Title: Explainable AI Methods for Multimodal Data in the Bioeconomy

Team: ATB: Marina Höhne
UOS: Martin Atzmüller

Abstract:
AI models can be trained on a variety of data, such as sensor data, satellite data, genomic data, image data, weather data, and climate data. This can be used in the context of different applications and several use cases in the bioeconomy domain, e.g., for improving performance regarding disease detection on plants, precise plant monitoring, smart water use, and optimizing waste recycling processes. Such models can then help, for example, to detect pests, identify weeds, and optimize the use of fertilizers and pesticides.

Although those AI models learn complex relationships between different components in the bioeconomy, those relations typically remain hidden within the often called black box model, leading to no or only poor knowledge about the underlying connections of the driving factors in the input data. Here, the field of explainable AI (XAI) has emerged, with the aim to provide insights into the decision-making strategies and to showcase the important features and complex relationships the decision is based on. In this project, methods of explainable AI for multimodal data will be developed for explaining AI models trained on multimodal data with applications in the bioeconomy. With the novel XAI methods we aim to showcase the complex relationships between different types of data, such as genetics, soil composition, weather, and climate data, learned by the AI model. We aim to contribute to the generation of novel knowledge that tackles the enormous challenge of an efficient realization of the bioeconomy.

Desired skills of the applicant:
Must: Excellent Knowledge in ML, especially Deep Learning, and in at least one of the following topics: Explainable AI, Graph Neural Networks, Object Detection

References:
Stefan Bloemheuvel, Jurgen van den Hoogen, Dario Jozinovic, Alberto Michelini, and Martin Atzmueller (2022) Graph neural networks for multivariate time series regression with application to seismic data. International Journal of Data Science and Analytics

Leonid Schwenke and Martin Atzmueller (2021) Constructing global coherence representations: Identifying interpretability and coherences of transformer attention in time series data. In Proc. IEEE International Conference on Data Science and Advanced Analytics, DSAA 2021, Porto, Portugal, October 6-9, 2021, pages 1–12. IEEE

Hedström, A., Weber, L., Krakowczyk, D., Bareeva, D., Motzkus, F., Samek, W., ... & Höhne, M. M. C. (2023). Quantus: An Explainable AI Toolkit for Responsible Evaluation of Neural Network Explanations and Beyond. Journal of Machine Learning Research, 24(34), 1-11.

Gautam, S., Boubekki, A., Hansen, S., Salahuddin, S. A., Jenssen, R., Höhne, M. M., & Kampffmeyer, M. (2022) ProtoVAE: A Trustworthy Self-Explainable Prototypical Variational Model. In Advances in Neural Information Processing Systems.