Automated Machine Learning

using Meta-Learning and Meta Features

Liliia Aliakberova - Maren Hoschek
Barcelona, Spain - eBISS 2023
liliia.aliakberova@estudiantat.upc.edu & maren.hoschek@estudiantat.upc.edu

1. Introduction

Machine learning is evolving, and Automated Machine Learning (**AutoML**), **meta-learning**, and **meta-features** play a vital role in driving innovation and expanding possibilities.This research overview emphasizes the significance of AutoML, meta-learning, and meta-learning using meta-features in improving the efficiency and effectiveness of **machine learning pipelines**. Through the exploration of meta-learning capabilities and the utilization of meta-features, our aim is to offer valuable insights and provide an overview of the latest advancements in AutoML.

2. Automated Machine Learning

AutoML is a actively developing field that aims to automate and optimize various stages of the machine learning pipeline [Feu15].

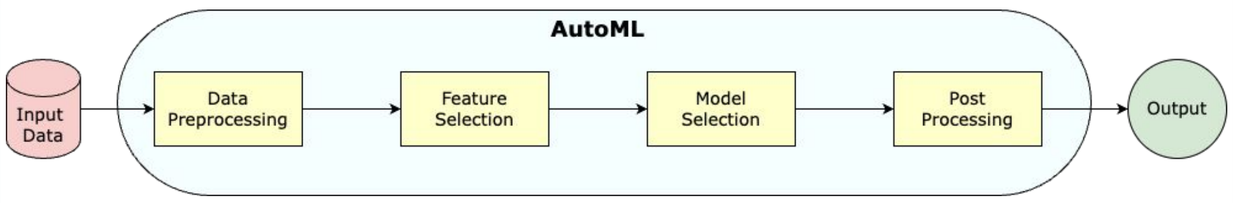

Figure 1: AutoML Concept

Automating the machine learning pipeline:

- **Automated preprocessing** ensures standardized data cleaning, transformation, and normalization, handling missing data, outliers, and categorical variables [Feu15].
- **Automated feature engineering** benefits from AutoML's automated feature selection, extraction, and engineering algorithms, enabling comprehensive exploration of the feature space and improving model performance [EMS19].
- AutoML empowers **automated selection** and **configuration of machine learning models** [Feu15]**.**
- **End-to-end AutoML systems** integrate all previous aspects streamlining workflows and ensuring reproducibility [EMS19].

3. Meta Learning

Meta-learning is a machine learning approach that focuses on automatically learning and adapting to new tasks by **leveraging knowledge gained from previous tasks**, enabling efficient transfer of information and improving learning performance [HKV19].
Meta-learning offers **adaptation to diverse datasets**, **efficient model selection**, **hyperparameter optimization**, **transfer learning** and **knowledge transfer**.

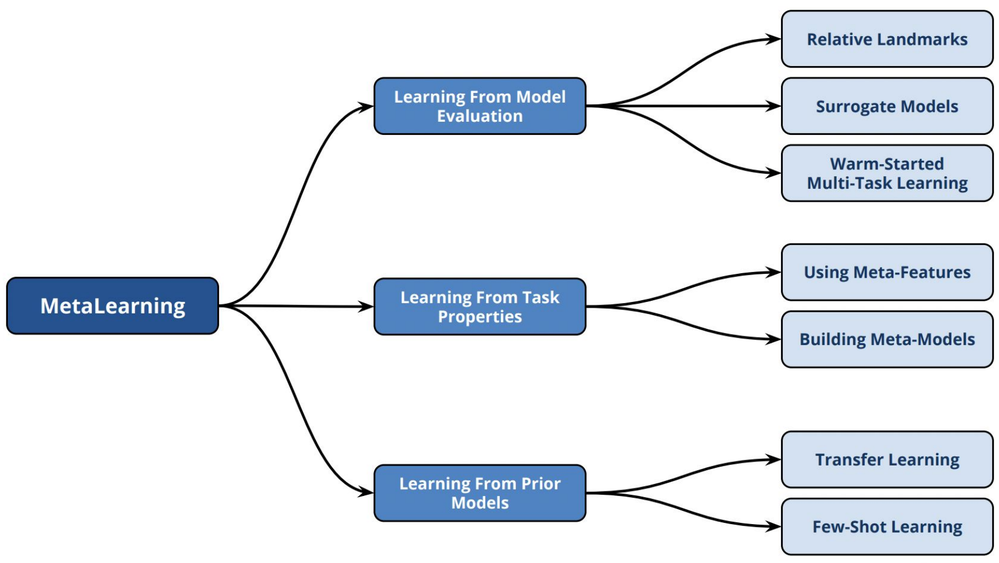

Figure 2: Meta-Learning Techniques Taxonomy [EMS19]

4. Meta Features

A **meta-feature** is a **quantitative or qualitative measure extracted from datasets** that captures higher-level properties of the data, serving as descriptors that effectively capture the structure and complexity of a dataset [Feu+15].

- **Extraction**: Most common are statistical measures such as mean, variance or correlation. Feature selection methods can be applied to identify the most informative meta-features [EMS19]. Alternative methods exists i.e. the use of relative landmarks [LBV12].
- **Application**: Meta model associates meta-features with optimal algorithm configurations, facilitating the selection of algorithms and hyperparameters for new datasets [HKV19]. They are not generated at once but like machine learning models iteratively improved over time [Feu+15].

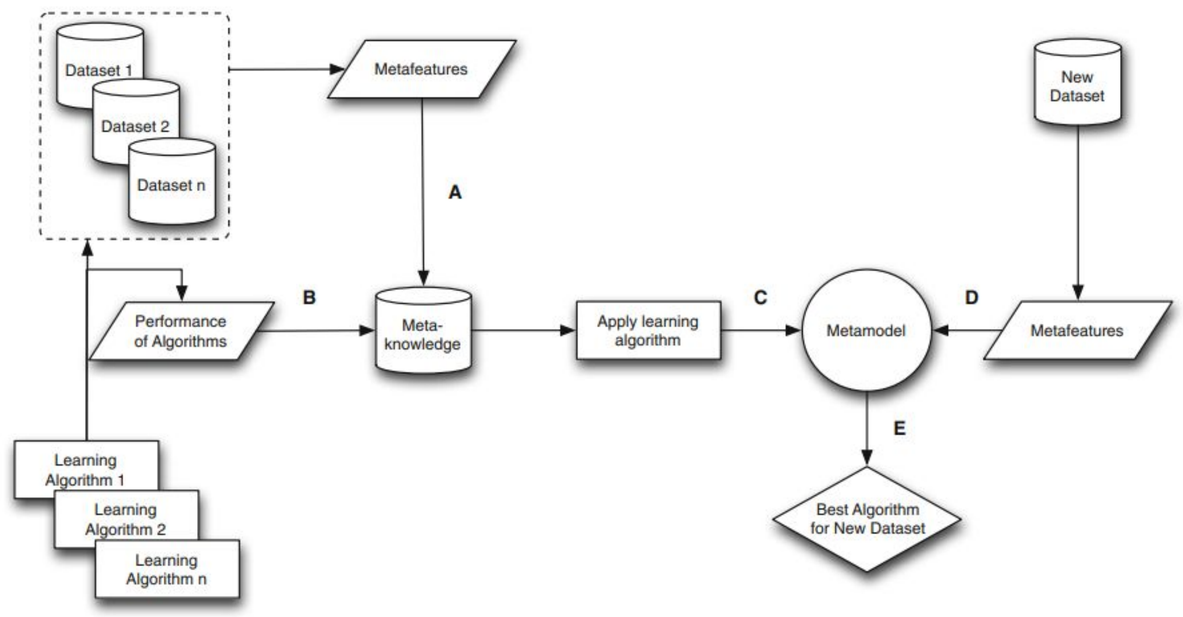

Figure 3: AutoML pipeline with meta-learning [PSM16]

- **Performance**: Default hyperparameters based on meta-features performed comparably to tuned hyperparameters in a study on SVMs and random forest models across 59 datasets [WMV20].

5. Conclusion

AutoML **democratizes machine learning** by minimizing human intervention and incorporating meta-learning, which utilizes meta-features to optimize model configurations. But their reliance on statistical measures may **limit** their ability to capture **contextual understanding**. With the increasing prominence of AI, AutoML is expected to **gain more popularity** in the coming years. To enhance the future applications of meta learning and meta features, it is important to **explore methods for integrating contextual understandings** into the meta learning process.

References

[Feu15] Matthias Feurer et al. "Efficient and Robust Automated Machine Learning". In: Advances in Neural Information Processing Systems. Ed. by C. Cortes et al. Vol. 28. Curran Associates, Inc., 2015
[EMS19] Radwa Elshawi, Mohamed Maher, and Sherif Sakr. Automated Machine Learning: State-of-The-Art and Open Challenges. 2019
[LBV12] Rui Leite, Pavel Brazdil, and Joaquin Vanschoren. "Selecting Classification Algorithms with Active Testing". In: Machine Learning and Data Mining in Pattern Recognition
[HKV19] Frank Hutter, Lars Kotthoff, and Joaquin Vanschoren, eds. Automated Machine Learning. The Springer Series on Challenges in Machine Learning
[WMV20] Hilde J. P. Weerts, Andreas C. Mueller, and Joaquin Vanschoren. Importance of Tuning Hyperparameters of Machine Learning Algorithms. 2020
[PSM16] Fábio Pinto, Carlos Soares, and João Mendes-Moreira. "Towards Automatic Generation of Metafeatures". In: Advances in Knowledge Discovery and Data Mining.