Financial Internet Quarterly 2025, vol. 21 / no. 1, p. 42-66
10.2478/fiqf-2025-0004

Financial Internet Quarterly

SECTOR-SPECIFIC FINANCIAL FORECASTING WITH MACHINE LEARNING ALGORITHM AND SHAP INTERACTION VALUES

Cansu Ergenç[1], Rafet Aktaş[2]

Abstract

This study examines the application of machine learning models to predict financial performance in various sectors, using data from 21 companies listed in the BIST100 index (2013-2023). The primary objective is to assess the potential of these models in improving financial forecast accuracy and to emphasize the need for transparent, explainable approaches in finance. A range of machine learning models, including Linear Regression, Ridge, Lasso, Decision Tree, Bagging, Random Forest, AdaBoost, Gradient Boosting (GBM), LightGBM, and XGBoost, were evaluated. Gradient Boosting emerged as the best-performing model, with ensemble methods generally demonstrating superior accuracy and stability compared to linear models. To enhance interpretability, SHAP (SHapley Additive exPlanations) values were utilized, identifying the most influential variables affecting predictions and providing insights into model behavior. Sector-based analyses further revealed differences in model performance and feature impacts, offering a granular understanding of financial dynamics across industries. The findings highlight the effectiveness of machine learning, particularly ensemble methods, in forecasting financial performance. The study underscores the importance of using explainable models in finance to build trust and support decision-making. By integrating advanced techniques with interpretability tools, this research contributes to financial technology, advancing the adoption of machine learning in data-driven investment strategies.

JEL classification: C51, C52, C53
Keywords: Machine Learning Models, SHAP, Financial Forecasting

Received: 31.07.2024 Accepted: 15.11.2024

Cite this:
Ergenç, C. & Aktaş, R. (2025). Sector-specific financial forecasting with machine learning algorithm and SHAP interaction values. Financial Internet Quarterly 21(1), pp. 42-66.

[1] Ankara Yildirim Beyazit University, Ankara, Turkey, e-mail: cansuergenc@aybu.edu.tr, ORCID: https://orcid.org/0000-0002-4722-0911.
[2] Ankara Yildirim Beyazit University, Ankara, Turkey, e-mail: raktas@aybu.edu.tr, ORCID: https://orcid.org/0009-0008-8033-4604.

Introduction

Financial performance has always been crucial for companies, impacting nations globally. It is crucial for all countries and companies (Perrini et al., 2011; Barauskaite & Streimikiene, 2020). In recent years, the combination of finance and artificial intelligence has not just led to progress, but a transformation in financial forecasting (Lin, 2019; Nguyen et al., 2022; Avelar & Jordão, 2024). Machine learning algorithms also play a major role in this transformation. Because machine learning algorithms have provided advanced techniques that can process large amounts of data, identify patterns, and make predictions with unprecedented accuracy (Zhou et al., 2017; Mahalakshmi et al., 2022; Bouchefry & De Souza, 2020). Learning from historical data and adapting to new information, which is a feature of machine learning models, and the performance of models that improve over time are very important developments for finance (Pandey & Sergeeva, 2022; Ionescu & Diaconita, 2023; George, 2024).

The place of accurate financial forecasting for financial markets is undeniable (Penman, 2002; Samonas, 2015; Kumar, 2017; Barnhizer & Barnhizer, 2019; Sastry, 2020; Massei, 2023). Investors reduce their financial risks and make informed investments by making the right investment decisions for accurate financial forecasts. Financial analysts, on the other hand, make recommendations to market participants in line with the results obtained from financial forecasts (Ramnath et al., 2008; Samonas, 2015; Magnan et al., 2015).

Policy makers use financial forecasts to prevent possible financial crises and guide the current economy. Managers can benefit from these financial forecasts in their strategic decisions regarding budgeting (Ramnath et al., 2008; Oliva & Watson, 2009; Magnan et al., 2015; Ballings et al., 2015; Geng et al., 2015). With such results, machine learning models are rapidly gaining acceptance in the field of finance.

When machine learning models used in financial forecasting are examined, it is seen that methods such as neural networks, decision trees and ensemble methods are used. Each method has its own advantages and disadvantages (Katal et al., 2013; Provost & Fawcett, 2013; Chen & Zhang, 2014; Najafabadi et al., 2015). The performances of these methods vary depending on the structure and size of the data used. The fact that these models give good results despite the complex structure of financial data has caused them to be preferred in areas such as credit risks, stock income, and estimating the total income of companies. In addition, the use of big data technologies has enabled the processing and analysis of large data sets, which has increased the precision and reliability of financial forecasts (Oliva & Watson, 2009; Provost & Fawcett, 2013; Chen & Zhang, 2014; Zhou et al., 2017).

BIST100 index is an important indicator of the Turkish stock market, which includes a wide range of sectors. Therefore, it provides a very comprehensive dataset in terms of financial data estimation, where the performance of machine learning models is evaluated. This research evaluates the performance of machine learning models in financial data. This evaluation is carried out both on the general performance of companies listed in BIST100 between 2013-2023 and on a sectoral basis. As a result, both the performance of machine learning models in financial markets and their performance on a sectoral basis are examined.

Our study uses SHAP (SHapley Additive Explanations) values in addition to traditional performance measures to interpret machine learning models. SHAP values increase the transparency and explainability of complex ML algorithms by providing insights into feature importance and interaction effects (Bhattacharya, 2022; Li, 2022; Baptista et al., 2022; Baptista, 2022). By examining SHAP values, this research not only evaluates the predictive accuracy of the models, but also clarifies the key factors affecting financial results. Thus, the importance of the variables in the models used for the model is also examined. The results of this study will be valuable for both academic research and real-world use and will provide important insights for investors, financial analysts, and policy makers. The rest of this paper is structured as follows. Section 2 provides a literature review of the machine learning forecasting models and factors that influence financial performance. Section 3 presents the methodology and summarizes nine machine learning models used to forecast financial performance. The results obtained are discussed in Section 4. Finally, the conclusion is presented in Section 5.

Literature review

In recent years, there has been considerable progress in financial forecasting using machine learning algorithms. Machine learning models are increasingly used in the financial sector to predict stock prices and classification (Sonkavde, 2023). Traditional models such as linear regression are still used (Gzar et al., 2022). Especially in predicting results based on input features, linear regression is a highly preferred model due to its simplicity and interpretability (Rosenbusch et al., 2019; Ryll ve Seidens, 2019; Seno, 2023). Machine learning models have been used in a wide range of financial domains for purposes such as credit default prediction and tourism demand forecasting, providing valuable insights for economic analysis and crisis detection, and have demonstrated the versatility and effectiveness of these algorithms in different sectors (Fan, 2023; Clavería et al., 2015; Afreen, 2020).

Linear regression is often complemented by other algorithms such as ridge regression, lasso regression and support vector regression to increase prediction accuracy (Xiao et al., 2020; Yoo et al., 2022). In addition, studies compare performance with models such as Random Forest, XGBoost and LSTM (Sonkavde, 2023). With the use of machine learning models in the financial sector, which model will fit the data better has become an important issue (Long et al., 2022; Akinrinola, 2024). Decision trees, which are a frequently used model among machine learning models, are preferred due to their effectiveness, interpretability and ease of visualization (Kourtellis et al., 2016; Moshkov, 1997; Azad et al., 2022; Poojitha & Kanagasabai, 2022). The structure of financial data is complex and variable, and Gradient Boosting, which has shown significant success in various practical applications due to its ability to handle complex relationships and produce accurate predictions in the use of such data, can be preferred (Natekin & Knoll, 2013; Chen, 2016; Kadiyala & Kumar, 2018; Davis et al., 2020). Along with this method, radiant Boosting algorithms such as XGBoost, LightGBM and others have become popular choices in the machine learning community due to their effectiveness in improving model performance and prediction accuracy (Mienye & Sun, 2022; Siringoringo et al., 2021; Zhang et al., 2011).

LightGBM has been compared with other machine learning models such as Random Forest, XGBoost, and traditional gradient boosting in the literature, and has outperformed these models in terms of performance, speed, accuracy, and efficiency (Fraz, 2024; Grissa et al., 2020; Unal et al., 2021; Jiang, 2024). LightGBM has been successfully used in various fields, including health, environmental science, finance, and geology (Rufo et al., 2021; Su et al., 2021; Park et al., 2021; Dong et al., 2022; Ko et al., 2022; Jiang, 2024; Xiang, 2024; Wang, 2024). Furthermore, the versatility of LightGBM is evident in its applications in various fields such as fault detection in wind turbines (Tang et al., 2020), intrusion detection in IoT systems (Zhao et al., 2023), fraud detection in banking data (Hashemi et al., 2023), and malware detection (Onoja et al., 2022). Another alternative to LightGBM is the XGBoost model. The XGBoost algorithm has been shown to exhibit very high accuracy and performance on various datasets (Chen, 2016; Kareem et al., 2023). It has been successfully used in various fields including election prediction (Suacana, 2024), aircraft icing severity assessment (Li et al., 2020), surface milling accuracy (Abbas, 2023), analysis of imbalanced data (Zhang et al., 2022), and prediction and optimization tasks (Zhang, 2024). It has been used in various applications such as jaundice detection in newborns (Abdulrazzak, 2024), fault detection in photovoltaic panels (Sairam, 2020), outcome prediction in healthcare (Deng et al., 2022) and cervical cancer screening (Edafetanure-Ibeh, 2024), line loss prediction (Wang et al., 2017), Arctic navigation risk assessment (Yao et al., 2023), PM2.5 concentration estimation (Pan, 2018) and permeate flux prediction in osmosis processes (Shi et al., 2022).

SHAP interaction values are very important for increasing the accuracy of machine learning models. They improve model interpretability by capturing local interaction effects between features, especially in models built on financial data (Orsini et al., 2022; Zern et al., 2023). In addition, SHAP interaction values ensure consistent individualized feature attribution for tree communities, providing consistent explanations for interaction effects in individual predictions (Lundberg et al., 2018; Mitchell et al., 2022). Using SHAP interaction values makes models more understandable and allows for a quantitative study of interaction effects (Long et al., 2022; Martini et al., 2022). As a result, it provides a unified approach to interpret complex model predictions and contributes to a more comprehensive understanding of model behavior (Li et al., 2020; Lundberg et al., 2020).

Methodology

In this section, we present the approach used to forecast the financial performance of companies listed on the BIST100 index from 2013 to 2023. The dataset consists of financial metrics such as Net Income, Total Assets, Total Liabilities, and Shareholders' Equity, which serve as the independent variables, while Total Revenue is the target variable. The data is split into a training set (80%) and a test set (20%) to ensure proper evaluation of model performance. We employ ten machine learning models: Linear Regression, Ridge Regression, Lasso Regression, Decision Tree, Bagging, Random Forest, AdaBoost, Gradient Boosting (GBM), LightGBM, and XGBoost. These models are chosen due to their varying complexity and ability to handle different types of financial data. We apply several evaluation metrics, including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), to assess the accuracy and robustness of the models. Each model's predictive performance is compared against the test set to evaluate its ability to generalize.

To enhance model interpretability, we use SHAP (SHapley Additive exPlanations) values, which allow us to assess the contribution of each input variable to the model's predictions. This helps in understanding the importance of financial metrics like Net Profit, Long-Term Liabilities, and Total Assets in driving financial performance outcomes. Additionally, we ensure that all models are configured to account for the temporal nature of the data, although no explicit time-series models were used. Neighboring vectors of data are considered within the framework of machine learning models to ensure that the time context is respected during training and predictions.

Machine learning models

Linear regression

Linear regression analysis assumes a linear relationship among multiple variables (Schroeder et al., 2016). The general Linear Regression model can be stated by the equation below:

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i \tag{1}$$

where, y_i dependent variable, x_i explanatory variables, β_0 constant term, β_k slope coefficients for each explanatory variable, ε_i the model's error term.

Ridge regression

Ridge regression is an extension of linear regression, known for its bias-variance trade-off control that provides a balance between model complexity and generalization performance, is a valuable technique used to address multicollinearity in datasets where independent variables are highly correlated (Malthouse, 1999; Kibria & Saleh, 2004; Khalaf, 2012; Kumar et al., 2021). By adding a penalty term to the OLS method, ridge regression helps to stabilize the predictions and prevent overfitting, making it a more reliable and consistent method for modeling relationships between variables (Xin & Khalid, 2018; Wei & Diğerleri, 2020; Li, 2024). Ridge regression minimizes the following cost function:

$$\widehat{\beta} = argmin\left\{\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}\beta_j^2\right\} \tag{2}$$

where λ is the regularization parameter.

Lasso regression

Lasso regression is a widely used technique in regression analysis known for its ability to perform variable selection and regularization (Rajaratnam et al., 2015; Signorino & Kirchner, 2018; Friedman et al., 2010). Lasso regression minimizes the following cost function:

$$\widehat{\beta} = argmin\left\{\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j x_{ij})^2 + \lambda\sum_{j=1}^{p}\left|\beta_j\right|\right\} \tag{3}$$

where λ is the regularization parameter.

Decision tree

A decision tree is a decision support tool that uses a tree-like graph to represent decisions and their potential outcomes, aiding in understanding complex scenarios and making predictions based on input data (Lo et al., 2014). The prediction for a decision tree is given by:

$$f(x) = \sum_{m=1}^{M} w_m I(x \in R_m) \tag{4}$$

where M is the number of terminal nodes, w_m is the predicted value in region R_m, and I(·) is an indicator function.

Bagging

Bagging, short for bootstrap aggregating, is a technique that involves generating multiple versions of a predictor by resampling the training data and then aggregating these predictors to create a more robust and accurate model (Breiman, 1996; Gianola et al., 2014; Soloff et al., 2023). Bagging prediction is:

$$\widehat{f}(x) = \frac{1}{B}\sum_{b=1}^{B} f_b(x) \tag{5}$$

where B is the number of bootstrap samples and $f_b(x)$ is the prediction from the b - th bootstrap sample.

Random forest

Random Forest is an ensemble supervised learning algorithm known for its high accuracy in classification tasks (Ilma et al., 2023; Sandhya & Padyana, 2021; Genuer, 2012). It generates multiple decision trees by resampling the data and aggregating the predictions, resulting in a robust and accurate model. (Genuer, 2012, Strobl et al., 2008; Mishina et al., 2015; Kulkarni & Sinha, 2012). Random Forest prediction is:

$$\widehat{f}(x) = \frac{1}{T}\sum_{t=1}^{T} f_b(x) \tag{6}$$

where T is the number of trees, and $f_t(x)$ is the prediction of the \$t\$ - th tree.

AdaBoost

AdaBoost, short for Adaptive Boosting, is an ensemble learning method that combines multiple weak learners to create a strong classifier (Paul et al., 2009; Meir & Rätsch, 2003). It iteratively adjusts the weights of incorrectly classified instances to focus on difficult cases, improving the overall model performance (Wang et al., 2022; Yin et al., 2017; Si et al., 2022). AdaBoost prediction is:

$$\widehat{f}(x) = \sum_{t=1}^{T}\alpha_t f_t(x) \tag{7}$$

where T is the number of trees, α_t is the weight assigned to the \$t\$ - th tree based on its accuracy, and $f_+(x)$ is the prediction of the t - th tree.

Gradient boosting

Gradient Boosting is a powerful ensemble machine learning technique that iteratively builds a series of weak learners, typically decision trees, to create a strong predictive model. By focusing on the errors of the previous models during training, Gradient Boosting aims to improve prediction accuracy by minimizing the overall loss function (Zhang et al., 2011; Mayr & Schmid, 2014; Johnson & Zhang, 2014). Gradient Boosting prediction is:

$$\widehat{f}(x) = \sum_{t=1}^{T} v f_t(x) \tag{8}$$

where T is the number of trees, v is the learning rate, and $f_t(x)$ is the t - th tree trained to predict the residuals of the previous trees.

LightGBM

LightGBM, short for Light Gradient Boosting Machine, is an extremely efficient algorithm designed for gradient-boosting decision trees (Jiang, 2024). LightGBM prediction is:

$$\widehat{f}(x) = \sum_{t=1}^{T} f_t(x) \tag{9}$$

where T is the number of trees, and $f_t(x)$ is the prediction of the t - th tree using the Light GBM framework, which employs gradient-based one-side sampling and exclusive feature bundling.

XGBoost

XGBoost, short for Extreme Gradient Boosting, is a powerful machine learning algorithm renowned for its scalability, speed, and accuracy (Chen, 2016). XGBoost prediction is:

$$\widehat{f}(x) = \sum_{t=1}^{T} f_t(x) \tag{10}$$

where T is the number of trees, and $f_t(x)$ is the prediction of the t - th tree using the XGBoost algorithm. which optimizes for a reqularized obiective to prevent overfitting.

Performance metrics of models

The evaluation of these models was conducted using several key performance metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and relative Root Mean Squared Error (rRMSE). The evaluation of the machine learning models in this study is based on several key performance metrics that quantify the accuracy and robustness of the predictions. The metrics used are as follows:

$$MSE = \frac{1}{n}\sum (Y_i - \widehat{Y}_i)^2 \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum (Y_i - \widehat{Y}_i)^2} \tag{12}$$

$$MAE = \frac{1}{n}\sum \left| Y - \widehat{Y}_i \right| \tag{13}$$

$$MAPE = \frac{100/\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y}_i}{y_i}\right| \tag{14}$$

$$rRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}}{\bar{y}} \tag{15}$$

SHAP (Shapley Additive Explanation) Approach

SHapley Additive exPlanations (SHAP) values are a method rooted in cooperative game theory that aims to provide a fair allocation of importance values to features in machine learning models (Uddin et al., 2022). This approach has been utilized in various studies to enhance the interpretability and transparency of machine learning models across different domains. For instance, SHAP values have been employed to interpret the outputs of support vector machines, random forests, convolutional neural networks, and long short-term memory models in forecasting climatic water balance (Uddin et al., 2022). Additionally, SHAP has been used to interpret models in predicting sepsis in-hospital mortality (Zhang, 2024), automating data center operations (Gebreyesus, 2024), and developing prognostic models for critically ill patients (Fan et al., 2023). The application of SHAP values extends to diverse areas such as predicting tropical cyclogenesis (Loi, 2024), evaluating hospital mobility (Santamato, 2024), predicting gout associated with dietary factors (Cao, 2024), and optimizing photodegradation rate predictions (Schossler, 2023). By leveraging SHAP values, researchers have gained deeper insights into model predictions, feature importance, and the specific contributions of variables to the outcomes of machine learning models (Cao, 2024). Furthermore, SHAP values have been instrumental in enhancing the interpretability, explainability, and fairness of machine learning models (Hickey et al., 2020).

For a model and input features , the SHAP value for a feature is given by:

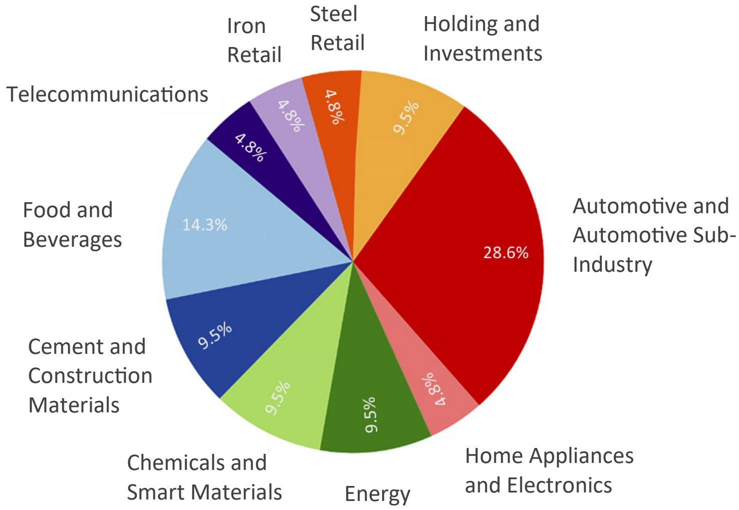
$$\phi i = \sum_{S \subseteq N} \frac{S!(|N| - |S| - 1)!}{|N|!}[f_s(x_s \cup \{i\}) - f_s(x_s)] \tag{16}$$

Where, N is the set of all features, S is any subset of N that does not include feature I, $f_s(x_s \cup \{i\})$ is the prediction of the model when feature i is included in subset S, $f_s(x_s)$ is the prediction of the model when feature i is not included.

Results

In this article, we investigate the impact of variables affecting financial performance on total revenue. The data consists of 21 companies listed in BIST100 for a 10-year period (2013-2023). Figure 1 shows the sectoral distribution of companies.

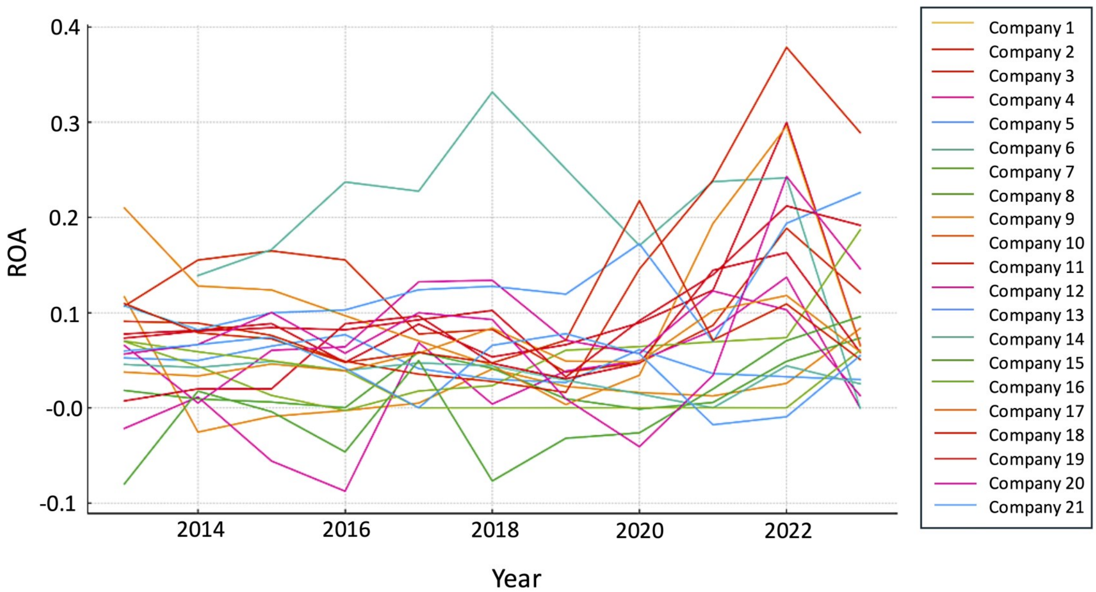
Figure 1: Sectoral Distribution of Companies



Source: Author's own work.

This study is separated into training (80%) and testing (20%) datasets to compare the performance of various machine learning models. The dataset is randomly split, with 80% used as the training dataset and the remaining 20% as the testing set. This approach is commonly used in prior studies (Abellán & Mantas, 2014; Antunes et al., 2017; Ben Jabeur et al., 2020). In our study, Net Income, Total Assets, Total Liabilities, and Shareholders' Equity, Short-term Liabilities, Long-term Liabilities were treated as independent features, while Total Revenue served as the output or target feature. Figure 2 shows the ROA for each company from 2013 to 2023.

Figure 2: ROA for each company from 2013 to 2023



Source: Author's own work.

Comparison of model performance

The performance of various machine learning models was evaluated using five key metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and relative Root Mean Squared Error (rRMSE). Table 1 summarizes the performance metrics for each model. Appendix 1 shows the performance of machine learning models over the test sample.

Table 1: MSE, RMSE, MAE, MAPE and rRMSE values

Model	MSE	RMSE	MAE	MAPE	rRMSE
Linear regression	0.0040	0.0635	0.0430	41.22%	0.736
Ridge regression	0.0040	0.0635	0.0430	41.22%	0.736
Lasso Regression	0.0040	0.0635	0.0430	41.22%	0.736
Decision Trees	0.0046	0.0676	0.0456	118.58%	0.783
Bagging	0.0016	0.0399	0.0239	4.19%	0.462
Random Forests	0.0016	0.0403	0.0246	4.55%	0.466
Adaboost	0.0018	0.0428	0.0269	2.83%	0.496
GBM	0.0014	0.0378	0.0235	13.92%	0.438
LightGBM	0.0044	0.0695	0.0719	69.84%	0.777
XGBoost	0.0046	0.0679	0.0616	73.54%	0.786

Source: Author’s own work.

The linear models, including Linear Regression, Ridge Regression, and Lasso Regression, exhibit identical performance metrics. These models are characterized by their simplicity and baseline nature, which is reflected in the relatively high values of MSE, RMSE, MAE, and rRMSE. The MAPE for these models is notably large at 41.22%, indicating that they may struggle to capture the complex relationships within the financial data effectively. This limitation suggests that while these models are straightforward to interpret, they are not well-suited for accurately predicting financial performance in this context. The Decision Tree model shows a higher MSE and RMSE compared to the linear models, with an exceptionally high MAPE of 118.58%. This high MAPE suggests that the Decision Tree model tends to overfit the data, making it less reliable for prediction despite its interpretability. The overfitting is likely due to the model's tendency to capture noise and fluctuations in the training data, leading to poor generalization to new data. On the other hand, ensemble methods such as Bagging and Random Forests demonstrate significantly better performance than the individual Decision Tree model. These models exhibit lower MSE, RMSE, and MAE values, with Bagging showing a slightly better performance than Random Forests. The MAPE values for Bagging and Random Forests are impressively low at 4.19% and 4.55%, respectively, indicating strong predictive accuracy and stability. These results highlight the effectiveness of ensemble methods in reducing variance and improving the robustness of predictions. AdaBoost performs well, with an MSE of 0.0018 and an RMSE of 0.0428. The model shows a remarkably low MAPE of 2.83%, underscoring its ability to handle complex data and improve overall prediction accuracy by focusing on misclassified instances. AdaBoost's iterative approach to adjusting the weights of misclassified instances contributes to its enhanced performance and reliability. Gradient Boosting (GBM) outperforms most models with the lowest MSE of 0.0014 and RMSE of 0.0378. The model's MAE and MAPE values indicate high accuracy and precision in predictions, making it a robust choice for financial forecasting. GBM's ability to iteratively improve upon errors made by previous models results in superior predictive capabilities and robustness. Also, LightGBM, known for its efficiency, shows higher error metrics compared to other boosting methods. This could be due to the model's sensitivity to the dataset characteristics or the hyperparameter settings used in this study. Its MAPE of 69.84% indicates considerable prediction errors in certain instances, suggesting that further tuning and adjustment may be needed to optimize its performance for this specific dataset. Similarly, XGBoost, another popular boosting algorithm, performs akin to Decision Trees, with an MSE of 0.0046 and an RMSE of 0.0679. However, it shows a relatively high MAPE of 73.54%, indicating that it may not be the best fit for this specific dataset without further tuning. The higher error metrics suggest that XGBoost's default settings might not be fully optimized for the financial forecasting task at hand. The variation in MAPE across these models can be attributed to their respective abilities to capture complex relationships in the financial data. Simpler models like Linear Regression, Ridge, and Lasso struggle with these intricacies, leading to higher error rates. On the contrary, ensemble methods like Bagging, Random Forests, and Gradient Boosting tend to mitigate overfitting and handle complex data relationships more effectively, resulting in lower MAPE and better predictive performance.

Comparison of model sector performance

The performance of the machine learning models was further analyzed across different sectors to understand how each model performed within specific industries. Table 2 presents the Mean Squared Error (MSE) values for each sector and model combination. Figure 4 shows Comparison of MSE Values Across Different Sectors and Models.

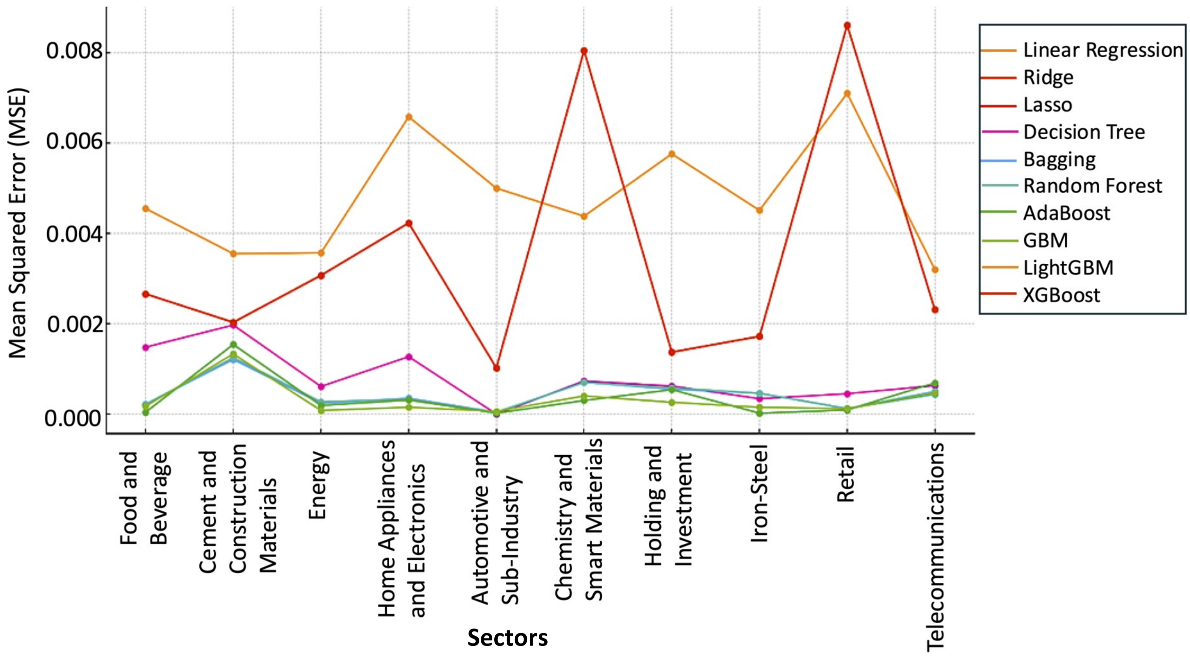
Table 2: MSE values for sector

Sectors	Linear regression	Ridge	Lasso	DT	Bagging
	MSE				
Food and Beverage	0.00266	0.00266	0.00266	0.00148	0.00022
Cement and Construction Materials	0.00203	0.00203	0.00203	0.00197	0.00121
Chemistry and Smart Materials	0.00307	0.00307	0.00307	0.00061	0.00024
Energy	0.00423	0.00423	0.00423	0.00127	0.00035
Home Appliances and Electronics	0.00102	0.00102	0.00102	0.00000	0.00004
Automotive and Automotive Sub-Industry	0.00805	0.00805	0.00805	0.00073	0.00070
Holding and Investment	0.00137	0.00137	0.00137	0.00062	0.00055
Iron-Steel	0.00172	0.00172	0.00172	0.00034	0.00046
Retail	0.00861	0.00861	0.00861	0.00045	0.00012
Telecommunications	0.00231	0.00231	0.00231	0.00063	0.00044
Sectors	**RF**	**Adaboost**	**GBM**	**LightGBM**	**XGBoost**
	MSE				
Food and Beverage	0.00021	0.00004	0.00017	0.00455	0.00266
Cement and Construction Materials	0.00125	0.00154	0.00133	0.00355	0.00203
Chemistry and Smart Materials	0.00027	0.00019	0.00008	0.00357	0.00307
Energy	0.00033	0.00031	0.00015	0.00658	0.00423
Home Appliances and Electronics	0.00004	0.00002	0.00006	0.00500	0.00102
Automotive and Automotive Sub-Industry	0.00070	0.00030	0.00040	0.00438	0.00805
Holding and Investment	0.00058	0.00054	0.00026	0.00576	0.00137
Iron-Steel	0.00046	0.00002	0.00015	0.00451	0.00172
Retail	0.00012	0.00009	0.00012	0.00710	0.00861
Telecommunications	0.00050	0.00069	0.00046	0.00320	0.00231

Source: Author's own work.

The Bagging model performs exceptionally well in this sector, achieving the lowest MSE of 0.00022. Random Forest follows closely with an MSE of 0.00021, indicating strong predictive performance. AdaBoost also performs well with an MSE of 0.00004, suggesting high accuracy in this sector. Bagging and Random Forests show better performance in this sector compared to other models, with MSE values of 0.00121 and 0.00125, respectively. GBM also performs well with an MSE of 0.00133, indicating good predictive capabilities in this industry. GBM and Bagging models exhibit the best performance in this sector with MSE values of 0.00008 and 0.00024, respectively. These results highlight the effectiveness of these ensemble methods in capturing the complexity of data in the chemistry and smart materials sector. In the energy sector, GBM stands out with an MSE of 0.00015, followed by Bagging with an MSE of 0.00035. These models demonstrate superior predictive accuracy, suggesting they are well-suited for forecasting financial performance in the energy industry.

Figure 4: Comparison of MSE values across different sectors and models



Source: Author's own work.

The Decision Tree model performs remarkably well in this sector with an MSE of 0.00000. Bagging and Random Forests also show good performance with MSE values of 0.00004 each, indicating their robustness in handling data from this sector. Bagging and Random Forests again show strong performance with MSE values of 0.00070 each. GBM also performs well with an MSE of 0.00040, highlighting the capability of these ensemble methods to accurately predict financial outcomes in the automotive industry. GBM and Bagging are the top performers in this sector with MSE values of 0.00026 and 0.00055, respectively. These models effectively capture the financial dynamics within holding and investment companies. Bagging and Random Forests achieve the lowest MSE values of 0.00046 each in the iron-steel sector. GBM follows closely with an MSE of 0.00015, demonstrating its robustness in this industry. Bagging and Random Forests exhibit superior performance in the retail sector with MSE values of 0.00012 each. GBM also performs well with an MSE of 0.00012, suggesting strong predictive capabilities for retail companies. In the telecommunications sector, GBM achieves the lowest MSE of 0.00046, followed by Bagging with an MSE of 0.00044. These results highlight the effectiveness of ensemble methods in predicting financial performance in the telecommunications industry.

Feature analysis

Appendix 2 shows the SHAP summary plot for the Linear Regression (a), Ridge Regression (b), Lasso Regression (c), Decision Tree (d), Bagging (e), Random Forest (f), AdaBoost (g), Gradient Boosting (h), LightGBM (i), and XGBoost (j), illustrating the impact of different features on the model's output. SHAP (SHapley Additive exPlanations) values help in understanding the contribution of each feature to the predictions made by the model.

Appendix 3 shows SHAP dependence plots for the relationship between feature pressure (x-axis) and SHAP values (y-axis). Red represents high values, while blue indicates low values. The SHAP summary plot provides insights into the importance and impact of various features on the model's predictions. The feature "Net Profit" has a significant impact on the model output. High values of net profit, indicated by red dots, tend to push the prediction towards positive values, while low values, indicated by blue dots, push it towards negative values. The dispersion of dots across the SHAP value axis indicates its substantial influence, making it a key predictor in the model. "Long-Term Liabilities" also play an important role in the model's predictions. High values of long-term liabilities are associated with a negative impact on the prediction, suggesting that greater long-term liabilities might decrease the predicted financial performance. This negative association highlights the potential risks associated with high long-term debts. The feature of "Total Assets" shows a varied impact on the model's output. High values of total assets generally contribute positively to the prediction, while low values contribute negatively, reflecting their significance in financial forecasting. The plot suggests that higher asset values enhance the company's financial outlook. "Total Income" has a mixed impact on the model's predictions. High values influence the prediction positively, whereas low values have a lesser but still notable impact. This indicates that total income is an important predictor of financial performance, with its fluctuations directly affecting the model's output. "Equity" shows a moderate impact on the model's predictions. The plot suggests that higher equity values slightly contribute to positive predictions, whereas lower values contribute to negative predictions. This moderate influence underscores the role of equity in providing financial stability and its contribution to the overall financial health of the company. Finally, "Short-Term Liabilities" appear to have the least impact among the listed features. While there is some dispersion, indicating variability, the overall influence on the model's output is less significant compared to other features. This suggests that short-term liabilities, while relevant, may not be as critical in determining long-term financial performance as other factors like net profit and total assets.

Discussion

The study uses different machine learning models to predict the financial performance of companies in different sectors in the BIS100 index. Among these models, Gradient Boosting (GBM) shows better results than other models in various performance metrics such as MSE, RMSE, MAE, MAPE and rRMSE. The conclusion that ensemble methods such as Gradient Boosting (GBM) and Bagging and Random Forests perform effectively in financial forecasting is supported by numerous research findings (Barboza et al., 2017; Zhan et al., 2021). Barboza et al. (2017) showed that Gradient Boosting and Bagging methods in machine learning models perform better than traditional statistical methods in bankruptcy prediction. Zhan et al. (2021) showed that the random forest bagging approach is a well performing model in pandemic prediction. In this study, it is seen that ensemble models show good results in sectors such as energy, chemicals and smart materials. Rohatgi et al. (2021) showed that the Gradient Boosting model performs well in predicting stock market movements in complex and volatile sectors.

SHAP values are very important for better performance of models and clearer interpretability. Prasad and Bakhshi (2022) showed the contribution of correct interpretation of these values to the model in their study. The potential of machine learning models to improve financial forecast accuracy and their importance in strategic decision making has been shown in many studies (Astrakhantseva & Gerasimov, 2023). The results of this study have far-reaching implications for many sectors. Machine learning models have great potential for more accurate financial forecasts.

The results of this study have far-reaching implications for the financial technology field. They highlight the potential of machine learning models to improve the accuracy of financial forecasts, which is invaluable for reducing risks and optimizing investment strategies. The findings also highlight the importance of model interpretability, especially in financial applications, where understanding the logic behind the forecasts can be as important as the forecasts themselves. This is particularly important given the increasing scrutiny of algorithmic transparency and accountability in finance.

While the study demonstrates the effectiveness of machine learning models in financial forecasting, it also reveals certain limitations. For example, models such as LightGBM and XGBoost did not perform as well as expected in certain sectors, suggesting the need for further tuning or adaptation to specific features of the financial data. Future research could explore the integration of additional features or alternative modeling approaches to improve predictive performance. Additionally, expanding the dataset to include more companies or a longer time period could provide a more comprehensive assessment of model effectiveness across different market conditions.

Conclusion

The study investigates the utilization of diverse machine learning models to forecast financial performance across various sectors using data from 21 companies listed on the BIST100 index spanning from 2013 to 2023. The models under evaluation encompass Linear Regression, Ridge Regression, Lasso Regression, Decision Trees, Bagging, Random Forests, AdaBoost, Gradient Boosting (GBM), LightGBM, and XGBoost. Model performance was evaluated using metrics such as MSE, RMSE, MAE, MAPE, and rRMSE.

The results reveal notable variations in model performance, both globally and within specific sectors. Gradient Boosting (GBM) emerged as the top performer, exhibiting the lowest MSE and RMSE values, along with high accuracy and precision in predictions. Bagging and Random Forests also displayed robust performance, highlighting the efficacy of ensemble methods in enhancing prediction accuracy and stability. Conversely, linear models struggled to capture the intricate relationships within the financial data, resulting in elevated error metrics.

The sector-specific analysis unveiled consistent strong performance of ensemble methods, particularly Bagging and Random Forests, across diverse industries. Furthermore, GBM demonstrated strong performance in sectors such as energy, chemistry and smart materials, and telecommunications. These findings underscore the criticality of selecting appropriate models based on sector-specific attributes to achieve precise financial forecasts.

The incorporation of SHAP (SHapley Additive exPlanations) values provided deeper insights into feature importance and interaction effects, thereby augmenting the interpretability of the machine learning models. The SHAP summary plot for the Random Forest model indicated that net profit and long-term liabilities were among the most influential features, significantly impacting the model's outcomes. This heightened level of interpretability is essential for stakeholders reliant on the predictions for informed decision-making, offering a clearer comprehension of the factors steering financial performance.

This study contributes to the burgeoning field of financial technology by showcasing the potential of machine learning models in enhancing the accuracy of financial forecasting. It also underscores the significance of model interpretability, particularly in pivotal domains such as finance, where understanding the underlying factors influencing predictions is imperative. The research's findings furnish valuable insights for investors, financial analysts, and policymakers, aiding them in making more informed decisions based on robust and transparent financial forecasts.

References

Abbas, A.T., Helmy, M.O., Al-Abduljabbar, A.A., Soliman, M.S., Hasan, A.S. & Elkaseer, A. (2023). Precision face milling of maraging steel 350: An experimental investigation and optimization using different machine learning techniques. Machines, 11(11), 1-20, https://doi.org/10.3390/machines11111001.

Abdulrazzak, A.Y., Mohammed, S.L., Al-Naji, A. & Chahl, J. (2024). Real-time jaundice detection in neonates based on machine learning models. BioMedInformatics, 4(1), 623-637, https://doi.org/10.3390/biomedinformatics4010034.

Afreen, M. (2020). Review paper on composite leading index creation for forecasting the Bangladeshi financial sector. International Journal of Finance & Banking Studies, 9(4), 23-32, https://doi.org/10.20525/ijfbs.v9i4.791.

Akinrinola, O., Addy, W.A., Ajayi-Nifise, A.O., Odeyemi, O. & Falaiye, T. (2024). Application of machine learning in tax prediction: A review with practical approaches. Global Journal of Engineering and Technology Advances, 18(2), 102-117, https://doi.org/10.30574/gjeta.2024.18.2.0028.

Avelar, E.A. & Jordão, R.V.D. (2024). The role of artificial intelligence in the decision-making process: A study on the financial analysis and movement forecasting of the world's largest stock exchanges. Management Decision, Preprint, 1-19, https://doi.org/10.1108/MD-09-2023-1625.

Azad, M., Chikalov, I., Hussain, S., Moshkov, M. & Zielosko, B. (2022). Greedy algorithms for decision trees with hypotheses. arXiv Preprint, https://arxiv.org/abs/2203.08848.

Ballings, M., Van den Poel, D., Hespeels, N. & Gryp, R. (2015). Evaluating multiple classifiers for stock price direction prediction. Expert Systems with Applications, 42(20), 7046-7056. https://doi.org/10.1016/j.eswa.2015.04.013.

Baptista, M.L., Goebel, K. & Henriques, E.M. (2022). Relation between prognostics predictor evaluation metrics and local interpretability SHAP values. Artificial Intelligence, 306, 1-22. https://doi.org/10.1016/j.artint.2021.103667.

Barauskaite, G. & Streimikiene, D. (2021). Corporate social responsibility and financial performance of companies: The puzzle of concepts, definitions and assessment methods. Corporate Social Responsibility and Environmental Management, 28(1), 278-287, https://doi.org/10.1002/csr.2033.

Barboza, F., Kimura, H. & Altman, E. (2017). Machine learning models and bankruptcy prediction. Expert Systems with Applications, 83, 405-417, https://doi.org/10.1016/j.eswa.2015.05.013.

Barnhizer, D. & Barnhizer, D. (2019). The Artificial Intelligence Contagion: Can Democracy Withstand the Imminent Transformation of Work, Wealth and the Social Order? SCB Distributors, West Rancho Dominguez.

Bhattacharya, A. (2022). Applied Machine Learning Explainability Techniques: Make ML models explainable and trustworthy for practical applications using LIME, SHAP, and more. Packt Publishing, Birmingham.

Breiman, L. (1996). Bagging predictors. Machine Learning, 24, 123-140.

Celen, B., Ozcelik, M.B., Turgut, F.M., Aras, C., Sivaraman, T., Kotak, Y., Geisbauer, C. & Schweiger, H.G. (2022). Calendar ageing modelling using machine learning: An experimental investigation on lithium-ion battery chemistries. Open Research Europe, 2(96), 1-24, https://doi.org/10.12688/openreseurope.14745.2.

Chen, C.P. & Zhang, C.Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on big data. Information Sciences, 275, 314-347, https://doi.org/10.1016/j.ins.2014.01.015.

Chen, T. & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794), New York.

Claveria, O., Monte, E. & Torra, S. (2016). Combination forecasts of tourism demand with machine learning models. Applied Economics Letters, 23(6), 428-431, https://doi.org/10.1080/13504851.2015.1078441.

Davis, J., Devos, L., Reyners, S. & Schoutens, W. (2020). Gradient boosting for quantitative finance. Journal of Computational Finance, 24(4), 1-30.

Deng, Y.H., Luo, X.Q., Yan, P., Zhang, N.Y., Liu, Y. & Duan, S.B. (2022). Outcome prediction for acute kidney injury among hospitalized children via eXtreme Gradient Boosting Algorithm. Scientific Reports, 12(1), 1-11.

Dong, Z., Wang, Q., Ke, Y., Zhang, W., Hong, Q., Liu, C., Liu, X., Yang, J., Xi, Y., Shi, J., Zhang, L., Zheng, Y., Lv, Q., Wang, Y., Wu, J., Sun, X., Cai, G., Qiao, S., Yin, C., Su, S. & Chen, X. (2022). Prediction of 3-year risk of diabetic kidney disease using machine learning based on electronic medical records. Journal of Translational Medicine, 20(1), 1-10, https://doi.org/10.1186/s12967-022-03339-1.

Edafetanure-Ibeh, F.T. (2024). Evaluating machine learning algorithms for cervical cancer prediction: A comparative analysis, Preprint, doi: 10.1109/ACCESS.2024.3469869.

Ekanayake, I.U., Meddage, D.P.P. & Rathnayake, U. (2022). A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP). Case Studies in Construction Materials, 16, 1-20, https://doi.org/10.1016/j.cscm.2022.e01059.

El Bouchefry, K. & de Souza, R.S. (2020). Learning in big data: Introduction to machine learning. In: Knowledge discovery in big data from astronomy and earth observation (pp. 225-249). Elsevier, Amsterdam.

Fan, J. (2023). Predicting credit default by SVM and decision tree model based on credit card data. BCP Business & Management, 38, 28-33.

Fraz, N. (2024). A study on comparison of various machine learning models for the best prediction of 305 days first lactation milk yield, Research Square, 1-16, https://doi.org/10.21203/rs.3.rs-4484720/v1.

Friedman, J., Hastie, T. & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, 33(1), 1-22, https://pmc.ncbi.nlm.nih.gov/articles/PMC2929880/.

Geng, R., Bose, I. & Chen, X. (2015). Prediction of financial distress: An empirical study of listed Chinese companies using data mining. European Journal of Operational Research, 241(1), 236-247.

Genuer, R. (2012). Variance reduction in purely random forests. Journal of Nonparametric Statistics, 24(3), 543-562.

George, A.S. (2024). Finance 4.0: The transformation of financial services in the digital age, Partners Universal Innovative Research Publication, 2(3), 104-125.

Gianola, D., Weigel, K.A., Krämer, N., Stella, A. & Schön, C.C. (2014). Enhancing genome-enabled prediction by bagging genomic BLUP. PLoS One, 9(4), 1-18, https://doi.org/10.1371/journal.pone.0091693.

Grissa, D., Nytoft Rasmussen, D., Krag, A., Brunak, S. & Juhl Jensen, L. (2020). Alcoholic liver disease: A registry view on comorbidities and disease prediction. PLoS Computational Biology, 16(9), 1-19.

Gupta, A., Sharma, A. & Goel, A. (2017). Review of regression analysis models. International Journal of Engineering Research & Technology, 6(8), 58-61.

Gzar, D.A., Mahmood, A.M. & Abbas, M.K. (2022). A comparative study of regression machine learning algorithms: Tradeoff between accuracy and computational complexity. Mathematical Modelling of Engineering Problems, 9(5), 1-8, https://doi.org/10.18280/mmep.090508.

Hashemi, S.K., Mirtaheri, S.L. & Greco, S. (2022). Fraud detection in banking data by machine learning techniques. IEEE Access, 11, 3034-3043, https://doi.org/10.1109/ACCESS.2022.3232287.

Ionescu, S.A. & Diaconita, V. (2023). Transforming financial decision-making: The interplay of AI, cloud computing, and advanced data management technologies. International Journal of Computers Communications & Control, 18(6), 1-9, https://doi.org/10.15837/ijccc.2023.6.5735.

Jalal Uddin, M., Li, Y., Abdus Sattar, M. & Mistry, S. (2022). Climatic water balance forecasting with machine learning and deep learning models over Bangladesh. International Journal of Climatology, 42(16), 10083-10106.

Jiang, X., Zhou, R., Jiang, F., Yan, Y., Zhang, Z. & Wang, J. (2024). Construction of diagnostic models for the progression of hepatocellular carcinoma using machine learning. Frontiers in Oncology, 14, 1-11.

Johnson, R. & Zhang, T. (2013). Learning nonlinear functions using regularized greedy forest. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(5), 942-954, https://doi.org/10.1109/TPAMI.2013.159.

Kadiyala, A. & Kumar, A. (2018). Applications of Python to evaluate the performance of decision tree-based boosting algorithms. Environmental Progress & Sustainable Energy, 37(2), 618-623.

Kareem, M.K., Aborisade, O.D., Onashoga, S.A., Sutikno, T. & Olayiwola, O.M. (2023). Efficient model for detecting application layer distributed denial of service attacks. Bulletin of Electrical Engineering and Informatics, 12(1), 441-450, https://doi.org/10.11591/eei.v12i1.3871.

Katal, A., Wazid, M. & Goudar, R.H. (2013). Big data: Issues, challenges, tools, and good practices. In 2013 Sixth International Conference on Contemporary Computing (IC3) (pp. 404-409). IEEE, New Jersey.

Khalaf, G. (2012). A proposed ridge parameter to improve the least square estimator. Journal of Modern Applied Statistical Methods, 11(2), 443-449, https://doi.org/10.22237/jmasm/1351743240.

Kibria, B.G. & Saleh, A.M.E. (2004). Performance of positive rule estimator in the ill-conditioned Gaussian regression model. Calcutta Statistical Association Bulletin, 55(4), 209-240.

Ko, P.C., Lin, P.C., Do, H.T. & Huang, Y.F. (2022). P2P lending default prediction based on AI and statistical models. Entropy, 24(6), 1-23, https://doi.org/10.3390/e24060801.

Kourtellis, N., Morales, G.D.F., Bifet, A. & Murdopo, A. (2016, December). VHT: Vertical Hoeffding Tree. In 2016 IEEE International Conference on Big Data (Big Data) (pp. 915-922). IEEE, New Jersey.

Kulkarni, V.Y. & Sinha, P. K. (2012, July). Pruning of random forest classifiers: A survey and future directions. In 2012 International Conference on Data Science & Engineering (ICDSE) (pp. 64-68). IEEE, New Jersey.

Kumar, B., Sharma, M., Bhat, A. & Kumar, P. (2021). An analysis of Indian agricultural workers: A ridge regression approach. Agricultural Economics Research Review, 34(1), 121-127.

Kumar, R. (2017). Machine learning and cognition in enterprises: Business intelligence transformed. Apress, New York, https://doi.org/10.5958/0974-0279.2021.00010.0.

Li, R., Shinde, A., Liu, A., Glaser, S., Lyou, Y., Yuh, B., Wong, J. & Amini, A. (2020). Machine learning–based interpretation and visualization of nonlinear interactions in prostate cancer survival. JCO Clinical Cancer Informatics, 4, 637-646, https://doi.org/10.1200/CCI.20.00002.

Li, S., Qin, J., He, M. & Paoli, R. (2020). Fast evaluation of aircraft icing severity using machine learning based on XGBoost. Aerospace, 7(4), 1-18, https://doi.org/10.3390/aerospace7040036.

Li, Z. (2022). Extracting spatial effects from machine learning models using local interpretation methods: An example of SHAP and XGBoost, Computers, Environment and Urban Systems, 96, 1-18.

Li, Z. (2024). Evaluation of sailing boat performance based on ridge regression and mathematical model optimization. Highlights in Science, Engineering and Technology, 85, 1275-1283.

Lin, T.C. (2019). Artificial intelligence, finance, and the law. Fordham Law Review, 88, 531-560.

Lo, W.T., Chang, Y.S., Sheu, R.K., Chiu, C.C. & Yuan, S.M. (2014). CUDT: A CUDA-based decision tree algorithm. The Scientific World Journal, 2014(1), 1-12, https://doi.org/10.1155/2014/745640.

Long, X., Kampouridis, M. & Jarchi, D. (2022). An in-depth investigation of genetic programming and nine other machine learning algorithms in a financial forecasting problem. In: 2022 IEEE Congress on Evolutionary Computation (CEC) (pp. 01-08). IEEE, New Jersey.

Lundberg, S.M., Erion, G.G. & Lee, S.I. (2018). Consistent individualized feature attribution for tree ensembles. arXiv Preprint 3, https://doi.org/10.48550/arXiv.1802.03888.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. & Lee, S.I. (2020). From local explanations to global understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56-67, https://doi.org/10.1038/s42256-019-0138-9.

Magnan, M., Menini, A. & Parbonetti, A. (2015). Fair value accounting: Information or confusion for financial markets? Review of Accounting Studies, 20, 559-591, https://doi.org/10.1007/s11142-014-9306-7.

Mahalakshmi, V., Kulkarni, N., Kumar, K.P., Kumar, K.S., Sree, D.N. & Durga, S. (2022). The role of implementing artificial intelligence and machine learning technologies in the financial services industry for creating competitive intelligence. Materials Today: Proceedings, 56, 2252-2255.

Malthouse, E.C. (1999). Ridge regression and direct marketing scoring models. Journal of Interactive Marketing, 13(4), 10-23, https://doi.org/10.1002/(SICI)1520-6653(199923)13:4%3C10::AID-DIR2%3E3.0.CO;2-3.

Martini, M.L., Neifert, S.N., Shuman, W.H., Chapman, E.K., Schüpper, A.J., Oermann, E.K., Mocco, J., Todd, M., Torner, C.J., Molyeux, A., Mayer, S., Le Roux, P., Vergouwen, M.D.I., Rinkel, G.J.E., Wong, G.K.C., Kirkpatrick, P., Quinn, A., Hänggi, D., Etminan, N., van der Bergh, W.M., Jaja, B.N.R., Cusimano, M., Shweizer, A.T., Suarez, J.I., Fukuda, H., Yamagata, S., Lo, B., Airton, L.O.M., Boogarts, H.D. & Macdonald, R.L. (2021). Rescue therapy for vasospasm following aneurysmal subarachnoid hemorrhage: A propensity score-matched analysis with machine learning. Journal of Neurosurgery, 136(1), 134-147.

Massei, G. (2023). Algorithmic trading: An overview and evaluation of its impact on financial markets. Finance Research Letters, 47, 1-101.

Mayr, A. & Schmid, M. (2014). Boosting the concordance index for survival data–A unified framework to derive and evaluate biomarker combinations. PLoS One, 9(1), 1-10, https://doi.org/10.1371/journal.pone.0084483.

Meir, R. & Rätsch, G. (2003). An introduction to boosting and leveraging. In: Advanced Lectures on Machine Learning: Machine Learning Summer School (pp. 118-183). Springer, Berlin.

Mienye, I.D. & Sun, Y. (2022). A survey of ensemble learning: Concepts, algorithms, applications, and prospects. IEEE Access, 10, 99129-99149, https://doi.org/10.1371/10.1109/ACCESS.2022.3207287.

Mishina, Y., Murata, R., Yamauchi, Y., Yamashita, T. & Fujiyoshi, H. (2015). Boosted random forest. IEICE Transactions on Information and Systems, 98(9), 1630-1636, https://doi.org/10.1587/transinf.2014OPP0004.

Mitchell, R., Frank, E. & Holmes, G. (2022). GPUTreeShap: Massively parallel exact calculation of SHAP scores for tree ensembles. PeerJ Computer Science, 8, 1-25, https://doi.org/10.7717/peerj-cs.880.

Mohammad, O.K.J., Seno, M.E. & Dhannoon, B.N. (2024). Detailed Cloud Linear Regression Services in Cloud Computing Environment. Informatica, 48(12), 1-10.

Moshkov, M. (1997). Algorithms for constructing of decision trees. In: Principles of Data Mining and Knowledge Discovery: First European Symposium (pp. 335-342). Springer, Berlin.

Najafabadi, M.M., Villanustre, F., Khoshgoftaar, T.M., Seliya, N., Wald, R. & Muharemagic, E. (2015). Deep learning applications and challenges in big data analytics. Journal of Big Data, 2(1), 1-21.

Natekin, A. & Knoll, A. (2013). Gradient boosting machines, a tutorial. Frontiers in Neurorobotics, 7, 21.

Nguyen, D.K., Sermpinis, G. & Stasinakis, C. (2023). Big data, artificial intelligence and machine learning: A transformative symbiosis in favor of financial technology. European Financial Management, 29(2), 517-548.

Oliva, R. & Watson, N. (2009). Managing functional biases in organizational forecasts: A case study of consensus forecasting in supply chain planning. Production and Operations Management, 18(2), 138-151.

Onoja, M., Jegede, A., Blamah, N., Abimbola, O.V. & Omotehinwa, T.O. (2022). EEMDS: Efficient and effective malware detection system with hybrid model based on XceptionCNN and LightGBM algorithm. Journal of Computing and Social Informatics, 1(2), 42-57, https://doi.org/10.33736/jcsi.4739.2022.

Orsini, N., Moore, A. & Wolk, A. (2022). Interaction analysis based on Shapley values and extreme gradient boosting: A realistic simulation and application to a large epidemiological prospective study. Frontiers in Nutrition, 9, 1-8.

Pan, B. (2018, February). Application of XGBoost algorithm in hourly PM2.5 concentration prediction. In: IOP Conference Series: Earth and Environmental Science (Vol. 113, p. 012127). IOP Publishing, Bristol.

Pandey, M.K. & Sergeeva, I. (2022). Artificial intelligence impact evaluation: Transforming paradigms in financial institutions. Mir Èkonomiki I Upravleniâ, 22(1), 147-164, https://doi.org/10.25205/2542-0429-2022-22-1-147-164.

Park, S., Son, S., Bae, J., Lee, D., Kim, J.J. & Kim, J. (2021). Robust spatiotemporal estimation of PM concentrations using boosting-based ensemble models. Sustainability, 13(24), 1-15.

Paul, B., Athithan, G. & Murty, M.N. (2009). Speeding up AdaBoost classifier with random projection. In: 2009 Seventh International Conference on Advances in Pattern Recognition (pp. 251-254), IEEE, New Jersey.

Penman, S. H. (2002). The quality of financial statements: Perspectives from the recent stock market bubble. Papers SSRN 319262, 1-44, http://dx.doi.org/10.2139/ssrn.319262.

Permana, S., Rosadi, R. & Nikki, N. (2022). Application of classification algorithm for sales prediction. TEKNOKOM, 5(2), 119-124, https://doi.org/10.31943/teknokom.v5i2.77.

Perrini, F., Russo, A., Tencati, A. & Vurro, C. (2011). Deconstructing the relationship between corporate social and financial performance. Journal of Business Ethics, 102, 59-76, https://doi.org/10.1007/s10551-011-1194-1.

Poojithaa, M. & Malathib, K. (2022). Decision tree over support vector machine for better accuracy in identifying the problem based on the Iris flower. Advances in Parallel Computing Algorithms, Tools and Paradigms, 41, 209-217.

Prasad, A. & Bakhshi, P. (2022). Forecasting the direction of daily changes in the India VIX index using machine learning. Journal of Risk and Financial Management, 15(12), 1-16, https://doi.org/10.3390/jrfm15120552.

Provost, F. & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. Big Data, 1(1), 51-59, https://doi.org/10.1089/big.2013.1508.

Rajaratnam, B., Roberts, S., Sparks, D. & Dalal, O. (2015). Lasso regression: Estimation and shrinkage via the limit of Gibbs sampling. Journal of the Royal Statistical Society: Series B Statistical Methodology, 78(1), 153-174.

Ramnath, S., Rock, S. & Shane, P. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. International Journal of Forecasting, 24(1), 34-75.

Rohatgi, S., Singh, K.K. & Jasuja, D. (2021). Comparative analysis of machine learning algorithm to forecast Indian stock market. In 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE) (pp. 278-283). IEEE, New Jersey, https://doi.org/10.1109/ICACITE51222.2021.9404642.

Rosenbusch, H., Soldner, F., Evans, A.M. & Zeelenberg, M. (2021). Supervised machine learning methods in psychology: A practical introduction with annotated R code. Social and Personality Psychology Compass, 15(2), 1-25.

Rufo, D., Debelee, T., Ibenthal, A. & Negera, W. (2021). Diagnosis of diabetes mellitus using gradient boosting machine (LightGBM). Diagnostics, 11(9), 1-14, https://doi.org/10.3390/diagnostics11091714

Ryll, L. & Seidens, S. (2019). Evaluating the performance of machine learning algorithms in financial market forecasting: A comprehensive survey, arXiv preprint, 1906, https://doi.org/10.48550/arXiv.1906.07786.

Sairam, S., Srinivasan, S., Marafioti, G., Subathra, B., Mathisen, G. & Bekiroglu, K. (2020). Explainable Incipient Fault Detection Systems for Photovoltaic Panels. arXiv preprint, 2011, https://doi.org/10.48550/arXiv.2011.09843.

Samonas, M. (2015). Financial forecasting, analysis, and modelling: A framework for long-term forecasting. John Wiley & Sons, Hoboken.

Sandhya, V. & Padyana, A. (2021). Machine learning based crop yield prediction on geographical and climatic data. In: 2021 Sixth International Conference on Image Information Processing (ICIIP) (pp. 186-191). IEEE, New Jersey.

Sastry, V.V.L.N. (2020). Artificial intelligence in financial services and banking industry. Idea Publishing, London.

Shi, F., Lu, S., Gu, J., Lin, J., Zhao, C., You, X. & Lin, X. (2022). Modeling and evaluation of the permeate flux in forward osmosis process with machine learning. Industrial & Engineering Chemistry Research, 61(49), 18045-18056.

Si, Z., Niu, H. & Wang, W. (2022). Credit Risk Assessment by a Comparison Application of Two Boosting Algorithms. In: Fuzzy Systems and Data Mining VIII (pp. 34-40). IOS Press, Amsterdam.

Signorino, C. & Kirchner, A. (2018). Using lasso to model interactions and nonlinearities in survey data. Survey Practice, 11(1), 1-10, https://doi.org/10.29115/SP-2018-0005.

Siringoringo, R., Perangin, R. & Jamaluddin, J. (2021). Model hibrid genetic-XGBoost dan principal component analysis pada segmentasi dan peramalan pasar. Methomika Jurnal Manajemen Informatika Dan Komputerisasi Akuntansi, 5(2), 97-103, https://doi.org/10.46880/jmika.Vol5No2.pp97-103.

Soloff, J. A., Barber, R.F. & Willett, R. (2024). Bagging provides assumption-free stability. Journal of Machine Learning Research, 25(131), 1-35.

Sonkavde, G. (2023). Forecasting stock market prices using machine learning and deep learning models: A systematic review, performance analysis and discussion of implications. International Journal of Financial Studies, 11(3), 1-22, https://doi.org/10.3390/ijfs11030094.

Strobl, C., Boulesteix, A., Kneib, T., Augustin, T. & Zeileis, A. (2008). Conditional variable importance for random forests. BMC Bioinformatics, 9(1), 1-11.

Su, H., Lu, X., Chen, Z., Zhang, H., Lu, W. & Wu, W. (2021). Estimating coastal chlorophyll-a concentration from time-series OLCI data based on machine learning. Remote Sensing, 13(4), 1-21.

Suacana, I. (2024). Optimizing the 2024 governor election quick count with extreme gradient boosting (XGBoost) to increase voting prediction accuracy. International Journal of Software Engineering and Computer Science, 4(1), 91-106, https://doi.org/10.35870/ijsecs.v4i1.2286.

Tang, M., Zhao, Q., Ding, S., Wu, H., Li, L., Wen, L. & Huang, B. (2020). An improved LightGBM algorithm for online fault detection of wind turbine gearboxes. Energies, 13(4), 807-823.

Uddin, M.N., Li, L.Z., Deng, B.Y. & Ye, J. (2023). Interpretable XGBoost–SHAP machine learning technique to predict the compressive strength of environment-friendly rice husk ash concrete. Innovative Infrastructure Solutions, 8(5), 147-168, https://doi.org/10.1007/s41062-023-01122-9.

Ünal, A. F., Kaleli, A. Y., Ummak, E. & Albayrak, Ö. (2021, August). A Comparison of State-of-the-Art Machine Learning Algorithms on Fault Indication and Remaining Useful Life Determination by Telemetry Data. In: 8th International Conference on Future Internet of Things and Cloud (pp. 79-85). IEEE, New Jersey.

Wang, L., Kern, R., Yu, E., Choi, S. & Pan, J. (2023). IntelliSleepScorer, a software package with a graphic user interface for automated sleep stage scoring in mice based on a light gradient boosting machine algorithm. Scientific Reports, 13(1), 1-11, https://doi.org/10.1038/s41598-023-31288-2.

Wang, M. (2024). Identification of mine water source based on TPE-LightGBM. Scientific Reports, 14(1), 1-11.

Wang, P., Xie, M., Wang, X., Yu, J., Chen, E., Zhou, Z., Niu, Y., Song, W., Ni, Q. & Zhu, J. (2022). Comparison of nomogram with machine learning techniques for prediction of overall survival in patients with retroperitoneal liposarcoma, Research Square, 1, 1-20.

Wang, S., Pengfei, D. & Tian, Y. (2017). A novel method of statistical line loss estimation for distribution feeders based on feeder cluster and modified XGBoost. Energies, 10(12), 1-17.

Wei, C. (2024). Comparison of different machine learning classification models for predicting deep vein thrombosis in lower extremity fractures. Scientific Reports, 14(1), 1-8.

Wei, L., Zhang, Y., Wang, Z., Zhao, L., Zhang, Y., Lu, X. & Cao, L. (2020). Hyperspectral inversion of soil organic matter content based on a combined spectral index model. Sensors, 20(10), 1-17.

Wu, Z., Lei, T., Shen, C., Wang, Z., Cao, D. & Hou, T. (2019). ADMET evaluation in drug discovery. 19. Reliable prediction of human cytochrome P450 inhibition using artificial intelligence approaches. Journal of Chemical Information and Modeling, 59(11), 4587-4601, https://doi.org/10.1021/acs.jcim.9b00801.

Xiang, Y. (2024). Enhancing non-invasive colorectal cancer screening with stool DNA methylation markers and LightGBM machine learning, Research Square, 1, 1-19, https://doi.org/10.21203/rs.3.rs-3857174/v1.

Xiao, D., Chen, J., Zhang, K. & Qian, H. (2020). Privacy-preserving locally weighted linear regression over encrypted millions of data. IEEE Access, 8, 2247-2257, https://doi.org/10.1109/ACCESS.2019.2962700.

Xin, S. & Khalid, K. (2018). Modelling house price using ridge regression and lasso regression. International Journal of Engineering & Technology, 7(4), 498-501.

Yao, S., Wu, Q., Kang, Q., Chen, Y. & Yi, L. (2023). An interpretable XGBoost-based approach for Arctic navigation risk assessment. Risk Analysis, 44(2), 459-476, https://doi.org/10.1111/risa.14175.

Yin, S., Ouyang, P., Xu, D., Liu, L. & Wei, S. (2017). An AdaBoost-based face detection system using parallel configurable architecture with optimized computation. IEEE Systems Journal, 11(1), 260-271.

Yoo, H., Lee, K., Woo, J., Park, S., Lee, S., Joo, J., Bae, J.-S., Hwong, H.-J. & Park, B. (2022). A genome-wide association study and machine-learning algorithm analysis on the prediction of facial phenotypes by genotypes in Korean women. Clinical, Cosmetic and Investigational Dermatology, 15, 433-445.

Zern, A., Broelemann, K. & Kasneci, G. (2023). Interventional SHAP values and interaction values for piecewise linear regression trees. Proceedings of the AAAI Conference on Artificial Intelligence, 37(9), 11164-11173.

Zhan, C., Zheng, Y., Zhang, H. & Wen, Q. (2021). Random-Forest-Bagging Broad Learning System with applications for COVID-19 pandemic. IEEE Internet of Things Journal, 8, 15906-15918.

Zhang, B., Sethy, A., Sainath, T.N. & Ramabhadran, B. (2011). Application specific loss minimization using gradient boosting. In: IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 4880-4883). IEEE, New Jersey, https://doi.org/10.1109/ICASSP.2011.5947449.

Zhang, J. (2024). Optimization and application of XGBoost logging prediction model for porosity and permeability based on k-means method. Applied Sciences, 14(10), 1-18, https://doi.org/10.3390/app14103956.

Zhang, J. (2024). Prediction of compressive strength of geopolymer concrete landscape design: Application of the novel hybrid RF-GWO-XGBoost algorithm. Buildings, 14(3), 1-32, https://doi.org/10.3390/buildings14030591.

Zhang, J., Mucs, D., Norinder, U. & Svensson, F. (2019). LightGBM: An effective and scalable algorithm for prediction of chemical toxicity-application to the Tox21 and mutagenicity data sets. Journal of Chemical Information and Modeling, 59(10), 4150-4158, https://doi.org/10.1021/acs.jcim.9b00633.

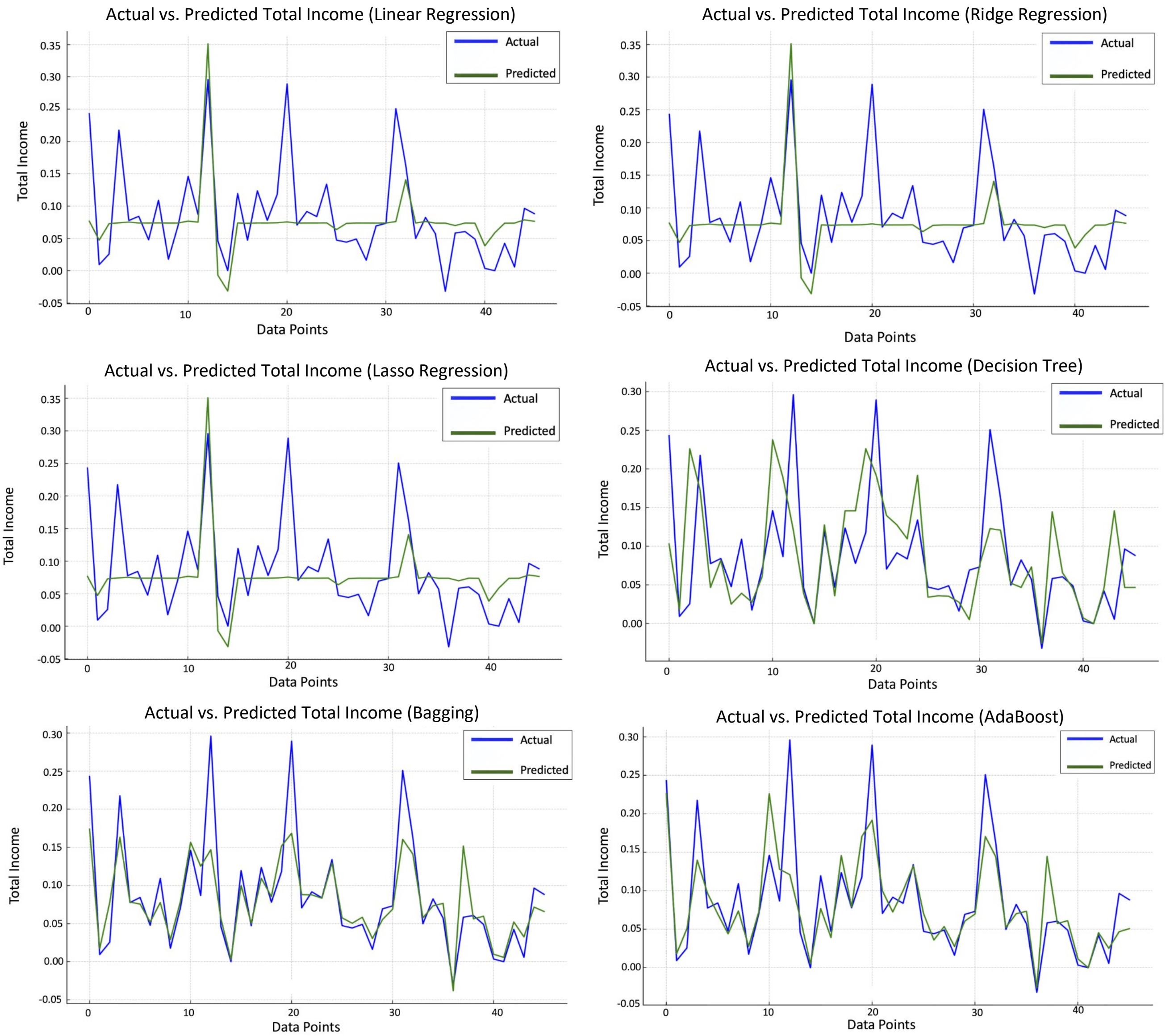
Zhang, P., Jia, Y. & Shang, Y. (2022). Research and application of XGBoost in imbalanced data. International Journal of Distributed Sensor Networks, 18(6), 1-10, https://doi.org/10.1177/15501329221106935.

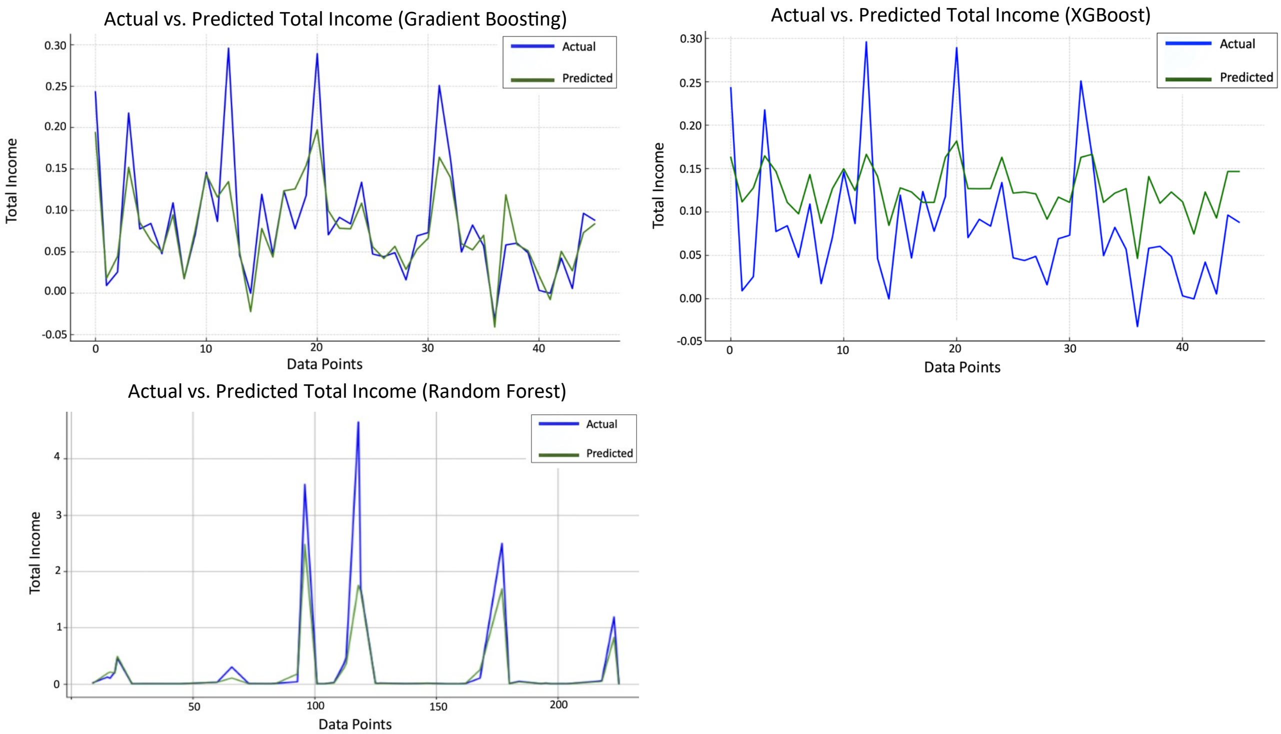
Zhao, G., Wang, Y. & Wang, J. (2023). Intrusion detection model of Internet of Things based on LightGBM. IEICE Transactions on Communications, 106(8), 622-634, https://doi.org/10.1587/transcom.2022EBP3169.

Zhou, L., Pan, S., Wang, J. & Vasilakos, A.V. (2017). Machine learning on big data: Opportunities and challenges. Neurocomputing, 237, 350-361, https://doi.org/10.1016/j.neucom.2017.01.026.

Appendix

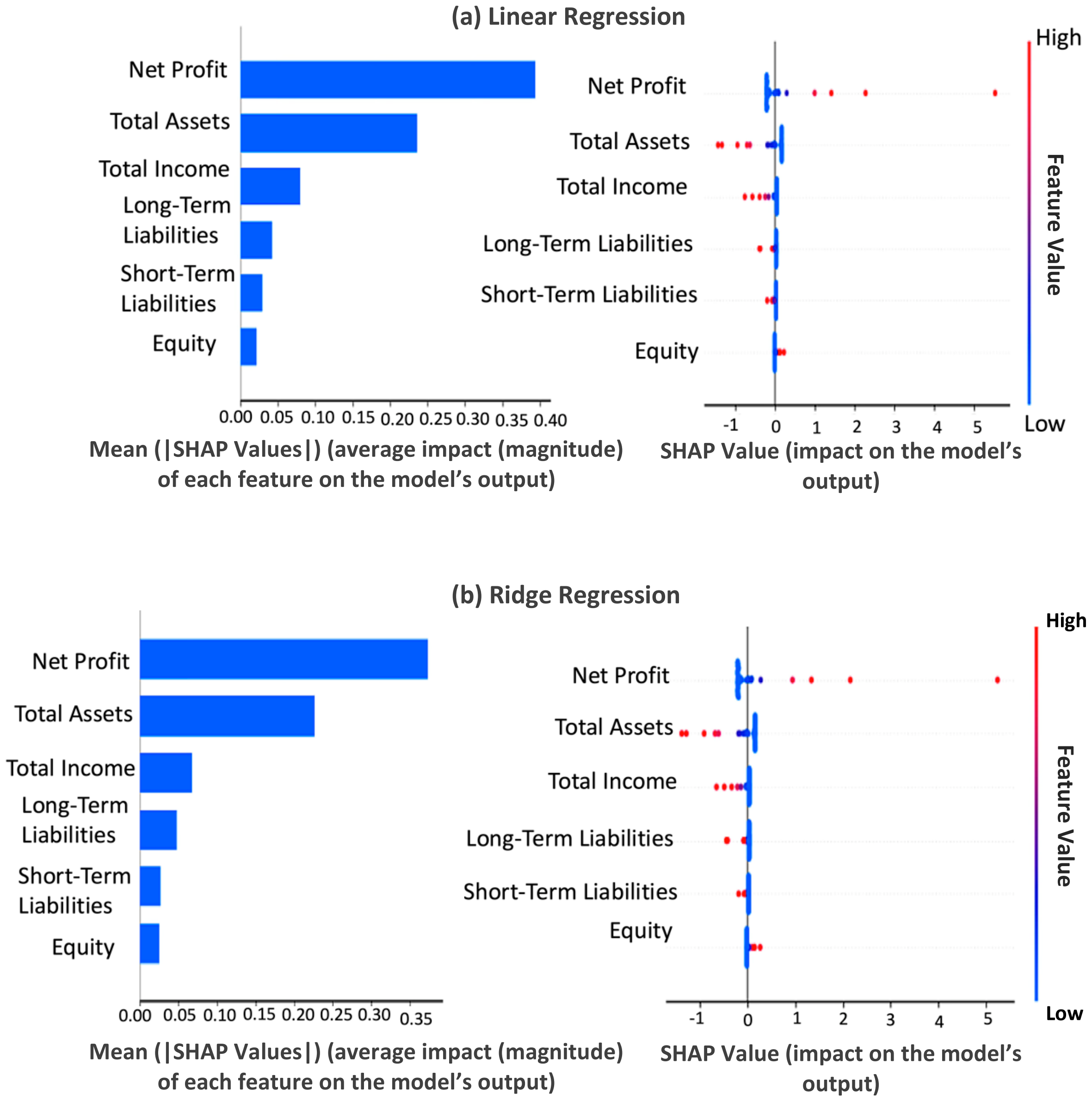
Appendix 1: Performances of machine learning models over test sample

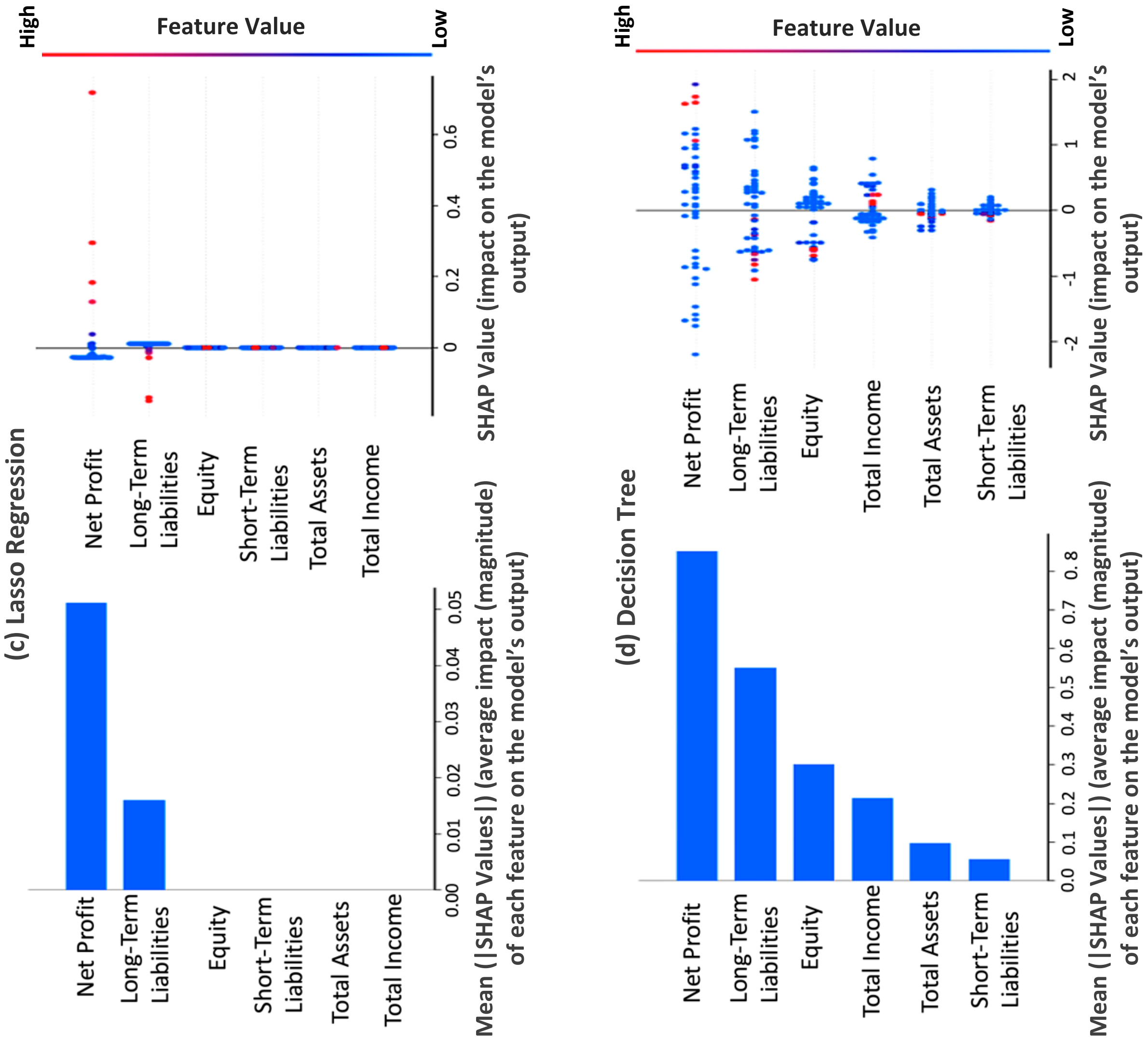


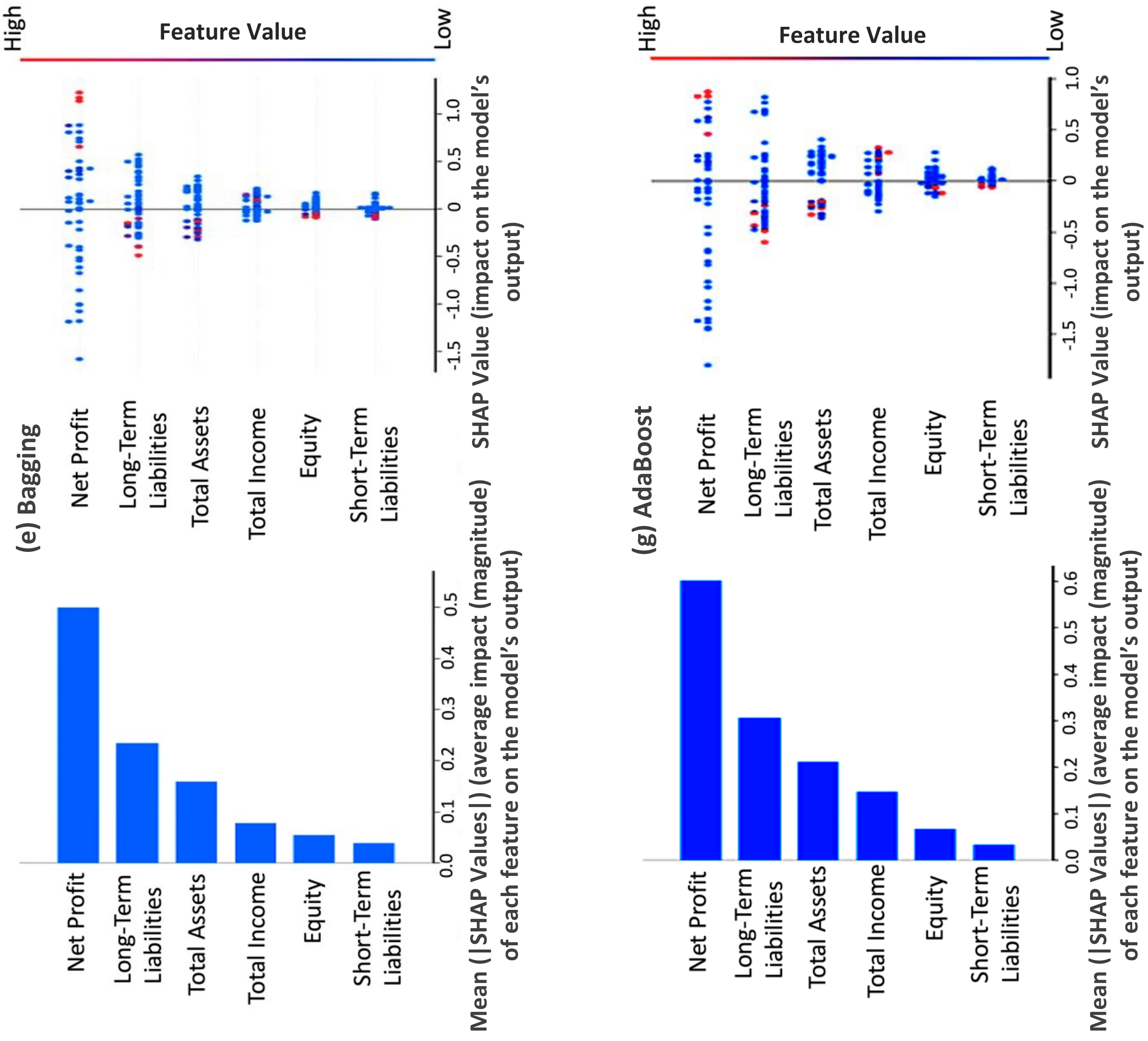


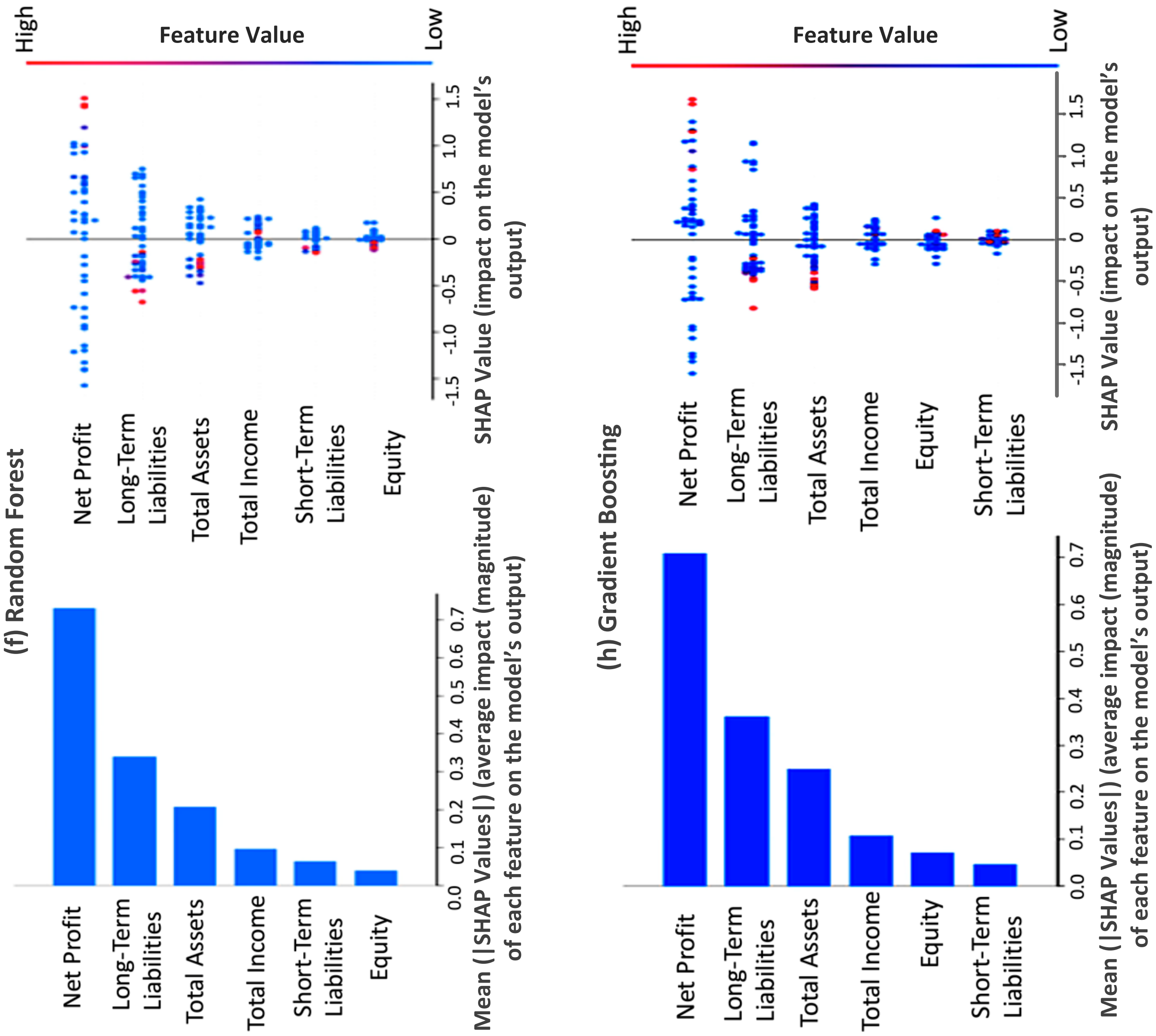
Source: Author's own work.

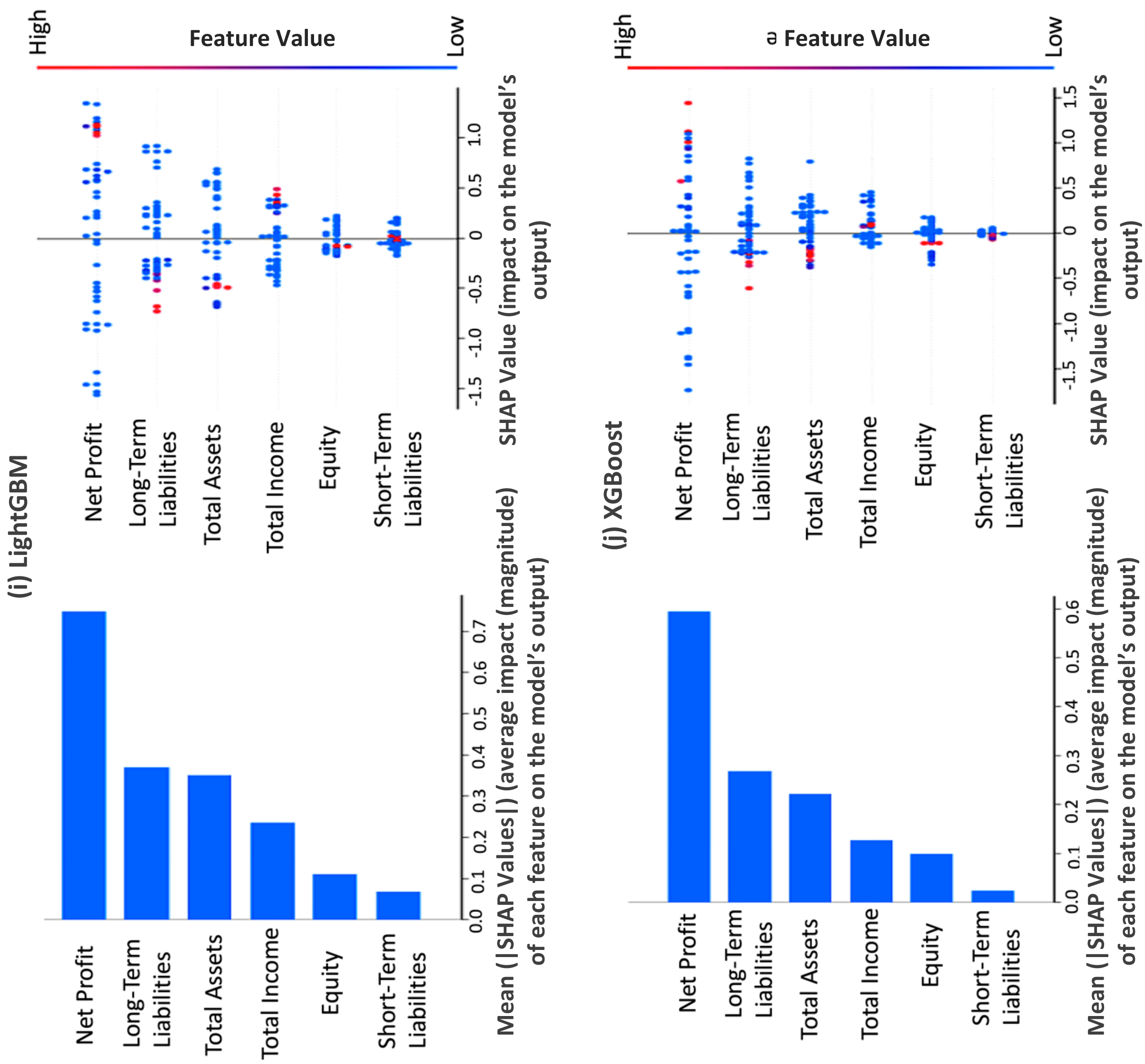
Appendix 2: SHAP feature importance and summary of the financial forecasting results for the selected machine learning models




(c) Lasso Regression
Net Profit
Long-Term Liabilities
Equity
Short-Term Liabilities
Total Assets
Total Income
Mean (|SHAP Values|) (average impact (magnitude) of each feature on the model's output)
SHAP Value (impact on the model's output)
Feature Value
High
Low
(d) Decision Tree
Net Profit
Long-Term Liabilities
Equity
Total Income
Total Assets
Short-Term Liabilities
Mean (|SHAP Values|) (average impact (magnitude) of each feature on the model's output)
SHAP Value (impact on the model's output)
Feature Value
High
Low

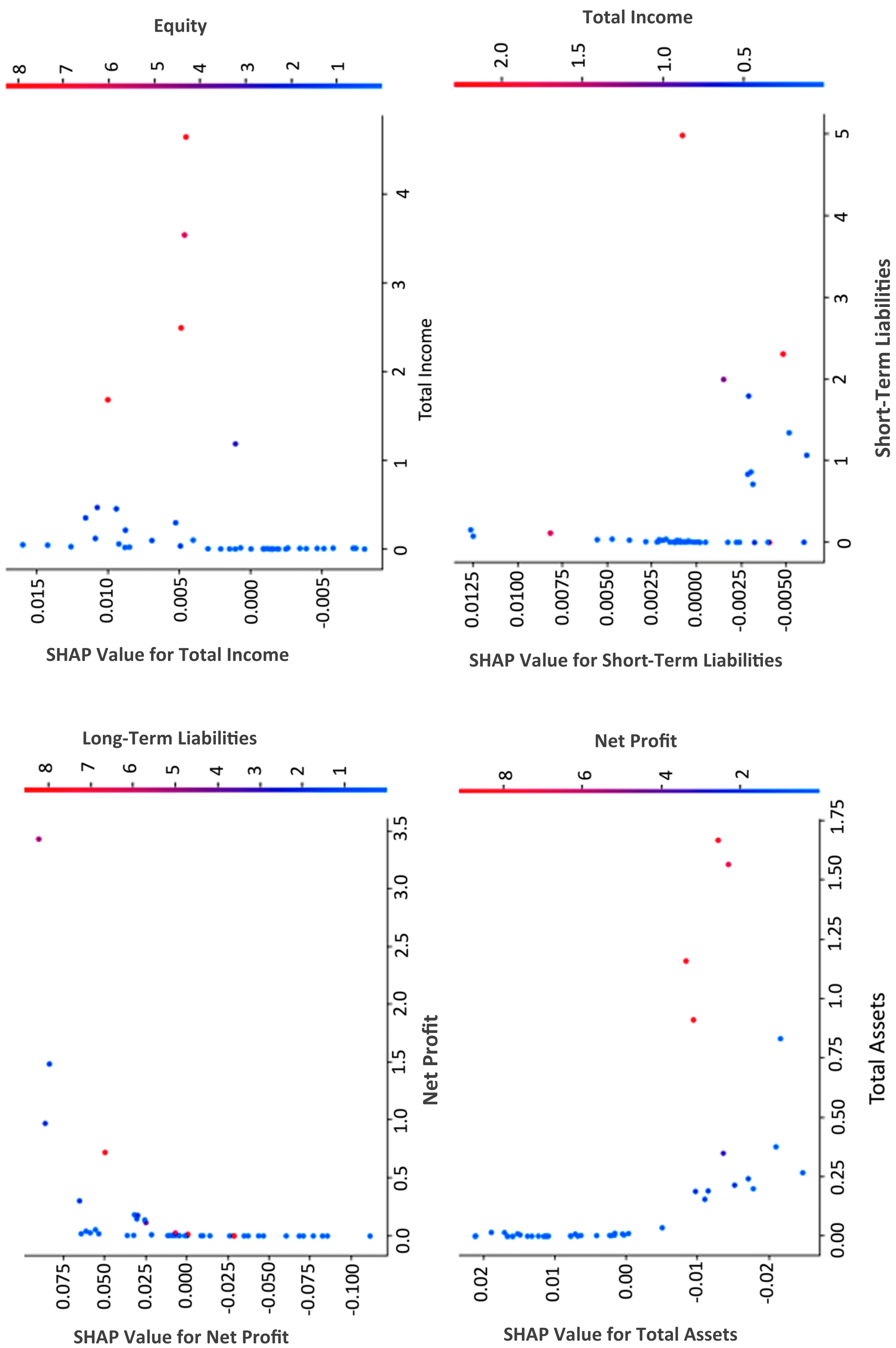

(e) Bagging
(g) AdaBoost
Net Profit
Long-Term Liabilities
Total Assets
Total Income
Equity
Short-Term Liabilities
Feature Value
High
Low
SHAP Value (impact on the model's output)
Mean (|SHAP Values|) (average impact (magnitude) of each feature on the model's output)

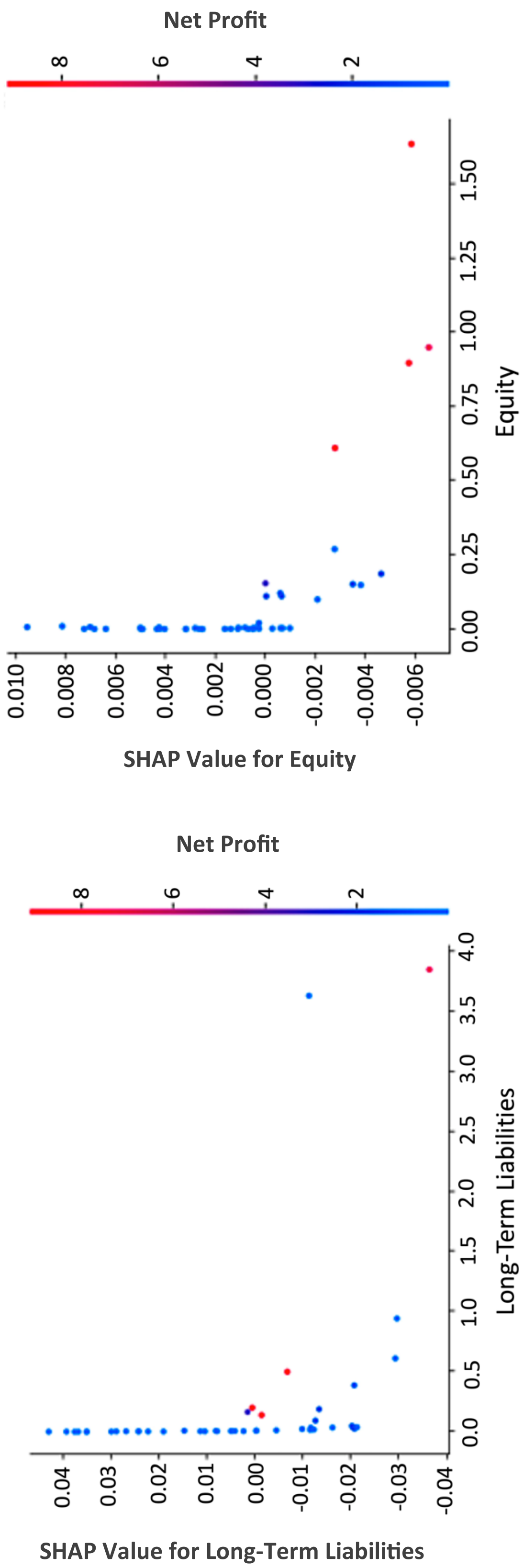

(f) Random Forest
(h) Gradient Boosting
Net Profit
Long-Term Liabilities
Total Assets
Total Income
Short-Term Liabilities
Equity
Mean (|SHAP Values|) (average impact (magnitude) of each feature on the model's output)
SHAP Value (impact on the model's output)
Feature Value
High
Low



Source: Author's own work.

Appendix 3: SHAP dependence plots for Gradient Boosting





Source: Author's own work.