PREDICTING CORPORATE GREENHOUSE GAS EMISSIONS

Cluster Innovation Centre
University of Delhi

Mentor : Dr. Sachin Kumar

Keshav Sharma (152122@cic.du.ac.in) | Mayank Arora 152129@cic.du.ac.in

Abstract

This study explores a novel approach employing machine learning techniques for predicting corporate GHG emissions, with a specific focus on the Indian subcontinent. By leveraging a comprehensive dataset and advanced regression models, the research aims to improve Scope 3 emission prediction accuracy. Findings indicate significant implications for corporate sustainability practices and environmental management.

Introduction

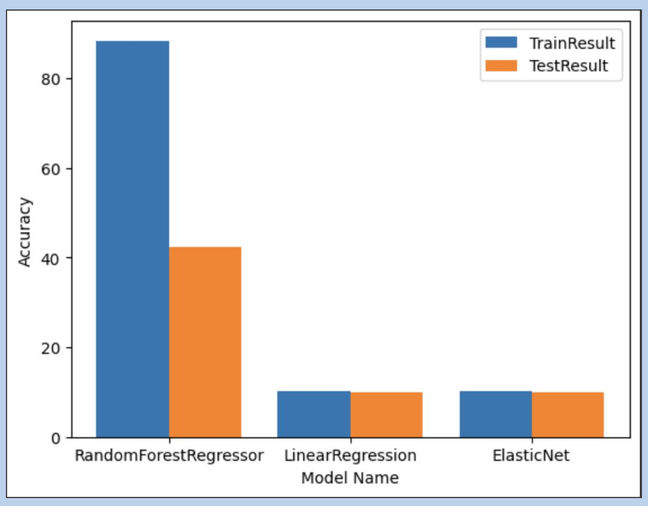
Greenhouse gas (GHG) emissions significantly impact the global environment, necessitating accurate prediction models for effective management. Traditional methods, particularly for Scope 3 emissions, face limitations in data accuracy and complexity. Recent studies have utilized machine learning to improve predictions, but challenges remain due to data scarcity and the intricate nature of indirect emissions. This research employs advanced machine learning techniques, including continuous models and LSTM networks, to enhance GHG emission predictions. By leveraging a comprehensive dataset from the CDP, this study aims to provide more accurate predictions, contributing to improved corporate sustainability practices and environmental management.

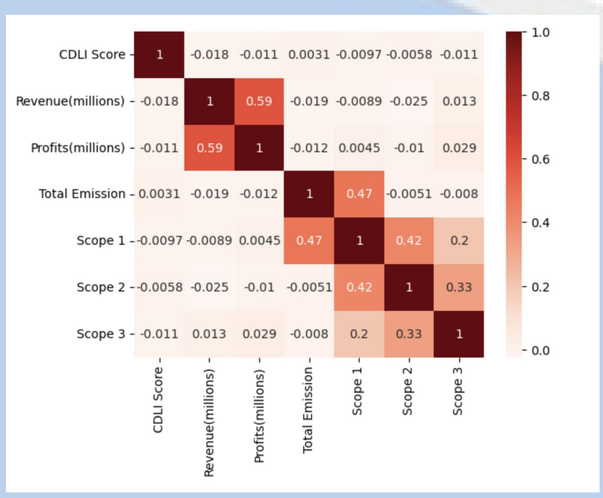
Research Methodology

This study utilizes data from the CDP, covering GHG emissions and revenue from 2006 to 2014. The data preprocessing involved merging datasets and handling missing values by replacing them with the mean. The independent variables included various financial and environmental metrics, with Scope 3 emissions as the dependent variable. This comprehensive dataset allowed for robust model training and evaluation. We employed a combination of continuous models like RandomForestRegressor, LinearRegression, ElasticNet and an LSTM model. The LSTM model, designed for time-series data, was trained with 200 epochs, a batch size of 20, and the Adam optimizer with a learning rate of 0.001. These models were evaluated using metrics such as MSE, RMSE, and MAE to ensure accurate and reliable GHG emission predictions.

Results

Performance was evaluated using MSE, RMSE, and MAE. The LSTM model demonstrated superior accuracy in predicting future GHG emissions compared to traditional continuous models. Key findings indicate significant future changes in emissions for companies like ANZ Bank and Talisman Energy, showcasing the effectiveness of the proposed methodology.





Conclusion

Our machine learning model successfully predicts GHG emissions, with significant implications for improving sustainability practices. However, the model faces limitations due to overfitting from mean replacement of missing values. Future work should explore alternative data handling techniques to further enhance prediction accuracy. Overall, this research enables small companies to estimate GHG emissions more accurately and cost-effectively, supporting better environmental management.

References

- Klaaßen, L., & Stoll, C. (2021). Harmonizing corporate carbon footprints. Nature communications, 12(1), 1-13.
- Han, Y., Gopal, A., Ouyang, L., & Key, A. (2021). Estimation of corporate greenhouse gas emissions via machine learning. arXiv preprint arXiv:2109.04318.
- Nguyen, Q., Diaz-Rainey, I., & Kuruppuarachchi, D. (2021). Predicting corporate carbon footprints for climate finance risk analyses: a machine learning approach. Energy Economics, 95, 105129.