Automated machine learning for dynamic energy management using time-series data

Wang, C.

Chapter 7

Conclusions and future work

In this thesis, we discussed the problem of automated machine learning for time-series forecasting and its application to short term load forecasting. Given a time-series data set, a set of machine learning models, and sets of hyperparameters of the machine learning models, we were interested in finding the machine learning model and the associated hyperparameters which achieve the best accuracy on the time-series data sets.

7.1 Summary

Chapter 1 provided an introduction to the topic of study in this thesis – machine learning, automated machine learning, and time-series forecasting. We provided the research questions we aimed to solve including how the current AutoML techniques are developed, how can AutoML techniques be used for short term load forecasting tasks and how can AutoML techniques be used for time-series forecasting tasks. We also provided the thesis outline in this chapter.

Chapter 2 introduced the application problem we aimed to solve. We provided the definition of automated machine learning for time-series forecasting problems by introducing key concepts and notations which are used later in the thesis. We also introduced the background of our application – short term load forecasting. To solve short term load forecasting tasks, we needed to understand what effects may influence load signals and select models to make the forecasting. In later chapters, we used AutoML techniques for model selection.

Chapter 3 provided an overview of the methodology of automated machine learning, including not only the techniques we later used to solve time-series forecasting problems, but also other interesting recent work in the field of automated machine learning, such as hyperparameter optimisation, automated machine learning systems, and neural architecture search. In this chapter, we answered the research question:

> *What AutoML techniques are available and potentially usable for time-series forecasting and other machine learning problems?*

Chapter 4 investigated the use of automated machine learning for short term load forecasting in order to answer the research question:

> *How can AutoML techniques be used efficiently for the regression problem to address problems such as dynamic energy management?*

Short term load forecasting, which is an application of time-series forecasting, is studied in this chapter. In our case study, we used AutoML techniques to do short term load forecasting with two real-world industrial data sets. One is a publicly available data set from the UCI repository (Candanedo, 2017; Dua and Graff, 2017). The data set indicates the energy usage of appliances in a household. The second one is the electricity consumption of an office building of Honda R&D with around 200 employees.We achieved higher accuracy than the grid search baseline from the original authors on the first data set, and on the second data set, we used only a few hours to achieve a result that closely matches the accuracy of a six-month human expert's project.

Chapter 5 looked at how to enhance AutoML for single-step ahead time-series forecasting by adding time-series features, and aimed to answer our third research question:

> *How can AutoML be used efficiently for single-step time-series forecasting?*

This chapter introduced enhanced automated machine learning with feature engineering techniques for time-series forecasting. This included (i) AutoML with automated window size selection, (ii) AutoML with time-series feature engineering and (iii) AutoML with both automated window size selection and time-series feature engineering. Specifically, we used a state-of-the-art automated machine learning system, auto-sklearn, and a well-known time-series feature extraction tool, tsfresh. To critically assess the approaches we studied, time-series forecasting was performed on a diverse set of benchmarks, covering synthetic and real-world problems using statistical techniques, machine learning techniques, and state-of-the-art automated machine learning techniques. Our empirical results indicated that our newly developed methods yield improved accuracy for 17 out of 20 data sets we studied in terms of *RMSE*.

Chapter 6 investigated the benefits of using AutoML with time-series features on multi-step forecasting, with the fourth research question in mind:

How can AutoML be used efficiently for multi-step time-series forecasting?

We studied automated machine learning with enhanced time-series features for multi-step time-series forecasting. In this chapter, we demonstrated how automated machine learning can be enhanced with feature engineering techniques for multi-step time-series forecasting. We combined auto-sklearn and tsfresh. We optimised both of the machine learning pipelines and the size of the window over which time-series data are used for predicting future time-steps. We evaluated these approaches with statistical techniques, machine learning techniques and state-of-the-art automated machine learning techniques, on a diverse set of benchmarks for multi-step time-series forecasting. Our empirical results indicated significant potential for improving the accuracy of multi-step time-series forecasting by using automated machine learning in combination with automatically optimised feature extraction techniques.

In summary, this thesis provided advanced techniques for time-series forecasting by applying AutoML. These results clearly demonstrate that the use of AutoML techniques on the time-series forecasting tasks allows to obtain highly accurate models fully automatically, without the need for making design choices by human experts.

7.2 Future work

Although considerable progress has been made in the context of this thesis, much work remains to be done Some perspectives for this future research are included here and are briefly discussed.

Meta-learning for time-series analysis: In time-series forecasting research, there is still huge uncertainty in the task of selecting an appropriate forecasting method for a problem. It is not only the individual algorithms that are available in great quantities; combinational approaches (the combinations of individual approaches, for example, the combination of meta-learning and machine learning (Feurer et al., 2015)) have been equally popular in the last decades. The question of whether to choose the most promising individual method or a combination is also not straightforward to answer. Usually, expert knowledge is needed to make an informed decision, but in many cases, this is not feasible due to lack of time, money and manpower. Some work has been done on meta-learning for time-series analysis with some demonstrable success (Oreshkin et al., 2021). The combination of meta-learning and AutoML is also not new (Feurer et al., 2015). However, meta-learning for both time-series analysis using AutoML approaches has not been studied yet. Time-series data sets from different domains may show different characteristics, so finding suitable data sets for meta-learning might be challenging.

AutoML for unstructured data: Most research in automated machine learning has focused on supervised learning (classic regression and classification tasks) for tabular data. In NAS, there is a major focus on optimising convolutional and recurrent neural networks. More recently, graph neural networks have also attracted attention (see, e.g., Ding et al., 2021; Gao et al., 2020; Li et al., 2021). Support for other types of structured data that are relevant in many practical applications, such as spatio-temporal data, is still limited. To develop AutoML methods for these types of data, more specialised search spaces need to be defined by adding other hyperparameters or preprocessing elements. The search space is where human expertise in designing specialised algorithms can be most easily incorporated in order to render the search process more efficient. Specialised search spaces have been created and used successfully for spatio-temporal data sets (see, e.g., Li et al., 2020c); however, there is substantial room for further work in this area.

Automated machine learning with limited amounts of labelled data:

For many machine learning models, especially for deep neural networks, the size of the labelled training data set has a significant influence on final performance. Platforms such as Google Cloud AutoML provides the service of generating high-quality training data through Google's human labelling service. However, many data instances cannot be easily labelled without domain expertise (e.g., medical diagnoses) or are very difficult to label even with such expertise (e.g., multi-spectral satellite images). There are a lot of time-series data sets without labels available online (e.g., audio). It is difficult to use them to train time-series classification models without the labels. Transfer learning can increase the efficiency of AutoML systems and reduce the need for labelled data sets (Wong et al., 2018). For instance, Salinas et al. (2021) study how incorporating transfer learning in AutoML systems can let NAS systems developed for natural images be used for remote sensing images. Active learning and semi-supervised learning are other related techniques that can help train models with smaller amounts of labelled data. Supporting these tasks within AutoML systems could be achieved by considering objective functions that can assess the quality of the trained models with a much smaller number of labelled instances. Changes to the objectives might render model selection approaches such as cross-validation obsolete, as demonstrated by Chen and Wujek (2020) for the case of active learning. Furthermore, some of the techniques used to improve the performance of AutoML systems for supervised learning may not be applicable anymore (e.g., meta-features, as demonstrated by Li et al. (2019)). There may also be instabilities in performance improvement when using a mixture of labelled and unlabelled instances in semi-supervised learning.

Multi-objective AutoML: A majority of AutoML systems focus on single-objective optimisation based on regression or classification accuracy. Considering more than a single objective can increase the potential of finding models or pipelines that are better suited for specific applications. For NAS, multi-objective optimisation has been considered to create neural network architectures that run on resource-constrained devices, such as mobile phones (Tan et al., 2019). While improving the energy efficiency of models such as CNNs that run on mobile devices is a well-explored area, there are still other objectives that can be considered to improve the quality of models. Solutions to most machine learning problems are often best assessed using multiple performance indicators, such as precision and recall. There are also machine learning problems of an inherently multi-objective nature, such as early time-series classification. As demonstrated by Ottervanger et al. (2021), applying a multi-objective optimisation approach to algorithm selection and hyperparameter optimisation for such problems can provide opportunities to present a more diverse set of Pareto-optimal solutions. Many of the HPO approaches mentioned previously based on reinforcement learning or evolutionary algorithms support multi-objective optimisation. However, in order to become applicable to general AutoML problems, these approaches need to be expanded to support tree-structured search spaces and conditional hyper-parameters.

Trustworthy AutoML: Next to the objectives that target computational performance metrics, it is increasingly desirable to consider objectives or constraints related to the trustworthiness and fairness of machine learning systems (e.g., (König et al., 2020, 2022; Perrone et al., 2021)). Goals such as trustworthiness are not yet crisply defined; therefore, automatically assessing the performance of machine learning systems based on these factors requires additional work on new performance metrics, whose optimisation increases widely accepted notions of trustworthiness.

Larger search spaces and additional design choices: Although research in

AutoML supports training a varied set of different types of model families, the main focus is still on optimising hyperparameters based on a specific search space. Still, a user should decide which kind of search space is relevant. Future research should consider extending these search spaces to allow a model to be trained automatically without much user involvement. Automating the process of configuring models in such a broad search space will make the process even more challenging and thus require even more focus on increasing the efficiency of the search strategy. From a time-series analysis perspective, AutoML should free users from data cleaning and data preprocessing work. AutoML should also be able to determine the relevant search space automatically.

More efficient search space design for NAS: As mentioned previously, a modular cell-based design can largely constrain the search space. However, a post-hoc analysis of various existing cell-based search spaces by Wan et al. (2022) still demonstrates a large amount of redundancy in the hyperparameters considered. Their analysis concludes that only a limited number of operations within cells are important while observing that commonly used cell types, such as reduction cells, are unimportant in increasing performance. This points to opportunities for further optimising the search space, which will hopefully allow for finding better architectures much faster.

AutoML for lifelong learning: So far, most AutoML systems consider a stable data generation process, and AutoML approaches for lifelong machine learning remain largely unexplored. For many data sets collected over time or in a streaming setting, it is possible that the best model found at a specific time is not the best one overall. Being able to identify changes in the data generation process can benefit machine learning models when used for streaming settings. For example, Mcfly (van Kuppevelt et al., 2020) is a library that uses deep learning on time-series data, but it only works on times series classification tasks. Celik and Vanschoren (2021) have tested six different concept drift adaptation strategies that can detect and adapt to the changes in the data generation process. Tetteroo et al. (2022) have implemented and compared these approaches in the context of Covid-19 time-series forecasting, where the data generation process undergoes various changes. We believe that AutoML systems can benefit from these kinds of strategies by incorporating them into the search process.

Automated data science: As discussed by De Bie et al. (2022), to automate data science, many steps, from the point of data generation to decision-making need to be optimised. These steps include (1) data exploration, (2) data engineering, (3) model building, and (4) exploitation. As evident in our extensive review, although the area of AutoML is thriving and has attracted much attention in recent years, most

AutoML research so far has focused on model building and mainly targeted supervised learning problems. To achieve the most ambitious goals of AutoML, problems beyond supervised learning and steps other than model building should be better supported by future AutoML systems. We also note that currently, the use of many AutoML approaches still involves choices that critically depend on expert knowledge, such as the selection of an AutoML system and, in many cases, its configuration for a given use case. Ideally, the levels of expertise required for carrying out these tasks should be reduced by further automation while maintaining a meaningful level of human insight and control.

AutoML for other application domains: While doing this research we realised that most of the work in AutoML focuses on specific tasks such as image classification. In this thesis, we fill the gap between machine learning and time-series forecasting. However, many other interesting problems can also be solved by machine learning, such as survival analysis. To solve these types of tasks, both machine learning knowledge and domain knowledge are important. For instance in engineering, medical and financial domains, different types of preprocessing methods and feature extractors (Coyle et al., 2005; Phinyomark et al., 2014) are used. So far we did not see many platforms that work for these domain knowledge-intensive tasks. These special methods from different domains may potentially be added to an AutoML pipeline to improve accuracy.

Detecting and preventing overfitting: An open problem in AutoML is overfitting. AutoML systems calculate the validation loss based on the training set, which has a finite number of data points. AutoML systems often find pipelines that work very well only on a particular finite set of data points. This has happened in our experiments, and happened to others as well (Cawley and Talbot, 2010). One simple solution to address this is shuffling the training set and the validation set during each evaluation (Lévesque, 2018). Ensemble methods are also used to prevent overfitting. However, there is no best technique to avoid overfitting. The best techniques also vary across different tasks.

Overall, AutoML is still a recent and fast-growing research area. AutoML approaches and systems will become even more powerful, versatile, and usable in future. The real-world use of AutoML systems will increase sharply in the near future for time-series forecasting and other domains. We hope that this thesis can help facilitate this development, by providing an overview to researchers and practitioners new to the area of AutoML and the use of AutoML in time-series forecasting. we see a great need for a focus on the design and principled deployment of AutoML systems and techniques that address real-world problems and application needs.