Prediction of carbon emissions from vehicles using interpretable neural networks

Ridho Sholehurrohman[1], Mochammad Reza Habibi[2,3], Joko Triloka[4]
[1]Department of Computer Science, Faculty of Mathematics and Natural Sciences, Lampung University, Bandar Lampung, Indonesia
[2]Department of Business Statistics, Faculty of Vocational Studies, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[3]Department of Computer Science and Electrical Engineering, Kumamoto University, Kumamoto, Japan
[4]Department of Master of Technology Management, Institut Informatika dan Bisnis (IIB) Darmajaya, Bandar Lampung, Indonesia

Article Info

Article history:

Received Dec 7, 2025
Revised Jun 25, 2026
Accepted Jul 23, 2026

Keywords:

Carbon emission
Neural network
Prediction of carbon emission
SHapley additive exPlanations model
Vehicle carbon dioxide emissions

ABSTRACT

Vehicle carbon dioxide emissions are an important environmental concern that requires prediction methods with both high accuracy and transparent explanations. This study proposes an interpretable neural network (INN) framework for estimating vehicle carbon dioxide emissions using a public Canadian fuel-consumption and emissions dataset. A feed-forward neural network is employed for regression, while SHapley additive exPlanations (SHAP) are used to interpret the contribution of each input feature. The proposed model is compared with linear regression, Bayesian regression, support vector regression (SVR), and extreme gradient boosting (XGBoost). Experimental results show that the neural-network model achieves a root mean squared error (RMSE) of 4.98, a mean absolute percentage error (MAPE) of 0.0128, and a coefficient of determination (R-squared) of 0.9926, outperforming all baseline models. SHAP analysis identifies combined fuel economy, city fuel consumption, and highway fuel consumption as the most influential predictors. Higher fuel economy is associated with lower predicted emissions, whereas higher city and highway fuel consumption increases predicted emissions. These findings demonstrate that the proposed framework provides accurate and interpretable vehicle carbon dioxide emission estimation.

Corresponding Author:

Ridho Sholehurrohman
Department of Computer Science, Faculty of Mathematics and Natural Sciences, Lampung University
Bandar Lampung, Indonesia
Email: ridho.sholehurrohman@fmipa.unila.ac.id

1. INTRODUCTION

Climate change has become one of the most pressing global challenges, primarily driven by the accumulation of greenhouse gases (GHGs) in the atmosphere. These gases, including carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O), play a crucial role in regulating Earth's temperature by absorbing and emitting infrared radiation [1]. However, when their concentrations exceed natural levels, they disrupt the atmospheric balance, leading to severe environmental consequences such as rising global temperatures, extreme weather patterns, and ecological disturbances. Among the key contributors to GHG emissions is the transportation sector, which continues to expand due to increased mobility and economic growth. Motorized vehicles, which predominantly rely on fossil fuels, are among the largest sources of carbon emissions. These emissions not only accelerate global warming but also degrade air quality, negatively impacting human health and ecological sustainability [1], [2]. As urbanization and population growth continue, the number of vehicles on the road is rising significantly, leading to a steady increase in emissions [3]. The main GHGs released from vehicle exhaust include CO_2, nitrogen oxides (NOx), and fine particulates, which contribute to both global climate change and localized air pollution [4], [5].

To illustrate the scale of vehicle emissions, Figure 1 (source https://www.bbc.com/news/science-environment-49349566) presents information on carbon emissions released by vehicles, highlighting their substantial contribution to total GHG emissions. This significant contribution emphasizes the urgent need for reliable monitoring and prediction methods to assess vehicle emissions and support effective mitigation strategies.

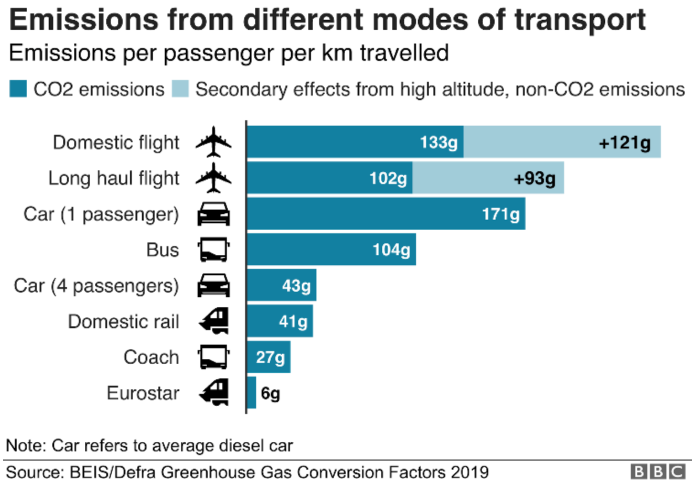


Figure 1. Carbon emissions from vehicles

Given the growing environmental concerns associated with vehicular emissions, there is a critical need to develop accurate and reliable methods for predicting carbon output. Monitoring and predicting emissions can play a crucial role in designing sustainable transportation policies and encouraging the transition to cleaner technologies. However, traditional methods used for emission prediction, such as statistical regression models, may have limited ability to capture complex and nonlinear relationships among vehicle characteristics, fuel-consumption patterns, and CO_2 emissions [4].

To overcome the limitations of conventional models, machine learning (ML) and deep learning (DL) techniques have been increasingly explored in recent years. DL models, particularly long short-term memory (LSTM), and bidirectional LSTM (BiLSTM), have shown strong predictive capabilities for CO_2 emissions. For instance, Al-Nefaie and Aldhyani [6] applied DL models to open vehicle emissions dataset and achieved accuracy rates exceeding 80%. Their results demonstrated the effectiveness of AI-driven models in improving prediction accuracy.

However, despite their high predictive performance, DL models suffer from a major drawback: a lack of interpretability. Many of these models operate as "black boxes," making it difficult to understand the reasoning behind their predictions. This lack of transparency poses challenges for policymakers, environmental regulators, and industry stakeholders, who require clear and interpretable insights to implement effective emission control strategies [7]. In practical applications, predictive accuracy alone is insufficient because decision-makers also need to understand the extent to which engine characteristics, fuel type, and fuel-consumption variables contribute to individual emission estimates.

While Al-Nefaie and Aldhyani [6] successfully improved emission prediction accuracy, it did not address the critical issue of model interpretability. The inability to explain how vehicle features, such as engine size, fuel type, and vehicle weight, affect emissions limits the practical application of these models in real-world decision-making. As a result, there is a growing demand for DL models that not only provide accurate predictions but also offer explainable insights into their decision-making process.

To bridge this gap, this study proposes an interpretable prediction framework based on interpretable neural networks (INNs), which combine neural networks with SHapley additive exPlanations (SHAP) to estimate carbon emissions from vehicles. The neural network is used for prediction, while SHAP is applied to provide feature-level explanations and quantify the contribution of each variable. This approach enables the identification of key factors influencing emissions and improves model transparency. As a result, the proposed framework achieves both high predictive performance and interpretability, making it suitable for environmental policy-making and regulatory applications.

In addition, many countries aim to achieve net-zero emissions by 2050. Given the significant contribution of vehicle emissions, as shown in Figure 1, there is a clear need for models that are both accurate and interpretable. Canada is one such country that has set a net-zero target for 2050 and provides publicly available data on vehicle fuel consumption and CO_2 emission measurement. This commitment is formally established through the Canadian net-zero emissions accountability act, which mandates a transparent and accountable process for achieving its national emissions targets [8]. Canada is therefore selected as the case-study context because it provides an appropriate combination of an established national emission-reduction commitment, publicly accessible vehicle data, and an increasing need for transparent evidence to support vehicle-emission assessment. The proposed framework is expected to be extendable to datasets from other countries in support of broader net-zero emission goals.

This study makes the following contributions to carbon emission analysis and AI-driven environmental monitoring: i) we develop an interpretable DL model that promotes transparency in emission prediction, addressing the limitations of black-box approaches, ii) we combine neural networks with SHAP to provide feature importance analysis, enabling a quantitative explanation of how vehicle characteristics influence emission outputs, iii) we demonstrate that the proposed approach effectively balances predictive accuracy and interpretability, making it suitable for real-world regulatory and policy-making applications, and iv) we identify key factors influencing CO_2 emissions, providing insights that support the development of sustainable transportation policies and eco-friendly vehicle technologies.

In summary, this study proposes an interpretable neural-network framework that combines a neural network for CO_2 emission prediction with SHAP for post-hoc explanation. The scientific contribution lies in the integration of comparative model benchmarking, global feature-importance analysis, and directional dependence analysis within a unified neural network centered framework for Canadian vehicle emission data. By addressing the interpretability limitations of neural-network models, the proposed approach provides accurate predictions and transparent feature-level evidence for environmental analysis and decision-making.

2. METHOD

This section describes the proposed methodology, including the overall framework, the neural-network model for prediction, the baseline models used for performance comparison, and the SHAP based interpretation approach. The proposed approach combines a feed-forward neural network with post-hoc SHAP explanations; therefore, the term INN in this study refers to an explainable prediction framework rather than a newly developed intrinsically interpretable architecture. The workflow of the proposed method is illustrated in Figure 2.

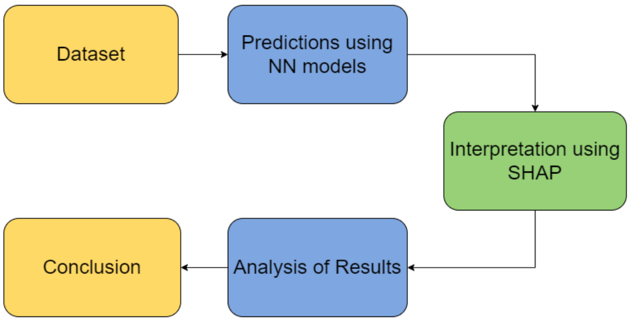


Figure 2. Proposed framework

2.1. Overall framework

The proposed framework consists of four main stages: data processing, prediction, interpretation, and result analysis, with relevant references ensuring scientific validity [9], [10]. First, the dataset is used to train a neural network model for CO_2 emission prediction. Next, SHAP is applied to interpret the model's predictions by quantifying the contribution of each input feature. The resulting explanations are then analyzed to identify key factors influencing emissions. Finally, conclusions are drawn based on both predictive performance and interpretability. The overall workflow is shown in Figure 2.

2.1.1. Neural networks

The first component of the proposed method, the INNs model, leverages a neural network for prediction where the input data undergo multiple layers of weighted transformations and nonlinear activations. In this process, understanding a prediction requires considering a large number of parameters interacting in a complex manner [11]. As a result, interpreting the behavior and predictions of neural networks is inherently challenging and often requires specialized interpretation methods.

While model-agnostic approaches such as local surrogate models or partial dependence plots can be applied, there are advantages to using interpretation techniques tailored for neural networks. First, neural networks learn hierarchical representations and abstract features within their hidden layers, which are not directly observable and therefore require dedicated methods to analyze. Second, the model's gradient information can be used to derive explanations efficiently, offering computational advantages over model-agnostic approaches that treat the model as a black box. More importantly, deep neural networks are particularly effective at learning high-level feature representations, reducing the need for extensive manual feature engineering [7], [12].

The neural network architecture used in this study is illustrated in Figure 3, which consists of an input layer, two hidden layers, and an output layer for regression. The input layer receives vehicle-related features from the dataset and serves as the entry point of the model. The hidden layers perform a series of nonlinear transformations, enabling the model to capture complex relationships between input variables and CO_2 emissions. The output layer produces the final prediction based on the learned representations. The specific structure of the output layer depends on the task [13]. In this study, it is designed to perform regression for CO_2 emission estimation. An example of the neural network architecture is shown in Figure 3.

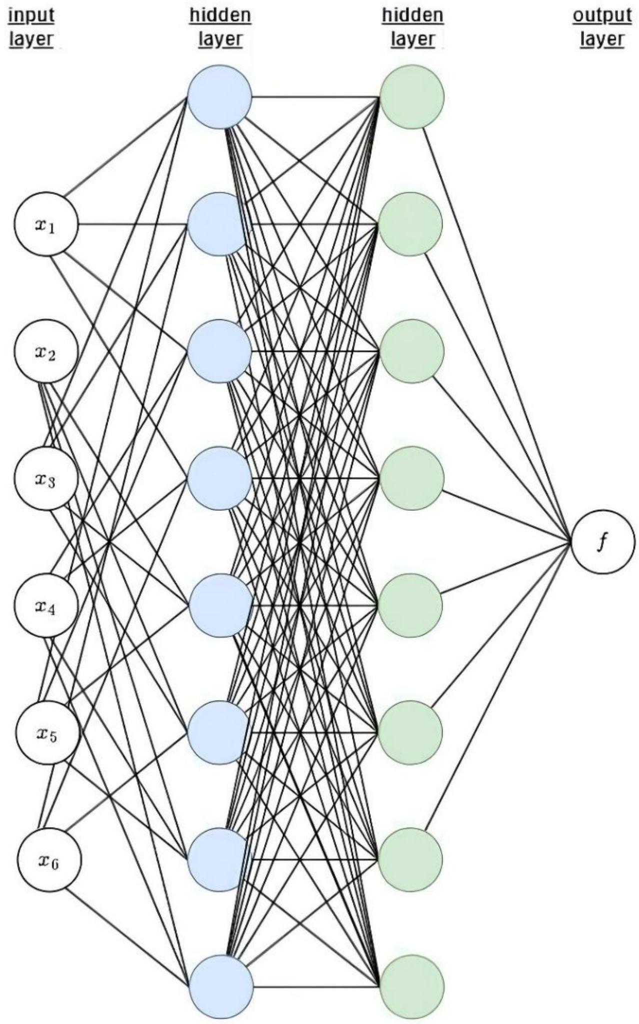


Figure 3. Basic architecture of a neural network

Despite their strong predictive capability, neural networks are often considered black-box models due to their limited interpretability. This limitation motivates the use of an additional interpretation method, as described in the next section.

2.1.2. SHapley additive exPlanations

The second component of the proposed method is SHAP, an interpretation technique that is not tied to a specific neural network architecture. Instead, it provides a unified framework for explaining predictions across various ML models, including DL and ensemble methods [14], [15]. SHAP values provide a quantitative measure of the contribution of each feature to a model's prediction. Specifically, SHAP explains the prediction for a given instance x by assigning an importance value to each feature, indicating its contribution to the final output. This approach is based on the concept of Shapley values from cooperative game theory [12], [16], where features are treated as players in a game, and the prediction is considered the total gain to be fairly distributed among them.

In this formulation, each feature contributes to the prediction depending on its marginal contribution across different combinations of features. For tabular data, each player corresponds to a feature or a set of attribute values, and SHAP evaluates how the inclusion or exclusion of each feature affects the prediction [16], [17]. A key advantage of SHAP is its formulation as an additive explanation model, which expresses the prediction as a linear combination of feature contributions. This formulation unifies ideas from Shapley values and local explanation methods such as LIME [16], [18]. The SHAP explanation model is defined as (1):

$$g(z') = \phi_0 + \sum_{j=1}^{M} \phi j z'_j \tag{1}$$

Where $g(z')$ denotes the additive explanation model, and $z' \in \{0,1\}^M$ represents a coalition vector indicating the presence or absence of input features. The variable z'_j indicates whether feature j is included in the coalition, where $z'_j = 1$ means that the feature is present and $z'_j = 0$ means that the feature is absent. The symbol M denotes the total number of input features or the maximum coalition size. The term ϕ_0 represents the expected model output or baseline prediction, while ϕ_j denotes the SHAP value assigned to feature j. The SHAP value ϕ_j quantifies the contribution of feature j to the difference between the baseline prediction and the final prediction for a particular instance.

In this study, SHAP is applied to interpret the predictions of the neural network model by quantifying the contribution of each input feature to CO_2 emissions. This enables a detailed understanding of how vehicle attributes influence model outputs. However, it should be noted that while SHAP provides valuable insights into feature importance, interpretation must be performed carefully, as feature contributions may vary depending on feature interactions and data distribution [16], [19]. Overall, SHAP enables a transparent analysis of model predictions by revealing the contribution of each feature, thereby supporting interpretable and reliable CO_2 emission estimation.

2.2. Experimental setup

2.2.1. Dataset

As stated in the Introduction, this study utilizes a vehicle emissions dataset from Canada, obtained from the Kaggle repository and originally curated from the Canadian government open data website. The dataset contains various vehicle attributes that influence CO_2 emissions, including engine characteristics and fuel consumption metrics. These features are used as inputs to the neural network model to predict CO_2 emissions. By analyzing these features, the study aims to identify key factors contributing to vehicular carbon emissions and provide insights into potential emission reduction strategies. The dataset consists of numerical variables represented as integer and floating-point values. Table 1 summarizes the key features used in this study.

Table 1. Datasets features

Variable	Type
Engine size (L)	Float
Cylinders	Int
Transmission	Int
Fuel type	Int
Fuel consumption city (L/100 km)	Float
Fuel consumption hwy (L/100 km)	Float
Fuel consumption comb (L/100 km)	Float
Fuel consumption comb (mpg)	Float
Engine size (L)	Float

The dataset includes a wide range of vehicle models with diverse characteristics, enabling a comprehensive analysis of CO_2 emissions in the transportation sector. Among these features, engine size is an important factor, as larger engines generally produce higher emissions due to increased fuel consumption. However, fuel consumption is identified as the most influential factor, as it directly reflects the amount of fuel used per unit distance. Since CO_2 emissions are proportional to fuel consumption, vehicles with higher fuel usage tend to produce higher emissions. This relationship highlights the importance of improving fuel efficiency to reduce carbon emissions.

2.2.2. Implementation details

The performance of a neural network model is influenced by the choice of hyperparameters, which determine both the network structure and the training process. In this study, key hyperparameters are selected empirically to ensure stable training and reliable predictive performance. The selected configuration is summarized in Table 2.

Table 2. Hyperparameter settings

Hyperparameter	Value
Number of hidden layers	2
Number of epochs	100
Number of nodes	{64,64}
Activation	ReLU
Optimizer	Adam

3. RESULTS AND DISCUSSION

3.1. Model training

Based on the selected hyperparameters, a neural network model was constructed to estimate CO_2 emissions using the features described in Table 1. This approach aligns with recent studies that have successfully applied artificial neural networks for fuel consumption modeling [19]. Figure 4 illustrates the training process of the model.

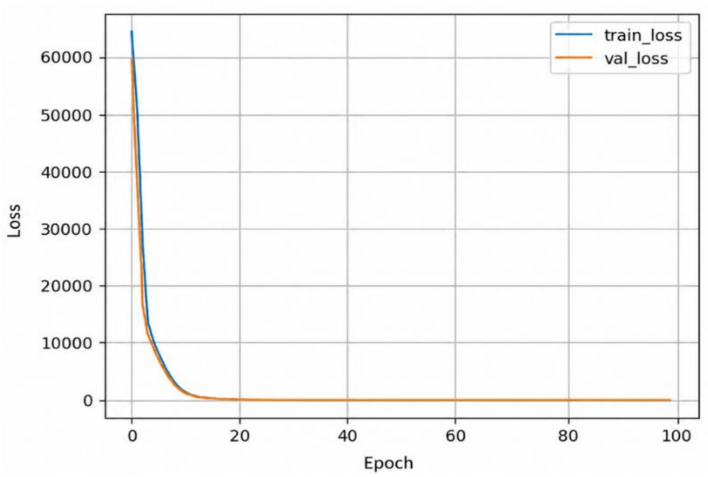


Figure 4. Loss curve

As shown in Figure 4, the loss value of the proposed model decreases as the number of epochs increases, indicating that the model learns effectively during training. In addition, the training and validation losses are closely aligned, suggesting that the model does not suffer from overfitting and demonstrates good generalization performance. Such behavior is consistent with well-optimized DL architectures where hyperparameter tuning and regular monitoring of loss curves ensure robust predictive capability [20].

3.2. Model comparison

To evaluate the effectiveness of the proposed approach, a comparative analysis was conducted against several baseline and advanced models, including linear regression, Bayesian regression, support vector regression (SVR), and extreme gradient boosting (XGBoost). These models are widely used for regression tasks and provide a benchmark for evaluating predictive performance [21]. Specifically, SVR is noted for its effectiveness in high-dimensional spaces [22], while XGBoost has gained prominence for its scalability and performance in structured data tasks [23].

The performance of the models was evaluated using root mean squared error (RMSE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2). The proposed model achieves an RMSE of 4.9818, a MAPE of 0.0128, and an R^2 of 0.9926, indicating low prediction error and a strong fit to the data. The high R^2 value suggests that the model explains 99.26% of the variance in CO_2 emissions. Additional statistical validation techniques, such as hypothesis testing (e.g., t-test or F-test) and cross-validation, may be considered to further examine the robustness and statistical significance of the observed performance. The comparison results are presented in Table 3.

Table 3. Benchmark results

Model	RMSE	MAPE	R^2
Linear regression	17.6115	0.04455	0.90167
Bayessian regression	17.6130	0.04454	0.90162
SVR	14.2814	0.02408	0.92539
XGBoost	16.2758	0.05560	0.89897
Proposed model	**4.9818**	**0.01280**	**0.99260**

As shown in Table 3, the proposed neural network model outperforms all baseline methods across the evaluation metrics. It achieves the lowest RMSE, indicating the smallest prediction error, and the lowest MAPE, reflecting high prediction consistency. In addition, the model attains the highest R^2 value, demonstrating a strong ability to explain the variance in CO_2 emissions. Compared to traditional statistical models such as linear and Bayesian regression, as well as advanced ML methods such as SVR and XGBoost, the proposed model demonstrates superior predictive capability. These results indicate that the proposed approach is effective in capturing the complex relationships between vehicle attributes and CO_2 emissions.

3.3. Model interpretation

Although the proposed neural network model demonstrates strong predictive performance, it remains inherently difficult to interpret due to its black-box nature. To address this limitation, SHAP are used to provide post-hoc interpretability. SHAP is based on cooperative game theory and enables the quantification of the contribution of each input variable to the model's predictions [24]. This allows for a better understanding of how different features influence CO_2 emission estimation.

Figure 5 illustrates the relative importance of each input variable based on their SHAP values. The results show that fuel consumption comb (X8), fuel consumption city (X5), and fuel consumption hwy (X6) are the most influential variables in the model. This indicates that fuel consumption-related features play a dominant role in determining CO_2 emissions, a finding consistent with recent studies analyzing vehicle emission datasets using explainable ML [25], [26].

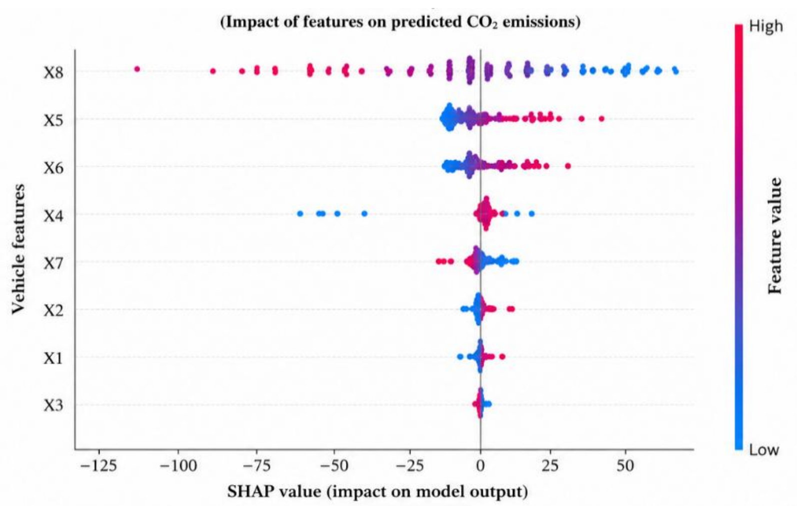


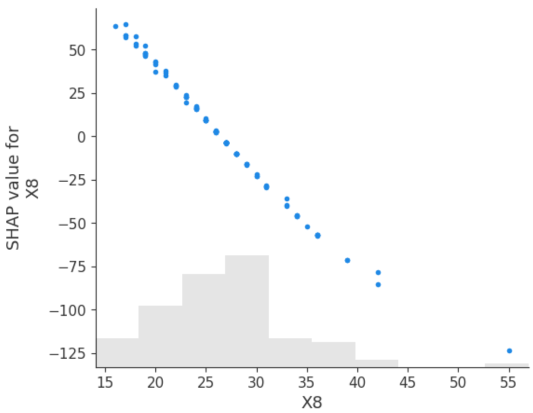
Figure 5. SHAP values

The SHAP values presented in Table 4 quantify the contribution of each feature to the model predictions. Specifically, the mean absolute SHAP value represents the average impact of a feature on the model output. A higher value indicates a greater influence on the prediction. For example, X8 has the highest SHAP value (28.76), indicating that it is the most influential feature in determining CO_2 emissions compared to the other variables.

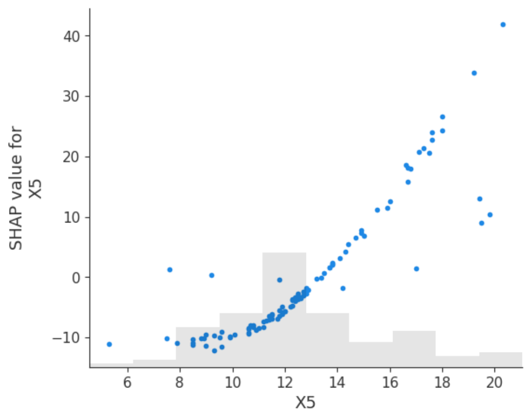
To further analyze the relationships between key variables and CO_2 emissions, Figure 6 illustrates the relationships between the three most influential variables (X8, X5, and X6) and the CO_2 estimates produced by the model. As shown in Figure 6(a), X8 is associated with lower predicted CO_2 emissions as fuel economy increases. Figures 6(b) and (c) show that X5 and X6 exhibit positive relationships with predicted CO_2 emissions. This indicates that higher fuel consumption under both city and highway driving conditions is associated with higher predicted emissions. This finding is consistent with domain knowledge, as greater fuel usage is closely associated with higher carbon output [27].

Table 4. Mean absolute SHAP values

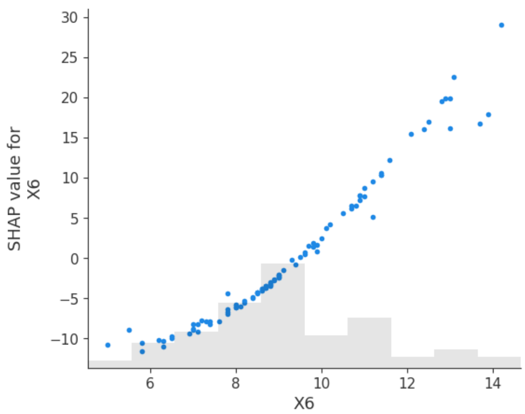
Variable/feature	SHAP value
X8	28.76
X5	8.58
X6	7.04
X4	5.24
X7	3.06
X2	1.52
X1	1.02
X3	0.43



(a)



(b)



(c)

Figure 6. Feature contributions of the top three variables; (a) fuel consumption comb (X8), (b) fuel consumption city (X5), and (c) fuel consumption hwy (X6)

For fuel consumption comb (X8), the relationship observed in Figure 6(a) indicates a more complex pattern. While variations in X8 significantly influence the model predictions, the relationship is not strictly linear and may reflect interactions with other variables or underlying data characteristics. This suggests that the combined fuel consumption metric captures additional information beyond individual driving conditions, contributing to its dominant importance in the model.

Overall, the interpretability analysis highlights that fuel consumption-related variables are the primary drivers of CO_2 emissions in the dataset. These findings are consistent with established knowledge in transportation and environmental studies, where fuel consumption is directly linked to carbon output. By providing both accurate predictions and interpretable insights, the proposed approach offers a practical tool for understanding and analyzing vehicle emissions. These results are particularly relevant for policymakers and vehicle manufacturers, as they emphasize the importance of improving fuel efficiency in both urban and highway driving conditions. Furthermore, the use of SHAP enhances transparency in the model, enabling more informed decision-making in environmental monitoring and sustainable transportation planning.

4. CONCLUSION

This study contributes an interpretable prediction framework for vehicle CO_2 emission estimation that effectively balances predictive accuracy and model transparency. The proposed framework combines a neural network for prediction with SHAP for post-hoc interpretability. Experimental results demonstrate that the proposed model achieves superior predictive performance with an RMSE of 4.98, a MAPE of 0.0128, and an R^2 of 0.9926, outperforming baseline methods including linear regression, Bayesian regression, SVR, and XGBoost. The SHAP-based analysis reveals that fuel consumption and fuel economy variables—particularly combined fuel economy, city fuel consumption, and highway fuel consumption—are the most influential predictors of CO_2 emissions. These findings indicate that improving vehicle fuel efficiency should remain an important consideration in emission-reduction strategies.

From a practical perspective, these findings offer valuable insights for policymakers, environmental regulators, and vehicle manufacturers. The ability to identify and quantify the contribution of specific vehicle attributes to CO_2 emission predictions supports more informed decision-making for sustainable transportation planning. The proposed framework is particularly relevant to the Canadian vehicle-emission context and aligns with Canada's commitment to achieving net-zero emissions by 2050. Despite these contributions, this study has several limitations. First, the interpretation relies on SHAP, which provides post-hoc explanations based on the trained model. Second, the current analysis is limited to a single dataset from Canada and does not account for regional variations. Third, the analysis does not explicitly model the dynamic impact of policy interventions over time. Therefore, future work will focus on extending the framework to support simulation-based policy analysis, particularly through integration with multi-agent systems or digital-twin environments. Additionally, future research will explore its applicability to other countries and the incorporation of real-time vehicle data for dynamic emission monitoring.

ACKNOWLEDGMENTS

The authors wish to thank the Basic Computing Laboratory, Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Lampung, and especially the Private Laboratory (2nd Floor, Computer Science Building, FMIPA Universitas Lampung), for their invaluable support and facilities.

FUNDING INFORMATION

This research received no external funding. The article processing charge (APC) for this publication was financially supported by Universitas Lampung.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Ridho Sholehurrohman	✓	✓	✓		✓	✓		✓	✓	✓	✓			
Mochammad Reza Habibi	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	
Joko Triloka				✓		✓	✓			✓	✓		✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are openly available in the Kaggle repository and were originally curated from the Canadian government open data portal. The dataset is publicly accessible and was used in accordance with the platform's terms of use. The corresponding author can provide further details regarding data access upon reasonable request.

REFERENCES

[1] C. Klemm and F. Wiese, “Indicators for the optimization of sustainable urban energy systems based on energy system modeling,” *Energy, Sustainability and Society*, vol. 12, no. 1, pp. 1-20, Dec. 2022, doi: 10.1186/s13705-021-00323-3.

[2] M. Menendez and L. Ambühl, “Implementing Design and Operational Measures for Sustainable Mobility: Lessons from Zurich,” *Sustainability*, vol. 14, no. 2, pp. 1-21, Jan. 2022, doi: 10.3390/su14020625.

[3] M. Aamir *et al.*, “Predicting the Environmental Change of Carbon Emission Patterns in South Asia: A Deep Learning Approach Using BiLSTM,” *Atmosphere*, vol. 13, no. 12, pp. 1-14, Nov. 2022, doi: 10.3390/atmos13122011.

[4] N. Niroomand, C. Bach, and M. Elser, “Segment-Based CO_2 Emission Evaluations From Passenger Cars Based on Deep Learning Techniques,” *IEEE Access*, vol. 9, pp. 166314–166327, 2021, doi: 10.1109/ACCESS.2021.3135604.

[5] R. Zhang *et al.*, “A Deep Learning Micro-Scale Model to Estimate the CO2 Emissions from Light-Duty Diesel Trucks Based on Real-World Driving,” *Atmosphere*, vol. 13, no. 9, pp. 1-12, Sep. 2022, doi: 10.3390/atmos13091466.

[6] A. H. Al-Nefaie and T. H. H. Aldhyani, “Predicting CO2 Emissions from Traffic Vehicles for Sustainable and Smart Environment Using a Deep Learning Model,” *Sustainability*, vol. 15, no. 9, pp. 1-21, May 2023, doi: 10.3390/su15097615.

[7] Q. Zhang, Y. N. Wu, and S.-C. Zhu, “Interpretable Convolutional Neural Networks,” in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 8827–8836, doi: 10.1109/CVPR.2018.00920.

[8] Government of Canada, “Canadian Net-Zero Emissions Accountability Act: S.C. 2021, c. 22,” Government of Canada.

[9] S. Altikat, “Prediction of CO2 emission from greenhouse to atmosphere with artificial neural networks and deep learning neural networks,” *International Journal of Environmental Science and Technology*, vol. 18, no. 10, pp. 3169–3178, Oct. 2021, doi: 10.1007/s13762-020-03079-z.

[10] L. Amarpuri, N. Yadav, G. Kumar, and S. Agrawal, “Prediction of CO 2 emissions using deep learning hybrid approach: A Case Study in Indian Context,” in *2019 Twelfth International Conference on Contemporary Computing (IC3)*, Noida, India: IEEE, Aug. 2019, pp. 1–6, doi: 10.1109/IC3.2019.8844902.

[11] C. Miles *et al.*, “Correlator convolutional neural networks as an interpretable architecture for image-like quantum matter data,” *Nature Communications*, vol. 12, no. 1, pp. 1-7, Jun. 2021, doi: 10.1038/s41467-021-23952-w.

[12] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Leanpub, 2019.

[13] N. Fortelny and C. Bock, “Knowledge-primed neural networks enable biologically interpretable deep learning on single-cell sequencing data,” *Genome Biology*, vol. 21, no. 1, pp. 1-36, Dec. 2020, doi: 10.1186/s13059-020-02100-5.

[14] K. Liu, R. Ye, L. Zhongzhu, and R. Ye, “Entropy-based discrimination between translated Chinese and original Chinese using data mining techniques,” *PLOS ONE*, vol. 17, no. 3, pp. 1-26, Mar. 2022, doi: 10.1371/journal.pone.0265633.

[15] R. I. Hamilton and P. N. Papadopoulos, “Using SHAP Values and Machine Learning to Understand Trends in the Transient Stability Limit,” *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 1384–1397, Jan. 2024, doi: 10.1109/TPWRS.2023.3248941.

[16] A. Cooper, “Explaining Machine Learning Models: A Non-Technical Guide to Interpreting SHAP Analyses,” Aidan Cooper, Nov 1, 2021, [Online]. Available: https://www.aidancooper.co.uk/a-non-technical-guide-to-interpreting-shap-analyses/. (Accessed: Mar. 25, 2026).

[17] S.-Y. Wang, S.-T. Liong, Y. S. Gan, and Y.-T. Sheng, “Cost-effective concrete fabrication for large irregularly shaped architectural structures,” *Automation in Construction*, vol. 156, Dec. 2023, doi: 10.1016/j.autcon.2023.105119.

[18] L. Jostins *et al.*, “Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease,” *Nature*, vol. 491, no. 7422, pp. 119–124, Nov. 2012, doi: 10.1038/nature11582.

[19] H. Xiao, Z. Wang, Z. Zhu, J. Zhou, and J. Lu, “Shapley-NAS: Discovering Operation Contribution for Neural Architecture Search,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2022, pp. 11882–11891, doi: 10.1109/CVPR52688.2022.01159.

[20] S. S. Haghshenas, S. S. Haghshenas, V. Astarita, and G. Guido, “Fuel consumption and CO_2 emissions prediction in road transport using a hybrid deep learning approach,” *Transportation Engineering*, vol. 21, pp. 1-13, Sep. 2025, doi: 10.1016/j.treng.2025.100364.

[21] G. Çınarer, M. K. Yeşilyurt, Ü. Ağbulut, Z. Yılbaşı, and K. Kılıç, “Application of various machine learning algorithms in view of predicting the CO 2 emissions in the transportation sector,” *Science and Technology for Energy Transition*, vol. 79, Mar. 2024, doi: 10.2516/stet/2024014.

[22] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*. in Springer Series in Statistics. New York, NY: Springer New York, 2009, doi: 10.1007/978-0-387-84858-7.

[23] T. Chen and C. Guestrin, “XGBoost,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge*

Discovery and Data Mining, New York, NY, USA: ACM, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
[24] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems*, pp. 4768-4777, 2017, doi: 10.5555/3295222.3295230.
[25] X. Guo, R. Kou, and X. He, "Towards Carbon Neutrality: Machine Learning Analysis of Vehicle Emissions in Canada," *Sustainability*, vol. 16, no. 23, pp. 1-18, Nov. 2024, doi: 10.3390/su162310526.
[26] A. Arifudin, N. Alamsyah, Budiman, A. Hendra, and E. Setiana, "Explainable AI for CO_2 Emission Prediction Using XGBoost Enhanced by Optuna and SHAP," in *2025 International Conference on Smart Computing, IoT and Machine Learning (SIML)*, IEEE, Jun. 2025, pp. 1–6, doi: 10.1109/SIML65326.2025.11081028.
[27] J. Sun, X. Mo, and S. Wang, "A Comparative Study of Explainable Machine Learning and Multivariate Regression for Predicting Fuel Consumption and CO2 Emissions in Multi-Brand Passenger Vehicles," *Prerint Research Square*, Feb. 2026, doi: 10.21203/rs.3.rs-8762253/v1.

BIOGRAPHIES OF AUTHORS

Ridho Sholehurrohman having graduated from Maulana Malik Ibrahim State Islamic University of Malang with a major in Developmental Mathematics Growth, he completed his master's degree from Sepuluh Nopember Institute of Technology in Surabaya. Currently, he is an active lecturer at the Department of Computer Science, University of Lampung. In addition to his teaching responsibilities, he is passionate about independent research and has outstanding publications. His research interests include data handling, object detection, and classification. Moreover, he actively contributes to community service activities facilitated by the university. He can be contacted at email: ridho.sholehurrohman@fmipa.unila.ac.id.

Mochammad Reza Habibi graduated from the Sepuluh Nopember Institute of Technology in Surabaya, successfully obtained his master's degree in computational mathematics. He is currently a Ph.D. student at the Department of Computer Science and Electrical Engineering, Kumamoto University, Japan. He has been actively engaged as a lecturer in the Department of Business Statistics at Institut Teknologi Sepuluh Nopember. Alongside his teaching responsibilities, he demonstrates a keen interest in conducting research, particularly in the field of machine learning and its applications. Additionally, he actively participates in community service activities facilitated by the university. He can be contacted at email: reza.habibi@its.ac.id.

Joko Triloka received his Ph.D. in computer science from Universiti Brunei Darussalam, Brunei Darussalam, in 2018. He is currently a lecturer at the Master of Informatics Engineering, Institut Informatika dan Bisnis (IIB) Darmajaya, Indonesia. His research interests include motion tracking systems, image processing, and pattern recognition, with applications in medical and sports science. His works have been published in journals such as Neural Computing and Applications and Journal of Musculoskeletal Research. He can be contacted at email: joko.triloka@darmajaya.ac.id.