x_{124}	L-Foot-OrientationV	0.076
40	x_{85}	L-HIP PWR	0.257	100	x_{69}	R-KNEE Flex MOM	0.075
41	x_{57}	L-KNEE Flex FRC	0.256	101	x_{161}	L-HIP Abd MOM V	0.074
42	x_{110}	L-Shl-Add-Ang V	0.254	102	x_{77}	L-KNEE Abd MOM	0.074
43	x_{116}	Trunk-Fwd-Tilt V	0.241	103	x_{119}	Pelvis-Fwd-Tilt V	0.074
44	x_{83}	R-KNEE PWR	0.226	104	x_{155}	R-KNEE Abd MOM V	0.072
45	x_5	R-KNEE Abd ANG	0.223	105	x_{129}	R-HIP Flex FRC V	0.071
46	x_{118}	Pelvis-Lat-Tilt V	0.211	106	x_{100}	L-HIP Rot ANG V	0.070
47	x_{92}	R-KNEE Abd ANG V	0.208	107	x_{91}	R-KNEE Rot ANG V	0.070
48	x_{134}	R-ANK Abd FRC V	0.207	108	x_{81}	L-ANK Flex MOM	0.061
49	x_{42}	R-HIP Flex FRC	0.206	109	x_{13}	L-HIP Rot ANG	0.060
50	x_{64}	R-HIP Rot MOM	0.202	110	x_{152}	R-HIP Abd MOM V	0.057
51	x_{143}	L-KNEE Abd FRC V	0.190	111	x_{106}	L-ANK Rot ANG V	0.057
52	x_{88}	R-HIP Rot ANG V	0.189	112	x_1	R-HIP Rot ANG	0.055
53	x_7	R-ANK Rot ANG	0.188	113	x_{117}	Trunk-Rotation V	0.054
54	x_{132}	R-KNEE Flex FRC V	0.179	114	x_{164}	L-KNEE Abd MOM V	0.050
55	x_{41}	R-HIP Abd FRC	0.168	115	x_{174}	L-ANK PWR V	0.050
56	x_{62}	L-GRF-LAT	0.165	116	x_{156}	R-KNEE Flex MOM V	0.048
57	x_{76}	L-KNEE Rot MOM	0.160	117	x_{137}	R-GRF-LAT V	0.045
58	x_{135}	R-ANK Flex FRC V	0.150	118	x_{71}	R-ANK Abd MOM	0.045
59	x_{98}	R-Shl-Add-Ang V	0.148	119	x_8	R-ANK Abd ANG	0.045
60	x_{72}	R-ANK Flex MOM	0.147	120	x_{12}	R-Shl-Flex-Ang	0.041

A.3 Univariate ID3-Merkmalsrelevanzen

Pos.	Nr.	Bezeichnung	Güte	Pos.	Nr.	Bezeichnung	Güte
1	x_6	R-KNEE Flex ANG	0.276	61	x_{107}	L-ANK Abd ANG V	0.062
2	x_{90}	R-HIP Flex ANG V	0.273	62	x_{128}	R-HIP Abd FRC V	0.061
3	x_3	R-HIP Flex ANG	0.261	63	x_{92}	R-KNEE Abd ANG V	0.059
4	x_{102}	L-HIP Flex ANG V	0.260	64	x_7	R-ANK Rot ANG	0.057
5	x_{93}	R-KNEE Flex ANG V	0.246	65	x_{132}	R-KNEE Flex FRC V	0.056
6	x_{15}	L-HIP Flex ANG	0.245	66	x_{62}	L-GRF-LAT	0.056
7	x_{21}	L-ANK Flex ANG	0.239	67	x_{146}	L-ANK Abd FRC V	0.054
8	x_{105}	L-KNEE Flex ANG V	0.233	68	x_{76}	L-KNEE Rot MOM	0.050
9	x_{18}	L-KNEE Flex ANG	0.212	69	x_{118}	Pelvis-Lat-Tilt V	0.050
10	x_{14}	L-HIP Abd ANG	0.211	70	x_{16}	L-KNEE Rot ANG	0.049
11	x_9	R-ANK Flex ANG	0.194	71	x_{154}	R-KNEE Rot MOM V	0.045
12	x_{44}	R-KNEE Abd FRC	0.170	72	x_{66}	R-HIP Flex MOM	0.044
13	x_{97}	R-Elbow-Ang V	0.156	73	x_8	R-ANK Abd ANG	0.043
14	x_{67}	R-KNEE Rot MOM	0.155	74	x_{157}	R-ANK Rot MOM V	0.042
15	x_{47}	R-ANK Abd FRC	0.154	75	x_{69}	R-KNEE Flex MOM	0.042
16	x_{20}	L-ANK Abd ANG	0.150	76	x_{135}	R-ANK Flex FRC V	0.042
17	x_{109}	L-Elbow-Ang V	0.148	77	x_{88}	R-HIP Rot ANG V	0.041
18	x_2	R-HIP Abd ANG	0.140	78	x_4	R-KNEE Rot ANG	0.038
19	x_{37}	L-Foot-Orientation	0.139	79	x_{73}	L-HIP Rot MOM	0.038
20	x_{96}	R-ANK Flex ANG V	0.132	80	x_{52}	L-HIP Rot FRC	0.038
21	x_{33}	Pelvis-Rotation	0.131	81	x_{79}	L-ANK Rot MOM	0.038
22	x_{31}	Pelvis-Lat-Tilt	0.128	82	x_{22}	L-Elbow-Ang	0.038
23	x_{34}	R-Foot-Orientation	0.126	83	x_1	R-HIP Rot ANG	0.037
24	x_{56}	L-KNEE Abd FRC	0.117	84	x_{86}	L-KNEE PWR	0.037
25	x_{82}	R-HIP PWR	0.113	85	x_{10}	R-Elbow-Ang	0.037
26	x_{120}	Pelvis-Rotation V	0.108	86	x_{74}	L-HIP Abd MOM	0.036
27	x_{85}	L-HIP PWR	0.104	87	x_{159}	R-ANK Flex MOM V	0.036
28	x_{108}	L-ANK Flex ANG V	0.104	88	x_{80}	L-ANK Abd MOM	0.036
29	x_{59}	L-ANK Abd FRC	0.102	89	x_{144}	L-KNEE Flex FRC V	0.036
30	x_{89}	R-HIP Abd ANG V	0.099	90	x_{81}	L-ANK Flex MOM	0.036
31	x_{41}	R-HIP Abd FRC	0.098	91	x_{168}	L-ANK Flex MOM V	0.035
32	x_{45}	R-KNEE Flex FRC	0.098	92	x_{71}	R-ANK Abd MOM	0.035
33	x_{83}	R-KNEE PWR	0.096	93	x_{30}	Trunk-Rotation	0.033
34	x_{50}	R-GRF-LAT	0.094	94	x_{28}	Trunk-Lat-Tilt	0.033
35	x_{60}	L-ANK Flex FRC	0.093	95	x_{77}	L-KNEE Abd MOM	0.032
36	x_{101}	L-HIP Abd ANG V	0.093	96	x_{147}	L-ANK Flex FRC V	0.031
37	x_{17}	L-KNEE Abd ANG	0.089	97	x_{173}	L-KNEE PWR V	0.031
38	x_{131}	R-KNEE Abd FRC V	0.089	98	x_{98}	R-Shl-Add-Ang V	0.031
39	x_5	R-KNEE Abd ANG	0.086	99	x_{58}	L-ANK Rot FRC	0.030
40	x_{151}	R-HIP Rot MOM V	0.086	100	x_{115}	Trunk-Lat-Tilt V	0.030
41	x_{70}	R-ANK Rot MOM	0.085	101	x_{54}	L-HIP Flex FRC	0.030
42	x_{64}	R-HIP Rot MOM	0.083	102	x_{133}	R-ANK Rot FRC V	0.028
43	x_{104}	L-KNEE Abd ANG V	0.083	103	x_{100}	L-HIP Rot ANG V	0.028
44	x_{134}	R-ANK Abd FRC V	0.080	104	x_{156}	R-KNEE Flex MOM V	0.028
45	x_{124}	L-Foot-OrientationV	0.080	105	x_{171}	R-ANK PWR V	0.028
46	x_{48}	R-ANK Flex FRC	0.078	106	x_{167}	L-ANK Abd MOM V	0.028
47	x_{111}	L-Shl-Flex-Ang V	0.078	107	x_{65}	R-HIP Abd MOM	0.027
48	x_{57}	L-KNEE Flex FRC	0.075	108	x_{87}	L-ANK PWR	0.027
49	x_{110}	L-Shl-Add-Ang V	0.074	109	x_{13}	L-HIP Rot ANG	0.026
50	x_{84}	R-ANK PWR	0.072	110	x_{91}	R-KNEE Rot ANG V	0.026
51	x_{99}	R-Shl-Flex-Ang V	0.072	111	x_{153}	R-HIP Flex MOM V	0.026
52	x_{140}	L-HIP Abd FRC V	0.071	112	x_{103}	L-KNEE Rot ANG V	0.026
53	x_{40}	R-HIP Rot FRC	0.070	113	x_{165}	L-KNEE Flex MOM V	0.025
54	x_{53}	L-HIP Abd FRC	0.069	114	x_{170}	R-KNEE PWR V	0.025
55	x_{42}	R-HIP Flex FRC	0.069	115	x_{106}	L-ANK Rot ANG V	0.025
56	x_{72}	R-ANK Flex MOM	0.069	116	x_{51}	R-GRF-VRT	0.024
57	x_{121}	R-Foot-OrientationV	0.067	117	x_{68}	R-KNEE Abd MOM	0.023
58	x_{143}	L-KNEE Abd FRC V	0.065	118	x_{162}	L-HIP Flex MOM V	0.022
59	x_{19}	L-ANK Rot ANG	0.065	119	x_{158}	R-ANK Abd MOM V	0.022
60	x_{116}	Trunk-Fwd-Tilt V	0.064	120	x_{55}	L-KNEE Rot FRC	0.022

A.4 Multivariate ID3-Merkmalsrelevanzen

Pos.	Nr.	Bezeichnung	Güte	Pos.	Nr.	Bezeichnung	Güte
1	x_6	R-KNEE Flex ANG	0.175	61	x_{76}	L-KNEE Rot MOM	0.038
2	x_3	R-HIP Flex ANG	0.161	62	x_{42}	R-HIP Flex FRC	0.038
3	x_{102}	L-HIP Flex ANG V	0.158	63	x_7	R-ANK Rot ANG	0.036
4	x_{93}	R-KNEE Flex ANG V	0.156	64	x_{143}	L-KNEE Abd FRC V	0.036
5	x_{15}	L-HIP Flex ANG	0.150	65	x_{128}	R-HIP Abd FRC V	0.034
6	x_{18}	L-KNEE Flex ANG	0.150	66	x_{99}	R-Shl-Flex-Ang V	0.034
7	x_{105}	L-KNEE Flex ANG V	0.147	67	x_{132}	R-KNEE Flex FRC V	0.032
8	x_{90}	R-HIP Flex ANG V	0.140	68	x_{168}	L-ANK Flex MOM V	0.032
9	x_{21}	L-ANK Flex ANG	0.135	69	x_{51}	R-GRF-VRT	0.032
10	x_{14}	L-HIP Abd ANG	0.122	70	x_{88}	R-HIP Rot ANG V	0.031
11	x_9	R-ANK Flex ANG	0.109	71	x_{87}	L-ANK PWR	0.031
12	x_{109}	L-Elbow-Ang V	0.093	72	x_{92}	R-KNEE Abd ANG V	0.031
13	x_{97}	R-Elbow-Ang V	0.093	73	x_{157}	R-ANK Rot MOM V	0.030
14	x_{44}	R-KNEE Abd FRC	0.093	74	x_{62}	L-GRF-LAT	0.030
15	x_{67}	R-KNEE Rot MOM	0.088	75	x_{69}	R-KNEE Flex MOM	0.030
16	x_{47}	R-ANK Abd FRC	0.085	76	x_1	R-HIP Rot ANG	0.029
17	x_{20}	L-ANK Abd ANG	0.084	77	x_8	R-ANK Abd ANG	0.029
18	x_{96}	R-ANK Flex ANG V	0.083	78	x_{159}	R-ANK Flex MOM V	0.029
19	x_{33}	Pelvis-Rotation	0.073	79	x_{118}	Pelvis-Lat-Tilt V	0.029
20	x_{37}	L-Foot-Orientation	0.072	80	x_{146}	L-ANK Abd FRC V	0.028
21	x_2	R-HIP Abd ANG	0.071	81	x_{173}	L-KNEE PWR V	0.028
22	x_{31}	Pelvis-Lat-Tilt	0.069	82	x_{52}	L-HIP Rot FRC	0.028
23	x_{56}	L-KNEE Abd FRC	0.066	83	x_{79}	L-ANK Rot MOM	0.028
24	x_{85}	L-HIP PWR	0.065	84	x_{162}	L-HIP Flex MOM V	0.028
25	x_{108}	L-ANK Flex ANG V	0.064	85	x_{71}	R-ANK Abd MOM	0.027
26	x_{83}	R-KNEE PWR	0.062	86	x_4	R-KNEE Rot ANG	0.026
27	x_{59}	L-ANK Abd FRC	0.060	87	x_{153}	R-HIP Flex MOM V	0.026
28	x_{34}	R-Foot-Orientation	0.058	88	x_{154}	R-KNEE Rot MOM V	0.026
29	x_{17}	L-KNEE Abd ANG	0.058	89	x_{167}	L-ANK Abd MOM V	0.026
30	x_{41}	R-HIP Abd FRC	0.058	90	x_{171}	R-ANK PWR V	0.025
31	x_{82}	R-HIP PWR	0.057	91	x_{58}	L-ANK Rot FRC	0.025
32	x_{45}	R-KNEE Flex FRC	0.056	92	x_{30}	Trunk-Rotation	0.025
33	x_{64}	R-HIP Rot MOM	0.056	93	x_{10}	R-Elbow-Ang	0.024
34	x_{50}	R-GRF-LAT	0.055	94	x_{155}	R-KNEE Abd MOM V	0.024
35	x_{120}	Pelvis-Rotation V	0.053	95	x_{73}	L-HIP Rot MOM	0.024
36	x_{84}	R-ANK PWR	0.052	96	x_{22}	L-Elbow-Ang	0.024
37	x_{101}	L-HIP Abd ANG V	0.052	97	x_{165}	L-KNEE Flex MOM V	0.024
38	x_{89}	R-HIP Abd ANG V	0.052	98	x_{86}	L-KNEE PWR	0.024
39	x_{70}	R-ANK Rot MOM	0.051	99	x_{74}	L-HIP Abd MOM	0.024
40	x_{60}	L-ANK Flex FRC	0.051	100	x_{135}	R-ANK Flex FRC V	0.023
41	x_{40}	R-HIP Rot FRC	0.049	101	x_{68}	R-KNEE Abd MOM	0.023
42	x_{134}	R-ANK Abd FRC V	0.049	102	x_{49}	R-GRF-FWD	0.023
43	x_{151}	R-HIP Rot MOM V	0.048	103	x_{55}	L-KNEE Rot FRC	0.022
44	x_{131}	R-KNEE Abd FRC V	0.047	104	x_{91}	R-KNEE Rot ANG V	0.022
45	x_{72}	R-ANK Flex MOM	0.046	105	x_{133}	R-ANK Rot FRC V	0.022
46	x_5	R-KNEE Abd ANG	0.045	106	x_{152}	R-HIP Abd MOM V	0.022
47	x_{111}	L-Shl-Flex-Ang V	0.045	107	x_{158}	R-ANK Abd MOM V	0.022
48	x_{57}	L-KNEE Flex FRC	0.045	108	x_{77}	L-KNEE Abd MOM	0.022
49	x_{124}	L-Foot-OrientationV	0.045	109	x_{63}	L-GRF-VRT	0.022
50	x_{48}	R-ANK Flex FRC	0.044	110	x_{144}	L-KNEE Flex FRC V	0.022
51	x_{110}	L-Shl-Add-Ang V	0.043	111	x_{65}	R-HIP Abd MOM	0.022
52	x_{107}	L-ANK Abd ANG V	0.043	112	x_{46}	R-ANK Rot FRC	0.021
53	x_{116}	Trunk-Fwd-Tilt V	0.043	113	x_{115}	Trunk-Lat-Tilt V	0.021
54	x_{53}	L-HIP Abd FRC	0.043	114	x_{80}	L-ANK Abd MOM	0.021
55	x_{19}	L-ANK Rot ANG	0.042	115	x_{156}	R-KNEE Flex MOM V	0.021
56	x_{121}	R-Foot-OrientationV	0.041	116	x_{43}	R-KNEE Rot FRC	0.020
57	x_{104}	L-KNEE Abd ANG V	0.041	117	x_{81}	L-ANK Flex MOM	0.020
58	x_{16}	L-KNEE Rot ANG	0.039	118	x_{147}	L-ANK Flex FRC V	0.020
59	x_{66}	R-HIP Flex MOM	0.038	119	x_{98}	R-Shl-Add-Ang V	0.018
60	x_{140}	L-HIP Abd FRC V	0.038	120	x_{54}	L-HIP Flex FRC	0.018

A.5 Zugehörigkeitsfunktionen der ausgewählten Merkmale

x_l	Bezeichnung	$a_{l,1}$	$a_{l,2}$	$a_{l,3}$	$a_{l,4}$	$a_{l,5}$
x_3	R-HIP Flex ANG	0.00	10.00	20.00	30.00	40.00
x_6	R-KNEE Flex ANG	0.00	10.00	20.00	30.00	40.00
x_{14}	L-HIP Abd ANG	-10.00	-2.00	0.00	2.00	10.00
x_{15}	L-KNEE Flex ANG	-10.00	0.00	10.00	20.00	30.00
x_{18}	L-HIP Flex ANG	0.00	10.00	20.00	30.00	40.00
x_{21}	L-ANK Flex ANG	-20.00	-10.00	0.00	6.00	10.00
x_{90}	R-HIP Flex ANG V	-1.00	0.00	1.00	2.00	3.00
x_{93}	R-KNEE Flex ANG V	-10.00	-3.00	0.00	1.00	10.00
x_{102}	L-HIP Flex ANG V	-2.00	-1.00	0.00	1.00	2.00
x_{105}	L-KNEE Flex ANG V	-10.00	-3.00	0.00	1.00	10.00

A.6 Vertrauensintervalle bei Binomialverteilungen

Die Vertrauensgrenzen $\hat{p}^l(y = B_j | \mathbf{x} = P_k)$ und $\hat{p}^u(y = B_j | \mathbf{x} = P_k)$ können statistischen Standardwerken entnommen oder über Näherungsformeln berechnet werden [34]:

$$\hat{p}^u(y = B_j | \mathbf{x} = P_k) = \min \left\{ 1; \frac{1}{0.7 + 0.25n_{P_k}} + \hat{p}(y = B_j | \mathbf{x} = P_k) (1 + 0.2347n_{P_k}^{0.1536} \times e^{-0.00375n_{P_k}}) - 0.571\hat{p}^2(y = B_j | \mathbf{x} = P_k)e^{-0.00317n_{P_k}^2} \right\} \quad (\text{A.1})$$

$$\hat{p}^l(y = B_j | \mathbf{x} = P_k) = \max \left\{ 0; 1 - \left(\frac{1}{0.7 + 0.25n_{P_k}} + (1 - \hat{p}(y = B_j | \mathbf{x} = P_k)) (1 + 0.2347n_{P_k}^{0.1536} e^{-0.00375n_{P_k}}) - 0.571(1 - \hat{p}^2(y = B_j | \mathbf{x} = P_k)) \times e^{-0.00317n_{P_k}^2} \right) \right\} \quad (\text{A.2})$$

Dabei wird die abkürzende Schreibweise $n_{P_k} = n(\mathbf{x} = P_k)$ verwendet.

Die Vertrauensgrenzen für $\hat{p}^l(y = B_j | \mathbf{x} = \bar{P}_k)$ und $\hat{p}^u(y = B_j | \mathbf{x} = \bar{P}_k)$ entstehen, indem $n_{\bar{P}_k}$ anstatt n_{P_k} sowie $\hat{p}(y = B_j | \mathbf{x} = P_k)$ anstatt $\hat{p}(y = B_j | \mathbf{x} = \bar{P}_k)$ eingesetzt werden.

A.7 Regelbasis nach Pruning

R_1 :	WENN (R-HIP Flex ANG V=NEG \cup ZE) \cap (L-KNEE Flex ANG V=NG \cup NK)	DANN y =TST-R
R_2 :	WENN (L-HIP Flex ANG V=ZE \cup PK \cup PG) \cap (L-KNEE Flex ANG V=NG \cup NK)	DANN y =TST-R
R_3 :	WENN (L-HIP Flex ANG=ZE \cup PK) \cap (R-HIP Flex ANG V=NEG \cup ZE \cup PK) \cap R-KNEE Flex ANG V=NK	DANN y =TSW-R
R_4 :	WENN (R-HIP Flex ANG V=NEG \cup ZE \cup PK) \cap R-KNEE Flex ANG V=NK \cap (L-HIP Flex ANG V=NG \cup NK \cup ZE)	DANN y =TSW-R
R_5 :	WENN R-HIP Flex ANG V=NEG \cap (L-HIP Flex ANG V=PK \cup PG) \cap (L-KNEE Flex ANG V=ZE \cup PK \cup PG)	DANN y =MST-R
R_6 :	WENN (L-HIP Flex ANG=PM \cup PG \cup PSG) \cap (L-HIP Flex ANG V=ZE \cup PK) \cap (L-KNEE Flex ANG V=NG \cup NK \cup ZE)	DANN y =TST-R
R_7 :	WENN (R-KNEE Flex ANG V=ZE \cup PK) \cap (L-KNEE Flex ANG V=NG \cup NK)	DANN y =TST-R
R_8 :	WENN (L-HIP Flex ANG V=PK \cup PG) \cap (L-KNEE Flex ANG V=ZE \cup PK)	DANN y =MST-R
R_9 :	WENN R-HIP Flex ANG=ZE \cap (L-KNEE Flex ANG=ZE \cup PK \cup PM) \cap (R-HIP Flex ANG V=NEG \cup ZE) \cap (L-KNEE Flex ANG V=ZE \cup PK)	DANN y =PSW-R
R_{10} :	WENN (R-HIP Flex ANG V=ZE \cup PK) \cap R-KNEE Flex ANG V=NK	DANN y =TSW-R
R_{11} :	WENN L-HIP Flex ANG=ZE \cap (R-KNEE Flex ANG V=NG \cup NK)	DANN y =TSW-R
R_{12} :	WENN (R-HIP Flex ANG=PM \cup PG \cup PSG) \cap R-HIP Flex ANG V=ZE \cap (R-KNEE Flex ANG V=ZE \cup PK \cup PG)	DANN y =LRE-R
R_{13} :	WENN (L-HIP Abd ANG=NG \cup NK) \cap R-HIP Flex ANG V=NEG \cap (L-KNEE Flex ANG V=ZE \cup PK)	DANN y =MST-R
R_{14} :	WENN R-HIP Flex ANG V=ZE \cap (R-KNEE Flex ANG V=NG \cup NK)	DANN y =TSW-R
R_{15} :	WENN L-HIP Flex ANG=ZE \cap L-ANK Flex ANG=PG \cap (R-KNEE Flex ANG V=ZE \cup PK \cup PG)	DANN y =LRE-R
R_{16} :	WENN (R-HIP Flex ANG=ZE \cup PK) \cap (R-KNEE Flex ANG=ZE \cup PK \cup PM) \cap R-KNEE Flex ANG V=PK \cap (L-KNEE Flex ANG V=PK \cup PG)	DANN y =PSW-R
R_{17} :	WENN L-HIP Flex ANG V=PG \cap (L-KNEE Flex ANG V=ZE \cup PK \cup PG)	DANN y =MST-R
R_{18} :	WENN (L-HIP Flex ANG=ZE \cup PK \cup PM) \cap L-HIP Flex ANG V=PG	DANN y =MST-R
R_{19} :	WENN L-KNEE Flex ANG V=NK	DANN y =TST-R
R_{20} :	WENN R-KNEE Flex ANG V=NK \cap (L-HIP Flex ANG V=NK \cup ZE)	DANN y =TSW-R
R_{21} :	WENN L-HIP Flex ANG=ZE \cap (L-KNEE Flex ANG=ZE \cup PK) \cap (R-KNEE Flex ANG V=PK \cup PG)	DANN y =LRE-R
R_{22} :	WENN R-HIP Flex ANG=ZE \cap (R-HIP Flex ANG V=NEG \cup ZE) \cap L-KNEE Flex ANG V=PK	DANN y =PSW-R
R_{23} :	WENN R-HIP Flex ANG=ZE \cap (L-KNEE Flex ANG=ZE \cup PK) \cap L-KNEE Flex ANG V=PK	DANN y =PSW-R
R_{24} :	WENN R-KNEE Flex ANG=PSG \cap L-ANK Flex ANG=PK	DANN y =MSW-R
R_{25} :	WENN (L-HIP Abd ANG=NK \cup ZE \cup PK) \cap R-HIP Flex ANG V=NEG \cap (L-KNEE Flex ANG V=NG \cup NK \cup ZE)	DANN y =TST-R
R_{26} :	WENN (R-KNEE Flex ANG V=ZE \cup PK) \cap L-HIP Flex ANG V=PG	DANN y =MST-R
R_{27} :	WENN (L-HIP Flex ANG=NICHT ZE) \cap (R-HIP Flex ANG V=PK \cup PM \cup PG) \cap R-KNEE Flex ANG V=PK	DANN y =ISW-R
R_{28} :	WENN (R-HIP Flex ANG=PM \cup PG \cup PSG) \cap (R-HIP Flex ANG V=PM \cup PG)	DANN y =MSW-R
R_{29} :	WENN (L-HIP Abd ANG=NICHT NG) \cap L-HIP Flex ANG=ZE \cap R-KNEE Flex ANG V=PK	DANN y =LRE-R
R_{30} :	WENN (R-KNEE Flex ANG=ZE \cup PK) \cap L-HIP Flex ANG=ZE \cap (R-KNEE Flex ANG V=PK \cup PG)	DANN y =LRE-R
R_{31} :	WENN (L-HIP Flex ANG=ZE \cup PK \cup PM) \cap R-HIP Flex ANG V=NEG \cap (L-KNEE Flex ANG V=ZE \cup PK)	DANN y =MST-R
R_{32} :	WENN L-HIP Abd ANG=NG	DANN y =MST-R

R_{33} :	WENN R-HIP Flex ANG=ZE \cap L-HIP Flex ANG V=ZE \cap (L-KNEE Flex ANG V=PK \cup PG)	DANN y=PSW-R
R_{34} :	WENN (R-HIP Flex ANG=PM \cup PG \cup PSG) \cap (L-KNEE Flex ANG=PM \cup PG) \cap (R-HIP Flex ANG V=NEG \cup ZE) \cap (R-KNEE Flex ANG V=ZE \cup PK)	DANN y=MST-R
R_{35} :	WENN (R-HIP Flex ANG=PM \cup PG) \cap (R-HIP Flex ANG V=PK \cup PM)	DANN y=MSW-R
R_{36} :	WENN R-HIP Flex ANG=ZE \cap (R-HIP Flex ANG V=PK \cup PM \cup PG)	DANN y=ISW-R
R_{37} :	WENN (R-HIP Flex ANG=PM \cup PG \cup PSG) \cap L-HIP Flex ANG V=PG	DANN y=MST-R
R_{38} :	WENN (R-HIP Flex ANG=PM \cup PG) \cap (L-HIP Abd ANG=NG \cup NK) \cap (L-ANK Flex ANG=NG \cup NK \cup ZE)	DANN y=MST-R
R_{39} :	WENN (R-HIP Flex ANG=PK \cup PM) \cap (R-HIP Flex ANG V=PK \cup PM)	DANN y=MSW-R
R_{40} :	WENN (R-HIP Flex ANG=ZE \cup PK) \cap (R-HIP Flex ANG V=PM \cup PG)	DANN y=ISW-R
R_{41} :	WENN (R-HIP Flex ANG=PK \cup PM \cup PG) \cap (L-HIP Abd ANG=NG \cup NK) \cap L-KNEE Flex ANG V=PK	DANN y=MST-R
R_{42} :	WENN (L-HIP Flex ANG=PM \cup PG \cup PSG) \cap R-KNEE Flex ANG V=ZE	DANN y=TST-R
R_{43} :	WENN (L-HIP Flex ANG=PM \cup PG \cup PSG) \cap (L-KNEE Flex ANG=PM \cup PG) \cap (R-KNEE Flex ANG V=PK \cup PG)	DANN y=ISW-R
R_{44} :	WENN (R-HIP Flex ANG=PM \cup PG \cup PSG) \cap L-HIP Flex ANG V=ZE \cap (L-KNEE Flex ANG V=PK \cup PG)	DANN y=LRE-R
R_{45} :	WENN R-KNEE Flex ANG=PSG \cap R-HIP Flex ANG V=PK	DANN y=MSW-R
R_{46} :	WENN (R-KNEE Flex ANG=PG \cup PSG) \cap (L-HIP Flex ANG=NICHT ZE) \cap R-KNEE Flex ANG V=PK	DANN y=ISW-R
R_{47} :	WENN R-HIP Flex ANG V=NEG \cap L-KNEE Flex ANG V=PK	DANN y=MST-R
R_{48} :	WENN R-HIP Flex ANG=PK \cap (R-HIP Flex ANG V=NEG \cup ZE) \cap (L-KNEE Flex ANG V=NG \cup NK \cup ZE)	DANN y=TST-R
R_{49} :	WENN R-HIP Flex ANG=ZE \cap (R-KNEE Flex ANG=PG \cup PSG)	DANN y=ISW-R
R_{50} :	WENN L-HIP Flex ANG=ZE \cap (L-ANK Flex ANG=PK \cup PG) \cap (R-KNEE Flex ANG V=ZE \cup PK)	DANN y=LRE-R
R_{51} :	WENN R-KNEE Flex ANG=PSG \cap L-HIP Flex ANG=PK	DANN y=MSW-R
R_{52} :	WENN (R-HIP Flex ANG=PK \cup PM) \cap (L-KNEE Flex ANG V=PK \cup PG)	DANN y=MST-R
R_{53} :	WENN R-KNEE Flex ANG V=PG	DANN y=ISW-R
R_{54} :	WENN (L-HIP Flex ANG=PM \cup PG) \cap (L-KNEE Flex ANG=NICHT NEG) \cap (L-HIP Flex ANG V=ZE \cup PK)	DANN y=TST-R
R_{55} :	WENN R-KNEE Flex ANG=PM \cap (L-HIP Abd ANG=NG \cup NK)	DANN y=MST-R
R_{56} :	WENN L-HIP Flex ANG=PSG \cap (R-HIP Flex ANG V=NEG \cup ZE) \cap (L-KNEE Flex ANG V=ZE \cup PK \cup PG)	DANN y=PSW-R
R_{57} :	WENN L-HIP Flex ANG=ZE \cap L-KNEE Flex ANG=ZE \cap (R-KNEE Flex ANG V=ZE \cup PK \cup PG)	DANN y=LRE-R
R_{58} :	WENN (L-HIP Abd ANG=NK \cup ZE) \cap L-HIP Flex ANG V=PG	DANN y=MST-R
R_{59} :	WENN (L-HIP Flex ANG=PG \cup PSG) \cap R-HIP Flex ANG V=ZE	DANN y=PSW-R
R_{60} :	WENN L-HIP Flex ANG=ZE \cap R-HIP Flex ANG V=NEG \cap (L-HIP Flex ANG V=NK \cup ZE)	DANN y=LRE-R
R_{61} :	WENN (R-HIP Flex ANG=PM \cup PG \cup PSG) \cap L-HIP Abd ANG=NK \cap R-HIP Flex ANG V=NEG	DANN y=MST-R
R_{62} :	WENN (L-HIP Flex ANG=PM \cup PG \cup PSG) \cap (L-ANK Flex ANG=NK \cup ZE \cup PK) \cap R-HIP Flex ANG V=ZE \cap (L-KNEE Flex ANG V=ZE \cup PK \cup PG)	DANN y=PSW-R
R_{63} :	WENN R-HIP Flex ANG=ZE \cap (R-KNEE Flex ANG=PK \cup PM) \cap (L-ANK Flex ANG=ZE \cup PK) \cap R-HIP Flex ANG V=ZE	DANN y=PSW-R
R_{64} :	WENN R-HIP Flex ANG=ZE \cap R-KNEE Flex ANG=ZE	DANN y=TST-R
R_{65} :	WENN (R-KNEE Flex ANG=NICHT ZE) \cap (L-HIP Flex ANG=PK \cup PM \cup PG) \cap (R-HIP Flex ANG V=ZE \cup PK \cup PM) \cap R-KNEE Flex ANG V=PK	DANN y=ISW-R
R_{66} :	WENN (R-HIP Flex ANG=PG \cup PSG) \cap L-HIP Flex ANG=PK \cap (R-HIP Flex ANG V=ZE \cup PK)	DANN y=TSW-R

R_{67} :	WENN R-HIP Flex ANG=PM \cap (L-HIP Abd ANG=NG \cup NK) \cap (L-ANK Flex ANG=NICHT PG)	DANN y=MST-R
R_{68} :	WENN R-KNEE Flex ANG=PSG \cap L-HIP Flex ANG=ZE	DANN y=MSW-R
R_{69} :	WENN L-HIP Flex ANG=PK \cap R-HIP Flex ANG V=NEG	DANN y=MST-R
R_{70} :	WENN (R-KNEE Flex ANG=ZE \cup PK) \cap L-HIP Flex ANG V=PK	DANN y=TST-R
R_{71} :	WENN (L-KNEE Flex ANG V=NG \cup NK)	DANN y=TST-R
R_{72} :	WENN L-HIP Flex ANG V=PG	DANN y=MST-R
R_{73} :	WENN R-HIP Flex ANG V=PM \cap (R-KNEE Flex ANG V=PK \cup PG)	DANN y=ISW-R
R_{74} :	WENN (R-KNEE Flex ANG=PK \cup PM \cup PG) \cap (R-HIP Flex ANG V=PK \cup PM \cup PG) \cap R-KNEE Flex ANG V=PK	DANN y=ISW-R
R_{75} :	WENN L-HIP Flex ANG=PK \cap L-ANK Flex ANG=PG	DANN y=TSW-R
R_{76} :	WENN (R-HIP Flex ANG=PG \cup PSG) \cap (L-HIP Flex ANG=ZE \cup PK) \cap R-HIP Flex ANG V=PK	DANN y=TSW-R
R_{77} :	WENN R-HIP Flex ANG=ZE \cap R-KNEE Flex ANG=PK \cap R-HIP Flex ANG V=ZE	DANN y=PSW-R
R_{78} :	WENN (R-KNEE Flex ANG=PG \cup PSG) \cap L-KNEE Flex ANG=NEG	DANN y=MSW-R
R_{79} :	WENN L-HIP Flex ANG=PM \cap (L-HIP Flex ANG V=ZE \cup PK)	DANN y=TST-R
R_{80} :	WENN R-HIP Flex ANG=ZE \cap R-KNEE Flex ANG=PM \cap (R-HIP Flex ANG V=NEG \cup ZE)	DANN y=PSW-R
R_{81} :	WENN R-KNEE Flex ANG V=NG	DANN y=TSW-R
R_{82} :	WENN R-HIP Flex ANG=PG \cap (R-HIP Flex ANG V=NEG \cup ZE \cup PK) \cap L-HIP Flex ANG V=NK	DANN y=TSW-R
R_{83} :	WENN (R-KNEE Flex ANG=NICHT ZE) \cap L-ANK Flex ANG=NK \cap (R-HIP Flex ANG V=NICHT NEG)	DANN y=ISW-R
R_{84} :	WENN R-HIP Flex ANG V=PM	DANN y=MSW-R
R_{85} :	WENN (R-HIP Flex ANG=PG \cup PSG) \cap L-HIP Flex ANG=PM	DANN y=MSW-R
R_{86} :	WENN R-HIP Flex ANG=PSG \cap R-HIP Flex ANG V=PK	DANN y=TSW-R
R_{87} :	WENN L-HIP Flex ANG=PK \cap R-KNEE Flex ANG V=PK	DANN y=ISW-R
R_{88} :	WENN L-HIP Flex ANG=ZE \cap R-KNEE Flex ANG V=PK	DANN y=LRE-R
R_{89} :	WENN (R-KNEE Flex ANG=PK \cup PM) \cap (L-HIP Flex ANG=PK \cup PM) \cap L-KNEE Flex ANG=NEG	DANN y=ISW-R
R_{90} :	WENN R-HIP Flex ANG V=ZE \cap (L-HIP Flex ANG V=PK \cup PG) \cap (L-KNEE Flex ANG V=NG \cup NK \cup ZE)	DANN y=TST-R
R_{91} :	WENN L-HIP Flex ANG=ZE \cap R-HIP Flex ANG V=NEG \cap L-KNEE Flex ANG V=ZE	DANN y=LRE-R
R_{92} :	WENN (R-KNEE Flex ANG=PK \cup PM) \cap L-KNEE Flex ANG=NEG \cap (L-ANK Flex ANG=NK \cup ZE)	DANN y=ISW-R
R_{93} :	WENN (R-KNEE Flex ANG=PK \cup PM) \cap L-KNEE Flex ANG=NEG \cap (R-KNEE Flex ANG V=ZE \cup PK)	DANN y=ISW-R
R_{94} :	WENN (R-HIP Flex ANG V=PK \cup PM \cup PG) \cap R-KNEE Flex ANG V=PK	DANN y=ISW-R
R_{95} :	WENN R-HIP Flex ANG=PG \cap R-HIP Flex ANG V=PK	DANN y=TSW-R
R_{96} :	WENN R-HIP Flex ANG V=PG	DANN y=ISW-R
R_{97} :	WENN R-HIP Flex ANG=PK \cap R-HIP Flex ANG V=ZE	DANN y=MST-R
R_{98} :	WENN L-HIP Flex ANG=ZE \cap (R-KNEE Flex ANG V=PK \cup PG)	DANN y=LRE-R
R_{99} :	WENN (R-HIP Flex ANG=PG \cup PSG) \cap (L-ANK Flex ANG=NG \cup NK) \cap (L-HIP Flex ANG V=NK \cup ZE \cup PK)	DANN y=LRE-R
R_{100} :	WENN R-HIP Flex ANG=PG \cap R-HIP Flex ANG V=ZE	DANN y=LRE-R
R_{101} :	WENN R-HIP Flex ANG=ZE \cap L-KNEE Flex ANG V=PK	DANN y=PSW-R
R_{102} :	WENN (R-HIP Flex ANG=PM \cup PG \cup PSG) \cap R-HIP Flex ANG V=ZE	DANN y=TSW-R
R_{103} :	WENN (L-HIP Flex ANG=PM \cup PG) \cap R-HIP Flex ANG V=NEG \cap L-KNEE Flex ANG V=ZE	DANN y=MST-R
R_{104} :	WENN R-KNEE Flex ANG V=NK	DANN y=TSW-R
R_{105} :	WENN R-HIP Flex ANG=ZE \cap L-ANK Flex ANG=NG	DANN y=TST-R
R_{106} :	WENN R-HIP Flex ANG=ZE \cap R-KNEE Flex ANG=ZE \cap L-KNEE Flex ANG=NEG \cap (R-HIP Flex ANG V=NEG \cup ZE)	DANN y=PSW-R
R_{107} :	WENN R-HIP Flex ANG=PK \cap (L-HIP Abd ANG=NG \cup NK) \cap L-HIP Flex ANG V=ZE	DANN y=MST-R

Regel-Nr.	B_j	Q	Q_P	Q_K	Q_S	Q_G	Fehler	Beispiele
R_1	B_3	0.52	0.52	1.00	0.98	1.00	13	209
R_2	B_3	0.50	0.50	1.00	0.98	1.00	13	202
R_3	B_7	0.47	0.47	1.00	0.98	1.00	21	157
R_4	B_7	0.46	0.46	1.00	0.98	1.00	31	173
R_5	B_2	0.46	0.46	1.00	0.97	1.00	8	159
R_6	B_3	0.45	0.45	1.00	0.97	1.00	19	202
R_7	B_3	0.45	0.45	1.00	0.98	1.00	9	190
R_8	B_2	0.42	0.42	1.00	0.97	1.00	16	179
R_9	B_4	0.39	0.39	1.00	0.96	1.00	15	84
R_{10}	B_7	0.39	0.39	1.00	0.97	1.00	29	143
R_{11}	B_7	0.37	0.37	1.00	0.97	1.00	14	123
R_{12}	B_1	0.37	0.37	1.00	0.95	1.00	11	81
R_{13}	B_2	0.37	0.37	1.00	0.98	1.00	31	185
R_{14}	B_7	0.37	0.37	1.00	0.95	1.00	0	77
R_{15}	B_1	0.36	0.36	1.00	0.95	1.00	7	84
R_{16}	B_4	0.35	0.35	1.00	0.95	1.00	7	74
R_{17}	B_2	0.34	0.34	1.00	0.96	1.00	7	129
R_{18}	B_2	0.33	0.33	1.00	0.96	1.00	11	131
R_{19}	B_3	0.32	0.36	0.99	0.98	1.00	33	218
R_{20}	B_7	0.31	0.31	1.00	0.97	1.00	38	145
R_{21}	B_1	0.31	0.31	1.00	0.94	1.00	1	64
R_{22}	B_4	0.31	0.31	1.00	0.95	1.00	9	69
R_{23}	B_4	0.30	0.30	1.00	0.94	1.00	9	69
R_{24}	B_6	0.30	0.30	1.00	0.96	1.00	27	123
R_{25}	B_3	0.30	0.30	1.00	0.97	1.00	43	187
R_{26}	B_2	0.29	0.29	1.00	0.96	1.00	14	127
R_{27}	B_5	0.29	0.29	1.00	0.97	1.00	34	146
R_{28}	B_6	0.28	0.28	1.00	0.95	1.00	9	89
R_{29}	B_1	0.28	0.28	1.00	0.96	1.00	18	90
R_{30}	B_1	0.28	0.28	1.00	0.93	1.00	2	53
R_{31}	B_2	0.28	0.28	1.00	0.98	1.00	57	202
R_{32}	B_2	0.27	0.27	1.00	0.96	1.00	21	115
R_{33}	B_4	0.27	0.27	1.00	0.94	1.00	13	64
R_{34}	B_2	0.26	0.26	1.00	0.96	1.00	14	128
R_{35}	B_6	0.26	0.28	0.99	0.96	1.00	24	112
R_{36}	B_5	0.26	0.26	1.00	0.95	1.00	14	89
R_{37}	B_2	0.26	0.26	1.00	0.96	1.00	12	117
R_{38}	B_2	0.25	0.25	1.00	0.96	1.00	17	122
R_{39}	B_6	0.25	0.25	1.00	0.95	1.00	9	85
R_{40}	B_5	0.24	0.27	0.99	0.95	1.00	10	91
R_{41}	B_2	0.24	0.24	1.00	0.96	1.00	32	152
R_{42}	B_3	0.24	0.33	0.97	0.97	1.00	58	244
R_{43}	B_5	0.24	0.24	1.00	0.95	1.00	17	84
R_{44}	B_1	0.23	0.23	1.00	0.94	1.00	15	61
R_{45}	B_6	0.22	0.22	1.00	0.94	1.00	12	79
R_{46}	B_5	0.22	0.22	1.00	0.96	1.00	28	118
R_{47}	B_2	0.21	0.23	0.99	0.97	1.00	53	185
R_{48}	B_3	0.21	0.21	1.00	0.95	1.00	13	96
R_{49}	B_5	0.20	0.20	1.00	0.93	1.00	7	65
R_{50}	B_1	0.20	0.42	0.93	0.95	1.00	24	127
R_{51}	B_6	0.20	0.20	1.00	0.93	1.00	17	64
R_{52}	B_2	0.20	0.20	1.00	0.95	1.00	9	97
R_{53}	B_5	0.19	0.19	1.00	0.52	1.00	3	7
R_{54}	B_3	0.19	0.22	0.98	0.96	1.00	40	148
R_{55}	B_2	0.18	0.18	1.00	0.95	1.00	30	112

Regel-Nr.	B_j	Q	Q_P	Q_K	Q_S	Q_G	Fehler	Beispiele
R_{56}	B_4	0.18	0.18	1.00	0.90	1.00	6	37
R_{57}	B_1	0.18	0.18	1.00	0.92	1.00	7	46
R_{58}	B_2	0.18	0.18	1.00	0.93	1.00	14	67
R_{59}	B_4	0.16	0.16	1.00	0.91	1.00	17	43
R_{60}	B_1	0.16	0.16	1.00	0.93	1.00	25	57
R_{61}	B_2	0.15	0.15	1.00	0.94	1.00	16	70
R_{62}	B_4	0.15	0.22	0.96	0.94	1.00	22	67
R_{63}	B_4	0.14	0.14	1.00	0.77	1.00	3	15
R_{64}	B_3	0.14	0.14	1.00	0.94	1.00	21	87
R_{65}	B_5	0.14	0.17	0.98	0.95	1.00	30	91
R_{66}	B_7	0.14	0.14	1.00	0.87	1.00	0	31
R_{67}	B_2	0.14	0.14	1.00	0.94	1.00	17	82
R_{68}	B_6	0.14	0.14	1.00	0.91	1.00	5	49
R_{69}	B_2	0.14	0.14	1.00	0.91	1.00	3	51
R_{70}	B_3	0.13	0.18	0.97	0.95	1.00	36	138
R_{71}	B_3	0.13	0.40	0.89	0.96	1.00	33	230
R_{72}	B_2	0.13	0.30	0.92	0.94	1.00	28	162
R_{73}	B_5	0.13	0.13	1.00	0.92	1.00	16	54
R_{74}	B_5	0.13	0.13	1.00	0.92	1.00	11	50
R_{75}	B_7	0.12	0.12	1.00	0.86	1.00	3	29
R_{76}	B_7	0.12	0.12	1.00	0.93	1.00	17	57
R_{77}	B_4	0.11	0.11	1.00	0.89	1.00	18	33
R_{78}	B_6	0.11	0.11	1.00	0.92	1.00	8	57
R_{79}	B_3	0.11	0.11	1.00	0.93	1.00	14	71
R_{80}	B_4	0.11	0.11	1.00	0.87	1.00	10	29
R_{81}	B_7	0.11	0.11	1.00	0.49	1.00	3	7
R_{82}	B_7	0.09	0.09	1.00	0.91	1.00	14	44
R_{83}	B_5	0.09	0.09	1.00	0.88	1.00	12	32
R_{84}	B_6	0.08	0.18	0.92	0.93	1.00	41	104
R_{85}	B_6	0.08	0.08	1.00	0.90	1.00	11	44
R_{86}	B_7	0.07	0.07	1.00	0.88	1.00	11	35
R_{87}	B_5	0.06	0.06	1.00	0.88	1.00	8	32
R_{88}	B_1	0.06	0.25	0.86	0.92	1.00	28	103
R_{89}	B_5	0.06	0.06	1.00	0.81	1.00	2	21
R_{90}	B_3	0.06	0.06	1.00	0.81	1.00	5	23
R_{91}	B_1	0.05	0.05	1.00	0.72	1.00	2	13
R_{92}	B_5	0.05	0.05	1.00	0.85	1.00	5	27
R_{93}	B_5	0.05	0.05	1.00	0.85	1.00	5	26
R_{94}	B_5	0.04	0.22	0.85	0.93	1.00	59	174
R_{95}	B_7	0.04	0.04	0.99	0.85	1.00	11	25
R_{96}	B_5	0.03	0.12	0.86	0.89	1.00	25	77
R_{97}	B_2	0.02	0.02	1.00	0.46	1.00	2	7
R_{98}	B_1	0.02	0.25	0.79	0.90	1.00	35	111
R_{99}	B_1	0.02	0.02	1.00	0.42	1.00	1	6
R_{100}	B_1	0.02	0.08	0.86	0.87	1.00	24	50
R_{101}	B_4	0.01	0.24	0.75	0.89	1.00	38	107
R_{102}	B_7	0.01	0.18	0.76	0.91	1.00	80	171
R_{103}	B_2	0.01	0.02	0.90	0.76	1.00	6	21
R_{104}	B_7	0.01	0.24	0.70	0.93	1.00	92	234
R_{105}	B_3	0.01	0.01	0.98	0.44	1.00	1	7
R_{106}	B_4	0.01	0.02	0.90	0.15	1.00	3	4
R_{107}	B_2	0.00	0.01	0.72	0.34	1.00	3	9

Anhang B

Implementierung

B.1 Funktion compute_features.m

```
function Y=compute_features(R_HIP_Flex,L_HIP_Flex,R_KNEE_Flex,L_KNEE_Flex);
%function Y=compute_features(R_HIP_Flex,L_HIP_Flex,R_KNEE_Flex,L_KNEE_Flex);
%computes features for 2D-projection of strides for left and right side.
%input vectors (N - number of sample rows, one column):
%R_HIP_Flex - flexion angle of the right hip (sagittal plane) in [°]
%L_HIP_Flex - flexion angle of the left hip (sagittal plane) in [°]
%R_KNEE_Flex - flexion angle of the right knee (sagittal plane) in [°]
%L_KNEE_Flex - flexion angle of the left knee (sagittal plane) in [°]
%output matrix (N - number of sample rows, 4 columns):
%Y - transformed features
% 1. column: 1. feature (right side)
% 2. column: 2. feature (right side)
% 3. column: 1. feature (left side)
% 4. column: 2. feature (left side)
%typical plot:
%plot(Y(:,1),Y(:,2),':');hold on;plot(Y(:,3),Y(:,4),'-')
%Copyright by Ralf Mikut, Forschungszentrum Karlsruhe, IAI, E-Mail: mikut@iai.fzk.de
%Version 1.0 (25.10.2000)
%
%Background:
%The function estimates angular velocities for all input features and uses the 8
%measured and estimated variables (knee + hip flexion left and right and its angular velocities)
%to project it in an optimal way for different gait phases. The criterion bases on a
%automatic feature selection and a discriminant analysis. In the resulting plane,
%the loading response is normally at the right-hand side of the projection. The stride
%consists of a clockwise rotation via the gait phases (mid and terminal stance are below,
%pre- and initial swing are on the left side, mid swing is on the top and terminal
%swing is top right) returning to the loading response. Interesting evaluations are
%differences between the body sides of one person, different persons, different velocities
%and so on. The location of differences show the gait phases of the differences.

%collect all values in one matrix X
X=[R_HIP_Flex L_HIP_Flex R_KNEE_Flex L_KNEE_Flex];
```



```
%compute velocities and append to matrix X
X=[X(2:size(X,1)-1,:) 0.5*(X(3:size(X,1),:)-X(1:size(X,1)-2,:))];

%prepare transformation matrix for right side
%(result of automatic feature selection and discriminant analysis)
phi_right=[ 0.4399 -0.2954 -0.1643 0.0000 0.0000 0.7849 -0.0071 -0.2760;
-0.0209 -0.0366 0.0908 0.0000 0.0000 -0.9538 -0.2111 0.1890]';

%prepare transformation matrix for left side - almost the same as for right side
%but with switched variables left ;-; right
phi_left = phi_right([2 1 4 3 6 5 8 7],:);

%compute the transformed features
Y=[X*phi_right X*phi_left];
```

B.2 Wichtige Symbole und Bezeichnungen

Anmerkung: Ein \hat{x} über einem Symbol x zeigt immer an, dass es sich um eine Schätzung handelt.

Symbol	Bezeichnung
$\mathbf{0}$	Matrix mit Null-Elementen
$\mathbf{1}$	Matrix mit Eins-Elementen
$a_{l,i}$	Parameter der Zugehörigkeitsfunktion des Terms $A_{l,i}$ ($i = 1$: rechtes Maximum Trapez-ZGF, $i = m_l$: linkes Maximum Trapez-ZGF, $i = 2, \dots, m_l - 1$: Maximum Dreieck-ZGF)
\mathbf{a}_l	Vektor der Parameter der Zugehörigkeitsfunktionen aller Terme des Merkmals x_l
$A_{l,i}$	i -ter linguistischer Term des l -ten Merkmals x_l
A_{l,R_k}	ODER-Verknüpfung linguistischer Terme des l -ten Merkmals x_l in der Teilprämisse der k -ten Regel
b_j	Parameter der Zugehörigkeitsfunktion des Terms B_j
\mathbf{b}	Vektor der Parameter der Zugehörigkeitsfunktionen aller Terme der Ausgangsgröße y
\mathbf{B}	Zwischenklassenvariationsmatrix der Merkmale - Dimension (s, s)
$\tilde{\mathbf{B}}$	Zwischenklassenvariationsmatrix der ausgewählten Merkmale - Dimension (s_m, s_m)
B_j	j -ter linguistischer Term der Ausgangsgröße y
B_{R_k}	Konklusion der k -ten Regel
det	Determinante
F_P	Klassifikationsfehler
F_P^0	Klassifikationsfehler der Trivialschätzung
$H(x), H(y)$	Entropie der Größen x (Eingangsentropie) bzw. y (Ausgangsentropie)
$H(x, y)$	Gesamtentropie der Größen x, y
$H(x; y)$	Transinformation
$H(x y)$	Äquivokation (Rückschlussentropie)
$H(y x)$	Irrelevanz
$H_I(x_l)$	Entropie der durch Intervalle diskretisierten Größe x mit $m_l + 1$ Werten
$H_{I,max}(x_l)$	maximal mögliche Entropie der durch Intervalle diskretisierten Größe x mit $m_l + 1$ Werten (bei Gleichverteilung)
I	Interpretierbarkeit
\mathbf{I}	Einheitsmatrix
i, j, k, l	Laufindizes
\mathcal{I}	Indexmenge
K	Faktor
ISW	Initial Swing (5. Schrittphase)
ln	natürlicher Logarithmus
\log_2	Logarithmus zur Basis 2
LRE	Loading Response (1. Schrittphase)
M	Merkmalsrelevanz
m_l	Anzahl der linguistischen Terme des l -ten Merkmals x_l

Symbol	Bezeichnung
m_y	Anzahl der linguistischen Terme (Klassen) der Ausgangsgröße
MST	Mid Stance (2. Schrittphase)
MSW	Mid Swing (6. Schrittphase)
N	Anzahl Messwerte
$n(x_l = A_{l,i})$	Anzahl des Auftretens der Klasse $A_{l,i}$ für das Merkmal x_l
$n(y = B_j)$	Anzahl des Auftretens der Ausgangsklasse B_j
$n(x_l = A_{l,i} \cap y = B_j)$	Anzahl des Auftretens der UND-Verknüpfung von $x_l = A_{l,i}$ und $y = B_j$
N_j	Anzahl Messwerte der j -ten Ausgangsklasse
N_k	Anzahl Messwerte im k -ten Knoten eines Entscheidungsbaums
NEG	linguistischer Term Negativ
NG	linguistischer Term Negativ Groß
NK	linguistischer Term Negativ Klein
NM	linguistischer Term Negativ Mittel
NSG	linguistischer Term Negativ Sehr Groß
P	Prämissenmatrix (r Zeilen, N Spalten)
P'	Prämissenmatrix zur Berücksichtigung von Redundanzen
PG	linguistischer Term Positiv Groß
PK	linguistischer Term Positiv Klein
POS	linguistischer Term Positiv
PM	linguistischer Term Positiv Mittel
PSG	linguistischer Term Positiv Sehr Groß
PSW	Pre Swing (4. Schrittphase)
$p(x_l = A_{l,i})$	relative Häufigkeit des Auftretens der Klasse $A_{l,i}$ für das Merkmal x_l
$p(\mathbf{x} y = B_j)$	bedingte relative Häufigkeit von \mathbf{x} für die Ausgangsklasse $y = B_j$ (\mathbf{x} mehrdimensional und reell)
$p(y = B_j \mathbf{x})$	bedingte relative Häufigkeit der Ausgangsklasse $y = B_j$ für \mathbf{x} (\mathbf{x} mehrdimensional und reell)
$p(y = B_j)$	relative Häufigkeit der Ausgangsklasse $y = B_j$
$p(x_l = A_{l,i} \cap y = B_j)$	relative Häufigkeit der UND-Verknüpfung von $x_l = A_{l,i}$ und $y = B_j$
$p(x_l = I_{l,i})$	relative Häufigkeit, dass der Wert des Merkmals x_l im Intervall $I_{l,i}$ liegt
$\hat{\mathbf{p}}(y x_l)$	Matrix der bedingten relativen Häufigkeiten für die Ausgangsklassen von y für die m_l Klassen des Merkmals x_l
$p^l()$	untere Schranke einer relativen Häufigkeit
$p^u()$	obere Schranke einer relativen Häufigkeit
Q	Gütefunktion
Q_D	Klarheit der Defaultregel
Q_K	Klarheit
Q_P	Klassifikationsgüte
Q_S	statistische Absicherungsgüte
r	Anzahl der Regeln
\mathbf{r}_0	Regelvektor der Trivialschätzung (Regel mit Prämisse Eins)
R	Regelmatrix, die Plausibilitäten der Konklusionen enthält (m_y Zeilen, r Spalten)

Symbol	Bezeichnung
R_k	k -te Regel
$R_k^{j,i}$	j -te Regel, die im i -ten Pruningschritt der k -ten Regel entsteht
s	Anzahl der Merkmale
s_m	Anzahl der ausgewählten Merkmale
s_d	Anzahl der transformierten Merkmale
S	Schätzung der Kovarianzmatrix der Merkmale
\tilde{S}	Schätzung der Kovarianzmatrix der ausgewählten Merkmale
S_j	Schätzung der Kovarianzmatrix der Merkmale, in die nur Beispiele der j -ten Ausgangsklasse eingehen
\tilde{S}_j	Schätzung Kovarianzmatrix der ausgewählten Merkmale, in die nur Beispiele der j -ten Ausgangsklasse eingehen
T	Gesamtvariationsmatrix der Merkmale – Dimension (s, s)
\tilde{T}	Gesamtvariationsmatrix der ausgewählten Merkmale – Dimension (s_m, s_m)
TST	Terminal Stance (3. Schrittphase)
TSW	Terminal Swing (7. Schrittphase)
v	Eigenvektor
\tilde{v}	Eigenvektor (Hauptkomponentenanalyse)
V	Transformationsmatrix zur Dimensionsreduzierung (im Kapitel Statistische Verfahren)
V	Redundanzmatrix (im Kapitel Fuzzy-Verfahren)
\tilde{V}	Transformationsmatrix zur Merkmalsselektion
V_{ges}	Gesamttransformationsmatrix mit Merkmalsselektion und Dimensionsreduzierung
x_l	l -tes Merkmal
$x_{i,l}$	i -tes Beispiel für das l -te Merkmal im Datensatz
$\tilde{x}_{D,l}$	l -tes transformiertes Merkmal (Diskriminanzanalyse)
$\tilde{x}_{H,l}$	l -tes transformiertes Merkmal (Hauptkomponentenanalyse)
\bar{x}	Vektor der Mittelwerte der ausgewählten Merkmale
\bar{x}_j	Vektor der Mittelwerte der ausgewählten Merkmale, in den nur Beispiele der j -ten Ausgangsklasse eingehen
X	Matrix der Merkmale (N Zeilen, s Spalten)
\tilde{X}	Matrix der ausgewählten Merkmale (N Zeilen, s_m Spalten)
\tilde{X}_j	Matrix der ausgewählten Merkmale (N Zeilen, s_m Spalten), in die nur Beispiele der j -ten Ausgangsklasse eingehen
\tilde{X}_D	Matrix der transformierten Merkmale (N Zeilen, s_d Spalten)
v_k	k -ter Knoten eines Entscheidungsbaums
$v_k(E)$	Endknoten eines Entscheidungsbaums
$v_k(E)$	Wurzelknoten eines Entscheidungsbaums
W	Innerklassenvariationsmatrix der Merkmale – Dimension (s, s)
\tilde{W}	Innerklassenvariationsmatrix der ausgewählten Merkmale – Dimension (s_m, s_m)
y	Vektor der Ausgangsgröße (N Zeilen, 1 Spalte)
y_i	i -tes Beispiel für die Ausgangsgröße y im Datensatz

Symbol	Bezeichnung
\mathbf{Z}	Zentriermatrix
ZE	linguistischer Term Null
α	statistisches Signifikanzniveau bzw. Wichtungsfaktor
β	Wichtungsfaktor
λ	Eigenwert
$\hat{\lambda}$	Eigenwert (Hauptkomponentenanalyse)
$\boldsymbol{\mu}_j$	Vektor der Erwartungswerte der Merkmale, in den nur Beispiele der j -ten Ausgangsklasse eingehen (im Kapitel Statistische Verfahren)
$\mu_{A_l,j}(x_{il})$	Zugehörigkeitsgrade zum j -ten linguistischen Term des Merkmals x_l für den i -ten Messwert
$\boldsymbol{\mu}_{A_l}(x_{il})$	Matrix der Zugehörigkeitsgrade zu allen linguistischen Termen des Merkmals x_l für den i -ten Messwert
$\mu_{B_j}(y_i)$	Zugehörigkeitsgrad zum j -ten linguistischen Term der Ausgangsgröße y für den i -ten Messwert
$\boldsymbol{\mu}_B(y_i)$	Vektor der Zugehörigkeitsgrade zu allen linguistischen Termen der Ausgangsgröße y für den i -ten Messwert
$\mu_{K_{k,i}}$	Zugehörigkeitsgrade der Aktivierung der k -ten Regel für den i -ten Messwert (r Zeilen, N Spalten)
$\mu_{P_{k,i}}$	Zugehörigkeitsgrade der Prämissenauswertung der k -ten Regel für den i -ten Messwert (r Zeilen, N Spalten)
$\boldsymbol{\mu}_X$	Matrix der fuzzifizierten Merkmale ($\sum_{l=1}^s m_l$ Zeilen, N Spalten)
$\boldsymbol{\mu}_{x_1}$	Matrix der Zugehörigkeitsgrade zu allen linguistischen Termen des Merkmals x_l für alle N Messwerte
$\boldsymbol{\mu}_X$	Matrix der fuzzifizierten Ausgangsgröße (m_y Zeilen, N Spalten)
σ^2	Streuung
$\boldsymbol{\Sigma}$	Kovarianzmatrix der Merkmale
$\tilde{\boldsymbol{\Sigma}}$	Kovarianzmatrix der ausgewählten Merkmale
$\boldsymbol{\Sigma}_j$	Kovarianzmatrix der Merkmale, in die nur Beispiele der j -ten Ausgangsklasse eingehen
\cup	ODER-Verknüpfung
\cap	UND-Verknüpfung

Literaturverzeichnis

- [1] Abel, R.; Rupp, R.; Siebel, A.; Döderlein, L.; Gerner, H. J.; Mikut, R.; Peter, N.; Bretthauer, G.: Classification of Gait Analysis Data Using Fuzzy Logic Based Rule Generation and Multivariate Statistical Analysis. *Gait and Posture* 12(1) (2000), S. 58 (Abstract ESMAC'2000, Lund).
- [2] Ahrens, H.; Läuter, J.: *Mehrdimensionale Varianzanalyse: Hypothesenprüfung, Dimensionserniedrigung, Diskrimination bei multivariaten Beobachtungen*. Berlin: Akademie-Verlag. 1974.
- [3] Barbeau, H.; Visintin, M.: The Effects of Body Weight Support on the Locomotor Pattern of Spastic paretic patients. *Canadian Journal of Neurological Science* 16 (1989), S. 315–325.
- [4] Barton, G.: Interpretation of Gait Data Using Kohonen Neural Networks. *Gait & Posture* 10 (1999), S. 85–86.
- [5] Barton, J.; Lees, A.: An Application of Neural Networks for Distinguishing Gait Patterns on the Basis of Hip-Knee Joint Angle Diagrams. *Gait & Posture* 5 (1997), S. 28–33.
- [6] Bauer, H.; Schöllhorn, W.: Self-Organizing Maps for the Analysis of Complex Movement Patterns. *Neur. Proc. Let.* 5 (1997), S. 193–199.
- [7] Baumann, J.: Clinical Experience of gait analysis in the management of cerebral palsy. *Prosthet Orthot Int* 8 (1984), S. 29–32.
- [8] Beckers, D.; Deckers, J.: *Ganganalyse und Gangschulung*. Berlin, Heidelberg: Springer-Verlag. 1997.
- [9] Bettenhausen, K. D.: *Automatische Struktursuche für Regler und Strecke. Beiträge zur datengetriebenen Analyse und optimierenden Führung komplexer Prozesse mit Hilfe evolutionärer Methoden und lernfähiger Fuzzy-Systeme*. Dissertation, TH Darmstadt. 1995.
- [10] Bonarini, A.: Evolutionary Learning of General Fuzzy Rules with Biased Evaluation Functions: Competition and Cooperation. In: *Proc. 1st IEEE Conf. on Evolutionary Computation*, Bd. 1, S. 51–56. Piscataway, NJ: IEEE Press. 1994.
- [11] Bontrager, E.: GAIT-ER-AID: An Expert System for Analysis of Gait with Automatic Intelligent Preprocessing of Data. In: *Proc., 4th Annual Symposium on Computer Applications in MED CARE*, S. 625–629. 1990.

- [12] Boyen, X.; Wehenkel, L.: Automatic Induction of Fuzzy Decision Trees and its Application to Power System Security Assessment. *Fuzzy Sets and Systems* 102 (1999), S. 3–19.
- [13] Bradl, U.; Scholle, H. C.; Hefter, H.; Dohle, C.; Freund, H.: Systemanalytische Techniken zur Bewertung motorischer Funktionen. In: *Motodiagnostik - Mototherapie II* (Scholle, H.; Struppler, A.; Freund, H.-J.; Hefter, H.; Schumann, N., Hg.), S. 77–82. Universitätsverlag Jena. 1994.
- [14] Cappozzo, A.: Gait Analysis Methodology. *Human Movement Science* 3 (1984), S. 27–50.
- [15] Chau, T.: A Review of Analytical Techniques for Gait Data. Part 1: Fuzzy, Statistical and Fractal Methods. *Gait and Posture* 13 (1) (2001), S. 49–66.
- [16] Chau, T.: A Review of Analytical Techniques for Gait Data. Part 2: Neural Network and Wavelet Methods. *Gait and Posture* 13 (2) (2001), S. 102 – 120.
- [17] Cios, K. J.; Pedrycz, W.; Swiniarski, R. W.: *Data Mining Methods for Knowledge Discovery*. Boston: Kluwer Academic. 1998.
- [18] Davis, R.: Clinical gait analysis. *Proc IEEE Eng Med Bio* 7 (1988) 3, S. 35–40.
- [19] Deluzio, K.; Wyss, U. P.; Zee, B.; Costigan, P. A.; Sorbie, C.: Principal Component Models of Knee Kinematics and Kinetics: Normal Vs. Pathological Gait Patterns. *Human Movement Science* 16 (1997), S. 201–218.
- [20] Dietz, V.: Neurophysiology of gait disorders: present and future applications. *Electroencephalogr Clin Neurophysiol*. 103 (1997), S. 333–55.
- [21] DIN EN 61131-7, Ausg. 1999-04 Speicherprogrammierbare Steuerungen - Teil 7: Fuzzy-Control-Programmierung (IEC 65B/368/CDV:1998); Dt. Fassung prEN 61131-7:1998.
- [22] Dorgan, S.; Schmidt, G.: Human Walking as a Model for Bipedal Walking Robot Design. In: *Proc. Euromech 375: Biology and Technology of Walking*, S. 260–267. 1998.
- [23] Dorgan, S. J.; Schmidt, G.: Human Walking as a Model for Bipedal Walking Robot Design. In: *Proc. Euromech 375: Biology and Technology of Walking*, S. 260–267.
- [24] Dung, L. T.; Otto, P.: FuzzyOpt – ein Werkzeug zum Entwurf optimaler Fuzzy-Systeme. *Automatisierungstechnik* 45 (1997) 11, S. 555–556.
- [25] Dzierzanowski, J.; Bourne, J.; Shiavi, R.; Sandell, H.; Guy, D.: GAITSPERT: An Expert System for the Evaluation of Abnormal Human Locomotion Arising from Stroke. *IEEE Trans Biomed Eng* 32(11) (1985), S. 935–942.
- [26] Eltze, J.: *Biologisch motivierte Entwicklung einer sechsbeinigen Laufmaschine*, Bd. 110 von *VDI-Fortschritt-Bericht, Reihe 17*. Düsseldorf: VDI-Verlag. 1994.
- [27] Erler, K.; Brückner, L.; Anders, C.; Scholle, H. C.: *Möglichkeiten der Qualitätssicherung in der Bewegungstherapie*, S. 139–143. dvs-Schriftenreihe, Czwalina-Verlag, Hamburg. 1997.

- [28] Fritsch, M.: *Baumorientierte Regel-Induktionsstrategie für das ROSA-Verfahren zur Modellierung komplexer dynamischer Systeme*. Fortschritt-Bericht VDI, Reihe 8, Nr. 565. Düsseldorf: VDI-Verlag. 1996.
- [29] Gage, J. R.: *Gait Analysis in Cerebral Palsy*. London: Mac Keith Press ; Oxford: Blackwell Scientific Publishers; New York: Cambridge University Press. 1991.
- [30] Gioftos, G.; Grieve, D. W.: The Use of Neural Networks to Recognize Patterns of Human Movement. *Clinical Biomechanics* 10 (4) (1995), S. 179–183.
- [31] Grieve, D. W.: Gait Patterns and the Speed of Walking. *J Biomed Eng* 3 (1968), S. 119–122.
- [32] Güth, V.; Abbink, F.; Reuken, R.: Comparison of pre- and postoperative electromyograms in children with cerebral palsy. *Electromyogr Clin Neurophysiol*. 4 (1985), S. 233–43.
- [33] Güth, V.; Steinhausen, D.; Abbink, F.: The investigation of walking of patients with cerebral palsy by the electromyogram (EMG) using surface electrodes. *Electromyogr Clin Neurophysiol* 3 (1984), S. 225–40.
- [34] Hartung, J.; Elpelt, B.; Klösener, K.-H.: *Lehr- und Handbuch der Angewandten Statistik*. München: Oldenbourg, 11. Aufl. 1998.
- [35] Hayashi, I.; Maeda, T.; Bastian, A.; Jain, L. C.: Generation of Fuzzy Decision Trees by Fuzzy ID3 with Adjusting Mechanism of AND/OR Operators. In: *Proc. IEEE Int. Conf. Fuzzy Systems*, S. 681–685. Piscataway, NJ. 1998.
- [36] Herrera, F.; Lozano, M.; Verdegay, J. L.: A Learning Process for Fuzzy Control Rules Using Genetic Algorithms. *Fuzzy Sets and Systems* 100 (1998), S. 143–158.
- [37] Hesse, S.; Uhlenbrock, D.; Sarkodie-Gyan, T.: Gait pattern of severely disabled hemiparetic subjects on a new controlled gait trainer as compared to assisted treadmill walking with partial body weight support. *Clin Rehabil*. 5 (1999), S. 401–10.
- [38] Hoffmann, F.; Pfister, G.: Optimierung hierarchischer Fuzzy-Regler mit Genetischen Algorithmen. In: *Fuzzy Logik: Theorie und Praxis, Proc. 4. Dortmunder Fuzzy-Tage* (Reusch, B., Hg.), S. 97–89. Berlin: Springer. 1994.
- [39] Holland, J. H.; Holyoak, K. J.; Nisbett, K. J.; Thagard, P. R.: *Induction: Processes of Inference*. Cambridge, MA: MIT Press. 1986.
- [40] Holve, R.: “The curse of Dimensionality” – und was man dagegen tun kann ... Hierarchische Fuzzy Systeme zur Musterklassifikation. In: Kiendl und Knicker [59], S. 195–208. 1998.
- [41] Holzreiter, S.; et al.: Assessment of gait patterns using neural networks. *Journal of Biomechanics* 26 (1993) 6, S. 645–51.
- [42] Hong, T. P.; Tseng, S. S.: A Generalised Version Space Learning Algorithm for Noisy and Uncertain Data. *IEEE Trans. Knowledge Data Engng.* 9 (1997), S. 336–340.
- [43] Huang, P.; Harris, C.; Nixon, M. S.: Recognising humans by gait via parametric canonical space. *Artificial Intelligence in Engineering* 13 (1999), S. 359–366.

- [44] Jäkel, J.: *Linguistische Fuzzy-Systeme mit verallgemeinerten Konklusionen und ihre Anwendung zur Modellbildung und Regelung*. Fortschritt-Berichte VDI, Reihe 8, Nr. 793. Düsseldorf: VDI-Verlag. 1999.
- [45] Jäkel, J.; Gröll, L.; Mikut, R.: Bewertungsmaße zum Generieren von Fuzzy-Regeln unter Beachtung linguistisch motivierter Restriktionen. In: Kiendl und Knicker [59], S. 15–28. 1998.
- [46] Jäkel, J.; Gröll, L.; Mikut, R.: Automatic Generation and Evaluation of Interpretable Rule Bases for Fuzzy Systems. In: *Computational Intelligence for Modelling, Control and Automation: Evolutionary Computation and Fuzzy Logic for Intelligent Control, Knowledge Acquisition and Information Retrieval; Proc. Conf., Wien, A, Feb. 17-19, 1999* (Mohammadian, M., Hg.), Bd. 2, S. 192–197. Amsterdam: IOS Press. 1999.
- [47] Jäkel, J.; Gröll, L.; Mikut, R.: Tree-Oriented Hypothesis Generation for Interpretable Fuzzy Rules. In: *Proc. 7th Europ. Congr. on Intelligent Techniques and Soft Computing EUFIT'99, Sep. 13-16, 1999*, S. 279–280. Aachen. CD-ROM. 1999.
- [48] Jäkel, J.; Gröll, L.; Mikut, R.: Automatic Generation and Evaluation of Interpretable Fuzzy Rules. In: *New Frontiers in Computational Intelligence and its Applications* (Mohammadian, M., Hg.), S. 1–10. Amsterdam: IOS Press. 2000.
- [49] Jäkel, J.; Mikut, R.; Malberg, H.; Bretthauer, G.: Datenbasierte Regelsuche für Fuzzy-Systeme mittels baumorientierter Verfahren. In: *Berichtsband 9. Workshop Fuzzy Control d. GMA-FA 5.22*, S. 1–15. Dortmund. 1999.
- [50] Jang, J.-S. R.: Structure Determination in Fuzzy Modeling: A Fuzzy CART Approach. In: *Proc. IEEE Int. Conf. on Fuzzy Systems*, S. 480–485. Orlando, Florida. 1994.
- [51] Janikow, C. Z.: Fuzzy Processing in Decision Trees. In: *Proc. Int. Symp. On Artificial Intelligence*, S. 360–367. Monterrey. 1993.
- [52] Jessen, H.; Slawinski, T.: Test and Rating Strategies for Data Based Rule Generation. Techn. Ber. CI-39/98, Universität Dortmund, Dept. Computer Science. 1998.
- [53] Johnson, K.; Denning, R.; Smith, P.; Smith, J.; Simon, S.: Dr. Gait III - A Multimedia System for Gait Analysis. *Gait & Posture* 4 (1996), S. 195–196.
- [54] Kadaba, M.; Ramakrishnan, H.; Jacobs, D.; Chambers, C.; Scarborough, N.; Goode, B.: Gait pattern recognition in spastic diplegia. *Dev Med Child Neurol* 33(9) (1991), S. 28.
- [55] Karr, C. L.: Design of an Adaptive Fuzzy Logic Controller Using a Genetic Algorithm. In: *Proc. Int. Conf. on Genetic Algorithms*, S. 450–457. San Mateo. 1991.
- [56] Kaufmann, K.; Chao, E.; Callahan, T.; Askew, L.; R.R., B.: Development of a functional performance index for quantitative gait analysis. *Biomed Sci Instrumentation* 23 (1987), S. 49–55.
- [57] Keller, H. B.: Learning Rules for Modelling Dynamic Systems Behaviour. In: *EUROSIM'95* (Breitenecker, F.; Husinsky, I., Hg.), S. 1205–1210. Amsterdam: Elsevier Science. 1995.

- [58] Kiendl, H.: *Fuzzy Control methodenorientiert*. München: Oldenbourg. 1997.
- [59] Kiendl, H.; Knicker, R. (Hg.): *Berichtsband 8. Workshop Fuzzy Control d. GMA-FA 5.22*, Forschungsbericht 0298 d. Fak. Elektrotechnik d. Univ. Dortmund. 1998.
- [60] Klose, A.; Nürnberger, A.: Applying Boolean Transformations to Fuzzy Rule Bases. In: *Proc. 7th Europ. Congr. on Intelligent Techniques and Soft Computing EUFIT'99, Sep. 13-16, 1999*, S. 215–217. Aachen. CD-ROM. 1999.
- [61] Krabs, M.; Kiendl, H.: Automatische Generierung von Fuzzy-Regeln mit dem ROSA-Verfahren. In: *Proc. Fuzzy-Control, GMA-Aussprachetag, Langen 22./23. März 1994*, VDI-Bericht 1113, S. 29–40. Düsseldorf: VDI-Verlag. 1994.
- [62] Krone, A.: *Datenbasierte Generierung von relevanten Fuzzy-Regeln zur Modellierung von Prozesszusammenhängen und Bedienstrategien*. Fortschritt-Berichte VDI, Reihe 10, Nr. 615. Düsseldorf: VDI Verlag. 1999.
- [63] Krone, A.; Bäck, T.; Teuber, P.: Evolutionäres Suchkonzept zum Aufstellen signifikanter Fuzzy-Regeln. *Automatisierungstechnik* 44 (1996) 8, S. 405–411.
- [64] Kühnapfel, M.; Fischer, M. S.: Lokomotion kleiner Säugetiere. *Verh. Dtsch. Zool. Ges.* (1994), S. 87–104.
- [65] Lines, G. T.; Kavli, T.: *The Equivalence of Spline Models and Fuzzy Logic Applied to Model Construction and Interpretation*, Kap. 11. Singapore: World Scientific. 1996.
- [66] Lutzenberger, C.; Pfeiffer, F.: A three-dimensional model of the human locomotor apparatus for gait analysis. In: *Proc., European Mechanics Colloquium*, S. 25–32. München. 1998.
- [67] Mainka, C.; Boenick, U.: Integrated gait analysis for future routine clinical use. *Biomed Tech* 38 (1993) 12, S. 325–31. German.
- [68] Mainka, C.; Friedrich, M.; Boenick, U.: Integrated Gait Analysis: A Complex Expert System Approach To Detailed Data Evaluation. *Biomedizinische Technik* 40 (1995) 12, S. 363–71. German.
- [69] Malberg, H.; Wessel, N.; Schirdewan, A.; Osterziel, K. J.; Voss, A.: Duale Sequenzmethode zur Analyse der spontanen Baroreflexsensitivität bei Patienten mit dilatativer Kardiomyopathie. *Z. Kardiologie* 88 (1999) 5, S. 331–337.
- [70] Menkveld, S. R.; Knipstein, E.; Quinn, J. R.: Analysis of gait patterns in normal schooled children. *J Ped Orthop* 8(3) (1988), S. 263–267.
- [71] Merkle, L.; Layne, C.; Bloomberg, J.; JJ, Z.: Using Factor Analysis To Identify Neuromuscular Synergies During Treadmill Walking. *Journal of Neuroscience Methods* (1998), S. 207–214.
- [72] Meyer, D.: Human Gait Classification Based on Hidden Markov Models. In: *3D Image Analysis and Synthesis '97* (Seidel, H.-P.; Girod, B.; Niemann, H., Hg.), S. 139–146. Erlangen. 1997.

- [73] Meyer, D.; Denzler, J.; Niemann, H.: Model Based Extraction of Articulated Objects in Image Sequences for Gait Analysis. In: *Fourth International Conference on Image Processing*, Bd. 3, S. 78–81. Santa Barbara, Californien, USA. 1997.
- [74] Meyer, D.; Niemann, H.: Automatische Bestimmung der Trajektorien ohne Marker zur Bewegungserkennung von Personen. In: *5. Workshop Digitale Bildverarbeitung in der Medizin*, S. 133–138. Berlin: Universitätsklinikum Benjamin Franklin. 1997.
- [75] Meyer, D.; Niemann, H.: Features for Optical Flow Based Gait Classification Using HMMs. In: *Image and Multidimensional Digital Signal Processing '98* (Niemann, H.; Seidel, H.-P.; Girod, B., Hg.), S. 75–79. Alpbach (Austria). 1998.
- [76] Meyer, D.; Pösl, J.; Niemann, H.: Gait Classification with HMMs for Trajectories of Body Parts Extracted by Mixture Densities. In: *British Machine Vision Conference (BMVC)*, S. 459–468. 1998.
- [77] Mikut, R.; Jäkel, J.; Gröll, L.: Inference Methods for Partially Redundant Rule Bases. In: *Proc. 7th Zittau Fuzzy Colloquium, Sep. 8-10, 1999*, S. 245–251. Hochschule Zittau/Görlitz. 1999.
- [78] Mikut, R.; Jäkel, J.; Gröll, L.: Automatic Design of Interpretable Membership Functions. In: *Proc. 8th Zittau Fuzzy Colloquium, Sep. 6-8, 2000*, S. 103–111. Hochschule Zittau/Görlitz. 2000.
- [79] Mikut, R.; Jäkel, J.; Gröll, L.: Inference Methods for Partially Redundant Rule Bases. In: *Fuzzy Control: Theory and Practice* (Hampel, R.; Wagenknecht, M.; Chaker, N., Hg.), Advances in Soft Computing, S. 177–185. Heidelberg: Physica. 2000.
- [80] Mikut, R.; Jäkel, J.; Gröll, L.: Informationstheoretische Maße zur Merkmalsauswahl, Generierung von Zugehörigkeitsfunktionen und Regeln für Fuzzy-Klassifikatoren. In: *Computational Intelligence und Industrielle Anwendungen, Baden-Baden, 11-12. Mai 2000*, VDI-Berichte 1526, S. 125–130. Düsseldorf: VDI-Verlag. 2000.
- [81] Mikut, R.; Jäkel, J.; Malberg, H.; Bretthauer, G.: Datenbasierter Entwurf von Fuzzy-Systemen für medizinische Diagnoseaufgaben. *Automatisierungstechnik* 48 (2000) 7, S. 317–326.
- [82] Mikut, R.; Peter, N.; Bretthauer, G.; Rupp, R.; Abel, R.; Siebel, A.; Döderlein, A.; Gerner, H.-J.: Fuzzy-Regelgenerierung und Multivariate Statistische Verfahren zur Schritt-Phasenerkennung in der Instrumentellen Ganganalyse. In: *Berichtsband 10. Workshop Fuzzy Control des GMA-UA 5.22*, S. 112–125. Karlsruhe: Forschungszentrum Karlsruhe, FZKA 6509. 2000.
- [83] Nelles, O.; Fischer, M.: Local Linear Model Trees (LOLIMOT) for Nonlinear System Identification of a Cooling Blast. In: *Proc. 4th Europ. Congr. on Intelligent Techniques and Soft Computing EUFIT'96*, S. 1187–1191. Aachen. 1996.
- [84] Ng, S. G.; Chizeck, H. J.: Fuzzy Model Identification for Classification of Gait Events in Paraplegics. *IEEE Transactions on Fuzzy Systems* 5 (4) (1997), S. 536–544.
- [85] N.N.: *OrthoTrack 4.1- Gait Analysis Software, Reference Manual*. MotionAnalysis. 1999.

- [86] Olree, K.; Vaughan, C.: Fundamental patterns of bilateral muscle activity in human locomotion. *Biological Cybernetics* 73 (1995) in press.
- [87] Patla, A.: Some Characteristics Of EMG Patterns During Locomotion: Implications For The Locomotor Control Process. *J Mot Behav* 17 (1985), S. 443–461.
- [88] Perry, J.: *Gait Analysis. Normal and Pathological Function*. Thorofare: Slack Inc. 1992.
- [89] Quinlan, J. R.: Induction of Decision Trees. *Machine Learning* 1 (1986), S. 81–106.
- [90] Quinlan, J. R.: *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann. 1993.
- [91] Riener, R.: *Neurophysiologische und Biomechanische Modellierung zur Entwicklung geregelter Neuroprothesen*. Dissertation, Technische Universität München, Utz-Verlag. 1997.
- [92] Riener, R.; Fuhr, T.: Patient-driven Control Of FES-supported Standing Up: A Simulation Study. *IEEE Transactions on Rehabilitation Engineering* 6 (1998) 2, S. 113–24.
- [93] Riener, R.; Fuhr, T.: Patient-driven control of FES-supported standing up: a simulation study. *Transactions on Rehabilitation Engineering* 6 (1998), S. 113–124.
- [94] Rives, J.: FID3: Fuzzy Induction Decision Tree. In: *Proc. 1st Int. Symp. Uncertainty, Modelling and Analysis*, S. 457–462. Los Alamitos, Calif.: IEEE Computer Soc. Press. 1990.
- [95] Runkler, T.; Roychowdhury, S.: Generating Decision Trees and Membership Functions by Fuzzy Clustering. In: *Proc. 7th Europ. Congr. on Intelligent Techniques and Soft Computing EUFIT'99, Sep. 13-16, 1999*, S. 128–129. Aachen. CD-ROM. 1999.
- [96] Rupp, R.; Schablowski, M.; Gerner, H.: Entwicklung und Evaluierung eines Diagnostiklaufbandes zur dreidimensionalen Erfassung der Gangdynamik. In: *Biomedizinische Technik, Band 43, Ergänzungsband* (Freyer, R.; Rabenau, M., Hg.), S. 192–93. Deutsche Gesellschaft für Biomedizinische Technik, Boenick, U and Schaldach, M. 1998.
- [97] Salatian, A.; Zheng, Y.: Gait Synthesis for a Biped Robot Climbing Surfaces Using Neural Networks - Part I: Static Learning. In: *Proc.; IEEE International Conference on Robotics and Automation*, S. 2601–2606. Nice. 1992.
- [98] Scholle, H. C.; Bradl, U.: *Zentrale Mechanismen der Bewegungsregulation - Einsatz moderner bildgebender Verfahren*, S. 255–270. Leipzig: monade Verlag und Agentur, Rainer Rodewald. 1996.
- [99] Schumann, N.: Grundlagen und Einsatz oberflächenelektromyographischer und biomechanischer Methoden im Rahmen der Diagnostik und Therapieverlaufskontrolle von Schlaganfall-Patienten. In: *Konzepte der Bewegungstherapie nach Schlaganfall* (Seidel, E. e. a., Hg.), S. 97–100. Reihe Praktische Physiotherapie / Sporttherapie, GFBB Bad Kösen. 1995.

- [100] Sepulveda, F.; Wells, D.; Vaughan, C.: A neural network representation of electromyography and joint dynamics in human gait. *Journal of Biomechanics* 26 (1993), S. 101–109.
- [101] Siebel, A.; Berghof, R.; Döderlein, L.: Modification of the Walking Pattern in Patients with Operated Anterior Cruciate Ligament Rupture Measured With and Without a Brace. *Gait and Posture* 3(2) (1995).
- [102] Smith, S. F.: *A Learning System Based on Genetic Adaptive Systems*. Dissertation, Univ. of Pittsburgh. 1980.
- [103] Song, S. M.; Waldron, K. J.: Analytical Approach for Gait Study and its Application on Wave Gaits. *International Journal of Robotics Research* 6(2) (1987), S. 60–71.
- [104] Steiner, R.: *Attraktorregelung für die Fortbewegung einer simulierten zweibeinigen Laufmaschine in drei Dimensionen*. Dissertation, Universität Düsseldorf. 1999.
- [105] Sutherland, D. H.; Kaufman, K. R. And. Wyatt, M. P.; Chambers, H. G.; Mubarak, S. J.: Double-Blind Study of Botulinum A Toxin Injections Into the Gastrocnemius Muscle in Patients with Cerebral Palsy. *Gait & Posture* 10 (1999), S. 1–9.
- [106] The MathWorks Inc., Natick, Mass.: *MATLAB – The Language of Technical Computing. Using MATLAB*. 1999.
- [107] Wang, C.-H.; Hong, T.-P.; Tseng, S.-S.: Inductive Learning from Fuzzy Examples. In: *Proc. the 5th IEEE Int. Conf. on Fuzzy Systems*, S. 13–18. New Orleans, LA, USA. 1996.
- [108] Wang, C.-H.; Liu, J.-F.; Hong, T.-P.; Tseng, S.-S.: A Fuzzy Inductive Learning Strategy for Modular Rules. *Fuzzy Sets and Systems* 103 (1999), S. 91–105.
- [109] Waters, R. L.; Frazier, J.; Garland, D.; Jordan, C.; Perry, J.: Electromyographic gait analysis before and after operative treatment for hemiplegic equinus and equinovarus deformity. *J Bone Joint Surg Am* 64A (1982), S. 284–288.
- [110] Weaver, W.; Shannon, C. E.: *The Mathematical Theory of Communication*. Techn. Ber., University of Illinois Press. 1949.
- [111] Weiss, A.; Krauss, P.; Hoch, G.; Quintern, J.: Design of a multi-sensor closed-loop neurostimulator. Part B: Input and output stage. In: *Neuroprosthetics - from Basic Research to Clinical Applications* (Pedotti, A. E. A., Hg.), S. 189–194. Springer-Verlag. 1996.
- [112] Wernig, A.; S., S. M.: Die Lokomotionstherapie am Laufband bei Querschnittlähmung, Ergebnisse einer fünfjährigen Studie. *Neurologische Rehabilitation* 1 (1995), S. 6–16.
- [113] Winter, D.: Use Of Kinetic Analyses In The Diagnostics Of Pathological Gait. *Physio Canada* 33 (1981), S. 209–214.
- [114] Winter, D.: Kinematic and kinetic patterns in human gait: Variability and compensating effects. *Hum Mvmt Sci* 3 (1984), S. 51–76.
- [115] Zadeh, L. A.: Fuzzy Sets. *Information and Control* 8 (1965), S. 338–353.

- [116] Zanchi, V.; Papic, V.; Cecic, M.: Quantitative Human Gait Analysis. *Simulation Practice and Theory* 8 (1-2) (2000), S. 127–140.
- [117] Zlatnik, D.: Intelligently Controlled Above Knee (A/K) Prosthesis. In: *Proc. 4th Int. Conf. On Motion and Vibration Control (MOVIC'98)*. Zürich. 1998.

Index

- Äquivokation, 44
- Abduktion, 8
- Adduktion, 8
- Akkumulation, 37
- Aktivierung, 37
- ANOVA-Verfahren, 21
- Ausgangsentropie, 43, 49
- Bayes-Klassifikator, 26–28
- C4.5-Verfahren, 52
- Defaultregel, 58–60, 63, 68, 69
- Defuzzifizierung, 38–39
- dimensionsreduzierende Transformation, 22–24
- disjunkte Partitionierung, 43
- Diskriminanzanalyse, 23–26, 72
- Eingangsentropie, 43, 45, 47, 49
- Endknoten, 50, 51
- Entropie, 44
- Entscheidungsbäume, 52, 55, 56
- Extension, 8
- Flexion, 8
- frühkindliche Hirnschädigung, 3, 6
- Frontalebene, 8, 11
- Fuzzifizierung, 34–35, 39, 50
- Fuzzy-Inferenz, 35–38, 58
- Fuzzy-Negation, 36
- Gesamtentropie, 43
- Gesamtvariationsmatrix, 21
- Häufigkeit, relative, 44–46
- Hauptkomponentenanalyse, 23–26, 72
- ID3-Verfahren, 52
- Infantile Zerebralparese (ICP), 3, 9–11
- Informationstheoretische Maße, 43–48, 52
- Initial Contact, 13
- Initial Swing, 14
- Innerklassenvariationsmatrix, 21
- Interpretierbarkeit, 49
- Irrelevanz, 44
- Kinematik, 8
- Kinetik, 9
- Klarheit, 41, 57, 60, 62–64
- Klassifikationsgüte
 - Fuzzy, 39, 54, 57, 58, 60, 62, 68, 72
 - Statistik, 23, 27, 72
- Kovarianzmatrix, 21
- linguistischer Term, 34–36, 43, 49, 57
- Loading Response, 13
- MANOVA-Verfahren, 19, 72
- Merkmalsrelevanz, 22, 52–54, 78, 80–82
- Merkmalsselektion, 19, 22, 54
- Mid Swing, 14
- Midstance, 13
- ODER-Verknüpfung, 36–37
- Prämissenauswertung, 37, 58
- Pre Swing, 14
- Querschnittlähmung, 3, 6
- Raum-Zeit-Parameter, 7
- Redundanzmatrix, 37, 69
- Regelmatrix, 58–60, 63, 68
- Regelplausibilität, 37, 58, 59
- Regelpruning, 57, 62–65
- Rotation, 8
- Sagittalebene, 8, 11
- Schrittphasen, 14
- Schwerpunktmethode, 38
- Schwerpunktmethode für Singletons, 38
- Singletons, 37
- statistische Absicherung, 41, 57, 58, 60–62, 84

- T-Konorm, 36
- T-Norm, 36
- t-Test, 21
- Terminal Stance, 13
- Terminal Swing, 14
- Transinformation, 44, 46, 48, 49, 52
- Transversalebene, 8, 11
- Trivialschätzung, 58

- UND-Verknüpfung, 36–37

- Wurzelknoten, 50, 51

- Zugehörigkeitsfunktion, 34–35, 38–39, 42–
48
- Zwischenklassenvariationsmatrix, 21