# Joint Meeting of the GfKl Working Groups AG MARKETING and AG DANK


## October 7th–8th, 2022
## TU Clausthal, Germany


**Program Committee and Organization:**
**Friederike Paetz, Gero Szepannek, Philipp Aschersleben**

# Final schedule

| Friday, October 7th, 2022 | | Saturday, October 8th, 2022 | |
|---|---|---|---|
| 13:15-14:00 | Registration / Coffee | 08:45-10:15 | Session C |
| 14:00-14:15 | Opening | 10:15-10:30 | Coffee break |
| 14:15-15:45 | Session A | 10:30-12:00 | Session D |
| 15:45-16:00 | Coffee break | 12:00-12:15 | Coffee break |
| 16:00-16:30 | Invited Talk: R. V. Kübler | 12:15-13:15 | Invited Talk: M. von Zahn |
| 16:30-18:00 | Session B | 13:15-13:45 | Data analysis competition |
| 18:00-18:15 | Coffee break | 13:45-14:00 | Closing |
| 18:15-19:00 | Invited Talk: CorrelAid e. V. | | |
| 19:30 | Dinner (optional/self-pay) | | |

**Venue:**
TU Clausthal
Institut für Wirtschaftswissenschaft
Seminarraum 1 (D2-306)
Julius-Albert-Straße 2
38678 Clausthal-Zellerfeld

**WiFi:**
The TU Clausthal is part of the eduroam network allowing you to use your home institution account. Alternatively, please inform us in advance if you want to use the TUC-Guest WLAN.

**Dinner on Friday evening:**
Hotel "Zum Harzer"
Treuerstr. 6
38678 Clausthal-Zellerfeld
https://www.zum-harzer.de/

**Bus shuttle:**
On Friday, we provide a bus shuttle from the central bus station ("ZOB") to the venue at 13:10 and 13:40 and on demand. Please let us know in advance, if you are interested in making use of this service (mail: philipp.aschersleben@tu-clausthal.de). On Friday, in case of unforeseeable changes, you can contact us via phone: 05323 72-7663.

# Sessions

## Friday, October 7th, 2022

### Session A

- Control of Shared Production Buffers: A Reinforcement Learning Approach (*Nora Krippendorff, Christoph Schwindt*)
- On the Relevance of Transaction Data: A Funnel Analysis of New Venture Survival (*Ulrich Müller-Funk, Christina Ungerer*)
- Investigating Deep Learning for Auditory Model Based Pitch Classification (*Jan-Paul Schulz, Gero Szepannek, Tamas Harczos*)

### Session B

- How efficient are German museums in using their public funding? An input-oriented network DEA approach (*Lea Hildebrand, Friederike Paetz*)
- Accounting for skewed distributions in modeling self-reported customer response data (*Nadine Schröder, Giampiero Marra, Rosalba Radice, Thomas Reutterer*)
- Enhance Conjoint with a Behavioral Framework (*Peter Kurz*)

## Saturday, October 8th, 2022

### Session C

- Demand-driven location planning using MNL, maximum likelihood estimation, and machine learning (*Knut Haase, Jannis Kück, Fiona Sauerbier, Martin Spindler*)
- Can't see the forest for the trees – explainability of random forests (*Gero Szepannek, Björn-Hergen von Holt*)
- Feature selection in high-dimensional data with tiny sample size (*Sigrun May*)

### Session D

- Immunophenotyping B-cell Lymphoma with a Trustworthy Artificial Intelligence System using Human-in-the-loop approach (*Michael C. Thrun, Jörg Hoffmann, Stefan W. Krause, Peter Krawitz, Cornelia Brendel, Alfred Ultsch*)
- Combining the Effects of ESG Ratings and Macroeconomics on Stock Returns Using Causal Inference and Worldwide Panel Data (*Tilo Zimmermann*)
- Automatic Fake News Detection to Ensure Quality of News Articles (*Adalbert F. X. Wilhelm, Bonaventure F. P. Dossou*)

# Abstracts

## SESSION A

### 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. 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.

- J. J. McMahon (2008) Time-dependence in Markovian decision processes, Ph.D. dissertation, Faculty of Mathematical and Computer Sciences, University of Adelaide, Australia.
- T. M. Mitchell (1997) Machine Learning. New York, NY: McGraw-Hill
- M. L. Puterman (1994) Markov decision processes: Discrete stochastic dynamic programming. New York: Wiley.

### 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.

- Jong, J. P. de, & Marsili, O. (2015). Founding a Business Inspired by Close Entrepreneurial Ties: Does it Matter for Survival? *Entrepreneurship Theory and Practice*, *39*(5), 1005–1025. https://doi.org/10.1111/etap.12086

- 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. https://doi.org/10.1007/s11628-012-0133-2
- Stinchcombe, A. L. (1965). Social structure and organizations. In J. March (Ed.), *Handbook of Organizations* (pp. 142–193). Chicago: Rand McNally.


**Investigating Deep Learning for Auditory Model Based Pitch Classification**


*Jan-Paul Schulz, Saarland University*
*Gero Szepannek, Stralsund University of Applied 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 successfull 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 mimicing human pitch perception.


- 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. DOI 10.48550/ARXIV.2107.05847, URL https://arxiv.org/abs/2107.05847
- Feldhoff F, Klefenz FM, Toepfer H, Harczos T (2022 (accepted)) Periodicity pitch perception part III: Sensibility and Pachinko volatility. Frontiers in Neuroscience
- Harczos T, Klefenz FM (2018) Modeling pitch perception with an active auditory model extended by octopus cells. Frontiers in Neuroscience 12, URL https://doi.org/10.3389/fnins.2018.00660
- Su H, Zhang H, Zhang X, Gao G (2016) Convolutional neural network for robust pitch determination. In: ICASSP 2016, pp 579–583, DOI 10.1109/ICASSP.2016.7471741
- Vecchi AO, Varnet L, Carney LH, Dau T, Bruce IC, Verhulst S, Majdak P (2021) A comparative study of eight human auditory models of monaural processing. ArXiv abs/2107.01753, URL https://arxiv.org/pdf/2107.01753.pdf

**SESSION B**

## 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.

- Barrio-Tellado, M. J. & Herrero-Prieto, L. C. (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, M. J. & Herrero-Prieto, L. C. (2022) Analysing the productivity and technical change in museums: A dynamic network approach. Journal of Cultural Heritage 53(2022):24-34. DOI: 10.1016/j.culher.2021.10.007

**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 & 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 & 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.

- Dunn & Smyth 2005 Series Evaluation of Tweedie Exponential Dispersion Model Densities, https://doi.org/10.1007/s11222-005-4070-y
- Peterson & Wilson 1992 Measuring Customer Satisfaction: Fact and Artifact, https://doi.org/10.1007/BF02723476
- Schoenmueller, Netzer, Stahl 2020 The Polarity of Online Reviews: Prevalence, Drivers and Implications, https://doi.org/10.1177/0022243720941832

**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?
bms use 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 improve 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.


- Allenby, G.M.; Rossi, P.E. (2006): Hierarchical Bayes Models, in: Grover, R.; Vriens, M. (Eds.): The Handbook of Marketing Research: Uses, Misuses, and Future Advances, 418-440, SAGE Publications Inc., Thousand Oaks.
- Kurz, P; Binner, S. (2011): Added Value through Covariates in HB Modeling?, Proceedings of the 2011 Sawtooth Software Conference.
- Liakhovitski, D.; Shmulyian, F. (2011): Covariates in Discrete Choice Models: Are They Worth the Trouble?" ART Form Presentation.
- Sentis, K.; Geller, V. (2011): The Impact of Covariates on HB Estimates, Proceedings of the 2011 Sawtooth Software Conference.

**SESSION C**

## 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.

- Train, K. (2009): Discrete Choice Methods with Simulation. Cambridge: Cambridge University Press.
- Friedman J, Hastie T, Tibshirani R. (2010): Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33(1),1-22.

## Can't see the forest for the trees – explainability of random forests

*Gero Szepannek, Stralsund University of Applied Science*
*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.

- Bannerjee, M., Ding, Y., Noone, A.-M. (2012). Identifying representative trees from ensembles. Stat in Med 31:1601-16.
- Szepannek, G., Luebke, K. (2022). Explaining Artificial Intelligence with Care. Künstl Intell. https://doi.org/10.1007/s13218-022-00764-8

**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 evalutation 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)


- 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), pp 18–20, DOI 10.1109/ICACA.2016.7887916
- Vabalas A, Gowen E, Poliakoff E, Casson AJ (2019) Machine learning algorithm validation with a limited sample size. PLOS ONE 14(11):e0224,365, DOI 10.1371/journal.pone.0224365
- Wahid A, Khan DM, Iqbal N, Janjuhah HT, Khan SA (2022) A generalized stability estimator based on inter-intrastability of subsets for highdimensional feature selection. Chemometrics and Intelligent Laboratory Systems 220:104,457, DOI 10.1016/J.CHEMOLAB.2021.104457

**SESSION D**

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

*Michael C. Thrun, Philipps University Marburg*
*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 sMarburg*

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.

- Bruggner RV, Bodenmiller B, Dill DL, Tibshirani RJ, Nolan GP. Automated identification of stratifying signatures in cellular subpopulations. Proceedings of the National Academy of Sciences. 2014;111(26):E2770-E7.
- Ultsch A, Hoffman J, Röhnert M, Von Bonin M, Oelschlägel U, Brendel C, Thrun, M.C. An Explainable AI System for the Diagnosis of High Dimensional Biomedical Data. 2022. 10.48550/arXiv.2107.01820.
- Zhao M, Mallesh N, Höllein A, Schabath R, Haferlach C, Haferlach T, et al. Hematologist-Level Classification of Mature B-Cell Neoplasm Using Deep Learning on Multiparameter Flow Cytometry Data. Cytometry Part A. 2020;97:1073-80. doi: 10.1002/cyto.a.24159.


**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.



- Liang, H., & Renneboog, L. (2020). Corporate social responsibility and sustainable fi-nance: a review of the literature. *SSRN Electronic Journal*, Finance Working Pa-per No. 701/2020. http://doi.org/10.2139/ssrn.3698631
- Mitton, T. (2022). Methodological variation in empirical corporate finance. *The Review of Financial Studies*, 35, 527-575. https://doi.org/10.1093/rfs/hhab030
- Pearl, J., Glymour, M., & Jewell, N. (2016). Causal inference in statistics, A primer, 1. Edition. West Sussex: John Wiley & Sons

**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, or the arising selfsupervised 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.

- Benkler, Y., Faris, R., & Roberts, H. (2018). *Network propaganda: Manipulation, disinformation, and radicalization in American politics*. Oxford University Press
- Pickard, V. (2016). Media failures in the age of Trump. *The Political Economy of Communication*, *4*(2). http://polecom.org/index.php/polecom/article/viewFile/74/264
- Zhang, T., Huang, M., and Zhao, L. (2018). Learning structured representation for text classification via reinforcement learning.