ESG Scores Predict Stock Performance in the Technology Sector

Jinglin Pei

Durham University, Durham, DH1 3LE, United Kingdom
Mljq48@durham.ac.uk

Abstract. Recently, Environmental, Social, and Governance (ESG) has become increasingly important in investment decisions. As a result, many investors are questioning whether ESG scores can help predict stock performance, especially in the technology sector. This study investigates whether ESG scores can enhance short-term return prediction for technology firms by using Bloomberg data from 2020 to 2025. This study applied Extreme Gradient Boosting (XGBoost) and Random Forest models to predict 20-day returns by combining ESG data with technical indicators. The results showed weak predictive performance with low accuracy and negative R^2 values. Ordinary Least Squares (OLS) regression revealed some significant ESG components, but their effects were inconsistent. Shapley Additive Explanations (SHAP) analysis confirmed ESG features had minimal impact compared to past return data. These findings suggest that traditional ESG scores may lack short-term predictive value. Future research should explore sentiment-based ESG measures, avoid inconsistency between different rating agencies, and improve model design to capture more meaningful insights.

Keywords: ESG scores, Stock return prediction, Technology sector

1 Introduction

Environmental, Social, and Governance (ESG) has become more and more popular for investors as a tool to determine if it is worthwhile to invest in a company. ESG score can, to some extent, indicate a company's responsibility to society. Technology companies nowadays have a huge impact on stock markets due to their rapid growth and frequent technical developments. Many top technology companies, such as Apple, Microsoft, and Google, always have a strong emphasis on their ESG initiatives and contributions. However, predicting stock price movements of technology companies accurately is still challenging, but it is very important. Therefore, investors always seek better ways to predict stock movements clearly and precisely, which makes this research extremely important.

The findings on whether ESG scores predict stock performance remain mixed. For instance, some studies found that there is no evidence supporting ESG's predictive power on stock performance, while others identified a slow but steady influence, particularly for MSCI scores [1-4]. ESG performance has a significantly stronger

A. J. Moshayedi (ed.), *Proceedings of the 2025 International Conference on Hybrid Commerce, Human Capital, and Economic Dynamics (ICHCH 2025)*, Advances in Economics, Business and Management Research 374,
https://doi.org/10.2991/978-2-38476-585-0_24

positive effect on stock returns once inconsistencies in ESG ratings among different rating agencies are corrected [5]. In China's A-share market, ESG performance can notably enhance returns, especially for non-state-owned firms and those in eastern China [6]. Also, incorporating ESG sentiment improves prediction accuracy, and stronger ESG performance generally leads to better long-term returns, whereas weak performance has adverse effects [7, 8]. Governance-related scores show some predictive power, but this is not a settled debate [9].

In terms of methodology, recent studies increasingly apply machine learning models, including XGBoost and Random Forest. For example, a study in China developed an interpretable XGBoost-SHAP framework to predict ESG ratings in the Chinese A-share markets, which achieved over 91% accuracy and an AUC of 97% [10]. This presents the model's robustness and transparency. In Taiwan, both models effectively forecasted ESG scores of non-financial firms, with XGBoost providing slightly better regression accuracy and Random Forest offering stability in Classification [11]. These results support the suitability of tree-based ensemble models for ESG-related stock return prediction.

However, a gap remains in the literature regarding the predictive power of ESG scores within specific sectors, particularly technology. This research aims to examine whether ESG scores can enhance short-term stock return prediction in the technology sector by applying XGBoost and Random Forest models that integrate ESG data with technical indicators to evaluate predictive performance.

2 Data and Methodology

2.1 Data Cleansing and Preparation

The dataset used for this research was obtained from the Bloomberg Terminal at Durham University [12]. Bloomberg is known for its reliability, standardisation, and comprehensive coverage of financial and ESG indicators, making it highly suitable for academic and forecasting research. The dataset included daily stock prices and ESG scores of 8 top technology firms in the S&P 500, and the time horizon is from 2020 to 2025. A data cleansing process was completed, and some additional features were added to help predict returns better. These features included moving averages of ESG Scores over the past 30 days, past returns, and cumulative returns over different time periods (1, 2, 5, 10, and 20 days). These steps will help understand how stock prices move in the short term.

2.2 Modelling Approaches

This research used two machine learning methods, including regression and classification. For each of the two methods, XGBoost and Random Forest were applied to predict the actual stock returns over the next 20 days and to predict whether stock prices would go up or down during the same period. XGBoost was chosen for its scalability and regularization capabilities. It builds decision trees sequentially, each correcting previous errors, which can help improve the model's accuracy [13]. The

results show that XGBoost can outperform traditional models in both accuracy and speed, especially in structured data prediction tasks such as financial time-series prediction [13]. Random Forest was also chosen because of its robustness and strong generalisation ability. Using randomly selected data and features helps prevent overfitting and improve model stability [14]. Therefore, Random Forest is competitive for many prediction tasks and often complements boosting methods like XGBoost in comparative analyses [14].

To test these models properly, the dataset was split using the TimeSeriesSplit method. This method can ensure that the dataset will be separated by time. This method is often used to preserve the time ordering in the data and can ensure that training is performed on past data and testing is performed on future data, which is important in time-dependent forecasting tasks [15].

For regression models, the research checked their accuracy by using common measures such as R^2, Mean Squared Error (MSE). For classification models, accuracy, precision, recall, F1-score, and AUC were measured. This research also adopts SHAP values for both models to clearly show which features mattered most in the models. As SHAP value can uniquely satisfy properties of local accuracy, consistency, and missingness, it makes it the most robust interpretability method for complex models such as XGBoost and Random Forest [16]. In addition, a basic linear regression model (OLS) was used to test if ESG scores showed statistically significant relationships with future returns, which can help to clarify the machine learning findings

3 Results

3.1 Regression Result

Table 1. Regression Metrics (XGBoost)

Fold	R^2	MSE
1	-1.5299	0.0172
2	-0.2042	0.0137
3	-0.1578	0.0132
4	-0.3353	0.0185
5	-0.0988	0.0329
Average	-0.4652	0.0191

Table 2. Regression Metrics (Random Forest)

Fold	R^2	MSE
1	-1.5137	0.017
2	-0.0419	0.0119
3	-0.0229	0.0117
4	-0.1162	0.0154
5	-0.0126	0.0303
Average	-0.3414	0.0172

Table 1 and Table 2 show the performance of the XGBoost and Random Forest regression models. Both models produce negative R^2 values across all folds with averages of -0.4652 and -0.3414 respectively, indicating their poor predictive performance. In addition, the average MSE was 0.0191 for XGBoost and 0.0172 for Random Forest, suggesting that ESG scores do not provide meaningful value in predicting short-term stock returns.

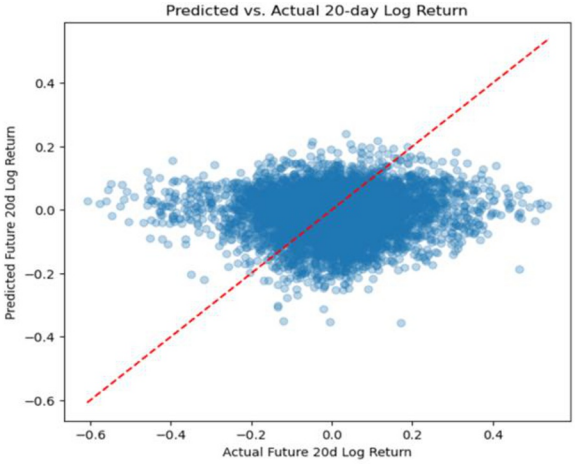

Fig. 1. Scatter plot of predicted vs actual 20-day returns (Picture credit: Original)

According to Fig. 1, the predictions are widely scattered and do not align closely with the red diagonal line, which represents perfect prediction. This visual pattern confirms that the model failed to capture the true variation in stock returns, which can further support the weak performance shown in the R^2 and MSE metrics.

Table 3. OLS Regression Results

Variable	Coefficient (Std. Error)
ESG Score	-0.2938***
	(0.101)
Env Score	0.0935*
	(0.050)
Soc Score	0.1138***
	(0.041)
Gov Score	0.1578***
	(0.033)
ESG Score MA30	0.2769***
	(0.101)
Env Score MA30	-0.0837*
	(0.050)
Soc Score MA30	-0.1043**
	(0.041)
Gov Score MA30	-0.1562***
	(0.033)
Log Return Lag1	0.0053

Continue **Table 3**

Variable	Coefficient (Std. Error)
	(0.053)
Log Return Lag2	-0.0213
	(0.053)
Log Return Lag5	0.0481
	(0.051)
Log Return Lag10	0.0172
	(0.049)
Log Return Lag20	-0.052
	(0.047)
Cum Return Lag5	0.0149
	(0.036)
Cum Return Lag10	0.0130
	(0.029)
Cum Return Lag20	-0.0171
	(0.016)

Significant at 1% *** 5% ** 10% *

Table 3 presents the OLS regression results. Variables including ESG score, social score, governance score, and ESG score MA30 are all positively related and statistically significant. In contrast, the moving averages of environmental, social, and governance scores show negative coefficients with governance score MA 30 but are also highly significant. These results mean that ESG components may influence stock returns. However, most lagged and cumulative return variables are not significant, which indicates limited short-term predictive power from past price movements.

3.2 Classification Results

Table 4. Classification Metrics (XGBoost)

Fold	Accuracy	Precision	Recall	F1-score	AUC
1	0.4636	0.5933	0.3144	0.411	0.5315
2	0.4795	0.4798	0.6528	0.5531	0.4706
3	0.5013	0.5961	0.5598	0.5774	0.4816
4	0.5119	0.5901	0.5063	0.545	0.4964
5	0.5529	0.5556	0.822	0.663	0.5638
Average	0.5019	0.563	0.5711	0.5499	0.5088

Table 5. Classification Metrics (Random Forest)

Fold	Accuracy	Precision	Recall	F1-score	AUC
1	0.4478	0.6173	0.19	0.2906	0.5164
2	0.4848	0.4879	0.8887	0.6299	0.5121
3	0.5688	0.6169	0.7685	0.6844	0.5086
4	0.4742	0.5453	0.5372	0.5413	0.4631
5	0.5364	0.5358	0.9988	0.6975	0.5282
Average	0.5024	0.5607	0.6766	0.5687	0.5057

Tables 4 and 5 present the classification performance of XGBoost and Random Forest across five folds. Both models indicate the average accuracy is close to random guessing, with XGBoost at 50.2% and Random Forest at 50.1%. In addition, although some folds showed relatively higher precision and recall, the overall F1-scores remained modest, which are 0.5499 for XGBoost and 0.5678 for Random Forest. In terms of AUC values, the results of 0.5088 and 0.5057 further confirm that the models had a weak ability to distinguish between positive and negative return cases. In conclusion, these results suggest that ESG scores do not effectively support binary classification of short-term stock return direction.

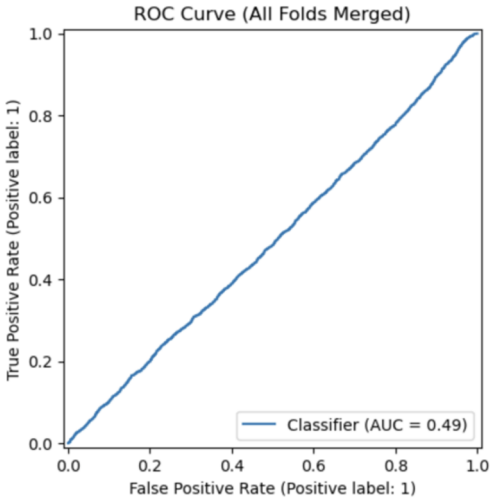

Fig. 2. ROC Curve (Picture credit: Original)

Fig. 2 shows the ROC curve for the classification model after merging all folds. According to the Fig. 1, the curve lies close to the diagonal, and the AUC is 0.49, which is slightly below the 0.5 threshold of random guessing. This result suggests that the classifier was unable to distinguish between positive and negative return cases based on ESG scores, which further reinforces the earlier finding that ESG data lack short-term predictive value in this context.

3.3 SHAP Analysis

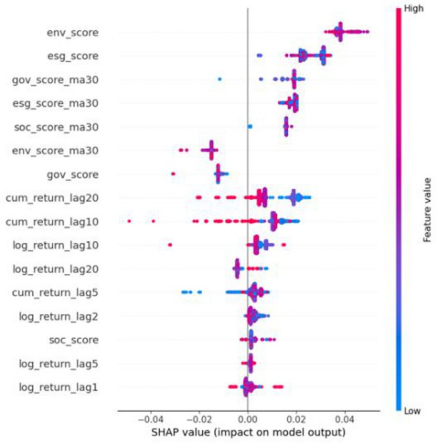

Fig. 3. SHAP value-Regression (Picture credit: Original)

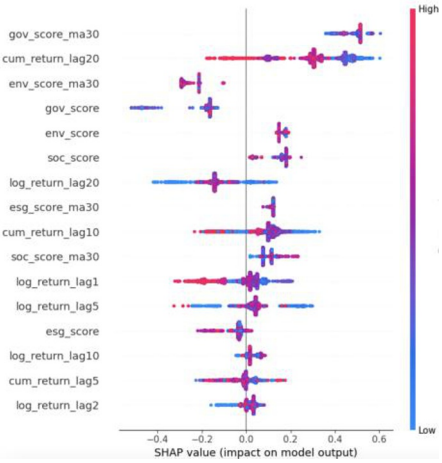

Fig. 4. SHAP value-Classification (Picture credit: Original)

Figs. 3 and 4 show SHAP value summaries for the regression and classification models, respectively. In both Figs. 3, 4, features based on past returns and volatility show the strongest influence on model predictions. ESG-related variables, including ESG score, environmental score, and their 30-day moving average, consistently appear in the lower part of the ranking with minimal SHAP values. This suggests that ESG scores contributed very little to either predicting the size of returns or the direction of movement. The results confirm that historical price data had much more predictive weight than ESG features in this short-term forecasting task.

4 Discussion

The results show that ESG scores do not effectively predict short-term stock returns in the technology sector. Both XGBoost and Random Forest models present low accuracy and negative R^2 values, which means weak model performance. This supports prior findings that argue that ESG scores provide little or no predictive power [1-3]. Thus, while these models are strong at capturing non-linear patterns and dealing with high-dimensional data, their performance here suggests a lack of short-term predictive power.

Although the OLS regression indicated some statistically significant ESG components, such as social and governance scores, the signs were mixed, and the effects were modest. This suggests that while ESG may have some association with returns, it is inconsistent and likely not useful for short-term forecasting. SHAP analysis further confirmed this, which presented that ESG-related features had very low influence compared to past return indicators.

A previous study found that ESG sentiment data can significantly improve stock return predictive accuracy, but this study shows that traditional ESG scores have limited continuation compared to historical price data [7]. The finding indicates that predictive power may depend on the type of ESG information used. For example, the incorporation of ESG sentiment and optimized feature engineering can improve XGBoost prediction accuracy, which suggests that future models could consider more sources of ESG inputs [10]. Furthermore, the inconsistency of ESG ratings across rating agencies may limit their utility in high-frequency prediction tasks [3]. Future research could address these limitations by applying a noise correction technique. For example, a previous study applies the noise-correction technique by instrumenting ESG scores using other agencies' ratings and this method significantly improved estimation accuracy with regression coefficients increasing by a factor of 2.1 on average [5].

5 Conclusion

This research explored whether ESG scores can be used as predictors of short-term stock returns in the technology sector using regression and classification models. The results showed that ESG scores had limited predictive power with an accuracy level is around 50% and R^2 values being negative across all folds. In addition, the traditional OLS regression confirmed that only a few ESG variables had weak and mixed relationships with returns. Finally, SHAP analysis further supported these findings by showing that ESG scores had very little influence on Model prediction compared to past return features. Overall, the results indicate that ESG scores are not suitable for predicting short-term stock movements in the tech sector.

However, this study has certain limitations. Firstly, it only used ESG scores from Bloomberg, which may lead to bias with other ESG rating agencies and may not fully capture the broader aspects of ESG information. Secondly, the sample is limited to eight major technology firms, reducing generalisability across sectors or market sizes. Therefore, in order to eliminate these limitations, future research could incorporate

ESG sentiment data or alternative ESG rating sources to enhance prediction accuracy. Expanding the sample size and testing across different industries or time horizons may also provide deeper insights into ESG's predictive role in the dynamic stock markets.

References

1. Alves, R., Krüger, P., van Dijk, M.: Drawing up the bill: Are ESG ratings related to stock returns around the world? Journal of Corporate Finance **93**, 102768 (2025)
2. Landi, G., Sciarelli, M.: Towards a more ethical market: the impact of ESG rating on corporate financial performance. Social Responsibility Journal **15**(1), 11–27 (2019)
3. Gibson, B.R., Krueger, P., Schmidt, P.S.: ESG Rating Disagreement and Stock Returns. Financial Analysts Journal **77**(4), 104–127 (2021)
4. Berg, F., Heeb, F., Kölbel, J.: The Economic Impact of ESG Ratings. SSRN Electronic Journal (2022)
5. Berg, F., Koelbel, J.F., Pavlova, A., Rigobon, R.: ESG Confusion and Stock Returns: Tackling the Problem of Noise. National Bureau of Economic Research (2022)
6. Yin, X.-N., Li, J.-P., Su, C.-W.: How does ESG performance affect stock returns? Heliyon **9**(5), e16320 (2023)
7. Lee, H., Kim, J.H., Jung, H.S.: Deep-learning-based stock market prediction incorporating ESG sentiment and technical indicators. Scientific Reports **14**(1), 10262 (2024)
8. Assael, J., Carlier, L., Challet, D.: Dissecting the Explanatory Power of ESG Features on Equity Returns by Sector, Capitalization, and Year with Interpretable Machine Learning. Journal of Risk and Financial Management **16**(3), 159 (2022)
9. Khan, M.: Corporate Governance, ESG, and Stock Returns around the World. Financial Analysts Journal **75**(4), 103–123 (2019)
10. Zhang, J., Zhao, Z.: Corporate ESG rating prediction based on XGBoost-SHAP interpretable machine learning model. Expert Systems with Applications **295**, 128809 (2025)
11. Lin, H.-Y., Hsu, B.: Empirical Study of ESG Score Prediction through Machine Learning—A Case of Non-Financial Companies in Taiwan. Sustainability **15**(19), 14106 (2023)
12. Bloomberg: S&P 500 stock returns and ESG scores (2020–2025). Bloomberg Terminal, accessed via Durham University Library Bloomberg Terminal (2025)
13. Chen, T., Guestrin, C.: XGBoost: a Scalable Tree Boosting System. In: Proceedings of the 22nd ACM SIGKDD Conference, pp. 785–794. ACM, New York (2016)
14. Bentéjac, C., Csörgő, A., Martínez-Muñoz, G.: A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review **54**(3) (2020)
15. Cerqueira, V., Torgo, L., Mozetič, I.: Evaluating time series forecasting models: an empirical study on performance estimation methods. Machine Learning **109** (2020)
16. Lundberg, S.M., Lee, S.I.: A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems **30** (2017)