Journal of Sustainable Development of Energy, Water and Environment Systems

http://www.sdewes.org/jsdewes

Year 2026, Volume 14, Issue 2, 1300653

Original Research Article

Predicting Environment, Social, and Governance Scores Using Firms' Financial Indicators: A Machine Learning Regression Approach

***Mariam Alsayyad**[*1,2], **Sayed Mohamed Fadel**[3]*

[1]Department of Mathematics, Bahrain Polytechnic, PO Box 33349, Isa Town, Kingdom of Bahrain
[2]American University of Bahrain, Building 1350 Road 4225, Riffa 942, Kingdom of Bahrain.
[3]Department of Finance, Bahrain Polytechnic, PO Box 33349, Isa Town, Kingdom of Bahrain
e-mail: mariam.alsayyad@polytechnic.bh, mohamed.fadhel@polytechnic.bh

Cite as: Alsayyad M, Fadel S.M, Predicting Environment, Social, and Governance Scores Using Firms' Financial Indicators: A Machine Learning Regression Approach, J.sustain. dev. energy water environ. syst., 14(2), 1300653, 2026, DOI: https://doi.org/10.13044/j.sdewes.d13.0653

ABSTRACT

The oil and gas industry plays a pivotal role in the global energy and financial markets. With increasing concerns surrounding Environmental, Social, and Governance scores, their impact on this sector has become a growing area of focus. This study aims to forecast Environment, Social, and Governance scores in the oil and gas sector using extensive datasets comprising publicly available financial and Environment, Social, and Governance indicators of firms. The research analysed data from 497 companies within the industry over 12 years. A total of 11 machine learning algorithms were utilised to predict Environment, Social, and Governance scores, including Decision Tree, three ensemble methods (Boosting, Bagging, and Voting), XGBoost, LightGBM, Random Forest, Extreme Random Trees, linear regression, robust linear regression, and Elastic Net. The analysis incorporated a one-year lag in Environment, Social, and Governance scores and employed panel data regression techniques in machine learning. The findings demonstrated a high predictive performance, with the best R^2 value reaching 0.922. These results provide a practical framework for investors and decision-makers to evaluate a firm's Environment, Social, and Governance performance, facilitating more informed and sustainable investment decisions within the oil and gas sector.

KEYWORDS

Sustainability, Machine Learning, ESG Score Prediction, Firm Financial Performance, Data Analysis, ESG Investments, Oil and Gas Industry.

INTRODUCTION

Environmental, Social, and Governance (ESG) serves as a corporate evaluation metric that assesses companies' non-financial performance. It encompasses key concerns such as corporate governance, climate change, social inequality, workplace conditions, and broader sustainability challenges facing society [1]. Additionally, ESG factors shape how companies engage in environmental innovation, uphold human rights, manage resources efficiently, adhere to social ethics, and fulfil their product responsibilities [2]. As a result, businesses worldwide are increasingly integrating ESG considerations into their managerial decision-making processes [1]. The oil and gas industry plays a vital role in the global energy and financial markets. It remains one of the most developed and transparent energy sectors [3] and continues to be a primary energy source worldwide [4]. However, compared to sectors such as finance, insurance,

[*] Corresponding author

and retail, this industry faces significant environmental and social challenges due to the nature of its operations. Consequently, investors, governments, environmental advocates, and the general public closely monitor the oil and gas sector, given its considerable ecological impact. In response to mounting stakeholder pressure, companies within this industry have implemented more sustainable policies and regulations, aiming to transition toward a more environmentally responsible energy model [4]. ESG performance is becoming a crucial factor in investment decisions. As corporate governance and environmental concerns gain prominence, investors are increasingly considering ESG metrics as key non-financial indicators [5]. The benefits of ESG investment are evident, as prioritising ESG initiatives can improve corporate behaviour and enhance access to external financial resources [6].

Environmental, Social, and Governance – Finance Analysis Utilising Machine Learning

Researchers have employed machine learning (ML) algorithms alongside firms' financial indicators to analyse ESG performance. The authors of [7] integrated annual financial indicators with natural language data in their study, utilising models such as XGBoost, CatBoost, a Keras sequential model, the K-Nearest Neighbours (KNN) classifier, and a simple linear regressor to predict firms' annual ESG ratings. The research conducted by [8] introduced a model for ESG score prediction by training a heterogeneous ensemble model that incorporates neural networks and the Random Forest algorithm. The authors of [9] not only examined the impact of financial indicators on ESG scores but also included governance factors to provide a more comprehensive assessment. Their study, which encompassed over 30 industries in Taiwan, including the oil, electricity, and gas sectors, tested four ML models: support vector machines (SVM), Random Forest, extreme learning machine (ELM), and extreme gradient boosting. Meanwhile, the authors of [10] initially implemented three standalone machine learning models: categorical gradient-boosted trees, feedforward neural networks, and gradient-boosted trees. They later proposed two ensemble models: a neural network ensemble and a heterogeneous ensemble combining all three individual models. Their findings revealed that the heterogeneous ensemble achieved the most accurate ESG predictions, as it leveraged the strengths of each particular model to uncover unique correlations between input variables and ESG scores. [3] investigated the external environmental factors and internal corporate attributes influencing ESG performance. They applied the Lasso method for data screening and employed ML models such as SVM, KNN, neural networks, and Random Forest to predict firms' ESG performance. Their study confirmed that both internal and external factors play a crucial role in enhancing corporate ESG performance. Furthermore, [11] and [12] incorporated social and governance datasets along with unstructured news and social media content, including Twitter data, to improve ESG score prediction. To analyse this unstructured text, they utilised natural language processing (NLP), which allowed the models to differentiate between pillar-specific input data and intermediate data representations. Their results highlighted the effectiveness of NLP in capturing the full scope of ESG-related data.

Fundamental Ratios

Existing literature has explored the use of financial indicators from firms in predicting ESG scores. In the study [7], various financial metrics were analysed, including the Return on Assets (ROA), Return on Equity (ROE), net income, total assets, price-to-free cash flow ratio, long-term debt, price-to-sales ratio, debt-to-equity ratio, quick ratio, earnings per share, current ratio, price-to-book ratio, price-to-earnings ratio, Earnings Before Interest, Taxes & Depreciation (EBITDA), return on investment, revenue, total liabilities, employee turnover rate, and the independence of the chair and Chief Executive Officer (CEO) roles. Additionally, [13] included the sales growth rate in their examination. The study [6] incorporated control variables, including the asset-liability ratio, Tobin's Q, operating cash flow, and firm size. Building on this, [14] added total return and dividend yield to the analysis. At the same time,

[4] also considered a company's market value to investigate its relationship with ESG performance within the oil and gas industry.

Environmental, Social, and Governance – Score Relevance in the Oil and Gas Industry

According to [4], due to the considerable environmental impact of the oil and gas sector, ESG (Environmental, Social, and Governance) scores have gained increased relevance. This study examines the impact of ESG strategies on various financial variables, aiming to develop a comprehensive model of a firm's financial system. Specifically, it assesses the potential effects of each ESG pillar on a company's market value. The research utilises ESG metrics along with financial and economic data obtained from Refinitiv Eikon DataStream. To evaluate operational and financial performance, the study employs the ROA and ROE metrics, respectively. The researchers employed partial least squares structural equation modelling (PLS-SEM) to develop a new global ESG score and examine its impact on firms' financial performance. Their findings indicate that ESG practices have a positive effect on financial performance within the oil and gas industry.

The related study [6] examined how sustainability performance affects the Chinese crude oil market using a panel regression approach. One of the key findings is that crude oil futures prices have a significant influence on ESG performance, suggesting that price fluctuations can encourage companies to enhance their ESG ratings.

A reverse analysis of the ESG-finance relationship in the oil and gas industry, as conducted by [15], forecasted oil prices during the COVID-19 pandemic by examining stock markets, global environmental indices (ESG), and green energy resources (GER). Advanced ML algorithms, including Random Forest, LightGBM, CatBoost, XGBoost, and neural network models, were used in this study. The study [14] examined how the projected dividend growth and discount rates of Latin American (Latam) oil and gas companies are impacted by their corporate ESG performance. The researchers built their research upon the work conducted by Campbell and Shiller to compute the present value decomposition.

Research Contribution

A noticeable gap exists in the existing literature, as limited studies have focused on predicting ESG scores in the oil and gas sector by integrating financial indicators, ESG metrics, and machine learning techniques. For instance, the research by [4] was limited to data from a single year (2020) and did not assess the predictive potential of their model or incorporate machine learning methods. Likewise, the studies [6] and [14] did not utilise machine learning in their analyses. While [15] did apply machine learning algorithms, ESG scores were merely used as inputs to analyse corporate financial performance, rather than being predicted. Moreover, past research has yet to establish a reliable prediction framework for ESG scores specifically within the oil and gas industry. As a result, investors and key decision-makers face difficulties in forecasting firms' ESG performance.

Environmental, Social, and Governance – Finance Analysis Utilising Machine Learning

The numerical results obtained from models utilising regression machine learning algorithms were evaluated based on their prediction accuracy. In the research reported in [7], [8], and [10], researchers used the mean average error (MAE) and the coefficient of determination (R^2) parameters. In the study [9], two additional performance variables were introduced: root mean square error (RMSE) and mean absolute percentage error (MAPE), for comparison of the models' results. The study [11] used the RMSE and MAE to examine the models obtained for predicting ESG scores.

Theoretical Background on Financial Indicators and Environmental, Social, and Governance Prediction

This report examines the performance of ESG scores through the lens of stakeholder theory. ESG scores and investments offer insights into various aspects of the company, including its internal governance competencies, interactions with local governments, and other social responsibilities. Furthermore, companies are inclined to embrace social responsibilities, including fulfilling debt obligations [16].

Previous research has shown that within areas featuring well-established ESG scoring systems, investors incorporate high ESG performance as a key factor in making investment decisions [17]. Additionally, stakeholders, as suggested by [18], believe that companies which prioritise strong ESG practices and a positive reputation are more likely to compete effectively in the market. Consequently, successful management, recognised by stakeholders, can lead to increased investment, productivity, and consumption.

Furthermore, according to [19], firms can improve their financial performance through the indirect benefits of maintaining solid stakeholder relationships. Companies that engage in socially and environmentally responsible practices, coupled with effective governance processes, align with the interests of stakeholders, thereby enhancing their financial performance. Furthermore, the resource-based hypothesis suggests that businesses with heightened ESG disclosure (ESGD) can gain competitive advantages, and that the financial performance appears to be positively correlated to ESGD. This suggestion indicates that companies can strengthen their corporate reputation and develop internal resources by engaging in ESGD initiatives.

Additionally, [7] demonstrated that financial ratios, such as ROA and ROE, can be accurately predicted using machine learning models based on ESG ratings, thereby reinforcing the strong link between ESG performance and financial outcomes. Similarly, [3] applied machine learning methods to fuse macroeconomic and microeconomic data, revealing that internal financial indicators play a significant role in shaping corporate ESG performance. In parallel, [12] established that strong financial results are positively associated with the adoption of sustainable business models, demonstrating that companies with solid financial standing tend to achieve better ESG outcomes while simultaneously reducing investment risk and securing competitive returns.

DATA

This section outlines the data sourcing and preprocessing procedures, along with the resulting datasets and the data analysis diagram.

Environmental, Social, and Governance Scores

LSEG Data & Analytics (LSEG) is a globally recognised provider of financial data, operating in 175 countries. Beyond its role as a key source of financial information, numerous research studies have analysed its ESG disclosures, which encompass data from over 12,000 public and private companies, covering more than 80% of global market capitalisation. The methodology used to determine sustainability ratings is entirely data-driven, carefully adjusted to reflect industry-specific criteria, and refined to mitigate biases related to market capitalisation and transparency [20]. For companies in the oil and gas industry, the parameters used in ESG score calculations were compared. The environmental (E) and social (S) pillar scores are assessed in relation to firms and their industry counterparts, whereas the governance (G) pillar score is evaluated based on the country of incorporation. The total ESG score is derived from the combined scores of these three pillars. These scores are not absolute but are instead calculated relative to the broader industry and market context in which the company operates [21]. A total of 186 selected metrics contribute to the ESG scores, which are further categorised into 10 groups, as illustrated in Figure 1. Consequently, the overall scores provided by Refinitiv are utilised for this analysis, in alignment with the methodologies applied by [10], [8], and [4].

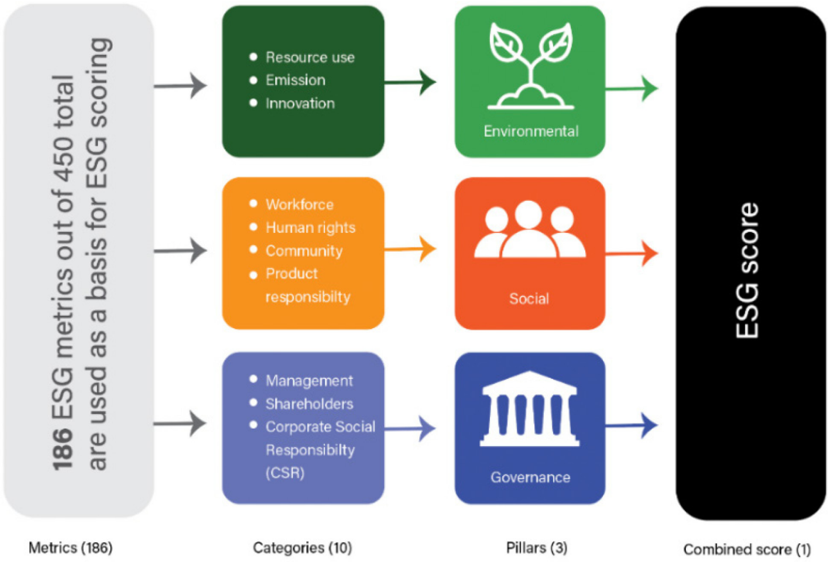

Figure 1. ESG score computation

Financial Indicators

The input variables, consisting of historical financial indicators relevant to firms in the oil and gas industry, are also obtained from LSEG. The retrieval of both financial indicators and ESG scores was carried out using two separate methodologies, Datastream and Refinitiv, both of which are well-established data sourcing methods provided by LSEG.

The data were extracted from the detailed examination of the literature. Notable differences can be observed between the financial indicators derived using the Datastream variables method and those selected via Refinitiv. These variations are primarily due to the differing data sources and the unique calculation methodologies applied to each indicator by the respective platforms. Appendix Table A1 outlines these discrepancies and highlights the indicators that were excluded from the analysis when Refinitiv was used as the data source.

Sample Size and Population

The dataset comprised companies from the oil and gas industry that had reported their ESG scores between 2010 and 2022. This selection process resulted in a total of 497 companies, whose data were acquired using both the Datastream and Refinitiv methodologies. Table 1 provides a breakdown of the number of companies from each continent included in this analysis.

Table 1. Number of companies included in the analysis

Region	Companies
America	260
Europe	83
Asia	122
Oceania	28
Africa	4
Total	497

Over the past decade and a half, the incorporation of ESG scores and related indicators into financial reporting has steadily increased within the oil and gas industry. To gain deeper insights into ESG score trends and patterns, a large sample size spanning 12 years, ranging from 2010 to 2022, was chosen for analysis. Machine learning algorithms are particularly effective in predicting outcomes from extensive datasets, as they can identify intricate and hidden relationships while improving prediction accuracy. Therefore, the comprehensive dataset covering 12 years is well-suited for ML analysis. The exclusion of ESG scores for 2023 was due to incomplete disclosures for the selected companies. Since ESG scores serve as the primary predictor, missing data were not processed to prevent introducing noise into the dataset, as such inconsistencies could compromise the accuracy and reliability of the prediction model [22].

Dummy Variables

Dummy variables were integrated into the dataset alongside ESG and financial indicators to enable a more detailed analysis. The explanatory variables considered in the study include country, firm size, and ESG scores. The purpose of dummy variables is to examine how firm size and geographic location influence ESG scores.

Country. The country dummy variable was created by identifying the headquarters location of each of the 497 companies included in the study. A total of 55 countries were represented in the dataset.

Firm size. To assess firm size, market capitalisation values were collected for all 497 companies. These were then grouped into six categories following the classification system proposed by [23], as detailed in **Table 2**. The decision to use market capitalisation as a measure of firm size aligns with the methodology of [24].

Table 2. Market capitalisation scale

Category	Market size [USD]		
Mega-cap	Equal to or greater than 200,000,000,000		
Big-cap	10,000,000,000	to	200,000,000,000
Mid-cap	2,000,000,000	to	10,000,000,000
Small-cap	250,000,000	to	2,000,000,000
Micro-cap	50,000,000	to	250,000,000
Nano-cap		up to	50,000,000

Environmental, Social, and Governance scores. Dummy variables for ESG scores were generated by categorising the continuous ESG values into three levels, based on the classification framework suggested by [7] and outlined in **Table 3**.

Table 3. ESG score classification (3 classes)

Class	Criterion
Leader	$70.143 \leq \text{ESG score} \leq 100$
Average	$40.286 \leq ESG\ score < 70.143$
Laggard	$ESG\ score < 40.286$

Data Preprocessing

Before applying machine learning algorithms, the extracted data from both sources required preprocessing. The following sections describe the preprocessing steps.

Data normalisation. Since ESG and financial indicators vary in scale, normalisation was essential to ensure that each feature contributes equally to the analysis. This step is critical, as the quality and consistency of input data directly influence the performance of machine learning algorithms. For this study, the Min-Max scaling method was employed, following the methodology outlined in [9]. This approach transforms all values to fall within a 0–1 range, helping to standardise the data and enhance model performance [25], as presented in equation (1).

$$X' = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where X' is the new data item that has been normalised, X is the old data item, X_{max}, and X_{min} are the maximum and minimum values in the data set, respectively.

Environmental, Social, and Governance score classification. Originally collected as continuous data, ESG scores were used for regression analysis. However, to expand the range of machine learning models that could be applied, ESG scores were also categorised into three classes for classification analysis. In supervised machine learning, this labelling enables classification tasks where ESG scores, serving as target variables, are divided into "leader," "average," and "laggard" groups as per work [7].

Lagged Environmental, Social, and Governance scores. To enhance the predictive performance of the models, feature engineering techniques were applied, specifically, the creation of lagged features. In this context, previous-year ESG scores were introduced as atime-lagged variable, based on the approach suggested by [26]. This accounts for time-dependent relationships between ESG scores and financial performance, helping to improve forecasting accuracy [27]. Lag represents the value of a variable from a previous time period [28].

Panel Data

This research employed panel data analysis, combining both cross-sectional and time-series dimensions. This method allows the modelling of unobserved heterogeneity and captures consistent characteristics over time, such as economic behaviour or management practices [29]. Following [30] and [31], a one-year lag between ESG scores and financial indicators was used to support this approach.

Data exclusion. Real-world datasets often contain missing, noisy, or irrelevant information, which can negatively impact the performance of models. During preprocessing, missing values in ESG and financial data were addressed. Rows or columns that were empty were excluded. Furthermore, in line with [8], any variable with over 65% missing data was removed from the dataset. As noisy data can significantly distort analytical outcomes and affect the effectiveness of normalisation techniques such as Min-Max scaling, companies with more than three missing ESG scores were excluded from both the Refinitiv and Datastream datasets to maintain data quality.

Training and testing data. In this analysis, cross-validation was employed as a resampling technique to assess model accuracy and reduce the risk of overfitting, a common issue that arises when training data is insufficient [27], [32]. This method ensured consistency by maintaining the same set of companies across all training and testing folds [33]. The default cross-validation configurations provided by MATLAB and Azure were used, with no further parameter adjustments.

Data analysis diagram. The data analysis process followed in this study is schematically shown in Figure 2. The process began with data collection, followed by multiple stages of data preprocessing. After preprocessing, the data were processed using two separate platforms, MATLAB and Azure AI/Machine Learning Studio, where ML algorithms were applied. The dataset was automatically split into training and testing subsets based on the default specifications of each platform. No parameter tuning was performed.

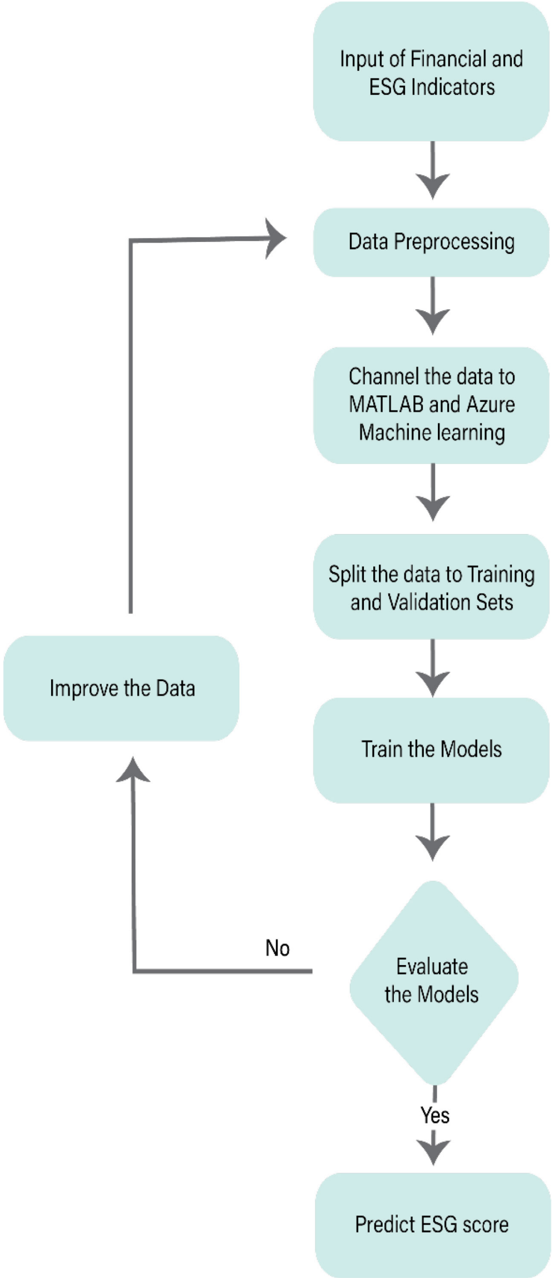

Figure 2. Data analysis diagram

Subsets. The datasets were organised into seven distinct groups. Figure 3 and Figure 4 illustrate the final sets used for both regression and classification ML analysis.

Dummy variables were integrated into the dataset alongside ESG scores and financial indicators. The explanatory variables selected for analysis included country, firm size, and ESG scores. The inclusion of dummy variables aimed to enable a comprehensive evaluation of how firm size and country influence ESG scores. Additionally, for datasets that produced the most promising results, further analyses were conducted. These included testing models without dummy variables, removing missing financial indicators, and assessing their impact, as detailed in Table 4. In total, six data sets were tested in this analysis, as shown in Figure 3 and Figure 4.

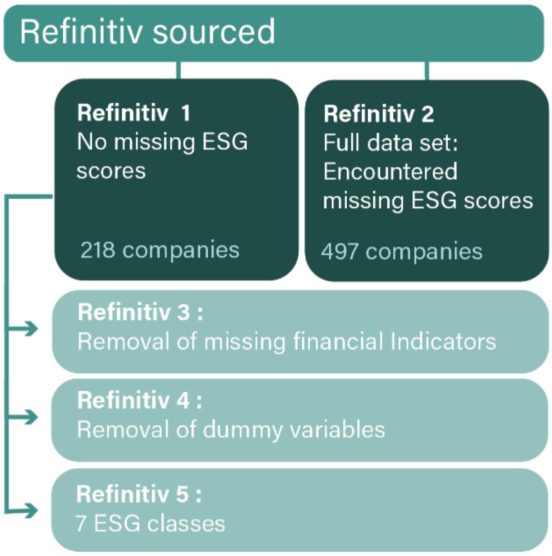

Figure 3. The four Refinitiv data sets

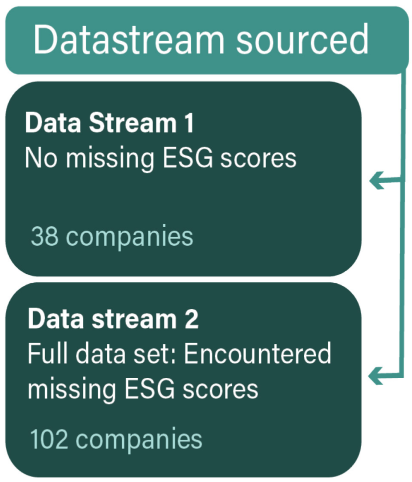

Figure 4. The two Datastream data sets

Table 4. Financial indicators with high missing values from Refinitiv

Indicator	Missing values (out of 2834)
Total Assets Actual	901
Total Liabilities	601
Return On Assets Actual	692
Tobin's Q	2514
Price to FCF per Share	1239
Price to Sales Ratio Mean	2493

DATA ANALYSIS AND FINDINGS

This section illustrates the results obtained by channelling the data to two distinct platforms for running the classification analysis: MATLAB and Azure.

MATLAB Results

The seven data sets were fed into regression learner analysis, using continuous ESG scores. The obtained R^2 values of Refinitiv 1 are provided in **Table 5**. The highest value of R^2 was found to be 0.892, obtained from the Bagged Tree (ensemble) model. R^2 is used to evaluate the effectiveness and fit of regression models; the higher the R^2 value (the closer to 1), the greater the explanation of the variance in the dependent variable (ESG scores). This is explained by the independent variables (financial indicators and ESG scores), according to [34]. Generally, a model with an R^2 value greater than 0.9 is considered reliable; hence, a value of 0.892 is slightly lower than being reliable.

Table 5. Refinitiv 1 (no missing ESG scores) regression R^2 values

Model Type	R^2
Bagged Tree	0.892
Boosted Tree	0.892
Decision Tree	0.865

To test the effect of companies with missing ESG scores, Refinitiv 2 was evaluated, and the R^2 values obtained were compared to those of Refinitiv 1, as shown in **Figure 5**. R^2 value was reduced from 0.892 to 0.812 for the ensemble Bagged Tree.

Figure 5. Comparison between R^2 values of Refinitiv 1 and 2 (encountered missing ESG scores)

In addition to evaluating the models based on the R^2 value, other performance metrics, such as MAE, MSE, and RMSE, were also analysed. **Table 6** and **Figure 6** present the RMSE, MSE and MAE values for both Refinitiv 1 and 2. According to [35], MAE and MSE are used to evaluate models by considering the distance between the predicted values and the actual training data points. MSE is more sensitive to outliers than MAE. The values obtained from the five models are promising, as the error values are small (the maximum error obtained is 0.0617 for MAE and 0.00717 for MSE). The ensemble Bagged Tree model outperforms the other two models, with an MAE value of 0.0535 and an MSE value of 0.00571. Furthermore, RMSE is commonly used to standardise the units of MSE. The distinct types of regularisation inherent to these metrics influence their relative effectiveness depending on the structure of the data [35]. The ensemble Bagged Tree outperforms other models, as it has the lowest RMSE value.

The values of MAE, MSE, and RMSE are zero if the regression model fits the data perfectly, and they are positive if the fit is not perfect. Therefore, models with lower values (closer to 0) exhibit better prediction performance [35]; accordingly, the error results obtained from the regression analysis of this dataset are excellent.

Table 6. Comparison between RMSE, MSE, and MAE values for Refinitiv 1 and 2

Model Type	RMSE		MSE		MAE	
	Refinitiv 1	Refinitiv 2	Refinitiv 1	Refinitiv 2	Refinitiv 1	Refinitiv 2
Bagged Tree	0.0756	0.1002	0.0057	0.01	0.0535	0.0674
Boosted Tree	0.0757	0.0949	0.0057	0.009	0.0553	0.0662
Decision Tree	0.0847	0.1068	0.0072	0.0114	0.0617	0.0745

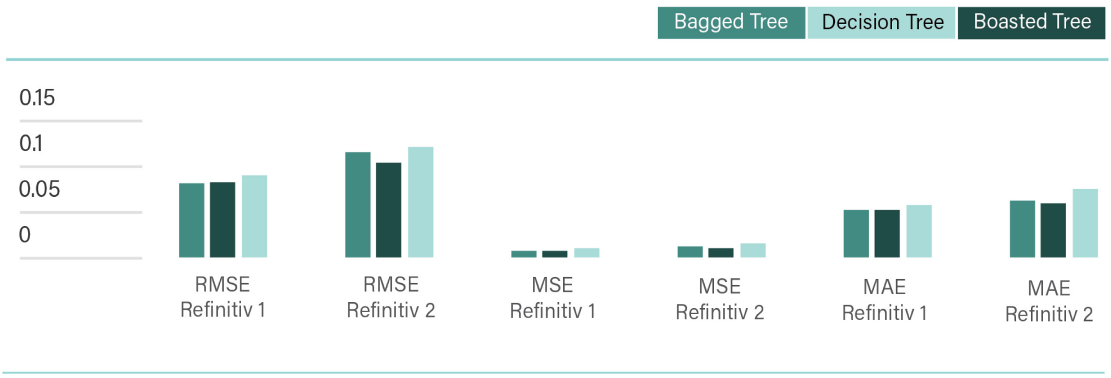

Figure 6. Comparison between RMSE, MSE, and MAE values for Refinitiv 1 and 2

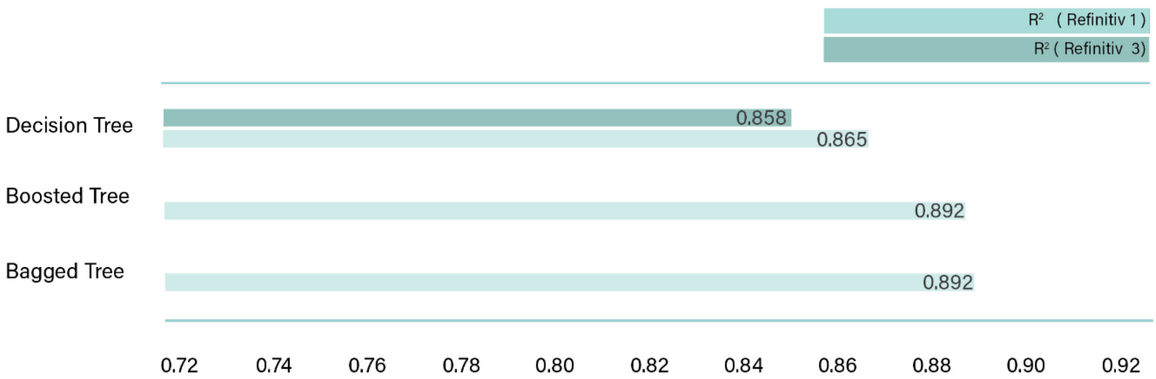

Figure 7. Comparison between R^2 values for Refinitiv 1 and 3 (no missing financial indicators)

Figure 7 presents a comparison of R^2 values for Refinitiv 1 and 3, testing the effect of removing the missing financial indicators from the dataset. The R^2 value was 0.892 before exclusion and 0.858 after exclusion.

Figure 8 presents a comparison between Refinitiv 1 and 4. It is observed that once the dummy variables are removed from the dataset, the performance of the regression analysis improves slightly in terms of R^2, from the highest accuracy of 0.892 obtained from the Bagged Tree in Refinitiv 1 to 0.920 obtained by the linear regression in Refinitiv 4 (an increase of 2.779%). This result introduced and made the robust linear regression models superior to the ensemble Bagged Tree, which obtained the highest accuracy. One possible explanation for the increase in performance after the dummy variables are removed is that the linear regression exhibits limitations in capturing the nonlinearity and complexity inherent in the data system with dummy variables [36]. Therefore, as a result of having complex sets with dummy variables, linearity is not assumed, so tree-based algorithms are more applicable than linear regression [33]

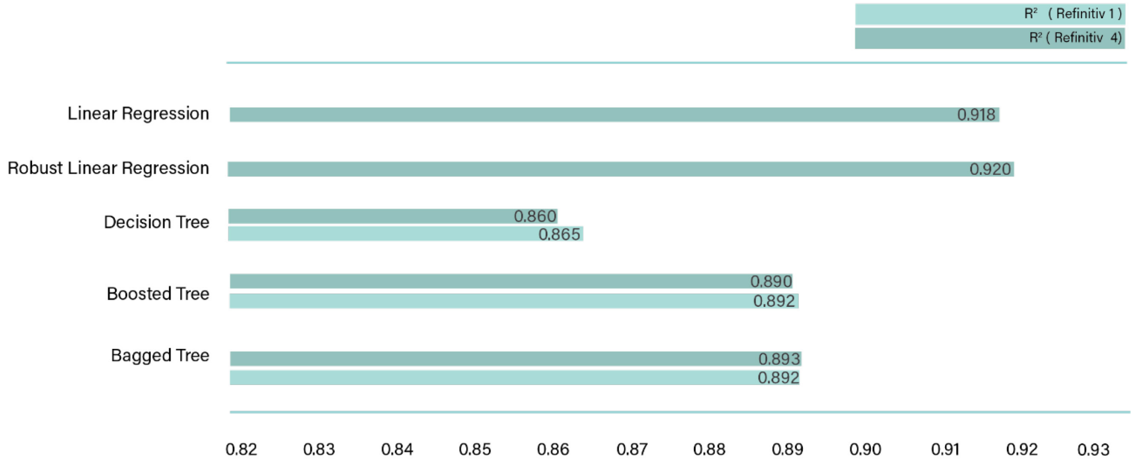

Figure 8. Comparison between R^2 values for Refinitiv 1 and 4 (no dummy variables)

Table 7. RMSE, MSE, and MAE values for Refinitiv 3 and 4

	Model type	MSE	RMSE	MAE
Refinitiv 3	Decision Tree	0.086537	0.007489	0.06246
	Decision Tree	0.086057	0.007406	0.062128
	Robust Linear	0.064558	0.004168	0.049058
Refinitiv 4	Linear	0.065173	0.004248	0.049689
	Bagged Tree	0.075014	0.005627	0.052935
	Boosted Tree	0.076111	0.005793	0.055162

Table 7 presents a comparison between Refinitiv 3 and 4 performance metrics in terms of the values of MAE, MSE, and RMSE. A similar regression analysis was performed on the data obtained from Datastream, and the results obtained are presented in Figure 9. The highest R^2 value of 0.878 was obtained from the Bagged Tree, compared to 0.892 from Refinitiv regression analysis.

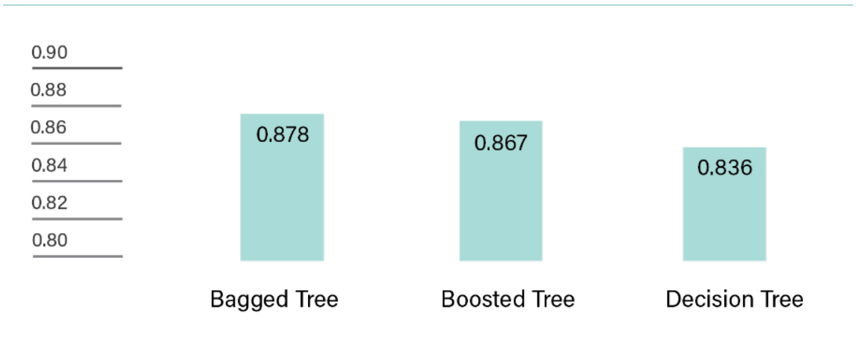

Figure 9. R^2 values of Datastream 1 (no missing ESG scores set)

To test the effect of utilising the full data set with missing ESG scores, Datastream 2 was used, and the values of R^2 obtained are compared to those of Datastream 1, as provided in Figure 10. The value of R^2 for the ensemble Bagged Tree increased slightly from 0.878 to 0.888 when the entire data set with missing values was used. The results indicate that sample size is crucial in machine learning. Hence, both classification and regression analysis would significantly benefit from a larger set of companies and historical data on ESG and financial indicators extended for a decade or more, along with robust features [7].

Table 8, **Figure 10**, and **Figure 11** represent the performance metrics of both Datastream 1 and 2 datasets for the regression analysis.

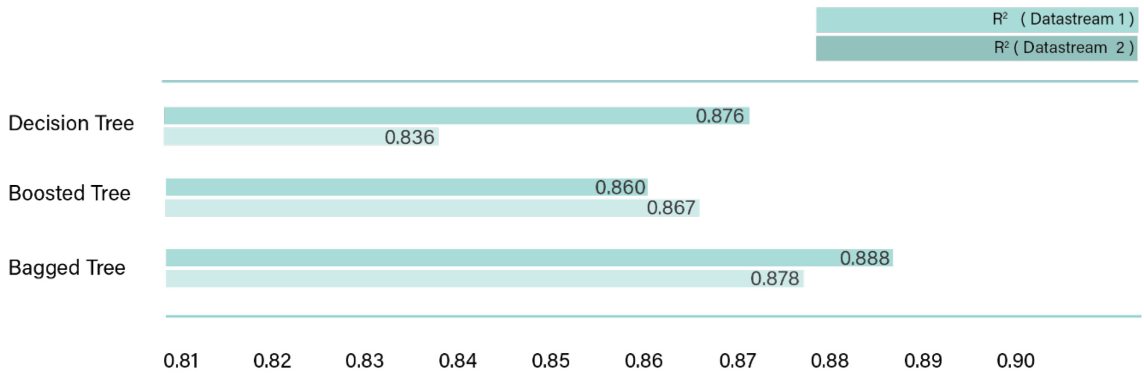

Figure 10. Comparison between R^2 values of Datastream 1 and 2 (encountered missing ESG scores)

Table 8. RMSE, MSE, and MAE values of Datastream 1 and 2

	RMSE		MSE		MAE	
Model type	Datastream 1	Datastream 2	Datastream 1	Datastream 2	Datastream 1	Datastream 2
Bagged Tree	0.0765	0.0784	0.0059	0.0062	0.0547	0.057
Boosted Tree	0.0798	0.0878	0.0064	0.0077	0.0604	0.0627
Decision Tree	0.0886	0.0826	0.0078	0.0068	0.0645	0.0601

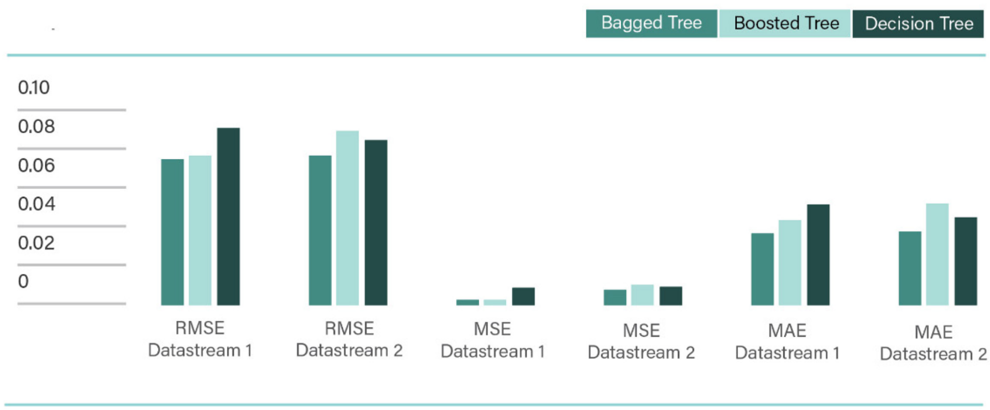

Figure 11. Comparison between RMSE, MSE, and MAE values of Datastream 1 and 2

Azure Results

Regression was applied to the Refinitiv 1 dataset, and the results are provided in **Table 9**. The highest R^2 value was obtained from the Voting Ensemble algorithm, at 0.922, followed by slightly reduced values from the other algorithms: Extreme Random Trees (0.920), Elastic Net (0.919), Random Forest (0.917), LightGBM (0.914), and XGBoost Regressor (0.906). Since a greater R^2 value suggests a better fit for the data set, its value is considered important. A value closer to 1 is aimed for, as when the R^2 core is 1, the regression predictions and the data are perfectly matched [37].

Table 9. Refinitiv 1 (3 classes of ESG scores) performance metrics

Performance metric	Voting Ensemble	Extreme Random Trees	Elastic Net	Random Forest	Light GBM	XGBoost Regressor
MAE	0.0466	0.0477	0.0472	0.0482	0.0496	0.0516
MAPE	20.0590	21.7140	18.7250	19.4550	20.5760	22.9710
R^2 value	0.9220	0.9199	0.9194	0.9171	0.9137	0.9065
RMSE	0.0640	0.0649	0.0650	0.0660	0.0674	0.0701

As explained by the results obtained from MATLAB, the closer the values of MAE and RMSE to zero, the better the model's performance. According to the obtained results, the highest MAE and RMSE values among the six models were 0.0516 and 0.0701, respectively, obtained by the XGBoost Regressor. Those values, although the highest, are promising and close to the error values of the other five models.

When relative variances in a regression job are more important than absolute values, the MAPE, which highlights percentage errors, is a suitable metric [35]. The values of MAPE obtained from this analysis showed the lowest value when the Elastic Net was used, with a percentage of 18.725, followed by Random Forest (19.455), Voting Ensemble (20.059), LightGBM (20.576), and XGBoost Regressor (22.971).

Testing the effect of removing the financial indicators, Table 10 represents a comparison between the R^2 values of Refinitiv 1, 2, and 3. The results show a slight reduction in the R^2 value for all the models, except for the XGBoost Regressor, whose R^2 value increased from 0.907 to 0.913 after the removal of the missing financial indicators. The influence of reducing the sample size in regression analysis on the performance of R^2 is consistent with the findings from MATLAB.

Table 10. Comparison between R^2 values for Refinitiv 1, 3, and 4

Model type	R^2 value		
	Refinitiv 1	Refinitiv 3	Refinitiv 4
Voting Ensemble	0.922	0.922	0.916
Extreme Random Trees	0.920	0.920	0.916
Elastic Net	0.919	0.919	0.913
Random Forest	0.917		0.912
Light GBM	0.914	0.913	0.906
XGBoost Regressor	0.906	0.913	
Decision Tree		0.906	0.905

Figure 12 illustrates the comparison of R^2 between Refinitiv 1, 3, and 4 for the regression analysis data set. First, the results of Refinitiv 4 show that the values of R^2 obtained for the models Voting Ensemble, Extreme Random Trees, Elastic Net, Random Forest, and LightGBM were slightly lower compared to the values obtained from Refinitiv 1. Using dummy variables in the dataset had a very slight negative impact on accuracy when the Azure platform was employed. Table 11 and Table 12 represent the obtained performance metrics for Refinitiv 1, 3, and 4, indicated by their values of MAE, RMSE, and MAPE.

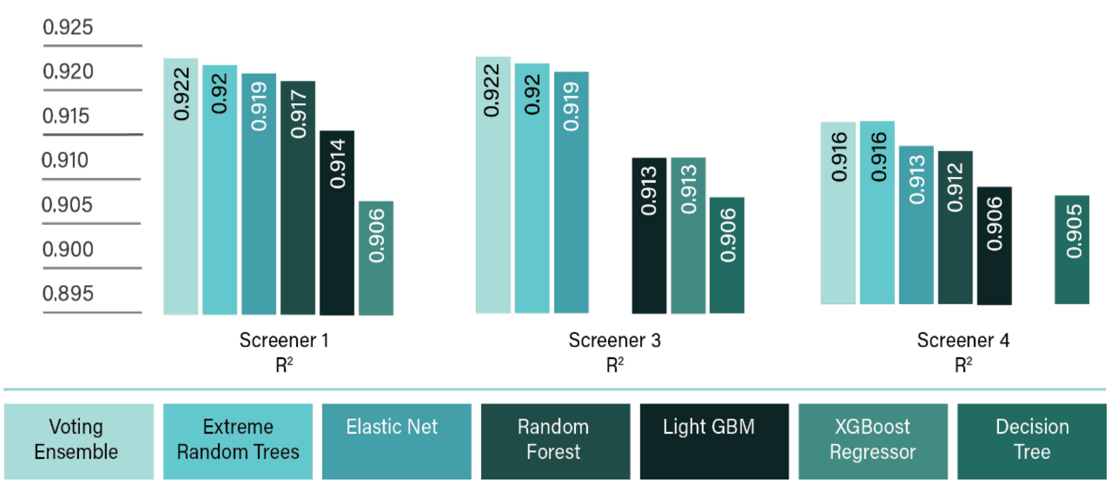

Figure 12. Comparison between R^2 values of Refinitiv 1, 3, and 4

Table 11. MAE and RMSE for Refinitiv 1, 3, and 4

Model type	MAE			RMSE		
	Refinitiv 1	Refinitiv 3	Refinitiv 4	Refinitiv 1	Refinitiv 3	Refinitiv 4
Voting Ensemble	0.0466	0.0468	0.0481	0.064	0.0642	0.0664
Extreme Random Trees	0.0477	0.0474	0.0483	0.0649	0.065	0.0666
Elastic Net	0.0472	0.0475	0.0489	0.065	0.0653	0.0676
Random Forest	0.0482		0.0501	0.066		0.0682
LightGBM	0.0496	0.0498	0.0511	0.0674	0.0677	0.0701
XGBoost Regressor	0.0516	0.0502		0.0701	0.0676	0.0709
Decision Tree		0.0514	0.0519		0.0703	

Table 12. MAPE for Refinitiv 1, 3, and 4

Model type	MAPE		
	Refinitiv 1	Refinitiv 3	Refinitiv 4
Voting Ensemble	20.06	19.46	19.04
Extreme Random Trees	21.71	20.04	19.78
Elastic Net	18.73	18.55	18.08
Random Forest	19.46		20.63
LightGBM	20.58	19.5	20.3
XGBoost Regressor	22.97	24.07	
Decision Tree		19.96	20.0

Feature Importance

Figure 13, Figure 14, and Figure 15 illustrate the feature importance of country level, firm size, and ESG scores for the Refinitiv 1 dataset, respectively. The minimum-redundancy maximum-relevancy (mRMR) feature selection method was applied in this analysis, as it has been shown to outperform traditional top-ranking methods in various recent studies, as noted by [38].

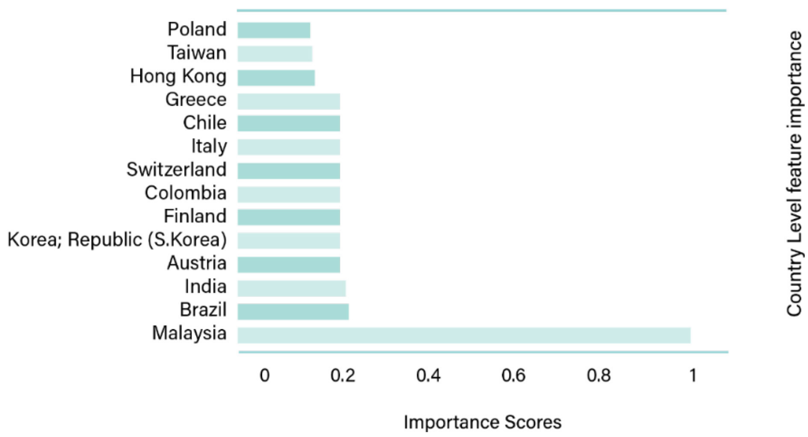

Figure 13. Importance score – country level

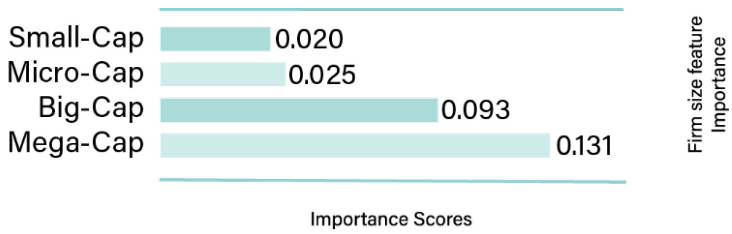

Figure 14. Importance scores – firm size

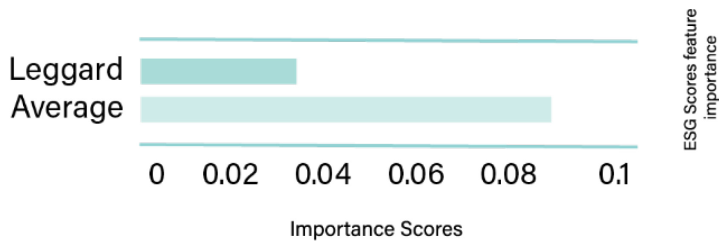

Figure 15. Importance scores – ESG score

DISCUSSION

This study examines the prediction of ESG scores in the oil and gas sector by leveraging historical financial and ESG indicators as input data. Additionally, machine learning has proven to be highly effective and well-suited for predicting complex scenarios within the ESG-finance domain [9].

MATLAB Analysis

The results obtained from Bagged, Boosted, and Decision Tree models for the Refinitiv 1 dataset were closely aligned. Each algorithm yielded high accuracy but emphasised different factors. In the case of the ensemble Boosted Tree model, one possible reason for its performance is that boosting operates sequentially, correcting errors from previous trees to minimise bias in the final prediction. The ensemble's final output is derived from the collective predictions of all trees in the model [39]. Decision Trees are widely recognised for their effectiveness in regression analysis [40], as demonstrated in the results across the datasets of Refinitiv 1, 2, 3, and 4. The strong performance of these models may be attributed to the efficiency of the tree traversal algorithm [40] and their high generalisation capability [25] Regarding ensemble Bagged Trees, their effectiveness lies in the fact that bagging generates multiple subsets of the original dataset and trains them independently [41] This approach utilises all available dataset features and integrates them simultaneously to produce a collective prediction [42]. Linear regression performed best for Refinitiv 4 when the dummy variables were removed from the data set. For the results obtained in MATLAB, linear regression failed to explain the results, except for this data set. The highest R^2 value and the lowest MAE, MSE, and RMSE values for Refinitiv 4 were obtained from the robust linear regression (a form of linear regression algorithms that utilises weighted least squares, including the weight as anadditional scaling factor in the fitting process), which improves the fit. The weights determine how much each response value affects the final parameter estimates, with low-quality data points, such as outliers, having less influence on the fit [38].

Azure Analysis

The Voting Ensemble delivered the highest performance among the regression models, though its advantage over other models was only marginal. The strong performance of the Voting Ensemble in machine learning tasks is attributed to its ability to combine multiple learning models rather than relying on a single predictor. It generates multiple decision trees and determines the preferred output class through a majority vote [43]. XGBoost was employed for regression analysis due to its advantages, including high accuracy, sensitivity, and precision. The model achieved a regression score of 0.906, indicating a reliable and well-performing outcome. The results obtained from the Random Forest algorithm were also reliable, likely due to its robustness against noise and resistance to overfitting [25]. Both Random Forest and LightGBM are well-suited for regression problems, effectively handling continuous values [44]. However, the Random Forest was unable to interpret the results for the Refinitiv 3 dataset accurately. Regarding the Extreme Random Trees algorithm, its results were very similar to those of other models, demonstrating its computational efficiency. According to [45], using all original training samples instead of bootstrap replicas in the randomised tree strategy helps reduce bias, thereby contributing to the model's effectiveness.

The Elastic Net algorithm, which is built upon traditional regression or logistic regression [46], was utilised for the regression analysis. The obtained R^2 values were also close to those obtained from the regression results of Voting Ensemble, Extreme Random Trees, Random Forest, LightGBM, XGBoost Regressor, and Decision Tree. However, the Elastic Net model achieved the lowest MAPE amongst the other algorithms. The performance of this model is attributed to its ability to assess the random effects of panel data regressions that incorporate multiple interactions [21]. Lastly, in Azure machine learning, the Decision Tree could interpret the results of the regression analysis of datasets Refinitiv 3 and 4 only, meaning that the Decision Tree was only applicable with the reduced datasets. This situation could be attributed to the fact that Decision Tree utilises a divide-and-conquer strategy, excelling when there are only a few highly relevant attributes. However, their performance declines when there are numerous complex interactions among attributes. Errors can propagate through the tree, especially as the number of classes increases. Additionally, as the tree grows larger, the records within leaf nodes often become too small to support statistically significant decisions [25]. Furthermore, another possible reason is that the Decision Tree could be unstable, and the difficulty in managing the tree size [40]

Datastream Results

The results obtained from the Datastream show a significant reduction in performance compared to the Refinitiv-sourced data. A comparison between the Datastream and Refinitiv regressions reveals that the highest R^2 value of 0.878 was obtained from the Bagged Tree, while the Refinitiv regression analysis yielded an R^2 of 0.892. Furthermore, the value of R^2 for the ensemble Bagged Tree increased slightly from 0.878 to 0.888 between Datastream 1 and 2. This reduction in performance for Datastream-sourced data could be attributed to overfitting; the sample dataset obtained from Datastream is too small compared to Refinitiv, and as a result, inadequate data are provided for training and testing the machine learning algorithms. According to [32], the more training data provided, the stronger the generalisation ability. The volume of training data available significantly influences the accuracy of learning outcomes, and improved performance correlates with larger datasets. Consequently, addressing data shortage becomes imperative to mitigate potential overfitting during training, thereby preserving the model's generalisation capability [32]

Effect of Country and Firm Size

At the country level, this study examined the impact of country-level sustainability performance. Feature importance analysis revealed that country-level factors had the strongest influence. Notably, Malaysia emerged as the most significant indicator, with a feature importance value of 0.934. Malaysia has been a focal point in ESG research since implementing its first corporate social responsibility (CSR) framework in 2006. The integration of CSR into the Tenth Malaysia Plan, along with mandatory CSR disclosure requirements under the Companies Act 2016, has contributed to high CSR scores among Malaysian firms. As a result, Malaysia has positioned itself as a global leader in CSR reporting, with approximately 97% of its top 100 companies disclosing their corporate sustainability performance [47]. These findings align with [48], who argued that countries with strong ESG improvement initiatives often experience positive economic outcomes from their sustainability efforts. Furthermore, this effect is more pronounced in nations with supportive regulatory, legal, financial, and cultural frameworks for ESG implementation.

At the firm level, feature importance was found to be lower than at the country level. Mega-capitalisation companies had the highest firm-level feature importance, with a value of 0.131. Company size plays a crucial role in determining access to ESG initiatives and financial stability [31]. Several factors underscore the importance of firm size in the relationship between ESG and financial performance. Larger firms generally possess greater financial resources than smaller ones, allowing them to invest more extensively in sustainability initiatives [10]. Additionally, large corporations typically have structured strategies and business monitoring mechanisms that facilitate effective sustainability management. Moreover, firm visibility plays a key role, as prominent companies are more likely to adopt superior sustainability practices to maintain their reputation among stakeholders [49]

Furthermore, the substantial difference in feature importance scores between the highest-ranked and second-ranked features suggests a high level of confidence in the feature selection process. This significant drop indicates that MATLAB is highly confident in identifying the most critical predictor [38].

Most studies highlighted in this work suggest a positive relationship between ESG performance and financial performance [10]. The results obtained from this study are aligned with the results obtained from [8] (where MAE = 11.2 and $R^2 = 0.54$) and [9] (see **Table 13**), which relied on quantitative approaches utilising firms' financial indicators to predict ESG scores, and are superior to the results of [10] (where MAE = 1.22 and $R^2 = 0.05$) and [11] (see **Table 14**). Furthermore, this research addresses the gap in the literature where the prediction of ESG scores, specifically in the oil and gas industry, utilising machine learning and firm financial indicators, is lacking.

Table 13. R^2 values obtained from [9]

Model type	R^2
XGBoost	0.9898
Random Forest	0.9678
Support Vector	0.9916

Table 14. RMSE and MAE values obtained from [11]

Model type	RMSE	MAE
Decision Tree	1.89	1.08
Linear Regression	5.29	4.42
LightGBM	5.9	34.82
XGB	1.18	1.4
Random Forest	5.44	3.94

According to stakeholder theory, adopting ESG practices and promoting transparency can contribute to enhancing a company's overall value [50]. Engaging with stakeholders and addressing their concerns through socially responsible actions can help mitigate externalities and build stronger, positive relationships. As a result, an increasing number of companies are now publicly disclosing their ethical, social, and environmental practices and actively working to reduce their negative environmental impacts [51]. The authors of [52] argue that ESG initiatives can create shareholder value by boosting firm cash flows; for instance, a strong corporate reputation can drive higher product sales, and skilled employees can improve productivity. In comparison, the authors of [10] emphasise that sustainability-oriented firms offer greater utility to socially conscious investors. These insights suggest that sustainability initiatives align more closely with the interests of stakeholders. Furthermore, firms exhibiting stronger financial performance are believed to have a greater capacity to undertake social responsibilities, as higher profitability enables more significant investment in socially responsible activities [51].

In particular, the oil and gas sector faces mounting ESG pressures from investors, regulatory bodies, and the broader public. Consequently, ESG scores have become an increasingly critical factor in investment decision-making, with their significance continuing to rise each year. In response to these trends, this research presents a comprehensive methodology designed to assist stakeholders and policymakers in predicting ESG scores, thereby supporting more sustainable investment decisions [9].

CONCLUSIONS

Companies and their stakeholders are increasingly prioritising ESG and sustainability, particularly in the oil and gas sector, where environmental and social risks are highly scrutinised. This study presents a predictive methodology for ESG scoring in the industry, leveraging publicly available financial and ESG indicators, as well as advanced machine learning algorithms. A total of eleven models, including Decision Tree, ensemble methods (Boosted, Bagged, and Voting), XGBoost, LightGBM, Random Forest, Extreme Random Trees, linear regression, robust regression, and Elastic Net, were evaluated. The Voting ensemble model achieved the highest predictive accuracy with an R^2 of 0.922.

In addition to firm-level indicators, the study incorporated country-level sustainability regulations and firm size, revealing that country-level ESG enforcement was the most influential factor in predicting ESG scores, followed by firm size.

This research contributes to the existing literature by presenting a data-driven approach specifically designed for ESG assessment in the oil and gas industry. In this area, standardised evaluation remains a challenge. The framework provides an accurate and reliable method to support ESG performance assessment, enabling more informed decision-making, supporting regulatory benchmarking, and enhancing transparency in ESG reporting and compliance practices across the energy sector.

NOMENCLATURE

Abbreviations

ESG	Environment, Social, and Governance
LightGBM	Light Gradient Boosting Machine (gradient-boosting framework)
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
ROA	Return On Assets
ROE	Return On Equity
XGBoost	Extreme Gradient Boosting (optimised gradient-boosting algorithm)

REFERENCES

1. O. Lee, H. Joo, H. Choi, and M. Cheon, "Proposing an Integrated Approach to Analyzing ESG Data via Machine Learning and Deep Learning Algorithms," *Sustainability (Switzerland)*, vol. 14, no. 14, Jul. 2022, https://doi.org/10.3390/su14148745.
2. M. Gamlath, C. Gunathilaka, A. Wijesinghe, S. Ahangama, I. Perera, and L. Sivaneasharajah, "An Integrated Approach to ESG Index Construction with Machine Learning," in *Moratuwa Engineering Research Conference, MERCon*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 252–257. https://doi.org/10.1109/MERCon60487.2023.10355516.
3. C. Zheng, M. Zhang, C. Zeng, F. Xiao, and M. Liu, "Analysis of the Influential Factors and Prediction of Corporate ESG Performance under Multi-source Data Fusion - Based on Frontier Machine Learning," European Alliance for Innovation n.o., Jun. 2023. https://doi.org/10.4108/eai.2-12-2022.2328731.
4. A. Ramírez-Orellana, Mc. C. Martínez-Victoria, A. García-Amate, and A. A. Rojo-Ramírez, "Is the corporate financial strategy in the oil and gas sector affected by ESG dimensions?," *Resources Policy*, vol. 81, Mar. 2023, https://doi.org/10.1016/j.resourpol.2023.103303.
5. H. Yu and T. Su, "ESG performance and corporate solvency," *Financ Res Lett*, vol. 59, Jan. 2024, https://doi.org/10.1016/j.frl.2023.104799.
6. D. Zhang, D. Bai, and X. Chen, "Can crude oil futures market volatility motivate peer firms in competing ESG performance? An exploration of Shanghai International Energy Exchange," *Energy Econ*, vol. 129, Jan. 2024, https://doi.org/10.1016/j.eneco.2023.107240.
7. M. Gamlath, C. Gunathilaka, A. Wijesinghe, S. Ahangama, I. Perera, and L. Sivaneasharajah, "An Integrated Approach to ESG Index Construction with Machine Learning," in *Moratuwa Engineering Research Conference, MERCon*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 252–257. https://doi.org/10.1109/MERCon60487.2023.10355516.

8. D. Si, "An Efficient Predicative Approach of ESG Invest Scoring Using Random Forest Algorithm," *BCP Business & Management GEBM*, vol. 45, pp. 382–392, Apr. 2023, https://doi.org/10.54691/bcpbm.v45i.4950.
9. H. Y. Lin and B. W. Hsu, "Empirical Study of ESG Score Prediction through Machine Learning—A Case of Non-Financial Companies in Taiwan," *Sustainability (Switzerland)*, vol. 15, no. 19, Oct. 2023, https://doi.org/10.3390/su151914106.
10. T. Krappel, A. Bogun, and D. Borth, "Heterogeneous Ensemble for ESG Ratings Prediction," Sep. 2021, http://arxiv.org/abs/2109.10085, [Accessed: Dec. 20, 2023].
11. O. Lee, H. Joo, H. Choi, and M. Cheon, "Proposing an Integrated Approach to Analyzing ESG Data via Machine Learning and Deep Learning Algorithms," *Sustainability (Switzerland)*, vol. 14, no. 14, Jul. 2022, https://doi.org/10.3390/su14148745.
12. A. Sokolov, J. Mostovoy, J. Ding, and L. Seco, "Building Machine Learning Systems to Automate ESG Index Construction," *The Journal of Financial Data Science*, vol. 3, no. 3, pp. 129–138, 2020, https://doi.org/10.3905/jesg.2021.1.010.
13. "Does ESG Score and Intangible Asset improve Financial performance?: Machine Learning Model of KOSPI Enterprises," in *Proceedings - 2022 IEEE International Conference on Big Data, Big Data 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 6621–6623. https://doi.org/10.1109/BigData55660.2022.10021050.
14. J. Rojo-Suárez, A. B. Alonso-Conde, and J. D. Gonzalez-Ruiz, "Does sustainability improve financial performance? An analysis of Latin American oil and gas firms," *Resources Policy*, vol. 88, Jan. 2024, https://doi.org/10.1016/j.resourpol.2023.104484.
15. S. Ben Jabeur, R. Khalfaoui, and W. Ben Arfi, "The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: Evidence from explainable machine learning," *J Environ Manage*, vol. 298, Nov. 2021, https://doi.org/10.1016/j.jenvman.2021.113511.
16. Y. J. Y. C. Y. S. Jinlin Zheng, "Green bond issuance and corporate ESG performance: Steps toward green and low-carbon development," *Res Int Bus Finance*, vol. 66, no. 2023, Oct. 2023.
17. M. H. Giovanni Cardillo, "Stay close to me: What do ESG scores tell about the deal timing in M&A transactions?," *Financ Res Lett*, vol. 51, no. 103498, Jan. 2023.
18. M. G. X.-Y. Z. L. K. Yiwei Li, "The impact of environmental, social, and governance disclosure on firm value: The role of CEO power," *The British Accounting Review*, vol. 50, no. 1, Jan. 2018.
19. K. Albitar, K. Hussainey, N. Kolade, and A. M. Gerged, "ESG disclosure and firm performance before and after IR: The moderating role of governance mechanisms," *International Journal of Accounting and Information Management*, vol. 28, no. 3, pp. 429–444, Jun. 2020, https://doi.org/10.1108/IJAIM-09-2019-0108.
20. LSEG Data & Analytics | Financial Technology & Data, "LSEG Data & Analytics," LSEG - London Stock Exchange Group, https://www.lseg.com/en/data-analytics, [Accessed: Jan. 13, 2024].
21. A. Del Vitto, D. Marazzina, and D. Stocco, "ESG ratings explainability through machine learning techniques," *Ann Oper Res*, 2023, https://doi.org/10.1007/s10479-023-05514-z.
22. A. L'Heureux, K. Grolinger, H. F. Elyamany, and M. A. M. Capretz, "Machine Learning with Big Data: Challenges and Approaches," *IEEE Access*, vol. 5, pp. 7776–7797, 2017, https://doi.org/10.1109/ACCESS.2017.2696365.
23. R. Wayman, "Understanding Small-Cap and Big-Cap Stocks," https://www.investopedia.com/insights/understanding-small-and-big-cap-stocks/, [Accessed: Apr. 01, 2024].
24. M. Jihadi, E. Vilantika, S. M. Hashemi, Z. Arifin, Y. Bachtiar, and F. Sholichah, "The Effect of Liquidity, Leverage, and Profitability on Firm Value: Empirical Evidence from Indonesia," *Journal of Asian Finance, Economics and Business*, vol. 8, no. 3, pp. 423–431, 2021, https://doi.org/10.13106/jafeb.2021.vol8.no3.0423.

25. A. Singh, N. Thakur, and A. Sharma, "A Review of Supervised Machine Learning Algorithms," in *2016 3rd international conference on computing for sustainable global development (INDIACom)*, IEEE, Mar. 2016, pp. 1310–1315.
26. V. Kumar, N. Kedam, K. V. Sharma, D. J. Mehta, and T. Caloiero, "Advanced Machine Learning Techniques to Improve Hydrological Prediction: A Comparative Analysis of Streamflow Prediction Models," *Water (Switzerland)*, vol. 15, no. 14, Jul. 2023, https://doi.org/10.3390/w15142572.
27. O. Surakhi *et al.*, "Time-lag selection for time-series forecasting using neural network and heuristic algorithm," *Electronics (Switzerland)*, vol. 10, no. 20, Oct. 2021, https://doi.org/10.3390/electronics10202518.
28. I. Y. Javeri, M. Toutiaee, I. B. Arpinar, T. W. Miller, and J. A. Miller, "Improving Neural Networks for Time Series Forecasting using Data Augmentation and AutoML," May 2021, http://arxiv.org/abs/2103.01992, [Accessed: May. 21, 2024].
29. L. A. B. C. Barros, D. R. Bergmann, F. Henrique Castro, and A. D. M. da Silveira, "Endogeneity in panel data regressions: Methodological guidance for corporate finance researchers," *Revista Brasileira de Gestao de Negocios*, vol. 22, no. Special Issue, pp. 437–461, 2020, https://doi.org/10.7819/rbgn.v22i0.4059.
30. L. Lam, "ESG Reporting and Its Effect on Financial Performance of Oil, Gas, and Utility Companies in the United States.," University of Richmond, Richmond, VA, USA, 2023, https://scholarship.richmond.edu/honors-theses/1697/, [Accessed: Jun. 03, 2024].
31. M. G. Bruna, S. Loprevite, D. Raucci, B. Ricca, and D. Rupo, "Investigating the marginal impact of ESG results on corporate financial performance," *Financ Res Lett*, vol. 47, Jun. 2022, https://doi.org/10.1016/j.frl.2022.102828.
32. H. Zheng, Z. Zhou, and J. Chen, "RLSTM: A New Framework of Stock Prediction by Using Random Noise for Overfitting Prevention," *Comput Intell Neurosci*, vol. 2021, 2021, https://doi.org/10.1155/2021/8865816.
33. J. Svanberg *et al.*, "Corporate governance performance ratings with machine learning," *Intelligent Systems in Accounting, Finance and Management*, vol. 29, no. 1, pp. 50–68, Jan. 2022, https://doi.org/10.1002/isaf.1505.
34. S. M. Baek, W. S. Kim, Y. S. Kim, S. Y. Baek, and Y. J. Kim, "Development of a simulation model for HMT of a 50 kW class agricultural tractor," *Applied Sciences (Switzerland)*, vol. 10, no. 12, Jun. 2020, https://doi.org/10.3390/APP10124064.
35. D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Comput Sci*, vol. 7, pp. 1–24, 2021, https://doi.org/10.7717/PEERJ-CS.623.
36. S. Jaydip, S. Mehtab, and A. Engelbrecht, *Machine Learning - Algorithms, Models and Applications*. IntechOpen, 2021. https://doi.org/10.5772/intechopen.94615.
37. K. Vaid and U. Ghose, "Predictive Analysis of Manpower Requirements in Scrum Projects Using Regression Techniques," in *Procedia Computer Science*, Elsevier B.V., 2020, pp. 335–344. https://doi.org/10.1016/j.procs.2020.06.039.
38. Mathworks, "Reduce Outlier Effects Using Robust Regression," https://www.mathworks.com/help/stats/robust-regression-reduce-outlier-effects.html, [Accessed: May. 28, 2024].
39. A. Shivanna and D. P. Agrawal, "Prediction of Defaulters using Machine Learning on Azure ML," in *2020 11th IEEE annual information technology, electronics and mobile communication conference (IEMCON)*, IEEE, Nov. 2020, pp. 0320–0325.
40. V. Sheth, U. Tripathi, and A. Sharma, "A Comparative Analysis of Machine Learning Algorithms for Classification Purpose," in *Procedia Computer Science*, Elsevier B.V., 2022, pp. 422–431. https://doi.org/10.1016/j.procs.2022.12.044.

41. P. Y. Taser, "Application of Bagging and Boosting Approaches Using Decision Tree-Based Algorithms in Diabetes Risk Prediction," MDPI AG, Mar. 2021, p. 6. https://doi.org/10.3390/proceedings2021074006.
42. N. Stepanov, D. Alekseeva, A. Ometov, and E. S. Lohan, "Applying Machine Learning to LTE Traffic Prediction: Comparison of Bagging, Random Forest, and SVM," in *2020 12th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, IEEE, 2020, pp. 119–123.
43. *Proceedings, 2020 16th IEEE International Colloquium on Signal Processing & its Application (CSPA 2020) : 28th-29th February 2020 : conference venue, Hotel Langkawi, Lot 1852 Jalan Penarak, Kuah 07000 Langkawi, Kedah, Malaysia*. IEEE, 2020.
44. I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," May 01, 2021, *Springer*. https://doi.org/10.1007/s42979-021-00592-x.
45. E. K. Ampomah, Z. Qin, and G. Nyame, "Evaluation of tree-based ensemble machine learning models in predicting stock price direction of movement," *Information (Switzerland)*, vol. 11, no. 6, Jun. 2020, https://doi.org/10.3390/info11060332.
46. R. A. de Lima Lemos, T. C. Silva, and B. M. Tabak, "Propension to customer churn in a financial institution: a machine learning approach," *Neural Comput Appl*, vol. 34, no. 14, pp. 11751–11768, Jul. 2022, https://doi.org/10.1007/s00521-022-07067-x.
47. W. M. W. Mohammad and S. Wasiuzzaman, "Environmental, Social and Governance (ESG) disclosure, competitive advantage and performance of firms in Malaysia," *Cleaner Environmental Systems*, vol. 2, Jun. 2021, https://doi.org/10.1016/j.cesys.2021.100015.
48. J. Wang, J. Yu, and R. Zhong, "Country environmental, social and governance performance and economic growth: The international evidence," *Accounting and Finance*, vol. 63, no. 4, pp. 3911–3941, Dec. 2023, https://doi.org/10.1111/acfi.13079.
49. C. F. Antonio D'Amato, "Corporate social responsibility and firm value: Do firm size and age matter? Empirical evidence from European listed companies," *Corp Soc Responsib Environ Manag*, vol. 27, no. 2, 2019.
50. Z. Chen and G. Xie, "ESG disclosure and financial performance: Moderating role of ESG investors," *International Review of Financial Analysis*, vol. 83, Oct. 2022, https://doi.org/10.1016/j.irfa.2022.102291.
51. E. Adegbite, Y. Guney, F. Kwabi, and S. Tahir, "Financial and corporate social performance in the UK listed firms: the relevance of non-linearity and lag effects," *Review of Quantitative Finance and Accounting*, vol. 52, no. 1, pp. 105–158, Jan. 2019, https://doi.org/10.1007/s11156-018-0705-x.
52. A. K. L. T. S. Stuart L. Gillan, "Firms and social responsibility: A review of ESG and CSR research in corporate finance," *Journal of Corporate Finance*, vol. 66, p. 101889, 2021, https://doi.org/10.1016/j.jcorpfin.2020.101889.

APPENDIX A

Final list of financial indicators obtained from the literature review.

Table A1. Datastream and Refinitiv extracted financial indicators

Data Stream Variables – Datastream	Refinitiv – Refinitiv
Return On Assets	Actual Return On Assets
Return On Equity Total %	Return On Equity – Actual
Tobin's Q	Tobin's Q (sourced from Datastream)
Market Value (Capital)	Market Value
Total Return Index	Total Return Index (Excluded)
Dividend Yield	Net Dividend Yield
Net Income Available To Common	Net Income – Mean Net Income Available To Common
Total Assets	Total Assets – Actual
Long Term Debt	Debt-Long-Term – Total
Price To Sales (Excluded)	Price To Sales Ratio – Mean
Quick Ratio	Quick Ratio
Earnings per share	Earnings per share – Mean
Current Ratio	Current Ratio
Price To Book Value	Price To Book Value per share
Price/Earnings Ratio (Adjusted)	Price/Earnings Ratio (Adjusted) (Excluded)
Employees	Employees (Excluded)
Voluntary Turnover of Employees	Voluntary Turnover of Employees (Excluded)
Involuntary Turnover of Employees	Involuntary Turnover of Employees (Excluded)
Earnings Before Interest, Taxes & Depreciation (EBITDA)	EBITDA – Actual
Net Sales Or Revenues	Revenues – Mean
Total Liabilities	Total Liabilities
CEO-Chairman Separation	CEO-Chairman Separation (Excluded)
Foreign Sales Growth 1 Year Annual Growth	Foreign Sales Growth 1 Year Annual Growth (Excluded)
Trailing Twelve Months Operating Margin	Trailing Twelve Months Operating Margin (Excluded)
Net Cash Flow Operating Activities	Net Cash Flow Operating Activities
Price/Cash Flow Ratio	Price/Cash Flow Ratio – per share
Total shareholders equity	Total shareholders equity (Excluded)
Total investment return	Total investment return (Excluded)
Net margin	Net margin (Excluded)

Paper submitted: 17.06.2025
Paper revised: 13.11.2025
Paper accepted: 17.11.2025