IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

Predictive Modeling of Vehicle Emissions and Driver Behaviour: An Integrated Approach

Sachin Singh[1], Naman Gaur[2], Aditya Sangar[3], Amit Chauhan[4], Dr. Deepak Gaur[5]
Department of computer science and engineering
Amity University, Noida, Uttar Pradesh, India

***Abstract*— The study examines the predictive modelling of car emissions and driver behaviour, using Canada Fuel Consumption Ratings data set from 2020-2024. Improving CO2 emissions, CO2 rating, and smog rating predictions kept as a primary focus in this foundation by deploying complex machine learning algorithms such as Decision Tree Regressor for CO2 emissions prediction; Random Forest Classifier for Smog Rating; and Support Vector Machine for CO2 Rating. These data sets are analysed to reveal our research integrates these models to help address significant environmental challenges associated with mobile source air pollutants pollution. In particular, our approach elucidates on the contribution of vehicle emission to global climate change hence enabling policy makers and urban developers to come up with appropriate decisions regarding emission limits, improvements on roads networks and ways of mitigating climatic changes.**

Additionally, this research examines driving style in terms of emission ratings which can promote road safety measures among users while encouraging sustainable transportation practices. The blended method not only enhances understanding of the intricate link between driving habits and exhaust gases but also provides practical measures towards minimizing environmental degradation as well as enhancing air quality overall.

***Keywords*— Vehicle emissions, Predictive modelling, Machine learning algorithms, Environmental impact, Driving behaviour analysis, Sustainability.**

I. Introduction

Predictive techniques to solve crucial environmental problems and increase safety on the road in recent years have been integrated in the transport systems. In particular, forecasting of emissions and driving behaviour for vehicles is a breakthrough with far-reaching implications for sustainable development as well as public health. The reduction of vehicular emissions which are among the main causes of air pollution and climate change is important in making sure that our environment remains clean.

This study establishes that there is a research gap necessitating comprehensive examinations into long-term effects and socio-economic impacts of predictive modelling interventions concerning driving behaviour, and emissions reduction. Although previous studies have constructed predictive models and used historical data, there exists no longitudinal studies assessing the continued effectiveness of these interventions in real world settings. Moreover, little attention has been given to socio-economic factors that might influence the adoption and efficacy of predictive modelling initiatives, thereby discouraging inclusive and targeted interventions.

The goal of this study is to fill these gaps by conducting a detailed analysis on the predictive modelling of vehicle emissions and driver behaviours within an integrated framework. By applying the Canada Fuel Consumption Ratings dataset between 2020 and 2024, this research intends to discuss the sustainability implications of interventions made through predictive modelling overtime alongside their adoption and effectiveness due to social economic factors. The study will use multiple dimensions such as environmental impact, driving behaviour, and socio-economic criteria to provide whichever advice can be implemented by policy makers, urban planners, or players in the transportation industry.

In addition, this research adds to the body of knowledge on environmentally friendly transportation practices and public health improvements through providing data-driven methods for predictive modelling in the study of vehicle emissions and driving behaviours. We seek to create a more sustainable and just means of transportation for future and current generations by collaborating with each other and using an interdisciplinary approach. Furthermore, the association of Django framework with user interface development facilitates smooth communication as well as better accessibility which encourages wider use of environmental sustainability insights based on predictive modelling.

II. Related Work

It is observed that Canada's natural gas and oil industry is acknowledged as a huge contributor to GHG emissions, which resonate with global concerns about environmental sustainability (5). This indicates a commitment by the industry towards cleaner energy sources and alignment of their goals with the growing worldwide need for sustainable methods of energy generation; in fact, this might be viewed as “progress”. (5)

Smart cities can greatly benefit from artificial intelligence (AI) through new approaches that provide smart solutions for sustainable urban ecosystems. Such proposals demonstrate the possible adoption of AI-informed prediction models aimed at addressing global warming through informed decisions made on transportation planning and logistics management regarding CO2 production levels from trucks in future time based on historical data. (3)

Around the world, transportation systems, specifically vehicles, have been instrumental in fostering economic growth and development. However, carbon dioxide (CO2) emissions have increased exponentially leading to major environmental issues that call for effective monitoring strategies as well as mitigation measures (1). In addition, with machine learning (ML) techniques it is now possible to track CO_2 emissions in real-time which can help take preventive measures to comply with regulatory standards on air quality. Consequently, there has been research on the use of different ML models for this purpose while focusing on achieving high predictive accuracy measured by R^2 metric (1).

Besides, sharing economy provides opportunities for reducing carbon footprints related to transport (2). These studies used hybrid unsupervised learning to classify vehicles according to emission levels and enable focused interventions for reducing CO_2 emissions (2).

For the examination of emissions, remote sensing data integration with complex techniques such as artificial neural networks (ANNs) have emerged as a highly promising means for predicting automotive exhaust fumes (4). By evaluating key factors and utilizing optimized ANN structures, reliable predictions based on these approaches have been made by researchers to exhibit the potential of this approach in emission modelling (4).

Broadly speaking, literature has suggested that advanced technologies, regulatory frameworks, and industry-led initiatives are among the multilateral approaches towards reducing environmental impacts and achieving sustainable development from transport and industrial sectors.

III. System Implementation and Methods

Here, we discuss about how machine learning models were implemented to predict CO_2 emission, CO_2 rating, smog rating and driving behaviour using the dataset explained above. We also explain our data preparation, pre-processing techniques, model selection strategies used in this research, calculation of driving behaviour and interactive system description.

A. Dataset

For this study a dataset was utilized consisting of fuel consumption ratings information from 2020 to 2024. Fuel consumption ratings are detailed as follows: Model Year; Make; Model; Vehicle Class; Engine Size, Cylinders, Transmission, Fuel Type, City Fuel Consumption, Highway Fuel Consumption, Comb Fuel Consumption, CO_2 Emissions, CO_2 Rating, Smog Rating. This compilation is credible because it comes from several sources with different types of vehicles including the whole spectrum of models.

B. Preprocessing

For training purposes extensive preprocessing steps were performed on datasets. First these datasets for each year had to be merged into one data frame enabling comprehensive analysis as well as model-training. In some datasets there was an irrelevant column “Unnamed:15” which was dropped.

```
Index(['Model year', 'Make', 'Model', 'Vehicle class', 'Engine size (L)',
       'Cylinders', 'Transmission', 'Fuel type', 'City (L/100 km)',
       'Highway (L/100 km)', 'Combined (L/100 km)', 'Combined (mpg)',
       'CO2 emissions (g/km)', 'CO2 rating', 'Smog rating', 'Unnamed: 15'],
      dtype='object')
```

Additionally, some processing steps are involved in the data analysis:

1) Data Cleaning: Any missing values or inconsistent data were well handled to ensure that the dataset was of quality.

2) Exploratory Data Analysis (EDA): The EDA method has been used to acquire meaningful insight about its distribution and nature. Histograms and boxplots showed how fuel consumption, engine size and emission rating attributes were spread across the classes.

3) Feature Extraction: In addition, essential features such as numerical attributes like engine size, fuel consumption and categorical attributes such as vehicle class and transmission type have been identified for model training.

Moreover, feature scaling done by using StandardScaler from scikit-learn library is applied to normalize input numerical data while OrdinalEncoder was used transforming categorical data converts categorical attributes into numeric format for easier model training.

C. Model Selection

To predict CO_2 emissions, CO_2 rating and smog rating, it was important to carefully select machine learning models. Below are the chosen models:

1) CO_2 Emission Prediction: A Decision Tree Regressor model was used to predict CO_2 emissions. As tree-based techniques, decision trees work best for regression since they can deal with both categorical and numerical data types.

2) CO_2 Rating Prediction: Support Vector Classifier (SVC) model was taken for the purpose of predicting CO_2 ratings. SVCs are appropriate when there is a complex or non-linear separation between classes allowing it to be also used in multi-class classification problems.

3) Smog Rating Prediction: For the prediction of smog ratings, we employed a Random Forest Classifier model. Being ensemble learning methods made up of multiple decision trees, random forests are strong classifiers which can work in classification tasks well too.

D. Driving Behaviour Calculation

The CO_2 rating predicted by the model is used in order to appraise driving behaviour. The CO_2 rating is a measure of how much pollution a vehicle causes, and lower ratings mean more eco-friendly driving styles. There are three classes of driving behaviour based on the forecasted CO_2 rating:

1) Good: Vehicles with minimal environmental impact as indicated by their CO_2 ratings.

2) Average: Vehicles that have moderate CO_2 ratings meaning that they have medium environmental impacts.

3) Bad: Vehicles with high CO_2 emissions implying significantly large environmental impacts.

This enables one to discern some patterns in driving habits and pinpoint areas where changes could be made to reduce emissions and promote sustainable practices of driving, ones that are environmentally friendly.

With these models, Python programming language was employed using popular libraries such as pandas, scikit-learn, and matplotlib. The training script follows typical industry practices with respect to model building while ensuring replicability.

It is hoped that through the use of the dataset, preprocessing techniques, model selection strategies outlined above, combined with the calculation for driving behavior this study aims at developing accurate predictive models for assessing vehicle emissions and driving behavior within towards Environmental Sustainability efforts and transportation movement.

E. Interactive System Description

In this section the application prototype will be discussed with the appropriate figure for the prototype.

Django framework and Bootstrap allows integration of predictive models into interactive User Interfaces (UI). A secure login screen is the first step to access authenticated users. Once successfully logged in, the user is directed to the input dashboard where they can enter vehicle specifications and receive instantaneous predictions on CO_2 emission, CO_2 rating, smog rating, driving behaviour and personalized recommendations. Django's backend functionality powers UI that enables you to upload user data for processing by predictive models which generate forecasts and analysis. Moreover, frontend design is improved by bootstrap components making it responsive to meet user we need with regard to visualization front-end designs.

User input data is accepted by the predictive models, which provide their results on the input dashboard. User can see predicted values and have their driving behaviour assessed as well as get some suggestions. Additionally, there is a model analysis results dashboard that shows how different predictive methods perform. Its integration with Django and Bootstrap facilitates its use and understanding, thereby enabling users to make smart decisions about the carbon emissions produced by their cars or about driving habits. This research paper aim at creating environmental awareness and fostering an environmentally sound transport system through this approach.

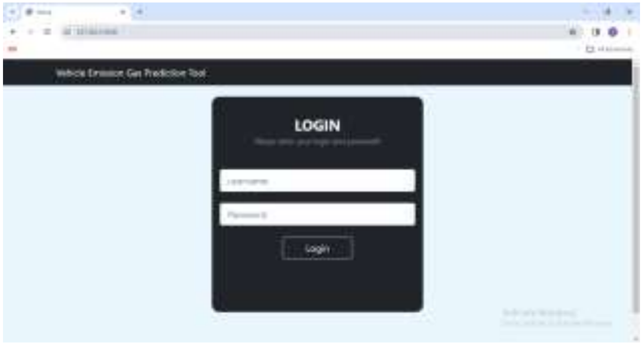

Fig. 1 Login Screen

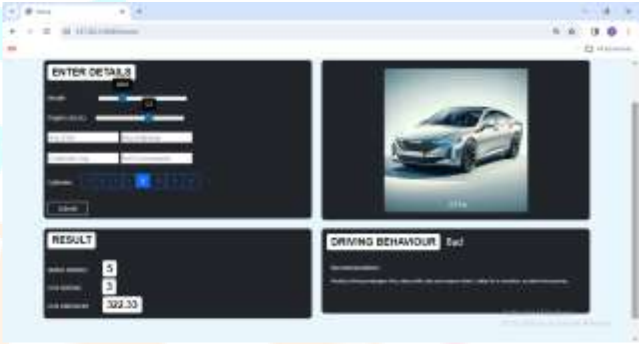

Fig. 2 Input Dashboard

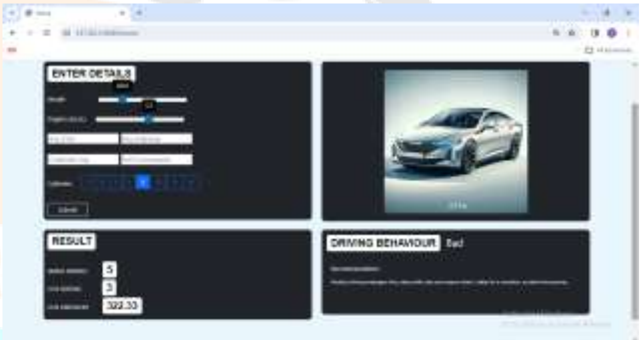

Fig. 2 Model Analysis Dashboard

F. Result and Discussion

A promising outcome was observed in combining UI framework with predictive models, indicating how effective it is in terms of real-time predictions and vehicle emission and driving behaviour analysis. The key findings and discussions from the application of the predictive models and user interface interaction are presented in the next sections.

1) Predictive Model Performance:

Accuracy scores obtained during model training were used to evaluate the performance of predictive models. It was found out that Decision Tree Regressor showed a good performance while predicting CO_2 emissions as Random

Forest Classifier achieved a high accuracy level for CO2 rating and smog rating.

The model analysis results dashboard provided users with detailed insights into how well the predictive models performed. This software allowed users to evaluate the reliability of their projections, enhancing trust on generated outputs.

2) User Interface interaction:

The user interface made users interact with the system easily, where they could enter vehicle specifications to obtain instant predictions for CO2 emissions, CO2 rating, smog rating, driving behavior and recommendations.

By having an intuitive design of input dashboards, users were able to be guided through a step-by-step process while at the same time allowing predicted values and analysis of driving behaviors to be visualized. Consequently, this approach enhanced ease of use and accessibility among diverse groups of people that might be involved in the product.

Authenticated or recognized users alone were allowed access into the login screen as per its function in maintaining privacy and security related to such confidential information.

3) Implications and Future Directions:

The successful incorporation of the predictive models within UI framework will have significant implications on promoting environmental consciousness as well as sustainable transport practices.

This can further involve refining predictive models further; integrating additional features for driving behavior analysis so that it will be more comprehensive; increasing or expanding the scope of UI addressing other wider environmental matters.

Moreover, usability and engagement may also be improved through user feedback as well as iterative enhancements on the UI interface thereby fostering a community-based approach towards environmental stewardship.

IV. Conclusions

In conclusion, a good example of this study is how predictive modeling techniques can be integrated into a user-friendly interface in order to allow real-time analysis of vehicle emissions and driving behavior. Any future eventuality can be predicted through the design of a simple interface, which shall also involve sturdy models for predicting future events hence allowing users to choose and adopt the right travel practices that are supportive of environmental sustainability. The findings underlined the potential offered by this approach towards reducing environmental impacts as well as promoting more environmentally friendly alternatives for transportation. Maintaining progressiveness requires enhancing predictive modeling methods, expanding user interface features, and addressing larger ecological matters. Technology will help us create a healthier environment and ensure our planet remains sustainable for generations to come. Continuous innovation and cooperation with other stakeholders are necessary prerequisites for improving on prediction models; augmenting the size of the user-interface; dealing with broader ecological issues so that we use technology as an instrument for generating healthier environments hence sustainable existence for tomorrow's children.

References

[1] G. Meccariello, L. Della Ragione, M. A. Costagliola and M. V. Prati, "Evaluation of emission factors for new technology S.I. Euro 4 cars," Proceedings of the ITI 2012 34th International Conference on Information Technology Interfaces, Cavtat, Croatia, 2012, pp. 441-448, doi: 10.2498/iti.2012.0407.

[2] M. Ghahramani and F. Pilla, "Analysis of Carbon Dioxide Emissions from Road Transport Using Taxi Trips," in IEEE Access, vol. 9, pp. 98573-98580, 2021, doi: 10.1109/ACCESS.2021.3096279.

[3] S. Yeasmin, S. N. J. Syed, L. A. Shmais and R. A. Dubayyan, "Artificial Intelligence-based CO2 Emission Predictive Analysis System," 2020 International Conference on Artificial Intelligence & Modern Assistive Technology (ICAIMAT), Riyadh, Saudi Arabia, 2020, pp. 1-6, doi: 10.1109/ICAIMAT51101.2020.9307995.

[4] Z. Xu, Y. Kang and W. Lv, "Analysis and prediction of vehicle exhaust emission using ANN," 2017 36th Chinese Control Conference (CCC), Dalian, China, 2017, pp. 4029-4033, doi: 10.23919/ChiCC.2017.8027988.

[5] Canadian Association of Petroleum Producers (CAPP). (Year, if available). Greenhouse Gas Emissions. CAPP. https://www.capp.ca/explore/greenhouse-gas-emissions/. [Accessed: 14-Mar-2024].