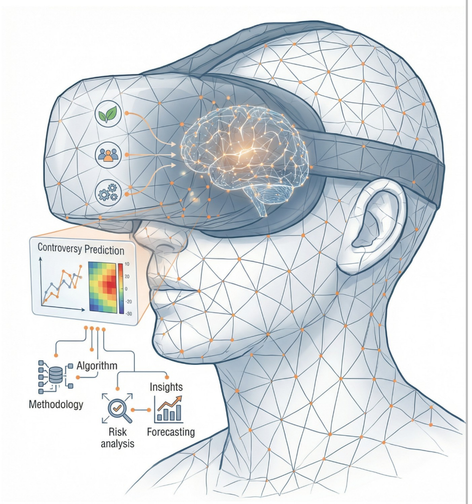
UC | Universidad de Cantabria

Modeling ESG Controversies: Methods and Evidence from Machine Learning

Elisa Baraibar-Diez, María D. Odriozola, Ignacio Llorente

University of Cantabria



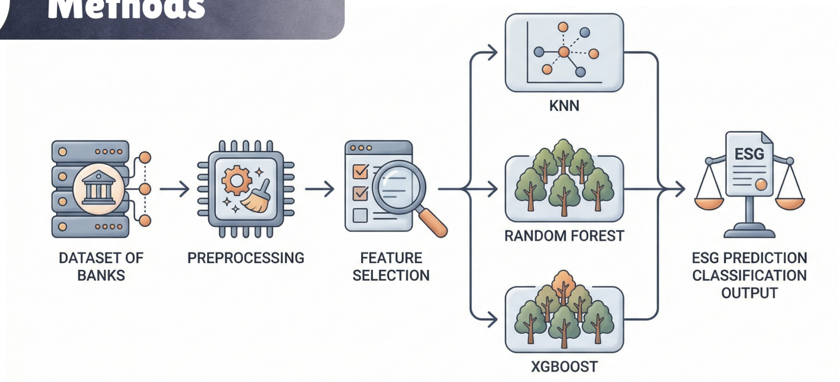
1 Introduction

Environmental, Social and Governance (ESG) criteria have become a central pillar in the evaluation of companies by investors, regulators, and society. ESG goes beyond traditional financial performance by focusing on how organizations manage environmental impact, social relationships, and governance practices in an ethical and transparent manner. In the European Union, this trend is reinforced by regulatory frameworks such as the EU Taxonomy, the Corporate Sustainability Reporting Directive (CSRD), and the Sustainable Finance Disclosure Regulation (SFDR).

The growing importance of ESG factors has generated a significant increase in the volume of ESG-related data and the number of information providers. However, strong ESG policies and detailed ESG reporting do not always reflect real corporate behavior. ESG controversies — including environmental accidents, labor conflicts, governance failures, and ethical scandals — represent negative media coverage that reveals failures in corporate ESG performance.

Artificial intelligence and data analytics are increasingly relevant for analyzing these issues and predicting controversy levels. Classification algorithms can be used to predict expected levels of controversy for a given company

2 Methods



This research adopts a quantitative approach based on supervised machine learning techniques to predict corporate ESG controversy levels. Supervised learning models are trained on labeled data to classify or predict outcomes.

The sample includes 215 banks from 34 countries using Refinitiv Eikon data from 2019–2023. The dependent variable is the company ESG controversy level, transformed into a four-level factor: very low, low, moderate, and high controversy.

Predictor variables include ESG ratings, firm size, board characteristics (diversity, skills, size, independence), CSR committee presence, whistleblowing protection, financial performance (ROA), and debt levels.

Three machine learning algorithms were applied:

- K-Nearest Neighbors (KNN)
- Random Forest
- XGBoost

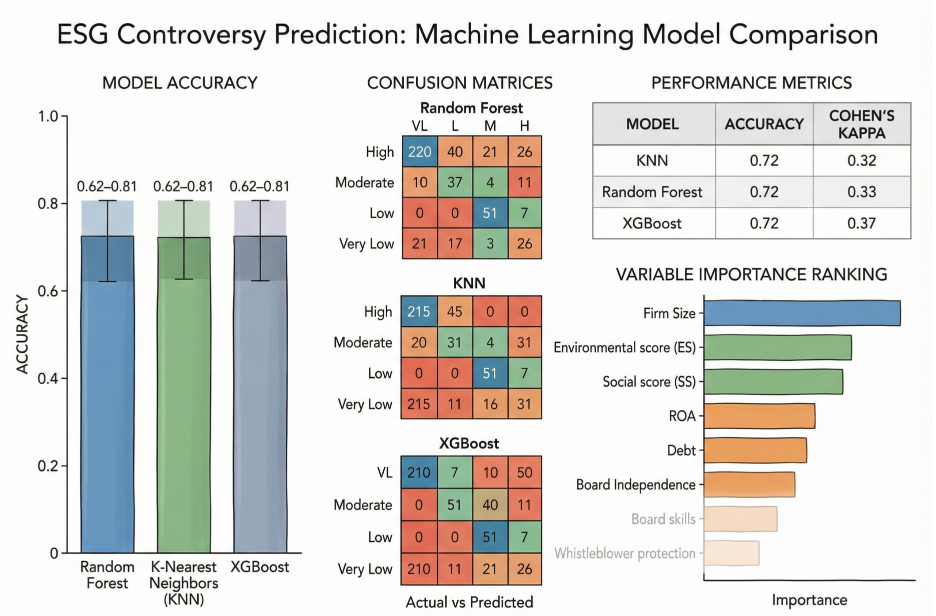
These models were used to classify companies according to predicted ESG controversy levels and evaluate predictive performance.

3 Results

All three algorithms achieved similar predictive accuracy of approximately 72%. However, the No Information Rate (69%) indicates that model performance only slightly exceeds baseline prediction. Cohen's Kappa values suggest weak but positive agreement between predicted and actual values.

Model performance varies across classes. All models perform well in identifying companies with very low controversy levels, but they struggle to predict moderate controversy cases. XGBoost performs better for low controversy classification, while Random Forest and XGBoost show better sensitivity for high controversy cases.

Variable importance analysis indicates that firm size, environmental score, and social score are among the most influential predictors. Board skills and whistleblowing protection appear less relevant. Financial indicators such as ROA and debt also contribute to predictive power.



4 Discussion

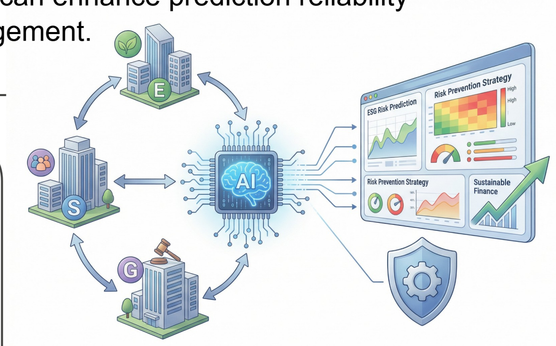
The results show that predicting ESG controversies using machine learning is feasible but challenging. The limited improvement over baseline prediction suggests that current data and models still face constraints. One major limitation is the imbalance in the dataset, with a high proportion of companies classified as low controversy.

Differences in algorithm performance reinforce the importance of using multiple predictive systems rather than relying on a single model. Combining approaches may improve robustness and predictive capacity. Additionally, expanding datasets and incorporating country or industry factors could improve model reliability and explanatory power.

5 Conclusion

Despite moderate predictive performance, the study highlights strong potential for using machine learning to anticipate ESG controversies. Companies can use predictive models as a form of "insurance policy" to identify risks early and implement preventive strategies.

The research emphasizes the need for improved data balance, larger samples, and integration of contextual factors such as country and industry. Using multiple algorithms can enhance prediction reliability and support better ESG risk management.



6 References

Agnese, P., Cerciello, M., Oriani, R., & Taddeo, S. (2024). "ESG controversies and profitability in the European banking sector." Finance Research Letters, 61, 105042.
Aouadi, A., & Marsat, S. (2018). "Do ESG controversies matter for firm value? Evidence from international data." Journal of Business Ethics, 151, 1027-1047.
Dipierro, A. R., Toma, P., & Frittelli, M. (2024). "Introducing ESG controversies as the polluting factor of banks' activity: a nonparametric efficiency approach." Journal of Economic Studies.
Elamer, A. A., & Boulhaga, M. (2024). "ESG controversies and corporate performance: The moderating effect of governance mechanisms and ESG practices." Corporate Social Responsibility and Environmental Management.
Mahyoub, M., Roslan Ja'afar, & Nur Laili, A. G. (2024). "ESG controversies and banking performance: The moderating effect of board activity." Asian Economic and Financial Review, 14(12), 895-913

RESEARCHERS' TUESDAY

The images included in this poster were generated using artificial intelligence through the software Illustrae