Report on the Joint Autumn Meeting of the GfKl Working Groups AG MARKETING and AG DANK

Friederike Paetz, Philipp Aschersleben and Gero Szepannek

Abstract This article reports on the joint working group meeting of the AG MARKETING and AG DANK within the GfKl Data Science Society. The meeting was held from October 7 to 8, 2022, hosted by the Clausthal University of Technology. The presented talks included topics from a great variety of fields from quantitative marketing and data analytics and numerical classification.

Friederike Paetz
Anhalt University of Applied Sciences, Strenzfelder Allee 28, 06406 Bernburg Germany
✉ friederike.paetz@hs-anhalt.de

Philipp Aschersleben
MEDICE Arzneimittel Pütter GmbH & Co. KG, Kuhloweg 37. 58638 Iserlohn Germany
✉ philipp.aschersleben@tu-clausthal.de

Gero Szepannek
University of Applied Sciences Stralsund, Zur Schwedenschanze 15, D-18435 Stralsund Germany
✉ gero.szepannek@hochschule-stralsund.de

ARCHIVES OF DATA SCIENCE, SERIES A
(ONLINE FIRST)
KIT SCIENTIFIC PUBLISHING
Vol. 8, No. 1, 2024

DOI 10.5445/IR/1000170657
ISSN 2363-9881

1 Introduction

Transnational, sustainable and close to business: The joint working group meeting on the topic of *Responsible Data Analysis and Machine Learning* offered international players from science and business a successful stage for mutual knowledge transfer.

For the first time, the two working groups *Data analysis and classification in marketing (AG MARKETING)* and *Data analysis and numerical classification (AG DANK)* met for a joint meeting, which took place on October, 7 to 8, 2022, at Clausthal University of Technology and was hosted by the Clausthal Executive School and the Institute of Management and Economics.

The invitation was accepted by participants from Germany, France, Austria and Switzerland, who discussed the twelve presentations on the topics of data analysis in marketing and data analysis and numerical classification. The lecture contents were very diverse, ranging from start-up survival probabilities (Müller-Funk, U. & Ungerer, C.), instrument recognition (Schulz, J.-P., Szepannek, G., & Harczos, T.), fake news detection (Wilhelm, A. & Dossou, B.F.P.) to medical (Thrun, M.C., Hoffmann, J., Krause, S.W., Krawitz, P., Brendel, C., & Ultsch, A.) and macroeconomic (Zimmermann, T.) use cases. Furthermore, optimization approaches in the fields of production (Krippendorff, N. & Schwindt, C.) and logistics (Haase, K., Kück, J., Sauerbier, F., & Spindler, M.) were discussed and a tool to determine the economic performance of museums (Hildebrand, L. & Paetz, F.) was introduced. Additionally, issues in the fields of customer responses (Kurz, P. as well as Schröder, N., Marra, G., Radice, R. & Reutterer, T.), small sample sizes (May, S.) and random forest models (Szepannek, G. & von Holt, B.-H.) were seized.

Besides these competitive talks, three invited talks complemented the portfolio of the joint meeting: Professor Dr. Raoul V. Kübler (ESSEC Business School Paris) gave a talk on the influence of social media in the past US presidential elections. Moritz von Zahn (Goethe University Frankfurt) discussed the influence of *green nudges* on return behavior in online shopping. In addition, Zoé Wolter and Philipp Bosch presented the association *CorrelAid e. V.*, which carries out data analyses and workshops for non-profits on a voluntary basis.

The two-day meeting concluded with the traditional AG DANK's data analysis competition, in which this year's task was to predict the results of a biathlon sprint race as accurately as possible based on the results of the 20 previous competition results. Here it was shown that even intuitive models, i.e. educated guessing, can lead to very accurate results and beat more complex approaches.

2 Control of Shared Production Buffers: A Reinforcement Learning Approach

Nora Krippendorff, Clausthal University of Technology
Christoph Schwindt, Clausthal University of Technology

We consider a buffer control problem arising in stochastic flow lines with dedicated and shared production buffers. Buffer control relies on decision rules which determine transfers of items between buffers and machines at the release or completion times of parts on the different production stages. We formulate a conceptual model of the problem for a basic scenario with one central buffer and explain how general system configurations and a tactical buffer allocation problem can be modeled within this framework. Under the assumption that the flow line can be represented as a Markovian production system, we provide a formulation as a continuous-time Markov decision problem admitting an optimal stationary policy. By applying the uniformization approach from McMahon (2008), the Markov decision problem is discretized in time and thus amenable to standard algorithms (Puterman, 2005). We propose a simple Q-learning implementation of reinforcement learning converging to an optimal stationary policy and validate the approach in a numerical experiment with a small toy problem (Mitchell, 1997).

3 On the Relevance of Transaction Data: A Funnel Analysis of New Venture Survival

Ulrich Müller-Funk, University of Münster
Christina Ungerer, University of Stuttgart

We study the impact of transaction relations - the relations built by new ventures to customers, partners, financiers, and human resources - on the chances of their survival. The development of start-ups is tracked over a fixed period subdivided into consecutive phases. Survival is viewed as a qualitative feature, analyzed within a state-phase- model via catenated classifiers and multiple tests. No censoring is involved. The model is applied to 482 new technology-based firms. Transaction relations turn out to be features significantly complementing standard factors, with a phase-wise varying importance. For additional information see De Jong and Marsili (2015); Simón-Moya et al (2012); Stinchcombe (1965).

4 Investigating Deep Learning for Auditory Model Based Pitch Classification

Jan-Paul Schulz, Saarland University
Gero Szepannek, Stralsund University of Applied Computer Science
Tamas Harczos, Fraunhofer IDMT

In order to increase the understanding of human sound perception auditory models have been developed to mimic the different steps of sound processing in the auditory system. Recent advances in neural networks have led to their successful application in many different contexts. In this work several established deep learning architectures are investigated with regard their ability to be used for pitch recognition in combination with auditory modelling. For this purpose, pitch estimation was modelled as a classification problem based on cochleograms. A particular emphasis has been laid on the appropriate choice of the model's hyperparameters. The results of the study are promising and can be interpreted as a step towards mimicking human pitch perception. For additional information see Bischl et al (2021); Feldhoff et al (2022); Harczos and Klefenz (2018); Su et al (2016); Vecchi et al (2022).

5 How efficient are German museums in using their public funding? An input-oriented network DEA approach

Lea Hildebrand, Ostfalia University of Applied Science
Friederike Paetz, Clausthal University of Technology

Data Envelopment Analysis (DEA) has become a progressive method of efficiency research. The non-parametric technique allows the performance measurement of peer groups facing production processes that include multiple input-output-structures. Conventional DEA models treat observed productions as black boxes without considering the efficiency of internal structures that facilitate the transformation of inputs into external outputs. In contrast to this, network DEA approaches base on the decomposition of the evaluated production processes. Recent studies of (Barrio-Tellado and Herrero-Prieto, 2019, 2022) first employed network DEA as a fertile methodical approach to measure and compare the (internal) performance of 23 Spanish museums. The present study extends the Spanish research model to analyse the efficiency of 51 German publicly funded museums on a two-stage network structure. Due to the input-orientation of the used model, particular emphasis is placed on the examination of the used public funding throughout the museums' production stages. The study grounds on data of the report year 2019 that was collected from museum websites, annual (financial) museum reporting and budget plans such as budget accounts of the integrated federal states, municipalities, and the national government. The research model assimilates the selected data to establish radial and input-oriented efficiency measures on two constructed production stages under the assumption of variable returns to scale. Thereby, the proposed research model initially evaluates the usage of the allocated public funding in consideration of the gained service ability level represented through a facility index, the personnel expense, weekly opening hours, special exhibitions and produced publications. The second stage assesses the adequacy of the provided service level with reference to the attained annual visitation level as external output of the decomposed production process. The results show substantially higher efficiency scores related to the first stage of the applied research model. Aside from that, museums that are assessed as optimum productions in stage one tend to gain higher efficiency scores in the following production stage. Overall, only three museums are detected as efficient throughout all production stages. The talk further discusses the results of the employed network DEA model in the light of the distribution of efficiency scores gained from the corresponding DEA black box model and broaches the issue of applied model building procedures to classify the observed DEA peer group.

6 Accounting for skewed distributions in modeling self-reported customer response data

Nadine Schröder, Vienna University of Economics and Business, Austria
Giampiero Marra, University College London, United Kingdom
Rosalba Radice, Bayes Business School – City, University of London, United Kingdom
Thomas Reutterer, Vienna University of Economics and Business, Austria

According to previous research (e.g., Peterson and Wilson, 1992) customers who take part in satisfaction surveys, to a vast extent, typically report that they were highly satisfied with a product or service. This phenomenon has gained a lot of attention in the field of product reviews as well. In most of these settings, customers are awarding top ratings leading to the classic j-shape distribution of star ratings (e.g., Schoenmueller et al, 2020). The distributional characteristics of such satisfaction metrics pose issues when the task is to model the relationship with other (potentially) influencing variables. In addition, since not all customers are participating in such surveys or are willing to provide a rating, the responses are typically subject to possible selection biases. As marketers are keen in understanding why customers give a certain rating, choosing the correct model is of vital importance to avoid biased and hence unreliable estimates leading to wrong managerial implications.

In our study, we model the drivers of customer review ratings in a lodging industry setting. When transforming the ratings into deviations from the top rating (which is of managerial interest in our application case), the j-shape pattern of the original distribution translates into a vast number of zeros. We hence propose a count model based on the Tweedie distribution (see Dunn and Smyth 2005 for applications and properties of the distribution) and account for sample selection to deal with the skewed distribution. We compare the Tweedie model to alternative models such as other count models and further candidate models that have been used in previous research. We evaluate the performance of the models regarding their distributional assumptions as well as further model selection criteria.

The hotel review data have been collected from two booking platforms between 2018 and 2019 and were enriched with booking data from a local hotel chain in a major European city. This gives us the opportunity to have a unique data set at hand enabling us to identify which customers have provided a review and observe the associated rating.

We find that our proposed Tweedie model performs better than other (sample selection) models. Our results show that coefficients vary in terms of significance and signs across the various models, which would imply different managerial conclusions and make the right model choice important.

7 Enhance Conjoint with a Behavioral Framework
Peter Kurz, bms marketing research + strategy

Shoppers are no stimulus-response machines, they are processing information and act accordingly.

Shopper perceptions of prices and values are therefore important to understand, the effect of price changes in a category: What goes on in the shopper's mind before he chooses a product?

The company bms marketing research + strategy uses 9 standard binary questions regarding shopping behavior in the category, upfront of each conjoint exercise. This helps to make the respondents remember their usual buying habits. These questions are based on principles from behavioral economics and guide our usage of this prior knowledge:

- Habits & heuristics
  People tend to simplify the task of decision making. An important *rule of thumb* for future choices - used consciously or not - is to revert to past experience.
- Frames & anchors
  When people make choices they look out for hints and references, which *frame* their decision making. The frame can be determined by memories of past decisions or by the context in which the new choice is presented.
- Brands provide orientation and guidance in all categories
- Price is a highly relevant landmark for all buying decision makers

We use the derived contextual information about each individual respondents disposition towards brand and price, knowledge (or lack of it), past behaviour and perceptions in the category. This prior knowledge of individual dispositions informs the following choice experiment.

In our paper we present the differences in the results of the choice model when respondents answer the 9 questions or doing the same experiment without the questions. Therefore, we have conducted 9 empirical studies where 50% of the respondent answered the 9 behavioral questions, upfront the choice experiment, whereas 50% do the choice exercise without the questions.

Our empirical results show that the framing (recall of past shopping experience) that takes place, has a positive impact on the answering behavior in the choice experiment and significantly improves hit rates and out of sample share estimates of the respondents that have answered the additional questions. Furthermore, we can use the 9 questions as co-variates to inform the CBC/HB estimation about the different shopping behavior of our respondents and further improve the results. Finally, we can segment our data into four consumer segments based on category involvement and brand switching disposition and gain insight for product development and pricing issues. For additional information see Allenby and Rossi (2006); Kurz and Binner (2010); Liakhovitski and Shmulyian (2011); Sentis and Geller (2010).

8 Demand-driven location planning using MNL, maximum likelihood estimation, and machine learning

Knut Haase, University of Hamburg
Jannis Kück, University of Hamburg
Fiona Sauerbier, University of Hamburg
Martin Spindler, University of Hamburg

Digitization leads to larger data sets available as input for optimization problems, in particular also for choice-based optimization problems which are usually combined with maximum likelihood estimation. Machine learning are particular useful for analyzing high-dimensional, complex data sets. We consider a high dimensional estimation problem and a location problem under the multinomial logit model. We integrate the machine learning methods Lasso regression and Ridge regression into the maximum likelihood method to estimate the multinomial logit model. We perform a computational study using synthetic data to determine the optimal solutions to location planning problems. The results are used to analyze the quality of the solutions of the location problems depending on the estimation method used. For additional information see Train (2009); Friedman et al (2010).

9 Analyzing Groves to Explain Random Forests

Gero Szepannek, Stralsund University of Applied Sciences
Björn-Hergen von Holt, Institut für Medizinische Biometrie und Statistik, Lübeck

Random forests show superior performance to decision trees in many machine learning problems. In contrast, the resulting models are no longer understandable and often called to be of black box nature. Different methods have been proposed to enhance interpretability of random forests in terms of tree-structured rule sets such as representative trees (Banerjee et al, 2012) or surrogate trees. These concepts can be extended to *groves* consisting of not only one single but a few trees. The explainability (Szepannek and Lübke, 2022) of a forest model by the different approaches is analyzed and juxtaposed to the complexity of the explanation.

10 Feature selection in high-dimensional data with tiny sample size

Sigrun May, Clausthal University of Technology

Fitting models on high-dimensional data with a tiny sample size often results in selecting many irrelevant features. In addition, nested crossvalidation is necessary to avoid biased performance evaluation. The selected feature subsets of each iteration usually differ even if they provide an equivalent prediction result. This leads to highly unstable feature subsets within the nested cross-validation. To address this lack of robustness we developed an alternative feature selection workflow: First, we cluster highly correlated features. To reduce the influence of outliers, we do not apply feature extraction. Instead, we calculate the best representative feature for each cluster. The degree of feature independence we adjust using a correlation threshold. Second, we reverse the feature selection. In the classical approach, applying LASSO to predict the label, features with nonzero coefficients are selected. Instead, we suggest using the potentially relevant feature itself as target feature.

All other features serve as training data. Furthermore, we eliminate features correlated to each respective target feature from each training set. Training is repeated twice, once with the label integrated into the training data and once without. We select a feature only if training data containing the label provides substantially better results than the same training data without the label. Based on this difference, a weight is assigned to each feature. To validate the selected feature subsets, we suggest a feature-weighted K-Nearest-Neighbor that considers these individual feature weights. Hence, both the subset and the individual relevance (weight) of each feature are included in the classification. Finally, we evaluate the results not only by an average evaluation metric but by 10 different metrics (micro matthews, micro accuracy, micro f1 score, micro balanced accuracy score, macro auc, macro logloss, macro brier score loss, macro top k accuracy score, stability, number of robust features). This way, we can show that reverse feature selection leads to an increased number of selected robust features. It reduces the influence of overfitting and unstable subsets. Reverse feature selection is an alternative method complementing state of the art ensemble techniques (Muthukrishnan and Rohini, 2016; Vabalas et al, 2019; Wahid et al, 2022).

11 Immunophenotyping B-cell Lymphoma with a Trustworthy Artificial Intelligence System using Human-in-the-loop approach

Michael C. Thrun, Philipps University Marbur
Jörg Hoffmann, Philipps University Marburg, University Hospital Giessen and Marburg
Stefan W. Krause, Universitätsklinikum Erlangen
Peter Krawitz, University Bonn
Cornelia Brendel, Philipps University Marburg, University Hospital Giessen and Marburg
Alfred Ultsch, Philipps University Marburg

Diagnostic immunophenotyping of lymphoma is often performed via multiparameter flow cytometry by measuring cell surface expression levels from several antigens on peripheral blood B cells. Artificial intelligence (AI) claims to diagnose lymphoma cases automatically. However, AIs harbor obstacles: They require many training examples, and - by their nature as sub-symbolic systems – trustworthiness is impaired because it is impossible to get either competence estimation or explanations about their decision.

Here we present a combined unsupervised and supervised artificial intelligence which closely resembles the stepwise human-in-the-loop diagnostic approach of medical experts. An unsupervised sample quality check through the tiles mining algorithm first allows one to identify core structures in data sets and recognize outliers. Next, a supervised explainable AI called ALPODS (Ultsch et al, 2022) is trained with only 256 lymphoma samples and an equal number of healthy controls. Subsequently, ALPODS was trained to differentiate between B-NHL and normal control samples in an explainable manner. Thereupon, the AI was trained to classify lymphoma on three levels in analogy to human experts:

1. Separation of normal controls from B-cell lymphoma,
2. Identification of CLL and HCL, and
3. Subclassification of other B-NHL, which often cannot be distinguished based on flow cytometry data alone.

In sum, this AI approach models the decision levels of human diagnostic experts and allows a human-in-the-loop to examine each step. Moreover, our AI is capable of calculating a value that indicates its own trustworthiness for the diagnosis of each sample.

The results show that our trustworthy learning artificial intelligence system is capable of diagnosing lymphoma from flow cytometric data with a tiny training cohort. The trustworthy AI system outperformed similar approaches on different levels. The AI system can extrapolate on the test set with an accuracy of 98.2% for the distinction between B-NHL and normal controls and harbors self-assessment properties about the trustworthiness of its decisions. It can be trained with only 256 lymphoma samples – making it applicable for single-center diagnostic laboratories that wish to work by AI support on a given lymphoma panel. The novel system was compared to similar published algorithms of a deep learning approach (Zhao et al, 2020) and Citrus (Bruggner et al, 2014). Our AI exhibited superior performance with a Mathews correlation coefficient of 87% within a seven-class system and 5904 test samples. To the best of our knowledge, it is the first AI capable of outperforming human experts (MCC=83%). The trustworthy AI system was validated on a different dataset from an independent diagnostic center.

12 Combining the Effects of ESG Ratings and Macroeconomics on Stock Returns Using Causal Inference and Worldwide Panel Data

Tilo Zimmermann, FOM Frankfurt

In this study, we investigate the diffuse effect of environmental, social, and governance (ESG) ratings on stock returns using a panel regression and a global sample of data. To our knowledge, this is the first study to integrate directed acyclic graphs (DAGs) as a method of causal modeling into the analysis. Specifically, we construct DAGs to overcome biases caused by control variables and demonstrate the combined impact of ESG ratings, macroeconomics, and the Fama-French (FF) factors on stock returns. The developed DAG suggests that the commonly used FF factors inhibit a mediating role within the influence of ESG ratings on stock returns. In a subsequent empirical analysis, we construct stock portfolios based on ESG ratings finding that worse ESG Ratings are associated with higher returns using the Kruskal-Wallis test and Welch's analysis of variance. However, the regression models developed in the study illustrate that not only the choice of the rating provider but also model specifications, such as sample selection and choice of control variables, influence the designated direction and significance of the impact of ESG ratings on stock returns. This is true even for a randomly created control variable. Therefore, our findings implicate that methodological variation can explain the heterogeneous results of previous studies. For additional information see Liang and Renneboog (2020); Mitton (2022); Pearl et al (2016).

13 Automatic Fake News Detection to Ensure Quality of News Articles

Adalbert F.X. Wilhelm, Jacobs University Bremen
Bonaventure F. P. Dossou, Jacobs University Bremen

Access to information is an inherent right of every living human being. Defined as factual information published in newspapers or broadcasted on radio or television, news helps us every day to keep informed about what is happening in our society and around the world. With the rise of the internet and social media during the last decades news generation and information broadcast had become possible without the regulatory action of media's gatekeeping and their editorial scrutiny facilitating generation and spread of fake news. While disinformation has circulated through media since the early days of mass communication, scholars and pundits have argued that recent years mark *the rise of the misinformation society* (Pickard, 2016) and the era of *alternative facts* and *post truth* (Benkler et al, 2018). The automatic detection of fake news, their reduction and eradication is therefore a great challenge with high relevance for the daily debates. Quite some research has been done in this regard, using machine learning techniques such as supervised learning, unsupervised learning, reinforcement learning(Zhang et al, 2018), or the arising self-supervised learning. In NLP, the battle to use the power of AI to detect, reduce and eradicate the propagation of fake news is an ongoing trend. In this presentation, using the ISOT fake news dataset, we test the importance of choosing the right embedding model. We integrate the most efficient embedding model, and we implement a reinforcement learning framework to perform binary classification of news articles.

14 How Social Media Drove the 2016 and 2020 U.S. Presidential Elections

Raoul V. Kübler, ESSEC Business School Paris

How did social media interactions versus the candidates' own actions drive the 2016 and 2020 U.S. presidential elections? Were candidates misled if they focused on traditional market research versus the newer probabilistic polls? Candidates have different political persuasion pathways, and different topics and media are better suited to achieve this end. The authors compose a daily data set combining donations, polls, and advertising with social media interactions on Twitter, Facebook, and Instagram. Structural topic modeling reveals nuances in topics and sentiment, after which persistence modeling reveals a very different impact on traditional versus the probabilistic polls, showing the danger of relying on one predictive metric. Candidates largely control their own fate, but the impact of their paid actions differs by channel, topic, and how they resonate in social media. Disinformation about the candidates hurts their chances on some topics, while boosting them on others. Applying the learnings from the 2016 election, the authors predict the 2020 election and draw insights to advise where, when, and how to drive the political conversation.

15 Reducing Product Returns through Green Nudges and Causal Machine Learning

Moritz von Zahn, Goethe University Frankfurt

As free customer deliveries are becoming a standard in E-commerce, product returns pose a growing challenge to online retailers and society. For retailers, product returns create considerable costs associated with transportation, labor, disposal and infrastructure to manage returns. From a societal perspective, increasing product returns contribute to increased pollution, additional trash, and often a waste of natural resources. Due to these costs, companies and society are interested in reducing product returns. However, retailers on a micro level possess only very few effective instruments to minimize product returns without harming customer demand and net sales. In this work, we propose a novel product return prevention instrument (Smart Green Nudging) that leverages Causal Machine Learning (CML) and the availability of rich customer and contextual data sources. Smart Green Nudging identifies and targets selected customers towards better shopping choices that will yield reduced product returns without diminishing customer demand and net sales. We evaluate the performance of Smart Green Nudging with real-world data from the German online shop of a large European retailer. Smart Green Nudging decreases product returns by 4.7% and increases profits by up to +12.9%. Moreover, we use methods from Explainable Artificial Intelligence (XAI) to reveal which customer characteristics drive the nudging effect. Thereby, XAI adds valuable managerial insights and helps improving personalization. Overall, this paper demonstrates the efficacy of using state-of-the-art CML and XAI to customize minimally invasive behavioral nudges in the digital environment as a means to tackle business and societal problems. For additional information see Wager and Athey (2018); Lundberg and Lee (2017).

16 Data4Good at CorrelAid – Impact for Civil Society and Volunteers

Zoé Wolter & Philipp Bosch, CorrelAid e. V.

CorrelAid e.V. is a non-partisan, non-profit network of data science enthusiasts who want to change the world using data science – in our presentation we want to show how CorrelAid has made the step from a student initiative to one of the pioneers in the field of Data4Good in civil society. In doing so, we elaborate on what Data4Good means to CorrelAid and how we pragmatically make a sustainable impact for civil society through our pro bono projects with other non-profit organizations while training and developing the skills of our volunteers. Our approach to education goes beyond simply passing on technical knowledge about models and programming. For us at CorrelAid, digital education also always means critically questioning technologies and methods.

References

Allenby GM, Rossi PE (2006) Hierachical bayes models. In: Grover R, Vriens M (eds) The Handbook of Marketing Research. Uses, Misuses, and Future Advances, SAGE Publications Inc., Thousand Oaks, California, chap 20, pp 418 – 440

Banerjee M, Ding Y, Noone AM (2012) Identifying representative trees from ensembles. Statistics in Medicine 31(15):1601 – 1616, DOI 10.1002/sim.4492

Barrio-Tellado MJ, Herrero-Prieto LC (2019) Modelling museum efficiency in producing inter-reliant outputs. Journal of Cultural Economics 43(3):485 – 512, DOI 10.1007/s10824-019-09347-2

Barrio-Tellado MJ, Herrero-Prieto LC (2022) Analysing productivity and technical change in museums: A dynamic network approach. Journal of Cultural Heritage 53:24 – 34, DOI 10.1016/j.culher.2021.10.007

Benkler Y, Faris R, Roberts H (2018) Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics. Oxford University Press, New York, DOI 10.1093/oso/9780190923624.001.0001

Bischl B, Binder M, Lang M, Pielok T, Richter J, Coors S, Thomas J, Ullmann T, Becker M, Boulesteix AL, Deng D, Lindauer M (2021) Hyperparameter optimization: Foundations, algorithms, best practices and open challenges. arXiv pp 1 – 70, DOI 10.48550/arXiv.2107.05847

Bruggner RV, Bodenmiller B, Dill DL, Nolan GP (2014) Automated identification of stratifying signatures in cellular subpopulations. Proceedings of the National Academy of Sciences 111(26):2770 – 2777, DOI 10.1073/pnas.1408792111

De Jong JP, Marsili O (2015) Founding a business inspired by close entreprenurial ties. does it matter for survival? Entrepreneurship Theory and

Practice 39(5):1005 – 1025, DOI 10.1111/etap.12086
Dunn PK, Smyth GK (2005) Series evaluation of Tweedie exponential dispersion model densities. Statistics and Computing 15:267 – 280, DOI 10.1007/s11222-005-4070-y
Feldhoff F, Klefenz FM, Toepfer H, Harczos T (2022) Periodicity pitch perception part III: Sensibility and Pachinko volatility. Frontiers in Neuroscience 16, DOI 10.3389/fnins.2022.736642
Friedman J, Hastie T, Tibshirani R (2010) Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software 33(1):1 – 22
Harczos T, Klefenz FM (2018) Modeling pitch perception with an active auditory model extended by octopus cells. Frontiers in Neuroscience 12, DOI 10.3389/fnins.2018.00660
Kurz P, Binner S (2010) Added value through covariates in HB modeling? In: Software S (ed) Proceedings of the Sawtooth Software Conference 2010, Sawtooth Software, Inc., Sequim, Washington, Proceedings of the Sawtooth Software Conference, vol 15, pp 269 – 282
Liakhovitski D, Shmulyian F (2011) Covariates in discrete choice models: Are they worth the trouble? ART Form Presentation
Liang H, Renneboog L (2020) Corporate social responsibility and sustainable finance: A review of the literature. SSRN 2020(701/2020):1 – 43, DOI 10.2139/ssrn.3698631
Lundberg SM, Lee SI (2017) A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, Curran Associates Inc., Red Hook, NY, USA, NIPS'17, pp 4768 – 4777
McMahon JJ (2008) Time-dependence in Markovian decision processes. PhD thesis, Faculty of Mathematical and Computer Science, Uniiversity of Adelaide, Adelaide
Mitchell TM (1997) Machine Learning. McGraw-Hill, New York
Mitton T (2022) Methodological variation in empirical corporate finance. The Review of Financial Studies 35(2):527 – 575, DOI 10.1093/rfs/hhab030
Muthukrishnan R, Rohini R (2016) Lasso: A feature selection technique in predictive modeling for machine learning. In: 2016 IEEE International Conference on Advances in Computer Applications (ICACA), Coimbatore, Tamil Nadu, India, pp 18 – 20, DOI 10.1109/ICACA.2016.7887916
Pearl J, Glymour M, Jewell NP (2016) Causal Inference in Statistics: A Primer, 1st edn. John Wiley & Sons, West Sussex

Peterson RA, Wilson WR (1992) Measuring customer satisfaction: Fact and artifact. Journal of the Academy of Marketing Science 20:61 – 71, DOI 10.1007/BF02723476

Pickard V (2016) Media failures in the age of Trump 4(2):118 – 122

Puterman ML (2005) Markov Decision Processes: Discrete Stochastic Dynamic Programming, 2nd edn. John Wiley, New York

Schoenmueller V, Netzer O, Stahl F (2020) The polarity of online reviews: Prevalence, drivers and implications. Journal of Marketing Research 57(5):853 – 877, DOI 10.1177/0022243720941832

Sentis K, Geller V (2010) The impact of covariates on HB estimates. In: Software S (ed) Proceedings of the Sawtooth Software Conference 2010, Sawtooth Software, Inc., Sequim, Washington, Proceedings of the Sawtooth Software Conference, vol 15, pp 255 – 267

Simón-Moya V, Revuelto-Taboada L, Ribeiro-Soriano D (2012) Are success and survival factors the same for social and business ventures? Service Business 6(2):219 – 242, DOI 10.1007/s11628-012-0133-2

Stinchcombe AL (1965) Social structure and organizations. In: March JG (ed) Handbook of Organizations, Rand McNally & Company, Chicago, Illinois, chap 4, pp 142 – 193

Su H, Zhang H, Zhang X, Gao G (2016) Convolutional neural network for robust pitch determination. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 579 – 583, DOI 10.1109/ICASSP.2016.7471741

Szepannek G, Lübke K (2022) Explaining artificial intelligence with care. Künstliche Intelligenz 36:125 – 134, DOI 10.1007/s13218-022-00764-8

Train K (2009) Discrete Choice Methods with Simulation, 2nd edn. Cambridge University Press, Cambridge, UK

Ultsch A, Hoffmann J, Röhnert M, Von Bonin M, Oelschlägel U, Brendel C, Thrun MC (2022) An explainable AI system for the diagnosis of high dimensional biomedical data. arXiv 2022:1 – 29, DOI 10.48550/arXiv.2107.01820

Vabalas A, Gowen E, Poliakoff E, Casson AJ (2019) Machine learning algorithm validation with a limited sample size. PloS one 14(11):1 – 20, DOI 10.1371/journal.pone.0224365

Vecchi AO, Varnet L, Carney LH, Dau T, Bruce IC, Verhulst S, Majdak P (2022) A comparative study of eight human auditory models of monaural processing. arXiv 2022:1 – 24, DOI 10.48550/arXiv.2107.01753

Wager S, Athey S (2018) Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association 113(523):1228 – 1242, DOI 10.1080/01621459.2017.1319839

Wahid A, Khan DM, Iqbal N, Janjuhah HT, Khan SA (2022) A generalized stability estimator based on inter-intrastability of subsets for high-dimensional feature selection. Chemometrics and Intelligent Laboratory Systems 220(104457):1 – 14, DOI 10.1016/j.chemolab.2021.104457

Zhang T, Huang M, Zhao L (2018) Learning structured representation for text classification via reinforcement learning. Proceedings of the AAAI Conference on Artificial Intelligence 32(1):6053 – 6060, DOI 10.1609/aaai.v32i1.12047

Zhao M, Mallesh N, Höllein A, Schabath R, Haferlach C, Haferlach T, Elsner F, Lüling H, Krawitz P, Kern W (2020) Hematologist-level classification of mature B-cell neoplasm using deep learning on multiparameter flow cytometry data. Cytometry Part A 97(10):1073 – 1080, DOI 10.1002/cyto.a.24159