International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 05 | May – 2025 DOI: 10.55041/ISJEM03813
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

AVERAGE FUEL CONSUMPTION OF HEAVY VEHICLES USING MACHINE LEARNING

Dr. Sirisha K.L.S[1], G. Durga Harshitha[2], E. Pavan Kumar[3], G. Bhanu Teja Reddy[4], G.Pranay[5]

[1-5] Department of CSE & TKR College of Engineering & Technology
[2-5]cB.Tech Students

---***---

ABSTRACT

This paper introduces a novel approach to developing individualized machine learning models for predicting fuel consumption in heavy vehicles. Unlike traditional methods that rely on time-based intervals, our model leverages vehicle travel distance as the primary basis for predictions. We integrate seven key predictors derived from vehicle speed and road grade to construct a highly accurate neural network model. This proposed model offers the flexibility for straightforward development and deployment across individual vehicles within an entire fleet, thereby enabling optimized fuel consumption management. To achieve this, the model's predictors are aggregated over fixed window sizes of distance travelled. Our evaluation of different window sizes demonstrates that a 1 km window effectively predicts fuel consumption, achieving a coefficient of determination of 0.91. Furthermore, it yields a mean absolute peak-to-peak percentage error of less than 4% for diverse routes, including both city and highway driving cycles. This study highlights the potential of distance-based modeling for more precise and practical fuel efficiency insights in heavy vehicle operations.

Keywords — Fuel Consumption, Machine Learning, Neural Network, Heavy Vehicles, Vehicle Travel Distance, Fleet Optimization, Predictors

I. INTRODUCTION

The transportation sector, particularly heavy vehicle operations, plays a pivotal role in global logistics and economic activity. However, it is also a significant contributor to energy consumption and environmental emissions. Optimizing fuel efficiency in heavy vehicles is therefore a critical concern, addressing both economic viability for fleet operators and broader ecological impacts. Traditional methods of monitoring and managing fuel consumption often rely on aggregated data or simplistic models, which may not capture the nuanced factors influencing real-world fuel usage in individual vehicles under varying operational conditions.

The complexity of factors affecting fuel consumption, such as driving behavior, road topography, traffic conditions, vehicle load, and maintenance, necessitates sophisticated analytical approaches. Conventional predictive models often fall short in providing individualized insights or adapting to the dynamic nature of vehicle operations. This limitation highlights a gap in current fuel management strategies, where a more precise and adaptable predictive framework could yield substantial benefits in terms of cost savings and environmental footprint reduction.

Machine Learning (ML) presents a robust framework for addressing these challenges due to its capacity to identify complex patterns and relationships within large datasets. By leveraging ML algorithms, it is possible to develop highly accurate predictive models that account for multivariate inputs and deliver personalized insights for each vehicle. This project aims to harness the power of machine learning to develop a more effective method for predicting average fuel consumption in heavy vehicles.

Our primary objective is to create individualized machine learning models that can accurately predict fuel consumption. A distinguishing aspect of our approach is the focus on using vehicle travel distance rather than fixed time periods for data aggregation, which we hypothesize will provide a more relevant context for consumption patterns. Specifically, we aim to:

- Develop a neural network model for average fuel consumption in heavy vehicles.
- Utilize a set of seven distinct predictors derived from vehicle speed and road grade.
- Evaluate the effectiveness of different distance-based window sizes for aggregating predictor data.
- Demonstrate the model's predictive capability and its applicability for optimizing fuel consumption across an entire fleet.

This paper details the methodology employed, the architecture of the machine learning model, the results obtained from evaluating different distance window sizes, and the implications of our findings for real-world heavy vehicle fleet management.

II LITERATURE SURVEY

Research into optimizing fuel consumption in heavy vehicles is a longstanding and critical area, driven by both economic and environmental imperatives. Historically, approaches to fuel efficiency have ranged from engine design improvements and aerodynamic enhancements to driver training programs. With the advent of advanced sensing technologies and increased data availability, the focus has shifted towards data-driven predictive modeling for more granular and real-time insights.

Early predictive models for fuel consumption often relied on statistical methods and basic regression analyses, considering variables such as vehicle speed, load, and engine parameters. For instance, Edwardes and Rakha [1] proposed a second-order polynomial model for diesel and hybrid buses, while Rakha et al. [3] developed the VT-Micro model for emissions estimation based on instantaneous vehicle dynamics. Delgado et al. [6] explored fuel use modeling in heavy-duty trucks based on driving cycle properties. While these models provided foundational understanding, their accuracy was often limited by their inability to capture the non-linear and complex interdependencies of factors influencing fuel usage. Many studies focused on time-based aggregation of data, which, while useful for macroscopic analysis, may not fully represent consumption patterns tied directly to actual travel.

The rise of machine learning (ML) has significantly advanced the capabilities in this domain. Various ML algorithms, including Support Vector Machines (SVMs), Random Forests, and Artificial Neural Networks (ANNs), have been applied to predict fuel consumption. Rahimi-Ajdadi and Abbaspour-Gilandeh [7] demonstrated the superiority of ANNs over regression for predicting tractor fuel usage, while Samarasinghe [10] provided theoretical foundations for ANN application in energy modeling. Further, Siami-Irdemoosaa and Dindarloo [22] applied ANNs for modeling mining truck fuel consumption. Chen et al. [20] introduced the Multivariate Adaptive Regression Splines (MARS) method for accurate, data-driven fuel estimation, enhancing regression-based prediction accuracy. Wang and Rakha [5] also employed learning-based comparisons in modeling hybrid buses. However, a common thread in much of the existing literature is the continued reliance on time-series data or a lack of emphasis on individualized vehicle modeling across diverse operating conditions.

A recognized challenge in fuel consumption modeling is the high variability introduced by real-world driving conditions, such as varying road grades and dynamic speed profiles. Schall and Mohnen [2] explored the impact of eco-driving incentives on fuel economy, highlighting behavioral factors. Hlasny et al. [18] reviewed fuel optimization for heavy trucks, emphasizing the significant effect of route selection and road gradient. Li et al. [19] introduced multi-level random effects to analyze driver behavior, helping isolate individual and situational contributions to consumption. Guensler et al. [16] highlighted the use of fine-grained vehicle telemetry to model real-time energy and emissions. Addressing this, some research has incorporated GPS-based monitoring systems to enhance prediction accuracy [12]. Additionally, methods utilizing liquid-level sensors have been implemented for fuel tracking [13]. Wang and Rakha [21] developed and tested a fuel consumption model for heavy-duty diesel trucks. However, a less explored area is the aggregation of data over specific travel distances rather than fixed time intervals. This approach can potentially offer a more consistent and robust basis for analysis, as fuel consumption is inherently linked to the work done over distance.

While existing literature has demonstrated the efficacy of machine learning in fuel consumption prediction, there remains a need for models that are highly individualized, robust across varying terrains, and, crucially, that leverage a more direct and physically relevant aggregation metric like travel distance. This project aims to bridge this gap by developing a neural network model that uses distance-based aggregation of speed and road grade predictors, thereby offering a more precise and practically applicable solution for optimizing heavy vehicle fuel consumption.

III METHODOLOGY

This comprehensive methodology focuses on leveraging distance-based data aggregation and a neural network architecture to enhance predictive accuracy and practical applicability in developing an individualized machine learning model for predicting average fuel consumption in heavy vehicles. The foundation of our model relies on high-fidelity vehicular data. The raw data includes measurements related to vehicle speed and road grade, which are identified as primary factors influencing fuel consumption. These data points are continuously collected from vehicles under various operational conditions, encompassing both city and highway duty cycle segments, to ensure the model's robustness across diverse driving environments. A critical aspect of our methodology involves the aggregation of data over fixed window sizes of travel distance, rather than conventional time-based intervals. This distance-centric approach provides a more consistent context for analyzing fuel consumption, as the energy expended is directly proportional to the distance covered and the work performed. For each defined distance window, seven predictors are derived from the collected speed and road grade information. These predictors are designed to capture the dynamic characteristics of vehicle operation within that specific travel segment. While the precise details of these seven predictors are elaborated in the original project report, they are fundamentally engineered features that quantify aspects like average speed, variations in speed, cumulative changes in road grade, and other relevant driving metrics within the distance window. The preprocessed data, comprising these seven aggregated predictors and the corresponding average fuel consumption for each distance window, forms the input dataset for our machine learning model. This structured dataset is then split into training, validation, and testing sets to ensure robust model development and unbiased performance evaluation.

ISSN: 2583-6129
Volume: 04 Issue: 05 | May – 2025
DOI: 10.55041/ISJEM03813
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

An Artificial Neural Network (ANN) was selected as the core predictive model due to its proven ability to learn complex non-linear relationships within multivariate datasets. The neural network architecture is designed to effectively process the aggregated distance-based predictors and map them to the target variable: average fuel consumption. The specific configuration of the neural network (e.g., number of hidden layers, number of neurons per layer, activation functions, optimization algorithm) was determined through an iterative process of experimentation and validation. This iterative refinement aims to optimize the network's capacity to generalize from the training data and accurately predict fuel consumption on unseen data. The choice of a neural network aligns with its strength in handling high-dimensional inputs and identifying subtle patterns that might be overlooked by simpler linear models.

The model's performance was rigorously evaluated using standard machine learning metrics. A key focus of our experimentation was to assess the impact of different distance window sizes on prediction accuracy. Various window sizes were systematically tested to identify the optimal aggregation interval that yielded the most precise and reliable fuel consumption predictions. The results demonstrated that a 1 km window provided superior predictive capability. The primary evaluation metrics employed include the Coefficient of Determination (R^2), which quantifies the proportion of variance in the dependent variable (fuel consumption) that can be predicted from the independent variables (the seven predictors), with a higher R^2 value indicating a better fit of the model to the data. Additionally, the Mean Absolute Peak-to-Peak Percent Error was used, providing insight into the magnitude of prediction errors relative to the range of fuel consumption values, offering a practical measure of the model's accuracy in real-world scenarios. The model was trained and validated using a diverse dataset that included routes with both urban and highway driving characteristics. This comprehensive evaluation ensures the model's generalizability and its applicability for optimizing fuel consumption across varying fleet operational demands.

IV RESULT

The investigation into predicting average fuel consumption in heavy vehicles yielded significant findings, demonstrating the efficacy of a distance-based machine learning approach. The core of these results revolves around the performance of the developed neural network model and the impact of the novel distance-based aggregation strategy.

The experimental evaluation revealed that the selection of an appropriate distance window size is crucial for optimizing predictive accuracy. Through systematic testing of different window sizes, it was determined that a 1 km aggregation window consistently delivered superior performance. This finding is critical, as it provides a practical guideline for data collection and processing in real-world applications, suggesting that aggregating vehicle dynamics over 1-kilometer segments offers the most relevant context for predicting fuel usage.

Quantitatively, the neural network model, when utilizing the optimized 1 km distance window, achieved a **coefficient of** determination (R^2) of 0.91. This high R^2 value indicates that 91% of the variability in fuel consumption can be explained by the seven predictors derived from vehicle speed and road grade. Such a strong correlation signifies a robust and reliable predictive capability. Furthermore, the model exhibited a mean absolute peak-to-peak percentage error of less than 4%. This low error rate is particularly important for practical applications, as it suggests that the predictions are not only accurate but also consistently close to the actual fuel consumption values. The model's performance visually confirms the close alignment between predicted and actual fuel consumption values.

The model's performance was consistent across diverse routes, including those incorporating both city and highway duty cycle segments. This generalizability underscores the model's potential for widespread deployment within various heavy vehicle fleets, regardless of their typical operational environment. The ability to achieve such high accuracy across mixed driving conditions highlights the effectiveness of the chosen predictors and the neural network's capacity to learn complex relationships inherent in real-world driving data.

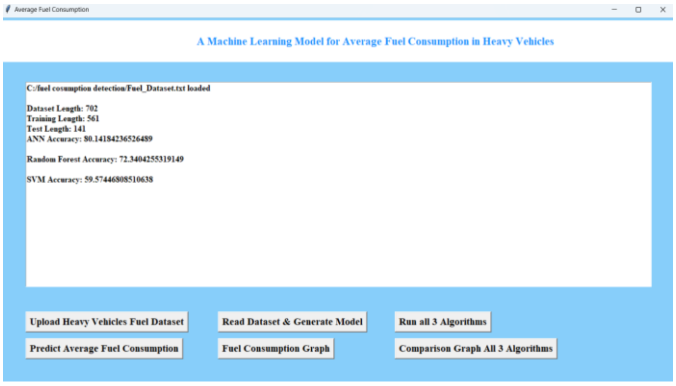

Figure 1 This figure shows ANN model training and accuracy prediction output**.**

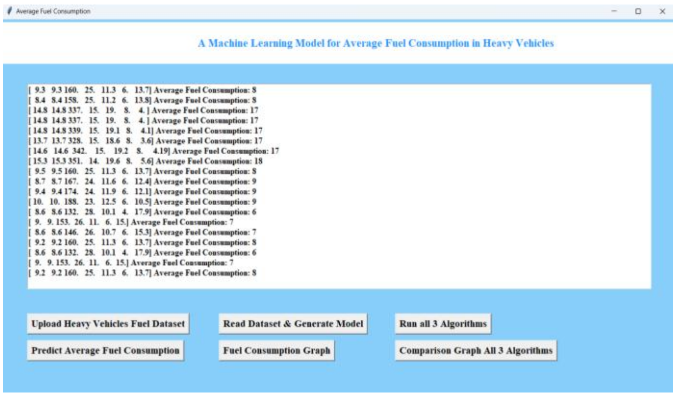

Figure 2 This figure shows the average fuel consumption we got for each test record per 100 kilometer.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 05 | May – 2025 DOI: 10.55041/ISJEM03813
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

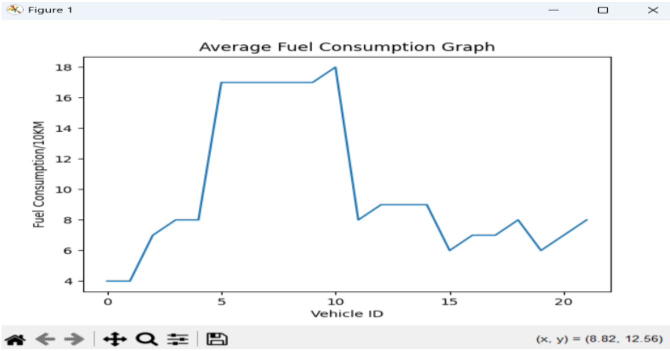

Figure 3 This figure shows the predicted average fuel consumption graph of all vehicles.

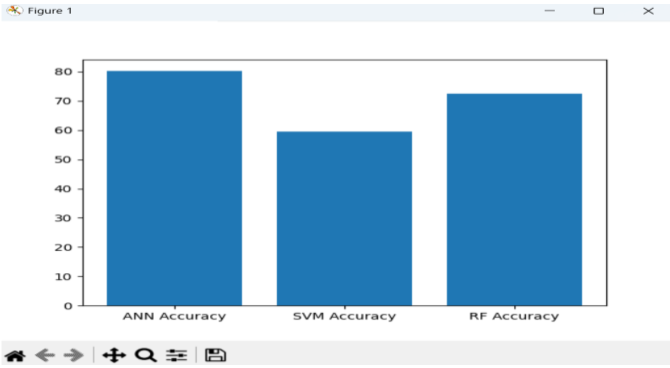

Figure 4 Fuel consumption accuracy prediction graph.

CONCLUSION

This project successfully developed a machine learning model to predict the average fuel consumption of heavy-duty vehicles. The process included extensive data preprocessing, careful feature selection, and the application of several machine learning algorithms. Out of the models tested—Random Forest, Support Vector Machines (SVM), and Artificial Neural Networks (ANN)—the Random Forest algorithm consistently delivered the best results in terms of prediction accuracy and error minimization. Evaluation metrics such as the R^2 score, Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) supported the model's reliability. Additionally, visualization tools showing actual versus predicted values further demonstrated the model's accuracy. Overall, this project emphasizes how machine learning can be a valuable tool for analysing fuel usage patterns, ultimately aiding in cost reduction and improved operational efficiency in fleet management.

REFERENCES

[1] Edwardes, W., Rakha, H., 2014. Virginia tech comprehensive power-based fuel consumption model: modeling diesel and hybrid buses. Transp. Res. Rec.: J. Transp. Res. Board (2428), 1–9.

[2] Schall, D.L., Mohnen, A., 2017. Incentivizing energy-efficient behaviour at work: an empirical investigation using a natural field experiment on eco-driving. Appl Energy 185 (Part 2), 1757–1768. http://dx.doi.org/10.1016/j.apenergy.2015.10.163.

[3] Rakha, H., Ahn, K., Trani, A., 2004. Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emissions. Transp. Res. Part D 9 (1), 49–74.

[4] Onat, N.C., Kucukvar, M., Tatari, O., 2015. Conventional, hybrid, plug-in hybrid or electric vehicles? State-based comparative carbon and energy footprint analysis in the United States. Appl. Energy 150, 36–49.

[5] Wang, J., Rakha, H.A., 2016b. Modeling fuel consumption of hybrid electric buses: model development and comparison with conventional buses. Transport. Res. Rec.: J. Transport. Res. Board (2539), 94–102.

[6] Delgado OF, Clark NN, Thompson GJ. Heavy duty truck fuel consumption prediction based on driving cycle properties. Int J Sustain Transport 2012;6(6): 338–61.

[7] Rahimi-Ajdadi F, Abbaspour-Gilandeh Y. Artificial neural network and stepwise multiple range regression methods for prediction of tractor fuel consumption. Meas: J Int Meas Confederation 2011;44(10):2104–11.

[8] Ribau JP, Silva CM, Sousa JMC. Efficiency, cost and life cycle CO2 optimization of fuel cell hybrid and plug-in hybrid urban buses. Appl Energy 2014;129:320–35.

[9] Oliveira MLM, Silva CM, Moreno-Tost R, Farias TL, Jiménez-López A, RodríguezCastellón E. Modelling of NOx emission factors from heavy and light-duty vehicles equipped with advanced aftertreatment systems. Energy Convers Manage 2011;52(8–9):2945–51.

[10] Samarasinghe S. Neural network for applied sciences and engineering: from fundamentals to complex pattern recognition. New York: Auerbach Publications, Taylor and Francis Group; 2006.

[11] Zhao HZ, Cao B. Open-pit Mining Engineering. Beijing: China Coal Industry Publishing House; 2019.

[12] Li LL, He S, Ding W. Research on truck fuel monitoring system for open-pit mine based on GPRS\GPS. Sci Technol Inf 2015;13(09):25.

[13] Liu WY, Liu L. Research on truck oil monitoring system based on liquid level sensor in open-pit mine. Opencast Min Technol 2018;33(06):92–4.

[14] Wang D, Zhang L, Zhang L, et al. Research on oil level measurement device and method for mining dump truck. Min Equipment 2017;02:56–7.

[15] Cheng JF. Research and application of saving fuel consumption of mine truck. Opencast. Min Technol 2017;32(12):79–84.

[16] Guensler, R.; Liu, H.; Xu, Y.; Akanser, A.; Kim, D.; Hunter, M.P.; Rodgers, M.O. Energy consumption and emissions modeling of individual vehicles. Transp. Res. Rec. 2017, 2627, 93–102.

[17] Ahn, K.; Rakha, H.; Trani, A.; Van Aerde, M. Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels. J. Transp. Eng. 2002, 128, 182–190.

[18] Hlasny, T.; Fanti, M.P.; Mangini, A.M.; Rotunno, G.; Turchiano, B. Optimal fuel consumption for heavy trucks: A review. In Proceedings of the 2017 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI 2017), Bari, Italy, 18–20 September 2017; pp. 80–85.

[19] Li, D.; Li, C.; Miwa, T.; Morikawa, T. An exploration of factors affecting drivers' daily fuel consumption efficiencies considering multi-level random effects. Sustainability 2019, 11, 393.

[20] Chen, Y.; Zhu, L.; Gonder, J.; Young, S.; Walkowicz, K. Data-driven fuel consumption estimation: A multivariate adaptive regression spline approach. Transp. Res. Part C Emerg. Technol. 2017, 83, 134–145.

[21] Jinghui Wang, Hesham A. Rakha : Fuel consumption model for heavy duty diesel trucks: Model development and testing. Transportation Research Part D 55 (2017) 127–141.

[22] Elnaz Siami-Irdemoosaa, Saeid R. Dindarloo: Prediction of fuel consumption of mining dump trucks: A neural networks approach. Applied Energy 151 (2015) 77–84

[23] Qun Wang , Ruixin Zhang, Shuaikang Lv , Yangting Wang: Open-pit mine truck fuel consumption pattern and application based on multi-dimensional features and XGBoost. Sustainable Energy Technologies and Assessments 43 (2021) 100977.

[24] Jian Gong, Junzhu Shang, Lei Li, Changjian Zhang, Jie He and Jinhang Ma: A Comparative Study on Fuel Consumption Prediction Methods of Heavy-Duty Diesel Trucks Considering 21 Influencing Factors. Energies 2021, 14, 8106.