Building a Model for Time Series Forecasting using AutoML Methods

Vladislav E. Kovalevsky
Laboratory of big data technologies in socio-cyberphysical systems
St. Petersburg Federal Research Center of the Russian Academy of Sciences
St. Petersburg, Russia
0000-0002-0414-906X

Nataly A. Zhukova
Laboratory of big data technologies in socio-cyberphysical systems
St. Petersburg Federal Research Center of the Russian Academy of Sciences
St. Petersburg, Russia
0000-0001-5877-4461

2024 XXVII International Conference on Soft Computing and Measurements (SCM) | 979-8-3503-6370-8/24/$31.00 ©2024 IEEE | DOI: 10.1109/SCM62608.2024.10554133

***Abstract*—The field of machine learning includes various algorithms that use data to train models that can afterwards deal with new previously unseen data. Selecting the most appropriate for a specific data algorithm and manually tuning its hyperparameters could be tedious and time-consuming. Methods and tools for the automation of this task form a field of Automated Machine Learning (AutoML). Most AutoML tools solve the problem of building models for classification and regression tasks based on data attributes without considering changes in a parameter over time. However, several systems already exist that automate the search of models for the time series forecasting problem. The peculiarity of the task of time series forecasting is that, in this case, it is necessary to consider the relationship between measurements and time, not just the diversity and other statistical characteristics of the samples. In this work, we explore several AutoML systems that capable to work with time series. The process of automated search of a model for time series forecasting is considered using the AutoGluon system and dataset containing data about temperature changes in different cities of the world. The effects of different limits and presets on the search of model is shown.**

Keywords—AutoML, Machine learning, Hyperparameters Optimization, Time Series Forecasting

I. INTRODUCTION

The field of machine learning includes various algorithms that can be used to build models that are capable of extracting knowledge from a large amount of data [1]. Such models could be used in dealing with different tasks, for example classification or regression problems. Given a large dataset, one can build and train a model using attributes values as input for the model. A particular task on this list is time series forecasting. The peculiarity of this task is that, in this case, it is necessary to consider the relationship between measurements and time and that the same attribute alternately serves as an input and a target.

Each machine learning algorithm has its own hyperparameters – values that determine its operation. Selecting a suitable algorithm and tuning its hyperparameters is a challenging task. In order to automate this tedious work, a number of solutions have been proposed, giving rise to the field of Automated Machine Learning (AutoML) [2]. Most AutoML tools are devoted to solving the problem of automated search of models for classification and regression and do not consider the relationship between parameters and time. However, to date, several systems have been developed that take into account these special needs and perform a search of models for the task of time series forecasting.

This work was supported by a state grant, project No. FFZF-2022-0006.

II. TIME SERIES FORECASTING

A time series consists of values of some parameters measured with equal time intervals, in the simplest case, measurements of one parameter. Time series data have a natural temporal ordering, which distinguishes time series analysis from cross-sectional studies, in which the observations do not have a natural ordering.

Time series analysis includes various statistical methods that enable the capture of the time series structure and forecasting [3]. Identification of the time series structure helps construct a mathematical model of the process on which the analyzed time series is based. Forecasting future time series values helps in effective decision-making.

Examples of time series are:

- Electrical activity of the brain;
- Precipitation measurements;
- Share prices;
- Annual retail sales;
- Number of heartbeats per minute.

Features of time series include:

- Dependency of data on time;
- The need for equal measurement intervals;
- Inability to change the order of data;
- The same data serves alternately as an input and as a target.

III. AUTOML SOLUTIONS FOR TIME SERIES FORECASTING

Nowadays, a number of solutions have been developed for the automated search of models for time series forecasting. Some of these solutions extend existing AutoML tools to the problem of time series forecasting, while others are specialized ones. The first approach includes solutions such as AutoGluon, AutoKeras, Auto-PyTorch, and FEDOT, which have in their arsenal both tools for building models based on time-independent data and for time series. The latter includes the AutoTS system, designed exclusively for building models for time series forecasting.

The AutoGluon system [4] allows the prediction of the future values of several time series contained in the data, taking into account historical data and other related covariates. AutoGluon combines the use of various machine

979-8-3503-6370-8/24/$31.00©2024 IEEE

Authorized licensed use limited to: SRM Institute of Science and Technology- RamaPuram. Downloaded on September 02,2024 at 09:02:05 UTC from IEEE Xplore. Restrictions apply.

learning algorithms such as ETS and ARIMA, LightGBM, as well as deep learning models such as DeepAR and Temporal Fusion Transformer. AutoGluon is based on the GluonTS library, which in turn is based on the PyTorch and MXNet libraries.

AutoKeras [5] is an AutoML system that searches deep learning algorithms from the Keras library, which is based on the TensorFlow machine learning library from Google.

The Auto-PyTorch system [6] also searches for suitable models among algorithms from the PyTorch library. Its feature is the use of a Portfolio - a set of pre-selected promising configurations of neural networks that serve as a starting point for further search.

AutoTS [7] is a specialized AutoML system designed for building time series models. This system searches for algorithms from several libraries, particularly StatsModels, GluonTS, Sklearn, and TensorFlow. A genetic algorithm is used to select a suitable model.

The FEDOT system [8] is another AutoML solution that uses an evolutionary approach for selecting models for time series forecasting.

A comparison of the above-mentioned AutoML solutions is given in Table 1.

TABLE I. AUTOML TOOLS COMPARISON

Solution	ML library	OS	Parallelization
AutoGluon	GluonTS, PyTorch, MXNet	Linux, MacOS, Windows	On GPU
AutoKeras	Keras, TensorFlow	Linux, Windows	On GPU
Auto-PyTorch	PyTorch	Linux	-
AutoTS	StatsModels, GluonTS, Scikit-Learn, TensorFlow	Linux, Windows	On CPU
FEDOT		Linux, MacOS, Windows	On GPU

The AutoGluon system was chosen for experiments with a suitable time series dataset to examine the capabilities of the AutoML approach in building models for time series forecasting.

IV. TIME SERIES DATA

A. Source data

Open data repositories such as Kaggle [9] and OpenML [10] contain many freely distributed datasets designed for training and testing various machine learning algorithms. However, only a small part of this data is time series and can be used to test the corresponding algorithms. There are also small repositories that store exclusively time series data, such as Time Series Classification [11].

While selecting a dataset for test purposes, preference was given to large datasets (>100000 instances) containing multiple time series simultaneously. The Daily Temperature of Major Cities dataset containing information about temperature changes in various world cities was chosen as a source of such data. This dataset contains data on the average temperature in Fahrenheit for 7 regions (Africa, Asia, Australia, Europe, Middle East, North America, South/Central America), 53 US states and 321 cities, and has the following characteristics:

- number of instances – 2906327;
- number of attributes – 8;
- interval – 1 day.
- measurement period – 26 years (1995 – 2020)

B. Data preprocessing

At the first step, the dataset was cleaned. Instances containing incorrect dates were removed, and outliers were removed using the Z-score. This reduced the dataset by 83990 records. Figure 1 shows an example of temperature measurements in three countries over two years after data cleaning. The figure shows how the average temperature in these countries changed depending on the time of the year.

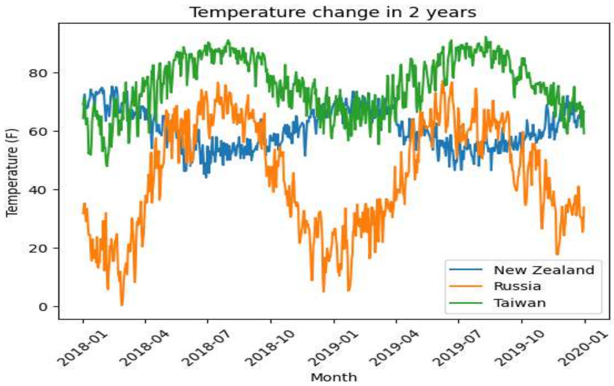

Fig. 1. Average temperature change in 3 countries over 2 years

Next, the data was transformed into a form suitable for use within the chosen AutoML system. The "State" column containing null values outside the US was removed, and the columns containing separate day, month, and year values were combined into a single "Date" column. Since the AutoGluon system requires the input data to be in numeric format, the categorical columns "Region" and "Country" were converted using the LabelEncoder function from the Scikit-Learn library.

C. Additional data preparation

For the work of AutoGluon, it is necessary to define a column containing datetime values and a column with time series identifiers in case the dataset contains several of them. The generated "Date" column was selected as the date column, and the "City" column was selected as the time series identifier, thus declaring that the temperature measurements in each city represent a separate time series.

V. EXPERIMENTS

The AutoGluon system requires specifying the target column containing the measured values, defining the forecast horizon, and setting the metric by which models will be evaluated. If the time series step is not recognized automatically, it could be defined manually. The column containing the average temperature was selected as the target column, and the step interval was defined as one day.

A. Presets

The search process is further customized by setting the time limits allocated for the search and selecting presets that

Authorized licensed use limited to: SRM Institute of Science and Technology- RamaPuram. Downloaded on September 02,2024 at 09:02:05 UTC from IEEE Xplore. Restrictions apply.

affect its quality. The system offers four preset options; see Table 2.

TABLE II. AUTOGLUON SYSTEM PRESETS

Preset	Description	Use cases
fast_training	Fit simple statistical and baseline models + fast tree-based models	Fast to train but may not be very accurate
medium_quality	Same models as in fast_training + deep learning model TemporalFusionTransformer	Good forecasts with reasonable training time
high_quality	More powerful deep learning, machine learning, and statistical forecasting models	Much more accurate than medium_quality, but it takes longer to train
best_quality	Same models as in high_quality, more cross-validation windows	Typically more accurate than high_quality, especially for datasets with few time series

B. Evaluation metrics

AutoGluon offers probabilistic and point metrics for evaluating models during the search process.

Probabilistic metrics:

- Scaled quantile loss (SQL);
- Weighted quantile loss (WQL).

Point forecast metrics:

- Mean absolute error (MAE);
- Mean absolute percentage error (MAPE);
- Mean absolute scaled error (MASE);
- Mean squared error (MSE);
- Root mean squared error (RMSE);
- Root mean squared scaled error (RMSSE);
- Symmetric mean absolute percentage error (SMAPE);
- Weighted absolute percentage error (WAPE).

To calculate these metrics, the following information about the time series and predicted values is used:

$y_{i,t}$ – observed value of time series i at time t.

$f_{i,t}$ – predicted value of time series i at time t.

N – number of time series in the dataset.

T – length of the observed time series.

H – length of the forecast horizon.

To assess impact of different presets and time limits on the search a search for a time series forecasting models was performed with a planning horizon of 60 (2 months) while changing various presets and evaluation metrics. In the experiments the following metrics were used: MAE, RMSE, WQL.

$$MAE = \frac{1}{N}\frac{1}{H}\sum_{i=1}^{N}\sum_{t=T+1}^{T+H}|y_{i,t} - f_{i,t}|$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\frac{1}{H}\sum_{i=1}^{N}\sum_{t=T+1}^{T+H}(y_{i,t} - f_{i,t})^2}$$

$$\text{WQL} = \frac{1}{\sum_{i=1}^{N}\sum_{t=T+1}^{T+H}|y_{i,t}|}\sum_{i=1}^{N}\sum_{t=T+1}^{T+H}\sum_{q} p_q(y_{i,t}, f_{i,t}^{q})$$

where $f_{i,t}^{q}$ – predicted quantile q of time series i at time t; $p_q(y,f)$ – quantile loss at level q.

The search for models was performed within the following time limits: 5, 10, 15, and 30 minutes. The results of the search runs are summarized in Table 3.

TABLE III. MODELS OBTAINED WITH DIFFERENT CONSTRAINTS

Preset	Time(s)	MAE	RMSE	WQL
fast_training	300	5.9093	7,7755	0.0741
fast_training	600	5.9929	7,8625	0.0747
fast_training	900	6.0363	7,8827	0.0754
fast_training	1800	6.0363	7.8827	0.0753
medium_quality	300	5.3815	7,3873	0.0986
medium_quality	600	5.3626	7,3514	0.0782
medium_quality	900	**5.3540**	7,3326	**0.07**
medium_quality	1800	5.3829	7.3205	0.0703
high_quality	300	5.9197	7,4377	0.1153
high_quality	600	5.4345	7,3975	0.0967
high_quality	900	5.4202	7,3617	0.0837
high_quality	1800	5.4057	**7.3162**	0.07
best_quality	300	8.9054	11,601	0.1158
best_quality	600	5.8145	7,8126	0.1158
best_quality	900	5.4362	7,4078	0.1158
best_quality	1800	5.4096	7.3562	0.0842

An example of a forecast for Moscow generated by a model found in 15 minutes using the medium_quality preset and the MAE evaluation metric is shown in Fig. 2.

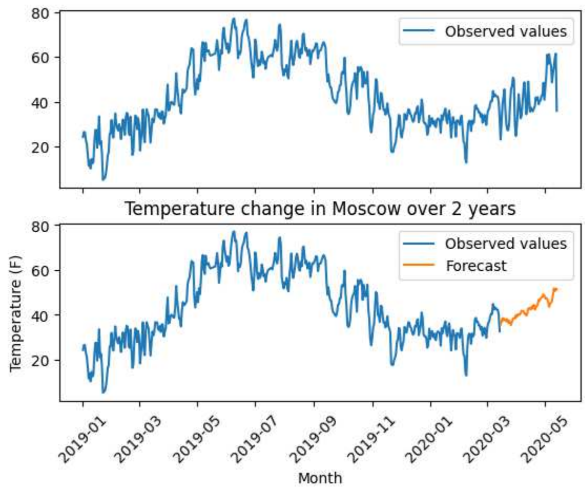

Fig. 2. Observed and predicted values of the time series

CONCLUSION

In this work we have shown the applicability of Automated Machine Learning to the problem of time series forecasting. For the experiments on automated model creation using AutoML tools, the AutoGluon system was used and the Daily Temperature of Major Cities dataset containing temperature data in 321 cities around the world.

In the experiments 3 metrics for models evaluation were used: MAE, RMSE, WQL and 4 time limits.

Authorized licensed use limited to: SRM Institute of Science and Technology- RamaPuram. Downloaded on September 02,2024 at 09:02:05 UTC from IEEE Xplore. Restrictions apply.

The experiments have shown that for the given dataset the most effective model evaluation metric is WQL, and that fast_training preset shows worse results as the time limit increases. With the given time limits of no more than 30 minutes for search, the preset that gives the most accurate model was medium_quality

REFERENCES

[1] Bishop, Christopher M. Pattern Recognition and Machine Learning. New York :Springer, 2006

[2] X. He, K. Zhao, X. Chu, “AutoML: A Survey of the State-of-the-Art”, Knowledge-Based Systems, vol. 212, Art. No. 106622, 27 p., 2021 doi: 10.1016/j.knosys.2020.106622.

[3] George E. P. Box, Gwilym M. Jenkins, Gregory C. Reinsel, Greta M. Ljung. Time Series Analysis: Forecasting and Control, 5th Edition. New York: Wiley, 2015

[4] N. Erickson, J. Mueller, A. Shirkov, H. Zhang, P. Larroy, M. Li, A. Smola, “AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data”. arXiv preprint arXiv:2003.06505 (2020) unpublished

[5] H. Jin, F. Chollet, Q. Song, X. Hu, “AutoKeras: An AutoML Library for Deep Learning”, Journal of Machine Learning Research, vol. 24. Art. No. 6., 6 p., 2023

[6] L Zimmer, M Lindauer, F. Hutter, “Auto-Pytorch: Multi-Fidelity MetaLearning for Efficient and Robust AutoDL”, IEEE Transactions On Pattern Analysis and Machine Intelligence, vol. 43, no. 9, pp. 3079-3090, 1 Sept. 2021, doi: 10.1109/TPAMI.2021.3067763

[7] AutoTS - time series package for Python, https://github.com/winedarksea/AutoTS (accessed 18 March 2024)

[8] I. Revin, V. A. Potemkin, N. R. Balabanov, N. O. Nikitin, “Automated machine learning approach for time series classification pipelines using evolutionary optimization”, Knowledge-Based Systems, vol. 268, 2023, ISSN 0950-7051

[9] Kaggle: Your Machine Learning and Data Science Community: http://www.kaggle.com (accesed 18 March 2024).

[10] OpenML. A worldwide machine learning lab. https://www.openml.org (accessed 18 March 2024).

[11] Time Series Classification Website. https://www.timeseriesclassification.com (accessed 18 March 2024).

Authorized licensed use limited to: SRM Institute of Science and Technology- RamaPuram. Downloaded on September 02,2024 at 09:02:05 UTC from IEEE Xplore. Restrictions apply.