OpenCarbon: An Open-Access Carbon Prediction Platform

[1]BH Bhujanga Reddy, [2]K Teja Sai, [3]G Sreeja, [4] M Seetharam

[1]*Assistant Professor Computer Science and Engineering Geethanjali College of Engineering and Technology Hyderabad, Telangana, India*

[2,3,4] *UG Student Computer Science and Engineering Geethanjali College of Engineering and Technology Hyderabad, Telangana, India*

Abstract- Climate change and rising carbon emissions are major global problems that need smart solutions for sustainable decision-making. This paper introduces OpenCarbon, an open-access AI-based platform for predicting carbon emissions and analyzing sustainability. It is designed to study carbon emission datasets, provide predictive insights, and raise environmental awareness. The system features a flexible machine learning pipeline that uses INFO optimization and an Optimized Extreme Learning Machine (ELM) model for precise emission predictions. The platform allows users to upload datasets dynamically, automate data processing, generate key performance indicators, visualize results with explainable AI using SHAP, and forecast future emissions. Built with FastAPI, HTML, JavaScript, and SQLite, the system offers an interactive dashboard and user-friendly analytics for those without advanced technical skills. For validation, it uses several real-world datasets, including OWID CO_2 data, World Energy Consumption, and country-specific carbon emission datasets. The findings show strong prediction accuracy, scalable performance, and valuable sustainability insights. OpenCarbon is a lightweight, smart tool for analyzing carbon emissions, monitoring the environment, and supporting sustainability decisions.

***Keywords*-Carbon Emissions, Sustainability Analytics, Explainable AI, ELM Model, Forecasting, FastAPI, SHAP, Environmental Data Analysis.**

I. INTRODUCTION

Global warming and climate change are significant environmental issues caused mainly by rising carbon dioxide (CO_2) emissions. Governments, researchers, and environmental groups need smart analytical tools to track emission trends, predict future values, and gain insights on sustainability. Traditional carbon analysis systems are often fixed, tied to specific datasets, and lack clarity, which limits their effectiveness for flexible environmental analytics.

Recent developments in Artificial Intelligence (AI) and Machine Learning (ML) have made predictive modeling and data-driven sustainability analysis possible. However, most current platforms either concentrate solely on visualization or need complex technical skills to use. There is a clear need for an open-access, smart, and flexible platform that can automatically analyze carbon datasets and give understandable predictions.

To tackle these limitations, this paper presents OpenCarbon, an AI-based platform for carbon prediction and sustainability analytics. The system is built to work with any dataset, is easy to use, and can produce predictive insights, KPI analytics, and clear visualizations from carbon emission data. The platform blends data preprocessing, improved machine learning, and forecasting methods to aid environmental decisions and research. Figure 1 shows the overall workflow and design of the OpenCarbon platform to help clarify the system's process.

II. LITERATURE SURVEY

Several studies have looked into using machine learning for analyzing environmental data and predicting carbon emissions. Traditional statistical models like linear regression and time-series forecasting have been commonly used to analyze emission trends. However, these models often struggle to capture the complex nonlinear relationships found in large-scale environmental datasets.

Recent research has focused on using AI-based predictive models for sustainability analytics. Machine learning techniques like Random Forest, Neural Networks, and Support Vector Machines have shown better results in emission prediction tasks. Also, explainable AI methods such as SHAP (SHapley Additive Explanations) have been introduced to help interpret model decisions and improve transparency in environmental analytics systems.

Open datasets like OWID CO_2 data and World Energy Consumption datasets are widely used in climate research to examine global emission patterns. However, many current systems do not combine forecasting, explainability, and interactive dashboards in one platform. Additionally, most platforms do not adjust to different dataset structures and need manual setup.

The proposed OpenCarbon platform addresses these gaps by combining machine learning, explainable AI, forecasting, and interactive dashboards into one open-access system for analyzing carbon emissions. The system uses an adaptive preprocessing pipeline, an INFO-optimized ELM prediction model, and SHAP-based explainability to provide clear, transparent, and scalable emission analytics.

III. PROPOSED SYSTEM

A. System Overview

OpenCarbon is a complete AI-driven web platform that allows for smart carbon emission analysis through automatic data processing, predictive modeling, and clear visualization. The system uses a modular design with frontend, backend, AI analytics engine, and database layers.

OpenCarbon is a full AI-driven web platform. It enables smart analysis of carbon emissions through automatic data processing, predictive modeling, and clear visualization. The system has a modular design. It includes frontend, backend, AI analytics engine, and database layers.

B. System Architecture

As shown in Fig. 1, the overall layout of the OpenCarbon platform brings together dataset upload, preprocessing, optimized ELM prediction, SHAP explainability, and forecasting in one AI-driven pipeline.

The backend manages dataset processing, model execution, forecasting, and API communication. The frontend offers interactive dashboards and visualization panels. The SQLite database keeps user sessions, dataset history, and analytical results for easy access.

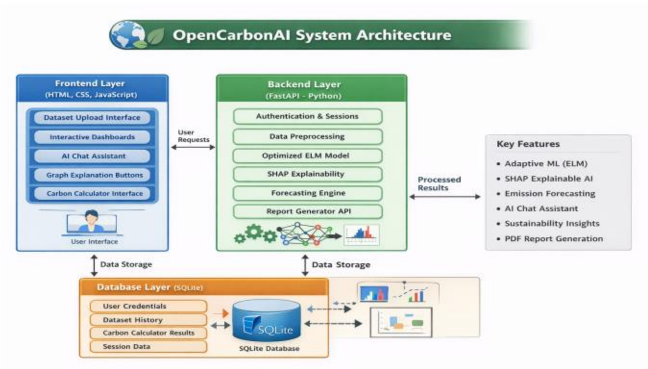

Fig. 1. Overall Architecture of OpenCarbon Platform

IV. METHODOLOGY

A. Data Collection and Datasets

The system uses several real-world carbon emission datasets, including:

- Carbon (CO2) Emissions by Country
- OWID CO_2 Dataset
- World Energy Consumption Dataset
- Custom carbon_data CSV datasets

These datasets include emissions, energy use, and environmental indicators for each country. They are used for predictive analysis.

B. Data Preprocessing

The preprocessing module automatically does:

- Missing value handling
- Data cleaning and normalization
- Feature selection
- Target column detection
- Dataset scaling and splitting

This preprocessing helps the platform use different carbon datasets without needing manual setup.

C. Machine Learning Model (INFO + Optimized ELM)

The main prediction engine uses an Optimized Extreme Learning Machine (ELM) model that incorporates the INFO optimization algorithm. The ELM model offers fast training speeds and effective generalization, making it a good fit for environmental datasets.

The INFO optimizer improves model parameter selection and boosts prediction accuracy. We evaluate model performance using metrics like MSE, RMSE, and R^2 score.

D. Explainable AI using SHAP

To ensure transparency, the system uses SHAP explainability to interpret feature importance and model predictions. SHAP summary plots and feature impact graphs help users see how variables like energy consumption, region, and historical emissions affect prediction outcomes.

E. Forecasting Engine

The forecasting module examines past emission trends and predicts future carbon emission values using time-series learning patterns. The system allows for global and country-specific forecasting, depending on the available data.

As shown in Fig. 2, the forecasting module examines past carbon emission trends and creates future emission predictions. This helps with long-term sustainability analysis and provides insights for policy-making.

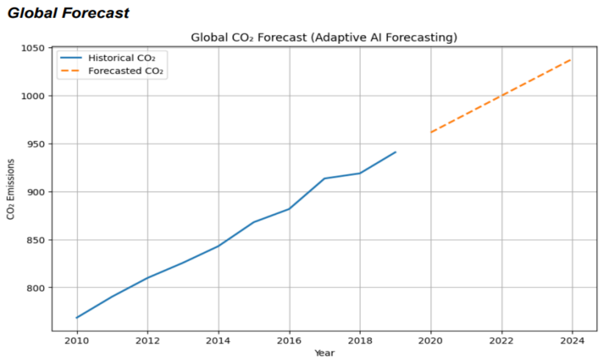

Fig. 2. *Future Carbon Emission Forecast Based on Historical Dataset*

V. System Modules

A. User Interface Module

The frontend interface is built with HTML, CSS, and JavaScript. It allows users to upload datasets and includes KPI dashboards, visualization panels, and features for interacting with AI. The design focuses on being simple, interactive, and easy to use.

As shown in Fig. 3, the dataset upload and analysis interface lets users upload carbon emission datasets and start automated AI-based analysis using an interactive and easy-to-use dashboard.

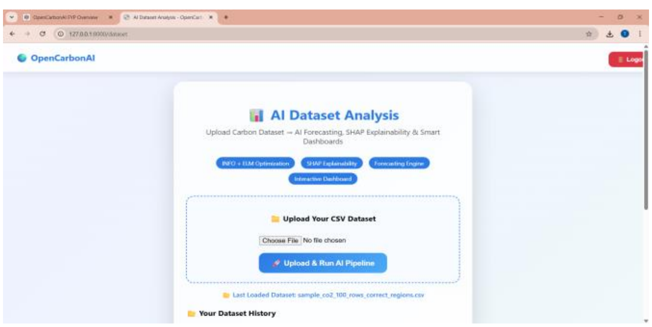

Fig. 3. Dataset Upload and Analysis Interface of OpenCarbon Platform

B. Backend Processing Module

The FastAPI backend manages dataset uploads, preprocessing, model execution, forecasting, and API responses. It also serves cached results to avoid reprocessing the dataset repeatedly.

C. AI Analytics Module

This module brings together ELM prediction, SHAP explainability, KPI generation, and sustainability insight analysis. It makes up the main intelligence layer of the OpenCarbon platform.

D. Database Module

SQLite database stores user credentials, dataset history, analytical results, and session-based information. This setup ensures lightweight and efficient storage.

VI. Results and Discussion

A. Dashboard Analysis

The OpenCarbon platform offers an interactive dashboard that shows important sustainability KPIs like average CO_2 emissions, maximum emissions, minimum emissions, emission trends, and the countries with the highest emissions. The dashboard updates automatically based on the uploaded dataset.

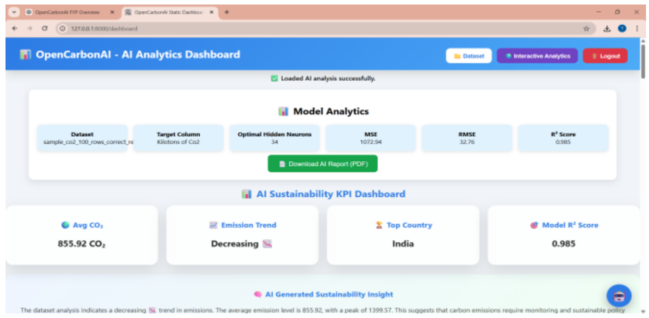

Fig. 4. Interactive Dashboard of OpenCarbon Showing KPI

As shown in Fig. 4, the interactive dashboard offers real-time KPI analytics. It includes average emissions, emission trends, and comparisons by country for a detailed environmental analysis.

B. Prediction and Model Performance

Experimental evaluation with several carbon emission datasets showed good prediction accuracy and stable model performance. The optimized ELM model reached reliable R^2 scores and low error metrics. This indicates a strong ability to predict outcomes for environmental data analysis.

As shown in Fig. 5, the optimized ELM model shows a close match between actual and predicted carbon emission values. This indicates strong prediction accuracy and reliable performance.

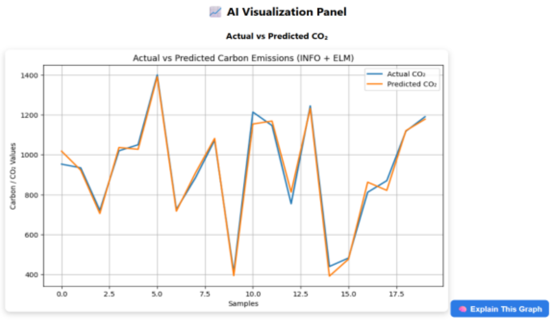

Fig. 5. Actual vs Predicted Carbon Emission Values Using Optimized ELM Model

C. Visualization and Explainability

The system produces several visual outputs, including:

- Actual vs Predicted Emission Graph
- CO_2 Trend Analysis
- SHAP Feature Importance Plot
- Forecast Graphs

As shown in Fig. 6, the SHAP feature importance plot explains how different environmental and energy-related features contribute to predicting carbon emissions. This improves the model's transparency and interpretability.

These visualizations improve understanding and support research-level environmental analysis.

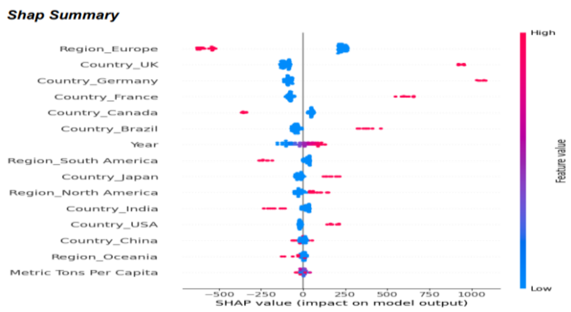

Fig. 6. *SHAP Feature Importance Analysis for Carbon Emission Prediction*

D. System Efficiency and Usability

The platform works well on regular laptop settings because of its lightweight structure and improved model design. The open-access and web-based interface makes it easy for students, researchers, and environmental analysts to use.

VII. Conclusion

This paper introduced OpenCarbon, an open-access platform that uses AI for predicting carbon emissions and analyzing sustainability. The system combines data preprocessing, machine learning, explainable AI, forecasting, and interactive dashboards into a single platform. Experimental results with real-world carbon emission datasets show good prediction performance and valuable sustainability insights, along with easy scalability. This platform helps climate analytics research by offering a clear, intelligent, and user-friendly solution for monitoring carbon emissions and supporting sustainability decisions.

VIII. Future Scope

Future improvements of the OpenCarbon platform may include the integration of real-time IoT environmental sensors, deep learning forecasting models, and cloud-based deployment for large-scale climate analytics. Additional features like geo-spatial emission mapping, policy recommendation systems, and multilingual sustainability insights can further enhance the platform's impact in environmental research and decision-making.

TABLE I. PROJECT MODULES AND DESCRIPTION

Project Module	Description
User Interface	Handles dataset upload, dashboard and visualization
Preprocessing Module	Data cleaning, normalization and feature selection
AI Prediction Module	INFO optimized ELM model for carbon prediction
Explainability Module	SHAP-based feature importance and interpretation
Forecasting Module	Future carbon emission prediction
Database Module	SQLite storage for datasets and results

TABLE II. TECHNOLOGY STACK

Component	Technology
Frontend	HTML, CSS, JavaScript
Backend	FastAPI (Python)

Database	SQLite
Data Processing	Pandas, NumPy
Visualization	Matplotlib
Explainable AI	SHAP
Machine Learning	Optimized ELM + INFO

TABLE III. TEAM CONTRIBUTION

Team Member	Contribution
Member I	AI Model Development, Data Preprocessing, and Prediction Module Implementation
Member II	Frontend Development and Documentation
Member III	Backend Support, Model Testing, Dataset Handling, Deployment, and System Performance Evaluation

ACKNOWLEDGMENT

The authors would like to thank Geethanjali College of Engineering and Technology for providing the academic environment and resources that supported this research work.

The authors express their sincere appreciation to BH Bhujanga Reddy for her valuable guidance, constructive feedback, and continuous support throughout the development of the Project.

The authors also acknowledge the support of the Department of Computer Science and Engineering for facilitating the successful completion of this work.

REFERENCES

[1] Y. Zhang et al., "An examination of daily CO_2 emissions prediction through a comparative analysis of machine learning, deep learning, and statistical models," Environmental Science and Pollution Research, 2025.

[2] J. Hofstetter et al., "Prediction of carbon dioxide levels in the European Alps based on machine learning algorithms," Applied and Computational Engineering, 2024.

[3] P. Mishra et al., "Neural network approach to carbon emission prediction in industrial and power sectors," Discover Applied Sciences, 2025.

[4] R. Singh et al., "Machine Learning Techniques for Multifactor Analysis of National CO_2 Emissions," arXiv preprint, 2025.

[5] M. Rahman and T. Chowdhury, "SHAP-Driven Explainable AI for Climate and Emission Forecasting Systems," Applied Energy, 2024.

[6] S. Patel and R. Kumar, "Performance Optimization of ELM-Based Carbon Emission Models Using Meta-Heuristic Algorithms," IEEE Access, 2024.

[7] L. Wang and R. Gupta, "Machine Learning Approaches for Environmental and Emission Forecasting: A Comprehensive Review," Environmental Modelling & Software, 2023.

[8] J. Liu, H. Chen, and X. Zhao, "AI-Based Carbon Emission Prediction Using Hybrid Machine Learning Models," Sustainable Computing: Informatics and Systems, 2023.

[9] K. Sharma and A. Verma, "Deep Learning and Time-Series Models for Global CO_2 Emission Forecasting," Journal of Cleaner Production, 2024.

[10] S. Das et al., "Explainable Artificial Intelligence in Environmental Monitoring and Climate Analytics," Environmental Research Letters, 2023.

[11] M. Ali, S. Khan, and N. Ahmad, "Data-Driven Carbon Emission Analysis Using Predictive Analytics and AI Techniques," Energy Reports, 2024.

[12] H. Zhang and Y. Li, "Sustainability Analytics Using Machine Learning for Climate Change Assessment," Renewable and Sustainable Energy Reviews, 2023.

[13] R. Kumar, P. Gupta, and S. Mehta, "Forecasting Carbon Emissions Using AI and Big Data Analytics," Sustainable Energy Technologies and Assessments, 2024.

[14] T. Nguyen and D. Tran, "Intelligent Environmental Decision Support Systems Using Explainable AI," Applied Soft Computing, 2023.