

Master Project

Scenario-based Evaluation of AutoML Platforms and their Suitability to Support Organizational Decision-Making

Term: Summer Term 2022

Language: English

Motivation

Decision-making processes in organizations are often supported by information technology such as machine learning (ML) approaches. The process of developing and maintaining ML models, often referred to as “ML pipeline”, includes steps such as data pre-processing, feature engineering, model selection, and model evaluation. All of those steps require in-depth knowledge and are labor-intensive. The multitude of possible variations in each of these steps poses multiple challenges concerned with, among others, the preparation of data, the selection of a suitable approach, or the lack of skilled experts.

The research field of automated machine learning (AutoML) addresses these challenges. AutoML can be understood as an umbrella term constituting a broad range of different techniques and tools that aim to automate various tasks within the ML pipeline. While these approaches may increase the measured performance of the resulting model, they might also contribute to increasing its opaqueness since core decisions about its design have been automated. This, then, might lead to challenges and limitations in the use of AutoML tools concerned with, e.g., concept drift (cf. Celik and Vanschoren 2021).

Description

The main aim of this master project is to evaluate, using a real-world scenario, selected AutoML platforms and their suitability to support organizational decision-making. To this aim the students should:

- (1) Make themselves familiar with the foundational concepts of ML and AutoML automated machine learning approaches and gain a rough overview of core elements in organizational decision-making.
- (2) Identify and select AutoML tools.
- (3) Select a real-world scenario to be used for the needs of comparative evaluation.
- (4) Design an evaluation framework as well as an evaluation protocol.
- (5) Conduct the evaluation of selected platforms.
- (6) Critically analyze obtained results and draw conclusions.

**Institut für Informatik und
Wirtschaftsinformatik (ICB)**

**Lehrstuhl für
Wirtschaftsinformatik und
Unternehmensmodellierung**

Dr. Monika Kaczmarek-Heß

Tel.: 0201 / 183 - 4330

monika.kaczmarek-hess@uni-due.de

Pierre Maier

Tel.: 0201 / 183 - 4563

pierre.maier@uni-due.de

R09 R04 H41

Universitätsstraße 9

45127 Essen

www.umo.wiwi.uni-due.de

Expected Outcomes

A report describing conducted state-of-the-art analysis, cf. point (1) of the project description, the selected scenario, proposed evaluation framework and protocol, a detailed analysis of the evaluation results, and final assessment of the maturity of the tools and their suitability to support decision-making processes. In addition, a final presentation of the project results is expected.

Introductory Literature

- Celik B, Vanschoren J (2021) Adaptation Strategies for Automated Machine Learning on Evolving Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43(9):3067-3078
- Drozdal J, Weisz J, Wang D, Dass G, Yao B, Zhao C, Muller M, Ju L, Su H (2020). Trust in AutoML: Exploring Information Needs for Establishing Trust in Automated Machine Learning Systems. *Proceedings of the 25th International Conference on Intelligent User Interfaces*. ACM, pp. 297–307.
- Ebadi A, Gauthier Y, Tremblay S, Paul P (2019). How can Automated Machine Learning Help Business Data Science Teams? *ICMLA 2019. 18th IEEE International Conference on Machine Learning and Applications: Proceedings*, pp. 1186-1191
- Truong A, Walters A, Goodsitt J, Hines K, Bruss CB, Farivar R (2019). Towards Automated Machine Learning: Evaluation and Comparison of AutoML Approaches and Tools. *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*. IEEE, pp. 1471–1479.
- Tuggener L, Amirian M, Rombach K, Lorwald S, Varlet A, Westermann C, Stadelmann T (2019). Automated Machine Learning in Practice: State of the Art and Recent Results. *2019 6th Swiss Conference on Data Science (SDS)*. IEEE, pp. 31–36.
- Wang Q, Ming Y, Jin Z, Shen Q, Liu D, Smith MJ, Veeramachaneni K, Qu H (2019). ATMSeer: Increasing Transparency and Controllability in Automated Machine Learning. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12.
- Weng Z. (2019). From Conventional Machine Learning to AutoML. *Journal of Physics: Conference Series* 1207:1-9
- Xanthopoulos I, Tsamardinos I, Christophides V, Simon E, Salinger A (2020). Putting the Human Back in the AutoML Loop. *Workshop Proceedings of the EDBT/ICDT 2020 Joint Conference*.

Application Procedure:

Please apply via email to the supervisors. Please attach a short letter of motivation (approximately one A4 page) and a recent transcript of record ('Leistungsnachweis'). You can apply individually or in a group of **2-4 participants** (in this case each person should still send a separate e-mail, however point to the other members of the group).

Application deadline: 20 April 2022, 23:59 h