Vol: 2, Issue: 2, 2024
Pages:86-95
Received: 1 August 2024
Accepted: 26 August 2024

CURRENT TRENDS IN COMPUTING
Research Article
DOI: 10.71074/CTC.1526313

APPLICATION OF AUTOMATED MACHINE LEARNING AND BAGGING TECHNIQUES TO CLASSIFY RICE VARIETIES

CIHAN BAYRAKTAR[1*]

[1] *Computer Technologies Department, Karabük University, 78050, Karabük, Türkiye*

ABSTRACT. **Nowadays, the research for digitalization in the agricultural sector recent years has significantly increased. In particular, machine learning and artificial intelligence have applications in agricultural product classification, quality control, and species identification. The fast processing times, high accuracy levels, and cost-effectiveness offered by digital solutions for quality control and classification accelerate these studies. Classifying rice species using traditional methods is a process that requires expertise, is time-consuming, and costly. Errors and differences of opinion due to human factors constitute essential limitations in this process. In order to eliminate these limitations, this study proposes a collaborative learning model utilizing Automated Machine Learning and Bagging techniques for rice species detection and classification. The model uses a dataset from the UCI Irvine Machine Learning Repository, which contains characteristics specific to the Osmancık and Cammeo rice varieties grown in Turkey. The dataset consists of 3810 data points, 2180 of which belong to Osmancık rice and 1630 to Cammeo rice. During the analysis, MLBox, an Automated Machine Learning library, was used to determine the optimal algorithm (Light Gradient Boosting Machine - LGBM) and its hyperparameters. Later, by applying the Bagging technique within the developed learning model, an accuracy rate of 93.54% was achieved in rice-type classification.**

1. INTRODUCTION

Agriculture is a sector of vital importance, especially for countries that are still developing. The reasons for this importance include ensuring sufficient and safe food supply, contributing to national income by providing job opportunities to large masses, encouraging the development of industry with the demand for agricultural inputs, developing exports, and contributing to the general development of the country. For these reasons, developing and supporting the agricultural sector should be among the priorities of every country [1]. In summary, a country's high agricultural productivity increases the economic welfare level of the country as a whole [2].

According to 2021 statistics, rice is one of the most important basic food products for the world population, producing more than 1 billion tons worldwide [3]. The criteria applied to detect quality rice vary according to regions and countries. However, among consumers, physical appearance, taste,

E-mail address: cihanbayraktar@karabuk.edu.tr (*).
Key words and phrases. Rice Types, Machine Learning, AutoML, Bagging, Classification.

aroma, smell, and cooking ability generally stand out as factors that are taken into consideration when determining the quality of rice [4].

In evaluating the quality of rice, the use of machine learning and artificial intelligence techniques instead of manual methods is increasing thanks to the low time and cost opportunities it provides. Manual methods at this stage result in long-term errors and high error rates due to human factors. In addition, the fact that evaluation can only be made by manual methods by experts of the relevant product brings another limitation. Differences in people's own opinions also cause differences in evaluation results. For this reason, the use of automatic systems instead of manual methods can enable more effective results in the quality evaluation of products [5].

Machine learning algorithms enable complex, high-dimensional data to be analyzed quickly and accurately. Fault detection, fraud detection, and product quality analysis can be given as examples. One of the main reasons for the widespread use of machine learning algorithms is that they enable the use of Graphics Processing Unit (GPU) on a large scale. Because GPUs can show much higher performance in data analysis operations than the Central Processing Unit (CPU) of computers. The development of these technologies in data analysis gives hope for solving evaluation problems in agricultural products [6]. Among the techniques used in machine learning applications, data analysis with Automated Machine Learning (AutoML) libraries can significantly impact different segments of agriculture and industry with the ease of hyperparameter optimization.

In the study conducted by Çınar et al., images of two rice species, Osmancık and Cammeo rice grains grown and registered in Turkey, were taken and processed, and a dataset was created by making feature extractions. The results were compared by applying machine learning algorithms and classification techniques to the produced dataset. As a result of the study, they stated that they got the best accuracy rate, 93.02%, from the model produced with the Logistic Regression algorithm [7].

In their previously published study on the dataset used in this study, İlhan et al. created a model with Deep Neural Networks for the classification of Osmancık and Cammeo rice varieties. As a result of the study, they reached an accuracy rate of 93.04% with the model prepared with Deep Neural Networks. It was stated that the model created in the study made successful classification [8].

In a study conducted by Jin et al., data analysis was carried out using deep learning algorithms such as LeNet, GoogleNet, and ResNet on the seeds imaged with hyperspectral imaging technology to classify rice seed varieties. As a result of the combination of hyperspectral imaging and deep learning algorithms, it has been determined that effective models can be produced in distinguishing rice seeds and the ResNet algorithm shows the best performance with 86.08% [9].

In their study, Jaithavil et al. created transfer learning models with VGG16, InceptionV3 and MobileNetV2 systems for the classification of paddy seeds and performed data analysis tests. It was announced that the proposed transfer learning model achieved high accuracy in classifying steel seeds, and the InceptionV3 model achieved the best result with an accuracy rate of 83.33%. Within the scope of the study, it was also stated that the MobileNetV2 model reached the same level of accuracy, but the classification performance of this model was not considered sufficient as the test loss occurred at 61.95% [10].

In the study conducted by Jumi et al. to classify rice types, the shape, color, and texture characteristics of rice were extracted using the Invariant Moment, Hue Saturation Value, and Local Dual Axis methods, and a dataset was created. Afterwards, the relevant dataset was analyzed with the k-nearest Neighbor classification algorithm, and an accuracy rate of 86.22% was obtained. It was stated that the data obtained within the scope of the study reached a promising result on the subject [11].

Hoang et al. investigated the difference between manual methods and the CNN algorithm that can be used to classify rice varieties. The VNRICE dataset was used in the study. After testing various CNN models, they found the best result of 99.04% with the learning model created by the DenseNet algorithm with 121 layers. In the study, it was stated that the 121-layer DenseNet model showed the highest performance, as the accuracy result, as well as the memory and resource consumption of the models, were taken into consideration [12].

In the study conducted by Mrutyunjaya and Harish Kumar, ensemble machine-learning algorithms were used to classify five different rice varieties with high accuracy. It has been stated that the learning model, which was created based on machine learning techniques and image processing methods, was successful in correctly classifying different rice varieties. In the study, the highest average classification accuracy among all tested algorithms was achieved by the Extreme Gradient Boosting (XGBoost) algorithm with 99.60% [13].

Köklü et al. carried out a classification study with deep learning algorithms for Arborio, Basmati, İpsala, Jasmine, and Karacadağ rice types. In this study, Artificial Neural Networks (ANN), Deep Neural Networks (DNN), and Conventional Neural Networks (CNN) algorithms were preferred for classification processes, and the results obtained as a result of the classification were compared. As a result of the analysis, ANN reached 99.87%, DNN 99.95%, and CNN 100% performance rates. The findings obtained in the study stated that the learning models used can be successfully applied in the classification of rice varieties and can help determine seed quality [5].

In this study, the difficulties encountered in the quality classification processes of agricultural products are discussed based on the classification of Osmancık and Cammeo rice varieties grown in Turkey. The slowness and high error rates of traditional methods require the use of more effective and accurate techniques in the quality control of agricultural products. In this context, the proposed solution is a learning model developed using AutoML and Bagging methods. This model aims to increase efficiency in agricultural production processes by classifying rice varieties more quickly and accurately. In the following sections of the study, detailed information is given about the dataset used, explanations are made regarding the production of the proposed models, and the findings obtained from the analysis of the data are conveyed.

2. Materials and Methods

In this study, in order to make a prediction and classify between Osmancık and Cammeo rice, the dataset was produced by Çınar and Köklü [7] and donated to the UC Irvine Machine Learning Pool, where analyses were carried out [14]. In the study conducted by [7], it was stated that the dataset used was created by transferring the images of 50g Osmancık and Cammeo rice to the computer environment and determining their properties, with a camera placed on a box that does not receive any external light but has an internal lighting mechanism. In this context, the dataset consists of 3810 lines of data in total. The attribute definitions determined for the dataset are given in Table 1.

TABLE 1. **Dataset attribute definitions**

Attribute Name	Attribute Definition
Area	Number of pixels within the boundaries of a grain of rice
Perimeter	The sum of the distances between pixels around the boundaries of the grain of rice
Major Axis Length	The longest line that can be drawn on a grain of rice
Minor Maxis Length	The shortest line that can be drawn on a grain of rice
Eccentricity	The degree of roundness of the ellipse with the same moments as a grain of rice
Convex Area	Number of pixels of the smallest convex hull of the area formed by the rice grain
Extend	The ratio of the area formed by the rice grain to the bounding box
Class	Result tag (Osmancık / Cammeo)

The class label distribution in the dataset was Osmancık 2180 and Cammeo 1630 (Figure 1).

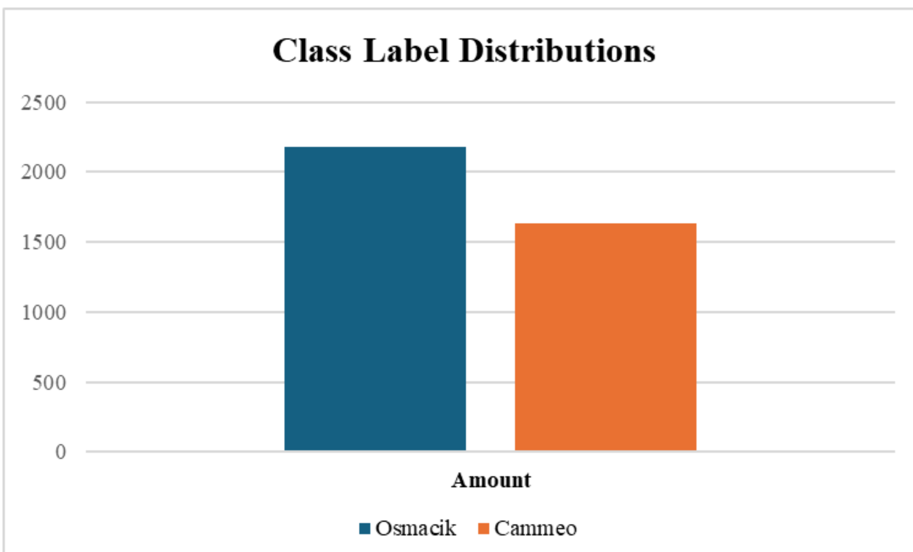

FIGURE 1. **Dataset class label distributions.**

In the analyses performed on the dataset, a computer with Intel Core i7 9750H CPU, 16 GB RAM Memory, Nvidia GeForce GTX 1050 3 GB Graphics Card, and Windows 11 operating system was used. Analyzes were carried out using the Python 3.7 programming language and the Jupyter Notebook editor.

```
> Number of common features : 7

gathering and crunching for train and test datasets ...
reindexing for train and test datasets ...
dropping training duplicates ...
dropping constant variables on training set ...

> Number of categorical features: 0
> Number of numerical features: 7
> Number of training samples : 2666
> Number of test samples : 1144

> You have no missing values on train set...

> Task : classification
Osmancik    1525
Cammeo      1141
Name: Class, dtype: int64
```

```
computing drifts ...
CPU time: 0.12216067314147949 seconds

> Top 10 drifts

('Minor_Axis_Length', 0.06610831029435671)
('Convex_Area', 0.02074032494137512)
('Eccentricity', 0.01564573835766625)
('Extent', 0.015288350059542655)
('Major_Axis_Length', 0.009507184488429798)
('Perimeter', 0.006318231655816131)
('Area', 0.0008354361317601811)

> Deleted variables : []
> Drift coefficients dumped into directory : save
```

FIGURE 2. **MLBox Data preprocessing results.**

2.1. Data Preprocessing:

Data preprocessing, also called data preparation, is the process of processing raw data and cleaning, modifying, and rearranging it before analysis. This step often requires formatting, adjustment, and integration to improve the information contained in the datasets. Data preprocessing is an important step in preparing data for processing and reducing the possibility of bias, but it can be a laborious task [15].

Preprocessing processes were run on the dataset analyzed in this study before analysis with MLBox, an AutoML library. In line with the results shown in Figure 2, no missing/erroneous data was found in the data, and no categorical data type was detected. In addition, in terms of the dataset order, it was seen that the attributes did not have a quality that would disrupt the order in reaching the result. Therefore, no attribute extraction was performed. With the current state of the dataset, the analysis process has begun for the creation and testing of learning models.

2.2. Data Analyses:

AutoML technology was used to determine the most suitable algorithms and hyperparameters before the bagging process to be used in the classification of Osmancık and Cammeo rice. AutoML is the preferred technique for performing complex analyses on large datasets. This technique significantly facilitates the analysis of large-scale data compared to traditional analysis methods [16]. AutoML aims to enable machine learning applications to produce better results by making it easier for data analysis experts to easily apply machine learning techniques, as well as to make appropriate hyperparameter adjustments for data analysis experts [17]. AutoML uses an automatic technique that allows machine learning algorithms to be configured at an optimal level. For this reason, the prevalence of its use among researchers continues to increase [18].

TABLE 2. **LGBM algorithm performance values**

Model	Precision	Sensitivity	Specifity	F1-Scrore	Accuracy
LGBM	0.912	0.925	0.928	0.918	0.927

AutoML technology covers the following processes in terms of data analysis [19]:

- Data Preprocessing: It ensures that the quality of the data is maintained by helping to perform various cleaning and preparation operations on the datasets before creating the learning models.
- Model Selection: AutoML allows the most appropriate model to be automatically selected according to the characteristics of the dataset and the classification method to be applied.
- Hyperparameter Optimization: AutoML technology automatically adjusts hyperparameter options that will optimize performance and accuracy without the need for manual intervention by the data analysis researcher.
- Binary, Multi-Class, and Multi-Label Classification: With AutoML tools, they can act in multiple ways in classification scenarios by creating effective solutions to such classification problems.

Many libraries implement AutoML techniques. Within the scope of this study, data analysis was carried out using the AutoML library named MLBox. MLBox is a library developed to perform distributed data processing, cleaning, and formatting processes. In order to provide these features, it supports state-of-the-art machine learning algorithms. In addition to individual algorithms, it can also work with ensemble learning algorithms such as LightGbm and XgBoost [20]. In addition, it can perform feature selection processes in an extremely robust manner and apply accurate hyperparameter optimizations in high-dimensional data structures [21]. MLBox performs data analysis with three basic sub-packages that work in a determined order. The first of these packages, preprocessing, ensures that the data is read and preprocessed if necessary. The second package, optimization, enables the application of appropriate hyperparameter optimizations and the testing processes of the created learning models. The third step, prediction, carries out the process of predicting the result using the obtained learning models and input data. The working order of MLBox occurs automatically since it is an AutoML library [22].

3. RESULTS

In our study, in order to achieve the best level of accuracy in classifying rice grains, the appropriate algorithm and hyperparameters were determined with the MLBox library, and then the detected algorithm was subjected to the Bagging process with hyperparameters, aiming to increase the performance rate incorrect predictions.

In the first stage, as a result of the analysis performed with the MLBox library, it was concluded that the Light Gradient Boosting Machine (LGBM) algorithm was suitable for the dataset with the hyperparameters given in Figure 3. The classification success of the LGBM algorithm on the dataset was determined as 92.70%, and it was observed that it achieved successful classification. Table 2 gives the performance values of the LGBM algorithm on the dataset.

```
LGBMClassifier(max_depth=3, boosting_type= 'gbdt', class_weight= None, colsample_bytree= 0.8,
               importance_type= 'split', learning_rate= 0.05, min_child_samples= 20, min_child_weight= 0.001,
               min_split_gain= 0.0, n_estimators= 500, n_jobs= -1, num_leaves= 31, objective= None,
               random_state= None, reg_alpha= 0.0, reg_lambda= 0.0, silent= True, subsample= 0.9,
               subsample for bin= 200000, subsample freq= 0, nthread= -1, seed= 0)
```

FIGURE 3. **LGBM classification algorithm hyperparameters.**

```
BaggingClassifier(LGBMClassifier(max_depth=3, boosting_type= 'gbdt', class_weight= None, colsample_bytree= 0.8,
                  importance_type= 'split', learning_rate= 0.05, min_child_samples= 20, min_child_weight= 0.001,
                  min_split_gain= 0.0, n_estimators= 500, n_jobs= -1, num_leaves= 31, objective= None,
                  random_state= None, reg_alpha= 0.0, reg_lambda= 0.0, silent= True, subsample= 0.9,
                  subsample_for_bin= 200000, subsample_freq= 0, nthread= -1, seed= 0),
                 max_samples=0.1, max_features=0.5, n_estimators=50)
```

FIGURE 4. **Bagging process hyperparameters.**

TABLE 3. **Performance of the learning model produced with bagging technique**

Tests	Precision	Sensitivity	Specifity	F1-Scrore	Accuracy
Test 1	0.905	0.951	0.925	0.927	0.936
Test 2	0.909	0.949	0.928	0.929	0.937
Test 3	0.916	0.938	0.932	0.927	0.935
Test 4	0.902	0.946	0.923	0.924	0.933
Test 5	0.907	0.938	0.926	0.922	0.931
Test 6	0.909	0.944	0.928	0.927	0.935
Test 7	0.909	0.951	0.928	0.930	0.938
Test 8	0.909	0.951	0.928	0.930	0.938
Test 9	0.902	0.944	0.923	0.923	0.932
Test 10	0.916	0.947	0.933	0.931	0.939
Mean	**0.909**	**0.946**	**0.927**	**0.927**	**0.9354**

In the second stage, the Bagging process was applied to the learning model using the LGBM algorithm and the determined hyperparameters, and the aim was to increase the classification performance rate. The bagging process was carried out using the hyperparameters shown in Figure 4, and the performance rate increased as targeted. As a result of the tests carried out using the learning model created with the bagging method, the correct classification success of the model reached an average level of 93.54% (Table 3). This result shows that the Bagging technique, one of the ensemble learning methods, has a positive effect on the model performance.

TABLE 4. **Performance values of models obtained from different studies**

Model	Precision	Sensitivity	Specifity	F1-Scrore	Accuracy
LR	0.915	0.923	0.937	0.918	0.9302
DNN	0.911	0.925	0.935	0.918	0.9304
Bagging	0.909	0.946	0.927	0.927	0.9354

Analyzes were carried out in different studies on the dataset we used in our study and the results were written. According to these studies, various classification algorithms were tested with the dataset by Çınar and Köklü [**?**] and it was stated that the best accuracy rate was determined as 93.02% with the Logistic Regression (LR) algorithm. In addition, in a study conducted by İlhan et al. [8], a learning model was created with Deep Neural Networks (DNN) and it was written that they reached an average accuracy rate of 93.04%. In our study, an average accuracy rate of 93.54% was achieved in the learning model created using the Bagging technique, which was carried out after the algorithm and Hyperparameter determination process with the MLBox library. Table 4 shows the comparison of performance values of the models obtained in the studies.

The structure of the dataset used in the study and the maintenance data specified in Table 4 stopped the initiation of the correct feature extraction process for the classification of the analyzed certificates. The high accuracy rates obtained in the analyses performed indicate these features. It is foreseen that analyses performed with different machine learning and deep learning techniques will also realize these developments in the near future. However, the proposed learning model structure can produce better results than preferring section structures instead of relying on a single model.

4. CONCLUSION

Within the scope of the study, the dataset containing the characteristics of Osmancık and Cammeo rice types, published as open source in the UCI Irvine Machine Learning Repository, was used. In the study, to perform analyses on the dataset, the MLBox library, one of the AutoML libraries, was used to determine the classification algorithm suitable for analysis and the hyperparameters that gave the most accurate results. At this stage, MLBox suggested the LGBM algorithm as a result, and an accuracy rate of 92.70% was achieved with the created learning model. In the next stage, the Bagging technique was applied to the LGBM algorithm to improve the learning model in order to achieve better results. An average accuracy level of 93.54% was obtained in the analyses made using the new learning model developed with the bagging technique.

When the performance values obtained in the study were examined, it was concluded that the learning model created was successful. In addition, a comparison was made with the performance values obtained in different studies on the same dataset (Table 4). As a result of this comparison, it was observed that all models had performance values close to each other and made successful classification. It can be said that the learning model obtained in our study gives slightly better results than the models in other studies.

In the future, in line with the results obtained from these studies, it is recommended to create automation structures for rapid classification of agricultural products and identification of their types. In this way, it will be possible to control the products produced in the agricultural sector faster and with minimum errors.

DECLARATIONS

- **Contribution Rate Statement:** Cihan BAYRAKTAR has conducted this study as a single author.
- **Conflict of Interest:** The author declares that they have no conflict of interest.
- **Data Availability:** Data are available at https://doi.org/10.24432/C5MW4Z

- **Statement of Support and Acknowledgment:** We would like to thank the researchers who carried out the necessary studies to prepare the dataset used in this study and made it available under the CC BY 4.0 license [7]. The dataset used in the study can be accessed at https://doi.org/10.24432/C5MW4Z Additionally, we thank the anonymous referees for their thoughtful comments and suggestions on the manuscript.

REFERENCES

[1] R. Alamyar, I. Boz, Marketing problems encountered by rice producers and their solutions: A case study of takhar-afghanistan, ISPEC Journal of Agricultural Sciences 5 (2) (2021) 381–392. doi:10.46291/ISPECJASVOL5ISS2PP381-392.

[2] E. Veziroglu, I. Pacal, A. Coskuncay, Derin evrişimli sinir ağları kullanılarak pirinç hastalıklarının sınıflandırılması, Journal of the Institute of Science and Technology 13 (2) (2023) 792–814. doi:10.21597/JIST.1265769. URL https://dergipark.org.tr/en/pub/jist/issue/77307/126576

[3] Food and Agriculture Organization of the United Nations, FAOSTAT dataset, accessed: 2021 (2021). URL https://www.fao.org/faostat/en/#data

[4] Z. C. Mutafcilar, Türkiye'de tescilli çeltik çeşitlerinin moleküler karakterizasyonu, Ph.d. thesis, Trakya University (2018).

[5] M. Koklu, I. Cinar, Y. S. Taspinar, Classification of rice varieties with deep learning methods, Computers and Electronics in Agriculture 187 (2021) 106285. doi:10.1016/j.compag.2021.106285.

[6] D. I. Patrício, R. Rieder, Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review, Computers and Electronics in Agriculture 153 (2018) 69–81. doi:10.1016/j.compag.2018.08.001.

[7] I. Cinar, M. Koklu, Classification of rice varieties using artificial intelligence methods, International Journal of Intelligent Systems and Applications in Engineering 7 (3) (2019) 188–194. doi:10.18201/IJISAE.2019355381.

[8] U. Ilhan, A. Ilhan, K. Uyar, E. I. Iseri, Classification of osmancik and cammeo rice varieties using deep neural networks, in: ISMSIT 2021 - 5th International Symposium on Multidisciplinary Studies and Innovative Technologies, Proceedings, Institute of Electrical and Electronics Engineers Inc., Ankara, Turkey, 2021, pp. 587–590. doi:10.1109/ISMSIT52890.2021.9604606.

[9] B. Jin, C. Zhang, L. Jia, Q. Tang, L. Gao, G. Zhao, H. Qi, Identification of rice seed varieties based on near-infrared hyperspectral imaging technology combined with deep learning, ACS Omega 7 (6) (2022) 4735–4749. doi:10.1021/acsomega.1c04102.

[10] D. Jaithavil, S. Triamlumlerd, M. Pracha, Paddy seed variety classification using transfer learning based on deep learning, in: Proceedings of the 2022 International Electrical Engineering Congress, iEECON 2022, Institute of Electrical and Electronics Engineers Inc., Khon Kaen, Thailand, 2022, pp. 1–4. doi:10.1109/IEECON53204.2022.9741677.

[11] J. Jumi, A. Zaenuddin, T. Mulyono, Identification of rice types based on shape, color and texture using k-nearest neighbors method as classifier, International Journal of Engineering Research Technology 9 (12) (2020). doi:10.17577/IJERTV9IS120013.

[12] V. T. Hoang, D. P. Van Hoai, T. Surinwarangkoon, H. T. Duong, K. Meethongjan, A comparative study of rice variety classification based on deep learning and hand-crafted features, ECTI Transactions on Computer and Information Technology (ECTI-CIT) 14 (1) (2020) 1–10. doi:10.37936/ECTI-CIT.2020141.204170.

[13] M. S. Mrutyunjaya, K. S. Harish Kumar, Non-destructive machine vision system based rice classification using ensemble machine learning algorithms, Recent Advances in Electrical Electronic Engineering (Formerly Recent Patents on Electrical Electronic Engineering) 16 (jul 2023). doi:10.2174/2352096516666230710144614.

[14] I. Cinar, M. Koklu, Rice (cammeo and osmancik) (2019). doi:10.24432/C5MW4Z.

[15] C. El Morr, M. Jammal, H. Ali-Hassan, W. El-Hallak, Data preprocessing, Springer International Publishing, Cham, 2022, pp. 117–163. doi:10.1007/978-3-031-16990-8_4.

[16] V. Kovalevsky, E. Stankova, N. Zhukova, O. Ogiy, A. Tristanov, Automl framework for labor potential modeling, in: Advances in Intelligent Systems and Computing, Vol. 13957, Springer, Cham, 2023, pp. 87–98. doi:10.1007/978-3-031-36808-0_6.

[17] Y. Sun, Q. Song, X. Gui, F. Ma, T. Wang, Automl in the wild: Obstacles, workarounds, and expectations, in: Conference on Human Factors in Computing Systems - Proceedings, Association for Computing Machinery, Hamburg, Germany, 2023, pp. 1–15. doi:10.1145/3544548.3581082.

[18] C. Wang, Z. Chen, M. Zhou, Automl from software engineering perspective: Landscapes and challenges, in: Proceedings - 2023 IEEE/ACM 20th International Conference on Mining Software Repositories, MSR 2023, IEEE Inc., Melbourne, Australia, 2023, pp. 39–51. doi:10.1109/MSR59073.2023.00019.

[19] M. Vinícius, C. Aragão, A. Guimarães Afonso, R. C. Ferraz, R. Gonçalves Ferreira, S. Gomes Leite, R. G. Ferreira, A practical evaluation of automl tools for binary, multiclass, and multilabel classification, TechRxiv (oct 2023). doi:10.36227/TECHRXIV.21792959.V1.

[20] S. Das, U. M. Cakmak, Hands-On Automated Machine Learning: A Beginner's Guide to Building Automated Machine Learning Systems Using AutoML and Python, Packt Publishing, Birmingham, UK, 2018.

[21] A. Aronio De Romblay, N. Cherel, M. Maskani, H. Gerard, Mlbox (2017).

[22] S. Ozdemir, S. Orslu, Makine ögrenmesinde yeni bir bakış açısı: Otomatik makine Ögrenmesi (automl), Journal of Information Systems and Management Research 1 (1) (2019) 23–30.