Uludağ University Journal of The Faculty of Engineering, Vol. 28, No. 2, 2023 RESEARCH

DOI: 10.17482/uumfd.1296479

INVESTIGATING THE EFFECT OF FEATURE SELECTION METHODS ON THE SUCCESS OF OVERALL EQUIPMENT EFFECTIVENESS PREDICTION

Ümit YILMAZ [*,**]
Özlem KUVAT [***]

Received: 12.05.2023; revised: 22.06.2023; accepted: 13.07.2023

Abstract: Overall equipment effectiveness (OEE) describes production efficiency by combining availability, performance, and quality and is used to evaluate production equipment's performance. This research's aim is to investigate the potential of the feature selection techniques and the multiple linear regression method, which is one of the machine learning techniques, in successfully predicting the OEE of the corrugated department of a box factory. In the study, six different planned downtimes and information on seventeen different previously known concepts related to activities to be performed are used as input features. Moreover, backward elimination, forward selection, stepwise selection, correlation-based feature selection (CFS), genetic algorithm, random forest, extra trees, ridge regression, lasso regression, and elastic net feature selection methods are proposed to find the most distinctive feature subset in the dataset. As a result of the analyses performed on the data set consisting of 23 features, 1 output and 1204 working days of information, the elastic net - multiple linear regression model, which selects 19 attributes, gave the best average R^2 value compared to other models developed. Occam's razor principle is taken into account since there is not a great difference between the average R^2 values obtained. Among the models developed according to the principle, the stepwise selection - multiple linear regression model yielded the best R^2 value among those that selected the fewest features.

Keywords: Feature selection, machine learning, overall equipment effectiveness

Öznitelik Seçim Yöntemlerinin Toplam Ekipman Etkinliği Tahmin Başarısı Üzerindeki Etkisinin Araştırılması

Öz: Toplam ekipman etkinliği (TEE); kullanılabilirliği, performansı ve kaliteyi birleştirerek üretim etkinliğini tanımlamaktadır ve üretim ekipmanının performansını değerlendirmek için kullanılmaktadır. Bu araştırmanın amacı, bir kutu fabrikasının oluklu mukavva departmanının TEE'sinin başarılı bir şekilde tahmin etmede, öznitelik seçim tekniklerinin ve makine öğrenmesi tekniklerinden biri olan çoklu doğrusal regresyon yönteminin potansiyelini araştırmaktır. Çalışmada girdi öznitelikleri olarak altı farklı planlı duruş süresi ve onyedi farklı gerçekleşecek faaliyetlere ilişkin önceden bilinen kavramlara ilişkin bilgiler kullanılmıştır. Ayrıca veri kümesinde en ayırt edici özellik alt kümesini bulmak için geriye doğru eleme, ileri doğru seçim, adımsal seçim, korelasyon tabanlı öznitelik seçim, genetik algoritma, rastgele orman, ekstra ağaç, ridge regresyon, lasso regresyon ve elastik net öznitelik seçim yöntemlerinden faydalanılmıştır. 23 öznitelikten, 1 çıktıdan ve 1204 iş günlük bilgiden oluşan veri seti üzerinde yapılan analizler neticesinde 19 adet öznitelik seçen elastik net – çoklu doğrusal regresyon modeli, geliştirilen diğer modellere kıyasla en iyi ortalama R^2 değerini vermiştir. Elde edilen ortalama R^2 değerleri arasında çok büyük bir fark

[*] Balıkesir University, Bigadiç Vocational School, Department of Management and Organization, 10440, Balıkesir
[**] Balıkesir University, Institute of Social Sciences, Department of Business Administration Doctorate Program, 10145, Balıkesir

[***] Balıkesir University, Faculty of Economics and Administrative Sciences, Department of Business, 10145, Balıkesir
Coresponding Author: Ümit Yılmaz (umityilmaz@balikesir.edu.tr)

olmaması dolayısıyla Occam'ın usturası ilkesi dikkate alınmıştır. İlkeye göre geliştirilen modellerden en az öznitelik seçenler arasında en iyi R^2 değerini stepwise selection - çoklu doğrusal regresyon modeli vermiştir.

Anahtar Kelimeler: Öznitelik seçimi, makine öğrenmesi, toplam ekipman etkinliği

1. INTRODUCTION

The OEE, expressed as a percentage, measures the performance and effectiveness of production operations and indicates how effective a company's production operations are (Chikwendu et al., 2020).

Businesses use performance evaluation systems to determine focus points to increase their performance and efficiency. OEE is a widely used metric in production as it describes production efficiency by combining availability, performance, and quality. Producers can measure the performance, availability, and quality of their machines and production lines using OEE metrics to identify which processes work most efficiently and which need improvement. This evaluation can also be conducted across the entire facility to compare OEE at a broader level (Ercan, 2020). OEE is a crucial measure of equipment efficiency, and careful analysis can determine its impact (Prasetyo and Veroya, 2020).

The increasing availability of data from businesses' equipment and operations is enabling the use of data analytics. With the growth of technology and information science, data has increased in significant quantities and dimensions. This huge amount of data contains a lot of useful information for humanity and allows for uncovering the past and looking to the future (Qu et al., 2023). Large datasets cause these sets to have noisy, irrelevant, and redundant features. This negatively affects the performance of regression methods in machine learning. In the training process of machine learning models, the quality output can be obtained from quality input. Therefore, the quality of input should be improved, and only relevant information should be given as input to the training and test process. In short, the high dimensionality leads to challenges of high computational cost, difficulties in model interpretation, and hampering prediction models' generalization abilities. For this reason, feature selection is required before the model is trained (Wang et al., 2023).

Machine learning aims to enhance the learning capabilities of machines, allowing them to uncover patterns within real-world data (Aydın, 2022). Feature selection, a fundamental topic of machine learning, is used extensively in prediction applications. The selection of contributing and informative features is essential to improve the working efficiency of machine learning techniques (Pasha and Mohamed, 2022). With the help of feature selection, redundant, irrelevant, and noisy features are removed from all features to identify an efficient subset of features (Wu et al., 2023). Feature selection aims to select the most suitable set of features to obtain optimal predictions regardless of the number of features used (Akman et al., 2023).

The aim of this study is to investigate the effect of various combinations of feature selection methods and multiple linear regression methods on OEE and to determine which feature selection method performs best in predicting OEE. None of the studies so far have systematically examined the impact of feature selection on the prediction of OEE.

This study consists of five chapters. This section provides an overwiev of OEE and feature selection. In the second section, a literatüre review of efficiency measures, In the second section, a literature review on the prediction of efficiency measures of feature selection methods is presented, especially the studies on the prediction of OEE by machine learning based feature selection methods. The third section presents the methodology to achieve the purpose of the study. The findings obtained in the study are described in the fourth section. Finally, the fifth section presents general conclusions and recommendations.

2. LITERATURE REVIEW

Feature selection is the process of selecting a subset of features by eliminating irrelevant or redundant features. Feature selection aims to reduce the difficulty of data availability and improve data quality. In the previous studies, feature selection methods are used before predictions of body fat (Lai et al., 2022), building energy performance (Olu-Ajayi et al., 2023), coal consumption (Zhou and Zhang, 2023), coffee yield (Barbosa et al., 2021), day-ahead electricity price (Li and Becker, 2021), earth skin temperature (Jamei et al., 2022), energy consumption (Moldovan and Slowik, 2021; Qiao et al., 2022; Sun et al., 2021; Zhao et al., 2023), mining blast vibration (Xu and Wang, 2023), ship motion (Wei et al., 2023), soybean yield (Corrales et al., 2022), stock index futures price (Yan, 2023), stock price and trend (Chaudhari and Thakkar, 2023), toxicity of chemical gases (Erturan et al., 2023), water quality (Kushwaha et al., 2023; Liu et al., 2023), etc.

This study's aim is OEE prediction using feature selection based machine learning techniques. A review of current literature on the aim identified is summarised as follows. Vogel-Heuser et al. (2017) utilized data-driven techniques to measure and elaborate on the identified causes and effects in their research. By employing logistic regression and decision-tree analysis, they investigated the connection between quality and process parameters.

Eroğlu (2019) several important features in a production environment that affect Overall Throughput Effectiveness (OTE) values, applying hybrid genetic algorithms and multiple linear regression methods on collected data.

Piran et al. (2020) found each feature's relevance and impact on the model's OEE with a stepwise selection method in their study, which aims to develop a method of evaluating the performance of operations in production systems.

Genç and Vupa Çilengiroğlu (2021) used logistic regression and Cart and Chaid decision tree algorithms from machine learning algorithms to find the variables that are thought to be effective on this score by categorizing OEE scores in a classifying and ranking manner.

This study added value to the strand of literature on the selection of features for the OEE prediction problem at the corrugated box area.

3. METHOD

Research and publication ethics were complied with in this study. There is no need for any ethical approval.

In this study, it is focused on the use of the multiple linear regression method, one of the machine learning techniques, with ten different feature selection methods to obtain an OEE prediction model with high prediction performance. The flowchart for the model developed for the prediction of OEE is shown in Figure 1 (Emanet et al., 2021).

3.1. Data Loading

In the research, the data obtained from the corrugated cardboard department of an enterprise was used. The data subject to the prediction covers a period of 4 years (1204 working days) between 2017 and 2020. The dataset consists of 23 features and one output. The features consist of planned stoppages and previously known data on the activities that will take place. The output is the percentage of OEE. The dataset obtained from the relevant department was taken directly from the ERP system. There are no missing or incorrect data in the dataset, and no inconsistencies were found in the dataset.

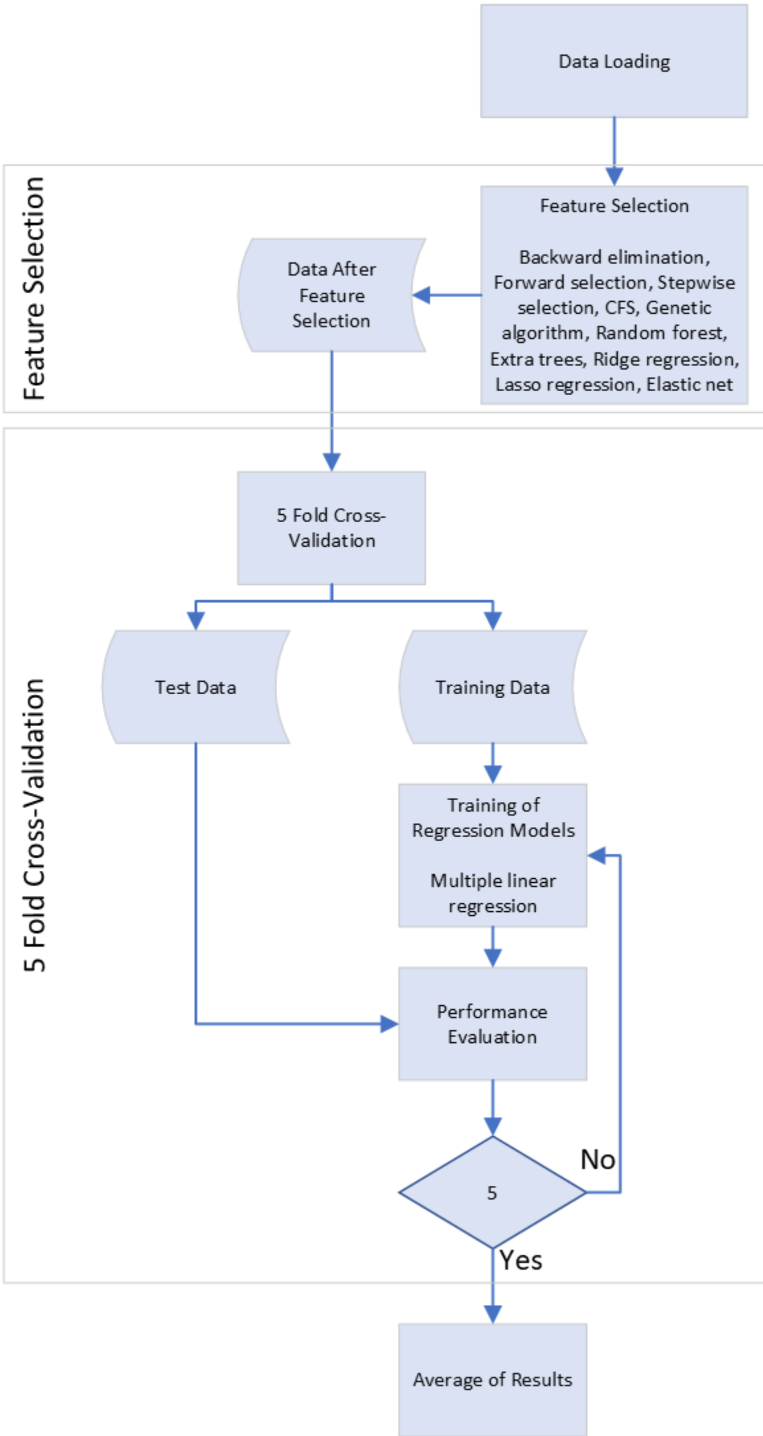

Figure 1:
Flowchart of the proposed model (Emanet et al., 2021)

3.2. Feature Selection Methods

In data analysis, various feature selection methods may be applied, each having its strengths and weaknesses. Stability is one of the most important features that characterize a feature selector, where they face irregularity in training data by sampling or deleting samples (da Costa et al., 2022).

Feature selection methods fall into three categories: filters, embedded methods, and wrappers. The filters rank the importance of features based on dependency, distance measure, or feature score and then select the higher-ranked features as the subset of features that are closest to the best. Filters are independent of learning algorithms and work fast. Embedded methods determine the subset of features that best contributes to the performance of the statistical learning model while building the model itself. Wrappers contain a statistical learner, such as a regressor, and search the feature subset space using the performance of the statistical learner (Tsanas, 2022; Yao et al., 2022).

In this study, backward elimination, forward selection, stepwise selection, correlation-based feature selection (CFS), genetic algorithm, random forest, extra trees, ridge regression, lasso regression, and elastic net methods were used as feature selection methods.

Backward Elimination begins with a full model containing all candidate features. Features are systematically removed from the model until a preset stop rule is satisfied. At a given stage of the elimination process, the variable causing the least decrease in a summary measure will be eliminated. Two possible summary measures can be utilized are deviation and coefficient of determination (R^2). The most frequently used rule for stopping is when all the remaining features in the model have a significant effect at a pre-determined level of significance.

At the start of forward selection, the model doesn't include any features. The model is updated step by step by adding features until a pre-established stopping rule is reached. During a specific stage of the selection process, the feature that produces the greatest increase in the summary measure is included in the model. If any added feature does not meet a predetermined level of significance, then no further variables are added to the model (Austin and Tu, 2004).

Stepwise selection is a method that combines forward selection and backward elimination. In essence, the process proceeds in a forward direction; however, at every stage of the feature selection process, once a feature has been included in the model, it may undergo a backward elimination process. Backward elimination is performed if any of the already selected features are found to have become redundant (Petersson et al., 2009).

The evaluation of feature subsets' value is performed by the CFS algorithm using a correlation-based metric. Good feature subsets, according to this method, consist of features that exhibit a strong correlation with the class, while not displaying a significant correlation with one another. The technique gives greater importance to subsets that have a strong relationship with the target variable but weak relationships among themselves. Thus, subsets that include redundant and irrelevant features are regarded as undesirable feature subsets (Liu and Schumann, 2005).

A population of solution structures, referred to as chromosomes, is generated by the genetic algorithm. Iteratively producing offspring with a greater level of fitness is achieved by choosing parents based on their mutation, crossover, and adaptation. To guide the selection of features, this method utilizes the performance of the estimator, which is used as the objective function (Sun et al., 2023).

The algorithm known as random forest evaluates the significance of each feature to decrease the quantity of features. The random forest is a method of creating an ensemble of decision trees. It does so by randomly selecting observations and variables, and then combining them to generate a variety of decision trees. Then, the features' importance score and the variables' predicted probability are calculated by summing up the votes generated by each decision tree. Features with high importance scores are typically selected, while others are ignored (Naseri et al., 2022).

The extra tree algorithm involves building a collection of trees on the dataset in an iterative manner, with each tree representing a specific subset of the data. In the trees, the features of a dataset are symbolized by the nodes, while the instances of a specific class are depicted by the leaves. According to the principle of information gain, the process of creating trees involves dividing nodes, and the features that have the highest relevance to the target are divided first. Thus, after the trees are generated, by sorting the nodes from the root to the leaves (excluding the leaf nodes) produced a sequence of features, which are ordered based on their relevance to the target. Once the features are ordered, the proportion of samples that reach node in the whole dataset is weighted with the purity of node. Thus, the importance of the feature is calculated (Xing et al., 2019).

Ridge regression is an improved least squares estimation technique that incorporates the least squares method's bias and a portion of the information to enhance the consistency and reliability of the regression coefficients' estimation, aligning it more closely with the real-world scenario (Zhao et al., 2022). Ridge regression is a method for selecting features based on statistical analysis. It examines multivariate cause-and-effect relationships within data (Toğaçar et al., 2020a, 2020b).

In Lasso regression, the best features are determined using the weight coefficient's absolute value as a penalty term in the cost function. The selection of the best features is made by giving a weight of zero to irrelevant input features and a weight of non-zero to relevant ones (Fathima et al., 2022).

Elastic net is a type of linear regression model that merges the strengths of both lasso and ridge regression to solve problems of high-dimensional feature selection. The elastic net is a middle ground between the penalties of lasso and ridge regression. These two penalties are combined linearly to form the elastic net penalty. The Lasso penalty promotes sparse of parameter predictions in the model, while the Ridge term averages the parameter prediction of the associated features applying a grouping effect. Therefore, elastic net performs both shrinkage and automatic feature selection. The weight given to lasso or ridge penalties by the Elastic net penalty can be adjusted smoothly based on the underlying problem's features and the preferences of the users. The penalty takes a value between 0 and 1. If the penalty is equal to zero, it functions as an elastic net ridge, and if it is equal to one, it functions as a lasso (Amini and Hu, 2021; Bai et al., 2023; Marafino et al., 2015).

3.3. Validation Method

Cross-validation is a partitioning strategy used in machine learning models. Cross-validation avoids the overfitting problem and improves prediction based on bias and/or variance (Almaghrabi et al., 2021). With a limited dataset, cross-validation is utilized to assess the efficacy of a network. The objective is to assess how effectively the model will perform in real-world scenarios. Cross-validation allows for predicting how well a model will perform on data that it hasn’t seen before, by using data that was not included in the training process. Cross-validation is commonly performed using the K-fold cross-validation method (Gunes et al., 2022).

K-fold cross-validation is a technique that splits the dataset into two subsets randomly as training and test datasets. The process of cross-validation involves dividing the dataset into K groups. The test set is comprised of one of these groups, while the training set is made up of the remaining K-1 groups. The K-fold cross-validation approach involves repeating the procedure K times, where the training and test datasets are replaced each time. The best model is selected based on the minimum error calculated using various statistical tools for error prediction (Saud et al., 2020). To assess the prediction error of a regression model in a traditional K-fold cross-validation method, the prediction errors of all folds in the cross-validation iterations are averaged (Wang et al., 2022).

3.4. Training Method

The study used the multiple linear regression method to train the regression model. This technique, known as linear regression, identifies the linear correlation between single or multiple predictor variables and the target variable. This approach assesses the impact of each independent variable on the dependent variable (Lap et al., 2023).

3.5. Performance Evaluation Metrics

There are several performance evaluation metrics employed to assess the predictive performance of the created models. During the testing phase, the model's performance was evaluated using performance metrics such as Mean Squared Error (MSE) and Coefficient of Determination (R^2). MSE is expressed as the mean of the squares of the differences between predicted and actual values (Fletcher et al., 2022). An MSE value close to zero indicates good performance of the model (Adak and Duralioğlu, 2023).

The R^2 value is a goodness-of-fit measure that measures the ratio of variance explained by the independent variable(s) for the dependent variable in a regression model with values in the range [0,1], indicating how well the model predicts unseen data samples. Where 0 indicates that there is no observed variance and 1 indicates that the variance in the dependent variable is 100% due to the movement of the independent variable(s) (Verma et al., 2022).

4. FINDINGS and COMMENTS

The dataset obtained from the enterprise was loaded into the software and feature selection was performed using backward elimination, forward selection, stepwise selection, CFS, genetic algorithm, random forest, extra trees, ridge regression, lasso regression, and elastic net feature selection methods. Five-fold cross-validation was applied to reveal the performance metrics of the models used in the study. Moreover, the problem was modeled with a multiple linear regression machine learning method. After feature selection, the reduced dataset was divided into 5 parts with the help of a 5-fold cross-validation function, training was performed on 4 parts, and testing was performed on 1 part. This was repeated for all combinations of training-test parts. The performances of the models on the test data on each fold were evaluated by MSE and R^2 metrics. With the help of the test set with the best MSE and the best R^2 values, a regression graph and a comparison graph were obtained using the predictions (outputs) and actual (target) values on this set. In addition, the averages of MSE and R^2 values obtained for all layers as a performance measure were also obtained as a result of the analysis. All analyses were performed using MATLAB R2022b.

As a result of the analyses performed, the features selected for each method and the average and best performance metrics values of the models developed with these features are summarized in Table 1.

Table 1. Features and performance metrics values selected according to the method

Methods	Selected Features	Number of Features	Best R^2	Best MSE	Average R^2	Average MSE
Ridge regression	2, 3, 4, 16, 17, 18, 19, 22	8	0.60813	0.0057910	0.60635	0.0058165
Stepwise selection	2, 4, 11, 18, 19, 20, 21, 22	8	0.63703	0.0055122	0.61927	0.0057845
Forward selection	2, 4, 10, 15, 17, 18, 19, 20, 21, 22, 23	11	0.60104	0.0053484	0.60076	0.0053765
Backward elimination	2, 4, 7, 9, 10, 16, 17, 18, 19, 20, 21, 22	12	0.61640	0.0051594	0.58927	0.0056655
Genetic algorithm	1, 4, 5, 10, 11, 13, 14, 15, 16, 18, 19, 20, 21, 22	14	0.65252	0.0048157	0.62059	0.0061646
Lasso regression	1, 2, 3, 4, 8, 10, 13, 15, 17, 18, 19, 20, 21, 22, 23	15	0.64292	0.0050668	0.62768	0.0058081
CFS	1, 2, 3, 4, 5, 6, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20	17	0.62557	0.0058483	0.61344	0.0060317
Extra trees	1, 2, 3, 4, 5, 6, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22	18	0.64054	0.0052140	0.62787	0.0057962
Elastic net	1, 2, 3, 4, 7, 8, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23	19	0.638555	0.005037	0.628593	0.005818
Random forest	1, 2, 3, 4, 5, 6, 9, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23	20	0.63028	0.0050817	0.62061	0.0058052

The regression plots between the output (predicted) OEE values and the target (actual) OEE values obtained as a result of the test processes of the models created by the feature selection methods are shown in Figure 2.

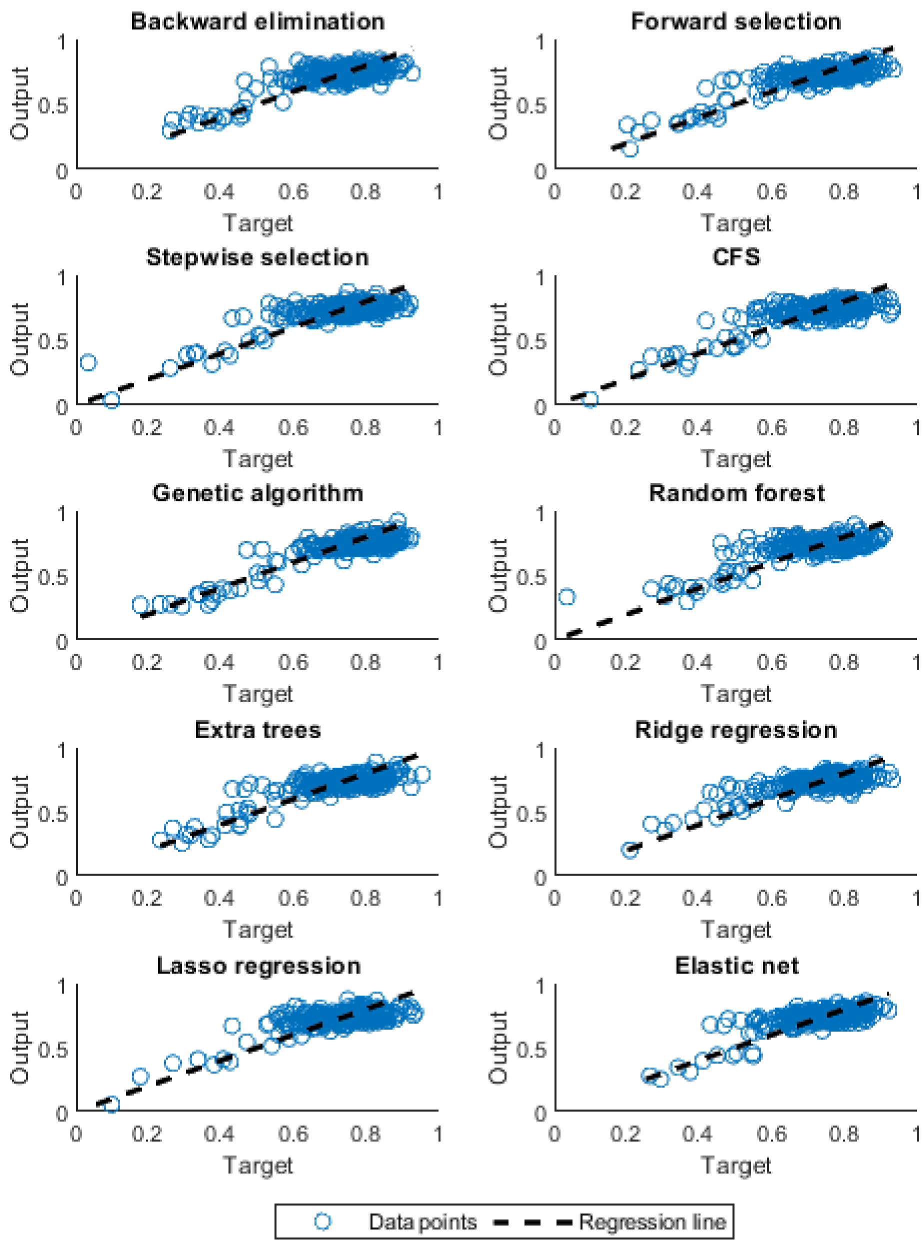

Figure 2:
Regression situations between the output values obtained after the test process and the target values

The course of the output (predicted) OEE values and the target (actual) OEE values obtained as a result of the test processes of the models created by the feature selection methods are shown in Figure 3.

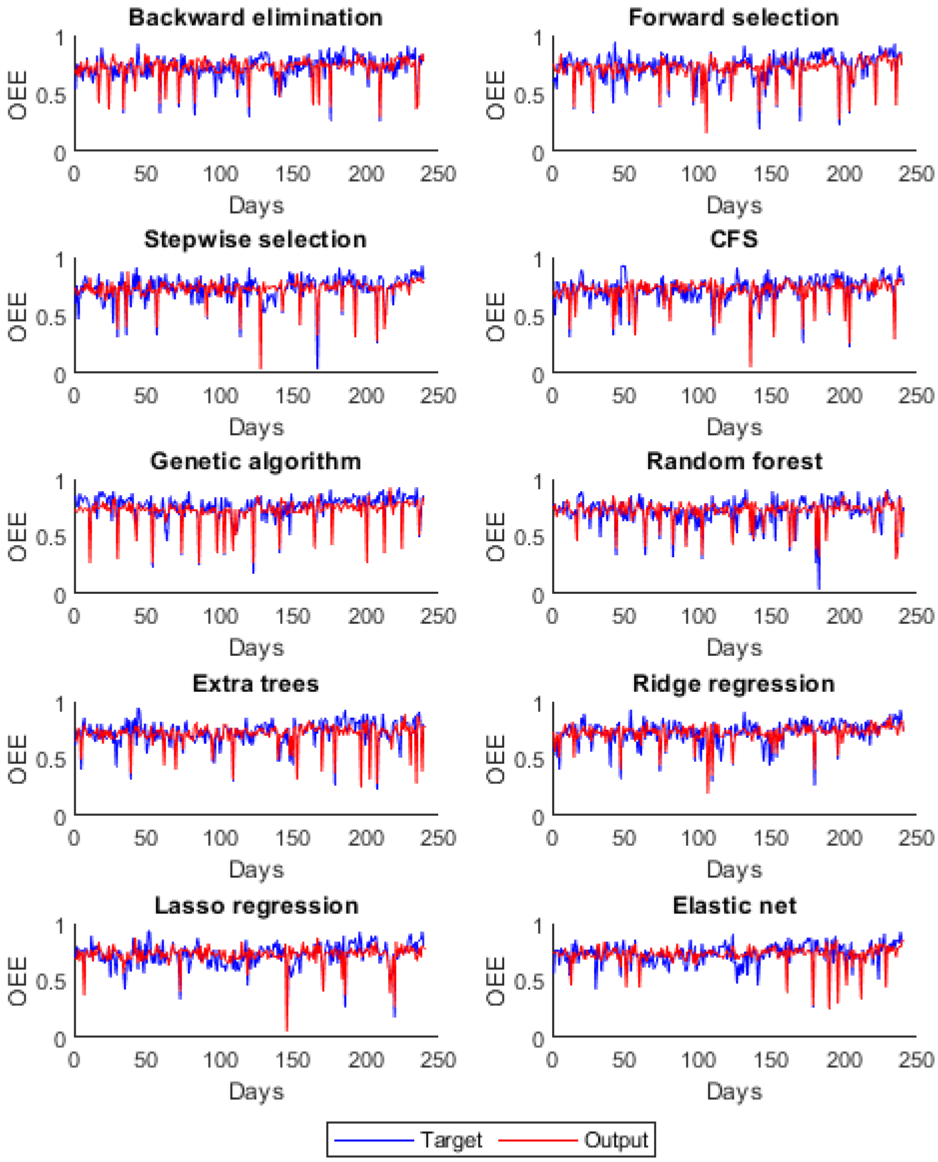

Figure 3:
Comparison of the output values obtained after the test process with the target values

Models were created as combinations of each of the machine learning methods used for feature selection and the multiple linear regression method. The models are analyzed for the prediction of OEE. The average MSE and average R^2 values obtained as a result of the analysis of the models were compared with each other. The statistical results shown in Table 1, Figure 2 and Figure 3 show that the forward selection - multiple regression model based on MSE and the elastic net - multiple regression model based on R^2 give the best values compared to the other models developed. Chicco et al. (2021) concluded in their study that the R^2 coefficient of determination was more informative than symmetric mean absolute percentage error (SMAPE), mean average error (MAE), mean absolute percentage error (MAPE), MSE and Root mean square error (RMSE) in regression analysis evaluations. Based on this finding, the elastic net-multiple regression model based on R^2 gives the best prediction performance compared to other models developed. This result emphasizes the necessity and contribution of appropriate feature selection in the OEE prediction process.

5. CONCLUSIONS AND RECOMMENDATIONS

Feature selection plays a crucial role in the data preprocessing phase of machine learning and data mining (Qu et al., 2023). Effectively improving prediction performance can be achieved in the dataset by identifying the most informative features and eliminating irrelevant or redundant ones. (Chamlal et al., 2022).

In this study, the objective is to examine the impact of different combinations of feature selection and regression techniques, and identify the most effective combination for predicting OEE. One of the primary achievements of the study is the introduction of a novel approach that employs machine learning algorithms in conjunction with feature selection techniques to predict the OEE value. Based on the experiments conducted on an actual dataset, the suggested approaches can generate a prediction of the OEE value with a maximum similarity of 0.65 and an average similarity of 0.63. Since the most important criterion is the average prediction performance, it can be concluded that the elastic net method gives the best result. However, the elastic net - multiple linear regression method pair has the disadvantage that more features are selected in the dataset than the other method pairs except one. In other words, increasing the number of features will make the model more complex, increasing the computation time and decreasing the generalization capability. This is where Occam's Razor principle can come into play. According to this principle, the simplest of the model alternatives with similar performance should be preferred (Korkmaz and Eroğlu, 2020). Since there is not a large gap between the average R^2 values obtained according to the methods, the stepwise selection method with the highest average R^2 value will be selected among the stepwise selection and ridge regression methods, which allow the selection of the least number of features (8 different features).

In future feature selection studies within the scope of this study, other machine learning methods can be used to improve prediction accuracy and reduce computation time.

CONFLICT OF INTEREST

The authors confirm that there is no known conflict of interest or common interest with any institution/organization or person.

AUTHOR CONTRIBUTION

Ümit YILMAZ contributed to the identification of the conceptual and design processes of the study, management of the conceptual and design processes of the study, data collection, data analysis and interpretation, drafting of the manuscript, and final approval and full responsibility. Özlem KUVAT contributed to the identification of the conceptual and design processes of the study, management of the conceptual and design processes of the study, data analysis and interpretation, drafting of the manuscript, critical review of the intellectual content, and final approval and full responsibility.

NOTE

This study is derived from Ümit YILMAZ's Ph.D. dissertation prepared at Balıkesir University, Institute of Social Sciences, Department of Business Administration, under the supervision of Özlem KUVAT.

REFERENCES

1. Adak, M. F. and Duralioğlu, Ö. (2023) Makine Öğrenmesi Yöntemleri Kullanılarak Öğrencilerin Kazanım Bilgileri ile Sınavlardaki Başarı Durumunun Tahmini, *Journal of Intelligent Systems: Theory and Applications*, 6(1), 43-51. doi:10.38016/jista.1183353
2. Akman, D. V., Malekipirbazari, M., Yenice, Z. D., Yeo, A., Adhikari, N., Wong, Y. K., Abbasi, B. and Gumus, A. T. (2023) k-best feature selection and ranking via stochastic approximation, *Expert Systems with Applications*, 213, 118864. doi:10.1016/j.eswa.2022.118864
3. Almaghrabi, F., Xu, D.-L. and Yang, J.-B. (2021) An evidential reasoning rule based feature selection for improving trauma outcome prediction, *Applied Soft Computing*, 103, 107112. doi:10.1016/j.asoc.2021.107112
4. Amini, F. and Hu, G. (2021) A two-layer feature selection method using Genetic Algorithm and Elastic Net, *Expert Systems with Applications*, 166, 114072. doi:10.1016/j.eswa.2020.114072
5. Austin, P. C. and Tu, J. V. (2004) Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality, *Journal of Clinical Epidemiology*, 57(11), 1138-1146. doi:10.1016/j.jclinepi.2004.04.003
6. Aydın, F. (2022) A New Instance Selection Method for Enlarging Margins Between Classes, *Journal of Intelligent Systems: Theory and Applications*, 5(2), 119-126. doi:10.38016/jista.1033354
7. Bai, H., Liu, P., Fu, X., Qiao, L., Liu, C., Xin, Y. and Ling, Z. (2023) Application of elastic net in quantitative analysis of major elements using Martian laser-induced breakdown spectroscopy datasets, *Spectrochimica Acta Part B: Atomic Spectroscopy*, 199, 106587. doi:10.1016/j.sab.2022.106587
8. Barbosa, B. D. S., Ferraz, G. A. e. S., Costa, L., Ampatzidis, Y., Vijayakumar, V. and dos Santos, L. M. (2021) UAV-based coffee yield prediction utilizing feature selection and deep learning, *Smart Agricultural Technology*, 1, 100010. doi:10.1016/j.atech.2021. 100010
9. Chamlal, H., Ouaderhman, T. and Aaboub, F. (2022) A graph based preordonnances theoretic supervised feature selection in high dimensional data, *Knowledge-Based Systems*, 257, 109899. doi:10.1016/j.knosys.2022.109899
10. Chaudhari, K. and Thakkar, A. (2023) Neural network systems with an integrated coefficient of variation-based feature selection for stock price and trend prediction, *Expert Systems with Applications*, 219, 119527. doi:10.1016/j.eswa.2023.119527
11. Chicco, D., Warrens, M. J. and Jurman, G. (2021) The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, *PeerJ Computer Science*, 7, e623. doi:10.7717/peerj-cs.623
12. Chikwendu, O. C., Chima, A. S. and Edith, M. C. (2020) The optimization of overall equipment effectiveness factors in a pharmaceutical company, *Heliyon*, 6(4), e03796. doi:10.1016/j.heliyon.2020.e03796
13. Corrales, D. C., Schoving, C., Raynal, H., Debaeke, P., Journet, E.-P. and Constantin, J. (2022) A surrogate model based on feature selection techniques and regression learners to improve soybean yield prediction in southern France, *Computers and Electronics in Agriculture*, 192, 106578. doi:10.1016/j.compag.2021.106578

14. da Costa, N. L., de Lima, M. D. and Barbosa, R. (2022) Analysis and improvements on feature selection methods based on artificial neural network weights, *Applied Soft Computing*, 127, 109395. doi:10.1016/j.asoc.2022.109395

15. Emanet, S., Karatas Baydogmus, G. and Demir, O. (2021) Öznitelik seçme yöntemlerinin makine öğrenmesi tabanlı saldırı tespit sistemi performansına etkileri, *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 12(5), 743-755. doi:10.24012/dumf.1051340

16. Ercan, E. (2020) Conduct asset performance management with a software-based approach, *Plant Engineering*, 74(3), 27-29.

17. Eroğlu, D. Y. (2019) Systematization, Implementation and Analysis of the Overall Throughput Effectiveness Calculation for the Finishing Processes after Weaving, *Journal of Textile & Apparel/Tekstil ve Konfeksiyon*, 29(2). doi:10.32710/tekstilvekonfeksiyon.457170

18. Erturan, A. M., Karaduman, G. and Durmaz, H. (2023) Machine learning-based approach for efficient prediction of toxicity of chemical gases using feature selection, *Journal of Hazardous Materials*, 455, 131616. doi:10.1016/j.jhazmat.2023.131616

19. Fathima, M. D., Samuel, S. J., Natchadalingam, R. and Kaveri, V. V. (2022) Majority voting ensembled feature selection and customized deep neural network for the enhanced clinical decision support system, *International Journal of Computers and Applications*, 44(10), 991-1001. doi:10.1080/1206212X.2022.2069643

20. Fletcher, L., Akhtar, N., Zhan, X., Jafarikia, M., Sullivan, B. P., Huber, L.-A. and Li, J. (2022) Identification of Candidate Salivary, Urinary and Serum Metabolic Biomarkers for High Litter Size Potential in Sows (Sus scrofa), *Metabolites*, 12(11), 1045. doi:10.3390/metabo12111045

21. Genç, İ. and Vupa Çilengiroğlu, Ö. (2021, June 4-6, 2021). Toplam Ekipman Etkinliği Puanının Lojistik ve Karar Ağacı Algoritmaları ile Modellenmesi. International Conference on Data Science & Applications, Online.

22. Gunes, H., Coramik, M., Bicakci, S., Citak, H. and Ege, Y. (2022) Crack identification system on MOH cold rolled grain oriented sheets: Application of K-fold cross validated BRANN, *Measurement*, 195, 111128. doi:10.1016/j.measurement.2022.111128

23. Jamei, M., Karbasi, M., Alawi, O. A., Kamar, H. M., Khedher, K. M., Abba, S. I. and Yaseen, Z. M. (2022) Earth skin temperature long-term prediction using novel extended Kalman filter integrated with Artificial Intelligence models and information gain feature selection, *Sustainable Computing: Informatics and Systems*, 35, 100721. doi:10.1016/j.suscom.2022.100721

24. Korkmaz, G. and Eroğlu, E. (2020) Model karmaşıklığının kontrolü, *İktisadi ve İdari Yaklaşımlar Dergisi*, 2(2), 146-162. doi:10.47138/jeaa.780031

25. Kushwaha, N. L., Rajput, J., Suna, T., Sena, D. R., Singh, D. K., Mishra, A. K., Sharma, P. K. and Mani, I. (2023) Metaheuristic approaches for prediction of water quality indices with relief algorithm-based feature selection, *Ecological Informatics*, 75, 102122. doi:10.1016/j.ecoinf.2023.102122

26. Lai, C.-M., Chiu, C.-C., Shih, Y.-C. and Huang, H.-P. (2022) A hybrid feature selection algorithm using simplified swarm optimization for body fat prediction, *Computer Methods and Programs in Biomedicine*, 226, 107183. doi:10.1016/j.cmpb.2022.107183

27. Lap, B. Q., Phan, T.-T.-H., Nguyen, H. D., Quang, L. X., Hang, P. T., Phi, N. Q., Hoang, V. T., Linh, P. G. and Hang, B. T. T. (2023) Predicting Water Quality Index (WQI) by feature selection and machine learning: A case study of An Kim Hai irrigation system, *Ecological Informatics*, 74, 101991. doi:10.1016/j.ecoinf.2023. 101991

28. Li, W. and Becker, D. M. (2021) Day-ahead electricity price prediction applying hybrid models of LSTM-based deep learning methods and feature selection algorithms under consideration of market coupling, *Energy*, 237, 121543. doi:https://doi.org/10.1016/j.energy.2021.121543

29. Liu, W., Liu, T., Liu, Z., Luo, H. and Pei, H. (2023) A novel deep learning ensemble model based on two-stage feature selection and intelligent optimization for water quality prediction, *Environmental Research*, 224, 115560. doi:https://doi.org/10.1016/ j.envres.2023.115560

30. Liu, Y. and Schumann, M. (2005) Data mining feature selection for credit scoring models, *Journal of the Operational Research Society*, 56(9), 1099-1108. doi:10.1057/palgrave.jors.2601976

31. Marafino, B. J., John Boscardin, W. and Adams Dudley, R. (2015) Efficient and sparse feature selection for biomedical text classification via the elastic net: Application to ICU risk stratification from nursing notes, *Journal of Biomedical Informatics*, 54, 114-120. doi:10.1016/j.jbi.2015.02.003

32. Moldovan, D. and Slowik, A. (2021) Energy consumption prediction of appliances using machine learning and multi-objective binary grey wolf optimization for feature selection, *Applied Soft Computing*, 111, 107745. doi:10.1016/j.asoc.2021.107745

33. Naseri, H., Waygood, E. O. D., Wang, B., Patterson, Z. and Daziano, R. A. (2022) A Novel Feature Selection Technique to Better Predict Climate Change Stage of Change, *Sustainability*, 14(1), 40. doi:10.3390/su14010040

34. Olu-Ajayi, R., Alaka, H., Sulaimon, I., Balogun, H., Wusu, G., Yusuf, W. and Adegoke, M. (2023) Building energy performance prediction: A reliability analysis and evaluation of feature selection methods, *Expert Systems with Applications*, 225, 120109. doi:10.1016/j.eswa.2023.120109

35. Pasha, S. J. and Mohamed, E. S. (2022) Advanced hybrid ensemble gain ratio feature selection model using machine learning for enhanced disease risk prediction, *Informatics in Medicine Unlocked*, 32, 101064. doi:10.1016/j.imu.2022.101064

36. Petersson, H., Klingvall, R. and Holmberg, M. (2009) Sensor array optimization using variable selection and a Scanning Light Pulse Technique, *Sensors and Actuators B: Chemical*, 142(2), 435-445. doi:10.1016/j.snb.2009.04.029

37. Piran, F. A. S., De Paris, A., Lacerda, D. P., Camargo, L. F. R., Serrano, R. and Cassel, R. A. (2020) Overall Equipment Effectiveness: Required but not Enough—An Analysis Integrating Overall Equipment Effect and Data Envelopment Analysis, *Global Journal of Flexible Systems Managemen*t, 21(2), 191-206. doi:10.1007/s40171-020-00238-6

38. Prasetyo, Y. T. and Veroya, F. C. (2020). An Application of Overall Equipment Effectiveness (OEE) for Minimizing the Bottleneck Process in Semiconductor Industry. 2020 IEEE 7th International Conference on Industrial Engineering and Applications (ICIEA).

39. Qiao, Q., Yunusa-Kaltungo, A. and Edwards, R. E. (2022) Feature selection strategy for machine learning methods in building energy consumption prediction, *Energy Reports*, 8, 13621-13654. doi:10.1016/j.egyr.2022.10.125

40. Qu, K., Xu, J., Hou, Q., Qu, K. and Sun, Y. (2023) Feature selection using Information Gain and decision information in neighborhood decision system, *Applied Soft Computing*, 136, 110100. doi:10.1016/j.asoc.2023.110100

41. Saud, S., Jamil, B., Upadhyay, Y. and Irshad, K. (2020) Performance improvement of empirical models for estimation of global solar radiation in India: A k-fold cross-validation approach, *Sustainable Energy Technologies and Assessments*, 40, 100768. doi:10.1016/j.seta.2020.100768

42. Sun, J., Liu, G., Sun, B. and Xiao, G. (2021) Light-stacking strengthened fusion based building energy consumption prediction framework via variable weight feature selection, *Applied Energy*, 303, 117694. doi:10.1016/j.apenergy.2021.117694

43. Sun, S., Hu, M., Wang, S. and Zhang, C. (2023) How to capture tourists' search behavior in tourism forecasts? A two-stage feature selection approach, *Expert Systems with Applications*, 213, 118895. doi:10.1016/j.eswa.2022.118895

44. Toğaçar, M., Ergen, B. and Cömert, Z. (2020a) Application of breast cancer diagnosis based on a combination of convolutional neural networks, ridge regression and linear discriminant analysis using invasive breast cancer images processed with autoencoders, *Medical Hypotheses*, 135, 109503. doi:10.1016/j.mehy.2019.109503

45. Toğaçar, M., Ergen, B. and Cömert, Z. (2020b) Waste classification using AutoEncoder network with integrated feature selection method in convolutional neural network models, *Measurement*, 153, 107459. doi:10.1016/j.measurement.2019.107459

46. Tsanas, A. (2022) Relevance, redundancy, and complementarity trade-off (RRCT): A principled, generic, robust feature-selection tool, *Patterns*, 3(5), 100471. doi:10.1016/j.patter.2022.100471

47. Verma, D., Jansen, D., Bach, K., Poel, M., Mork, P. J. and d'Hollosy, W. O. N. (2022) Exploratory application of machine learning methods on patient reported data in the development of supervised models for predicting outcomes, *BMC Medical Informatics and Decision Making*, 22(1), 227. doi:10.1186/s12911-022-01973-9

48. Vogel-Heuser, B., Karaseva, V., Folmer, J. and Kirchen, I. (2017) Operator Knowledge Inclusion in Data-Mining Approaches for Product Quality Assurance using Cause-Effect Graphs, *IFAC-PapersOnLine*, 50(1), 1358-1365. doi:10.1016/j.ifacol.2017. 08.233

49. Wang, T., Hu, Z. and Liu, H. (2023) A unified view of feature selection based on Hilbert-Schmidt independence criterion, *Chemometrics and Intelligent Laboratory Systems*, 104807. doi:10.1016/j.chemolab.2023.104807

50. Wang, Z., Lei, Y., Cui, H., Miao, H., Zhang, D., Wu, Z. and Liu, G. (2022) Enhanced RBF neural network metamodelling approach assisted by sliced splitting-based K-fold cross-validation and its application for the stiffened cylindrical shells, *Aerospace Science and Technology*, 124, 107534. doi:10.1016/j.ast.2022.107534

51. Wei, Y., Chen, Z., Zhao, C., Chen, X., He, J. and Zhang, C. (2023) A time-varying ensemble model for ship motion prediction based on feature selection and clustering methods, *Ocean Engineering*, 270, 113659. doi:10.1016/j.oceaneng.2023.113659

52. Wu, T., Hao, Y., Yang, B. and Peng, L. (2023) ECM-EFS: An ensemble feature selection based on enhanced co-association matrix, *Pattern Recognition*, 139, 109449. doi:10.1016/j.patcog.2023.109449

53. Xing, W., Pei, B., Li, S., Chen, G. and Xie, C. (2019) Using learning analytics to support students' engineering design: the angle of prediction, *Interactive Learning Environments*, 1-18. doi:10.1080/10494820.2019.1680391

54. Xu, G. and Wang, X. (2023) Support vector regression optimized by black widow optimization algorithm combining with feature selection by MARS for mining blast vibration prediction, *Measurement*, 218, 113106. doi:10.1016/j.measurement.2023. 113106

55. Yan, W.-L. (2023) Stock index futures price prediction using feature selection and deep learning, *The North American Journal of Economics and Finance*, 64, 101867. doi:10.1016/j.najef.2022.101867

56. Yao, G., Hu, X. and Wang, G. (2022) A novel ensemble feature selection method by integrating multiple ranking information combined with an SVM ensemble model for enterprise credit risk prediction in the supply chain, *Expert Systems with Applications*, 200, 117002. doi:10.1016/j.eswa.2022.117002

57. Zhao, L., Li, Y., Li, S. and Ke, H. (2023) A frequency item mining based embedded feature selection algorithm and its application in energy consumption prediction of electric bus, *Energy*, 271, 126999. doi:10.1016/j.energy.2023.126999

58. Zhao, Z., Wu, J., Cai, F., Zhang, S. and Wang, Y.-G. (2022) A statistical learning framework for spatial-temporal feature selection and application to air quality index forecasting, *Ecological Indicators*, 144, 109416. doi:10.1016/j.ecolind.2022.109416

59. Zhou, J. and Zhang, W. (2023) Coal consumption prediction in thermal power units: A feature construction and selection method, *Energy*, 273, 126996. doi:10.1016/j.energy.2023.126996