2024 International Conference on Intelligent and Innovative Computing Applications (ICONIC)
7th-9th November 2024, Pearle Beach Resort & Spa, Flic en Flac, Mauritius
https://doi.org/10.59200/ICONIC.2024.024

Analysing the Impact of Engine Size on CO_2 Emissions in Five-Seater Vehicles Using Machine Learning Techniques

Antonio Stefano Cristaudi
School of Information Technology
Varsity College, Independent Institute of Education
Johannesburg, South Africa
ST10090982@vcconnect.edu.za

Mfowabo Maphosa
School of Engineering, EBIT
University of Pretoria
Pretoria, South Africa
mfowabo.maphosa@up.ac.za

Abstract
The release of carbon emissions into the atmosphere has a negative impact on the environment. Carbon emissions are considered a significant factor in global warming. These harmful emissions trap heat and lead to an increase in the earth's average temperatures, which in turn causes melting polar ice caps, rising sea levels, and severe weather patterns. Five-seater vehicles represent a significant portion of the global automotive fleet, making their emissions impactful in contributing to these environmental changes. This study investigates the relationship between engine size and CO2 emissions in five-seater vehicles using machine learning algorithms. Linear regression, ElasticNet, and neural networks were applied to a dataset of 38, 500 observations obtained from Kaggle. The study found a significant correlation between engine sizes and CO2 emissions. Factors such as engine displacement and fuel type were analysed and identified as contributors to CO2 emissions. The results indicate that advanced engine technologies such as turbochargers and hybrid systems can mitigate emissions by improving efficiency although the impact varies. These findings highlight the importance of engine downsizing and technological integration to reduce the automotive sector's carbon footprint. This study offers insights that can help guide environmentally-friendly strategies in the automotive industry.

Keywords: CO2 emissions, automotive industry, climate change, machine learning, engine size

1. Introduction

In 2015, the United Nations adopted the 2030 Agenda for Sustainable Development, which outlines 17 goals and 169 targets structured around five pillars (people, planet, prosperity, peace, and partnership) to promote sustainable, resilient development (United Nations, 2015). There is an urgent global need to mitigate climate change and reduce greenhouse gas emissions from the transportation sector. The automotive industry is a major source of CO2 emissions globally, prompting the USA and the EU to establish stringent emission regulations aimed at mitigating environmental impacts (Blumberg & Posada, 2015; O'Driscoll et al., 2018).

The relationship between CO2 emissions and climate change has been a topic of interest among researchers for a long time. The foundation was laid by Gilbert Plass's seminal work which emphasised the need for environmentally sustainable strategies to combat the challenges of greenhouse gas emissions on the environment (Plass, 1956). This understanding was further contextualised by the European Union's (EU) Green Deal and the USA's updated regulatory frameworks on vehicle emissions (European Commission, 2020; United States Government, 2024). The EU aims to become the world's first climate-neutral continent by 2050, underscoring the urgency of reducing CO2 emissions through stringent standards like the Euro 6 regulations (European Commission, 2024). These regulations have prompted a global shift in automotive innovation, including the adoption of smaller displacement engines with forced induction, hybrid technologies, or electric vehicles to meet both performance and environmental goals (Vedran & Božica, 2018).

Engine downsizing, often coupled with forced induction technologies such as turbocharging, has become a key strategy for improving fuel efficiency and reducing emissions while maintaining performance (International Council on Clean Transportation, 2019). Despite these advancements, it remains challenging to accurately assess the impact of these technologies on CO2 emissions due to the variability in engine designs, vehicle types and driving conditions. Machine learning offers new opportunities to examine these relationships and estimate the influence of engine size on CO2 emissions with greater precision (Zhang et al., 2017).

Issues of climate change must be taken seriously as they are critical to humanity's survival (Younus et al., 2023). Change must be prioritised for the better of humanity. This all starts with the challenge of tackling air pollution. Air pollution such as carbon emission harms the environment, and drives climate change, whilst posing serious health risks (Younus et al., 2023). Addressing these concerns through greener transportation solutions is vital for transitioning to a more globally sustainable development model. Currently, Asia is the biggest CO2 emitter globally contributing just over 68% of emissions, followed by the US with 25%. The EU ranks 3rd with 18%. Africa and South America only account for 4% of global emissions each (Macknick, 2011; Roser & Ritche, 2020).

The EU and the United States have introduced stringent emission regulations which are government-imposed standards aimed at reducing the number of harmful pollutants; in this case, CO2 and nitrogen oxides that are emitted by vehicles to protect the environment and public health. (European Commission, 2020, 2024; United States Government, 2024) Despite these efforts, there remains a gap in the literature regarding the impact of engine size on CO2 emissions, particularly for five-seater vehicles (Galvin & Healy, 2020). While prior research has explored engine downsizing and forced induction technologies, there is limited use of machine learning techniques to analyse these relationships in detail (Vedran & Božica, 2018).

Machine learning provides a unique approach to predictive analysis (Zhang et al., 2017). This study uses machine learning models to predict the impact of engine size on CO2 emissions and provide insights that could inform more effective legislation and engine design, contributing to the reduction of the automotive sector's carbon footprint.

Understanding the relationship between engine size and CO2 emissions is crucial for developing effective strategies for emission reductions and improving fuel efficiency. The study sought to answer the following research questions:

- What is the relationship between engine size and CO2 emissions levels in five-seater vehicles, considering factors such as engine displacement, number of cylinders, and fuel type?
- How do different engine technologies, such as turbochargers, superchargers, or hybrid systems, impact the relationship between engine size and CO2 emissions?

The subsequent sections are organised as follows: the following section outlines a literature review. Next, the methodology is outlined, explaining the approach taken in the study. The results are then presented, followed by a discussion that interprets the findings. Finally, the conclusion summarises the key points and implications of the study.

2. Literature Review

The impacts of global warming and climate change have become more pronounced in recent decades. Rising global temperatures and extreme weather patterns are posing a significant threat to global ecosystems and human societies (Kahn Ribeiro et al., 2007). The transportation sector makes up a significant portion of global greenhouse gas emissions and has therefore become a major focus for policymakers worldwide. Recent studies emphasize the urgent need for emission reductions to meet the goals of the Paris Agreement. This is due to the transportation sector accounting for nearly 25% of global CO2 emissions (NRDC, 2017).

The EU's ambitious Green Deal (European Commission, 2020) has emerged as a hallmark initiative, aiming to make Europe the world's first climate-neutral continent by 2050. This was initiated after the Paris Agreement was signed by countries responsible for 97% of global emissions on the 12th of December 2015 (NRDC, 2017). The vision of this agreement is to have all developed and developing countries, make significant commitments to address climate change. This vision is being enforced drastically by stringent emission standards, such as the Euro 6 regulations and The US government's emission standards (European Commission, 2024; NRDC, 2017; United States Government, 2024). This imposes strict limits on pollutants such as Carbon Oxides (COx) and Nitrogen Oxides (NOx) emitted by gasoline, petroleum, and diesel-powered vehicles.

The new EU and US transportation and emission regulations are responsible for the automotive industry's shift towards smaller engines. Forcing the automotive industry to equip itself with a renewed take on forced induction technologies and other alternate technologies like electrical motors. The innovation push in this region has seen a massive intensification in harnessing the power of forced induction. As a result, companies within the automotive industry can now downsize engines whilst simultaneously maintaining and enhancing the overall engine performance. It remains an imperative goal that automotive manufacturers and car companies meet the rigorous Euro 6 standards while satisfying consumer demands for both power and fuel efficiency (Vedran & Božica, 2018).

The regulatory pressure and technological innovation underscore the pivotal role that engine size and forced induction play within the transition towards cleaner, greener transportation solutions in the EU (European Commission, 2020, 2024; United States Government, 2024). These regulations serve as an important driving force within the automotive landscape as the global community witnessed a paradigm shift towards more sustainable solutions. Thus, setting the stage for a future where engine efficiency and environmental consciousness go hand in hand (Galvin & Healy, 2020; United States Government, 2024).

Forced induction technologies such as turbocharging, or supercharging are powerhouse technologies in petroleum and diesel engines. The enhancements can drastically improve overall engine performance and fuel efficiency. This can be done by tapping into exhaust gas energy, and turbochargers, boosting engine torque and power across various speeds. They enable downsizing engines without sacrificing power, meeting stringent regulations effectively. Moreover, forced induction engines offer design flexibility, enabling smaller, lighter engines with comparable or better performance. In essence, driving innovation within the automotive industry to deliver more powerful, fuel-efficient, and environmentally friendly vehicles (Vedran & Božica, 2018).

Technological Advancements in Engine Efficiency

Technological advancements play a critical role in meeting stringent regulatory standards and consumers' environmental expectations. Forced induction technologies, particularly turbocharging and superchargers, have become integral to modern engine design due to their ability to enhance power output and fuel efficiency simultaneously whilst remaining environmentally friendly (O'Driscoll et al., 2018). Turbocharging works by utilising exhaust gases to drive a turbine, which then compresses the intake air, allowing more fuel to be burned more efficiently (Vedran & Božica, 2018).

Recent studies have highlighted the benefits of turbocharging and improving engine fuel efficiency and overall engine performance. Vedran & Božica (2018) found that turbocharging can increase engine torque by over 62%. This improvement is particularly significant in the context of stringent emissions regulations. This is based on the notion that stringent emissions regulations have forced manufacturers to produce engines that meet regulatory standards without sacrificing performance (Automobile Association Developments, 2017; O'Driscoll et al., 2018; Vedran & Božica, 2018).

Lighter and more efficient engines have been made possible by developments in materials science and engine design in addition to turbocharging (González et al., 2012). For instance, engine weight has decreased overall because of the introduction of modern steel and aluminium alloys in engine components, improving performance and fuel economy (Costiuc & Anghel, 2017).

Consumer Preferences and Market Trends

The way that engine technology has evolved has also been greatly affected by consumer tastes. A rising number of people are looking for cars that combine low emissions, excellent performance, and fuel efficiency (Vedran & Božica, 2018). Studies have indicated that because of increasing fuel prices and worries about the environment, buyers are prepared to pay more for cars with higher fuel efficiency and fewer pollutants (Eckardt, 2022). The current shift in consumer preferences has once again propelled the automotive market into a completely different direction as automakers have been prompted to prioritise the development of fuel-efficient, high-performance engines that comply with emissions regulations.

Furthermore, the popularity of hybrid and electric cars has expanded the range of options available to consumers. Reducing emissions without sacrificing performance is possible with the combination of conventional internal combustion engines and electric drivetrains. According to Toyota Motor Corporation (Toyota Motor Corporation, 2017), hybrid electric vehicles are now more than ever viewed as a bridge of technologies that combine the advantages of electric power with the range and practicality of conventional engines.

Critical Review of Current Knowledge

While the benefits of turbocharging and other forced induction technologies are well-documented, there are still gaps in understanding their long-term impacts on overall emissions performance and engine durability (Vedran & Božica, 2018). Recent literature agrees on the efficiency gains and emission reductions associated with these technologies, but there is less consensus on their potential trade-offs. One of the issues to be considered is durability and maintenance costs for turbocharged and supercharged engines. Compared to normally aspirated engines, turbochargers and superchargers are subjected to higher temperatures and pressures during operation which may result in faster wear and tear (Vedran & Božica, 2018). The cost of ownership and the lifetime reliability of vehicles with turbochargers and superchargers are affected by this (Eckardt, 2022; Pratt, 2023).

Nonetheless, several restrictions and difficulties in this field are highlighted by the volume of research that is now available. A significant obstacle is needed for extensive datasets that encompass a broad spectrum of factors to produce accurate prediction models. Weather, driving conditions, maintenance procedures, and other factors might also significantly impair accuracy. This is one of the main reasons that a robust set of procedures is needed to guarantee model validation and the accuracy of prediction models (Vedran & Božica, 2018).

The potential environmental impact of hybrid and electric vehicles is the final and possible critical area for further research. The overall environmental benefits of these technologies are dependent on factors such as the source of electricity and the lifecycle emissions of battery production and disposal, although they offer significant reductions in tailpipe emissions (Hawkins et al., 2012; Toyota Motor Corporation, 2017).

Emerging Trends and Future Directions

The future of automotive engine technology lies in the continued evolution of hybrid and electric powertrains (Hawkins et al., 2012; Toyota Motor Corporation, 2017). The shift towards electrification is driven by both regulatory pressures and advancements in battery technology, which is forcing hybrid vehicles to be more viable for mainstream adoption (Berkeley et al., 2017; European Commission, 2020, 2024).

Elaborating on the development of how new combustion technologies as well as electrification, can further improve the efficiency and emissions performance of internal combustion engines available (Duan et al., 2021; Paykani et al., 2015). These technologies aim to combine the best features of gasoline and diesel engines, achieving high efficiency with low emissions (Vedran & Božica, 2018).

However, major mountains remain to be climbed because of the challenges in areas regarding the long-term durability and environmental impact of these technologies. Whilst it is clear what can be achieved in the short term, the question remains in the long term as consumers may forget to service their vehicles or fill up with a lower quality fuel than what the vehicle requires. A response to this, even though it is very limited in its current capacity, is the introduction of hybrid and electric vehicles.

Hybrid and electric propulsion systems are the latest form of technological advancements, often used in conjunction with conventional internal combustion engines. These systems have proven to be extremely technologically advanced and have become a viable strategy for cutting emissions and increasing efficiency. Although this new technology is very expensive, the technology is still extremely new and further research should concentrate on improving the existing technologies, investigating possible materials, and reviewing the manufacturing processes. This is so facilitation of the continuous development of sustainable vehicle solutions can be maintained. This study aims to highlight the significance of continuous green technological innovation and improvement within the automotive industry. The omission of European and American emission laws raises a question about their potential role as contributors to global warming.

3. Methodology

This study makes use of a combination of two out of the four primary paradigms. Positivism quantifies data for reliable estimates. On the other end, post-positivism acknowledges subjectivity (Kamal, 2019). For this study, the positivist paradigm was deemed the most appropriate because it aligns with the principles of objectivity, quantifiability, and generalisability.

The positivist paradigm emphasizes using empirical evidence and statistical methods to determine the relationships that exist within the data. By following the positivist paradigm, the emphasis on predictions based on objective data becomes vital. Vehicle data, such as engine size, fuel type, CO2 emissions, and the presence of turbochargers or superchargers, can be quantitatively analysed to obtain reliable estimates and insights. By employing a positivist paradigm and quantitative approach, this research aims to provide valuable insights into trends in vehicle technology, engine size and emissions. The focus on generalisability ensures that the study's findings can be applied to a broader context beyond the specific dataset used. The process involved in the study used data preprocessing, training and testing the models. This process is explained in more detail in the subsections below.

Dataset

The study used data from a dataset sourced from Kaggle, provided by the Environmental Protection Agency (US Environmental Protection Agency, 2017). It comprises records of five-seater vehicles with four wheels and includes detailed information on engine size, CO2 emissions, fuel economy, and other relevant variables. The dataset spans across various vehicles throughout the years, ranging from 1984 to 2017. The dataset offers a very detailed view of emission trends over the years. Fuel economy estimates provided in the dataset are derived from laboratory testing under standardised conditions, ensuring fair comparisons between vehicles.

To ensure accuracy and reliability, the dataset underwent pre-processing that involved cleaning and verification processes. Handling missing values, removing outliers, checking for inconsistencies and removing them. This approach aims to produce insights that can help identify strategies, common practices as well as alternative practices for reducing CO2 emissions and improving vehicle efficiency.

Data Analysis

The study was completed with the Jupyter Notebook Software using Python as a language in a Python 3 environment. The study was done using a Dell G15 computer. Powered by an AMD Ryzen 7 5800H CPU with Radeon Graphics 3.20 GHz and 32GB RAM of 31,9 GB usable. The following libraries were used: Pandas, NumPy, Seaborn, Matplotlib, and Skit Learn. Libraries help provide various functions and methods to handle data.

The data analysis for this study employed machine learning algorithms to determine the relationship between engine size and CO2 emissions in five-seater vehicles. In this process, the initial step removed all forms of irrelevant columns that did not add value to the data analysis and duplicate entries. This has enabled the next process to begin where all unwanted outliers have been filtered out and missing data handled by either standardising data, dropping rows or simply using a one-hot encoder. The final step ran through all the variables and entries within the cleaned dataset and questioned if the data made sense so that key variables could be used to identify trends and patterns. Key variables such as engine size, CO2 emissions, fuel economy, and forced induction technologies were extracted and prepared for analysis.

Various machine-learning techniques were employed to model the relationship between engine size and CO2 emissions. Special focus was initially put on linear regression but then progressed to the use of ElasticNet penalties, and neural networks to help identify significant predictors. Furthermore, the use of special metrics like Mean Residual Sum of Squares (MRSe) and R-squared (r^2) have been used to evaluate the model's performance. This provided insights into the variance explained by the predictors (Mahesh, 2018).

Linear Regression was initially chosen for its simplicity and interpretability. This traditional model offered a clear view of the linear relationship between predictors and emissions but could not capture the deep layering relationships that a neural network can. Neural networks were incorporated to capture more complex and non-linear relationships. Neural networks helped assess interactions between features. This modern neural network technique even allowed the modelling of intricate patterns that the traditional linear regression model might have missed. This approach aims to enhance the analysis by potentially revealing deeper insights into how engine size impacts CO2 emissions (Mahesh, 2018; Nti et al., 2021).

Cross-validation techniques have also been utilised to ensure the robustness and reliability of all models, including the neural networks. ElasticNet was then used to combine the benefits of both Lasso and Ridge regression. The purpose was to find a balance between the trade-off feature selection and coefficient shrinkage. This approach combining Cross-validation and ElasticNet penalty is particularly useful in handling datasets with multicollinearity or a high number of predictors. This comprehensive approach has led to a better understanding of the relationships between engine characteristics and CO2 emissions.

4. Experiments and Results

In the pursuit of the objectives outlined in the study, a rigorous evaluation and plotting of data points were undertaken. The dataset is comprised of historical records and includes vehicles, fuel type, CO2 emissions, engine displacement, and other relevant metrics. The analysis and visualisations are summarized as follows: Figure 1 illustrates the trend of engine displacement over time. The y-axis represents engine displacement in litres, while the x-axis displays the years. The graph shows that from the year 1984 to the year 1995, there was a noticeable upward trend in engine displacement. After the year 1995, the trend reversed its course with a steady decline in engine sizes until around the year 2000.

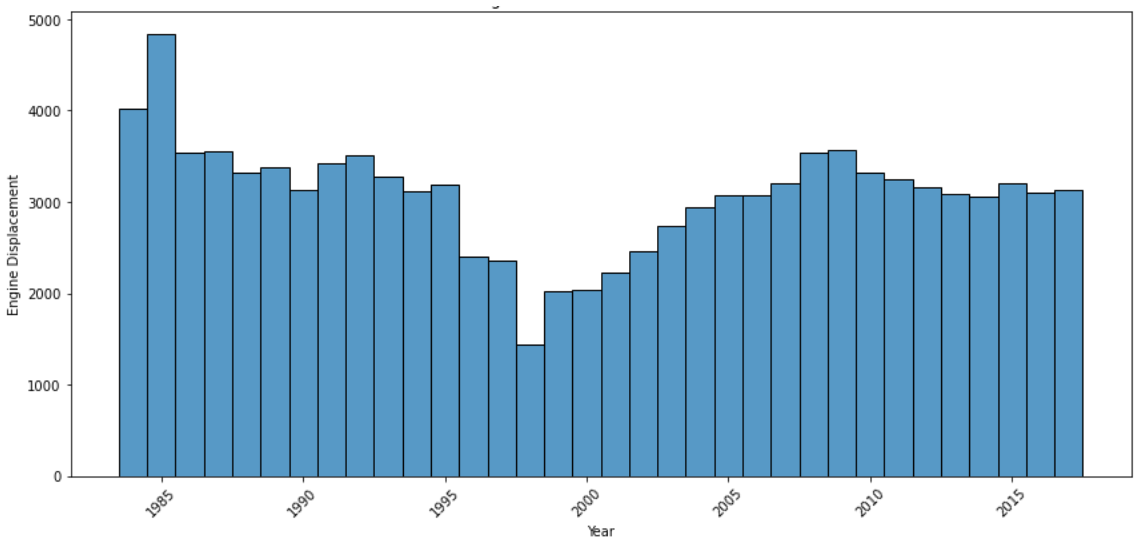

Figure 1. Engine size data over the years

A resurgence in engine size then occurred from 2007 to 2008. Post-2008, a significant downward trend was observed, with engine sizes decreasing steadily. The two metrics measured in Figure 2 (Supercharger = Green, Turbocharger = Blue) are then shown again in Figure 3 as ‘Turbocharger or Supercharger’ (Blue line). In Figure 2 it can be seen there is no exponential growth for superchargers but rather a steady increase in growth. Superchargers account for just under 5% of all forced induction vehicles on the road to date according to the dataset. Turbochargers in Figure 2 are the line in blue which, from the year 1985 have always occupied at least 5% of all the forced induction vehicles sold. Figure 3 shows that there has been a steady decrease in turbocharged cars from 1987 to 1996. Since the year 1997, turbocharged cars have grown exponentially occupying just under 50% of all forced induction vehicles on the road in 2017.

The graphs in Figure 2 and Figure 3 depict the prevalence of turbochargers and superchargers in vehicles from 1985 to the year 2017. The two metrics measured in Figure 2 are represented again in Figure 3 as 'Turbocharger or Supercharger' (Blue line) and the naturally aspirated vehicles (Red line). The y-axis ranges from 0 to 1, indicating the percentage of vehicles equipped with these technologies. Forced induction vehicles have shown a steady increase that transformed into an exponential increase from the year 1997 to 2017. Reaching over 50% of the market in 2017, after the exponential boom from 1997 depicted in Figure 3, even with superchargers that have remained below 10% in Figure 2, despite an initial presence.

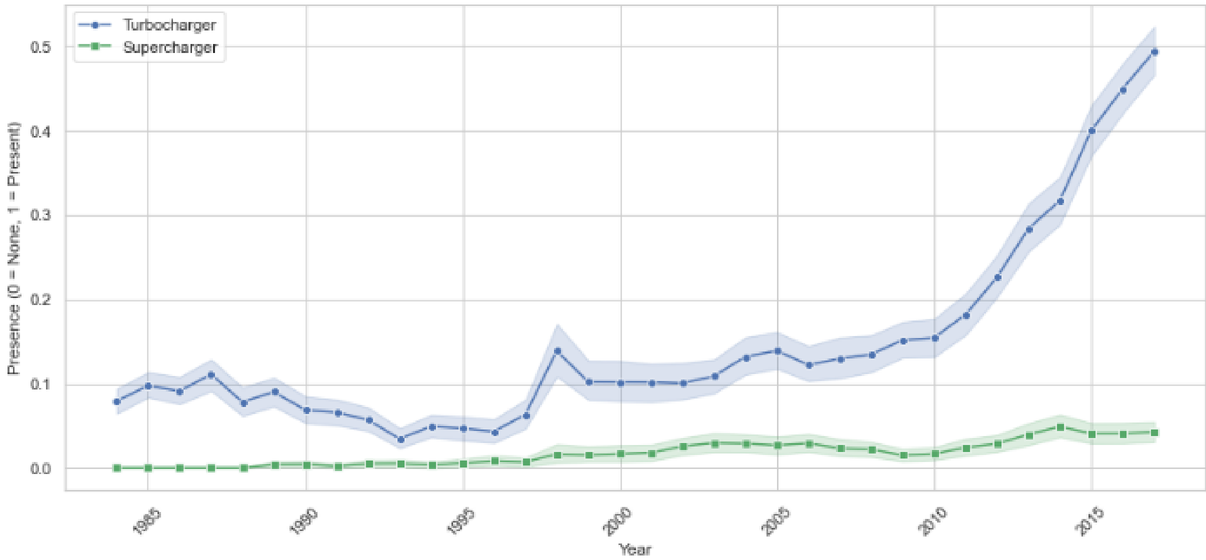

Figure 2. The presence of turbochargers and superchargers in five-seater vehicles

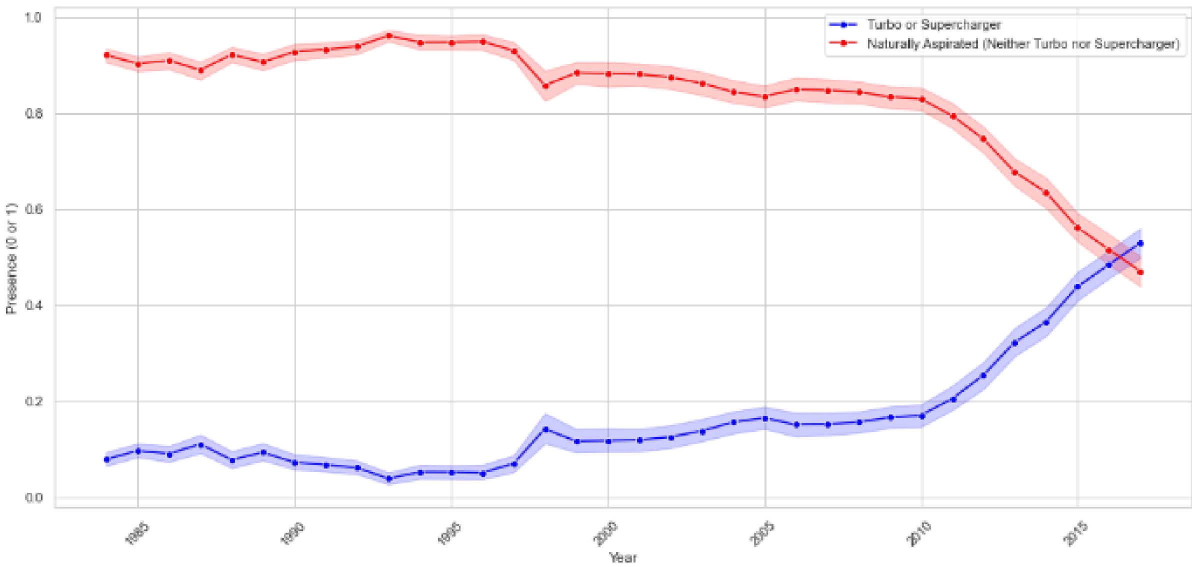

Figure 3. The presence of turbochargers or superchargers and naturally aspirated engines in five-seater vehicles

Figure 4 shows the distribution of different fuel types used in vehicles over time. Regular petroleum (Red) remains the most common and constant fuel type. However, there has been a growing shift towards premium petroleum (Green), E85 gas (Orange), and, to a lesser extent, diesel (Blue). Diesel was once more prevalent but has since declined as other fuel types have gained popularity.

The aim was to predict tailpipe CO2 emissions based on a very carefully curated set of vehicle features. The dataset spanned across the years 1984 to 2017, and the primary objective was to develop an accurate model capable of estimating CO2 emissions. By selecting the most relevant features, such as ‘Engine Cylinders', 'Engine Displacement', 'Turbocharger', 'Supercharger', 'Fuel Type', 'Start Stop Technology', 'Class', and 'Transmission'. These features were carefully chosen to capture essential aspects of vehicle performance and design.

A pre-processing pipeline was then established above and beyond the cleaned data. The purpose of this was to standardise numerical features such as engine specifications and start-stop technology using StandardScaler. Categorical features in this research such as fuel type and vehicle class were encoded using OneHotEncoder. This pre-processing step ensured that the model could effectively handle different feature types. The heart of the study lay in the ElasticNet regression model. By combining Lasso and Ridge regularisation techniques, the model struck a balance between feature selection and feature grouping. The neural network model was then fine-tuned and the model’s hyperparameters were then put through a grid search with cross-validation.

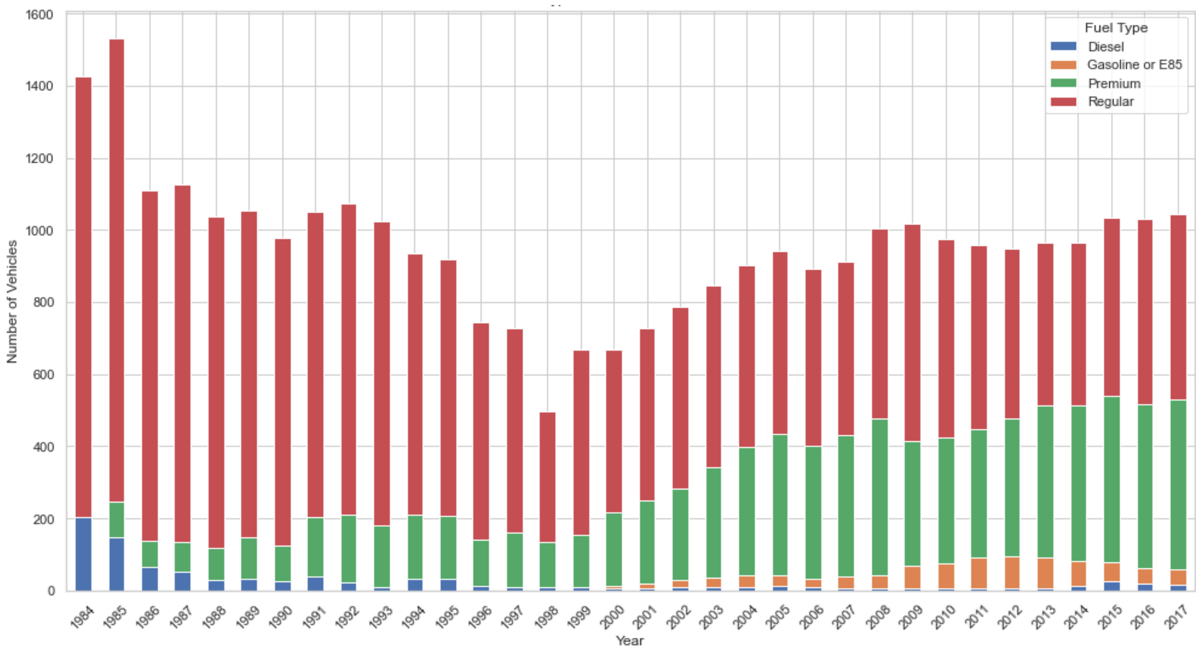

Figure 4. Fuel type distribution over the years

Once the best model had been identified, model training then commenced using the training data as an 80/20 split was determined. Predictions were then made on the testing data and the model’s performance was evaluated using Mean Squared Error (MSE) and r2. These metrics provided insights into the model’s accuracy and ability to explain variance in CO2 emissions. The purpose of the cross-validation was to assess the model’s robustness across different folds of data. The cross-validated MSE and standard deviation served as valuable indicators of overall performance. Table 1 presents the performance metrics.

Table 1. Performance results

Metric	**Value**
Mean Squared Error	2017.77
R-squared	0.83
ElasticNet - Cross-validated MS	2169.28
ElasticNet - Cross-validated Standard Deviation	763.58

The MSE is what is used to measure the average squared difference between the actual and predicted values. A lower MSE indicates a better fit of the model to the data. The MSE is often used to assess the accuracy of predictions (Nti et al., 2021).

R2 is known as the coefficient of determination. This is a representation of the model's proportion of the variance in the dependent variable. In this case, the tailpipe CO2 emissions are made predictable from the independent variables. The independent variables contain features like engine displacement, fuel type, and forced induction technologies (Nti et al., 2021). An r2 value ranges from 0 to 1, where 1 indicates a perfect fit. The value of 0.83 means that 83% of the variance in CO2 emissions can be explained by the model. Showing a strong correlation.

This technique assesses the model's performance on different subsets of the data. Cross-validation divides the dataset into different datasets and trains the model multiple times using a different subset each time (Nti et al., 2021). The reason for this is that the cross-validated MSE provides a more reliable estimate of the model's performance on unseen data. The value of 2169.28 suggests a slightly higher error than the initial MSE indicating the model's robustness. The standard deviation of the MSE values is obtained during the cross-validation process. It shows how much the MSE varies across different folds of the cross-validation process. A lower standard deviation indicates more consistent model performance across different subsets of the data (Nti et al., 2021).

5. Evaluation and discussion

This study provides a comprehensive analysis of the relationship between engine size and CO2 emissions. Additionally, the study provides a nuanced analysis of how engine displacement, cylinder count, and fuel type influence this relationship. There is a big emphasis on the impact of technological advancements like forced induction and engine downsizing. This study identifies significant trends and relationships using machine learning techniques that have helped to shape the automotive industry over the decades. The results indicate that engine downsizing and the adoption of forced induction have contributed to improved fuel efficiency and reduced emissions. Larger engine displacements correlate with higher CO2 emissions due to the increased volume of air and fuel combusted. Although the complexity of these relationships does warrant further exploration. This study contributes valuable insights into how vehicle features influence vehicle emissions whilst also offering a solid foundation for future research in this area.

The predictive models developed in this study demonstrated a strong correlation between engine size and CO2 emissions. This accounts for key vehicle features such as fuel type, engine displacement, and the presence of turbochargers and superchargers. The use of the ElasticNet technique provided a balanced approach to feature selection. The reason for the use of this technique was to ensure that only the most relevant variables were considered so the risk of an overfitting data model could be minimised.

The inclusion of neural networks helped to further enhance the predictive power of the study's models. However, this can be taken a step further to enhance predictive accuracy and capture temporal patterns in engine performance and emissions data. Model and predictive accuracy improvement can be done using Recurrent Neural Networks (RNNs). RNNs offer a significant advancement over plain neural networks because RNNs are designed to handle sequential data and learn from temporal dependencies.

This offers a more nuanced understanding of how different engine parameters interact to influence emissions. Findings show that CO2 emissions are higher on city roads compared to highways due to the 'Stop-Start' motion that continuously occurs on city roads. City driving involves frequent stops and starts. This means the vehicle is constantly accelerating from a standstill. Accelerating from a stop consumes more fuel compared to maintaining a steady speed because the engine must work harder to overcome inertia.

The number of cylinders inside an engine is a crucial factor which affects CO2 emissions. Engines with more cylinders generate more power but also tend to have higher fuel consumption rates that in turn result in greater vehicle emissions. There is a direct relationship between the number of cylinders and CO2 emissions meaning more cylinders equals significantly more CO2 emissions (Francisco & Posada, 2017). Explored in related studies, there have been promising results in predicting CO2 emissions using various models, indicating their effectiveness in emission forecasting (Chadha et al., 2023; Robertson, 2019).

During the study, other findings helped to reveal the limitations of the current approach, particularly in terms of the granularity of the data used. Even though the machine learning models captured the broad trends, they lack detailed engine parameters such as boost ratio, fuel injection type, fuel-to-air ratio, vehicle weight, engine design, compression ratios as well as many other engine technologies. It was likely that the accuracy of the prediction model was constrained.

A more in-depth view of this can be seen with the type of fuel used for vehicle commuting. The fuel type of an engine also plays a significant role in CO2 emissions. Different fuels have varying combustion efficiencies and emission profiles. The dataset does not account for this. However, more detailed data would enhance the analysis and model accuracy. Diesel engines typically emit more CO2 compared to gasoline engines due to their higher energy content and combustion characteristics. The study considers fuel types as a key variable. Results indicate that the type of fuel influences emissions levels. Therefore, aligning with prior research suggests that alternative fuels, such as biofuels offer reduced emissions compared to traditional fossil fuels. These factors play a critical role in determining an engine's efficiency and emissions output.

Therefore, the omission of these key factors from the dataset presents a significant area for improvement. Incorporating these finer details in future research would enable more precise modelling and deeper insights into how specific engine technologies impact CO2 emissions. Previous studies have demonstrated that turbochargers and superchargers improve engine efficiency by increasing air intake, and smaller turbocharged engines may emit less CO2 than larger naturally aspirated ones (Mahmoudi et al., 2017).

Integrating biofuels, renewable energy sources, and fuel cell technologies could offer viable strategies for reducing vehicle emissions and advancing towards a zero-emission future (Harvey, 2018; Ramli et al., 2011; Veziroglu & MacArio, 2011). From an environmental perspective, this research highlights the urgency needed for continued innovation in vehicle design and regulatory frameworks. The insights gained from this study can inform the development of more stringent emissions standards and guide the automotive industry towards greener technologies.

6. Conclusion

This study offers significant contributions to the understanding of the relationship between engine size and CO2 emissions in five-seater vehicles, which represent a significant portion of the global automotive market. This study opens the door for future research that delves deeper into the intricacies of engine design. The incorporation of more detailed engine parameters in predictive models will not only enhance the accuracy of emissions forecasts but also provide a more comprehensive view of how various technologies interact and influence vehicle performance as well as environmental impact. As automotive industry and policymakers continue to address the challenges of climate change. The insights from this research will play a crucial role in driving the future of automotive innovation. Steering the industry towards a new era of cleaner, more efficient vehicles that not only meet but set the standard for global environmental regulations.

7. References

Automobile Association Developments. (2017). Euro emissions standards AA. https://www.theaa.com/driving-advice/fuels-environment/euro-emissions-standards

Berkeley, N., Bailey, D., Jones, A., & Jarvis, D. (2017). Assessing the transition towards battery electric vehicles: A multi-level perspective on drivers of, and barriers to, take up. Transportation Research Part A: Policy and Practice, 106, 320–332. https://doi.org/10.1016/j.tra.2017.10.004

Blumberg, K., & Posada, F. (2015). Comparison of US and EU programs to control light-duty vehicle emissions. https://theicct.org/sites/default/files/ICCT_comparison%20Euro%20v%20US.pdf

Chadha, A. S., Shinde, Y., Sharma, N., & De, P. K. (2023). Predicting CO2 emissions by vehicles using machine learning. Lecture Notes on Data Engineering and Communications Technologies, 137, 197–207. https://doi.org/10.1007/978-981-19-2600-6_14

Costiuc, I, & Anghel, C. (2017). Evolution of the pressure wave supercharger concept. IOP Conf. Series: Materials Science and Engineering, 252, 1–17. https://doi.org/10.1088/1757-899X/252/1/012081

Duan, X., Lai, M. C., Jansons, M., Guo, G., & Liu, J. (2021). A review of controlling strategies of the ignition timing and combustion phase in homogeneous charge compression ignition (HCCI) engine. Fuel, 285, 119142. https://doi.org/10.1016/J.FUEL.2020.119142

Eckardt, D. (2022). 'Connections' and early turbo-jet developments 1935–1939. Jet Web, 247–399. https://doi.org/10.1007/978-3-658-38531-6_6

European Commission. (2020). The European Green Deal - European Commission. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

European Commission. (2024). European Climate Law - European Commission. https://climate.ec.europa.eu/eu-action/european-climate-law_en

Francisco, S., & Posada, F. (2017). South Africa's new passenger vehicle CO 2 emissions baseline analysis. The International Council on Clean Transportation, Final Report, 1–40. www.theicct.org

Galvin, R., & Healy, N. (2020). The Green New Deal in the United States: what it is and how to pay for it. Energy Research & Social Science, 67, 101529. https://doi.org/10.1016/J.ERSS.2020.101529

González, P. J. C., Furubayashi, T., & Nakata, T. (2012). Energy use and CO2 emissions reduction potential in passenger car fleet using zero emission vehicles and lightweight materials. Energy: The International Journal, 48(1), 548–565. https://doi.org/10.1016/J.ENERGY.2012.09.041

Harvey, L. D. D. (2018). Resource implications of alternative strategies for achieving zero greenhouse gas emissions from light-duty vehicles by 2060. Applied Energy, 212, 663–679. https://doi.org/10.1016/J.APENERGY.2017.11.074

Hawkins, T. R., Gausen, O. M., & Strømman, A. H. (2012). Environmental impacts of hybrid and electric vehicles review. International Journal of Life Cycle Assessment, 17(8), 997–1014. https://doi.org/10.1007/S11367-012-0440-9/METRICS

International Council on Clean Transportation. (2019). Gasoline vs. diesel comparing CO2 emission levels of a modern medium-sized car model under laboratory and on-road testing conditions background. ICCT Fact Sheet: Europe, 1, 1–3. https://theicct.org/sites/default/files/Gas%20_v%20_Diesel_%20CO2_emissions_%20EN_%20Fact%20_Sheet%202019_05_07_0.pdf

Kahn Ribeiro, S., Kobayashi, S., Hata, H., Sims, R., Olav Skjolsvik, K., Bose, R., Kheshgi, H., Ribeiro, K., Kobayashi, S., Beuthe, M., Gasca, J., Greene, D., Lee, D. S., Muromachi, Y., Newton, P. J., Plotkin, S., Sperling, D., Wit, R., & Zhou, P. J. (2007). Climate Change 2007. Mitigation of Climate Change, 4, 1–863. https://www.ipcc.ch/site/assets/uploads/2018/03/ar4_wg3_full_report-1.pdf

Macknick, J. (2011). Energy and CO2 emission data uncertainties. Carbon Management, 2(2), 189–205. https://doi.org/10.4155/CMT.11.10

Mahesh, B. (2018). Machine learning algorithms. International Journal of Science and Research, 9(1), 381-386 https://doi.org/10.21275/ART20203995

Mahmoudi, A. R., Khazaee, I., & Ghazikhani, M. (2017). Simulating the effects of turbocharging on the emission levels of a gasoline engine. Alexandria Engineering Journal, 56(4), 737–748. https://doi.org/10.1016/J.AEJ.2017.03.005

NRDC. (2017). The Paris Agreement on Climate Change. A(11). http://edgar.jrc.ec.europa.eu/overview.

Nti, I. K., Nyarko-Boateng, O., & Aning, J. (2021). Performance of machine learning algorithms with different K values in K-fold Cross-Validation. Article in International Journal of Information Technology and Computer Science, 6, 61–71. https://doi.org/10.5815/ijitcs.2021.06.05

O'Driscoll, R., Stettler, M. E. J., Molden, N., Oxley, T., & ApSimon, H. M. (2018). Real world CO2 and NOx emissions from 149 Euro 5 and 6 diesel, gasoline and hybrid passenger cars. Science of The Total Environment, 621, 282–290. https://doi.org/10.1016/J.SCITOTENV.2017.11.271

Paykani, A., Kakaee, A. H., Rahnama, P., & Reitz, R. D. (2015). Progress and recent trends in reactivity-controlled compression ignition engines. Sage Journals, 17(5), 481–524. https://doi.org/http://dx.doi.org/10.1177/1468087415593013

Plass, G. (1956). The Carbon Dioxide Theory of Climatic Change. Tellus, 8(2), 140–154. https://doi.org/10.1111/J.2153-3490.1956.TB01206.X

Pratt, D. (2023). Are turbocharged engines less reliable? - Consumer reports. https://www.consumerreports.org/cars/car-reliability-owner-satisfaction/are-turbocharged-engines-less-reliable-a5104151401/

Ramli, W., Daud, W., Najafpour, G., & Rahimnejad, M. (2011). Clean energy for tomorrow: towards zero emission and carbon free future. Official Peer Reviewed Journal of Babol Noshirvani University of Technology Iranica J. Energy & Environ, 2(3), 262–273. https://doi.org/10.5829/idosi.ijee.2011.02.03.0001

Robertson, L. S. (2019). Motor vehicle CO2 emissions in the United States: potential behavioral feedback and global warming. Weather, Climate, and Society, 11(3), 623–628. https://doi.org/10.1175/WCAS-D-18-0128.1

Roser, M., & Ritche, H. (2020, June). CO2 emissions - our world in data. How Much CO2 Does the World Emit? Which Countries Emit the Most? https://ourworldindata.org/co2-emissions

Kamal, S. S. L. B. A. (2019). Research paradigm and the philosophical foundations of a qualitative study. International Journal of Social Sciences, 4(3), 1386–1394. https://doi.org/10.20319/pijss.2019.43.13861394

Toyota Motor Corporation. (2017). The evolution of the Prius | Toyota Motor Corporation Official Global Website. https://global.toyota/en/prius20th/evolution/

United Nations. (2015). Transforming our world: the 2030 agenda for sustainable development preamble. Transforming Our World: The 2030 Agenda for Sustainable Development.

United States Government. (2024). Basic Information about the Emission Standards Reference Guide for On-road and Nonroad Vehicles and Engines US EPA. https://www.epa.gov/emission-standards-reference-guide/basic-information-about-emission-standards-reference-guide-road

US Environmental Protection Agency. (2017). Vehicle fuel economy estimates, 1984-2017. Kaggle. https://www.kaggle.com/datasets/epa/fuel-economy

Veziroglu, A., & MacArio, R. (2011). Fuel cell vehicles: state of the art with economic and environmental concerns. International Journal of Hydrogen Energy, 36(1), 25–43. https://doi.org/10.1016/J.IJHYDENE.2010.08.145

Younus, M., Purnomo, E. P., Husein, R., & Khairunnisa, T. (2023). Towards a greener future: promoting green and sustainable development in transportation operation. E3S Web of Conferences, 440, 1–17. https://doi.org/10.1051/e3sconf/202344001004

Vedran, M., & Božica, Z. (2018). Naturally aspirated gasoline engine upgrade with turbocharger - numerical investigation of change in operating parameters. International Scientific Journal ‘Machines. Technologies. Materials.’, 12(5), 204–207.

Zhang, J., Hu, T., Zhan, Z., He, J., & Zhang, C. (2017). Experimental investigation of performance of GDI engine with exhaust variable valve timing system. Chinese Internal Combustion Engine Engineering, 38(2). https://doi.org/10.13949/J.CNKI.NRJGC.2017.02.014