

ESG Score Prediction with Financial Metrics

Aravindkrishna Arivudainambi

Lynbrook High School, 1280 Johnson Ave, San Jose, California, 95129, U.S.A.; aravindkrishna.nambi@gmail.com

ABSTRACT: This study investigates the relationship between various financial metrics and companies' Environmental, Social, and Governance (ESG) scores. Based on a comprehensive analysis of financial data, the research aims to develop a machine-learning model capable of predicting ESG scores. Leveraging datasets from Kaggle and Yahoo Finance, we extracted relevant information, including cash flow statements, income statements, balance sheets, and company-specific attributes such as sector, location, and employee count. The methodology involved data preprocessing, feature engineering, and evaluating multiple regression models, including multilayer perceptron regressor, linear regression, lasso regression, ridge regression, and elastic net regression. The results demonstrate that the Multilayer Perceptron Regressor outperformed the other models, achieving a mean squared error of 0.145 and an R-squared value of 0.989. This indicates its superior ability to capture the complex relationships between financial metrics and ESG scores. The study highlights the significance of financial data in assessing a company's ESG performance and provides insights into the most influential factors contributing to ESG scores. Additionally, it underscores the potential of machine learning techniques in predicting sustainability and ethical practices based on quantitative financial information. The findings have practical implications for investors, analysts, and companies seeking to align their financial strategies with environmental, social, and governance considerations.

KEYWORDS: Robotics and Intelligence Machines, Machine Learning, Finance, Environmental Science, Sustainability.

■ Introduction

Environmental, Social, and Governance (ESG) factors have gained significant prominence in the investment landscape in recent years. Investors, stakeholders, and regulatory bodies increasingly seek to align their financial decisions with sustainable and ethical practices. As a result, there is a growing demand for accurate assessment and prediction of companies' ESG performance.¹ This study aims to leverage the power of machine learning techniques to develop a predictive model capable of estimating ESG scores based on comprehensive financial data.

Historically, ESG evaluations have relied heavily on qualitative assessments and subjective criteria.² However, this approach may overlook the intricate relationships between quantitative financial metrics and a company's commitment to sustainability and ethical governance.³ By harnessing the wealth of information contained within financial statements, such as cash flow statements, income statements, and balance sheets, this research endeavors to uncover the quantitative underpinnings of ESG performance.

The methodology employed in this study involves data acquisition from reputable sources, meticulous feature engineering, and the evaluation of multiple regression models, including advanced techniques like the Multilayer Perceptron Regression. The overarching goal is to identify the most influential financial factors contributing to a company's ESG score and develop a robust predictive model that can assist investors, analysts, and companies in aligning their financial strategies with environmental, social, and governance considerations.

The study found that financial metrics are important in contributing to a company's ESG score. Specifically, the MLP regressor is the model with the most prominent result due to its more complex structure, which can detect more minute patterns within the data. Other regressors used, including linear, Lasso, Ridge, and Elastic Net, provided wretched results with high error and low R-squared values.

■ Literature Review

One of the first pieces of literature I looked at was the "Empirical Study of ESG Score Prediction through Machine Learning—A Case of Non-Financial Companies in Taiwan" by Hsio-Yi Lin and Bin-Wei Hsu. They use machine learning models to predict ESG scores for Taiwan pre-COVID times. Their results are about the performances of different models, and the models that came to the top were RF, ELM, SVM, and XGBoost. The article was created before 2019 and uses Taiwanese datasets, so the results might not completely match our results, but should align closely. They split each category of importance into different features, e.g., the ESG score is divided into environmental, social, and governance scores. Likewise, the financial metric they used was long-term capital adequacy ratio, current ratio, quick ratio, fixed assets turnover ratio, return on operating assets ROA before tax, and operating profit to paid-up capital ratio. The document also splits governance indicators into multiple categories, including but not limited to stock earnings deviation, earnings per share deviation ratio, and total shares. Their results showed significantly high R-squared values, greater than 0.98, for most of their test data amounts.⁴ The metrics they used are closely related to what we are using.

The following article I looked at was “Capital Structure and Speed of Adjustment: The Impact of Environmental, Social, and Governance (ESG) Performance” by Yusuf Babatunde Adenye and Ines Kammoun. This study tries to see how a company’s efforts in being environmentally, socially, and ethically responsible affect how much debt it takes on and how quickly it adjusts its debt to meet its needs. Their results showed that companies with higher ESG must borrow more but are perceived as less reliant on debt by the market. Those excelling in environmental aspects adjust their debt levels faster to meet their finances.⁵ This made me think about debt and its impact on the ESG score, and I realized that it was an important metric to consider when calculating ESG scores based on financial metrics.

The final article I believed impacted the experiment was an economic article for predicting ESG Scores (<https://www.economy.com/products/esg/esg-score-predictor>). This article showed how they can accurately predict ESG scores using sector and location data. This was my first thought of using sector and location, but then I realized that it wouldn’t be as useful because I only have fewer than 500 data points, and they are all in a relatively nearby area. This data covers many countries and jurisdictions, so they have more overall data. This would result in a more accurate model that predicts ESG scores. The article does not provide much information about the training process other than the amount of data they used.⁶

■ Methods

To achieve this goal, data was acquired from Kaggle and Yahoo Finance.^{7,8} The Kaggle dataset provided information about S&P 500 companies and their respective ESG ratings.⁷ The dataset also provides the sector and industry of a company, which has a relationship with the ESG scores. The dataset also includes the address, full-time employees, controversy score, and controversy level. A business’s location affects a company’s environmental score and seems to be an important feature of the model. Other information, including full-time employees, controversy score, and controversy level, could provide insight into the ESG score. We extracted each company’s cash flow, balance sheet, and income statement from the Yahoo Finance dataset.^{7,9} This data was necessary because ESG is commonly used to get investors and capital into a company. Using this knowledge, there can be a correlation between data from cash flow, balance sheet, income statement, and ESG Scores. In the world of corporate sustainability and responsible investing, evaluating Environmental, Social, and Governance (ESG) scores is crucial.¹⁰ Different industries have unique criteria for assessing performance within the ESG framework. For example, energy companies are closely examined for carbon emissions, while the technology sector focuses on data privacy and security for the social aspect of ESG scores. However, analyzing specific sectors, like the S&P 500, can be challenging due to the limited number of companies. Striking a balance between sector-specific insights and broader trends is necessary to thoroughly assess a company’s ESG performance, highlighting the complex relationship between industry dynamics and sustainability metrics.

Geographical factors significantly influence a company’s ESG position, with location data playing a crucial role.¹¹ The region of operation can impact environmental, social, and governance practices. For example, companies in areas with strict environmental regulations may have lower carbon footprints, positively affecting their ESG scores in the environmental category. Social practices are influenced by cultural and societal norms in the company’s location, and governance standards vary based on regional regulatory frameworks. Integrating location data into ESG assessments enhances understanding by considering contextual factors that shape a company’s sustainability and ethical practices.

Full-time employees, controversy score, and controversy level may not provide meaningful insights into the ESG score as they often lack direct correlations with the environmental, social, and governance factors that comprise the ESG framework. While important operationally, the number of full-time employees may not inherently reflect a company’s sustainability practices. Similarly, while indicating potential ethical concerns, the controversy score and controversy level may not align with the criteria evaluated within the ESG framework, making them less directly relevant to the comprehensive assessment of a company’s ESG performance. Therefore, these factors may offer limited value when evaluating a company’s sustainability and ethical standing within the ESG context.

Examining the Cash Flow statement on Yahoo Finance is crucial for calculating a company’s ESG score. This financial metric provides insights into environmental, social, and governance (ESG) aspects.⁵ Positive cash flow not only supports environmentally sustainable initiatives (Environmental Score) but also allows investments in social programs (Social Score). Efficient cash flow management indicates strong governance practices, enhancing transparency and financial stewardship (Governance Score). Analyzing the Cash Flow statement is essential for comprehensively understanding a company’s commitment to sustainability, social responsibility, and effective governance within the ESG framework.

The income statement is crucial in establishing a link between a company’s financial performance and ESG score. Providing a comprehensive overview of revenues, expenses, and profits, the income statement is instrumental in evaluating a company’s environmental, social, and governance (ESG) aspects. The profits and expenditures disclosed in this statement allow the assessment of a company’s commitment to environmentally sustainable practices, which influence the environmental score. Additionally, the income statement provides insights into financial allocations for social responsibility initiatives, impacting the Social Score. As the income statement reflects, transparent financial reporting and responsible resource allocation contribute to strong governance practices, positively affecting the Governance Score. A thorough examination of the income statement facilitates a holistic understanding of a company’s financial health and alignment with ESG principles, guiding investors and analysts in assessing its sustainability and ethical practices.

Completing the financial triad, the balance sheet is crucial for comprehensively understanding a company’s ESG stand-

ing.¹² Detailing assets, liabilities, and equity provides key insights into a company's financial health and its implications for environmental, social, and governance (ESG) considerations. The asset side may reveal investments in sustainable resources, contributing to the Environmental Score. Liabilities and equity illuminate a company's financial obligations and ownership structure, influencing social responsibility and governance practices, respectively. A strong and transparent balance sheet reflects sound financial management, positively impacting the Governance Score. Thoroughly analyzing the balance sheet is integral for investors and analysts, as it assesses a company's overall financial stability and alignment with ESG principles and completes the tripartite evaluation of sustainability and ethical practices.

To achieve the best use of all our data, we extracted multiple columns from all three of the financial triad. The first item we want is the return on equity. We will use assets, liabilities, and net income to determine the shareholders' equity. We want to know how much of the company's worth is made by the investors of the capital. The next item we look for is the Current Ratio. The current ratio is useful for finding a company's ability to cover its liabilities with its assets in the short term.¹³ The next item we look for is earnings per share. This value shows how many shares there are to purchase in the market. The more you must split your ownership with others.¹⁴ The next item we concatenated is the net profit margin. This shows how much profit you earn from each product. The final items we used are free cash flow, which checks if there is money to fund accidental incidents, and operating cash flow, which shows how much cash we have aside for operating.^{15,16} The table below shows how we computed all these values and how they are defined (Table 1).

Table 1: Derived financial metrics and definitions. The table shows the importance of each financial metric.

	Derivation	Description
Return on Equity (ROE)	$ROE = \frac{\text{Net Income}}{\text{Total Assets} - \text{Total Liabilities}}$	Used to find the shareholder's equity. Will provide how much of the company's total worth is made of the investor's capital
Current Ratio (CR)	$CR = \frac{\text{Current Asset}}{\text{Current Liability}}$	Finding the company's ability to cover its liabilities with its assets in the short term
Long-Term Debt to Net Income (LTDNI)	$LTDNI = \frac{\text{Long Term Debt}}{\text{Net Income}}$	This tells how many years it will take for a company to pay back its long-term debt with its income. A more sufficient company should pay it back faster.
Earnings per Share (EPS)	$EPS = \frac{\text{Net Income}}{\text{Shares Issued}}$	How many shares are there to purchase in the market, the more you must split your ownership with others
Net Profit Margin (NPM)	$NPM = \frac{\text{Net Income}}{\text{Total Revenue}}$	How much does the company earn from each product
Free Cash Flow	N/A	Check if there is money to fund accidental incidents
Operating Cash Flow	N/A	How much cash does the company have set aside for operating
Net Earnings	N/A	Profit the company makes

Next was preprocessing the data. To achieve this goal, we first removed all the missing data, which removed six rows from the data. We didn't use imputation because only a small amount of data needed to be removed; imputation seemed un-

necessary. We also did not need to remove any outliers because we were taking data from the last year. Because of this, there is no outlier comparison to conduct.

Standardizing the data with sklearn's StandardScaler, where each data point is transformed by subtracting the mean and dividing by the standard deviation (represented as $z = \frac{x_i - \mu}{\sigma}$), is an essential preprocessing step. This step is important because it makes all the data on the same scale. Standardizing facilitates accurate model training and improves the model's adaptability to varying datasets, fostering a more resilient and dependable predictive performance. Additionally, it enhances the model's ability to handle outliers and ensures more effective convergence during the training process.

Next, we removed the unimportant features of the dataset. We only focused on the data we derived and deemed all the other information unnecessary. The main idea behind this is that all the data we thought necessary, as discussed in Table 1, have incorporated all the required data between each financial triad. One of the main reasons for this is that many columns in the other parts of the data were missing, so removing this would provide very small amounts of data to train with. We cherry-picked the most important features based on two qualities.

Additionally, the feature importance of the given variables was indicative of the factors most responsible for the variance in ESG scores.

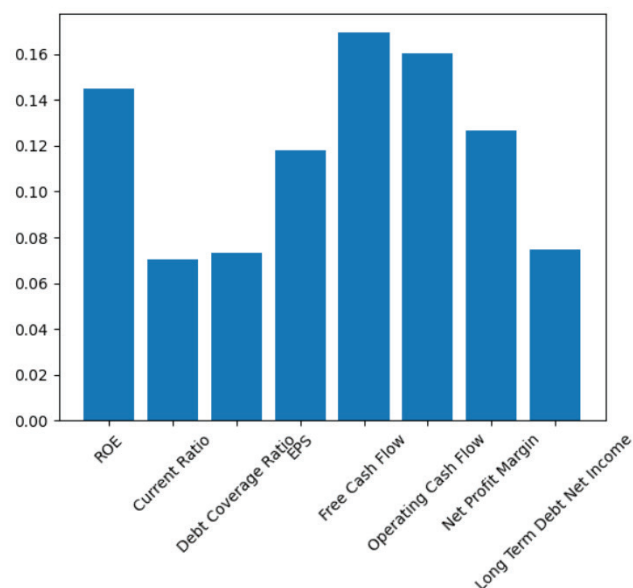


Figure 1: Financial metric feature importance bar chart. The most important feature is free cash flow. All features are significant to the model because each feature's importance is above 0.05.

If the feature importance is above 0.05, the feature is valuable in the training process. This graph shows the different features and their corresponding importance, and all the values are above 0.05, showing their contribution to the model. The highest of these is the free cash flow, which shows its highest importance. We first train a model to calculate these scores and then list the weights for each feature. The importance of the weights is listed in the graph above. The data we believe will

provide optimal results includes ROE, Current Ratio, Debt Coverage Ratio, EPS, Free Cash Flow, Operating Cash Flow, Net Profit Margin, Current Ratio, Long Term Debt Net Income, Location, and Sector. After further analysis, we decided to remove location from the data because all the companies are inside the United States and can skew the data more than help it. We also removed the sector from the dataset because the range of the sector was too broad and would be too diluted in the dataset to provide less gain for the model to pick up on.

The next step to train the model was splitting the dataset into training and testing piles. We used an 80:20 ratio for training and testing. We tested multiple regression-based models, including MLP Regressor, Linear Regression, Lasso Regression, Ridge Regression, and Elastic Net Regression.

The Multilayer Perceptron Regressor is a neural network model that excels in capturing intricate nonlinear relationships within data. In our situation, where we are dealing with a diverse set of financial features and aiming to predict complex ESG scores, the MLP Regressor is well-suited. Its ability to automatically learn and adapt to patterns in large datasets aligns with the complexity of factors influencing ESG performance. However, caution is required as MLPs are prone to overfitting, especially without careful hyperparameter tuning. The computational expense is another consideration, but given the significance of the task, the MLP Regressor offers a powerful tool for modeling the intricate connections between financial metrics and ESG scores. However, unlike many of the models discussed below, the MLP regressor can fit a nonlinear relationship between the explanatory and response variables.¹⁷

Linear Regression is a fundamental and interpretable model that assumes a linear relationship between input features and the target variable. While this model might seem simplistic, its straightforward nature makes it an important baseline for comparison. The linearity may not fully capture the complex connections present in our data. This is calculated using the formula $Y_i = f(X_i, \beta) + e_i$, where Y_i , which is the dependent variable, f is the function, X_i the independent variables, β are the unknown parameters and e_i is the error terms.¹⁸

Lasso Regression, a variant of linear regression, introduces regularization by adding a penalty term based on the absolute values of the coefficients. This encourages sparsity, effectively performing feature selection. In our case, where we have diverse financial features, Lasso Regression becomes valuable for identifying the most influential variables for predicting ESG scores. It aids in streamlining the model by focusing on the most relevant metrics and preventing overfitting. However, careful tuning of the regularization parameter is crucial to balance the feature selection and maintain model performance.^{19,20}

Ridge Regression is another variant of linear regression, but with regularization based on the squared values of the coefficients. It effectively handles multicollinearity, making it suitable for situations where financial metrics may be correlated. In our context, when dealing with a triad of financial data – cash flow, income statement, and balance sheet – Ridge Regression becomes a valuable tool. It helps prevent overfitting

and ensures that the model generalizes well to the complex relationships between financial health and ESG scores.²¹

Elastic Net Regression combines the strengths of L1 (Lasso) and L2 (Ridge) regularization, offering a balanced approach. It is beneficial when dealing with datasets that may have correlated features. In our scenario, where we explore the relationship between various financial metrics and ESG scores, Elastic Net Regression becomes relevant. Simultaneously, addressing feature selection and multicollinearity provides a robust framework for modeling the intricate factors contributing to ESG performance. However, the challenge lies in tuning two hyperparameters, making them more complex than individual Lasso or Ridge models. Careful optimization is required to strike the right balance and harness the strengths of both regularization techniques. The table below summarizes all the models and their positives and negatives (Table 2).^{22,23}

Table 2: Model trade-off analysis based on machine learning theory.

	Advantages	Disadvantages
MLP Regressor	Capable of handling complex patterns	Prone to overfitting, Computational expense, requires careful hyperparameter tuning
Linear Regression	Easy to interpret and understand	Limited to linear patterns
Lasso Regression	Performs feature selection	May struggle with highly correlated features
Ridge Regression	Effective in handling multicollinearity	May not detect patterns that are not easily visible
Elastic Net Regression	Balances feature selection and multicollinearity	May not detect patterns that are not easily visible, More complex than Lasso or Ridge alone

We then trained the models and checked for two different values to see their accuracy. The first value we looked for is the mean squared error. The mean squared error (MSE) measures how close a regression line is to a set of data points. This means the higher the MSE, the less accurate the model is, and the lower the MSE, the more accurate the model is. The subsequent measurement we used to test the accuracy of our model is the R-squared, which we used for the line of best fit at the start of this section. It is a statistical measure that provides information about the goodness of fit of a regression model. The closer the model is to 1.0, the better the model performs.

■ Results and Discussion

We first measured the individual correlations between each item and its correlations with their ESG score. The results for some scatterplots and their corresponding R-squared measurements are shown.

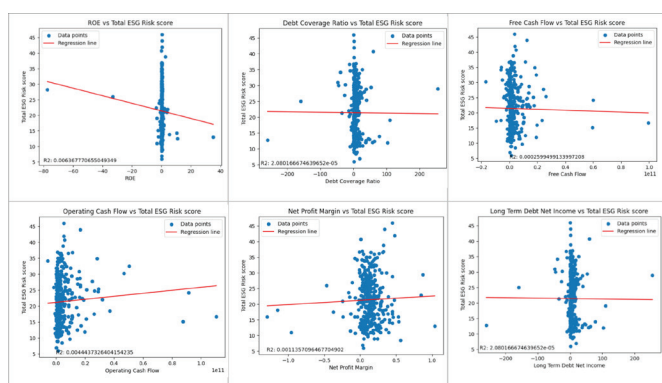


Figure 2: Scatterplot of Financial Metrics vs ESG Score. No apparent correlation is seen between the values.

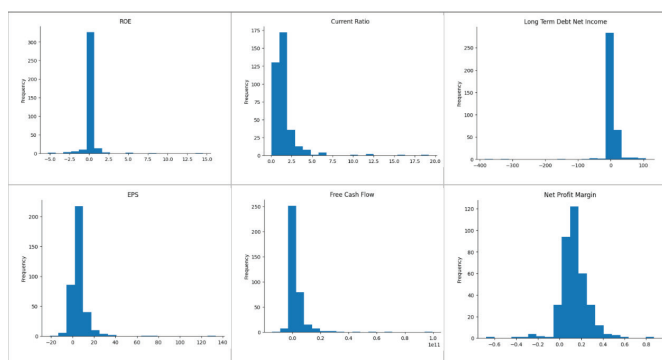


Figure 3: Distribution of financial metrics. Most are normally distributed; the current ratio and EPS are skewed right.

Most graphs, including the current ratio and free cash flow, are skewed to the left. ROE looks evenly distributed, with most items close to 0.0 (Figure 3). The Net Profit Margin is also normally distributed, with most items close to 0.1. The current ratio is skewed to the right. If the items are skewed to the right, they have more power, like for EPS. Skewed left means that the values are closer to 0.

Using all this knowledge from the data histograms and scatter plots, we trained the data with an MLP Regressor, Linear Regression, Lasso Regression, Ridge Regression, and Elastic Net Regression. The results in Table 3 show corresponding R-squared and MSE values and indicate that the MLP regressor is the clear best model. The subsequent section analyzes why the models performed the way they did.

Table 3: Mean Squared Error and R-squared value attained by each model.

	Mean Squared Error (MSE)	R ²
MLP Regressor	0.145	0.989
Linear Regression	18.394	-0.419
Lasso Regression	22.217	-0.714
Ridge Regression	18.475	-0.426
Elastic Net Regression	18.475	-0.426

The R-squared is negative in some cases because this signifies that the model does not follow the trend of the data.²⁴ Another way to put this is that the selected model is worse than choosing a horizontal line for this data. This shows how badly the regressor fits the data. The Mean Squared Error (MSE) and R-squared (R-squared) values provide insights into the performance of different regression models. The MLP

Regressor achieved a low MSE of 0.145 and a high R-squared of 0.989, indicating its better predictive accuracy and ability to explain variance in the data.

In contrast, Linear Regression, Lasso Regression, Ridge Regression, and Elastic Net Regression all exhibited higher MSE values and negative R-squared scores, suggesting poor model fit and limited explanatory power. These traditional linear models may struggle to capture the underlying patterns in the data, resulting in higher prediction errors and negative R-squared values that indicate their inability to outperform a simple mean-based model. The strong performance of the MLP Regressor could be attributed to its ability to capture non-linear relationships in the dataset, making it a more suitable choice for the given task. It can generalize to many shapes like sinusoidal, parabolic, and exponential, for example, allowing detection of more intricate patterns.

The MLP Regressor, deemed the most accurate model, is also practical. After further testing by passing in new data, the model provided close to accurate ESG scores for most cases, but was sometimes off by a lot. Overall, the model delivered believable results, not creating negative numbers or numbers not associated with ESG.

■ Conclusion

This study has demonstrated the potential of machine learning techniques in predicting Environmental, Social, and Governance (ESG) scores based on comprehensive financial data. By leveraging datasets from Kaggle and Yahoo Finance and using the feature engineering process, we identified key financial metrics that exhibit a strong relationship with ESG performance.^{7,8}

The Multilayer Perceptron Regressor emerged as the most accurate model, outperforming traditional linear regression models and variants such as Lasso, Ridge, and Elastic Net regression. Its ability to capture complex, non-linear patterns within the data enabled it to achieve a low mean squared error of 0.145 and a high R-squared value of 0.989, indicating predictive power and variance explanation.

While the model exhibited promising results, there is still room for improvement. The small dataset does not allow for generalization to all industries in the world, but only to the information technology sector. Future research could incorporate larger datasets, encompassing a broader range of companies and industries, to enhance the model's generalizability. For example, including more industries like companies in the Dow Jones Industrial Average will generalize the findings even further. Additionally, exploring ensemble techniques or incorporating qualitative factors may further refine the model's predictive capabilities. Furthermore, using more advanced neural networks like Deep Neural Networks might result in stronger results because they can detect more nuanced patterns in the data.

This study's findings have significant implications for investors, analysts, and companies. By providing a quantitative framework for assessing ESG performance, this research enables more informed decision-making and facilitates the alignment of financial strategies with environmental, social, and governance considerations. Ultimately, this work contrib-

utes to the broader goal of promoting sustainable and ethical business practices, underscoring the potential of data-driven approaches to driving positive change.

■ Acknowledgments

I want to thank my mentor, John Lee, for guiding me to complete this research. He helped me learn more about finance, emphasizing each financial metric's importance. He steered me away from rabbit holes and kept me on the right track. He provided me with motivation and constant support throughout this project. I am truly grateful for the assistance and encouragement throughout this study.

■ References

- Op 5 ESG Investment Trends in 2023. KnowESG. <https://www.knowesg.com/featured-article/top-5-esg-investment-trends-in-2023> (accessed 2024-12-09).
- Adame, L.; Soto, R.; Ballinez, R.; Prieto, M.; Villaseñor, M.; Boni, F. ESG Evaluation Criteria: Environmental, Social and Governance Analysis; HR Ratings: December 2021. Available at: https://www.hrratings.com/docs/metodologia/ESG_Evaluation_Criteria_December_2021.pdf (accessed Dec 2024).
- Kim, M.-S.; Lee, H.-J.; Jung, S.-H.; Lee, S.-Y.; Kim, H.-J. Predicting Future ESG Performance Using Past Corporate Financial Information: Application of Deep Neural Networks. *J. Intell. Inf. Syst.* 2023, 29 (2), 85–100. <https://doi.org/10.13088/JIIS.2023.29.2.085>. (1)
- Lin, H.-Y.; Hsu, B.-W. Empirical Study of ESG Score Prediction Through Machine Learning—A Case of Non-Financial Companies in Taiwan. *Sustainability* 2023, 15, 14106. <https://doi.org/10.3390/su151914106>. (2)
- Adeneye, Y.; Kammoun, I.; Aqilah, S. Capital Structure and Speed of Adjustment: The Impact of Environmental, Social, and Governance (ESG) Performance. *Sustainability Accounting, Management and Policy Journal* 2022, 14. <https://doi.org/10.1108/SAMPJ-01-2022-0060>. (3)
- Moody's Analytics. ESG Score Predictor. <https://www.economy.com/products/esg/esg-score-predictor> (accessed Apr 6, 2024). (4)
- Kaggle. S&P 500 ESG Risk Ratings. <https://www.kaggle.com/datasets/pritish509/s-and-p-500-esg-risk-ratings> (accessed Apr 6, 2024). (5)
- Arrousi, R. Yfinance. <https://pypi.org/project/yfinance/> (accessed Apr 6, 2024). (6)
- D'Amato, V.; D'Ecclesia, R.; Levantesi, S. Fundamental Ratios as Predictors of ESG Scores: A Machine Learning Approach. *Decisions Econ. Finan.* 2021, 44, 1087–1110. <https://doi.org/10.1007/s10203-021-00364-5>.
- Huber, B. M.; Comstock, M. ESG Reports and Ratings: What They Are, Why They Matter. *The Harvard Law School Forum on Corporate Governance*. <https://corpgov.law.harvard.edu/2017/07/27/esg-reports-and-ratings-what-they-are-why-they-matter/> (accessed Apr 6, 2024). (7)
- A.M.O. Hmouda; G. Orzes; Sauer, P. C.; Molinaro, M. Determinants of Environmental, Social and Governance Scores: Evidence from the Electric Power Supply Chains. *Journal of Cleaner Production* 2024, 471, 143372–143372. <https://doi.org/10.1016/j.jclepro.2024.143372>.
- D'Amato, V.; D'Ecclesia, R.; Levantesi, S. ESG Score Prediction Through Random Forest Algorithm. *Comput. Manag. Sci.* 2022, 19, 347–373. <https://doi.org/10.1007/s10287-021-00419-3>.
- Current Ratio: Assess Your Ability to Pay Short-Term Liabilities. Allianz Trade Corporate. https://www.allianz-trade.com/en_US/insights/current-ratio.html.
- Fernando, J. Earnings Per Share (EPS): What It Means and How to Calculate It. Investopedia. <https://www.investopedia.com/terms/e/eps.asp>.
- Fernando, J. Free Cash Flow (FCF). Investopedia. <https://www.investopedia.com/terms/f/freecashflow.asp>.
- Tuovila, A. Operating Cash Flow (OCF) Definition. Investopedia. <https://www.investopedia.com/terms/o/operatingcashflow.as>
- What is Multilayer Perceptron? Deepchecks. <https://www.deepchecks.com/glossary/multilayer-perceptron/>.
- Introduction to Simple Linear Regression 1. Regression Equation. https://murraylax.org/eco307/notes/regression_intro.pdf.
- Agrawal, S. Feature Selection Using Lasso Regression. Medium. <https://medium.com/@agrawalsam1997/feature-selection-using-lasso-regression-10f49c973f08>.
- IBM. What is regression? | IBM. [www.ibm.com. https://www.ibm.com/topics/lasso-regression](https://www.ibm.com/topics/lasso-regression).
- Murel, J. What is ridge regression? | IBM. [www.ibm.com. https://www.ibm.com/topics/ridge-regression](https://www.ibm.com/topics/ridge-regression).
- Brownlee, J. How to Develop Elastic Net Regression Models in Python. Machine Learning Mastery. <https://machinelearningmastery.com/elastic-net-regression-in-python/>.
- Kumar, V. Regularization Methods: Lasso (L1), Ridge (L2) and Elastic Net. Medium. <https://medium.com/@ipvikas/regularization-methods-lasso-l1-ridge-l2-and-elastic-net-2bf5bb35d5ee> (accessed 2024-12-09)
- [https://www.graphpad.com/support/faq/how-can-rsup-2sup-be-negative/#:~:text=R2%20will%20be%20negative%20when%20the%20line%20or%20curve,different%20data%20set\)%2C%20or%20](https://www.graphpad.com/support/faq/how-can-rsup-2sup-be-negative/#:~:text=R2%20will%20be%20negative%20when%20the%20line%20or%20curve,different%20data%20set)%2C%20or%20)
- De Lucia, C.; Pazienza, P.; Bartlett, M. Does Good ESG Lead to Better Financial Performances by Firms? Machine Learning and Logistic Regression Models of Public Enterprises in Europe. *Sustainability* 2020, 12, 5317. <https://doi.org/10.3390/su12135317>.
- https://github.com/aravindkrishna2008/ESG-Research-Paper/blob/main/Training_%2B_Graphs.ipynb

■ Author

Aravindkrishna Arivudaianmbi is a sophomore at Lynbrook High School in San Jose, California. He is interested in machine learning and artificial intelligence and wants to use his knowledge to generate new and innovative ideas for using machine learning to predict the future better.