WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Data-Driven Prediction of CO_2 Emissions

Suganth R[1], Mr. K. Vivekanadan[2]

[1]III B.SC.CT, Department of Computer Technology, Sri Krishna Adithya College of Arts &Science Coimbatore .

[2]M.C.A.,M.Phil.,Ph.D Assistant professor, Department of Computer Technology, Sri Krishna Adithya college of Arts &Science Coimbatore.

ABSTRACT:

Climate change remains a critical global challenge, and the reduction of carbon emissions is essential in mitigating its effects. The increase in global temperatures, extreme weather patterns, and rising sea levels have made carbon emission reduction an urgent necessity. This study explores the use of machine learning models to predict carbon dioxide (CO2) emissions, identifying key factors contributing to environmental pollution. The research implements Random Forest Regression, Gradient Boosting Regressor, and Linear Regression models to forecast emission trends based on industrial activities, energy consumption, and socio-economic factors. The study also compares the performance of different machine learning models using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared values. The results indicate that machine learning techniques provide accurate insights for policymakers and industries in formulating strategies to reduce carbon footprints. Furthermore, this study highlights the potential of artificial intelligence in sustainability, enabling the development of automated monitoring and prediction systems that can support environmental conservation efforts.

1. INTRODUCTION:

The rapid industrialization and economic expansion of the past century have significantly increased greenhouse gas emissions, particularly CO2, which is a major contributor to global warming. Various international agreements, such as the Paris Agreement, aim to limit emissions, but achieving these targets remains a challenge due to policy limitations and economic constraints. Traditional statistical methods have been used for emission forecasting, but machine learning (ML) offers an advanced approach for analyzing vast datasets and predicting future trends.

The demand for energy has continuously increased due to urbanization and the growing global population. Fossil fuels remain the dominant source of energy, leading to the release of massive amounts of carbon dioxide. The development of alternative energy sources such as wind, solar, and hydroelectric power can contribute to reducing emissions, but these technologies are not yet widespread enough to replace fossil fuels entirely. This paper aims to bridge the gap by applying machine learning models to analyze existing emission trends and provide actionable insights for reducing emissions.

2. LITERATURE STUDY

Several studies have investigated CO2 emission trends using statistical and ML techniques. Traditional time-series models, such as the Seasonal Autoregressive Integrated Moving Average (SARIMA), have been effective in analyzing historical emission patterns. However, machine learning models, such as Random Forest and Gradient Boosting, provide enhanced predictive accuracy by incorporating multiple variables, including energy consumption, GDP, and industrial production. Research highlights that AI-based carbon tracking tools, such as Carbontracker and eco2AI, play a crucial role in reducing emissions by optimizing energy consumption in computing processes.

A study conducted by Bhatt et al. (2023) emphasized the significance of using artificial intelligence in climate research. Their research suggested that ML models could be used not only for forecasting emissions but also for analyzing policy impacts. Similarly, Giannelos et al. (2024) explored the effectiveness of machine learning in predicting CO2 emissions from the building sector, highlighting the potential role of AI in sustainable construction and urban planning.

3. METHODOLOGY

The methodology involves data collection, preprocessing, model selection, and evaluation.

3.1 Data Collection:

The dataset includes historical emission records spanning several years, sourced from industrial reports, government records, and environmental databases. It consists of variables such as production values, energy consumption, and CO2 emissions.

Data from agencies such as the Environmental Protection Agency (EPA), World Bank, and United Nations Climate Change Division are utilized to ensure accuracy and reliability.

3.2 Data Preprocessing:

Data cleaning involves handling missing values, normalizing numerical data, and performing feature selection to retain the most relevant variables for model training. Feature engineering is applied to create meaningful attributes, and outlier detection techniques are used to eliminate anomalies that could impact the accuracy of predictions.

3.3 Machine Learning Models:

Linear Regression: Used to establish a relationship between emission factors and CO2 levels. This model serves as a baseline for comparison with more complex algorithms.

Random Forest Regressor: A decision-tree-based model that improves predictive accuracy through ensemble learning by aggregating multiple decision trees.

Gradient Boosting Regressor: A powerful model that optimizes prediction by iteratively reducing errors in previous models. This model is particularly effective in handling non-linear relationships in data.

3.4 Model Evaluation:

Models are assessed based on MSE, MAE, R-squared (R^2), and Root Mean Squared Error (RMSE) to determine the best-performing model for emission prediction. Cross-validation techniques such as k-fold validation are employed to ensure the robustness of the models.

4. IMPLEMENTATION

- The implementation phase includes developing a system that integrates ML models for real-time emission forecasting.

Front-end Development: The interface is developed using Python-based frameworks such as Flask or Streamlit to allow user interaction with the prediction system. Users can input parameters such as industrial production values and energy consumption levels to generate emission forecasts.

Back-end Processing: Google Colab and Kaggle are used for model training and computation. Cloud-based solutions such as AWS and Google Cloud Platform are explored for scalable deployment.

Deployment: The trained models are deployed using cloud platforms to facilitate dynamic data input and real-time predictions. A web-based dashboard is developed to visualize emission trends using interactive graphs and reports.

5. RESULT

- The machine learning models were trained and tested using emission datasets, and their performance was evaluated. The results demonstrated:
- **Linear Regression:** Provided moderate accuracy but struggled with complex relationships.
- **Random Forest:** Achieved higher accuracy due to its ability to handle multiple variables and non-linear relationships.
- **Gradient Boosting:** Delivered the best performance with the lowest error rate and highest R-squared value, making it the most suitable model for emission prediction.

- The findings indicate that machine learning techniques can effectively predict carbon emissions, providing essential insights for environmental planning and policy development. Furthermore, visualization techniques such as heatmaps and trend analysis graphs demonstrated the variation in emissions over time, reinforcing the importance of proactive policy-making.

6. CONCLUSION

This research demonstrates the potential of machine learning in predicting CO_2 emissions, helping policymakers and industries adopt data-driven strategies for emission reduction. The ability of machine learning models to analyze large datasets, detect complex patterns, and provide accurate forecasts makes them invaluable tools in combating climate change.

Future research should focus on incorporating real-time data streams from various sources, including satellite imagery, IoT sensors, and industrial monitoring systems. The integration of such data sources will enhance the model's predictive accuracy and provide more timely insights for decision-makers. Additionally, enhancing model interpretability through explainable AI (XAI) techniques will help policymakers trust and effectively utilize machine learning insights for regulatory purposes.

Moreover, developing hybrid models that combine deep learning, reinforcement learning, and traditional statistical approaches can further improve long-term forecasting accuracy. Implementing AI-driven decision support systems will enable governments and industries to simulate different policy scenarios and predict the effectiveness of emission reduction strategies before implementation.

By integrating AI-driven approaches, sustainable environmental practices can be achieved more effectively. Machine learning applications in emissions reduction extend beyond forecasting; they can also optimize industrial processes, improve energy efficiency, and enhance carbon capture technologies. Policymakers, researchers, and industries must collaborate to harness these technological advancements to accelerate the global transition to carbon neutrality. In conclusion, AI and machine learning hold significant potential in shaping a greener future, making data-driven

7. REFERENCES:

(1) IPCC, "Climate Change 2023: Mitigation of Climate Change," Intergovernmental Panel on Climate Change, 2023.

(2) Pielke, R. A., "Climate Prediction and Climate Policy," Nature Climate Change, 2022.

(3) Goodfellow, I., Bengio, Y., Courville, A., "Deep Learning," MIT Press, 2016.

(4) Xu, R., et al., "Machine Learning Applications in Environmental Sustainability," Renewable Energy, 2023.

(5) Kumar, S., et al., "Carbon Footprint Reduction Strategies: A Data-Driven Approach," Journal of Cleaner Production, 2024.

(6) Anderson, T., et al., "Predicting Greenhouse Gas Emissions with AI," Environmental Modelling & Software, 2023.

(7) Le Quéré, C., et al., "Trends in Global CO2 Emissions," Earth System Science Data, 2023.

(8) Smith, J., "The Role of AI in Climate Change Mitigation," Environmental Research Letters, 2022.

(9) Chen, B., et al., "Integrating Machine Learning with Climate Science," Journal of Climate Change, 2024.

(10) Ray, S., "Artificial Intelligence for Sustainable Development," IEEE Transactions on Sustainable Computing, 2023.

(11) Zou, H., et al., "Application of Deep Learning in Predicting Environmental Pollution," Science of the Total Environment, 2024.

(12) Yeo, J., "Carbon Trading and AI-Based Forecasting," Journal of Environmental Economics, 2023."Climate Change 2023: Mitigation of Climate Change," Intergovernmental Panel on Climate Change, 2023.