Indonesian Journal of Sustainability Accounting and Management
ISSN 2597–6214 | e–ISSN 2597–6222
DOI: 10.20448/ijsam.v9i2.7942

# Carbon Emissions and Firm Profitability: A SHAP and ALE-Based XAI Approach on the BIST Sustainability Index

**Demet Ever[1]* | Ahmet Çankal[2]**

[1]*Osmaniye Korkut Ata University, Osmaniye Vocational High School, Department of Accounting and Tax, Osmaniye, Turkey*
[2]*Osmaniye Korkut Ata University, Faculty of Economics and Administrative Sciences, Management Information Systems, Osmaniye, Turkey*

*Correspondence to: Demet Ever, Osmaniye Korkut Ata University, Osmaniye Vocational High School, Department of Accounting and Tax, Osmaniye, *Turkey*.
E-mail: demetever@osmaniye.edu.tr

**Abstract:** The primary objective of this study is to examine the relationship between carbon emissions and firm profitability using Explainable Artificial Intelligence (XAI) methods. The research analyzes various machine learning (ML) techniques to predict firm profitability. Additionally, XAI methods such as Shapley Additive Explanations (SHAP) and Accumulated Local Effects (ALE) are employed to enhance interpretability for decision-makers. The study focuses on firms listed in the Borsa Istanbul Sustainability Index (XUSRD), utilizing 189 firm-year data points collected from sustainability, operational, and integrated reports between 2021 and 2023. The findings indicate that low carbon emissions positively influence Return on Assets (ROA), while high emissions have mixed effects, positive in some firms and negative in others. Regarding Return on Equity (ROE), the analysis reveals a negative trend. Furthermore, Earnings Per Share (EPS) emerged as the variable with the highest contribution to both profitability models. The Random Forest algorithm was identified as the most effective method for predicting both ROA and ROE. This study contributes to the emerging literature by applying XAI techniques, specifically SHAP and ALE, to interpret machine learning-based profitability models within the context of carbon accounting, offering valuable insights for stakeholders and policymakers.

**Keywords:** BIST Sustainability Index, carbon accounting, carbon emissions, firm profitability, machine learning, XAI, Shapley Additive Explanation, Accumulated Local Effects.

**Article info:** Received 31 October 2025 | revised 15 December 2025 | accepted 19 December 2025 | published 24 December 2025

**Recommended citation:** Ever, D., & Çankal, A. (2025). Carbon Emissions and Firm Profitability: A SHAP and ALE-Based XAI Approach on the BIST Sustainability Index. *Indonesian Journal of Sustainability Accounting and Management*, 9(2), 158–178. https://doi.org/10.20448/ijsam.v9i2.7942

## INTRODUCTION

One of the most critical steps in combating climate change is the reduction of carbon emissions (CE). In recent years, global awareness of climate change has led businesses to report environmental impacts more transparently. In this regard, CE has become one of the most important indicators in terms of both environmental and financial performance. Indeed, reducing CE has become a fundamental component of sustainability policies. The Kyoto Protocol, which came into force in 2005, has set the reduction of greenhouse gas (GHG) emissions as a necessary target for many countries and firms worldwide. According to (World

Meteorological Organization [WMO], 2024), worldwide GHG emissions are projected to rise to 41.6 billion tons of $CO_2$ by 2024. These developments have increased the importance of the net-zero emission goal adopted in the Paris Agreement. To limit global warming to 1.5°C, GHG emissions must be reduced by 45% by 2030, and net-zero emissions must be achieved by 2050. The implementation of this agreement is crucial for achieving sustainable development goals (United Nations, 2025). Countries with high CE, such as China, aim to reach peak emissions by 2030 and achieve carbon neutrality by 2060 (Liu et al., 2022). While international agreements such as the Kyoto Protocol and the Paris Agreement are among the most important tools for reducing GHG emissions, carbon accounting is one of the most vital mechanisms for ensuring the enforceability of these goals (Demircioğlu & Ever, 2020; Li et al., 2024). Carbon accounting enables firms to analyze environmental impacts in an integrated manner with financial information and serves as the foundation for a sustainability-focused corporate management approach (Chang, Lee, & Lee, 2024).

In Turkey, total greenhouse gas (GHG) emissions reached 598.9 million tons of $CO_2$ equivalent. Despite this, emissions per person were calculated at 7.0 tons of $CO_2$ (TUIK, 2025). This significant increase suggests that the relationship between carbon emissions (CE) and firm profitability (FP) warrants a more comprehensive examination. Accurate measurement, reporting, and management of CE by firms have become fundamental components of accounting practices aimed at combating climate change (Hassan, 2019; Imoniana, Soares, & Domingos, 2018; Tuesta, Naranjo, & Ripoll, 2021). Accounting and reporting for CE not only serve as indicators of environmental sensitivity but also act as vital tools that enhance transparency in strategic planning and decision-making processes (Li et al., 2024). Therefore, establishing a relationship between carbon accounting data, such as CE, and financial performance holds strategic importance for both environmental and financial sustainability.

The negative effects of global climate change and the emergence of solid carbon disclosure regulations are exerting pressure on the transformation, calculation, and reporting of carbon emissions beyond the operational limits of businesses (Onat, Mandouri, Kucukvar, Kutty, & Al-Muftah, 2025). In this context, a broad global trend has emerged toward sustainability-oriented performance measurement (Shui, Zhang, Wang, & Smart, 2025). These developments have made carbon emissions a critical determinant of corporate performance (Chang, 2025; Han & Wei, 2025; Kurnia, Agustia, Soewarno, & Ardianto, 2025; Onat et al., 2025). Simultaneously, international regulatory frameworks such as the EU Taxonomy and the Carbon Border Adjustment Mechanism (CBAM) have strengthened the strategic and financial importance of integrating carbon-related criteria into corporate reporting systems (Pan & Liu, 2024; Tonnarello, Vermiglio, Migliardo, & Naciti, 2025).

Recently, the relationship between CE and FP has been explained using traditional regression models. However, new approaches based on machine learning (ML) have been developed. Although these techniques provide high accuracy, they often raise questions in the field of accounting and finance, especially regarding the transparency of decision-making processes, as they are often considered a "black box" structure (Guidotti et al., 2018; Rudin, 2019). Therefore, to make the internal functioning of the model more understandable and to create significant value for users involved in decision-making, techniques of Explainable Artificial Intelligence (XAI) have been developed (Arrieta et al., 2020; Wu et al., 2025; Zhang, Cho, & Vasarhelyi, 2022).

Although the literature contains numerous studies on the relationship between CE and FP, there are relatively few investigations into how this relationship can be explained using XAI techniques. Traditional regression-based models are often insufficient for revealing complex and nonlinear relationships between variables. This limitation results in a lack of understanding among decision-makers involved in strategy development processes. However, XAI techniques, which make models interpretable and explainable, serve as valuable analytical tools for decision-makers in the business environment (Černevičienė & Kabašinskas, 2024; Holzinger, Saranti, Molnar, Biecek, & Samek, 2022; Lundberg et al., 2019; Zhang, Wu, Qu, & Chen, 2022). XAI is important in interdisciplinary contexts, such as accounting and sustainability management, especially because it provides clear explanations of accuracy and responsibility in financial studies.

The primary aim of this study is to elucidate the relationship between Explainable Artificial Intelligence (XAI) methods and CE and Financial Performance (FP). Specifically, the research evaluates the influence of environmental performance, as measured by CE, on profitability, aiming to present this relationship more transparently and understandably to decision-makers. The analysis is based on 189 firm-year observations, including data from the Borsa Istanbul Sustainability Index (XUSRD) for the period 2021-2023. In this study, various machine learning (ML) algorithms were employed to assess their predictive capabilities. These algorithms include Random Forest, ExtraTrees, XGBoost, LightGBM, Multi-Layer Perceptron (MLP), Support Vector Regression (SVR), and Linear Regression. To compare the predictive accuracy of these models, the Diebold-Mariano test was applied, utilizing Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) metrics. Furthermore, to enhance interpretability, the study utilized XAI methods such as Shapley Additive Explanations (SHAP) and Accumulated Local Effects (ALE). These techniques help ensure that the model outputs are comprehensible and interpretable by decision-makers. The variables incorporated in the models include both non-financial indicators, such as carbon emissions, and financial accounting indicators, including Return on Assets (ROA), Return on Equity (ROE), leverage, price-to-book value (PB), earnings per share (EPS), and firm size. The CE data were sourced from sustainability reports, annual reports, and integrated reports published by the enterprises, while financial indicators were obtained from the Finnet2000Plus database. Overall, this study aims to answer specific research questions related to the impact of environmental performance on financial outcomes and the effectiveness of different ML models and XAI techniques in explaining these relationships.

1. How does the CE role in firms listed on the XUSRD influence financial performance indicators such as Return on Assets (ROA) and Return on Equity (ROE)?
2. Which machine learning methods are the most successful in terms of false positive predictions?
3. How do XAI methods explain the relationship between the CE and FP of firms in the XUSRD?

This study contributes to the accounting literature by applying the XAI methods SHAP and ALE to explain the relationship between CE and financial performance (FP) for firms listed on the XUSRD in Turkey. XAI, which has emerged as an important technology in recent years, is rapidly gaining popularity, especially in the fields of finance and sustainability. It offers transparency in decision-making processes by providing interpretability of machine learning (ML) models. However, the integration of this technology into accounting research in Turkey remains limited. In this study, the application of XAI methods addresses the increasing need for transparency in financial analyses of CE and offers contextual explanations for the mixed findings in the literature regarding the relationship between CE and FP. This study is divided into five sections. Firstly, the aim and significance of the research are summarized in the introduction. Then, the literature review examines previous studies on the relationship between CE and FP. The methodology section explains the dataset, dependent and independent variables, machine learning algorithms, and XAI-based methods such as SHAP and ALE. The findings section compares the performance of the machine learning models and evaluates the effects of CE on ROA and ROE. Finally, the conclusion section addresses the research questions and provides recommendations regarding the limitations of the study and suggestions for future research.

## LITERATURE REVIEW

Awareness of sustainability and climate change issues is rapidly increasing globally. This awareness has made environmental performance (CE) measurement and reporting processes increasingly important. In this context, the relationship between CE and FP has been the subject of extensive discussions and diverse perspectives in the literature. Specifically, the impact of CE on FP has often been analyzed across different countries and using various indices. Although there are inconsistent findings on this relationship in the literature, both positive and negative effects have been observed.

Many studies have shown that CE firms negatively impact financial performance. For example, Lee, Min, and Yook (2015) found that, in a panel data analysis covering the years 2003-2010, the company value of CE remains stable. Similarly, Delmas, Nairn-Birch, and Lim (2016) analyzed 1,095 US companies and argued that environmental performance negatively affects indicators such as ROA in the short term. However, it could generate potential value through Tobin's Q in the long term. Miah, Hasan, and Usman (2021) found that firms of CE in developing economies have adversely affected accounting and market-based indicators such as ROE, credit note, Z-skoru and Tobin's Q. Desai, Raval, Baser, and Desai (2022) state that high Indian firms have a more negative impact on the financial performance of CE. Houqe, Opare, Zahir-ul-Hassan, and Ahmed (2022), using the data of US firms, found that CE and agent costs reduce market reactions by adversely affecting company performance. Laskar, Kulshrestha, Bahuguna, and Adichwal (2022) reported that carbon density is negative and has significant effects on financial performance. Koç and Deran (2024) from Turkiye show that carbon emissions affect the cost of borrowing, while Sakın and Kefe (2024) have negative effects on ROA and ROE. Güneysu and Atasel (2022) find that CE has a negative impact on EPS and ROA in non-funded BIST100 firms inTurkiye, but there is no significant relationship with other indicators.

However, some academic studies argue that firms can contribute positively to the reduction of CE in financial performance. Lee and Min (2015) revealed that green R&D investments in Japanese manufacturing firms reduce CE and increase financial performance. Gallego-Álvarez, Segura, and Martínez-Ferrero (2015) discovered that the reduction of CE has a positive effect on financial performance. Khatib et al. (2023) used panel data analysis obtained from ten European countries and found that management improved carbon release as a regulatory mechanism of environmental education and that this positively affected firm performance. Aruga, Chiba, and Goshima (2023) reported that low CE contributes to long-term performance and lower capital costs in Tokyo Exchange firms. Tong, Li, and Yang (2025) in China show that digital transformation improves enterprise carbon performance and is positively associated with financial performance. Nichita, Nechita, Manea, Irimescu, and Manea (2021) have analyzed the effects of CE on sales profitability in enterprises operating in the chemical industry in Central East Europe. The findings indicate that lower levels of CE are associated with increased sales profitability, which in turn positively impacts the financial performance related to environmental performance. Ergün (2024) found that CE can increase financial performance in a study on the manufacturing sector in Turkiye. Ferrat (2021) shows that carbon performance negatively affects corporate financial performance in the short term, but can positively affect it in the long term. In recent years, XAI-based methods such as SHAP and ALE have been widely used in research in finance, sustainability, and accounting. These tools offer valuable predictions to decision-makers by increasing the interpretation of model outputs. For example, Srivalli and Sumanthi (2025) highlighted the critical role of XAI-based financial risk models in providing trust and decision-making for stakeholders. Yildirim, Okay, and Özdemir (2024) analyzed the financial performance of firms operating in the energy sector in Turkiyewith three different ML algorithms and gained a significant contribution to XAI applications by providing the disclosure of models with the LIME method. Similarly, Çankal and Ever (2025) analyzed the relationship between XUSRD firms' financial performance and renewable energy consumption using SHAP-supported ML models and showed that XAI provides transparency in performance evaluations based on accounting. Lachuer and Jabeur (2022) explored the connection between Corporate Social Responsibility (CSR) and financial performance in the US market using (2022) XAI and revealed that CSR only improved financial results on certain sustainability thresholds. Zhang et al. (2022) suggested a comprehensive XAI (SHAP) framework for the prediction of financial troubles. The study shows how commentable models can be used in financial decision-making processes that meet the high accuracy, trust, and transparency needs of external stakeholders. Similarly, Thompson, Buertey, and Kim (2025) reevaluate the CSR-FP relationship using SHAP-based ML. Finally, Topuz, Bajaj, Coussement, and Urban (2025) state that, thanks to XAI methods such as SHAP and LIME, increasing the disclosure of complex black box models has become an important requirement in terms of compliance and shareholder trust. The reviewed literature revealed both positive and negative findings regarding the influence of CE on FP. This situation indicates that the relationship between these variables varies depending on the sector, time period, and methods employed. However, a significant portion of existing studies is limited to linear modeling techniques, leaving complex, nonlinear relationships between variables insufficiently explained. Although the application of Explainable Artificial Intelligence (XAI) methods in finance, accounting, and sustainability has increased in recent years, these approaches are primarily confined to theoretically based and localized studies that aim to elucidate the CE-FP relationship. In this context, this study, which proposes an XAI-supported approach in Turkey, seeks to address a notable gap in both methodological and contextual literature. The relationship between CE and FP can be explained theoretically using several complementary frameworks. According to Stakeholder Theory, firms are expected to manage environmental impacts in accordance with the expectations of various stakeholders, such as contractors, regulators, and customers, who may punish or reward firms based on their environmental performance (Cordeiro & Tewari, 2015; Elijido-Ten, 2007). Similarly, Legitimacy Theory states that firms engage in social responsibility activities, including CE reduction, to protect their legitimacy and secure continuous access to resources (Crossley, Elmagrhi, & Ntim, 2021; Solikhah, Yulianto, & Suryarini, 2020). From the perspective of Resource-Based Theory, environmental responsibility can be viewed as a valuable, rare, and immutable resource that enhances long-term competitive advantage (McWilliams & Siegel, 2010; Sharma, Bhattacharya, & Thukral, 2019). These theories serve as both risk channels of CE, such as regulatory penalties and reputation loss, and opportunity channels. For example, they explain how CE can influence financial performance through cost savings, brand value, and corporate reputation.

## MATERIAL AND METHODS

### Research Design

In this study, beyond traditional regression methods, various machine learning algorithms were employed to assess the predictability of FP. Explainable AI (XAI) techniques, such as SHAP and ALE, were utilized to ensure that the model outputs were understandable and interpretable by decision-makers. Unlike conventional econometric models, this research adopts a machine learning-based analysis approach that not only guarantees high predictive accuracy but also provides explainable insights that support decision-making processes.

### Data Set and Variables

The study included 89 firms listed in the XUSRD by the Public Disclosure Platform as of December 2024. To access carbon emission data, the Sustainability, Annual, and Integrated Reports on each firm's website were analyzed individually. However, carbon emission data for only 63 firms could be accessed, and the remaining firms were excluded from the scope of the study.

Data on carbon emission amounts are generally included in the environmental performance indicators of reports and consist of the sum of scopes 1, 2, and 3 of GHG emissions. All data were recorded in tons of carbon dioxide equivalent (tons $CO_2e$), and data presented in different units (e.g., kilograms, tons, or millions) were standardized and converted to $CO_2e$. Data on the financial indicators were obtained from the Finnet2000Plus database. The dataset used in this study consists of 189 firm-year observations covering the years 2021-2023. The BIST firm codes for these 63 firms are presented in Table 1.

**Table 1:** BIST codes of the XUSRD firms.

| Number | BIST Code | Number | BIST Code | Number | BIST Code | Number | BIST Code |
|---|---|---|---|---|---|---|---|
| 1 | AGHOL | 17 | BIMAS | 33 | KORDS | 49 | TCELL |
| 2 | AGESA | 18 | BIZIM | 34 | LOGO | 50 | TUPRS |
| 3 | AKBNK | 19 | DOHOL | 35 | MAGEN | 51 | THYAO |
| 4 | AKCNS | 20 | DOAS | 36 | MAVI | 52 | TTKOM |
| 5 | AKENR | 21 | ENJSA | 37 | MGROS | 53 | TTRAK |
| 6 | AKFGY | 22 | ENKAI | 38 | MPARK | 54 | GARAN |
| 7 | AKSGY | 23 | EREGL | 39 | NATEN | 55 | HALKB |
| 8 | AKSA | 24 | ESEN | 40 | NUHCM | 56 | ISCBTR |
| 9 | AKGRT | 25 | FROTO | 41 | OTKAR | 57 | TSKB |
| 10 | ALBRK | 26 | GWIND | 42 | PETKM | 58 | SISE |
| 11 | AEFES | 27 | ISDMR | 43 | SASA | 59 | VAKBN |
| 12 | ANHYT | 28 | ISMEN | 44 | SKBNK | 60 | ULKER |
| 13 | ARCLK | 29 | KERVT | 45 | SOKM | 61 | VESBE |
| 14 | ASELS | 30 | KMPUR | 46 | TATGD | 62 | VESTL |
| 15 | AYDEM | 31 | KCHOL | 47 | TKFEN | 63 | ZOREN |
| 16 | AYGAZ | 32 | KONTR | 48 | TOASO | | |

The main reason for choosing the firms in the XUSRD as the dataset in this study is that these firms have a relatively higher level of transparency and responsibility in reporting their environmental, social, and governance (ESG) performance (BIST, 2014). Because this index includes firms that comply with sustainability principles and regularly disclose environmental performance indicators such as CE, it provides a suitable sample in terms of data quality and comparability.

In the literature, both accounting-based (ROA, ROE and Leverage) and market-based (PB, EPS and Firm Size) performance indicators are frequently used (Aruga et al., 2023; Chang et al., 2024; Houqe et al., 2022; Khatib et al., 2023; Koç & Deran, 2024; Laskar et al., 2022; Le & Nguyen-Phung, 2024; Sakın & Kefe, 2024). The variables used in this study are presented in Table 2 in detail with their sources.

**Table 2:** Dependent and Independent Variables

| Variable Type | Variable Description | Formula | Source |
|---|---|---|---|
| Dependent Variables | Return on Assets (ROA) | Net Income/Total Assets | Finnet2000Plus |
| | Return on Equity (ROE) | Net Income/Total Equity | |
| Independent Variables | Carbon Emission (tons CO2e) | Scope 1+Scope 2+Scope 3 | Sustainability Report, Annual Report, Integrated Report |
| | Leverage | Total Depts/Total Assets | Finnet2000Plus |
| | Price to Book Value (PB) | Market Price/Book Value | |
| | Earnings Per Share (EPS) | Earning After Tax/Outstanding Stocks | |
| | Firm Size (FS) | Total Assets | |

Table 2 lists the dependent and independent variables. In this study, return on assets (ROA) and return on equity (ROE), which are accounting-based performance measures, were used as dependent variables. The independent variables include carbon emissions (measured in tons of $CO_2$ equivalent), leverage, PB, EPS, and firm size.

## Data Pre-Processing

Data preprocessing steps were carried out using Python, along with the pandas, numpy, and scikit-learn libraries. Missing data (NaN) were excluded from the analysis, and extreme values particularly for the ROA and ROE variables were statistically evaluated. In the initial stage, eight observations were manually removed, and the remaining outliers were addressed using the Interquartile Range (IQR) method. As a result of this process, the final dataset comprised 145 observations.

## Transformation and Scaling

The Yeo-Johnson transformation was applied to adjust the distributions of certain variables, making them closer to a normal distribution. Specifically, the variables ROA, Carbon Emission, and Firm Size TL were transformed using this method. Additionally, the independent variables "Trans Carbon Emission," "Leverage," "Price to Book," "Earnings Per Share," and "Trans Firm Size" were standardized using StandardScaler, which adjusted them to have a mean of zero and a standard deviation of one. This standardization step was implemented to enhance the performance of machine learning models by ensuring that all variables contributed equally to the analysis.

## The Machine Learning Algorithms Used

In this study, the performances of seven regression-based algorithms were examined. Complex characteristics such as non-linearity in the structure of financial data, variability over time (non-stationarity), and strong connections between consecutive values (serial correlation) make the prediction of these data quite challenging (Karaboğa, Karaboğa, Şekeroğlu, Kızıloğlu, & Acılar, 2024).

In this study, seven algorithms representing different learning paradigms were selected to model the nonlinear effects of CE on FP: Linear Regression, LightGBM, Multi-Layer Perceptron (MLP), Support Vector Regression (SVR), Random Forest, Extra Trees, and XGBoost. Linear Regression was used as the basic model for identifying linear relationships. The other methods were preferred because they do not rely on the linear assumption and demonstrate high prediction performance for nonlinear relationships. Additionally, SHAP and ALE methods from Explainable AI (XAI) were employed to make the decision mechanisms of these models understandable. This approach ensured transparency regarding the prediction models of the most successful methods.

The small size of the dataset used in this study (n=145) presents a risk of overfitting, especially in machine learning methods. To address this issue, we employed a 5-fold cross-validation approach. The effectiveness of machine learning methods is highly sensitive to parameter settings; therefore, GridSearchCV hyperparameter optimization was utilized to identify the most effective model for each method. The limited size of the dataset restricts the generalizability of these findings. Future studies with larger datasets could help mitigate these limitations.

### *Linear Regression*

Linear regression is a linear method that estimates a specific y output value using n x variables. It assumes that there is a linear relationship between y and x. The mathematical Equation 1 expresses this linear relationship between the input and output, representing a line or plane.

$$y = w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n + b \quad (1)$$

Here, y is the predicted outcome (dependent variable or output value) and $x_1$ is one of the factors influencing the outcome (independent variable or input value). "$w_1$" is a coefficient that indicates the strength of "$x_1$"'s influence on the outcome (weight), and "b" is a term that represents the fixed effect of other factors (bias). This approach works well when there is a linear relationship between each input and output variable and when the input variables are not strongly correlated with each other (Kim, Bae, & Jang, 2022).

### *Support Vector Machines (SVM)*

Support Vector Regression (SVR) is based on Support Vector Machines (SVM), which were originally developed for binary classification. The primary objective of this method is to establish an error margin ε around the data, focusing only on the margin below a specific threshold, known as ε-sensitivity. The fundamental approach of Support Vector Regression (SVR), similar to the Support Vector Machine (SVM) algorithm, involves transforming the data and mapping it into a high-dimensional feature space using kernel functions. A hyperplane with a high generalization capacity and a good fit to the data can be identified (López, López, & Crossa, 2022).

### *Random Forest*

The Random Forest algorithm is a robust, tree-based ensemble method with significant modeling capacity for nonlinear structures, as highlighted by Breiman (2001). It is notable for its high performance, particularly in datasets characterized by complex nonlinear relationships and numerous variables. This method operates by constructing a large number of decision trees on randomly selected subsets of the dataset; each tree provides independent predictions, and the final output is obtained by averaging these predictions (Karaboğa et al., 2024).

### *Extremely Randomized Trees*

Recommended by Geurts, Ernst, and Wehenkel (2006), ET is an ensemble learning method that combines multiple decision trees similar to the Random Forest. ET is a choice of features for a random subset in node divisions. Unlike Random Forest, it makes the division process more random by selecting random division points for selected features instead of looking for the best division point Geurts et al. (2006). The algorithm uses all the training datasets to choose a random feature subset (K) for each tree. It forms division rules by identifying random cutting points for the features. One of these random divisions is selected if it is to split a knot. This process is repeated until the leaf nodes are reached (Alfian et al., 2022). It also reduces the variance of the model while reducing randomness, calculation time, and memory usage (Mastelini, Nakano, Vens, & De Leon Ferreira, 2022). After training is completed, the Gini importance can be calculated to evaluate the importance of properties (Alfian et al., 2022). The random structure of ET improves the rigidity of the model against noisy data, and the parameter adjustment ensures the adaptability of the model to the problem characteristics (Geurts et al., 2006).

### *Extreme Gradient Boosting (XGBoost)*

XGBoost is a tree-based gradient boosting algorithm that stands out due to its high efficiency. A specialized sparsity-sensitive method has been developed to improve performance on sparse datasets. An important factor contributing to the success of the model is the selection of hyperparameters. The algorithm creates a new tree based on the errors remaining from the previous trees, thereby systematically reducing the estimation error. The final model is obtained by either minimizing the residual values or reaching the maximum number of trees (Bulut & Korkmaz, 2024; Chen & Guestrin, 2016; Karamanou, Kalampokis, & Tarabanis, 2022).

### *Light Gradient Boosting Machine (LightGBM)*

LightGBM is a gradient boosting method developed by Microsoft in 2017 to enhance the performance of GBDT algorithms. The algorithm achieves high-accuracy information gain by reducing feature dimensions (EFB) and data volume (GOSS). Unlike XGBoost, which uses a leaf-wise tree growth strategy, LightGBM employs a leaf-wise approach that can lead to deeper trees and potentially higher accuracy. It uses fewer nodes, which improves computational efficiency. Additionally, LightGBM can directly process categorical features without requiring encoding, simplifying the preprocessing pipeline and reducing training time (Bakır, Orak, & Yüksel, 2024; Ke et al., 2017).

### *Multi-Layer Perceptron Regression (MLP)*

A type of artificial neural network (ANN) designed as an artificial model of the human brain neuron; the multi-layer perceptron architecture is an ANN method created for use in regression problems. The method consists of an output layer, input layer, and at least one hidden layer. The output layer processes the weights of the values calculated in the hidden layers and produces the final output by adjusting the weights. ANNs include activation functions that apply the weighted sum of the input values to neurons and produce the output value (Orhan, Kilinc, Albayrak, Aydin, & Torun, 2022).

## Evaluation of the Machine Learning Algorithms

The performance of the ML algorithms varied depending on the characteristics of the dataset used in the comparative analysis. The performance metrics commonly used to compare machine-learning models are presented in Table 3, along with their formulas (Chang et al., 2024). The algorithms were trained using hyperparameter optimization with the GridSearchCV method, employing a training/test split of 80% to 20%.

**Table 3:** Performance Metrics Used to Evaluate Machine Learning Models

| Metric | Description | Calculation |
|---|---|---|
| $R^2$ | Determination Coefficient | $R^2 = 1 - \frac{\sum_{i=1}^{n}(A_i - F_i)^2}{\sum_{i=1}^{n}(A_i - \bar{A})^2}$ |
| MAE | Mean absolute error | $\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} \lvert F_i - A_i \rvert$ |
| Adjusted $R^2$ | Adjusted $R^2$ | $\text{Adjusted } R^2 = 1 - \left(\frac{(1-R^2)(n-1)}{n-k-1}\right)$ |
| RMSE | Root means square error | $\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (A_i - F_i)^2} = \sqrt{\text{MSE}}$ |

These metrics ($R^2$, MAE, Adjusted $R^2$, and RMSE) were calculated separately for the ROA and ROE forecast models. The forecasting success of different machine learning algorithms on both financial accounting performance indicators was then comparatively analyzed.

### *Diebold-Mariano Test*

In this study, the Diebold-Mariano (DM) test developed by Diebold and Mariano (1995) was used to determine whether the differences in the performance of the most successful prediction models obtained as a result of hyperparameter optimization and cross-validation processes were statistically significant. The Diebold-Mariano (DM) test is a non-parametric method used to measure the significance of differences between the prediction errors of two different models (Rathod et al., 2021; Uğurlu, Öksüz, & Taş, 2018). In this context, between the reference models with the highest $R^2$ values, bilateral comparisons were made based on both the root mean square error (RMSE) and mean absolute error (MAE) criteria.

### *Explainable Artificial Intelligence (XAI)*

Artificial intelligence has a significant impact on social life today. However, most artificial intelligence methods are based on machine learning algorithms, which serve as the decision mechanisms of internal functioning. These algorithms are often considered a non-transparent "black box." Although models created with deep learning and ensemble methods achieve high accuracy, their complexity makes them difficult to interpret. This situation necessitates models that can be explained both reliably and understandably. The general term for methods developed to clarify the decision-making processes of artificial intelligence and machine learning models is XAI (Explainable Artificial Intelligence). Prominent methods for transparency in decision-making include SHAP, LIME, ALE, and PDP. These methods aim to elucidate the reasons behind model predictions, thereby enhancing trust and interpretability in AI systems (Černevičienė & Kabašinskas, 2024).

We used SHAP and ALE to investigate how our best-performing ROA and ROE prediction models actually work internally. Using SHAP, we created feature importance bar charts, beeswarm plots, and partial dependence plots, while ALE helped us understand the precise marginal impact of CE on both ROA and ROE.

### *SHapley Additive Explanations (SHAP)*

Lundberg and Lee (2017) proposed SHAP's method to explain the forecasts of black box ML models that cannot be interpreted. This method is a powerful explanation technique used to reveal the causes of the outputs of various machine learning models based on collaborative game theory, such as classification and regression. SHAP is widely preferred to make model decisions more transparent (Zhang et al., 2023). The SHAP method aims to accurately measure the contribution of each feature to the model prediction. This contribution is calculated using the Shapley value, which represents the average effect of a feature across all possible combinations of features (Yıldırım, Rençber, & Yıldırım, 2024). The SHAP formula is defined in Equation 2.

$$g(x') = \Phi_0 + \sum_{j=1}^{M} \Phi_j x'_j \quad (2)$$

Here, $g(x')$ the final prediction of the model represents the average forecast value derived from $\Phi_0$ the training data, while the number of attributes used by the M model refers to the value of Shapley, which contributes to predicting the corresponding feature. Thus, the model output is made more understandable by evaluating the effects of self-qualities individually (Yildirim et al., 2024).

The contributions of individual variables to the model's output can be visualized on both an average and a sample basis. SHAP enables analysis of both global (overall importance) and local (individual prediction) aspects. In this study, SHAP analysis is used to examine the effects of financial and environmental variables on both the ROA and ROE models.

### *Accumulated Local Effects (ALE)*

ALE charts produce visualizations based on local information developed to overcome security issues related to Partial Addiction Graphics and high-correlation variables. ALE measures the average marginal effect of a variable on the model using local gradients. These charts clearly show the direction and magnitude of a variable's effect across different models. This is important for detecting nonlinear relationships within the data (Apley & Zhu, 2020; Christoph, 2020). In this study, ALE graphics were preferred to visualize and analyze the structure of the nonlinear relationship between ROA and ROE, which are financial performance indicators.

## ANALYSIS AND FINDING

The flow diagram of the research is presented in Figure 1. In line with the purpose of the study, both financial and non-financial data were initially collected. These data were then organized and prepared through a data preprocessing step. Subsequently, the dataset was randomly divided into an 80% training set and a 20% test set.

During the model development process, hyperparameter optimization and 5-fold cross-validation were performed using the GridSearchCV method. Machine learning methods such as Random Forest, ExtraTrees, XGBoost, LightGBM, Multi-Layer Perceptron (MLP), Support Vector Regression (SVR), and linear regression algorithms were employed. The performance of the developed prediction models was evaluated using metrics including $R^2$, adjusted $R^2$, Root Mean Square Error (RMSE), and Mean Absolute Error (MAE). Ultimately, the most successful model was identified by comparing the performance metrics across different tests. Additionally, the prediction accuracies of the models were statistically compared using the Diebold-Mariano test, based on MAE and RMSE metrics. To enhance the interpretability of the optimal model, SHAP (SHapley Additive exPlanations) and ALE (Accumulated Local Effects) analyses were conducted. These analyses revealed the effects and importance of independent variables on model outputs, specifically for Return on Assets (ROA) and Return on Equity (ROE) predictions.

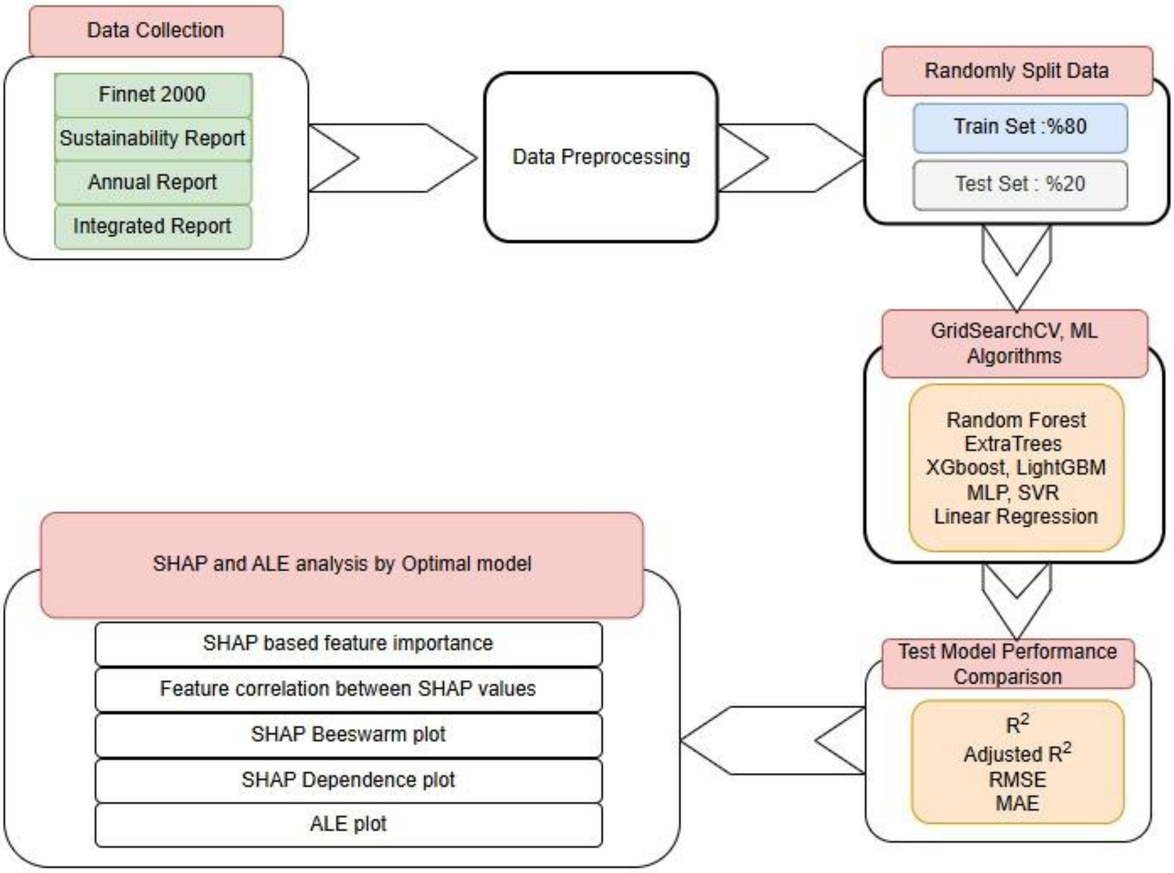



**Figure 1:** Analysis flowchart for the development and evaluation of the models.

## Machine Learning Prediction Results

In this section, the performance of various ML algorithms in predicting the dependent variables ROA and ROE is compared. Table 3 shows the regression performance metrics of the best ROA prediction models for different ML algorithms on the test dataset. Table 4 lists the regression performance metrics of the best ROE prediction models for the different ML algorithms on the test dataset. The results obtained for both dependent variables are presented in detail.

**Table 4:** ROA Regression Results

| Model | Test $R^2$ | Adjusted $R^2$ | RMSE | MAE |
|---|---|---|---|---|
| Random Forest | **0.7224** | **0.6620** | **0.4293** | **0.3329** |
| XGBoost | 0.6500 | 0.5739 | 0.4820 | 0.3949 |
| LightGBM | 0.6381 | 0.5595 | 0.4901 | 0.3889 |
| SVR | 0.6059 | 0.5202 | 0.5115 | 0.3972 |
| ExtraTrees | 0.6005 | 0.5137 | 0.5150 | 0.4053 |
| MLP | 0.5569 | 0.4605 | 0.5424 | 0.4319 |
| Linear Regression | 0.5190 | 0.4144 | 0.5651 | 0.4606 |

As shown in Table 4, Random Forest demonstrates the best overall performance, with the highest $R^2$ value of 0.7224, the lowest Root Mean Square Error (RMSE) of 0.4293, and the lowest Mean Absolute Error (MAE) of 0.3329. These results indicate that the Random Forest model is the most effective in explaining the variance in Return on Assets (ROA) and in minimizing absolute errors. In conclusion, the Random Forest model appears to be the most promising approach for this dataset, offering reliable predictive accuracy and robustness.

In line with the strong performance shown in Table 4, Random Forest was accepted as the reference model, and statistical comparisons were made with other models based on the RMSE and MAE loss functions. In this context, binary model comparisons were performed using Diebold-Mariano test statistics, and the results are presented in Table 5.

**Table 5:** Diebold-Mariano Test Statistics for the ROA Regression Models (Reference Model: Random Forest)

| Compared Model | DM (RMSE) | DM (MAE) |
|---|---|---|
| LightGBM | −2.725*** (p=0.006) | −1.998** (p=0.046) |
| LinearReg | −2.508** (p=0.012) | −2.206** (p=0.027) |
| XGBoost | −1.047 (p=0.295) | −1.188 (p=0.235) |
| ExtraTrees | −1.730 *(p=0.084) | −1.054 (p=0.292) |
| MLPReg | −1.269 (p=0.205) | −0.794 (p=0.427) |
| SVRReg | −1.194 (p=0.232) | −0.150 (p=0.881) |

**Note:** Negative DM statistics indicate that the reference model had a lower prediction error than the model being compared. Significance levels are indicated by * for 10%, ** for 5%, and *** for 1%. Only significant levels present in the table are shown.

Table 5 demonstrates that when the Random Forest model was used as the reference model, it produced statistically significantly lower prediction errors than the LightGBM and Linear Regression models. These differences are statistically significant at the 5% and 1% significance levels for both the RMSE and MAE metrics. Although no significant difference was observed compared to other models, a result close to the 10% significance level was obtained in the comparison with the ExtraTrees model based on the RMSE.

Table 6 summarizes the performances of the different ML models in the test. XGBoost performed the best, with the highest $R^2$ (0.6953) and the lowest RMSE (0.4256). This indicates that XGBoost successfully explained the variance in the dependent variable and minimized the root mean square error. Random Forest has a similarly high $R^2$ (0.6942) and a low RMSE (0.4263). Therefore, both the XGBoost and Random Forest models are strong candidates for predicting ROE.

**Table 6:** ROE Regression Results

| Model | Test $R^2$ | Adjusted $R^2$ | RMSE | MAE |
|---|---|---|---|---|
| XGBoost | 0.6953 | 0.6318 | 0.4256 | 0.3245 |
| Random Forest | 0.6942 | 0.6305 | 0.4263 | 0.3578 |
| ExtraTrees | 0.6698 | 0.6010 | 0.4430 | 0.3551 |
| LightGBM | 0.6544 | 0.5824 | 0.4532 | 0.3899 |
| SVR | 0.5912 | 0.5060 | 0.4929 | 0.3841 |
| MLP | 0.5810 | 0.4937 | 0.4990 | 0.4119 |
| Linear Regression | 0.5440 | 0.4490 | 0.5206 | 0.4444 |

In the comparisons made for ROE estimates, XGBoost was selected as the reference model, and Diebold-Mariano tests were performed with other models. The results are listed in Table 7.

**Table 7:** Diebold-Mariano (DM) Test Statistics for ROE Regression Models (Reference Model: XGBoost)

| Compared Model | DM (RMSE) | DM (MAE) |
|---|---|---|
| LinearReg | −2.102** (p=0.035) | −2.501** (p=0.012) |
| RandomForest | −0.797 (p=0.426) | −1.491 (p=0.136) |
| LightGBM | −1.213 (p=0.225) | −1.498 (p=0.134) |
| ExtraTrees | −0.903 (p=0.366) | −1.468 (p=0.142) |
| MLPReg | −1.840* (p=0.066) | −1.608 (p=0.108) |
| SVRReg | −1.850* (p=0.064) | −1.426 (p=0.154) |

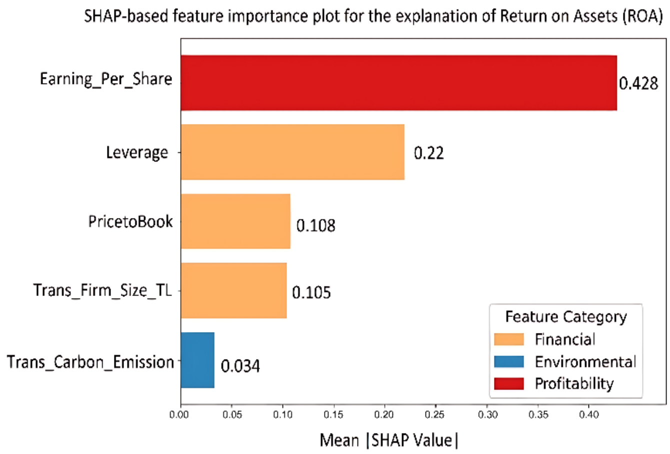
**Note:** Negative DM statistics indicate that the reference model had a lower prediction error than the model being compared. Significance levels are indicated by * for 10%, ** for 5%, and *** for 1%. Only significant levels present in the table are shown.

As shown in Table 7, the XGBoost model was found to perform significantly better than the linear regression model in terms of both RMSE and MAE. Although significance levels were not reached in comparisons with other models, the p-values in the tests against the MLPReg and SVRReg models were very close to the 10% significance threshold.

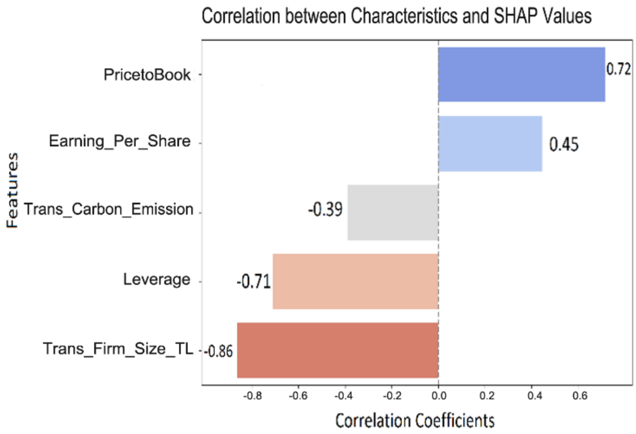
Overall, the results obtained from both ROA and ROE predictions indicate that among the seven machine learning algorithms tested, the Random Forest and XGBoost models demonstrate the strongest generalization capabilities based on $R^2$ and error metrics. The Random Forest model proved to be the most accurate in predicting ROA, while the XGBoost model excelled in predicting ROE. Consistent with these findings, the Random Forest model, optimized with the most appropriate hyperparameters, was selected as the baseline model for subsequent SHAP and ALE analyses.

### *SHAP-Based Interpretations*

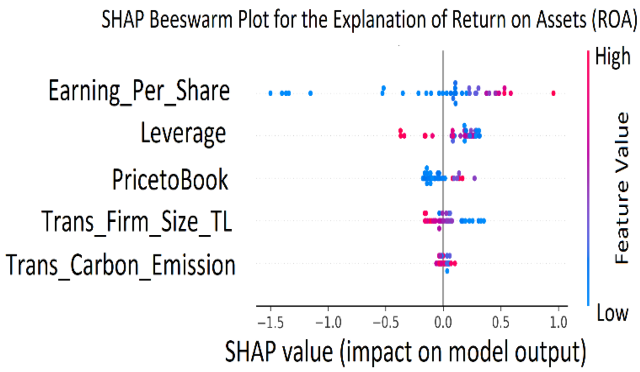
First, the optimal random forest model for ROA estimation was configured with the following hyperparameters: n_estimators = 100, max_depth = None, min_samples_split = 10, and min_samples_leaf = 2. Next, the explainability of the model and the analysis of variable importance levels were calculated using the TreeExplainer method from the SHAP library.



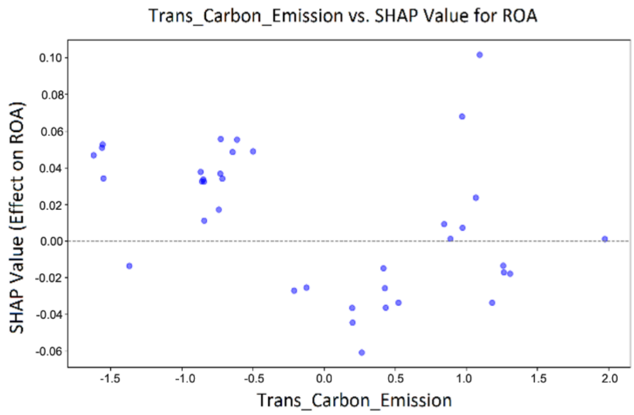

(a) Shap-based feature importance for ROA




(b) Feature Correlation between SHAP values




(c) Shap beeswarm plot for ROA




(d) SHAP dependence plot for ROA

**Figure 2.** SHAP plots for ROA with RandomForest

Figure 2 shows the SHAP plots obtained for the ROA predictor model created using the random forest algorithm. Figure 2(a) illustrates the contribution levels of the ROA forecast model for each individual variable, based on the mean SHAP (Shapley Additive exPlanations) values. The horizontal axis represents the average impact of each variable on ROA prediction. The analysis indicates that EPS and Leverage are significant variables contributing to the ROA prediction model. Additionally, the Trans_Carbon_Emission variable has the lowest contribution, with a mean SHAP value of 0.034. This suggests that Trans_Carbon_Emission has a smaller effect on the model compared to financial accounting indicators, such as EPS and Leverage, which have a more substantial influence on ROA prediction.

Figure 2(b) shows the relationship between the variables and SHAP values, which contributes to the understanding of interactions in the model. The PB variable has shown a strong positive correlation at 0.72 levels with ROA. This situation indicates that firms with a higher PB ratio generally have higher ROA values. The EPS variable also revealed a positive correlation at 0.45 levels. On the other hand, Trans_Carbon_Emission has shown a negative correlation at the -0.39 level with ROA. The correlation between Leverage and ROA indicates a highly negative relationship, approximately -0.71. These results suggest that CE negatively affects ROA, but this effect is not as strong as that of leverage.

Figure 2(c) demonstrates the SHAP beeswarm plot, submitting a detailed visualization of the contributions made to the ROA forecast across individual observations in the dataset. The horizontal axis represents the SHAP values. Each point in the plot corresponds to an observation in the dataset. The colors indicate the size of the variables: blue represents low values, and red represents high values. Variables are sorted from above according to the size of their effects on the model. According to the plot, the EPS variable usually produces positive SHAP values, which highlights its decisive role in the model. A wide range of SHAP values for this variable indicates that its effect remains high if it differs between observations. The leverage and PB variables showed linear effects, producing both positive and negative SHAP values. Conversely, low Trans_Carbon_Emissions values can be positive or negative depending on the company, influenced by the impact of high values, while positively affecting ROA. This finding supports the idea that CE can be effective in ROA.

Figure 2(d) visualizes the relationship between SHAP values, representing the impact of the Trans_Carbon_Emission variable and ROA, using the SHAP dependence plot. Such charts are a powerful tool for analyzing how a certain variable contributes to an ML model and how this additive changes according to the variable value. The plot reveals that lower levels of CE often have a positive effect on ROA. This finding suggests that firms that reduce CE may be more advantageous in terms of their financial performance. However, while high CE levels positively impact ROA in some firms, they can have negative results in others. This reveals that the effect of CE is contextual and may vary between firms.

In summary, according to the findings of the SHAP analysis, CE has a lower effect than financial accounting indicators on ROA. Although the impact of CE on FP is very low, it should not be ignored in these variable strategic decision-making processes for operating activities.

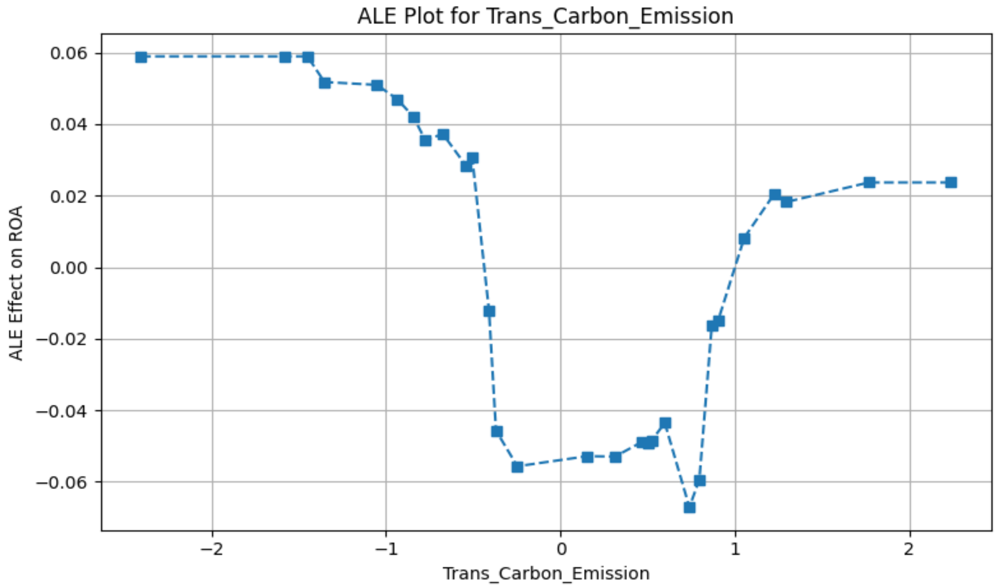
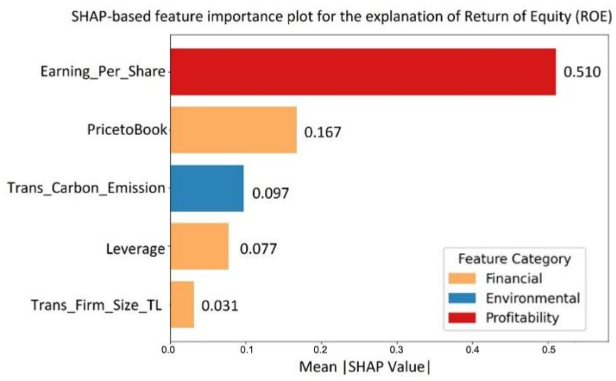



**Figure 3:** ALE plot of Trans_Carbon_Emission variable on ROA

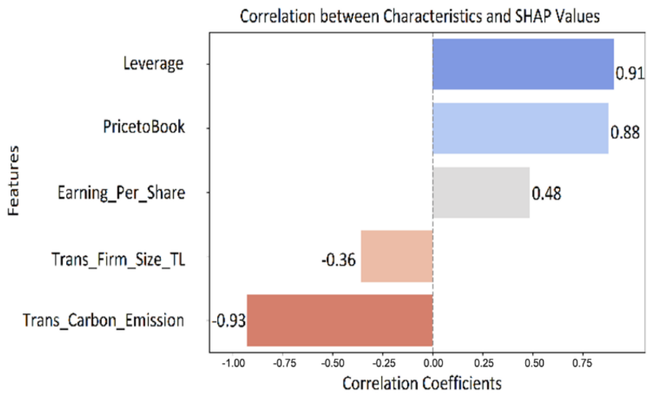
Figure 3 visualizes the ALE on ROA in different value ranges of the Trans_Carbon_Emission variable. ALE plots were used to explain the decisions of the ML models and to show the effect of a specific variable on the model. The chart clearly shows that the relationship between Trans_Carbon_Emission and ROA is non-linear.

When Trans_Carbon_Emission X < ~-0,4, a positive ALE effect was observed. This indicates that firms with low CE tend to have higher their ROA. In other words, the model predicts that these firms are more profitable. When the Trans_Carbon_Emission -1 < X < 1, the ALE value quickly reduces a negative value (~-0,05). This indicates that businesses with moderate CE will have a lower ROA value. When the Trans_Carbon_Emission value is high (X > 1), the ALE effect passes to the positive zones. This indicates that some businesses with high CE can have a negative or partially positive effect on ROA.

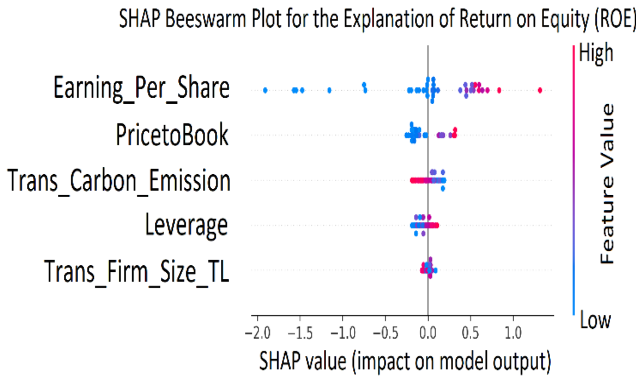
The ALE chart reveals that the effect of CE on FP is not linear; it is positive for extremely high and extremely low values, and medium levels are much more negatively effective.



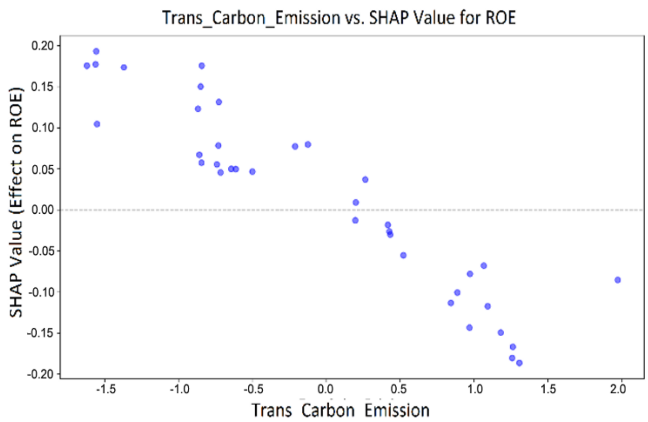

(a) Shap-based feature importance for ROE




(b) Feature Correlation between SHAP values




(c) Shap beeswarm plot for ROE




(d) SHAP dependence plot for ROE

**Figure 4:** SHAP plots for ROE with Random Forest

Figure 4 shows the SHAP plots of the ROE forecast model created using the Random Forest algorithm. According to the data in Figure 4(a), the variable that provides the highest contribution to the ROE regression model is EPS. This finding reveals that EPS plays a decisive role in predicting ROE. In contrast, the Trans_Carbon_Emission variable showed a moderate effect on the model with an average SHAP value of 0.097. This indicates that the impact of CE on ROE is not at a dismissible level.

When Figure 4(b) is analyzed, the leverage variable exhibits a strong positive correlation with SHAP values at the 0,91 level. This indicates that firms with high leverage ratios are more positively evaluated by the ROE model. Similarly, the PB variable showed a positive correlation at 0.88 levels. This finding demonstrates that this variable contributes to the model's performance for firms with high market values. The EPS variable showed a moderate positive correlation (0.48). This suggests that the model considers EPS an element that increases profitability. Conversely, the Trans_Carbon_Emission variable exhibited a high negative correlation (-0.93). This indicates that carbon emissions significantly reduce its impact on ROE in businesses with high emission levels.

In the SHAP beeswarm plot shown in Figure 4(c), the EPS variable has the greatest effect on ROE. Observations with high EPS values (red spots) often show strong negative effects when producing positive SHAP values, and low EPS values (blue spots) are associated with negative effects. Similarly, while the PB variable creates a mild positive effect on ROE at high values, a slight negative effect is observed at low values (blue spots). High-emission values of the Trans_Carbon_Emission variable (red spots) are often associated with negative SHAP values, leading to a decrease in ROE. Conversely, low Trans_Carbon_Emission values (blue spots) are associated with positive SHAP values and tend to increase ROE. These findings indicate a negative relationship between CE and ROE. Regarding the leverage variable, low leverage levels show a slight negative impact, while medium levels create a mild positive effect. In contrast, the effect of Trans_Firm_Size on ROE is minimal.

When the SHAP dependency plot for ROE in Figure 4(d) is analyzed, the transCE values are reduced, and the SHAP values are increased. These data reveal that the model associates a lower CE with a higher ROE value. In contrast, the negative trend of SHAP values increasing the CE values shows that higher emission levels are associated with lower ROE estimates by the model.

Similar to the ROA model, lower CE levels in the ROE model correspond to relatively less negative SHAP values. This suggests that firms with lower emissions can alleviate the negative effects of CE on ROE. While the overall trend is negative, the SHAP dependency chart shows lower effects at low and medium CE levels, indicating a certain threshold effect.

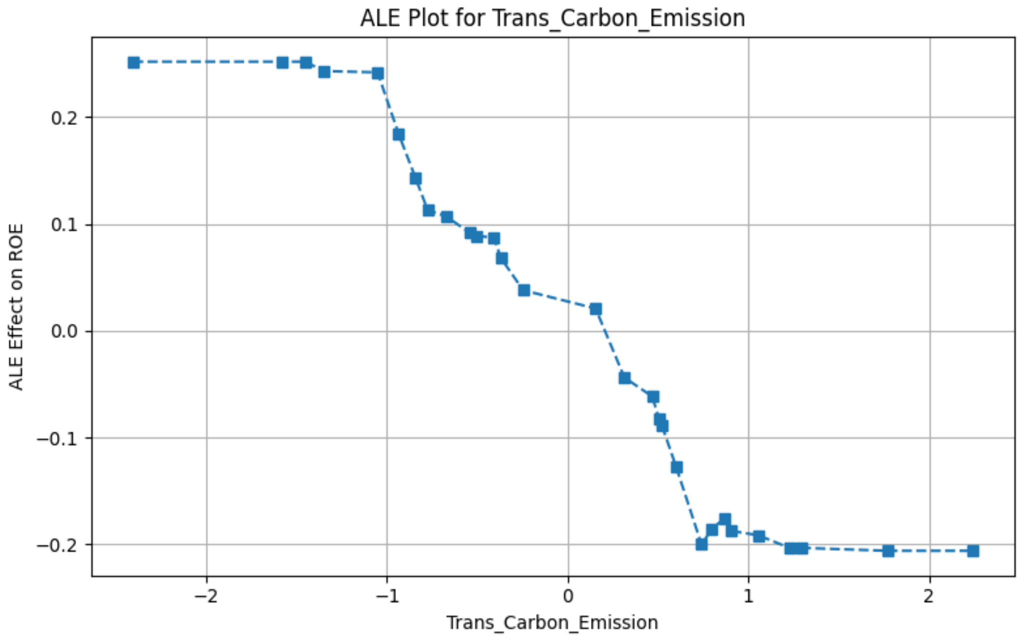



**Figure 5:** ALE plot of Trans_Carbon_Emission variable on ROE

As shown in Figure 5, the Trans_Carbon_Emission variable was analyzed using the ALE plot on ROE. This analysis revealed that the impact of Trans_Carbon_Emission values on ROE had no linear structure. As shown in the ALE plot, for firms with low emission levels (-2<X<-1), the effect of CE on ROE is positive or neutral. However, as the CE level increases, especially in the range (-1<X<1), the impact on ROE is significantly reduced. This indicates that an increase in CE in this range particularly reduces FP. When the Trans_Carbon_Emission value reaches high levels (X>1), the negative impact on ROE is relatively stable. Hence, a negative effect was observed horizontally after a certain level. In summary, Trans_Carbon_Emission values have a positive effect on ROE, but the effect of these values increases rapidly. At high CE levels, this negative impact on ROE continues and puts pressure on FP. This analysis shows that CE levels may have an important impact on ROE, and low-emission levels may be more suitable for financial performance.

The ALE plot clearly indicates that the effect of CE on FP is not linear. At low emission levels, emissions have a positive or neutral effect on profitability; however, this effect becomes increasingly negative when emission values exceed a certain threshold. This suggests that high emission levels can lead to negative outcomes for FP. Therefore, it is crucial for firms to determine the optimal CE level for their activities to achieve sustainable development.

## CONCLUSION

This study investigated the effect of CE on FP using SHAP and ALE from XAI methods. In the context of this study, annual data of firms listed in the XUSRD for the period 2021-2023 were used. The CE data of these firms are provided in sustainability, annual, and integrated reports, while financial accounting data are obtained through Finnet2000plus. Among the variables in the study, ROA and ROE are the dependent variables, while the other variables are assigned as independent variables (CE in tons CO2e, leverage, PB, EPS, and FS).

The data produced in the study were applied to seven different machine learning algorithms (Random Forest, ExtraTrees, XGBoost, LightGBM, MLP, SVR, and linear regression), and these algorithms were compared based on their predictive performances on ROA and ROE. Among these, Random Forest demonstrated the strongest overallization capacity. Therefore, the Random Forest model trained with the most suitable hyperparameters for SHAP and ALE analyses is preferred.

Shap analysis revealed that the most effective variables in terms of ROA are EPS and leverage, while for ROE, EPS is a decisive variable. In summary, in both models, the EPS variable emerged as the variable that provided the highest contribution to the predictive and profitability outcomes. The CE variable has a low impact on the ROA prediction model, whereas the ROE model exhibits a moderate and negative relationship with FP. CE has a negative effect on both ROA and ROE, and this effect is more pronounced in ROE. This finding shows that CE can be directly associated with FP, and investors perceive CE firms as riskier. This indicates that CE may be directly related to FP and that investors perceive firms with high CE as riskier.

SHAP beeswarm plots reveal that low CE levels have a positive effect on ROA, while at high CE levels, some firms create positive effects, while others show negative effects. These findings show that ROA is not directly affected by CE and can yield different results at the industry and firm levels of carbon strategies. ROE is negatively correlated. High CE values tend to reduce profitability, whereas low CE values tend to increase profitability. The findings related to these relationships were consistent with the SHAP dependence plot analysis.

The ALE analysis results indicated that the effect of CE on FP (ROA and ROE) was not linear. Additionally, CE is positive or neutral up to a certain level, but a negative effect is markedly observed when specific thresholds are exceeded. This indicates that determining the optimal CE level for firms is critical for FP and that low CE strategies can support FP.

Consequently, the findings demonstrate that CE and FP have a non-linear and complex relationship. Based on SHAP and ALE analyses, high CE levels often have a negative impact on profitability, but under certain conditions, low CE levels often positively affect profitability. These results are supported by a meta-analysis by Busch and Lewandowski (2018), which highlights the methodological differences, such as the measurement and context of this relationship. In contrast to Lee et al. (2015), this study does not achieve a general conclusion that CE consistently reduces firm value. What's more, EPS emerged as the most effective factor for both performance indicators, and Random Forest was determined to be the most effective ML algorithm in financial forecasts. Financial accounting indicators are decisive in profitability prediction, whereas the impact of CE is low, yet it is an important variable. These findings show that CE requires more effective environmental and financial regulations to account for its impact on FP. Policymakers must report CE in a standard manner and encourage firms to integrate this information into their financial reports. This will also support the adoption of low-carbon firm strategies and increase transparency. On the contrary, the increased global emphasis on CE financial reporting is supported by new international frameworks such as the Carbon Border Adjustment Mechanism (CBAM) and the EU Taxonomy. These advancements further highlight the strategic importance of reporting CE information for firms that want to maintain transparency, comparability, compatibility, and financial stability.

This study contributes to the advancement of the literature by utilizing XAI methods (SHAP and ALE) to review ML-based FP models within the scope of carbon accounting. Moving beyond traditional statistical methods, these approaches enhance the understanding of ML model outputs. They can assist firms in evaluating the financial impacts of CE strategies and support the development of optimal carbon-efficient policies. Furthermore, they can significantly aid in integrating sustainability accounting policies, particularly in developing markets, and promote more interpretable and consistent ESG reporting. Although such applications are increasingly common in international studies, they remain relatively rare in research conducted in Turkey. The findings of this study have numerous practical and policy implications. For companies, especially those in sustainability indices, the negative relationship between high CE and profitability criteria (notably ROE) underscores the financial significance of carbon management strategies. This information may influence strategic decisions related to investment planning, risk management, and ESG disclosure.

This study had some limitations. The research focuses on the CE data of firms, including XUSRD for the 2021-2023 period, and the sample is limited to only 63 firms that broadcast public data during this period. Sustainability reporting has not been able to create a larger sample due to data accessibility issues, as it is still a developing area in Turkey and covers a limited number of related index firms. Therefore, only firms with accessible data are included in the analysis. This situation partially limits the overall ability of the results and makes it difficult to analyze long-term trends. CE data were collected manually from sustainability, integrated, and operating reports of firms, and some firms were excluded from the analysis due to shortcomings and access challenges in data standardization. 2024 data have not been included in the analysis as they had not yet been disclosed to the public when the work was executed. Sectoral analysis has been removed from the evaluation because the study would exceed the scope of the existing work.

For future research, high-emission sectors (e.g., energy, cement, manufacturing, transportation) and low-emission sectors (e.g., software and services) are recommended for industry-specific SHAP and ALE analyses to better understand the differences. Expansion of the dataset to include international firms can increase comparability among regulatory environments. In addition, scenario-based models can be used to simulate the effects of Carbon Pricing Mechanisms (such as Carbon Tax or Emission Trading Systems).

**FUNDING:** This study received no specific financial support.
**INSTITUTIONAL REVIEW BOARD STATEMENT:** Not applicable.
**TRANSPARENCY:** The authors confirm that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
**COMPETING INTERESTS:** The authors declare that they have no competing interests.
**AUTHORS' CONTRIBUTIONS:** Both authors contributed equally to the conception and design of the study. Both authors have read and agreed to the published version of the manuscript.

## REFERENCES

Alfian, G., Syafrudin, M., Fahrurrozi, I., Fitriyani, N. L., Atmaji, F. T. D., Widodo, T., . . . Rhee, J. (2022). Predicting breast cancer from risk factors using SVM and extra-trees-based feature selection method. *Computers*, *11*(9), 136. https://doi.org/10.3390/computers11090136

Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(4), 1059-1086. https://doi.org/10.1111/rssb.12377

Arrieta, B. A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., . . . Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, *58*, 82-115. https://doi.org/10.1016/j.inffus.2019.12.012

Aruga, R., Chiba, T., & Goshima, K. (2023). *CO2 emissions and corporate performance: Japan's evidence with double machine learning*. IMES Discussion Paper Series No. 22-E-01. Institute for Monetary and Economic Studies, Bank of Japan.

Bakır, R., Orak, C., & Yüksel, A. (2024). Optimizing hydrogen evolution prediction: A unified approach using random forests, lightGBM, and Bagging Regressor ensemble model. *International Journal of Hydrogen Energy*, *67*, 101-110. https://doi.org/10.1016/j.ijhydene.2024.04.173

BIST. (2014). *Borsa Istanbul sustainability index*. Istanbul, Turkey: Borsa Istanbul.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Bulut, S., & Korkmaz, A. (2024). Comparative analysis of machine learning models for android malware detection. *Sakarya University Journal of Science*, *28*(3), 517-530. https://doi.org/10.16984/saufenbilder.1350839

Busch, T., & Lewandowski, S. (2018). Corporate carbon and financial performance: A meta-analysis. *Journal of Industrial Ecology*, *22*(4), 745-759. https://doi.org/10.1111/jiec.12591

Çankal, A., & Ever, D. (2025). The effects of renewable energy consumption on financial performance: An explainable artificial intelligence (XAI)-based research on the BIST sustainability index. *International Journal of Energy Economics and Policy*, *15*(4), 204-213. https://doi.org/10.32479/ijeep.19602

Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. *Artificial Intelligence Review*, *57*(8), 216. https://doi.org/10.1007/s10462-024-10854-8

Chang, A., Lee, T.-S., & Lee, H.-M. (2024). Applying sustainable development goals in financial forecasting using machine learning techniques. *Corporate Social Responsibility and Environmental Management*, *31*(3), 2277-2289. https://doi.org/10.1002/csr.2694

Chang, J. (2025). Implementation of the EU carbon border adjustment mechanism and China's policy and legal responses. *Environmental Impact Assessment Review*, *110*, 107683. https://doi.org/10.1016/j.eiar.2024.107683

Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. Paper presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Christoph, M. (2020). *Interpretable machine learning: A guide for making black box models explainable. Leanpub*. Retrieved from http://dlib.hust.edu.vn/handle/HUST/24999

Cordeiro, J. J., & Tewari, M. (2015). Firm characteristics, industry context, and investor reactions to environmental CSR: A stakeholder theory approach. *Journal of Business Ethics*, *130*(4), 833-849. https://doi.org/10.1007/s10551-014-2115-x

Crossley, R. M., Elmagrhi, M. H., & Ntim, C. G. (2021). Sustainability and legitimacy theory: The case of sustainable social and environmental practices of small and medium-sized enterprises. *Business Strategy and the Environment*, *30*(8), 3740-3762. https://doi.org/10.1002/bse.2837

Delmas, M. A., Nairn-Birch, N., & Lim, J. (2016). Dynamics of environmental and financial performance: The case of greenhouse gas emissions. *Organization & Environment*, *28*(4), 374-393. https://doi.org/10.1177/1086026615620238

Demircioğlu, E., & Ever, D. (2020). A case study for carbon cost determination in iron and steel company. *Journal of Business Research-Turk*, *12*(1), 649–662. https://doi.org/10.20491/isarder.2020.868

Desai, R., Raval, A., Baser, N., & Desai, J. (2022). Impact of carbon emission on financial performance: Empirical evidence from India. *South Asian Journal of Business Studies*, *11*(4), 450-470. https://doi.org/10.1108/SAJBS-10-2020-0384

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *13*(3), 253-263. https://doi.org/10.1080/07350015.1995.10524599

Elijido-Ten, E. (2007). Applying stakeholder theory to analyze corporate environmental performance: Evidence from Australian listed companies. *Asian Review of Accounting*, *15*(2), 164-184. https://doi.org/10.1108/13217340710823378

Ergün, H. O. (2024). The impact of carbon emissions on financial performance: The example of the manufacturing sector *Gümüşhane University Journal of Social Sciences*, *15*(2), 383-395.

Ferrat, Y. (2021). Carbon emissions and firm performance: A matter of horizon, materiality and regional specificities. *Journal of Cleaner Production*, *329*, 129743. https://doi.org/10.1016/j.jclepro.2021.129743

Gallego-Álvarez, I., Segura, L., & Martínez-Ferrero, J. (2015). Carbon emission reduction: The impact on the financial and operational performance of international companies. *Journal of Cleaner Production*, *103*, 149-159. https://doi.org/10.1016/j.jclepro.2014.08.047

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3-42. https://doi.org/10.1007/s10994-006-6226-1

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, *51*(5), 1–42. https://doi.org/10.1145/3236009

Güneysu, Y., & Atasel, O. Y. (2022). Examining the relationship between carbon emissions and financial performance: A study on BIST100 index. *Fırat Üniversitesi Sosyal Bilimler Dergisi*, *32*(3), 1183–1193. https://doi.org/10.18069/firatsbed.1125859

Han, Y., & Wei, T. (2025). Supply chain digitization and corporate carbon emissions: A chain mediation examination based on digital transformation and green innovation. *Journal of Environmental Management*, *379*, 124825. https://doi.org/10.1016/j.jenvman.2025.124825

Hassan, A. (2019). Do renewable energy incentive policies improve the performance of energy firms? Evidence from OECD countries. *OPEC Energy Review*, *43*(2), 168-192. https://doi.org/10.1111/opec.12146

Holzinger, A., Saranti, A., Molnar, C., Biecek, P., & Samek, W. (2022). Explainable AI methods - A brief overview. In Holzinger, A., Goebel, R., Fong, R., Moon, T., Müller, KR., Samek, W. (Eds.), xxAI - Beyond Explainable AI. xxAI 2020. Lecture Notes in Computer Science. In (pp. 13-38). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-04083-2_2

Houqe, M. N., Opare, S., Zahir-ul-Hassan, M. K., & Ahmed, K. (2022). The effects of carbon emissions and agency costs on firm performance. *Journal of Risk and Financial Management*, *15*(4), 152. https://doi.org/10.3390/jrfm15040152

Imoniana, J. O., Soares, R. R., & Domingos, L. C. (2018). A review of sustainability accounting for emission reduction credit and compliance with emission rules in Brazil: A discourse analysis. *Journal of Cleaner Production*, *172*, 2045-2057. https://doi.org/10.1016/j.jclepro.2017.11.217

Karaboğa, K., Cici Karaboğa, E. N., Şekeroğlu, G., Kızıloğlu, E., & Acılar, A. M. (2024). Price prediction of dual-listed stocks with RF and LSTM algorithms: NYSE and BIST comparison. *Mathematical Modelling and Numerical Simulation with Applications*, *4*(5), 207-230. https://doi.org/10.53391/mmnsa.1577228

Karamanou, A., Kalampokis, E., & Tarabanis, K. (2022). Linked open government data to predict and explain house prices: The case of Scottish statistics portal. *Big Data Research*, *30*, 100355. https://doi.org/10.1016/j.bdr.2022.100355

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., . . . Liu, T.-Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), Advances in Neural Information Processing Systems. In (Vol. 30, pp. 3146–3154). Red Hook, NY, USA: Curran Associates, Inc.

Khatib, S. F. A., Ismail, I. H. M., Salameh, N., Abbas, A. F., Bazhair, A. H., & Sulimany, H. G. H. (2023). Carbon emission and firm performance: The moderating role of management environmental training. *Sustainability*, *15*(13), 10485. https://doi.org/10.3390/su151310485

Kim, S.-J., Bae, S.-J., & Jang, M.-W. (2022). Linear regression machine learning algorithms for estimating reference evapotranspiration using limited climate data. *Sustainability*, *14*(18), 11674. https://doi.org/10.3390/su141811674

Koç, F. Ö., & Deran, A. (2024). The impact of carbon emissions on corporate cost of debt (COD): A research on Borsa İstanbul (BIST) sustainability index. *Sosyoekonomi*, *33*(60), 33-50. https://doi.org/10.17233/sosyoekonomi.2024.02.02

Kurnia, P., Agustia, D., Soewarno, N., & Ardianto, A. (2025). The mediating role of carbon emission disclosure in the relationship between structure of corporate governance and firm performance. *Journal of Applied Accounting Research*, *26*(5), 1106–1134. https://doi.org/10.1108/JAAR-01-2023-0015

Lachuer, J., & Jabeur, S. B. (2022). Explainable artificial intelligence modeling for corporate social responsibility and financial performance. *Journal of Asset Management*, *23*(7), 619–630. https://doi.org/10.1057/s41260-022-00291-z

Laskar, N., Kulshrestha, N., Bahuguna, P. C., & Adichwal, N. K. (2022). Carbon emission intensity and firm performance: An empirical investigation in Indian context. *Journal of Statistics and Management Systems*, *25*(5), 1073–1081. https://doi.org/10.1080/09720510.2022.2067392

Le, H., & Nguyen-Phung, H. T. (2024). Assessing the impact of environmental performance on corporate financial performance: A firm-level study of GHG emissions in Africa. *Sustainable Production and Consumption*, *47*, 644-654. https://doi.org/10.1016/j.spc.2024.04.024

Lee, K. H., & Min, B. (2015). Green R&D for eco-innovation and its impact on carbon emissions and firm performance. *Journal of Cleaner Production*, *108*, 534-542. https://doi.org/10.1016/j.jclepro.2015.05.114

Lee, K.-H., Min, B., & Yook, K.-H. (2015). The impacts of carbon (CO2) emissions and environmental research and development (R&D) investment on firm performance. *International Journal of Production Economics*, *167*, 1-11. https://doi.org/10.1016/j.ijpe.2015.05.018

Li, Y., Yang, X., Du, E., Liu, Y., Zhang, S., Yang, C., … Liu, C. (2024). A review on carbon emission accounting approaches for the electricity power industry. *Applied Energy*, *359*, 122681. https://doi.org/10.1016/j.apenergy.2024.122681

Liu, Z., Sun, T., Yu, Y., Ke, P., Deng, Z., Lu, C., & Ding, X. (2022). Near-real-time carbon emission accounting technology toward carbon neutrality. *Engineering*, *14*, 44-51. https://doi.org/10.1016/j.eng.2021.12.019

López, O. A. M., López, A. M., & Crossa, J. (2022). Support vector machines and support vector regression. In O. A. M. López, A. M. López, & J. Crossa (Eds.), Multivariate statistical machine learning methods for genomic prediction. In (pp. 337–378). Cham, Switzerland: Springer International Publishing.

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., & Lee, S. I. (2019). Explainable AI for trees: From local explanations to global understanding. *arXiv preprint arXiv:1905.04610*. https://doi.org/10.48550/arXiv.1905.04610

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), Advances in Neural Information Processing Systems. In (Vol. 30, pp. 4765–4774). United States: Curran Associates, Inc.

Mastelini, S. M., Nakano, F. K., Vens, C., & De Leon Ferreira, A. C. P. (2022). Online extra trees regressor. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(10), 6755-6767.

McWilliams, A., & Siegel, D. S. (2010). Creating and capturing value: Strategic corporate social responsibility, resource-based theory, and sustainable competitive advantage. *Journal of Management*, *37*(5), 1480-1495. https://doi.org/10.1177/0149206310385696

Miah, M. D., Hasan, R., & Usman, M. (2021). Carbon emissions and firm performance: Evidence from financial and non-financial firms from selected emerging economies. *Sustainability*, *13*(23), 13281. https://doi.org/10.3390/su132313281

Nichita, E. M., Nechita, E., Manea, C. L., Irimescu, A. M., & Manea, D. (2021). Are reported greenhouse gas emissions influencing corporate financial performance. *Journal of Accounting and Management Information Systems*, *20*(4), 585-606.

Onat, N. C., Mandouri, J., Kucukvar, M., Kutty, A. A., & Al-Muftah, A. A. (2025). Driving sustainable business practices with carbon accounting and reporting: A Holistic framework and empirical analysis. *Corporate Social Responsibility and Environmental Management*, *32*(2), 2795–2814.

Orhan, U., Kilinc, E., Albayrak, F., Aydin, A., & Torun, A. (2022). Ultrasound penetration-based digital soil texture analyzer. *Arabian Journal for Science and Engineering*, *47*(8), 10751-10767. https://doi.org/10.1007/s13369-022-06766-w

Pan, X., & Liu, S. (2024). The development, changes and responses of the European Union carbon border adjustment mechanism in the context of global energy transition. *World Development Sustainability*, *4*, 100148. https://doi.org/10.1016/j.wds.2024.100148

Rathod, S., Saha, A., Patil, R., Ondrasek, G., Gireesh, C., Anantha, M. S., & Sundaram, R. M. (2021). Two-stage spatiotemporal time series modelling approach for rice yield prediction & advanced agroecosystem management. *Agronomy*, *11*(12), 2502. https://doi.org/10.3390/agronomy11122502

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206-215. https://doi.org/10.1038/s42256-019-0048-x

Sakın, İ., & Kefe, İ. (2024). The impact of carbon emissions on firms' financial performance: An application in BIST sustainability index. *Trends in Business and Economics*, *38*(1), 39-47.

Sharma, D., Bhattacharya, S., & Thukral, S. (2019). Resource-based view on corporate sustainable financial reporting and firm performance: Evidences from emerging Indian economy. *International Journal of Business Governance and Ethics*, *13*(4), 323-344. https://doi.org/10.1504/IJBGE.2019.099565

Shui, X., Zhang, M., Wang, Y., & Smart, P. (2025). Do climate change regulatory pressures increase corporate environmental sustainability performance? The moderating roles of foreign market exposure and industry carbon intensity. *British Journal of Management*, *36*(1), 223-239. https://doi.org/10.1111/1467-8551.12841

Solikhah, B., Yulianto, A., & Suryarini, T. (2020). Legitimacy theory perspective on the quality of carbon emission disclosure: Case study on manufacturing companies in Indonesia Stock Exchange. *IOP Conference Series: Earth and Environmental Science*, *448*(1), 012063. https://doi.org/10.1088/1755-1315/448/1/012063

Srivalli, K. S., & Sumanthi, D. (2025). *Enhancing financial risk assessment through explainable ai: A shap-based approach for transparent decision-making*. SSRN Scholarly Paper No. 5190418, Social Science Research Network.

Thompson, E. K., Buertey, S., & Kim, S.-Y. (2025). How important is corporate social responsibility for corporate financial performance?: A machine learning prediction and model interpretability approach. *Business Ethics, the Environment & Responsibility*. https://doi.org/10.1111/beer.12820

Tong, Z., Li, B., & Yang, L. (2025). Digital transformation, carbon performance and financial performance: Empirical evidence from the Chinese Stock Market. *Environment, Development and Sustainability*, *27*(7), 17129-17146. https://doi.org/10.1007/s10668-024-04564-y

Tonnarello, F., Vermiglio, C., Migliardo, C., & Naciti, V. (2025). The impact of EU taxonomy for sustainable activities on european utilities' performance. *Business Strategy and the Environment*, *34*(3), 2848–2862. https://doi.org/10.1002/bse.4128

Topuz, K., Bajaj, A., Coussement, K., & Urban, T. L. (2025). Interpretable machine learning and explainable artificial intelligence. *Annals of Operations Research*, *347*(2), 775-782. https://doi.org/10.1007/s10479-025-06577-w

Tuesta, Y., Naranjo, C. S. C., & Ripoll, F. V. (2021). Carbon management accounting and financial performance: Evidence from the European Union emission trading system. *Business Strategy and the Environment*, *30*(2), 1270–1282. https://doi.org/10.1002/bse.2683

TUIK. (2025). *TUIK institutional*. Turkey: Turkish Statistical Institute.

Uğurlu, U., Öksüz, İ., & Taş, O. (2018). Electricity price forecasting using recurrent neural networks. *Energies*, *11*(5), 1255. https://doi.org/10.3390/en11051255

United Nations. (2025). *The Paris agreement*. United States: United Nations.

World Meteorological Organization (WMO). (2024). *Record carbon emissions highlight urgency of global greenhouse gas watch*. Geneva, Switzerland: World Meteorological Organization.

Wu, Y., Li, F., Shu, H., Li, S., Cui, L., Tan, M., ... Wei, X. (2025). Identifying the key influencing factors of psychological birth trauma in primiparous women with interpretable machine learning. *International Journal of Nursing Sciences*, *12*(3), 253-260. https://doi.org/10.1016/j.ijnss.2025.04.008

Yıldırım, H. H., Rençber, Ö. F., & Yıldırım, C. Y. (2024). Ranking the determinants of financial performance using machine learning methods: An application to BIST energy companies. *Mathematical Modelling and Numerical Simulation with Applications*, *4*(5-Special Issue: ICAME'24), 165-186. https://doi.org/10.53391/mmnsa.1594426

Yildirim, M., Okay, F. Y., & Özdemir, S. (2024). A comparative analysis on the reliability of interpretable machine learning. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, *30*(4), 494-508.

Zhang, C. A., Cho, S., & Vasarhelyi, M. (2022). Explainable artificial intelligence (XAI) in auditing. *International Journal of Accounting Information Systems*, *46*, 100572. https://doi.org/10.1016/j.accinf.2022.100572

Zhang, J., Ma, X., Zhang, J., Sun, D., Zhou, X., Mi, C., & Wen, H. (2023). Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model. *Journal of Environmental Management*, *332*, 117357. https://doi.org/10.1016/j.jenvman.2023.117357

Zhang, Z., Wu, C., Qu, S., & Chen, X. (2022). An explainable artificial intelligence approach for financial distress prediction. *Information Processing & Management*, *59*(4), 102988. https://doi.org/10.1016/j.ipm.2022.102988