3rd Conference on Production Systems and Logistics

Function Analysis For Selecting Automated Machine Learning Solutions

Günther Schuh[1], Max-Ferdinand Stroh[1], Justus Benning[1], Stefan Leachu[1], Katharina Schmid[1]

[1]*FIR, Institute for Industrial Management at RWTH Aachen University, Aachen, Germany*

Abstract

Methods of machine learning (ML) are notoriously difficult for enterprises to employ productively. Data science is not a core skill of most companies, and acquiring external talent is expensive. Automated machine learning (Auto-ML) aims to alleviate this, democratising machine learning by introducing elements such as low-code / no-code functionalities into its model creation process. Multiple applications are possible for Auto-ML, such as Natural Language Processing (NLP), predictive modelling and optimization. However, employing Auto-ML still proves difficult for companies due to the dynamic vendor market: The solutions vary in scope and functionality while providers do little to delineate their offerings from related solutions like industrial IoT-Platforms. Additionally, the current research on Auto-ML focuses on mathematical optimization of the underlying algorithms, with diminishing returns for end users. The aim of this paper is to provide an overview over available, user-friendly ML technology through a descriptive model of the functions of current Auto-ML solutions. The model was created based on case studies of available solutions and an analysis of relevant literature. This method yielded a comprehensive function tree for Auto-ML solutions along with a methodology to update the descriptive model in case the dynamic provider market changes. Thus, the paper catalyses the use of ML in companies by providing companies and stakeholders with a framework to assess the functional scope of Auto-ML solutions.

Keywords

Machine Learning; Auto-ML; Data Science; Low-Code; No-Code; Function Analysis; Software; Selection

1. Introduction

Production and logistics continue to be one of the most promising fields for the application of machine learning; however, as data science is not one of the core skills of manufacturing companies, they are severely affected by the scarcity of experts in this field [1]. Automated machine learning (Auto-ML) addresses this problem by democratizing and simplifying the value creation process of machine learning, from data collection to model validation [2]. In recent years, software providers have created a plethora of user friendly software solutions that aim to support companies without expertise in this field to create value from data [3]. This potentially helps companies creating their own machine learning models for production, logistics and supporting processes, for example with Natural Language Processing (NLP). However, the multitude of solutions and use cases gives rise to another problem: the challenge of selecting the correct software for the specific needs of a company.

This paper presents the first in a series of models that aim to ultimately yield a structured selection process for manufacturing companies. To design a process that integrates both the most recent trends in Auto-ML and the individual requirements of the company using it, the provider perspective as well as the user perspective need to be integrated (see Figure 1). This paper touches upon the provider perspective (Model I) and thus lays the groundwork for said selection process by describing the functions (or features, synonymously used in this paper) currently offered by Auto-ML-solutions. The overall research goal is to provide businesses with a practicable approach to unlock this key technology.

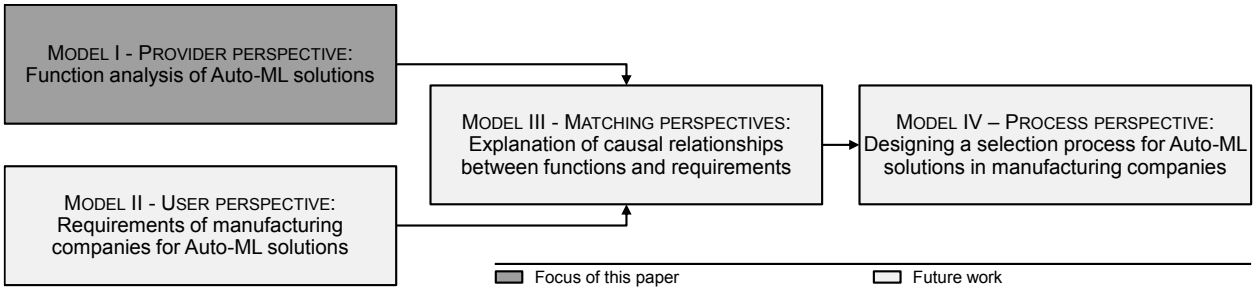

Figure 1: Context of this paper in the overall research goal

The interim results of Model I provide an up-to-date overview of the dynamic vendor market for auto ML solutions. The complexity of this market is increasing as the range of functions offered by solutions is growing rapidly. [4]. Although surveys on Auto-ML solutions are available (see chapter 2.1), companies are overburdened with building up know-how about AI and ML. Their current average level of knowledge is still lacking [5]. Especially in the case of emerging technologies, companies find it difficult to free up the resources to conduct targeted technology research [6,7]. The model developed here addresses this challenge by aggregating multiple current sources and structuring the information hierarchically.

The state of the art of Auto-ML literature focuses mainly on performance benchmarking and applying Auto-ML to existing prediction problems, as will be shown in Section 2.1. While these efforts help advance the maturity of the technology and aid scientific progress, they are of little concern for end users looking for a business solution. However, some sources focusing on the functions provided to potential users can be found [8–10]. Still, the resources found only mention parts of the functionality range that modern Auto-ML solutions provide. Their goal is to inform executive stakeholders on a surface level. Thus, the goal of this paper is to analyse and aggregate existing literature to build an extensive descriptive model. The paper proposes a hierarchical structure of the model, so that even laypersons can quickly draw information from the results without losing any of the information depth.

2. Methods

The descriptive model was built using a two-step-approach: First, an integrative literature review was conducted to source information in a structured manner. Then, the ARIS-toolset (Architecture of integrated information systems) was used to build a hierarchical model of the functions offered by Auto-ML solutions.

2.1 Literature review

The literature was sourced and selected using the integrative literature review technique by TORRACO [11], specifically the approach of synthesizing new knowledge about an emerging topic. Currently, literature about the user-facing functions of Auto-ML solutions is sparse. Presently, research mostly focuses on optimizing the performance of underlying algorithms or applying Auto-ML in a novel way.

Literature was collected from five scientific publication portals: arXiv [12], SpringerLink [13], ScienceDirect [14], Packtpub [15] and ETH Research Collection [16]. The search terms used were “automl

features" and "automl survey" (except on the Packtpub site, where the more general "automated machine learning" was used due to smaller total publication volume). The collection was conducted from 20th of December 2021 till 3rd of March 2022. A funnel-type approach was used, narrowing down the results from the initial results list to a short list of relevant sources that were then used in the second phase of the research, building the model. The selection process is presented in Figure 2.

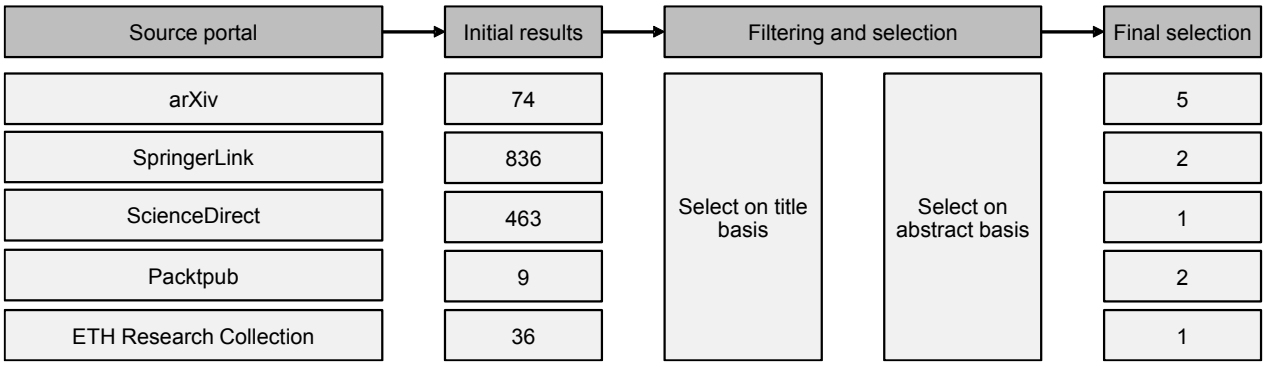

Figure 2: Literature sourcing and selection

The selection on title level was conducted by discarding all results that just presented or assessed the performance of underlying algorithms (e.g., "Performance review…", "A new algorithm…"), which was most of the initial results. The remaining potential sources were screened on an abstract level to ensure that they were directly or indirectly describing features or functions of Auto-ML solutions. In this step, a lot of sources were discarded that used the term "feature" or "function" in a mathematical sense (e. g. "Feature space" or similar expressions). In total, eleven publications were selected to build the hierarchical function model. They are listed in Table 1. After deciding on the final selection, all mentions of functions or features of Auto-ML tools were extracted from the items in the table. The total number of mentions found in the respective publications are listed in the table as well (column "Number of functions"). All sources together yielded 275 functions and features (at this point including duplicates).

Table 1: Final selection of sources for the descriptive model

Source	Type	Sourced from platform	Number of functions
Truong et al. [10]	Journal paper	arXiv	31
Li et al. [17]	Journal paper	ETHZ	16
Humm, Zender [18]	Journal paper	Springerlink	5
Lee et al. [19]	Journal paper	Sciencedirect	8
Das [20]	Monograph (book)	Packtpub	42
He et al. [21]	Journal paper	arXiv	24
Hutter et al. [22]	Compilation (book)	Springerlink	22
Masood, Sherif [23]	Monograph (book)	Packtpub	42
Elshawi et al. [24]	Journal paper	arXiv	32
Zöller, Huber [25]	Journal paper	arXiv	41
Yao et al. [26]	Journal paper	arXiv	12

2.2 Building the hierarchical model

The model was built using the ARIS-toolkit by SCHEER [27], specifically its "function perspective", which provides an interdisciplinary framework for modelling the functions of information systems and business processes for decision makers in IT as well as in management. To achieve the hierarchical structure of the

model, which will help satisfy the individual information depth requested by stakeholders in potential integration projects, the functions are classified into four levels: function bundles, functions, partial functions and at the most granular level, elementary functions [28]. The syntax used in modeling the functions was "verb" + "noun" [28]. The first step of model building was removing all duplicate mentions of functions between the sources. Then, a preliminary classification of the functions was conducted, dividing them across the four levels according to SCHEER. Thirdly, the root node of the hierarchy was defined (the "function bundle" at the topmost level) and divided into functions, following the top-down approach recommended by SCHEER [27]. The structuring criterion used was the value creation process [27]. Following this, a bottom-up approach was used to connect the more granular functions to the above, aggregated layers. This combined approach has the benefit of a validation of the model's internal consistency. The final model comprises 149 functions divided across four levels.

3. Results

The 149 functions will be presented in nine parts, following the structure of the second level (the "functions"). The root node of the hierarchical model (the first level, called the "function bundle") is not visualized in the following figures, as it would simply repeat. This root node includes all subordinate functions and was named "Automate machine learning process" to provide a high amount of generality for the subordinate functions.

The first cluster starts at the data collection step. The main differentiating partial and elementary functions of Auto-ML solutions appear to be the different data types they can handle. Some are even fit to process unstructured data like images and videos (see Figure 3).

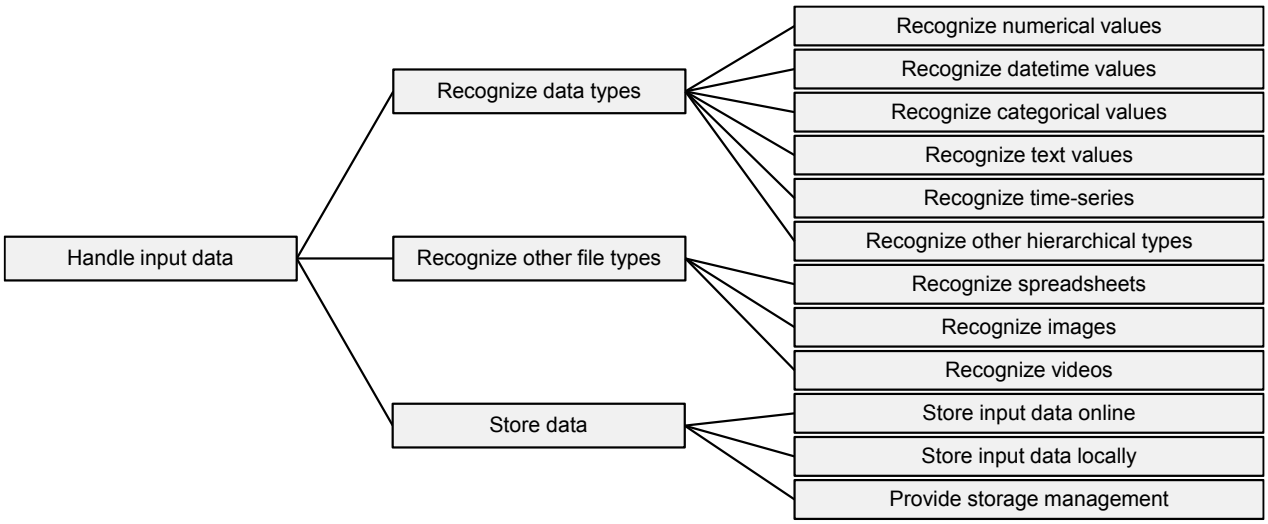

Figure 3: Handle input data

The second group comprises the functions concerned with the pipeline structure (see Figure 4). Three different approaches were found. The first (and most complex) function is only offered by a few tools. These tools can automatically create an entire machine learning pipeline and are not bound by a rigid or even sequential structure. The more common approach is to specify a fixed pipeline structure, in which the solution searches for the optimal pre-processing / model combination for the given problem. Lastly, some tools do not automate any tasks on the pipeline level but guide the user through the respective steps and recommend options that can be chosen.

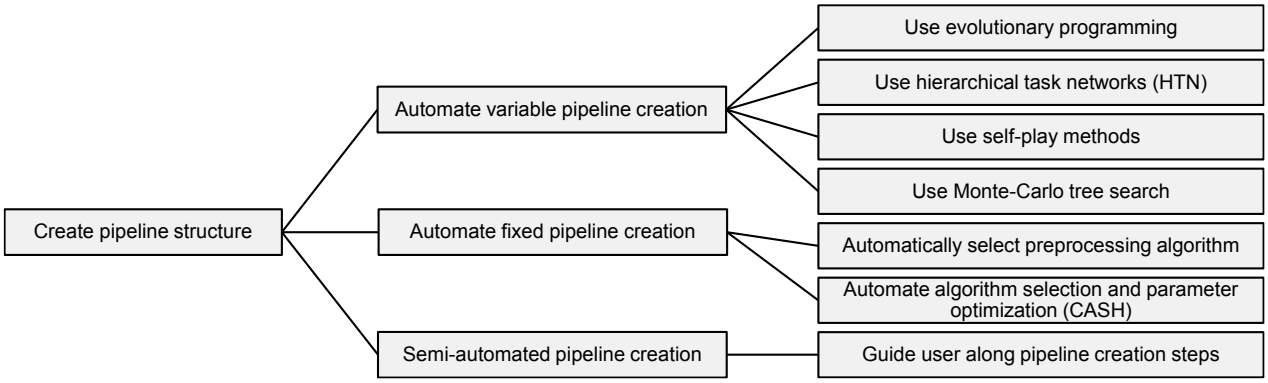

Figure 4: Create pipeline structure

The third cluster consists of partial functions that help prepare the stored data for further processing by cleaning unwanted anomalies, transforming and substituting data types and imputing missing values. This task is optionally simplified by some systems by visualizing the data in terms of different properties (see Figure 5).

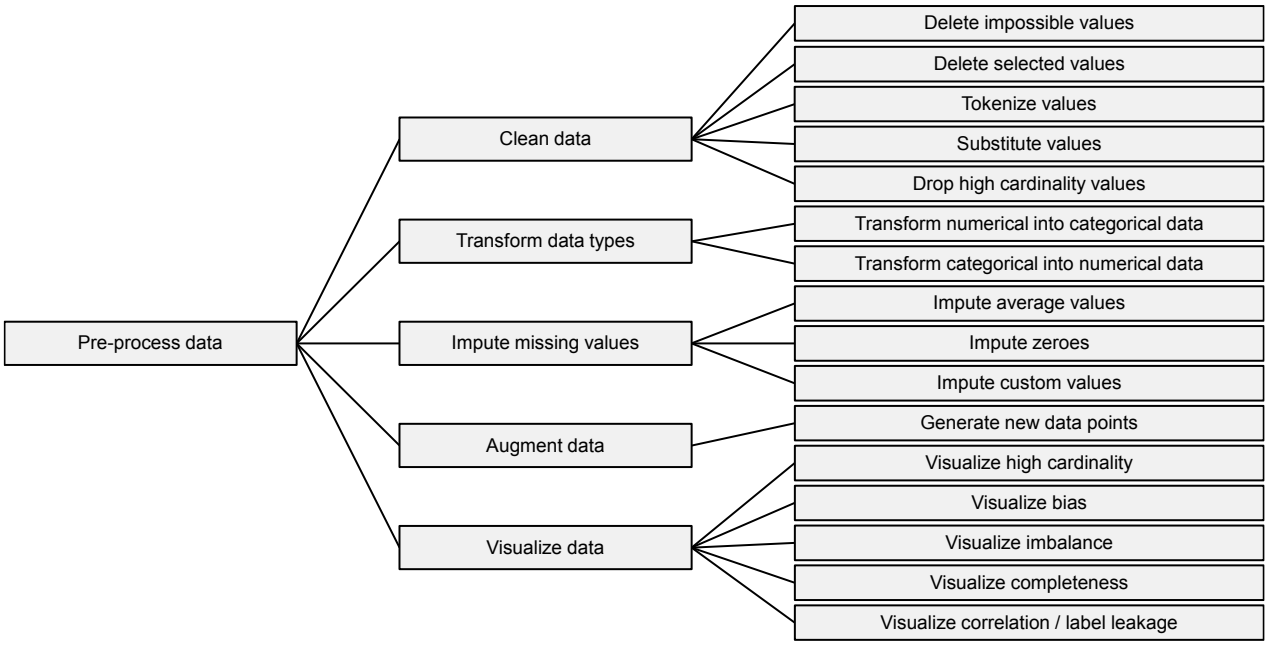

Figure 5: Pre-process data

The fourth cluster is occupied with feature engineering, a task that is particularly important for machine learning and normally requires experience in the field of data science. Well-selected features make the subsequent machine learning process more precise and improve the model's performance. Here, the Auto-ML solutions support users by providing functions for rescaling and grouping and aiding in a pre-selection of possible features. Some also feature more advanced techniques like feature extraction, generation, or dimensionality reduction, which could help unexperienced users in training valid models on sparse, noisy, or highly dimensional datasets (see Figure 6).

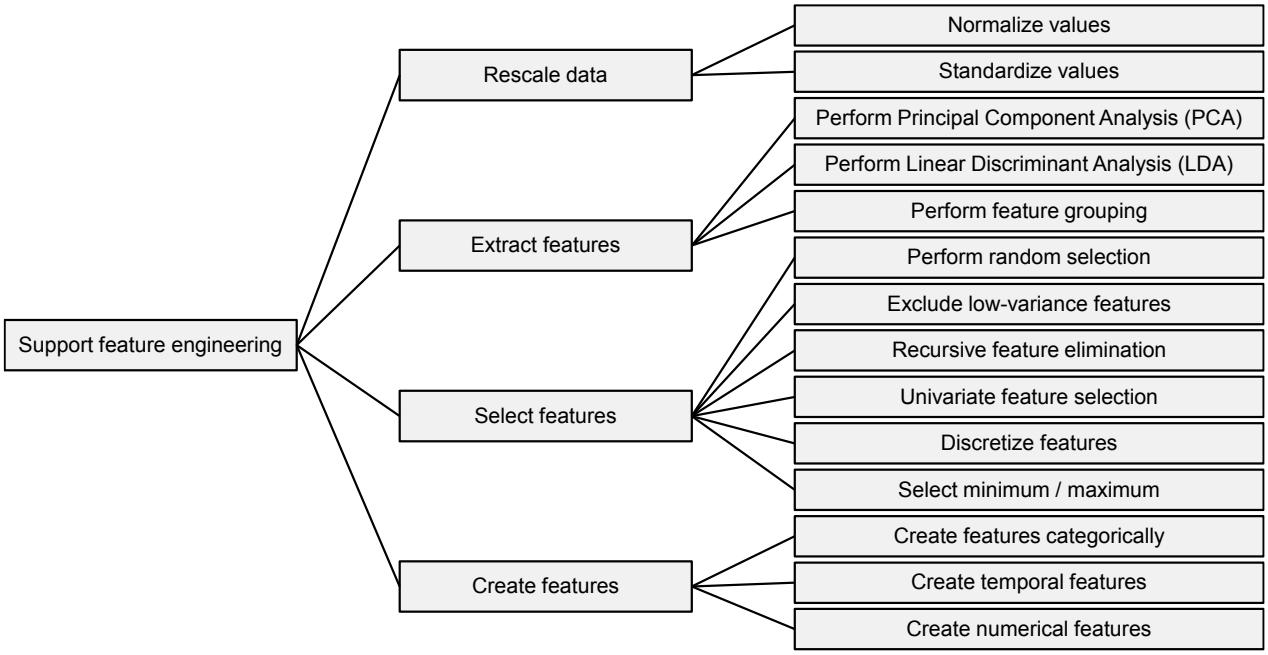

Figure 6: Support feature engineering

Approaching the actual modelling step, the research showed that different tools offer support for different kinds of machine learning problems. While the most popular problem type for Auto-ML appears to be supervised problems, some solutions offer a wider range of learning and modelling types (see Figure 7).

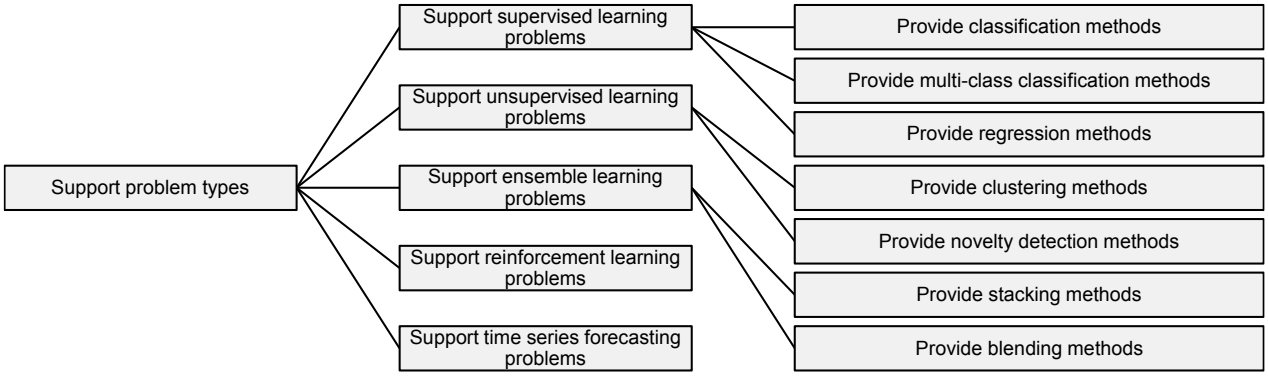

Figure 7: Support problem types

The sixth cluster is the biggest, as it represents the core back-end function of Auto-ML-Software. Here, the models are selected and their hyperparameters (i.e., the structural characteristics) are calculated based on the dataset (see Figure 8). A variety of techniques is employed by the different solution providers. Next to the variety of different model types, solutions also differ in their automation approach: Some choose combined algorithm selection and hyperparameter optimization (CASH) while others opt for letting the user choose a model and then optimizing the hyperparameters separately (conventional hyperparameter optimization or HPO). The generation of neural networks (NAS, short for neural architecture search) is closely related to HPO but can employ different search algorithms and thus is listed separately. Furthermore, some solutions provide quality-of-life-features (such as early stopping) to improve usability of the solution.

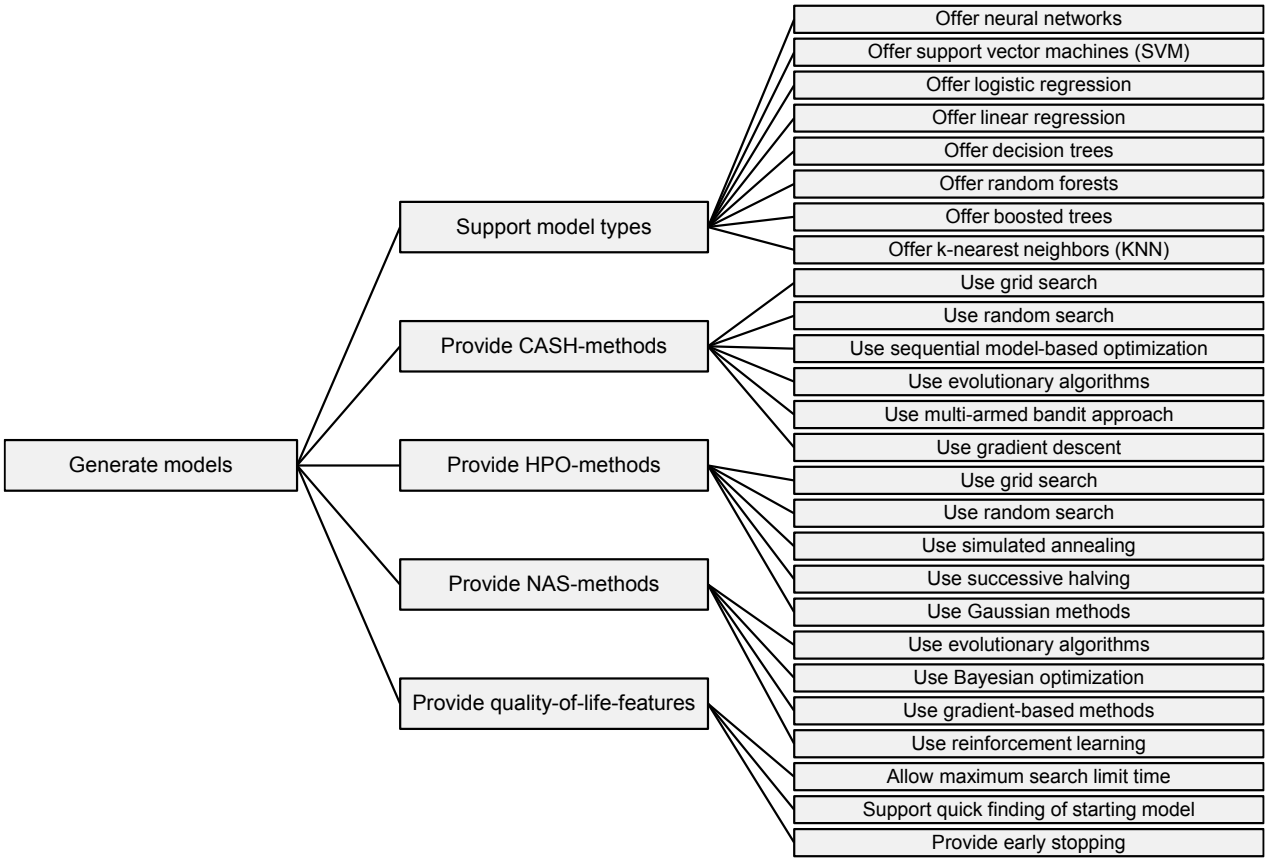

Figure 8: Generate models

The seventh cluster lists functions with which the user can review the results and choose a model to go forward with. The functions of Auto-ML software in this step revolve around helping the user make an informed decision about what is the optimal model for their use case. This can be done by ranking the models according to different criteria and visualizing their differences (see Figure 9).

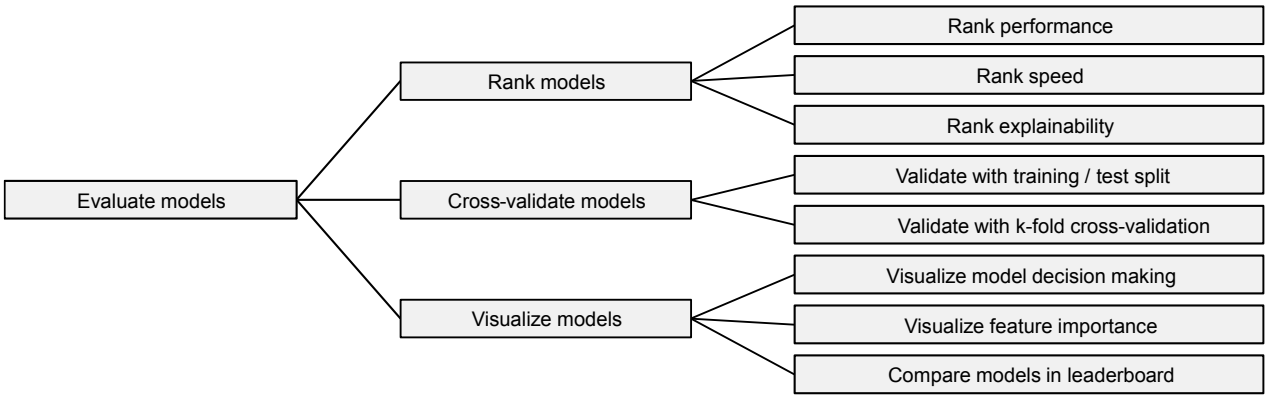

Figure 9: Evaluate models

After a model has been chosen and validated, some solutions provide support for end user in employing the models productively. Normally, this would demand some expertise in software engineering. However, some solutions even include their own cloud-based service that lets users host models online, significantly reducing the effort of providing the model as a service (see Figure 10).

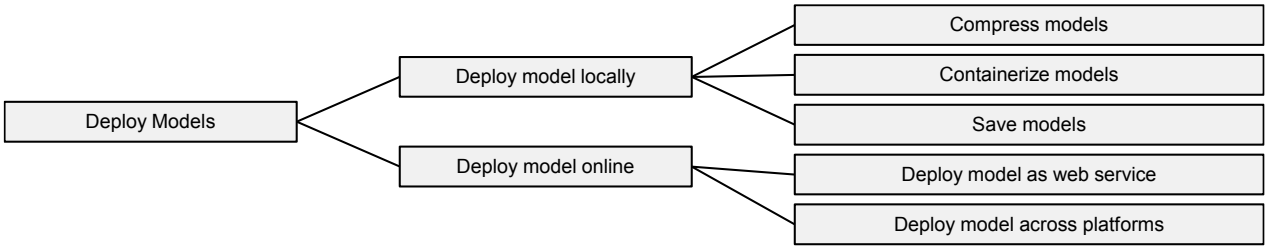

Figure 10: Deploy models

To offer support for users in the entire lifecycle of a machine learning model, some solutions even provide features for machine learning operations (ML-Ops), ensuring the correct and safe use of the model predictions in the field. Features include guardrails to minimize unexpected behaviour and providing model alerts to boost safety when using the model to steer sensitive processes. Lastly, meta-learning functionalities help to create new models more efficiently by learning best-practices from previous modelling efforts (see Figure 11).

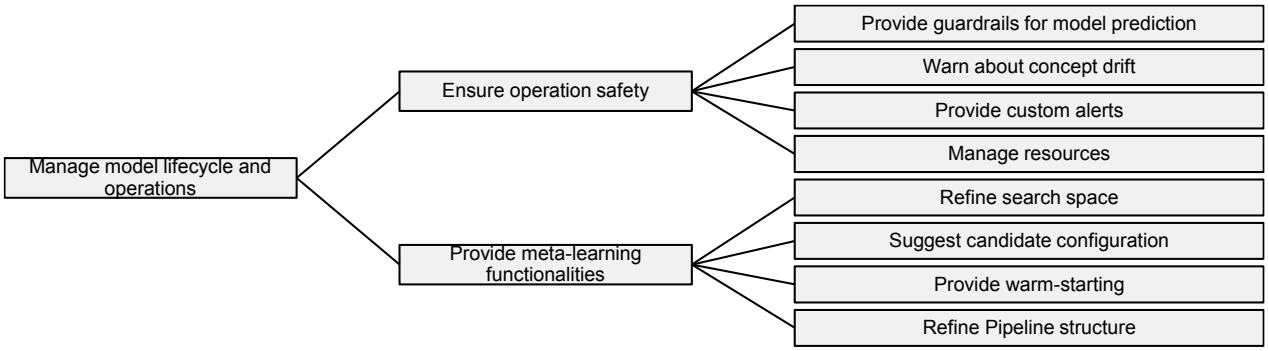

Figure 11: Optimize model operations

4. Discussion

The provision of automated machine learning software is an emerging, highly competitive market and thus the above results are subject to change. New functions may be implemented into established solutions, or a new competitor could enter the market and provide a disruptive set of features that changes the playing field. Thus, the descriptive model provided represents a snapshot of the feature sets available for a limited amount of time. However, care was taken to implement a modular model structure as well as a reproducible approach. This means that the model can easily be augmented and adapted in case of changes.

5. Summary and Outlook

The aim of the paper was to develop a new descriptive model of the functions of Auto-ML software solutions. A hierarchically structured model was chosen to give a dynamic depth of information. The information for building the model was sourced using the integrative literature review technique by TORRACO, while the model itself was built with the ARIS methodology described by SCHEER. Like mentioned in the introduction of the paper, the descriptive model of Auto-ML functions is only the first step to create a structured process, with which companies can select a fitting Auto-ML solution for their needs. To enhance practical usability, the user perspective must be considered as well. Companies have differing requirements and needs regarding machine learning, from performance to transparency and security. These research questions will be tackled in future publications.

Acknowledgements

This effort has been funded by the Leitmarkt.NRW program (EFRE-0801386) and the European Union in the European Regional Development Fund (EFRE) under the name “TechRad”. The authors wish to acknowledge the EFRE for their support.

References

[1] Reder, B., 2021. Studie Machine Learning 2021. https://www.lufthansa-industry-solutions.com/de-de/studien/idg-studie-machine-learning-2021. Accessed 2 September 2021.

[2] Chen, Y.-W., Song, Q., Hu, X., 2021. Techniques for Automated Machine Learning. SIGKDD Explor. Newsl. 22 (2), 35–50.

[3] Das, P., Perrone, V., Ivkin, N., Bansal, T., Karnin, Z., Shen, H., Shcherbatyi, I., Elor, Y., Wu, W., Zolic, A., Lienart, T., Tang, A., Ahmed, A., Faddoul, J.B., Jenatton, R., Winkelmolen, F., Gautier, P., Dirac, L., Perunicic, A., Miladinovic, M., Zappella, G., Archambeau, C., Seeger, M., Dutt, B., Rouesnel, L., 2020. Amazon SageMaker Autopilot: a white box AutoML solution at scale. http://arxiv.org/pdf/2012.08483v2.

[4] Statista Research Department, 2022. Number of AI/ML service offerings at hyperscale CSPs worldwide 2020-2021, by provider. Statista Research Department. https://www.statista.com/statistics/1268286/worldwide-ai-machine-learning-service-offerings-hyperscalers/. Accessed 7 March 2022.

[5] Kaul, A., Schieler, M., Hans, C., 2019. Künstliche Intelligenz im europäischen Mittelstand: Status quo, Perspektiven und was jetzt zu tun ist. https://www.uni-saarland.de/fileadmin/upload/lehrstuhl/kaul/Universita%CC%88t_des_Saarlandes_Ku%CC%88nstliche_Intelligenz_im_europa%CC%88ischen_Mittelstand_2019-10_digital.pdf. Accessed 7 March 2022.

[6] Cetindamar, D., Phaal, R., Probert, D.R., 2016. Technology management as a profession and the challenges ahead. Journal of Engineering and Technology Management 41 (4), 1–13.

[7] Schuh, G., 2012. Innovationsmanagement. Springer Berlin Heidelberg, Berlin, Heidelberg, 422 pp.

[8] Carlsson, K., Gualtieri, M., 2019. The Forrester New Wave™: Automation-Focused Machine Learning Solutions, Q2 2019: The Nine Providers That Matter Most And How They Stack Up, Cambridge, 19 pp. Accessed 17 February 2021.

[9] Dilmegani, C., 2022. AutoML Tech: Products of 2022 Compared: in-Depth Guide. https://research.aimultiple.com/automl-comparison/. Accessed 12 January 2022.

[10] Truong, A., Walters, A., Goodsitt, J., Hines, K., Bruss, C.B., Farivar, R., 2019 - 2019. Towards Automated Machine Learning: Evaluation and Comparison of AutoML Approaches and Tools, in: 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI). 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), Portland, OR, USA. 04.11.2019 - 06.11.2019. IEEE, pp. 1471–1479.

[11] Torraco, R.J., 2005. Writing Integrative Literature Reviews: Guidelines and Examples. Human Resource Development Review 4 (3), 356–367.

[12] Cornell University, 2022. arXiv. https://arxiv.org/. Accessed 12 January 2022.

[13] Springer Verlag, 2022. SpringerLink. https://link.springer.com/. Accessed 12 January 2022.

[14] Elsevier B. V., 2022. Sciencedirect. https://www.sciencedirect.com/. Accessed 12 January 2022.

[15] Packt Verlag. Packtpub. https://www.packtpub.com/. Accessed 12 January 2022.

[16] ETH Zürich, 2022. ETH Research Collection. https://www.research-collection.ethz.ch/. Accessed 12 January 2022.

[17] Li, Y., Wang, Z., Ding, B., Zhang, C., 2021. AutoML: A Perspective where Industry Meets Academy, in: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. KDD '21: The

27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event Singapore. 14.08.2021-18.08.2021. ACM, New York, NY, USA, pp. 4048–4049.

[18] Humm, B.G., Zender, A., 2021. An Ontology-Based Concept for Meta AutoML, in: Maglogiannis, I., Macintyre, J., Iliadis, L. (Eds.), Artificial Intelligence Applications and Innovations, vol. 627. Springer International Publishing, Cham, pp. 117–128.

[19] Lee, K.M., Yoo, J., Kim, S.-W., Lee, J.-H., Hong, J., 2019. Autonomic machine learning platform. International Journal of Information Management 49, 491–501.

[20] Das, S., 2018. Hands-On Automated Machine Learning: A beginner's guide to building automated machine learning systems using AutoML and Python, 1st ed. Packt Publishing Limited, Birmingham, 282 pp.

[21] He, X., Zhao, K., Chu, X., 2021. AutoML: A Survey of the State-of-the-Art. Knowledge-Based Systems 212 (3), 106622.

[22] Hutter, F., Kotthoff, L., Vanschoren, J., 2019. Automated Machine Learning: Methods, Systems, Challenges. Springer International Publishing; Imprint: Springer, Cham, 219).

[23] Masood, A., Sherif, A., 2021. Automated Machine Learning, 1st edition ed. Packt Publishing; Safari, Erscheinungsort nicht ermittelbar, Boston, MA, 312 pp.

[24] Elshawi, R., Maher, M., Sakr, S., 2019. Automated Machine Learning: State-of-The-Art and Open Challenges. http://arxiv.org/pdf/1906.02287v2.

[25] Zöller, M.-A., Huber, M.F., 2021. Benchmark and Survey of Automated Machine Learning Frameworks. Journal of Artificial Intelligence Research 70 (1), 409–474.

[26] Yao, Q., Wang, M., Chen, Y., Dai, W., Li, Y.-F., Tu, W.-W., Yang, Q., Yu, Y. Taking Human out of Learning Applications: A Survey on Automated Machine Learning.

[27] Scheer, A.-W., 1998. ARIS - Modellierungsmethoden, Metamodelle, Anwendungen, Dritte, völlig neubearbeitete und erweiterte Auflage ed. Springer Berlin Heidelberg, Berlin, Heidelberg, 219 pp.

[28] Krcmar, H., 2015. Informationsmanagement, 6. überarb ed. Springer, Berlin [u. a.].

Biography

Günther Schuh (*1958) is head of the chair of Production Systems (WZL-PS) at RWTH Aachen University and member of the directorates of the Machine Tool Laboratory (WZL) at the RWTH, Fraunhofer Institute for Production Technology (IPT) and Director of the FIR at RWTH Aachen University.

Max-Ferdinand Stroh (*1991) is a researcher at FIR at RWTH Aachen University since 2017 in the department Information Management. He is head of the department Information Management. His scientific work is focused on the practical application of AI, smart products, and IT-OT-Integration.

Justus Aaron Benning (*1995) is a researcher at FIR at RWTH Aachen University in the department of Information Management since 2019. He is leading the group Information Logistics and is head of software development in his department. While his degree is in mechanical engineering and business administration, he spent a semester abroad at Korea University in Seoul to focus on the business applications of artificial intelligence.

Stefan Leachu (*1996) is a researcher at FIR at RWTH Aachen University in the department of Information Management since 2021. With his degree in computer science, he enjoys linking the technical issues of various technologies with the value-added potential that can be achieved in the business world.

Katharina Schmid (*1992) studied business economics at DHBW Ravensburg, specializing in logistics and controlling. After graduation, she worked as a project manager in supply chain projects for an international automotive company. To further improve her skills, she then studied computer science at RWTH Aachen University. She joined FIR at RWTH Aachen University in 2019 in the department of Information Management.