Chair of Explainable AI-Based Business Information Systems
Prof. Dr. Ulrich Gnewuch

Bachelor's Thesis: Empirical Comparison of the Effectiveness of XAI Methods in Fostering Explainability for Lay Users

Supervisor: M.Sc. Philipp Hansen (philipp.hansen@uni-passau.de)
Start date: at the next possible date

Motivation and Goals

When decisions are made based on the output of a high-risk Artificial Intelligence (AI) system, the affected person has a right to receive explanations on the decision-making process and corresponding result (European Parliament & Council of the European Union, 2024). This however poses a significant challenge, as AI algorithms are oftentimes opaque in nature, making it difficult to provide reasons on why, e. g., a credit application was denied (Adadi & Berrada, 2018). The field of Explainable AI (XAI) has set out to address the lack of transparency of AI algorithms and provide model explanations (Adadi & Berrada, 2018). However, proposed XAI solutions are oftentimes not evaluated with human users, which raises the question to how useful they really are in practice (Suh et al., 2025).

In light of this, the goal of the thesis is to empirically compare contemporary XAI methods, such as LIME and SHAP, concerning their ability to explain the prediction of a Machine Learning (ML) classifier to lay users. For this purpose, the student shall provide a brief overview on state-of-the-art XAI frameworks, conduct a small-scale user study for a given use case (e. g., credit applications), and evaluate two frameworks with lay users.

Required Skills

- Interest in AI and ML methods
- Good English skills
- Prior coding experience (e. g., Python, MATLAB, etc.)

Starting Literature (Topic)

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 of the European Parliament and of the Council. *Official Journal of the European Union*. http://data.europa.eu/eli/reg/2024/1689/oj

Martínez, M. A. M., & Mädche, A. (2023). Designing Interactive Explainable AI Systems for Lay Users. *ICIS 2023 Proceedings*, 5. https://aisel.aisnet.org/icis2023/dab_sc/dab_sc/5

Suh, A., Hurley, I., Smith, N., & Siu, H. C. (2025). Fewer Than 1% of Explainable AI Papers Validate Explainability with Humans (No. arXiv:2503.16507v1). *arXiv*. https://doi.org/10.48550/arXiv.2503.16507

Starting Literature (Method)

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly*, 28(1), 75–105. https://doi.org/10.2307/25148625

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems*, 24(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, 26(2), xiii–xxiii. https://www.jstor.org/stable/4132319