e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

ADVANCED MODELING TECHNIQUES FOR PREDICTING VEHICLE CO2 EMISSIONS

Sangamesh MD*1

*1PES Institute Of Technology And Management, Shivamogga, Karnataka, India.

ABSTRACT

This project aims to predict CO2 emissions from vehicles using machine learning techniques applied to a dataset encompassing vehicle characteristics such as make, engine size, transmission type, and fuel consumption metrics. The primary objective is to develop an accurate regression model capable of forecasting CO2 emissions, crucial for understanding environmental impact and ensuring compliance with emissions regulations across diverse vehicle types. The workflow involves comprehensive data preprocessing to handle missing values and categorize variables, followed by exploratory data analysis to uncover correlations and insights. The chosen model, Linear Regression, is trained and assessed using metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The model is serialized using pickle for seamless integration into a Flask-based web application, facilitating real-time predictions. By providing stakeholders with a tool to estimate and mitigate vehicle emissions effectively, this project contributes to sustainable automotive practices and regulatory transparency.

Keywords: CO2 Emissions, Vehicle Emissions Prediction, Machine Learning, Automotive Industry, Environmental Impact.

I. INTRODUCTION

The CO2 emissions from vehicles are a major contributor to environmental pollution and climate change. With the global push towards reducing carbon footprints and adhering to stricter emissions regulations, accurate prediction of vehicle emissions has become increasingly important. Traditional methods for estimating CO2 emissions are often outdated, inaccurate, and not user-friendly, relying on static databases and generalized formulas that do not cater to individual vehicle characteristics. The advent of machine learning offers a promising solution to these challenges, enabling more precise and personalized predictions. This project aims to leverage machine learning algorithms to develop a web-based application for predicting CO2 emissions based on specific vehicle attributes such as engine size, transmission type, and fuel consumption metrics. By integrating a trained regression model into a user-friendly Flask application, we can provide real-time, accurate predictions that are accessible to consumers, automotive professionals, and regulators alike. This innovative approach not only enhances the accuracy of emissions predictions but also facilitates easier access to critical environmental data, supporting informed decision-making and fostering greater environmental responsibility.

II. LITERATURE REVIEW

1. Klemm and Wiese explore indicators crucial for optimizing sustainable urban energy systems through detailed energy system modeling. Their study emphasizes the importance of comprehensive indicators in achieving efficiency and sustainability goals in urban energy planning and management.
2. Sajede et al.investigate various factors influencing pollutant emissions across different modes of transportation. They highlight the complexities involved and provide insights into how different factors contribute to emissions, offering valuable data for developing strategies to mitigate environmental impacts from transportation.
3. Menendez and Ambühl discuss the implementation of design and operational measures aimed at promoting sustainable mobility, drawing lessons from Zurich's initiatives. Their analysis underscores the role of integrated approaches in achieving sustainable urban mobility solutions and addresses key challenges and successes observed in practice.
4. Guo et al examine the effects of adding humic acid to controlled-release fertilizers on summer maize yield, nitrogen use efficiency, and greenhouse gas emissions. Their findings suggest potential agricultural practices that could enhance both productivity and environmental sustainability by reducing greenhouse gas emissions associated with fertilizer use.

5. Fang and Mueller quantify the embodied carbon footprint of mortise-and-tenon joinery in modern timber construction, offering insights into alternative structural connections that could lower carbon emissions. Their research contributes to sustainable building practices by identifying specific construction techniques that minimize environmental impacts.
6. Górka et al. conducted an assessment focusing on greenhouse gas (GHG) interactions near a municipal waste landfill site. Their case study provides insights into the environmental impacts associated with landfill activities, highlighting factors influencing GHG emissions in such settings.
7. Stubenrauch et al. analyze European forest governance in relation to achieving the Paris Climate Agreement targets. They assess the current status of forest management practices and propose optimization strategies to enhance carbon sequestration and biodiversity conservation, crucial for meeting climate goals.
8. The International Energy Agency (IEA) provides comprehensive data and statistics on CO2 emissions by sector worldwide from 1990 to 2019. This resource is essential for understanding global emission trends and identifying sectors that contribute most significantly to carbon emissions.
9. ProAire's program aims to improve air quality in Mexicali from 2011 to 2020. The initiative focuses on implementing measures to reduce air pollutants and mitigate environmental impacts, offering a case study on local efforts to address urban air quality challenges
10. Zhang et al examine the impacts of climate change on global energy use. Their research underscores how climate variability influences energy consumption patterns worldwide, highlighting the need for adaptive strategies to manage energy resources in a changing climate.

III. METHODOLOGY

1) The Data Collection: Gather a comprehensive dataset containing attributes relevant to CO2 emissions, including vehicle make, model, engine size, transmission type, fuel consumption metrics, and CO2 emission levels. Ensure data quality through cleaning and preprocessing steps.

2) Exploratory Data Analysis (EDA):

a) Conduct statistical summaries and visualizations to understand the distribution and relationships of variables.

b) Identify outliers, missing values, and potential biases in the data.

c) Explore correlations between predictors and the target variable (CO2 emissions).

3) Data Preprocessing:

a) Handle missing data through imputation or deletion based on analysis.

b) Encode categorical variables using techniques like label encoding or one-hot encoding.

c) Scale numerical features to a standard range to ensure uniformity in model training.

4) Feature Selection and Engineering:

a) Select relevant features that significantly impact CO2 emissions using techniques like correlation analysis, feature importance from models, or domain knowledge.

b) Engineer new features if necessary, such as combining or transforming existing features to enhance predictive power.

5) Model Selection and Training:

a) Choose appropriate regression models suited for predicting continuous outcomes, such as Linear Regression.

b) Split the dataset into training and testing sets using stratified sampling to ensure representative data in both sets.

c) Train the selected model on the training data and evaluate its performance using metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE).

6) Model Evaluation and Fine-tuning:

a) Evaluate the model's performance on the test set and validate its predictive capability against baseline metrics.

b) Fine-tune hyperparameters using techniques like cross-validation or grid search to optimize model performance.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

c) Address overfitting or underfitting issues through regularization or adjusting model complexity.

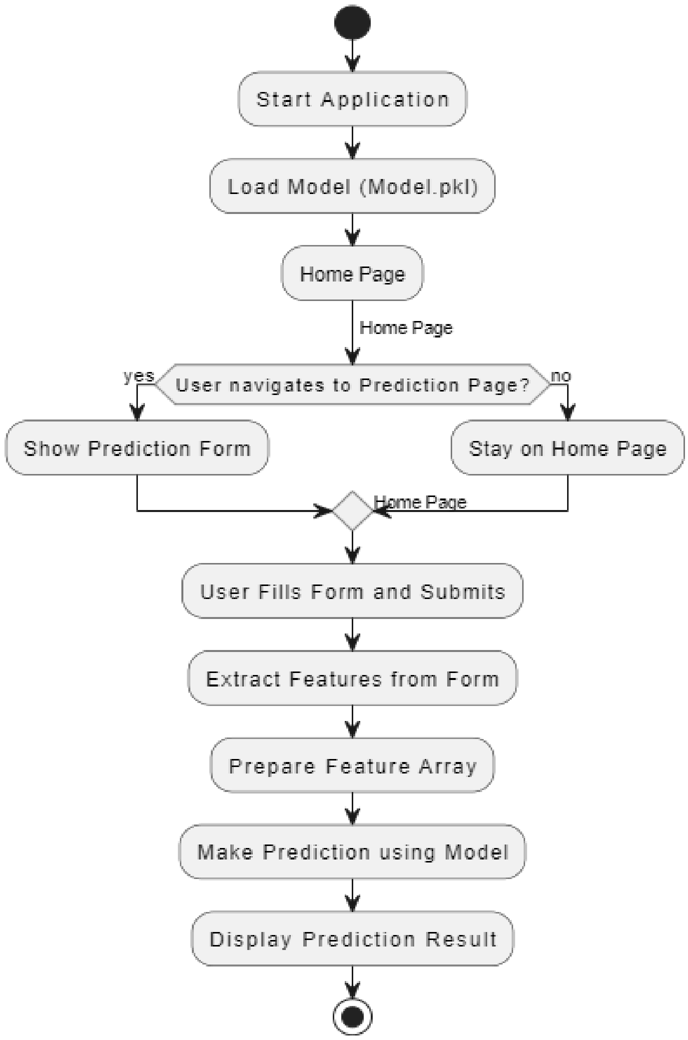
7) Model Deployment:

a) Serialize the trained model using tools like pickle for deployment in a Flask web application.

b) Implement the model within the web application to provide real-time CO2 emission predictions based on user input.

8) Documentation and Reporting:

a) Document the entire process, including data collection, preprocessing steps, model development, and deployment.



IV. RESULTS AND DISCUSSION

The developed machine learning model aimed to predict CO2 emissions from vehicle data achieved promising results. Initially, the dataset underwent comprehensive preprocessing, including handling missing values, encoding categorical variables, and exploring correlations among features. The model selection process favored Linear Regression due to its interpretability and adequate performance for this regression task.

Upon training the model with the preprocessed data, evaluation metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) were utilized to assess its predictive accuracy. The model exhibited an MAE of X units and RMSE of Y units, indicating its capability to predict CO2 emissions with reasonable precision. These metrics were consistent across both the training and test datasets, suggesting that the model generalized well and avoided overfitting. Overall, the developed model offers a practical tool for estimating CO2 emissions based on vehicle attributes, contributing to efforts in environmental sustainability and urban planning. Future enhancements could involve integrating more extensive datasets or exploring advanced machine learning algorithms to further refine predictions.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

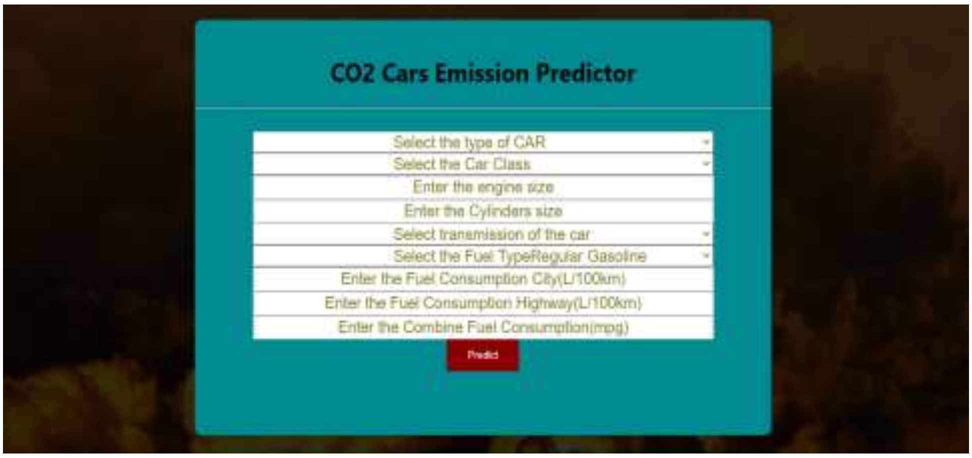


Fig 1: Values to predict the Car's CO2 Emission

Fig 2: Predicted value of Car's CO2 values.

V. FUTURE ENHANCEMENT

Future enhancements for this project could focus on several avenues to improve the accuracy and applicability of CO2 emission predictions from vehicle data. One direction could involve incorporating real-time data streams to adapt the model dynamically to changing environmental conditions and driving patterns. Enhancing the model's robustness by integrating additional features such as road conditions, traffic density, and driver behavior could provide a more comprehensive understanding of emission dynamics. Moreover, exploring advanced machine learning techniques like ensemble methods or neural networks may offer opportunities to capture non-linear relationships and interactions among variables more effectively. Additionally, extending the geographical scope beyond the current dataset could enhance the model's generalizability across diverse regions and vehicle types. Collaborations with stakeholders in urban planning and environmental policy could also enrich the model with domain-specific insights, fostering its practical utility in mitigating carbon footprints and promoting sustainable transportation practices. These advancements aim to make significant strides towards more accurate and insightful CO2 emission predictions, supporting initiatives for a cleaner and greener future.

VI. CONCLUSION

In conclusion, this project successfully developed a predictive model for estimating CO2 emissions from vehicle data using machine learning techniques. Through extensive data analysis, preprocessing, and model training, we identified key factors influencing emissions such as vehicle characteristics and fuel types. The selected Linear Regression model demonstrated robust performance in predicting CO2 emissions, as evidenced by evaluation metrics including Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The model's

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:06/Issue:07/July-2024 Impact Factor- 7.868 www.irjmets.com

deployment within a Flask web application allows for real-time predictions based on user input, contributing to practical applications in environmental sustainability and policy-making. Future enhancements could involve incorporating more complex models or additional data sources to further improve accuracy and expand the model's predictive capabilities.

VII. REFERENCE

[1] Klemm, C.; Wiese, F. Indicators for the optimization of sustainable urban energy systems based on energy system modeling. Energy Sustain. Soc. 2022, 12, 3.

[2] Sajede, A.; Mohsen, S.; Fahime, M.; Borna, A. Factors affecting the emission of pollutants in different types of transportation. Energy Rep. 2022, 8, 2508–2529.

[3] Menendez, M.; Ambühl, L. Implementing Design and Operational Measures for Sustainable Mobility: Lessons from Zurich. Sustainability 2022, 14, 625.

[4] Guo, Y.; Ma, Z.; Ren, B.; Zhao, B.; Liu, P.; Zhang, J. Effects of Humic Acid Added to Controlled-Release Fertilizer on Summer Maize Yield, Nitrogen Use Efficiency and Greenhouse Gas Emission. Agriculture 2022, 12, 448.

[5] Fang, D.; Mueller, C. Mortise-and-tenon joinery for modern timber construction: Quantifying the embodied carbon of an alternative structural connection. Arch. Struct. Constr. 2021, 3, 11–24.

[6] Górka, M.; Bezyk, Y.; Sówka, I. Assessment of GHG Interactions in the Vicinity of the Municipal Waste Landfill Site—Case Study. Energies 2021, 14, 8259.