JOURNAL OF CULTURAL ANALYSIS AND SOCIAL CHANGE

Journal of Cultural Analysis and Social Change,
2025, 10(3), 2042-2056
ISSN: 2589-1316

Predicting Corporate ESG Scores from Financial Performance and Environmental Indicators: A Machine Learning Framework

Olfa CHAOUECH[1*]

[1]*Faculty of Economic Sciences and Management of Tunis, University of Tunis El Manar, Tunisia*

*Corresponding Author: chaouecholfa@gmail.com

Citation: CHAOUECH, O. (2025). Predicting Corporate ESG Scores from Financial Performance and Environmental Indicators: A Machine Learning Framework. *Journal of Cultural Analysis and Social Change*, *10*(3), 2042–2056. https://doi.org/10.64753/jcasc.v10i3.2710

Published: December 02, 2025

ABSTRACT

As investors, regulators, and the public increasingly emphasize sustainable investment amid growing climate concerns, the accurate prediction of Environmental, Social, and Governance (ESG) metrics has become a crucial complement to traditional assessment methods. This study analyzes 1,000 companies across nine industries and seven regions between 2015 and 2025 to predict overall ESG scores using key financial and environmental indicators. To ensure robust predictive performance, a diverse set of machine learning algorithms—including Linear Regression, Random Forests, and four boosting models (AdaBoost, LightGBM, XGBoost, and CatBoost)—was employed. To address potential bias in panel data, a panel-aware machine learning framework incorporating GroupKFold cross-validation was implemented. The results show that boosting algorithms consistently outperform traditional linear approaches in predicting ESG scores. Among them, CatBoost achieved the best overall performance, with the lowest RMSE (4.608), MAE (2.222), and MSE (21.234), and the highest R^2 (0.913), indicating strong predictive accuracy. Overall, this study presents an innovative and transferable framework for predicting ESG scores, thus contributing to both empirical research and quantitative modeling practices. Furthermore, it advances the sustainability field by providing a machine learning–based application that enables companies to predict their ESG scores in real time.

Keywords: ESG, Machine Learning, Boosting Algorithms, Sustainable Development, Predictive Modeling.

INTRODUCTION

The classical economic perspective posits that a firm's primary objective is to maximize profits for its owners, a notion deeply rooted in Adam Smith's concept of the "invisible hand" (Smith, 1776) and later crystallized by Milton Friedman's shareholder value theory (Friedman, 1970). However, despite its enduring influence, this doctrine has faced increasing challenge from those advocating that corporations should serve a broader societal purpose. This shift is codified in the stakeholder theory, which emphasizes that firms must consider a wider range of actors—including employees, customers, local communities, suppliers, and the natural environment—in their decision-making processes (Freeman, 1984; Donaldson & Preston, 1995). In line with this expanded view of corporate responsibility, global initiatives and investor demand have driven the integration of Environmental, Social, and Governance (ESG) factors into mainstream finance. The launch of the United Nations Global Compact in 2000, which encourages firms to align operations with principles on human rights, labor, the environment, and anti-corruption, marked a key turning point. Concurrently, the "Who Cares Wins"[1] initiative popularized ESG integration, emphasizing that firms prioritizing sustainability tend to exhibit greater resilience, operational efficiency, and stakeholder loyalty (Eccles et al., 2014; Friede et al., 2015). Investor demand for sustainability

[1] United Nations Global Compact. (2004). Who Cares Wins.

information has since exploded, demonstrated by the fact that 98% of S&P 500 companies published ESG disclosures in 2022, a dramatic increase from only 20% in 2011[2].

Existing ESG literature consistently validates the financial materiality of ESG performance, particularly during periods of market stress. While earlier studies documented the safe-haven qualities of ESG investments in past crises (e.g., Nofsinger & Varma, 2014; Lins et al., 2017), compelling evidence from the COVID-19 shock reinforces this argument. Albuquerque et al. (2020) demonstrated that firms with stronger environmental and social profiles earned higher returns and faced lower volatility in early 2020. This performance advantage was also observed at the portfolio level (Díaz et al., 2021), Broadstock et al. (2021) showing that ESG funds in both developed and emerging markets maintained superior drawdown protection throughout the pandemic. However, translating this robust financial evidence into scalable, reliable investment decisions faces considerable practical obstacles. Despite the widely accepted financial benefits of ESG integration, the process of measuring and assessing ESG performance remains fraught with systemic challenges. This primarily stems from the complexities inherent in the data-generation process. First, the manual collection of data is often prone to human error and highly time-consuming (Segal, 2024). Second, the level of reporting standardization remains critically low. Because ESG disclosure is still largely voluntary (or semi-mandatory) in many jurisdictions, publicly listed firms can choose among various regulatory frameworks and reporting standards. This discretion leads to inconsistencies that complicate comparative analysis for stakeholders (Russell Investments, 2023). Consequently, the authenticity and credibility of ESG reports are frequently questioned due to limited transparency, and the publicly available data often fail to meet investor expectations (Zhao et al., 2025). Furthermore, a significant portion of companies still fail to provide relevant ESG indicators (ESG Book, 2023). In addition to data collection issues, there are the methodological complexities inherent in assessing and scoring ESG performance. Rating systems must determine both the scoring mechanisms and the weighting schemes for each component (Environmental, Social, and Governance). Recent research highlights that aggregated ESG scores often mask substantial variation at the sub-category level, and that the weighting and aggregation assumptions matter significantly for investment outcomes (Linklater et al., 2024). Determining the optimal weighting structure therefore remains a major methodological challenge, particularly since these weights may require dynamic adjustment to reflect evolving market, regulatory, and technological conditions (Wang et al., 2025). These compounding data and methodological hurdles create an urgent need for advanced, automated solutions. Therefore, contemporary research has focused on the integration of artificial intelligence (AI) as a transformative tool capable of rapidly generating predictive insights from complex organizational data. Using AI to predict ESG scores presents a cost-effective alternative to traditional ESG reporting practices, which are often resource-intensive and data-heavy. Predictive models can estimate ESG performance based on a limited set of key inputs, reducing the need for exhaustive data collection and manual metric calculations. This approach streamlines sustainability monitoring while maintaining accuracy and regulatory alignment.

Research employing AI-based techniques to predict ESG scores by integrating key financial and environmental performance indicators remains limited. Most prior studies have relied primarily on time-series approaches that use past ESG scores to forecast future outcomes (Taskin et al., 2025).

Building on these transformative developments, this study makes a substantive contribution to the literature by introducing an integrated machine learning framework for ESG score prediction. The framework leverages a suite of advanced algorithms, including linear regression, random forests, and and four boosting algorithms—AdaBoost, LightGBM, XGBoost, and CatBoost—to capture complex, nonlinear relationships between financial performance and environmental indicators. By utilizing these key variables, the proposed models accurately predict overall ESG scores for previously unseen companies in real time. Furthermore, a rigorous GroupKFold cross-validation strategy is employed to enhance model generalizability and eliminate data leakage, thereby ensuring the robustness and reliability of the predictive framework.

LITERATURE REVIEW

Sustainable investing has rapidly become a central trend in finance, driven by research indicating a strong correlation between robust ESG practices and superior financial returns (Aydoğmuş et al., 2022; Bancu, 2024; da Cunha al., 2025). Specifically, studies suggest that companies demonstrating stronger environmental stewardship, higher social standards, and more robust governance tend to outperform those that do not (Yu et al., 2024; Chen et al., 2024).

Giese and Shah (2025) conducted long-term assessments of ESG ratings across global markets, meticulously mapping the mechanisms through which ESG characteristics impact financial performance. Their analysis revealed

[2] See https://www.ga-institute.com/research/ga-research-directory/sustainability-reporting-trends/2023-sustainability-reporting-in-focus/

that firms with higher ESG ratings consistently exhibited stronger long-term earnings growth, which subsequently translated into superior stock market performance, even when controlling for traditional factors like sector, region, and company size. Furthermore, a risk-based assessment showed that high-ESG firms experienced lower earnings variability and reduced stock price volatility. Ultimately, the authors demonstrated that ESG ratings positively influenced the performance of real-world ESG indices tracked by exchange-traded funds (ETFs), firmly underscoring the material financial relevance of ESG considerations. Despite the clear financial imperative demonstrated by these findings, the process of ESG data gathering and analysis presents a significant and resource-intensive challenge (Ibrahim et al. 2025). The primary obstacle is the current absence of full automation, which imposes a high manual burden across the industry (Zou et al., 2025). Traditional manual methods for collecting, aggregating, and analyzing this data are inherently resource-heavy, requiring substantial investment in internal teams and complex sourcing of information from multiple, often inconsistent, data providers (Rozanski, 2023; OECD, 2025; Guerrero & Viteri, 2025).

The scale of this manual effort is staggering. For instance, the London Stock Exchange Group (LSEG, 2024) reports employing over six hundred analysts dedicated to manual ESG data collection across more than 870 indicators. Similarly, the OECD (2025) highlights that institutional investors often spend nearly half a million dollars annually on external ESG data services, while the Financial Reporting Council (FRC, 2023) noted that many investors maintain dedicated ESG teams solely responsible for aggregating and standardizing data from a wide range of sources. This resource-heavy and highly manual process—which requires hundreds of analysts and access to numerous data sources—underscores a pressing need for a fundamental shift. In response to this significant data challenge, the rise of Artificial Intelligence (AI) has begun to reshape the landscape of sustainability assessment over the past two decades. What was once a highly manual, resource-intensive process is now evolving into a data-driven and automated practice (Zou et al., 2025; FRC, 2023).

The integration of AI with advanced data analytics represents a transformative step toward automating ESG evaluation, enabling organizations to process vast, complex datasets with greater speed, accuracy, and consistency (Zeng et al., 2024). Building upon this automation is the emerging field of ESG forecasting. By leveraging machine learning models to predict future ESG performance, forecasting offers a forward-looking, scalable, and cost-effective alternative to traditional, backward-looking measurement approaches (Sariyer et al., 2024). As Giese & Shah (2025) suggest, this provides decision-makers with timely insights into a company's future sustainability outcomes. The academic literature has increasingly validated the power of machine learning in addressing the complexity of ESG data. For example, Garcia et al. (2020) demonstrated the utility of a rough set model for analyzing clustering scenarios of publicly traded European companies, highlighting the importance of financial variables in predicting ESG performance.

Further solidifying this link, D'Amato et al. (2021) investigated whether structural financial data could be used to predict firms' ESG ratings. By applying the Random Forest machine learning algorithm to balance sheet and income statement items, they found that traditional financial statements serve as an effective tool to explain a firm's ESG score. Extending this work into market performance, Suprihadi and Danila (2024) explored the use of machine learning—combining algorithms like Support Vector Machine, Random Forest, and Long Short-Term Memory—for forecasting ESG stock indices, demonstrating that non-linear and ensemble-based models yield more accurate predictions that reflect the complex nature of ESG-related financial data.

Other researchers have focused on identifying the specific factors that drive ESG scores Rozanski et al. (2024) developed a multilayer perceptron (MLP) artificial neural network to examine the relationship between ESG disclosures and ESG scores. Their model achieved a strong predictive power (R^2= 79%) and revealed that the social and environmental dimensions exerted the most substantial influence on overall scores, concluding that machine learning is superior to linear models for capturing the complexity of ESG determinants.

Similarly, Taskin et al. (2025) demonstrated the feasibility of ESG prediction by solely using historical ESG scores with machine learning, proving that past ratings can provide accurate and low-cost forecasts, particularly in emerging markets. However, a limitation remains: their reliance on ESG history alone overlooks the crucial contribution of financial and environmental indicators, underscoring the ongoing research effort to develop more comprehensive predictive frameworks that integrate various data streams for a holistic view of future ESG performance. The aim of this work is to explore and evaluate the application of various machine learning regression models in predicting corporate ESG scores, with particular emphasis on advancing sustainable development through AI-driven insights. The analysis examines the performance of multiple regression algorithms—including Linear Regression, Random Forest Regression, Adaptive Boosting, Light Gradient Boosting Machine (LGBM), Extreme Gradient Boosting, and CatBoost —to identify the model that demonstrates superior predictive accuracy in capturing the underlying patterns and nuances of ESG scores. A comprehensive evaluation of key performance metrics, such as Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE) and R-squared, is conducted to facilitate effective model comparison and selection. Extending previous research, this study further integrates financial and environmental performance indicators to enhance ESG score prediction.

Moreover, by employing GroupKFold cross-validation, the study mitigates data leakage and ensures a more rigorous assessment of model generalizability. The insights derived from this research contribute to a deeper understanding of predictive modeling for corporate ESG scores and its implications for sustainable development and economic policy.

METHODOLOGY

This section examines a range of machine learning algorithms specifically adapted for regression analysis, aimed at uncovering patterns within data and predicting continuous target variables. In this study, a diverse ensemble of regression models was implemented, including Linear Regression (LR), Random Forest (RF), AdaBoost, Light Gradient Boosting Machine (LGBM), Extreme Gradient Boosting (XGBoost), and CatBoost. Although many of these algorithms can also be applied to classification problems, the present analysis focuses exclusively on their regression capabilities to forecast firms' ESG scores. The models were trained and validated on a balanced panel dataset comprising 1,000 firms over the period 2015–2025. The proposed methodology ensures that prediction performance is assessed under realistic forecasting conditions, where models are trained only on past data and tested on unseen companies.

The pipeline begins by cleaning the dataset through outlier removal using the IQR method. Missing values are then imputed using the KNN imputer for numeric features and the mode for categorical ones, ensuring data completeness for modeling. To capture temporal dependencies within each company, one-year lag features were generated for both the target variable and the selected numerical predictors. Year dummy variable was included to capture time-specific effects. After feature engineering, the dataset was dynamically analyzed to identify numerical and categorical variables, which were then preprocessed using a scalable ColumnTransformer. Numerical features were standardized with a StandardScaler, while categorical features were one-hot encoded. The model was trained and evaluated using GroupKFold cross-validation, ensuring that each company was entirely contained within either the training or test set in each fold. This approach prevents data leakage across both time and entities, thereby providing a robust and panel-aware evaluation of model performance.

Machine Learning Regression Models

Machine learning regression models utilize large datasets to capture complex patterns and generate continuous predictions. They have been extensively applied across domains such as finance, sustainability, and environmental forecasting. In this study, a suite of regression algorithms—including Linear regression, Random Forest, AdaBoost, LGBM, XGBoost, and CatBoost was employed to predict ESG performance.

Linear regression (LR), is a fundamental predictive modeling technique that establishes a linear relationship between input features and a continuous outcome variable. In its multiple form, known as multiple linear regression (MLR), the model allows us to examine how several independent variables $(X_1, X_2, \ldots., X_k)$ simultaneously influence a dependent variable (Y). The computational task of MLR involves fitting a straight line (or hyperplane in higher dimensions) to a set of observed data points. The most commonly employed estimation method is the ordinary least squares (OLS) approach, which adjusts the regression parameters to minimize the sum of squared differences between the observed values and the values predicted by the model.

The general form of the regression equation is expressed as:

$$LR: Y_{it} = \beta_0 + \sum_{j=1}^{k} \beta_j X_{it} + \varepsilon_{it} \qquad (1)$$

In this framework, β_0 represents the intercept (constant term), β_j (for $j = 1, \ldots, k$) denotes the parameters associated with each input feature X_i, and ε_i is the error term captures the unexplained variation.

The Random Forest (RF) algorithm is an ensemble regression method that aggregates the predictions of multiple decision trees to enhance accuracy and robustness. Let $X_{it} \in \mathbb{R}^p$ denote the feature vector for firm i at time t, where p is the number of predictors, and let $Y_{it} \in \mathbb{R}$ represent the corresponding ESG score. A Random Forest constructs K regression trees, each trained on a bootstrap sample of the original dataset and a randomly selected subset of predictors. The prediction from the $k - th$ tree is denoted as $h_k(X_{it})$. The ensemble prediction for firm i at time t is then obtained by averaging the individual tree predictions:

$$RF: \widehat{ESG_{it}} = \frac{1}{K} \sum_{k=1}^{K} h_k\,(X_{it}) \qquad (2)$$

To enhance predictive performance, this study employs four boosting-based ensemble regressors: AdaBoost, LGBM, XGBoost, and CatBoost. Although all four algorithms follow the general principle of boosting (sequentially combining weak learners to form a strong predictor), they differ in their specific mechanisms and areas of strength.

The Adaptive Boosting (AdaBoost) Regressor is an ensemble learning model that combines multiple base learners to improve predictive performance compared to a single learner. Originally proposed by Freund and

Schapire (1997), AdaBoost is a classical ensemble learning algorithm that enhances the accuracy of weak learners by adaptively adjusting sample weights.

Initially developed for classification and later extended to regression tasks, AdaBoost iteratively combines weak learners—such as decision stumps or simple regressors—into a single, strong predictive model. At each iteration, sample weights are updated to emphasize harder-to-predict cases, enabling subsequent learners to focus on challenging examples.

For regression problems, the ensemble prediction after N iterations is expressed as:

$$F(X) = \sum_{j=1}^{N} \alpha_j h_j(X) \qquad (3)$$

where $h_j(X)$ denotes the j^{th} weak regressor, and α_j represents its corresponding weight, reflecting its contribution to the final ensemble prediction. AdaBoost performs well on complex datasets and can achieve high predictive accuracy even when individual base regressors are relatively weak.

The Light Gradient Boosting Machine (LightGBM), introduced by Ke et al. (2017), is an open-source gradient boosting framework based on decision-tree learning. It is widely adopted due to its high efficiency, scalability, and ability to handle both large-scale and sparse datasets. LightGBM improves computational speed and memory efficiency through two key innovations: Gradient-based One-Side Sampling (GOSS) and Exclusive Feature Bundling (EFB).

GOSS retains all instances with large gradients and randomly samples those with small gradients, ensuring that the most informative samples are preserved. EFB, on the other hand, bundles sparse features into a single composite feature without losing critical information, thereby reducing dimensionality. Together, GOSS and EFB substantially accelerate training while maintaining accuracy.

Formally, LightGBM minimizes an objective function that combines a convex loss function ℓ, which measures the difference between observed and predicted values, with a regularization term Ω that penalizes model complexity:

$$\mathcal{F} = \sum_{i=1}^{n} \ell(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k), \qquad (4)$$

where n is the number of training samples, K is the number of trees, and f_k denotes the k^{th} regression tree. The regularization term is defined as:

$$\Omega(f_k) = \gamma T_k + \frac{1}{2}\lambda \sum_{j=1}^{T_k} w_{j,k}^2, \qquad (5)$$

where T_k is the number of leaves in the k^{th} tree, $w_{j,k}$ is the score of the j^{th} leaf, and γ and λ are regularization parameters that control model complexity.

The LightGBM algorithm trains trees iteratively by computing first- and second-order gradients of the loss function and selecting splits that maximize predictive gain. Unlike traditional level-wise boosting—where all nodes at the same depth are expanded simultaneously—LightGBM employs a leaf-wise growth strategy, expanding the leaf with the highest potential gain at each step. This approach improves both accuracy and efficiency, particularly for large and sparse datasets, though it increases the risk of overfitting, which is mitigated through depth constraints and regularization (Ke et al., 2017).

The Extreme Gradient Boosting (XGBoost) algorithm, proposed by Chen and Guestrin (2016), is a scalable and efficient gradient boosting framework designed for both regression and classification tasks. Belonging to the boosting family, XGBoost builds models sequentially, where each new tree corrects the errors of its predecessors. Specifically, it constructs a sequence of classification and regression trees (CARTs), with each iteration trained on the residuals of the previous model. This iterative refinement enables the model to minimize the loss function more effectively than standard boosting techniques.

Formally, let $X_{it} \in \mathbb{R}^p$ denote the predictor vector for firm i at time t, and let $Y_{it} \in \mathbb{R}$ represent the target ESG score. After D boosting rounds, the ensemble prediction is expressed as:

$$\hat{Y}_{it} = F(X_{it}) = \sum_{d=1}^{D} f_d(X_{it}), \quad f_d \in \mathcal{F} \qquad (6)$$

where $\mathcal{F}$ denotes the space of CART functions. Each f_d corresponds to an independent regression tree with structure q, which assigns each observation to a leaf node. Each leaf $j \in \{1, \dots, T\}$ is associated with a weight $w_{j,d} \in \mathbb{R}$, representing its contribution to the final prediction.

The training process minimizes an objective function that combines a differentiable convex loss function $\mathcal{L}$, measuring the difference between observed and predicted values, with a regularization term $\Omega(f)$, which penalizes model complexity:

$$\mathcal{L} = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{d=1}^{D} \Omega\,(f_d) \qquad (7)$$

where, n is the number of training samples, y_i is the observed value, and $\hat{y}_i$ is the predicted value.

The regularization term $\Omega\,(f_d)$ is defined as:

$$\Omega\,(f_d) = \gamma T_d + \frac{1}{2}\lambda \sum_{j=1}^{T_d} w_{j,d}^2 \qquad (8)$$

where T_d is the number of leaves in the d^{th} tree, $w_{j,d}$ is the score of the j^{th} leaf in the d^{th} tree, and γ and λ are the regularization parameters $(\gamma, \lambda > 0)$ that control the model complexity.

The Categorical Boosting (CatBoost) algorithm, introduced by Prokhorenkova et al. (2018), is a gradient boosting decision tree (GBDT) framework specifically designed to efficiently handle categorical and mixed-type data. Belonging to the boosting family, CatBoost constructs an ensemble of trees sequentially, where each successive tree attempts to minimize the residual errors of its predecessors. Its architecture incorporates several novel mechanisms, most notably Greedy Target-Based Statistics (Greedy TBS) and Ordered Boosting, that address key issues in traditional gradient boosting methods, such as target leakage, conditional shift, and overfitting.

Let $X_{it} \in \mathbb{R}^p$ denote the predictor vector for firm i at time t, and let $Y_{it} \in \mathbb{R}$ represent the target ESG score. After D boosting iterations, the ensemble predictor is given by:

$$\hat{Y}_{it} = F(X_{it}) = \sum_{d=1}^{D} f_d(X_{it}), \quad f_d \in \mathcal{F} \qquad (9)$$

where $\mathcal{F}$ denotes the space of regression trees.

Unlike standard GBDT algorithms that require pre-encoding of categorical features, CatBoost dynamically transforms categorical variables during training using Greedy TBS. For each categorical feature $x_{\sigma_p,t}$, the transformation is computed as:

$$x_{\sigma_p,t} = \frac{\sum_{j=1}^{p-1} \mathbb{I}\left(x_{\sigma_j,t} = x_{\sigma_p,t}\right) . Y_{\sigma_j} + c . P}{\sum_{j=1}^{p-1}\left[x_{\sigma_j,t} = x_{\sigma_p,t}\right] + c} \qquad (10)$$

where P represents a prior value (typically the mean target value) and $c > 0$ coefficient controlling the strength of the prior. This encoding reduces noise for low-frequency categories and mitigates overfitting caused by conditional shift.

Furthermore, CatBoost utilizes oblivious (symmetric) decision trees as base learners, where a uniform splitting criterion is applied across all nodes at the same depth (Kohavi & Li, 1995; Langley & Sage, 1994). This structure produces balanced trees that enhance computational efficiency and minimize overfitting. The model also incorporates a regularization parameter to penalize excessive complexity, thereby maintaining robustness and preventing the ensemble from over-fitting the training data (Nagassou et al., 2023).

Finally, CatBoost replaces conventional gradient computation with Ordered Boosting, an approach that updates model estimates sequentially on progressively accumulated data subsets. This strategy mitigates gradient bias and further strengthens the model's generalization ability. Collectively, these innovations make CatBoost a highly effective algorithm for both classification and regression tasks involving categorical and mixed-type data.

Performances Metrics

To robustly evaluate the predictive performance of the proposed machine learning regression models, a set of established evaluation metrics was employed: Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and the Coefficient of Determination (R^2). Following the framework outlined by Bourdeau et al. (2019), these complementary metrics provide a comprehensive and multidimensional assessment of model predictive performance. MAE and RMSE measure the average magnitude of prediction errors, with RMSE assigning greater weight to larger deviations. MSE further emphasizes variance in prediction accuracy, while R^2 quantifies the proportion of variance in the observed data that is explained by the model.

Together, these indicators offer a robust and balanced appraisal of both accuracy and explanatory power, enabling a nuanced comparison of model reliability and generalization performance across different regression algorithms (see Figure 4).

The Mean Absolute Error (MAE) is a widely adopted evaluation metric that measures the average magnitude of prediction errors by computing the absolute differences between predicted and observed values. Unlike other error measures, MAE accounts solely for the size of the errors, disregarding their direction (whether positive or negative). Consequently, it provides a straightforward interpretation of the model's overall accuracy by measuring how close predictions are to the observed observations on average.

Mathematically, the MAE is expressed as:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \qquad (12)$$

where N denotes the sample size, y_i represents the observed value, and $\hat{y}_i$ is the corresponding predicted value. Lower MAE values indicate higher predictive accuracy, as they reflect smaller average deviations between predicted and observed outcomes.

The Mean Squared Error (MSE) measures the average squared deviation between the model's predictions and the observed outcomes, penalizing larger errors more heavily due to the squaring operation.

Mathematically, MSE is defined as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad (13)$$

where N denotes the number of data points, y_i represents the observed value, and $\hat{y}_i$ is the predicted value. The MSE value ranges from 0 to infinity (∞), where lower values indicate higher predictive accuracy. Since the errors are squared prior to averaging, MSE is particularly sensitive to outliers, as larger deviations contribute disproportionately to the total error.

The Root Mean Square Error (RMSE) assesses the spread of errors by quantifying the standard deviation of prediction errors, also referred to as residuals. Residuals represent the deviations between the observed values and the values predicted by the model. As such, RMSE measures the extent to which these residuals are dispersed, thereby indicating how closely the data points are concentrated around the model's best-fit regression line.

The computation of RMSE involves three primary steps: (1) squaring each residual, (2) computing the mean of the squared residuals, and (3) taking the square root of this mean value. This produces a single metric that reflects the model's overall predictive accuracy, where lower RMSE values indicate a better fit. The mathematical expression for RMSE is given as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \qquad (11)$$

where N denotes the sample size, y_i represents the observed target value, and $\hat{y}_i$ is the corresponding predicted value.

The Coefficient of Determination, denoted as R^2, measures the proportion of variance in the dependent variable that is explained by the regression model. In other words, it assesses how well the model's predictions approximate the actual observed data. A R^2 value indicates that the model successfully captures a larger share of the variability in the dependent variable, reflecting superior model performance and explanatory power.

Mathematically, R^2 is expressed as:

$$R^2 = 1 - \frac{SSE}{SST} \qquad (14)$$

where SSE (Sum of Squared Errors) represents the sum of squared differences between the observed and predicted values, and SST (Total Sum of Squares) denotes the total variance in the observed data, calculated as the sum of squared deviations between each observed value and the mean of the dependent variable.

Alternatively, this relationship can be expressed as:

$$R^2 = \frac{SSR}{SST} \qquad (15)$$

where SSR (Regression Sum of Squares) quantifies the variation explained by the model. The value of R^2 ranges between 0 and 1, with values closer to 1 indicating a better model fit.

DATA DESCRIPTION

This study utilizes a balanced annual panel dataset of 1,000 firms from nine industries (see Figure A1) across seven regions (see Figure A2) over the period 2015–2025. The dataset, obtained from the Kaggle[3] platform, integrates key financial indicators—including revenue, profit margin, market capitalization, and growth rate—with environmental metrics such as carbon emissions, water usage and energy consumption, and overall ESG scores.

Table 1 presents the summary statistics for all variables using untransformed values. On average, firms report a revenue of 4,670.85, with a relatively large standard deviation (9,969.95), indicating substantial variability in firm size. The mean profit margin is 10.9%, ranging from –20% to 50%, reflecting differences in profitability across firms. Market capitalization shows a high mean value of 13,380.62 with a large spread, suggesting the presence of both small and large market players. The average growth rate is 4.83%, but with notable dispersion, as some firms experience significant decline while others grow rapidly. The mean ESG overall score is 54.62, indicating moderate sustainability performance across firms. Environmental indicators show high variability: carbon emissions average 1.27 million units, water usage 560,044 units, and energy consumption 11.7 million units, all with large standard deviations that suggest pronounced disparities in environmental impact.

Table 1. Summary statistics

Index	Revenue	Profit Margin	Market Cap	Growth Rate	ESG_ Overall	Carbon Emissions	Water Usage	Energy Consumption
count	11000.0	11000.0	11000.0	10000.0	11000.0	11000.0	11000.0	11000.0
mean	4670.85	10.9	13380.62	4.83	54.62	1271462.35	560044.15	11658393.75
std	9969.95	8.76	39922.87	9.42	15.89	5067760.3	1565686.22	50958356.02
min	35.9	-20.0	1.8	-36.0	6.3	2042.2	1021.1	5105.5

[3] https://www.kaggle.com/datasets/shriyashjagtap/esg-and-financial-performance-dataset.

25%	938.78	5.3	1098.52	-1.32	44.1	122852.95	64884.68	306916.12
50%	1902.3	10.5	3096.45	4.9	54.6	292073.45	203880.45	1221745.45
75%	4342.62	16.3	9995.5	11.0	65.6	740731.12	525187.98	5616436.72
max	180810.4	50.0	865271.7	38.0	98.8	174104721.4	52231416.4	1741047214.3

For data preprocessing, the KNN (k-nearest neighbors) imputation method was applied to handle missing values, while the interquartile range (IQR) method was used to identify and remove outliers. Following these cleaning procedures, the dataset was restructured for panel data analysis by setting a multi-index consisting of CompanyID and Year.

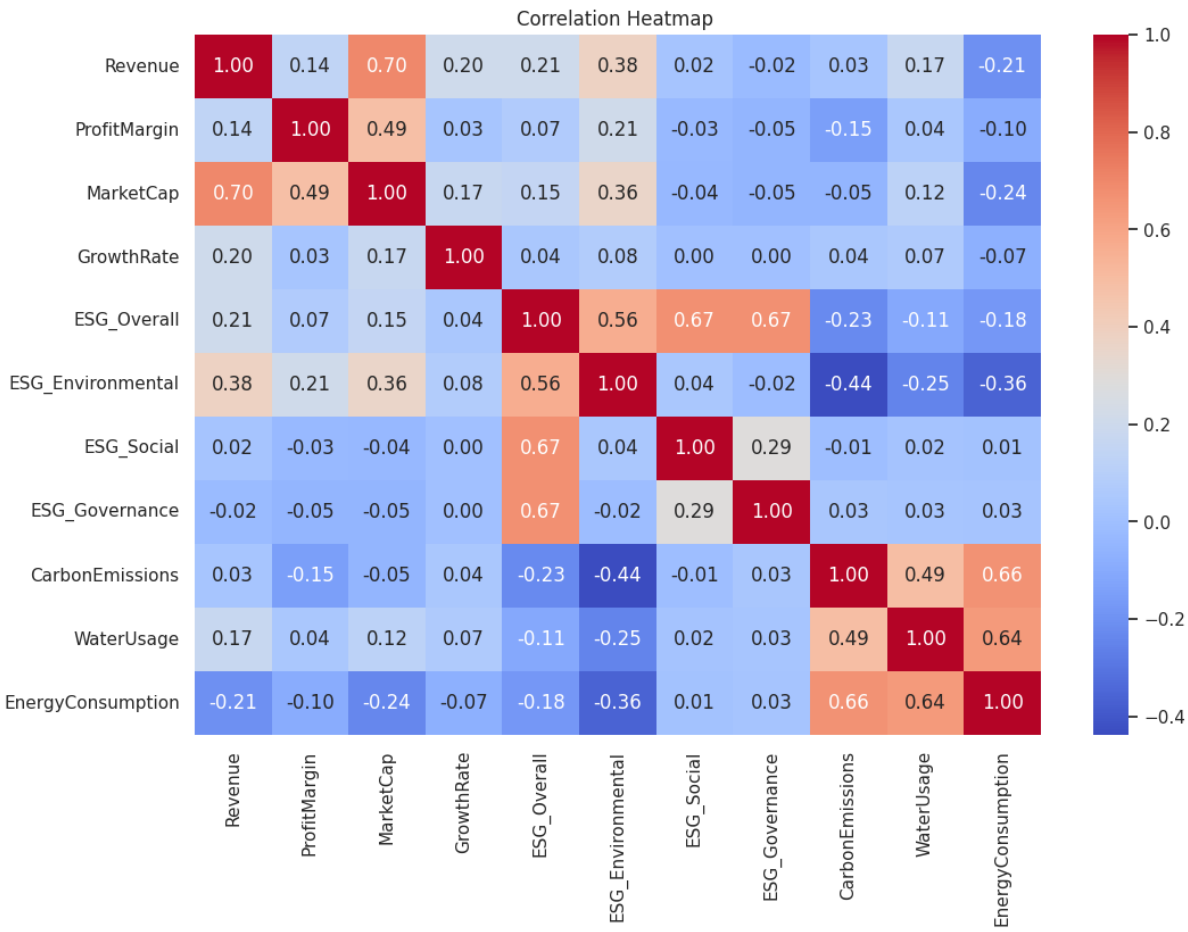

Figure 1. Correlation Heatmap.

Figure 1 displays the correlation heatmap illustrating the relationships among the variables. The results reveal generally low to moderate correlation coefficients, indicating the absence of multicollinearity.

RESULTS

In this study, a range of machine learning models was applied to predict corporate ESG overall scores based on key financial and environmental indicators, with the goal of enhancing sustainable decision-making through AI-driven insights. The algorithms evaluated include Linear Regression, Random Forest Regression, AdaBoost, LightGBM, XGBoost, and CatBoost. To ensure rigorous and unbiased model assessment, Group K-Fold cross-validation was employed, using companies as the grouping variable to prevent data leakage across folds.

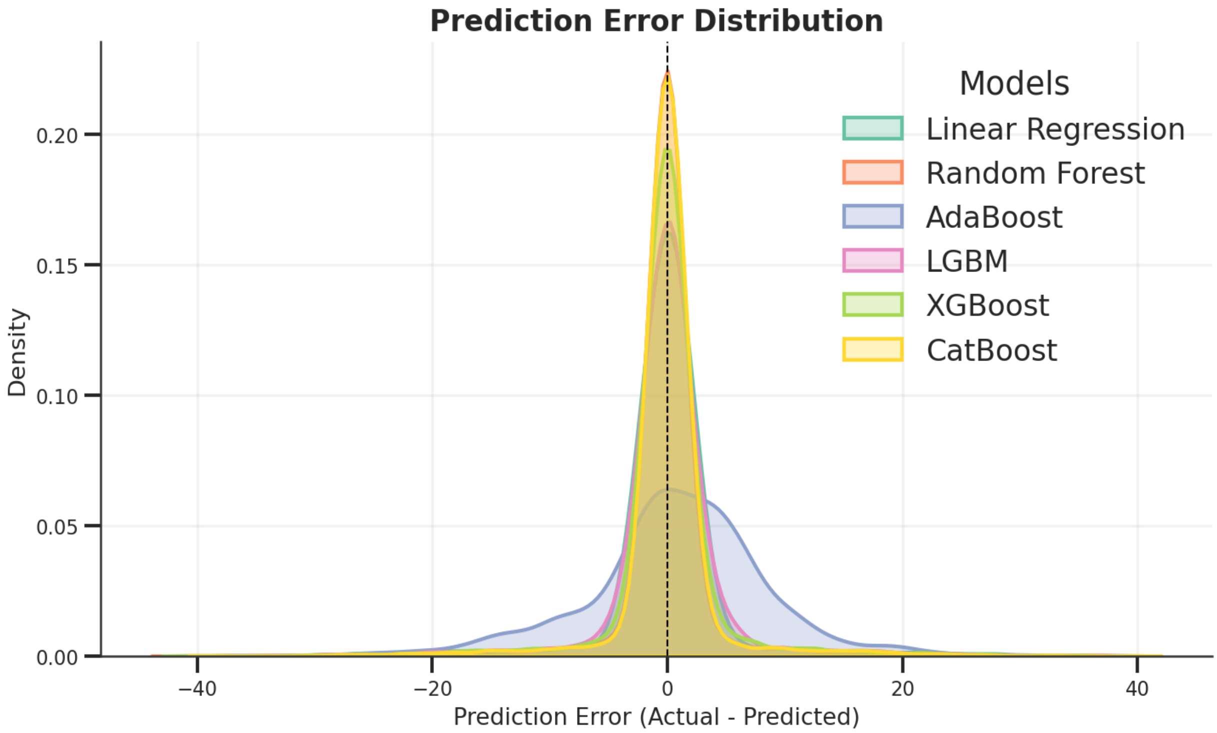

Figure 2. Error distribution plot.

The prediction error distributions of the evaluated regression models are presented in *Figure 2*. All models exhibit error distributions centered around zero, indicating minimal systematic bias. However, distinct differences emerge in the spread and sharpness of these distributions. The advanced gradient boosting algorithms—LightGBM, XGBoost, and CatBoost—demonstrate narrow, sharply peaked distributions, reflecting higher predictive accuracy and lower error variance. In contrast, AdaBoost, despite being an earlier boosting approach, displays a noticeably wider error spread, indicating greater variability and less stable performance. Linear Regression and Random Forest Regression exhibit intermediate behavior, with error distributions broader than those of modern gradient boosting models but narrower than that of AdaBoost.

Overall, these findings highlight that, with the exception of AdaBoost, the advanced gradient boosting frameworks deliver the most consistent, accurate, and reliable predictive performance among the evaluated models.

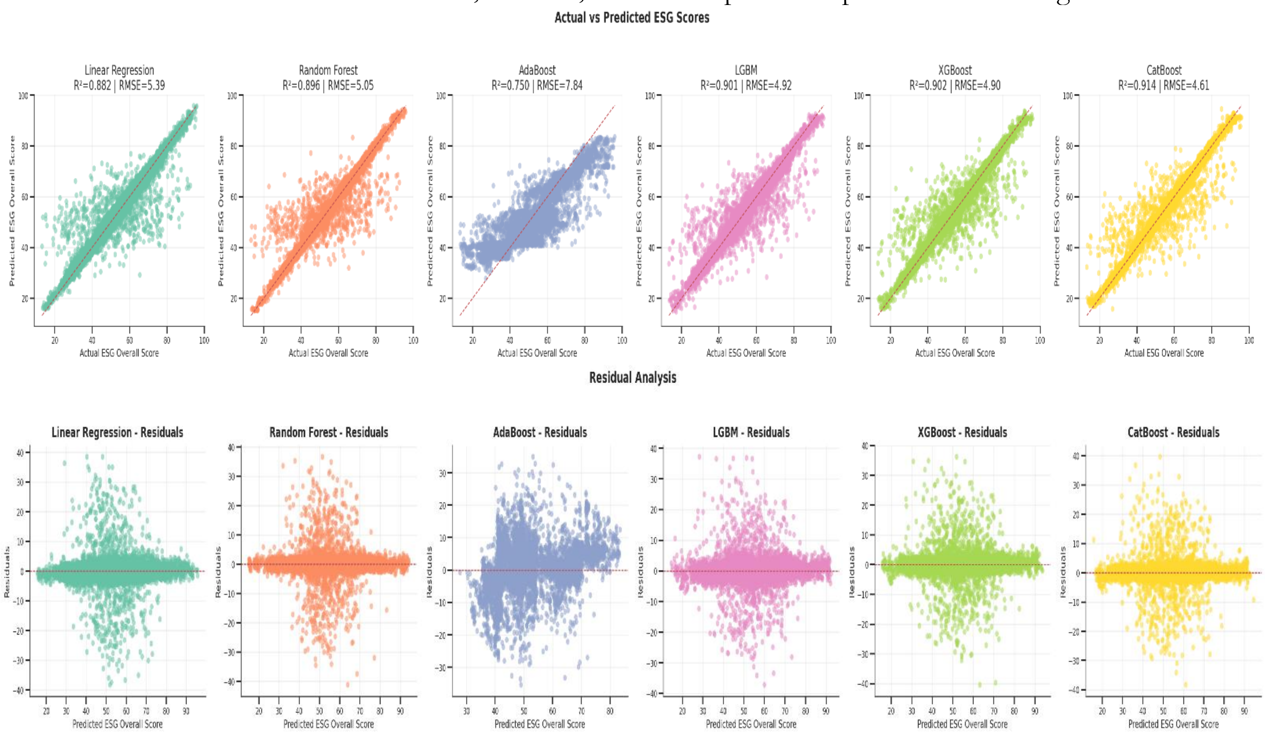

Figure 3. Scatter and Residual plots.

As illustrated in Figure 3, the close alignment of predicted versus observed values along the diagonal line across all models indicates strong overall predictive performance. The Linear Regression model demonstrates a generally linear fit but shows greater dispersion at the extremes, suggesting a limited capacity to capture nonlinear patterns. The Random Forest model improves prediction consistency, producing tighter clustering around the reference line. In contrast, AdaBoost exhibits heteroscedasticity and a noticeable bias in the lower score ranges, reflecting reduced stability and potential sensitivity to data variability. The advanced gradient boosting models—LightGBM, XGBoost, and CatBoost—show the highest fidelity to actual values, with minimal deviation from the 1:1 line, underscoring their superior ability to model the complex, nonlinear relationships underlying ESG scores.

The corresponding residual plots further emphasize these differences. The Linear Regression model displays clear heteroscedasticity, with residuals showing greater dispersion around mid-range predicted values, indicating an inability to capture nonlinear relationships. The Random Forest model mitigates this issue, yielding a more symmetric distribution of residuals around zero. However, AdaBoost continues to exhibit irregular residual patterns and a wider spread, suggesting potential overfitting or heightened noise sensitivity. In contrast, the gradient boosting algorithms, particularly LightGBM and XGBoost, produce well-centered and homoscedastic residuals, reflecting minimal systematic bias and improved calibration. Collectively, these results confirm that modern gradient boosting methods provide the most robust and accurate predictive performance for forecasting corporate ESG overall scores**.**

Table 2. Evaluation metrics.

Model	RMSE	MAE	MSE	R^2
Linear Regression	5.3885	2.7995	29.0364	0.8818
Random Forest	5.0487	2.3451	25.4899	0.8963
AdaBoost	7.8404	5.7949	61.4730	0.7498
LightGBM	4.9212	2.7175	24.2191	0.9014
XGBoost	4.8976	2.5085	23.9868	0.9024
CatBoost	4.6080	2.2224	21.2342	0.9136

As shown in Table 1, the performance comparison reveals that ensemble models outperform traditional linear methods in predicting ESG scores. CatBoost achieved the best overall results, with the lowest RMSE (4.6080), MAE (2.2224), and MSE (21.2342), alongside the highest R^2 value (0.9136), indicating strong predictive accuracy and explanatory power. Leveraging its automatic handling of categorical variables and reduced overfitting. Moreover, gradient boosting frameworks—LightGBM and XGBoost—proved highly efficient in managing large and complex datasets. both achieved high predictive power (R^2 = 0.9014 and 0.9024, respectively), confirming their strength in optimizing bias–variance trade-offs. Random Forest produced moderate results, offering a reasonable balance between bias and variance reduction. AdaBoost performed the weakest, with the largest prediction errors and lowest explanatory capacity. This relatively poor performance can be attributed to AdaBoost's sensitivity to noisy data and outliers, as the algorithm places higher weights on misclassified or high-error instances in successive iterations.

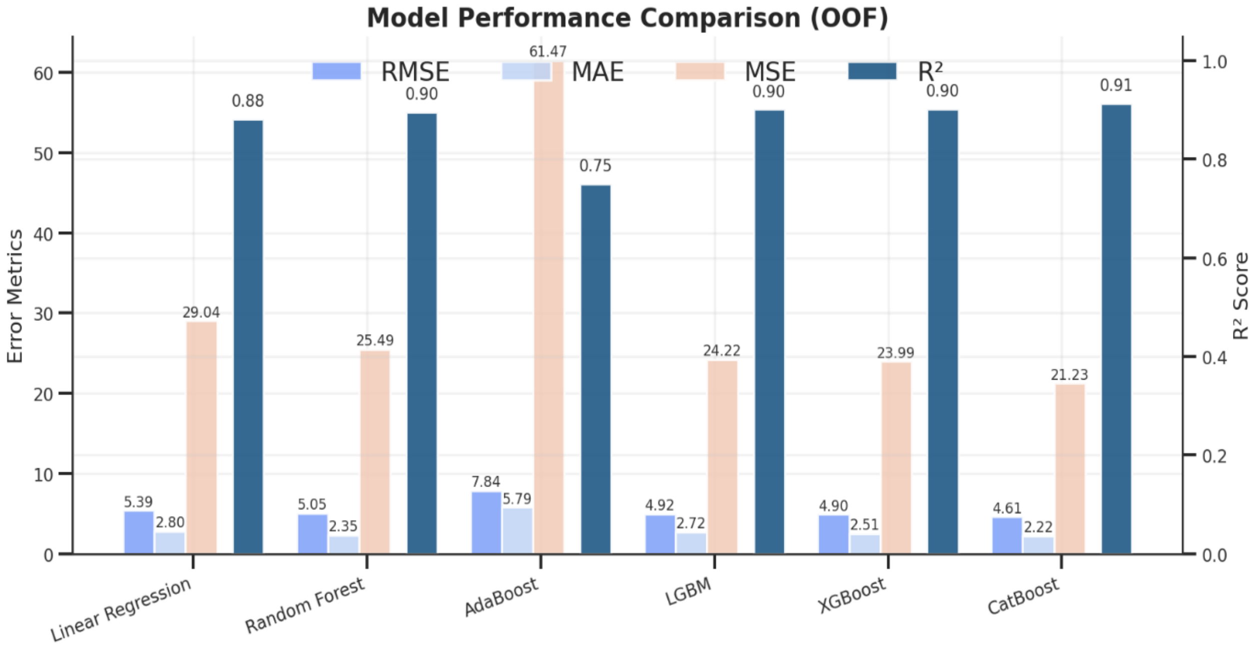

Figure 3. Model Performance Comparison.

DISCUSSION

This study demonstrates that advanced machine learning models—specifically gradient boosting frameworks such as LightGBM, XGBoost, and CatBoost—achieve superior predictive accuracy in estimating corporate ESG overall scores compared with traditional linear and simpler ensemble methods. This finding aligns with emerging evidence in the literature that highlights the effectiveness of data-driven and non-linear modeling approaches in sustainability assessment (Guerrero & Viteri, 2025; Sariyer et al., 2024). The strong performance of these adaptive learning frameworks underscores their capacity to capture the complex, multidimensional interdependencies inherent in ESG dynamics, where financial and environmental indicators interact with firm-specific characteristics and broader industry contexts to shape overall ESG outcomes. While linear models remain valuable for interpretability and exploratory correlation analysis, they appear fundamentally limited in representing such intricate non-linear relationships. The consistently narrow error distributions and high R^2 values obtained from gradient boosting algorithms provide compelling empirical support for the superiority of non-linear approaches in modeling the underlying drivers of ESG performance, corroborating prior observations that traditional econometric techniques often oversimplify these dynamics (Friede et al., 2015).

The outcomes of this research contribute to the growing intersection between artificial intelligence and sustainability analytics in two key ways. Methodologically, they validate advanced ensemble algorithms as robust and reliable tools for ESG forecasting, offering practical implications for sustainability assessment platforms and decision-support systems (Chen & Guestrin, 2016; Ke et al., 2017). Theoretically, they reinforce the notion that ESG performance arises from complex, non-linear interactions extending beyond financial fundamentals, thereby necessitating adaptive learning frameworks for comprehensive operationalization and interpretation (D'Amato et al., 2021; Taskin et al., 2025). Collectively, these findings underscore the transformative potential of AI-driven analytics to enhance both the reliability and conceptual grounding of sustainability evaluation, thereby supporting the ongoing digital transformation of corporate ESG assessment processes (Aydoğmuş et al., 2022; Bancu, 2024).

CONCLUSION

The purpose of this study is to predict corporate ESG overall scores using financial and environmental indicators for a sample of 1,000 firms over the period 2015–2025. To achieve this, the study harnesses the predictive power of diverse machine learning models to forecast ESG performance and support firms in developing more sustainable strategic pathways. The comprehensive analysis encompasses Linear Regression, Random Forest Regression, AdaBoost, LightGBM, XGBoost, and CatBoost enabling a comparative evaluation of traditional and advanced ensemble learning techniques. The findings demonstrate that the CatBoost model delivers superior predictive accuracy, as evidenced by a detailed evaluation across multiple performance metrics. Moreover, the comparative analysis of the machine learning models revealed distinct differences in predictive performance.

The gradient boosting frameworks—LightGBM and XGBoost— also proved highly effective in handling large and complex datasets, achieving strong predictive accuracy ($R^2 = 0.9014, R^2 = 0.9024$, respectively). Random Forest demonstrated moderate predictive capability, offering a reasonable balance between bias reduction and variance control. In contrast, Linear Regression yielded relatively higher error values and lower explanatory power, reflecting its limited ability to capture nonlinear relationships within the data. Among the evaluated algorithms, AdaBoost performed the weakest, exhibiting the largest prediction errors and the lowest explanatory capacity.

This study offers meaningful contributions to the advancement of predictive modeling approaches for corporate ESG overall scores. By integrating artificial intelligence with principles of financial and environmental stewardship, this research establishes a foundation for more informed, data-driven decision-making and the development of proactive sustainability strategies. In an era characterized by heightened environmental uncertainty and evolving stakeholder expectations, the adoption of advanced machine learning techniques serves as a critical enabler of transparency, resilience, and long-term value creation. Ultimately, the convergence of AI-driven analytics and sustainability governance represents a pivotal step toward fostering a more sustainable, equitable, and resilient global economy.

REFERENCES

Albuquerque, R., Koskinen, Y., Yang, S., Zhang, C., 2020. Resiliency of environmental and social stocks: an analysis of the exogenous COVID-19 market crash. Rev. Corp. Financ. Stud. 9 (3), 593–621. https://doi.org/10.1093/rcfs/cfaa011.

Aydoğmuş, M., Gülay, G., & Ergun, K. (2022). Impact of ESG performance on firm value and profitability. *Borsa Istanbul Review*, *22*, S119-S127. https://doi.org/10.1016/j.bir.2022.11.006

Bancu, A. (2024). A Meta-Analysis of ESG Disclosure and Company's Economic Performance. In *Proceedings of the International Conference on Business Excellence* (Vol. 18, No. 1, pp. 2042-2056). Sciendo. https://doi.org/10.2478/picbe-2024-0173

Bourdeau, M., Qiang Zhai, X., Nefzaoui, E., Guo, X., & Chatellier, P. (2019). Modeling and forecasting building energy consumption: A review of data-driven techniques. Sustainable Cities and Society, 48, 101533. https://doi.org/10.1016/j.scs.2019.101533

Broadstock, D.C., Chan, K., Cheng, L.T.W., Wang, X., 2021. The role of ESG performance during times of financial crisis: evidence from COVID-19 in China. Financ. Res. Lett. 38, 101716. https://doi.org/10.1016/j.frl.2020.101716

Chen, S., Song, Y., & Gao, P. (2023). Environmental, social, and governance (ESG) performance and financial outcomes: Analyzing the impact of ESG on financial performance. *Journal of environmental management*, *345*, 118829. https://doi.org/10.1016/j.jenvman.2023.118829

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

D'Amato, V., D'Ecclesia, R., Levantesi, S., 2021. Fundamental ratios as predictors of ESG scores: a machine learning approach. Decis. Econ. Financ. 44, 1087–1110. https://doi.org/10.1007/s10203-021-00364-5

da Cunha, Í. G. F., Policarpo, R. V. S., de Oliveira, P. C. S., Abdala, E. C., & do Nascimento Rebelatto, D. A. (2025). A systematic review of ESG indicators and corporate performance: proposal for a conceptual framework. *Future Business Journal*, *11*(1), 106. https://doi.org/10.1186/s43093-025-00539-1

Díaz, V., Ibrushi, D., Zhao, J., 2021. Reconsidering systematic factors during the COVID-19 pandemic – the rising importance of ESG. Financ. Res. Lett. 38, 101870. https://doi.org/10.1016/j.frl.2020.101870

Donaldson, T., & Preston, L. E. (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review, 20*(1), 65–91. https://doi.org/10.5465/amr.1995.9503271992

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science, 60*(11), 2835–2857. https://doi.org/10.1287/mnsc.2014.1984

ESG Book. (2023, June 27). *Industry-relevant ESG disclosure levels remain low despite rise in sustainability reporting. ESG Book.* https://www.esgbook.com/insights/press-releases/esg-disclosure-levels-remain-low-despite-rise-sustainability-reporting

Financial Reporting Council. (2023). *ESG data distribution and consumption*. FRC. https://www.frc.org.uk/library/digital-reporting/esg/phase-2-esg-data-distribution-and-consumption

Freeman, R. E. (1984). *Strategic Management: A Stakeholder Approach*. Boston: Pitman. https://doi.org/10.2139/ssrn.263511

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. Journal of Computer and System Sciences, 55, 119–139. https://doi.org/10.1006/jcss.1997.1504

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment, 5*(4), 210–233. https://doi.org/10.1080/20430795.2015.1118917

Friedman, M. (1970, September 13). The Social Responsibility of Business is to Increase Its Profits. *The New York Times Magazine*. https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html

García, F., Gonz´alez-Bueno, J., Guijarro, F., Oliver, J., 2020. Forecasting the environmental, social, and governance rating of firms by using corporate financial performance variables: a rough set approach. Sustain 12 (8), 3324. https://doi.org/10.3390/su12083324

Giese, G., & Shah, D. (2025). ESG Ratings in Global Equity Markets: A Long-Term Performance Review. *Journal of Impact & ESG Investing*, *6*(1). 10.3905/jesg.2025.1.132

Guerrero, S., & Viteri, J. P. (2025). What are environmental, social, and governance scores measuring? The role of outcome and impact indicators in ESG scores. Finance Research Letters, 72, 106529. https://doi.org/10.1016/j.frl.2024.106529

Ibrahim, *, Razib, N. H., & Rasel, I. H. (2025). The role of data analytics in enhancing ESG transparency in the corporate sector of Bangladesh. *Global Journal of Engineering and Technology Advances, 2025, 22(01), 081-093*. https://doi.org/10.30574/gjeta.2025.22.1.0245

Kartal, M. T., Kılıç Depren, S., Pata, U. K., Taşkın, D., & Şavlı, T. (2024). Modeling the link between environmental, social, and governance disclosures and scores: the case of publicly traded companies in the Borsa Istanbul Sustainability Index. *Financial Innovation*, *10*(1), 80. https://doi.org/10.1186/s40854-024-00619-1

Ke, G., et al. (2017) LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, 4-9 December 2017, 314s9-3157. https://dl.acm.org/doi/10.5555/3294996.3295074

Langley, P., & Sage, S. (1994). Oblivious decision trees and abstract cases. In Working Notes of the AAAI-94 Workshop on Case-Based Reasoning, Seattle, 113–117.

Linklater, R., Caprioli, S., Foschi, J., Crupi, R., & Sabatino, A. (2024). *Deconstructing ESG scores: Investing at the category score level. Journal of Asset Management, 25*(3), 222–244. https://doi.org/10.1057/s41260-024-00356-1

Lins, K.V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: the value of corporate social responsibility during the financial crisis. J. Financ. 72 (4), 1785–1824. https://doi.org/10.1111/jofi.12505

London Stock Exchange Group. (2024). *ESG scores*. LSEG. https://www.lseg.com/en/data-analytics/sustainable-finance/esg-scores

Nagassou, M., Mwangi, R. W., & Nyarige, E. (2023). A hybrid ensemble learning approach utilizing light gradient boosting machine and category boosting model for lifestyle-based prediction of type-II diabetes mellitus. *Journal of Data Analysis and Information Processing*, *11*(4), 480-511. https://doi.org/10.4236/jdaip.2023.114025

Nofsinger, J., Varma, A., 2014. Socially responsible funds and market crises. J. Bank. Financ. 48, 180–193. https://doi.org/10.1016/j.jbankfin.2013.12.016

Organisation for Economic Co-operation and Development. (2025). Behind ESG ratings: Unpacking sustainability metrics. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/02/behind-esg-ratings_4591b8bb/3f055f0c-en.pdf

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, *31*.

Ron Kohavi & Chia-Hsin Li (1995) *Oblivious Decision Trees, Graphs, and Top-Down Pruning*. https://www.ijcai.org/Proceedings/95-2/Papers/008.pdf

Rozanski, M. (2023, September 18). *Challenges of manual data collection for ESG reporting*. Osense. https://osense.ai/resources/challenges-of-manual-data-collection-for-esg-reporting

Russell Investments. (2023, October 18). *2023 manager ESG survey: Key insights*. Russell Investments. https://russellinvestments.com/uk/insights/articles/2023-esg-manager-survey

Sariyer, G., Mangla, S. K., Chowdhury, S., Sozen, M. E., & Kazancoglu, Y. (2024). Predictive and prescriptive analytics for ESG performance evaluation: A case of Fortune 500 companies. *Journal of Business Research*, *181*, 114742. https://doi.org/10.1016/j.jbusres.2024.114742

Segal, M. (2024, May 1). *83% of companies say collecting accurate data for CSRD reporting requirements will be a challenge: Workiva survey. ESG Today*. https://www.esgtoday.com/83-of-companies-say-collecting-accurate-data-for-csrd-reporting-requirements-will-be-a-challenge-workiva-survey

Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. https://www.gutenberg.org/ebooks/3300

Suprihadi, E., Danila, N., 2024. Forecasting ESG stock indices using a machine learning approach. Glob. Bus. Rev. 0 (0). https://doi.org/10.1177/09721509241234033
Taskin, D., Sariyer, G., Acar, E., & Cagli, E. C. (2025). Do past ESG scores efficiently predict future ESG performance? *Research in International Business and Finance*, *74*, 102706. https://doi.org/10.1016/j.ribaf.2024.102706
United Nations Global Compact. (2000). *The Ten Principles of the UN Global Compact*. https://www.unglobalcompact.org/what-is-gc/mission/principles
United Nations Global Compact. (2004). *Who Cares Wins*. https://www.unglobalcompact.org/library/145
Wang, Y., Li, H., & Zhang, X. (2025). *ESG rating system in sustainable finance: Challenges and suggestions for improvement. Advances in Economics, Management and Political Sciences, 3*(2), 45–58. https://www.ewadirect.com/proceedings/aemps/article/view/20659
Yu, Z., Farooq, U., Alam, M. M., & Dai, J. (2024). How does environmental, social, and governance (ESG) performance determine investment mix? New empirical evidence from BRICS. *Borsa Istanbul Review*, *24*(3), 520-529. https://doi.org/10.1016/j.bir.2024.02.007
Zeng, X., Zheng, L., & Cui, C. (2024). *Leveraging AI and machine learning for ESG data analysis and sustainable investment decision-making.* ACE, Vol. 87. https://doi.org/10.54254/2755-2721/87/20241590
Zhao, L., Li, J., & Liu, Y. (2025). *Are companies really awakening in ESG? Analyzing ESG disclosure toward corporate sustainability text data. International Journal of Disclosure and Governance, 22*(1), 12–29. https://doi.org/10.1057/s41310-025-00309-z
Zou, Y., Shi, M., Chen, Z., Deng, Z., Lei, Z., Zeng, Z., ... & Zhou, W. (2025). ESGReveal: An LLM-based approach for extracting structured data from ESG reports. *Journal of Cleaner Production*, *489*, 144572. https://doi.org/10.1016/j.jclepro.2024.144572

APPENDIX A

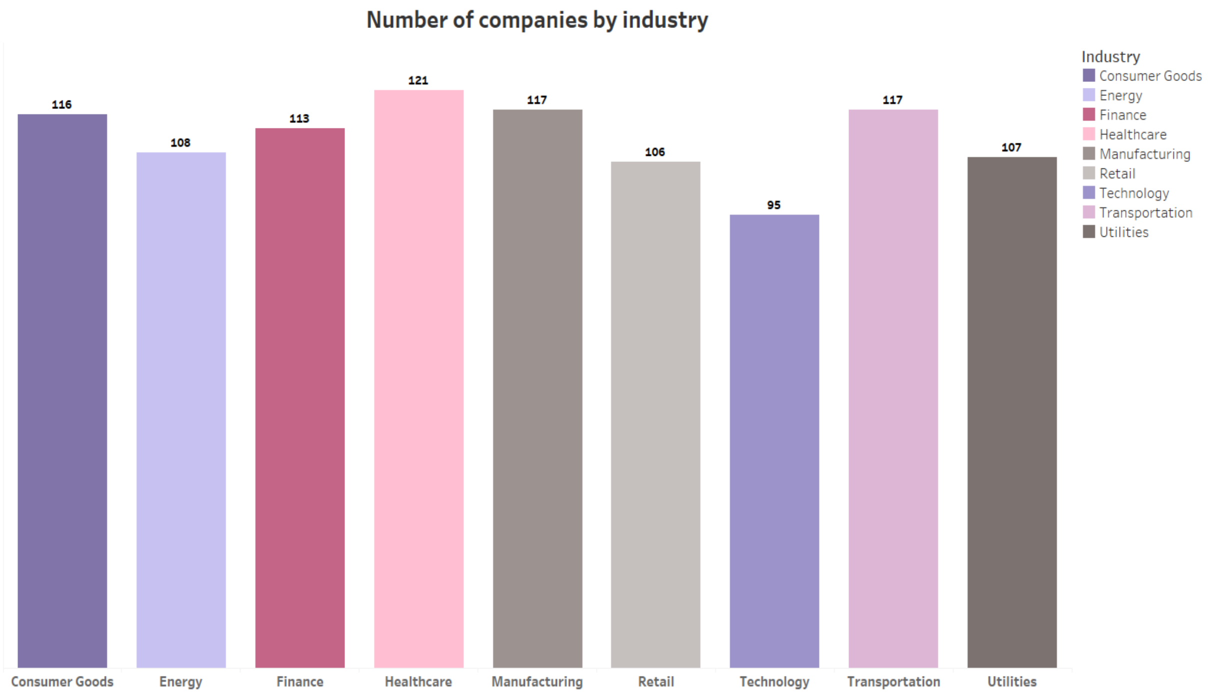

Figure A1:

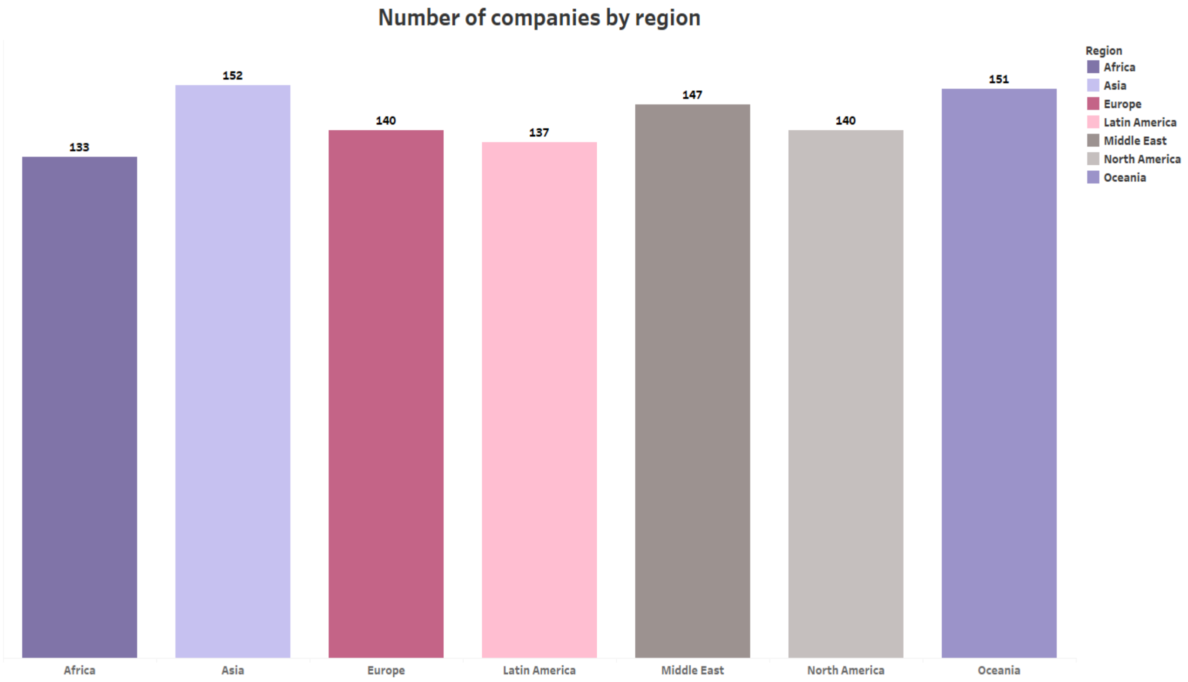
Number of companies by region
Region
Africa
Asia
Europe
Latin America
Middle East
North America
Oceania
133
152
140
137
147
140
151
Africa
Asia
Europe
Latin America
Middle East
North America
Oceania

Figure A2:

.