



## OPEN ACCESS

## EDITED BY

Vishal Dagar,  
Great Lakes Institute of Management,  
India

## REVIEWED BY

Muhammad Owais Khan,  
University of Agriculture, Peshawar,  
Pakistan  
Muhammad Imran Khan,  
University of Agriculture, Peshawar,  
Pakistan  
Mahejabin Sayyad,  
Agasti Arts, Commerce And Dadasaheb  
Rupwate Science College, India

## \*CORRESPONDENCE

Muhammad Anees Khan,  
lookanees123@gmail.com

## SPECIALTY SECTION

This article was submitted to  
Environmental Economics and  
Management,  
a section of the journal  
Frontiers in Environmental Science

RECEIVED 22 June 2022

ACCEPTED 13 July 2022

PUBLISHED 04 October 2022

## CITATION

Raza H, Khan MA, Mazliham MS,  
Alam MM, Aman N and Abbas K (2022),  
Applying artificial intelligence  
techniques for predicting the  
environment, social, and governance  
(ESG) pillar score based on balance  
sheet and income statement data: A  
case of non-financial companies of  
USA, UK, and Germany.  
*Front. Environ. Sci.* 10:975487.  
doi: 10.3389/fenvs.2022.975487

## COPYRIGHT

© 2022 Raza, Khan, Mazliham, Alam,  
Aman and Abbas. This is an open-access  
article distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original  
publication in this journal is cited, in  
accordance with accepted academic  
practice. No use, distribution or  
reproduction is permitted which does  
not comply with these terms.

# Applying artificial intelligence techniques for predicting the environment, social, and governance (ESG) pillar score based on balance sheet and income statement data: A case of non-financial companies of USA, UK, and Germany

Hassan Raza<sup>1</sup>, Muhammad Anees Khan<sup>1\*</sup>, M. S. Mazliham<sup>2</sup>,  
Muhammad Mansoor Alam<sup>2</sup>, Nida Aman<sup>1</sup> and Kumail Abbas<sup>2</sup>

<sup>1</sup>Bahria Business School, Bahria University, Islamabad, Pakistan, <sup>2</sup>Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia

Due to globalization, environment, social, and governance (ESG) issues have gained importance over the last few decades. ESG is a worldwide issue, which clarifies that organizations throughout the world are lacking in contribution to the environment, society, and corporate governance characteristics for sustainable development. The problem of ESG spread over all stakeholders needs to be addressed. In this regard, rating agencies also have a close eye on ESG issues and have developed the methodology of score that aims to provide disclosure on ESG metrics which, in return, help investors and asset managers better differentiate between responsible and irresponsible companies. The ESG score has become an important tool among asset managers but is highly questioned in terms of reliability. The study objective was to develop machine learning algorithms to assess how balance sheet and income statement data impact the Thomson Reuters ESG score for non-financial public companies of USA, UK, and Germany from 2008 to 2020. In addition, the study also has an objective to assess which machine learning (ML) algorithm better predicts the ESG score using structural data, that is, return on assets (ROA), return on equity (ROE), earning per share (EPS), earnings before interest and taxes (EBIT), dividend yield, and net sales. The results concluded that balance sheet and income statement data are critical in explaining the ESG score, and the ANN algorithm outperforms with minimum RMSE and MAE values. All in all, the results of the study, based on the concept of artificial intelligence, bring suggestion for improvement to regulatory bodies, researchers, academia, practitioners, publicly listed companies around the globe, and last but not the least to the US, UK, and Germany markets. Moreover, it also provides suggestions for up-to-date compliance of ESG-relevant activities for boosting the firm performance.

## KEYWORDS

ESG score, machine learning, prediction, sustainability, balance sheet

## Introduction

Over the last two decades, ESG has significantly influenced businesses, stakeholders, and shareholders around the world and caused damage, including a decline in the worldwide reputations of companies, auditors, security analysts, regulators, and financial markets based on ESG compliance (Ball, 2009). Despite the effort made for minimizing the issue of ESG, they have not stopped (Vladu & Cuzdriorean, 2013). Moreover, due to lack of overall compliance with ESG issues, corporate frauds have been on the increase across the world in the form of false ESG reporting, environmental ignorance, and social embezzlement (MingChia, 2012). Ultimately, these accounting frauds and violations have created ripples in the corporate world. As a result, investors and other stakeholders have lost their trust and confidence in the ESG reporting processes and management teams (Hamid, Hashim, & Salleh, 2012). The previously discussed ESG issues have led to the global collapse of many high-profile businesses (Moncarz, Moncarz, Cabello, & Moncarz, 2006). In this regard, researchers used machine learning techniques to predict the firm performance and success based on ESG scores (Ahmed & Hamdan, 2015), which involves the automatic deduction of patterns in data. It is described as a program that learns to automatically accomplish a task rather than programmed explicitly. It has the ability to process a vast volume of data and extract meaningful information through different programming techniques. They are widely used nowadays for predicting or forecasting the stock market, commodity market, and foreign exchange market. The digital transformation journey is at the heart of the fourth industrial revolution, and analysts and decision-makers aim to make daily processes simpler and more efficient. The process of data gathering, analysis, planning, implementation, tracking and automation, and reporting can be overwhelming. Artificial intelligence is a major technological breakthrough that everyone is talking about its exciting potential. Artificial intelligence is described as a machine's ability to make intellectual human-like decisions and continue to improve. John McCarthy came up with the term "artificial intelligence" in 1956. Deep learning, natural language processing (NLP), machine learning, image recognition, sound recognition, cognitive computing, and enhanced intelligence are all terms used to explain artificial intelligence (Yaninen, 2017). Machine learning comprises the creation of models, especially statistical models that can be developed and predict outcomes. According to a recent Bank of America Corporation study, ESG investments based in the United States will grow the size of the US stock market over the next two to three decades. Data are becoming more readily available and of higher quality, which provide useful tools for analyzing sustainable investing. Artificial intelligence

can swiftly reveal hidden risks and possibilities that traditional analysis may oversee. In a comparison of conventional techniques that are quickly becoming obsolete, artificial intelligence techniques may offer significant benefits to the world of finance, by automating certain tasks and boosting analytical capability. Artificial intelligence is a critical part of modern finance because it makes it cheaper, quicker, bigger, more available, profitable, and competitive in a variety of ways.

The benefit of this research is that the non-linearity and complexity of historical data require adoption of new technologies which better analyze the data because traditional statistical methodologies have various flaws that might put financial organizations at risk and negatively impact their performance (Chen Y. et al., 2021). The benefit of this research is that the ML technique predicts the future, while the old regression techniques estimate the data rather than predicting it.

On the other hand, the benefit of this research will contribute to minimizing the bad practices of ESG. As it is reported by the previous authors that sometimes ESG reports are unethical and executed to mislead the stakeholders and investors. Thus, it can be added that current research will for sure improve the services committed by companies to the stakeholders and will also improve the overall treatment and process based on the stakeholder's theory concept, which in turn will help in future research, thus developing the smart system. This research will for sure help in developing a new and latest idea with certain data diagnosis, treatments, and processes that could eventually help one in industry, academia, and scholar in the future.

This study will fill the exhibited gap of these variables (ESG and FFP), in the limited and incongruent literature on ESG and FFP by extending the previous conventional model to the novel artificial intelligence model. The problem statement and literature review sections, respectively, identify the literature gaps of the study. Moreover, as discussed previously in the problem statement, the ESG score and FFP in developing and emerging economies are in their infancy stage and have no exception for developed countries, that is, why working on the ESG score for development became very important. Although many previous academic researchers have contributed to knowledge of the ESG and FFP, however, through traditional methodology but in this study, we worked on ESG based on the latest novel methodology, that is, artificial intelligence based on machine learning techniques. The aim of the study was to sum up the gaps in the study depicted on the basis of the inclusive-exclusive theory with a wide range of different datasets for the latent period based on the machine learning technique for the ignored non-financial sectors. Therefore, the previously discussed gaps create a problem in understanding that rationalizes the need to continue research

using the latest data available in concerned company's annual reports.

## Environment, social, and governance (ESG)

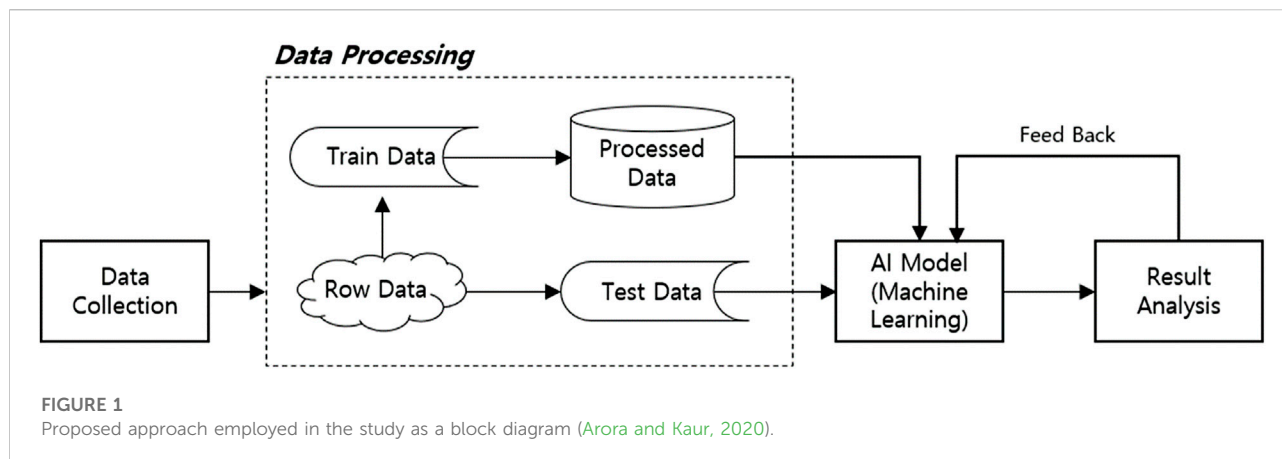
Environmental, social, and governance (ESG) are three key aspects in determining a company's long-term viability and ethical influence (Zheng et al., 2021a; Zheng et al., 2021b; Zheng et al., 2021c). An ESG score has become a popular phrase used by investors and in capital markets to assess a firm's behavior and determine a company's financial performance. ESG is an umbrella concept for firms and businesses that aim to generate favorable returns while simultaneously having a long-term influence on society, the environment, and the company's governance referred to as "sustainable investing" (Lei et al., 2021; He et al., 2022; Yao et al., 2022). It is based on the growing belief that environmental and social factors are increasingly impacting the financial performance of organizations. In other words, so-called "non-financial" risk becomes financial risk (Antonicic, 2020). Following the guidelines for responsible investments (PRI) in 2006, a collection of principles for incorporating sustainability aspects, ESG investing has formally entered mainstream investment discourse (Cantele et al., 2020). A variety of causes may be attributed to the increase in ESG investments. The business sector is growing more aware of social, labor, and human rights issues as supply chains become more complicated. The ESG score is determined by considering several factors for an environmental pillar, social pillar, and governance pillar (Choi et al., 2021; Yang et al., 2022; Zhao and Wang 2022). Greenhouse gas emissions, deforestation, waste material and pollution, climate change, and resource depletion are all factors that impact a company's environmental pillar. Equal treatment of all corporate employees, including gender concerns, health and safety, job security, and human rights is a prevalent topic related to the social side of sustainability (Cardillo and Longo, 2020). Tax planning, executive remuneration, board diversity, political donations, corruption, lobbying, and bribery are among the topics that focus on the governance side of sustainability. The fundamental question is whether AI technology can tell the difference between responsible and irresponsible businesses. For the years 2008–2020, the study observed the role of structural variables including balance sheet and income statement data as predictors of ESG scores of public corporations in the United Kingdom, the United States, and Germany using machine learning techniques. KNN, polynomial regression, naive Bayes, random forest, decision tree, support vector machine (linear), support vector machine (Rbf), and artificial neural networks are examples of supervised machine learning techniques used in the study. These algorithms were created using Thomson Reuters Refinitiv ESG scores to forecast

the ESG score. A total of 300 non-financial enterprises from the technology and communication, chemical, fertilizer, pharmaceutical, power generation and distribution, textile, and cement industries were chosen for the study. After the initial data preprocessing stage, the dataset was partitioned into two subsets: a training set (3,120) for training the models and a testing set (780) for testing the models, totaling 3,900 rows. Testing data are used to evaluate the training algorithm progress. Although investing in environmental, social, and governance (ESG)-driven portfolios account for a sizable and rising share of global assets under management, quantitative methodologies for improving and standardizing ESG rating and portfolio creation remain neglected (Sokolov et al., 2021). Internal approaches and practices are currently being used to incorporate ESG concerns into institutions' business strategy, procedures, and risk management. Inconsistency and data scarcity, described as "a scarcity of relevant, comparable, reliable, and user-friendly data," threaten the validity of ESG risk measurements (EBA, 2020a). The major risk-based approaches depend on historical data and are analyzed with traditional statistical techniques. According to the EBA (2020b), traditional models do not integrate ESG aspects, and the majority of ESG risks are non-linear. The non-linearity and complexity of historical data require adoption of new technologies which better analyze the data. Traditional statistical methodologies have various flaws that might put financial organizations at risk and negatively impact their performance (Chen Y. et al., 2021).

## Methodology of the study

### Theory, sample, and data

The study used the computational learning theory, with the purpose to identify that which machine learning technique best predicts the ESG score. The computational learning theory refers to a formal mathematical framework that aims to quantify learning tasks and algorithms. It is also called "statistical learning theory." The basic purpose of the computational learning theory is to learn the machine learning algorithm and determine what is understandable. It will help us to know the required data sufficient for the training of a specific algorithm. This theory is related to the design and analysis of machine learning algorithms (association for computational learning). Thus, the study's overall concept is based on the computational learning theory, which is an area of theoretical computing that discusses the design of computer programs and their ability to learn, as well as the identification of computing limitations with machines (Chen et al., 2021a; Lei et al., 2021; Wu and Zhu 2021). The computational learning theory aids in posing and answering concerns about the performance of learning algorithms. Data were collected for non-financial public companies of USA, UK, and Germany for the period of



13 years from 2008 to 2020. Data were collected from Thomson Reuter’s data stream. The population for the research study includes all non-financial companies listed on the New York Stock Exchange (NYSE), London Stock Exchange (LSE), and Frankfurt Stock Exchange. Moreover, we used a sample of 300 companies (100 from each country) from the non-financial sector in the United States, the United Kingdom, and Germany to establish the accuracy of ESG score prediction.

### Variables used

Following accounting, market and sales performance was used for the prediction of the ESG pillar score:

### ML technique and algorithms used

In this study, we used the supervised machine learning techniques like K-nearest neighbor (KNN), polynomial regression, naïve Bayes, random forest, artificial neural networks (ANNs), support vector machine (linear), and support vector machine (Rbf) algorithms to predict ESG score for non-financial companies of UK, USA, and Germany. The dataset for all six predictor features as shown in Table 1 were divided into train set and a test set for processing into the algorithm. The training set for the training model and a testing set for testing models has a ratio of 80:20, respectively. The whole dataset consisted of 3,900 rows, of which 3,120 rows (80%) were used for training and 780 rows (20%) were used for testing. Moreover, Python version 3.9.2 was used to assess values for ESG score prediction. Based on these prediction results, the criteria for selection of algorithms are based on the value of error terms; the less the error term, the higher the accuracy of the algorithm. Figure 1

depicts the proposed approach employed in the study as a block diagram (Arora and Kaur, 2020).

### Performance evaluation of ML techniques

The performance of different supervised technique models used in this study was examined using two evaluation metrics, that is, root mean square error (RMSE) and mean absolute error. These measurements demonstrate how accurate our projections are and how much they differ from the actual data.

### Root mean square error

RMSE is used for measuring the error rate of the regression models. It is an effective metric to compare the forecasting errors and is calculated as follows:

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

where n is the number of test samples,  $y_i$  is the true target value of the  $i$ th sample, and  $\hat{y}_i$  is the forecasted value by the regressor.

TABLE 1 Variables used as predictor features for the ESG score.

Variable	Formula
Return of asset (ROA)	Net income/total assets
Return on equity (ROE)	Net income/total equity
Earnings before interest and taxes (EBIT)	Revenue – cost of goods sold – operating expenses
Earnings per share (EPS)	Net income/outstanding number of shares
Dividend yield (DY)	Annual dividend per share/current share price
Net sales	Gross sales – (sales discounts + sales allowance + sale returns)

## Mean absolute error

MAE is a performance metric used to evaluate the performance of a regressor. It is calculated as follows:

$$\text{MAE} = \frac{\sum_{i=1}^n |y_i - x_i|}{n},$$

where  $n$  is the number of test samples,  $y_i$  is the true target value of the  $i$ th sample and  $x_i$  is the forecasted value by the regressor and  $|\cdot|$  represents the absolute value.

## Analysis and results: ESG pillar score

### Model evaluation based on RMSE and MAE using artificial intelligence models

Table 2 reports the results of different artificial intelligence models like *KNN*, *naïve Bayes*, *SVM linear*, and *SVM RBF* that are based on the final evaluation of RMSE and MAE for selecting the best model. The relationship is built on the basis of the computational learning theory, which states that which machine learning technique best predicts the ESG score. Thus, we used four different techniques of machine learning as cited in Table 2. We run RMSE and MAE measurements on all the four models in order to select the best model out of these four based on the figures of error measurement for the variable ESG pillar score. The objective of the analysis was to select one of these four techniques, which will be the best and most accurate predictor of the ESG pillar score with less error value. We applied two types of SVM kernel functions for the prediction of the ESG pillar score. Considering the aforementioned phenomena, the result in Table 2 indicates that the values of RMSE are less in SVM RBF which is 0.1247 than SVM linear which is 0.4589, while the value of MAE is also less in SVM RBF, that is, 0.01882 than SVM linear which is 0.0598. If the prediction is based on RMSE and MAE, we can say that SVM RBF performs better than SVM linear. Similarly, subject to Table 2, the result obtained from the naïve Bayes technique shows the value of 0.25266 for RMSE and the value of 0.7913 for MAE. In line with previous results, if we compare these results with SVM techniques, we can see that naïve

TABLE 2 Result of error measurement using KNN, SVM linear, SVM RBF, and naïve Bayes.

Technique	RMSE	MAE
KNN	0.1243	0.1456
Naïve Bayes	0.2526	0.7913
SVM linear	0.4589	0.0598
SVM RBF	0.1247	0.0188

TABLE 3 Result of ESG pillar score error measurement using the ANN Technique.

Technique	RMSE	MAE
ESG pillar score ANN 1-2	0.2745	0.0117
ESG pillar score ANN 1-4	0.2894	0.0871
ESG pillar score ANN 1-8	0.9852	0.1122
ESG pillar score ANN 1-16	0.1444	0.2255
ESG pillar score ANN 1-32	0.1942	0.2356
ESG pillar score ANN 1-64	0.3245	0.6667
ESG pillar score ANN 1-128	0.3257	0.1013
ESG pillar score ANN 2-2	0.3985	0.2315
ESG pillar score ANN 2-4	0.2345	0.6419
ESG pillar score ANN 2-8	0.5713	0.9874
ESG pillar score ANN 2-16	0.3197	0.0253
ESG pillar score ANN 2-32	0.1285	0.0125
ESG pillar score ANN 2-64	0.1456	0.0127
ESG pillar score ANN 2-128	0.1287	0.0275
ESG pillar score ANN 3-2	0.1124	0.2369
ESG pillar score ANN 3-4	0.1245	0.1236
ESG pillar score ANN 3-8	0.1265	0.1235
ESG pillar score ANN 3-16	0.1258	0.1297
ESG pillar score ANN 3-32	0.1284	0.3975
ESG pillar score ANN 3-64	0.1278	0.1245
ESG pillar score ANN 3-128	0.1945	0.2356
ESG pillar score ANN 4-2	0.7536	0.3641
ESG pillar score ANN 4-4	0.2314	0.2314
ESG pillar score ANN 4-8	0.8974	0.9746
ESG pillar score ANN 4-16	0.1235	0.0485
ESG pillar score ANN 4-32	0.1244	0.0214
ESG pillar score ANN 4-64	0.0149	0.0124
ESG pillar score ANN 4-128	0.1865	0.0974
ESG pillar score ANN 5-2	0.1973	0.1975
ESG pillar score ANN 5-4	0.1973	0.0971
ESG pillar score ANN 5-8	0.6713	0.2451
ESG pillar score ANN 5-16	0.4597	0.0966
ESG pillar score ANN 5-32	0.0974	0.0971
ESG pillar score ANN 5-64	0.0931	0.0235
ESG pillar score ANN 5-128	0.5491	0.1245

Bayes is better than SVM linear based on less error values for RMSE; however, if we compare the values of naïve Bayes with SVM RBF, we can conclude that SVM RBF outperforms the naïve Bayes, and if the prediction is performed on MAE value, then SVM RBF is also better than naïve Bayes based on the value comparisons in Table 2. Likewise, the results of RMSE obtained from the KNN technique show the least values among the other three techniques. The RMSE value obtained from KNN is 0.1243 which is less than the RMSE value of naïve Bayes (that is, 0.25266), SVM linear (i.e., 0.4589), and SVM RBF (that is,



TABLE 4 Result of ESG pillar score error measurement using hidden layer 1 with 2, 4, 8, 16, 32, 64, and 128 neurons.

Technique	RMSE	MAE
ESG pillar score ANN 1-2	0.2333	0.0117
ESG pillar score ANN 1-4	0.2223	0.0871
ESG pillar score ANN 1-8	0.8738	0.1122
ESG pillar score ANN 1-16	0.1444	0.2255
ESG pillar score ANN 1-32	0.1872	0.2356
ESG pillar score ANN 1-64	0.2244	0.6667
ESG pillar score ANN 1-128	0.6661	0.1013

0.1247). Similarly, the MAE value obtained by applying the KNN technique is 0.1456 which outperforms the results of naïve Bayes (that is, 0.7913). So, overall, we can say that from the aforementioned techniques in the table, the KNN technique outperforms the rest of the three techniques (naïve Bayes, SVM linear, and SVM RBF) based on RMSE, whereas SVM RBF outperforms the rest of the three techniques, that is, naïve Bayes, SVM linear and KNN based on MAE.

## Model evaluation based on RMSE and MAE using artificial neural networks (ANNs)

Table 3 shows the results of ESG pillar score error measurement using the ANN Technique. The results are in the form of ANN, illustrated as ANN w-z means, where w represents the hidden layers and each layer comprises z neurons. Each neuron in the hidden layer receives input from all neurons of previous layers. We found prediction errors in every hidden layer with the help of these neurons. In our study, we predicted the results of the ANN technique by using five hidden layers and neurons up to 128. Here, we first discussed the results of each hidden layer individually with their multiple numbers of neurons and determined prediction results based on the number of neurons in the respective hidden layer. After that, the comparison will be made based on the hidden layer that which layer gives fewer error values.

### Model evaluation based on RMSE and MAE using hidden layer 1

We have summarized the results of hidden layer 1 in Table 4. According to our error measurement criteria, hidden layer 1 with 16 neurons shows less prediction error than the other neurons within hidden layer 1. As illustrated in the aforementioned Table 4, the results show that RMSE and MAE values are less in hidden layer 1 with 16 and 2 neurons, that is, 0.1444 and

TABLE 5 Result of ESG pillar score error measurement using hidden layer 2 with 2, 4, 8, 16, 32, 64, and 128 neurons.

Technique	RMSE	MAE
ESG pillar score ANN 2-2	0.3985	0.2315
ESG pillar score ANN 2-4	0.2345	0.6419
ESG pillar score ANN 2-8	0.5713	0.9874
ESG pillar score ANN 2-16	0.3197	0.0253
ESG pillar score ANN 2-32	0.1285	0.0125
ESG pillar score ANN 2-64	0.1456	0.0125
ESG pillar score ANN 2-128	0.1287	0.0275

TABLE 6 Result of ESG pillar score error measurement using hidden layer 3 with 2, 4, 8, 16, 32, 64, and 128 neurons.

Technique	RMSE	MAE
ESG pillar score ANN 3-2	0.1945	0.2369
ESG pillar score ANN 3-4	0.1245	0.1245
ESG pillar score ANN 3-8	0.1265	0.1235
ESG pillar score ANN 3-16	0.1258	0.1297
ESG pillar score ANN 3-32	0.1284	0.3975
ESG pillar score ANN 3-64	0.1278	0.1236
ESG pillar score ANN 3-128	0.1124	0.2356

0.01177, respectively, than hidden layer 1 with 4, 8, 32, 64, and 128 neurons.

### Model evaluation based on RMSE and MAE using hidden layer 2

Evaluation based on RMSE and MAE using hidden layer 2 is summarized in Table 5. According to our error measurement criteria, hidden layer 2 with 32 neurons shows less prediction error than the other neurons within hidden layer 2. As illustrated in the aforementioned Table 5, the results show that RMSE and MAE values are less in hidden layer 2 with 32 neurons, that is, 0.1285 and 0.0125, respectively, than hidden layer 2 with 2, 4, 8, 16, 64, and 128 neurons.

### Model evaluation based on RMSE and MAE using hidden layer 3

The results of hidden layer 3, based on RMSE and MAE, are shown in Table 6 alongside the AI techniques and figures. According to our error measurement criteria, hidden layer 3 with 128 neurons shows less prediction error for RMSE value than the other neurons within hidden layer 3. As illustrated in aforementioned Table 6, the

TABLE 7 Result of ESG pillar score error measurement using hidden layer 4 with 2, 4, 8, 16, 32, 64, and 128 neurons.

Technique	RMSE	MAE
ESG pillar score ANN 4-2	0.7536	0.3641
ESG pillar score ANN 4-4	0.2314	0.2314
ESG pillar score ANN 4-8	0.8974	0.9746
ESG pillar score ANN 4-16	0.1235	0.0485
ESG pillar score ANN 4-32	0.1244	0.0214
ESG pillar score ANN 4-64	0.0149	0.0124
ESG pillar score ANN 4-128	0.1865	0.0974

TABLE 8 Result of ESG pillar score error measurement using hidden layer 5 with 2, 4, 8, 16, 32, 64, and 128 neurons.

Technique	RMSE	MAE
ESG pillar score ANN 5-2	0.1973	0.1975
ESG pillar score ANN 5-4	0.1973	0.0971
ESG pillar score ANN 5-8	0.6713	0.2451
ESG pillar score ANN 5-16	0.4597	0.0966
ESG pillar score ANN 5-32	0.0974	0.0971
ESG pillar score ANN 5-64	0.0931	0.0235
ESG pillar score ANN 5-128	0.5491	0.1245

results show that the RMSE value is less in hidden layer 3 with 128 neurons, i.e., 0.1124 as compared to hidden layer 3 with 2, 4, 8, 16, 32, and 64 neurons. But if we look at the MAE value, then we can see that it is less in hidden layer 3 with 64 neurons, i.e., 0.1236 as compared to hidden layer 3 with 2, 4, 8, 16, 32, and 128 neurons.

## Model evaluation based on RMSE and MAE using hidden layer 4

We have summarized the results of hidden layer 4 in Table 7. Therefore, according to our error measurement criteria, hidden layer 4 with 64 neurons shows less prediction error than the other neurons within hidden layer 4. As illustrated in aforementioned Table 7, the results show that RMSE and MAE values are less in hidden layer 4 with 64 neurons, that is, 0.01497 and 0.01249, respectively, than hidden layer 4 with 2, 4, 8, 16, 32, and 128 neurons.

## Model evaluation based on RMSE and MAE using hidden layer 5

Table 6 reports the results of hidden layer 5. According to our error measurement criteria, hidden layer 5 with 64 neurons shows less prediction error than the other neurons within

TABLE 9 Result of ESG pillar score error measurement using polynomial regression.

Technique	RMSE	MAE
ESG pillar score Poly 1	0.2455	0.3412
ESG pillar score Poly 2	0.7845	0.1249
ESG pillar score Poly 3	0.8974	0.9713
ESG pillar score Poly 4	0.2314	0.7849
ESG pillar score Poly 5	0.1593	0.1786

hidden layer 5. As illustrated in aforementioned Table 8, the results show that RMSE and MAE values are less in hidden layer 5 with 64 neurons, that is, 0.09312 and 0.0235, respectively, than hidden layer 5 with 2, 4, 8, 16, 32, and 128 neurons.

## Model evaluation based on RMSE and MAE using polynomial regression

In Table 9, the results of polynomial regression up to degree 5 have been shown which are applied to our dataset for accurate prediction of the ESG pillar score using RMSE and MAE as evaluation matrices. We can see that polynomial regression with degree 5 has the least value for RMSE, that is, 0.1593 as compared to RMSE values for polynomial regression with degrees 1, 2, 3, and 4, i.e., 0.2455, 0.7845, 0.8974, and 0.2314, respectively. Likewise, for MAE, polynomial regression with degree 2 outperforms as it has the least value, i.e., 0.1249 among MAE values for polynomial regression with degrees 1, 3, 4, and 5, that is, 0.3412, 0.9713, 0.7849, and 0.17849, respectively. So, for both evaluation parameters, polynomial regression with degrees 5 and 2 outperforms. Therefore, polynomial degree 2 is the best for predicting the ESG pillar score because it has the minimum prediction error during the test.

## Models' evaluations based on RMSE and MAE using decision tree and random forest

In Table 10, the results of two techniques of machine learning, that is, decision tree (individual) and random forests with 2, 3, 5, 7, 10, and 15 decision trees have been shown which are applied to our dataset for accurate prediction of the ESG pillar score using RMSE and MAE as evaluation matrices. We can see that the decision tree model has RMSE and MAE values of 0.6976 and 0.7591, respectively. But the random forest technique with different numbers of decision trees gives relatively lower values of RMSE and MAE than the single decision tree model. If we make predictions based on the RMSE parameter, the random forest technique with 15 decision trees gives the lowest value of RMSE, that is, 0.2245 among the single decision tree and random

TABLE 10 ESG pillar score error measurement using the decision tree and random forest technique.

Technique	RMSE	MAE
ESG pillar score decision tree	0.6976	0.7591
ESG pillar score random forest 2	0.3649	0.3713
ESG pillar score random forest 3	0.2550	0.4923
ESG pillar score random forest 5	0.3599	0.3316
ESG pillar score random forest 7	0.2255	0.3311
ESG pillar score random forest 10	0.5467	0.2915
ESG pillar score random forest 15	0.2245	0.3018

TABLE 11 Comparison of results based on RMSE and MAE of all the techniques used for predicting ESG pillar score.

Technique	RMSE	MAE
ESG pillar score KNN	0.1243	0.1456
ESG pillar score naive Bayes	0.2526	0.7913
ESG pillar score SVM linear	0.4589	0.0598
ESG pillar score SVM RBF	0.1247	0.0188
ESG pillar score decision tree	0.6976	0.7591
ESG pillar score ANN	0.0149	0.0117
ESG pillar score poly	0.1593	0.1249
ESG pillar score random forest	0.2245	0.2915

forest with 2, 3, 5, 7, and 10 decision trees. If we make predictions based on the MAE parameter, the random forest technique with 10 decision trees gives the lowest value of MAE, that is, 0.2915 as compared to a single decision tree and random forest with 2, 3, 5, 7, and 15 decision trees. Therefore, a random forest with a greater number of decision trees gives more accurate results, and the reason for this technique to perform better than a single decision tree is that it has the power of numerous decision trees combined in it and it does not rely on a single decision tree's feature relevance. We can say collectively that for both evaluation parameters, the random forest technique outperforms with the least values of RMSE and MAE.

TABLE 12 Comparison of the order of all the techniques used for predicting ESG pillar score.

ESG pillar score using RMSE parameter
ANN > KNN > SVM RBF > POLY regression > random forest > naive Bayes > SVM linear > decision tree
ESG pillar score MAE parameter
ANN > SVM RBF > SVM linear > naive Bayes > POLY regression > KNN > random forest > naive Bayes

## Comparison of all techniques

As we have discussed the results of all the techniques with their different degrees, nodes, and hidden layers within the same technique and then compared their results with the result of other relevant techniques, now we determine which technique is overall the best fit for accurate prediction of the ESG pillar score among all the techniques used.

From Table 11, we can see that ANN, KNN, and SVM RBF techniques outperform compared to the rest of the techniques as these give the least values of RMSE and MAE. The ANN technique is the best fit for accurate prediction of the ESG pillar score as it is giving the minimum error values, and the reason this technique performs better than other techniques is that it has the power of numerous hidden layers and number of neurons, i.e., 2, 4, 8, 16, 32, 64, and 128. The more the number of hidden layers is added, the more accurate the results with minimum error values are achieved. Finally, the order of accuracy of ESG pillar score prediction using different machine learning techniques is as follows in Table 12.

## Contribution

This study investigates the relationship between ESG and FFP based on artificial intelligence-supervised techniques. Hence, by adopting this latest advancement and convergence of traditional to artificial intelligence techniques, this study is of immense value to investors, academia, regulatory bodies, policymakers, the accounting, and auditing profession, and other relevant stakeholders. In addition, the results of this study will reveal evidence on whether FFP can be enhanced through ESG score and compliance. Apart from these, the results of the study are significantly beneficial for establishing rules pertaining to board of directors' attributes, particularly those in developing capital markets. Importantly, the obtained results will be of utmost help to both developed and developing countries and can be a benchmark for planning and formulating ESG scores and compliance policies in the future.

## Conclusion and recommendations

In the current study, we have used artificial intelligence techniques based on different machine learning algorithms, including random forest, decision tree, artificial neural networks (ANN), K-nearest neighbor (KNN), naive Bayes, support vector machine (linear), support vector machine (radial basis function), and various degrees of polynomial regression. These machine learning algorithms were used to predict future ESG scores using balance sheet and income statement items, that is, return on asset (ROA), return on equity (ROE), and earnings before interest and taxes (EBIT), earnings per share (EPS), net sales (NS),



and dividend yield (DY). The target features in the study are ESG pillar score; the values of these three features are predicted using the balance sheet and income statement items. The findings of this study will help in developing the newly proposed ESG score code following the aforementioned methods, and this study will contribute to the ongoing debate about ESG compliance, issues, and its incorporations into the companies' reporting standards. All in all, the results of this study have significance for regulatory bodies, researchers, academic researchers, practitioners, industries, publicly listed companies, and the security commission to uplift their ESG performance and curb the issues pertinent to ESG. Therefore, consideration should be given to this study to help the stakeholders and improve the overall performance of the organizations of the world. In short, our findings will be of interest to academics and practitioners interested in sustainable finance and sustainable investing. Moreover, talking about the limitation of the study, due to time constraints, this study could not include other developed countries like France, Norway, Denmark, Italy, Canada, Australia, and Holland. This research might be conducted on other countries with similar potential to further authenticate its findings or to identify changes in results about different markets and comprehend this variation about them, to widen its reach and verify its findings. Furthermore, two assessment parameters are used in this work to make predictions. To further evaluate the results, additional assessment factors such as accuracy, f1 score, and confusion matrix might be included. In this current study, only ML algorithms were analyzed, discussed, and compared. In the future, researchers might use both ML approaches and econometric forecasting models to get more exact findings and might be able to determine whether the hybrid of both methods outperforms or which of the individual methods is the best. Future researchers might consider a comparison of ML techniques using examples of developed and developing countries, whereas in this study, we have taken examples of UK, USA, and Germany. Future researchers might compare these findings by taking examples of developing countries like India, China, and Malaysia.

## Study limits and prospective research areas

Apart from several contributions, this study has some limitations. First, as discussed previously, the scope of the study is limited by its population, which included only three developed countries. For this reason, the sample should be expanded to developing countries, and

## References

Ahmed, E., and Hamdan, A. J. I. M. R. (2015). The impact of corporate governance on firm performance: Evidence from. *Bahrain Bourse* 11 (2), 21.

the results should be compared with the developed main market countries. Second, due to data availability and time constraints, this study reports only the selected years for this study; however, future studies could be extended to the proceeding years of future years subject to the availability of data. Last but not the least, future studies should include a purposive sampling procedure for small, medium, and large countries for better insight, understanding, and comparison. Moreover, this study is limited to the machine learning-supervised models which may restrict the generalizability of the findings; therefore, future researchers should utilize unsupervised machine learning techniques. In addition to previous findings, future researchers can take different years, samples, and sectors for examining the impact of ECG mechanisms on FFP based on machine learning and other techniques.

## Data availability statement

Publicly available datasets were analyzed in this study. The data can be found at REFINITIV DATASTREAM.

## Author contributions

MK performed the analysis and data collection. HR contributed to writing the abstract and introduction. MM contributed to the literature, and MA worked on the methodology and final draft of the manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Antonic, m. (2020). Uncovering hidden signals for sustainable investing using Big Data: Artificial intelligence, machine learning and natural language processing. *J. Risk Manag. Financ. Inst.* 13 (2), 106–113.

- Arora, N., and Kaur, P. D. (2020). A Bolasso based consistent feature selection enabled random forest classification algorithm: An application to credit risk assessment. *Appl. Soft Comput.* 86, 105936. doi:10.1016/j.asoc.2019.105936
- Ball, R. (2009). Market and political/regulatory perspectives on the recent accounting scandals. *J. Account. Res.* 47 (2), 277–323. doi:10.1111/j.1475-679x.2009.00325.x
- Cantele, S., Moggi, S., and Campedelli, B. (2020). Spreading sustainability innovation through the co-evolution of sustainable business models and partnerships. *Sustainability* 12 (3), 1190. doi:10.3390/su12031190
- Cardillo, E., and Longo, M. C. (2020). Managerial reporting tools for social sustainability: Insights from a local government experience. *Sustainability* 12 (9), 3675. doi:10.3390/su12093675
- Chen, J., Wang, Q., Huang, J., and Chen, X. (2021a). Motorcycle ban and traffic safety: Evidence from a quasi-experiment at zhejiang, China. *J. Adv. Transp.* 2021, 1–13. doi:10.1155/2021/7552180
- Chen, Y., Kumara, E. K., and Sivakumar, V. (2021b). Investigation of Finance industry on risk awareness model and digital economic growth. *Ann. Oper. Res.*, 1–22. doi:10.1007/s10479-021-04287-7
- Choi, Y., Wang, J., Zhu, Y., and Lai, W. (2021). Students' perception and expectation towards pharmacy education: A qualitative study of pharmacy students in a developing country. *Indian J. Pharm. Educ. Res.* 55 (1), 63–69. doi:10.5530/ijper.55.1.9
- EBA (2020a). Discussion paper on management and supervision of ESG risks for credit institutions and investment firms. Consultation Paper (EBA/DP/2020/03).
- EBA (2020b). Risk assessment of the european banking system. Available at: [https://www.eba.europa.eu/sites/default/documents/files/document\\_library/Risk%20Analysis%20and%20Data/Risk%20Assessment%20Reports/2020/December%202020/961060/Risk20Assessment\\_Report\\_December2020.pdf](https://www.eba.europa.eu/sites/default/documents/files/document_library/Risk%20Analysis%20and%20Data/Risk%20Assessment%20Reports/2020/December%202020/961060/Risk20Assessment_Report_December2020.pdf).
- Hamid, F., Hashim, H. A., and Salleh, Z. (2012). Motivation for earnings management among auditors in Malaysia. *Procedia - Soc. Behav. Sci.* 65, 239–246. doi:10.1016/j.sbspro.2012.11.117
- He, L., Mu, L., Jean, J. A., Zhang, L., Wu, H., Zhou, T., et al. (2022). Contributions and challenges of public health social work practice during the initial 2020 COVID-19 outbreak in China. *Br. J. Soc. Work.* bcac077. doi:10.1093/bjsw/bcac077
- Lei, W., Hui, Z., Xiang, L., Zelin, Z., Xu-Hui, X., and Evans, S. (2021). Optimal remanufacturing service resource allocation for generalized growth of retired mechanical products: Maximizing matching efficiency. *IEEE access* 9, 89655–89674. doi:10.1109/ACCESS.2021.3089896
- Lei, X. T., Xu, Q. Y., and Jin, C. Z. (2021). Nature of property right and the motives for holding cash: Empirical evidence from Chinese listed companies. *MDE. Manage. Decis. Econ.* 43, 1482–1500. doi:10.1002/mde.3469
- MingChia, C. (2012). The influence of workplace spirituality on motivations for earnings management: A study in taiwan's hospitality industry. *J. Hosp. Manage. Tour.* 3 (1), 1–11. doi:10.5897/jhmt11.017
- Moncarz, E. S., Moncarz, R., Cabello, A., and Moncarz, B. (2006). The rise and collapse of enron: Financial innovation, errors and lessons. *Contaduría Adm.* 218, 17–37. doi:10.22201/fca.24488410e.2006.579
- Sokolov, V., Mostovoy, J., Ding, J., and Seco, L. (2021). Building machine learning systems for automated ESG scoring. *J. Impact ESG Invest.* 1 (3), 39–50. doi:10.3905/jesg.2021.1.010
- Vladu, A. B., and Cuzdriorean, D. D. (2013). Financial transparency and earnings management: Insights from the last decade leading journals published research. *Rev. Contab. Dirección* 16, 129–160.
- Wu, Y., and Zhu, W. (2021). The role of CSR engagement in customer-company identification and behavioral intention during the COVID-19 pandemic. *Front. Psychol.* 12, 721410. doi:10.3389/fpsyg.2021.721410
- Yaninen, D. (2017). Artificial intelligence and the accounting profession in 2030. *J. Account. Finance*, 3–29.
- Yang, J., Liu, H., Ma, K., Yang, B., and Guerrero, J. M. (2022). An optimization strategy of price and conversion factor considering the coupling of electricity and gas based on three-stage game. *IEEE Trans. Autom. Sci. Eng.*, 1–14. doi:10.1109/TASE.2022.3171446
- Yao, L., Li, X., Zheng, R., and Zhang, Y. (2022). The impact of air pollution perception on urban settlement intentions of young talent in China. *Int. J. Environ. Res. Public Health* 19 (3), 1080. doi:10.3390/ijerph19031080
- Zhao, L., and Wang, L. (2022). A new lightweight network based on MobileNetV3. *KSII Trans. Internet Inf. Syst.* 16 (1), 1–15. doi:10.3837/tiis.2022.01.001
- Zheng, W., Liu, X., Ni, X., Yin, L., and Yang, B. (2021b). Improving visual reasoning through semantic representation. *IEEE access* 9, 91476–91486. doi:10.1109/ACCESS.2021.3074937
- Zheng, W., Liu, X., and Yin, L. (2021c). Sentence representation method based on multi-layer semantic network. *Appl. Sci.* 11 (3), 1316. doi:10.3390/app11031316
- Zheng, W., Yin, L., Chen, X., Ma, Z., Liu, S., and Yang, B. (2021a). Knowledge base graph embedding module design for Visual question answering model. *Pattern Recognit.* 120, 108153. doi:10.1016/j.patcog.2021.108153