Fuel Consumption Prediction Using Bayesian Neural Networks

Syarifah Diana Permai
Statistics Department, School of Computer Science, Bina Nusantara University

Jurike V Moniaga
Computer Science Department, School of Computer Science, Bina Nusantara University

Zener Sukra Lie
Automotive and Robotics Engineering Program, Computer Engineering Department, Faculty of Engineering, Bina Nusantara University

Ferry Jie
Markets and Services Research Centre, School of Business and Law, Edith Cowan University

https://doi.org/10.5109/7395669

出版情報：Proceedings of International Exchange and Innovation Conference on Engineering & Sciences (IEICES). 11, pp.1239-1245, 2025-10-30. International Exchange and Innovation Conference on Engineering & Sciences
バージョン：

Fuel Consumption Prediction Using Bayesian Neural Networks

Syarifah Diana Permai[1], Jurike V Moniaga[2], Zener Sukra Lie[3], Ferry Jie[4]
[1]Statistics Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480, [2]Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480, [3]Automotive and Robotics Engineering Program, Computer Engineering Department, Faculty of Engineering, Bina Nusantara University, Jakarta, Indonesia 11480, [4]Markets and Services Research Centre, School of Business and Law, Edith Cowan University, Joondalup, Australia 6027.
Corresponding author email: syarifah.permai@binus.ac.id

Abstract: *Transportation is one of the necessities of life. Because humans need transportation to move from one location to another. Transportation requires fuel. On the other hand, fuel consumption is important and must be controlled. This is because fuel can come from both renewable and non-renewable energy sources, depending on the type and process of its formation. Several factors influence the fuel efficiency of a car, including the type of engine, vehicle weight, aerodynamics, driving habits, and other vehicle conditions. This research aims to predict car fuel consumption and identify the factors that affect fuel consumption. Several Machine Learning and Statistical Learning methods were used to predict fuel consumption in this research. The best method for predicting fuel consumption is the Bayesian Neural Network because it produces the lowest error value compared to other methods.*

Keywords: Fuel Consumption; Energy; Machine Learning; Bayesian Neural Networks.

1. INTRODUCTION

Vehicles have become an important aspect of people's lives today. Due to the importance of vehicles, the options available today include public transportation and private vehicles. However, despite the congestion, private vehicles have become the top choice for many. One of the needs in a vehicle is fuel. Fluctuations in fuel prices are one of the problems because if fuel consumption increases, the community's need for fuel consumption also increases. Fuel prices also depend on fuel supply compared to fuel consumption [1]. In addition, the energy crisis due to the war between Russia and Ukraine, also soaring fuel demand due to the economy recovering after COVID-19 resulted in an increase in fuel prices [2].

Fuel consumption is one of the biggest costs in vehicle operation. Accurate prediction helps optimize spending. One of the benefits of fuel consumption prediction is energy efficiency. Because predicting fuel consumption can optimize energy use, reduce waste, and reduce operational costs. This is beneficial for producers and consumers. In addition, it can also help consumers choose vehicles that suit their needs based on fuel efficiency. Global emissions regulations encourage car manufacturers to produce more fuel-efficient vehicles. Several countries have established fuel efficiency standards as one of the most effective climate change mitigation measures [3].

More efficient fuel consumption means lower CO_2 emissions. Before the Industrial Revolution, carbon dioxide emissions were very low, but after 1950 carbon dioxide emissions continued to grow rapidly. Carbon dioxide emissions are one of the main causes of global climate change [4]. To energy efficiency, it also encourages to control of climate change to avoid the adverse impacts of extreme climate change. This supports the goals of Sustainable Development Goals (SDGs) 12, namely responsible consumption and production, and also SDGs 13, namely Climate Change.

Fuel can come from renewable or non-renewable energy sources. Non-renewable fuels come from natural sources that take millions of years to form and cannot be renewed in a short period, such as gasoline and diesel. Non-renewable fuels are still the main consumption and contribute to high greenhouse gas emissions. One way to replace fossil fuels is to use renewable energy fuels such as biomass [5]. However, it is still important to control fuel consumption.

In addition, the impact is carbon dioxide emissions. There is research linking between renewable energy consumption, non-renewables, CO_2 emissions, and economic growth. Economic activity and energy consumption have a major influence on the release of CO2 into the atmosphere [6]. The results of the research show a two-way causality between economic growth and renewable energy consumption, there is also a unidirectional causality between renewable energy and non-renewable energy consumption as well as renewable energy and CO_2 emissions [7].

Fuel consumption prediction not only saves costs, but also supports sustainability, regulatory compliance, and automotive innovation. Therefore, the right method is needed to predict fuel consumption. The statistical method is a simple and easy-to-implement method. Although linear regression is a traditional statistical method, some studies have shown that linear regression methods can predict better than machine learning methods such as Neural Network and Ensample methods [8].

However, along with the development of science, there are also Machine Learning methods that can produce better predictions. Therefore, in this research, modeling was carried out to predict fuel consumption using several methods such as Linear Regression, Decision Tree, Neural Network, and Bayesian Neural Networks. Furthermore, a comparison of prediction results is carried out using model evaluation on error values to produce the best model. In addition, the purpose of this research is to find out the factors that affect fuel consumption.

2. RELATED RESEARCH

Fuel Efficiency prediction has also been carried out by Vasmi, et al (2022) using machine learning methods including Linear Regression, Ridge, Lasso, KNN, XGBoost Regressor, Random Forest, Decision Tree, and Neural Network. The results show that the Decision Tree method produces the best predictions among other methods [9]. In addition, fuel prediction was also carried out by Manjunath and Ashok (2023) using the Linear Regression and Support Vector Regression (SVR) methods. In the SVR model, four different kernel functions are used, including linear, polynomial, radial base function, and sigmoid. The results show that the Linear Regression model is better than the SVR model on all types of kernel functions [10].

Fuel consumption prediction was also carried out by Hamed, Khafagy, and Badry (2021) using several machine learning methods including Decision Tree, Random Forest, Support Vector Machine (SVM) and Neural Network (NN). The results suggest that SVM can predict better than other methods [11]. In addition, Zhao et al (2023) also predicted Vehicle Fuel Consumption using SVM, Random Forest and Neural Network methods. The results showed that the Neural Network method provided better predictions than other methods [12].

3. DATA

The data used in this research was obtained from Kaggle [13]. The dataset provides a list of car specifications and is accompanied by fuel consumption for each type of car. There are several variables in the dataset. However, in this research, only 10 independent variables and one dependent variable were used. The total sample used in this research was 840 samples as a training data. The dependent variable in this research is fuel consumption per 100 km. The independent variables used are presented in Table 1.

Table 1. The Independent Variables.

Variable	Description	Type
Width	The width of the car (mm)	Numeric
Height	The height of the car (mm)	Numeric
Front track	The axle track of a vehicle (mm)	Numeric
Weight	The full weight of the car (kg)	Numeric
Torque	Maximum Torque	Numeric
CC	Capacity cm3	Numeric
Fuel tank capacity	Fuel tank capacity (liters)	Numeric
Hp	Horsepower	Numeric
Fuel grade	The grade of fuel used by the car (RON95, RON98, Diesel)	Categoric
Rear brakes	The rear brakes (Disk, Disk ceramic, Drum, Ventilated disc, Perforated)	Categoric

4. METHODOLOGY

The following are the stages of the research conducted in this research.

1. Data pre-processing. At this stage, data cleaning is carried out. There are some missing values in the dataset obtained. Then this condition must be overcome by deleting missing data.
2. Exploratory data analysis. At this stage, several graphs are made, including boxplots, histograms, and scatter plots for numerical data. As for categorical data, a bar chart is made. In addition, a correlation test was also carried out to determine whether there is a linear relationship between dependent variables and each independent variable.
3. Linear regression. At this stage, a Linear Regression model is made, then a parameter significance test is carried out to obtain factors that affect fuel consumption variables. Next, an assumption test was carried out on the model to find out whether the assumption of the linear regression model was fulfilled or not.
4. Decision Tree. At this stage, the optimal max depth parameters must be sought to obtain the best model in the Decision Tree method. In addition, feature importance was also obtained to find out the most important factors in predicting fuel consumption variables using the Decision Tree method.
5. Neural Networks. At this stage, the optimal parameters must be sought, namely the optimal number of nodes in the input layer and hidden layer.
6. Bayesian Neural Networks. At this stage, the optimal parameters must also be sought, namely the optimal number of nodes in the input layer and hidden layer.

5. RESULTS

The results of this research consist of exploratory data analysis and modeling. In exploratory data analysis, several graphs and correlation tests were made to determine the condition of the data. Meanwhile, in modeling, several models were built, including Linear Regression, Decision Tree, Neural Networks, and Bayesian Neural Networks.

5.1 Exploratory Data Analysis

Data exploration is carried out to identify the data. There are two types of data in this research, namely numerical data and categorical data. In numerical data, outlier identification was carried out using Boxplot. Meanwhile, in the categorical data, a bar chart is made for each variable.

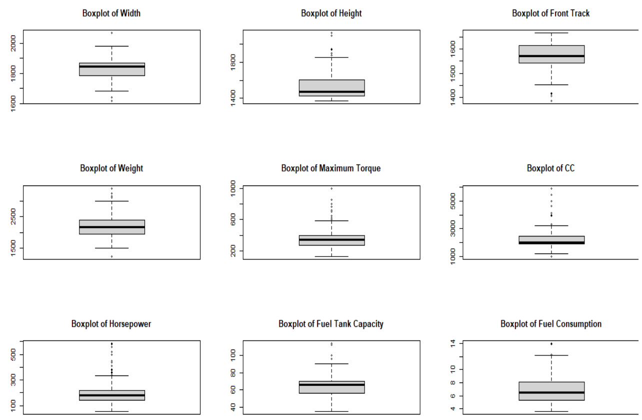

Fig. 1. Boxplot of the numerical variables

Fig.1 shows that all variables for numerical variables have outliers. In addition, the length of the whisker is also not the same and the median value is also not in the middle of the box. This suggests that the data is skewed and may not be normally distributed. Therefore, it is necessary to carry out modeling that can handle this problem in order to obtain good prediction results for fuel consumption.

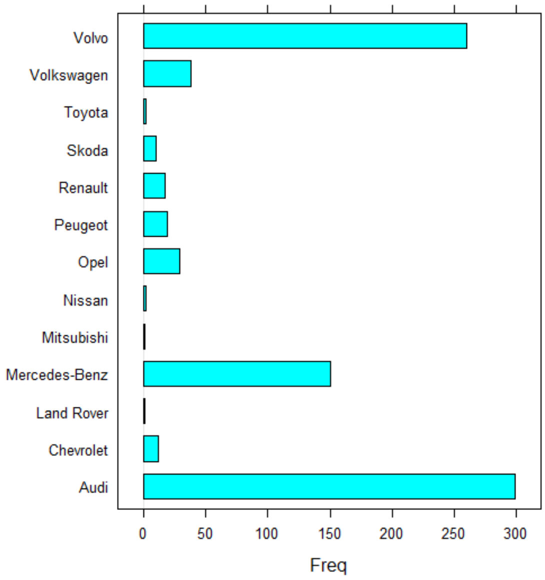

Fig. 2. Bar chart of the Brand variable

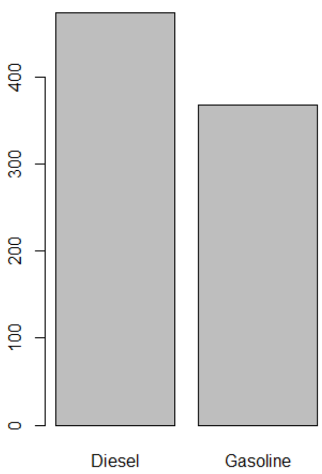

Fig. 3. Bar chart of the Engine Type variable

Fig. 2 shows the bar chart of the Brand variable used in this research. There are several car brands used, the three brands with the most frequency as samples are Audi, Volvo, and Mercedes-Benz. Besides that, there are also other brands such as Volkswagen, Toyota, Skoda, Renault, Peugeot, Opel, Nissan, Mitsubishi, Land Rover, and Chevrolet. While Fig. 3 shows the Engine type variable. There are two types of engines, namely Gasoline and Diesel. In this research, the sample used the most engine type is Diesel.

Furthermore, exploration was carried out using histograms, scatter plots, and correlation values on the numerical variables as presented in Fig.4.

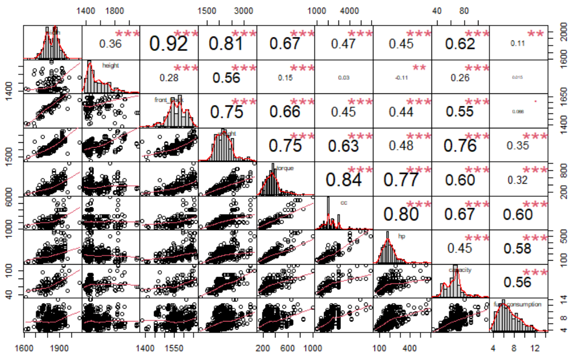

Fig. 4. Histograms, Scatter Plots, and Correlation values

Based on Fig.4, it can be seen the histogram for each numerical variable. Each variable shows a skewed pattern. This is in line with the results of the Boxplot which shows that there are outliers on all numerical variables. Meanwhile, in the scatter plot, it can be seen that there is a linear pattern between the dependent variable (fuel consumption) and all independent variables. This is in line with the high correlation value between dependent variables and each independent variable. The high correlation value between the fuel consumption variable and the independent variable was found in the cc, hp, and capacity variables where the values were 0.6, 0.58, and 0.56, respectively. Meanwhile, the lowest correlation is the correlation between the fuel consumption variable and the height variable. All correlations have positive values, this shows that if each independent variable increases, the fuel consumption variable will also increase.

However, to prove whether there is a linear relationship between dependent variables and all independent variables, it is necessary to use correlation tests and evaluate through p-values. Therefore, the results of the correlation test are presented in Fig.5 which shows the p-value of the correlation test.

	width	height	front_track	weight	torque	cc	hp	capacity	fuel_consumption
width	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
height	0.00	0.00	0.00	0.00	0.00	**0.38**	0.00	0.00	**0.66**
front_track	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.06
weight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
torque	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
cc	0.00	**0.38**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
hp	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
capacity	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
fuel_consumption	0.00	**0.66**	0.06	0.00	0.00	0.00	0.00	0.00	0.00

Fig. 5. The p-value of the correlation test

Based on Fig.5, the p-value in the correlation test results between the fuel consumption variable and each independent variable showed that the p-value was the highest of 0.66. This value is the result of a correlation test between the fuel consumption variable and the height variable. In addition, the p-value of the correlation test between the fuel consumption variable and the front track variable showed a value of 0.06. Both values are greater than 0.05. It can be concluded that the two variables do not have a linear relationship with the fuel consumption variable. Therefore, based on these results, feature selection can be carried out in the modeling to predict fuel consumption by not using height and front track variables. In addition, one of the things that can be used to perform feature selection is that both variables have a significant correlation with other independent variables. The highest correlation is found in the front track and width variables, which is 0.92 as presented in Fig. 4. Other models can also be used by involving all independent variables, because the results may show that variables that do not have a linear relationship affect fuel consumption. This is done to obtain better results in predicting variable fuel consumption.

5.2 Linear Regression

Table 2 shows the results of linear regression modeling using the variables that are significant in the model. Due to the results of parameter significance testing on all independent models, there are some insignificant variables in the model.

Table 2. Linear Regression Model

Variable	Estimate	P-value
Width	-0.0067140	2.18×10^{-5}
Height	0.0017860	1.11×10^{-6}
Front track	-0.0107504	3.27×10^{-9}
Weight	0.00182	2.43×10^{-10}
Torque	-0.0042549	2.12×10^{-10}
CC	0.0021521	2×10^{-16}
Fuel tank capacity	0.0635293	2×10^{-16}
Fuel grade: RON95	2.0156019	2×10^{-16}
Fuel grade: RON98	2.5787992	2×10^{-16}
Rear brakes: Disk ceramic	-2.8946626	0.001208
Rear brakes: Drum	-0.9453443	3.25×10^{-5}
Rear brakes: Ventilated disc	-0.3540932	0.000172
Rear brakes: Perforated	-4.6643475	2×10^{-16}

Based on Table 1, the coefficients in the regression model obtained positive and negative values. Independent variables that have positive values include height, full weight, capacity cm3 (cc), fuel tank capacity, and fuel grade. This shows that if these variables increase in value, fuel consumption will also increase. Meanwhile, the independent variables that have negative coefficients are width, front track, maximum torque, and rear brakes. It means that if these variables increase, fuel consumption will decrease. In addition, in the linear regression model, a p-value of less than 0.05 was obtained for all independent variables. This shows that these variables are significant in the linear regression model. These variables are width, height, front track, full weight, maximum torque, capacity cm3 (cc), fuel tank capacity, fuel grade, and rear brakes. It can be concluded that these independent variables affect fuel consumption. Based on the 10 independent variables used in this research, only nine variables affect fuel consumption. One variable that does not affect fuel consumption based on the linear regression model is the horsepower variable.

The linear regression model has the assumption that there is no multicollinearity in independent variables and the classical assumption is that it is residual independent, identical, and normally distributed. Therefore, it is necessary to check the assumptions in the linear regression model. Based on the results of the Variance Inflation Factors (VIF) test, the VIF value was obtained for the full weight variable of 10.2825, the width variable of 9.5272, the front track variable of 8.7268, the maximum torque variable of 6.8684, and the cc variable of 5.6979. Meanwhile, other independent variables have a VIF value of less than 5. This shows that there are independent variables that have a VIF value of more than 10. It can be concluded that multicollinearity occurs in independent variables.

Furthermore, the independent test was carried out using the Durbin-Watson (DW) test. The results of the DW test obtained a test statistical value of 1.0067 and a p-value of 2.2×10^{-16}. Since the p-value is less than 0.05, it can be concluded that there is an autocorrelation in the residual. This shows that the residual is not independent. Identical tests were conducted using the studentized Breusch-Pagan (BP) test. The results of the BP test obtained a test statistical value of 112.26 and a p-value of 2.2×10^{-16}. It can be concluded that heteroscedasticity occurs in residuals because the p-value is less than 0.05. This shows that the residual is not identical. A normally distributed residual test is performed using the Kolmogorov-Smirnov (KS) test. The results of the KS test obtained a test statistical value of 0.056293 and a p-value of 1.313×10^{-6}. The null hypothesis is residual has normally distributed and the alternative hypothesis is residual has not normally distributed. The H_0 is rejected because the p-value is smaller than 0.05. This shows that the residual is not normally distributed. Based on the results of the test of these assumptions, the conclusion was obtained that all assumptions were not met. Therefore, the linear regression model obtained is still not good enough in predicting fuel consumption. In this research, modeling was carried out using machine learning as an alternative to statistical learning.

5.3 Decision Tree

One of the machine learning methods used in this research is the Decision Tree. In the Decision Tree model, 10 independent variables are used as inputs for the model. The optimum model obtained is a model with a max depth of 30. Because if the input is less than 10 and the max depth is less than 30, bad results are obtained based on the evaluation of the error value in the model. Then, it can be concluded that the optimal number of inputs is 10 and the optimal number of max depths is 30. The results of feature importance analysis in the Decision Tree model were obtained as shown in Fig. 6.

Fig.6 shows the magnitude of the importance level of each variable based on the Decision Tree method. In the Decision Tree model, it was concluded that the most important variable in the model was capacity cm3 (cc). The CC variable has the highest importance level value of 1778.45758. The second position is the horsepower variable which is also important in the Decision Tree model. Because the horsepower variable has an importance level of 1008.40773. This is in contrast to the results of the linear regression model, where the horsepower variable is not significant in the model. Meanwhile, the two variables that are not very important in the Decision Tree model are the height and rear brakes variables which have feature importance values of 174.20257 and 21.44133, respectively. These two variables have the lowest importance value among other independent variables. Other independent variables are also important in the model because all other variables have a value of importance level above 500.

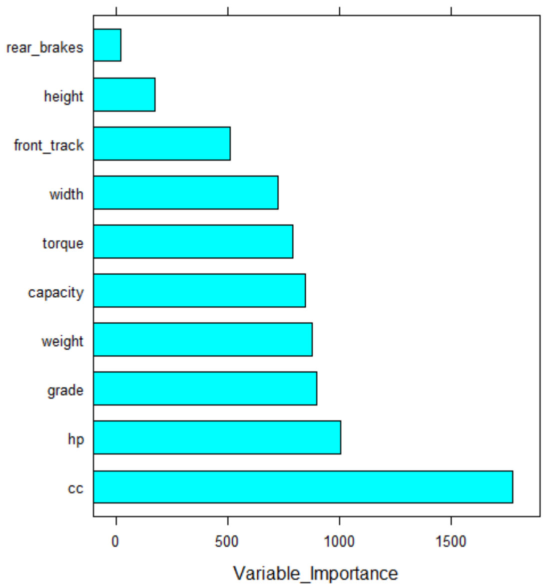

Fig. 6. The feature importance analysis of the Decision Tree

5.4 Neural Networks

Moreover, machine learning modeling was carried out using the Neural Networks (NN) method. In this method, the optimal number of nodes must be found in the input layer and hidden layer. Therefore, several Neural Network models were built and the optimal model was obtained, namely a model with the number of nodes on the input layer of 8 and the number of nodes on the hidden layer of 5. The following is Fig. 7 of the results of the NN model formed.

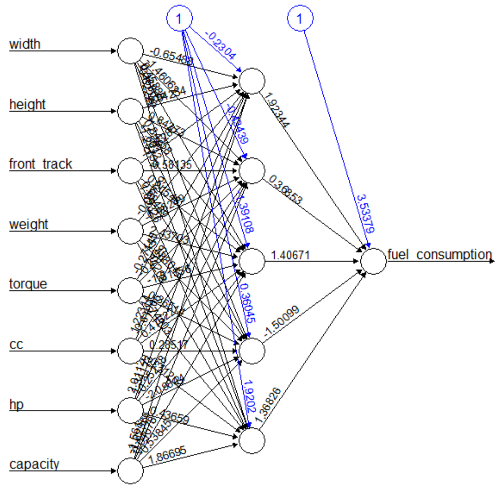

Fig. 7. The Result of the Neural Networks Model

Fig. 7 shows the results of the Neural Networks model with 8 variables in the input layer and 5 nodes in the hidden layer. In Fig. 7, it can be seen the weight of each node from the input layer to the hidden layer, and the weight of each node in the hidden layer to the output layer. Nodes with the number 1 on the hidden layer and the output layer indicate a bias or constant value in the Neural Networks model. The number of variables on the input layer as well as the number of nodes in the hidden layer will be used as a reference for the Bayesian Neural Networks model.

5.5 Bayesian Neural Networks

In the Bayesian Neural Networks method, 8 variables are used as inputs and the optimal number of nodes in the hidden layer is searched. A comparison of results was carried out for several nodes. Table 3 shows the results of a comparison of the Bayesian Neural Networks model.

Table 3. Bayesian Neural Networks Model

Number of Nodes	RMSE	MAE	MAPE
5	0.68442	0.51510	7.8786
10	0.55492	0.41499	6.3759
20	0.42523	0.31969	4.9583
22	**0.41311**	**0.31164**	**4.8478**
25	0.51308	0.38085	5.9089
30	0.44679	0.33696	5.2294

There are several number of nodes that are tried in the hidden layer, ranging from 5 nodes to 30 nodes. The results of the Bayesian Neural Networks model comparison obtained that the optimal number of nodes for the hidden layer is 22. It can be seen in Table 3 that the smallest RMSE, MAE, and MAPE values are found in the number of nodes 22 in the hidden layer.

5.6 Comparison

Based on the results of modeling that have been built to predict fuel consumption, a comparison of all the models that are formed is carried out. The results of the model comparison used RMSE, MAE, and MAPE values. This criterion uses an error value, so a good model is the one with the least value. The results of the comparison of all models are presented in Table 4.

Table 4. Comparison of Fuel Consumption Prediction Models

Method	RMSE	MAE	MAPE
Linear Regression	0.84302	0.63863	9.4327
Decision Tree	0.81474	0.63816	9.5822
Neural Networks	1.97375	1.60243	24.847
Bayesian Neural Networks	**0.41311**	**0.31164**	**4.8478**

The linear regression model is not good enough because the model does not meet the assumptions. Hence the model cannot be used. The machine learning methods used, namely the Decision Tree, Neural Networks, and Bayesian Neural Networks models. The model with the highest RMSE, MAE, and MAPE values is the Neural Networks model. It was concluded that the best model is the Bayesian Neural Networks model. This is because the BNN model has relatively small RMSE, MAE, and MAPE values compared to other models. Therefore, the BNN model can be used to predict fuel consumption.

6. CONCLUSION

Fuel consumption prediction is important because it can help the country's economy. In addition, minimal fuel consumption can help to reduce carbon dioxide emissions and the greenhouse gas effect. Therefore, this research contributes to helping control energy efficiency and climate change. In addition, the development of machine learning can also be applied in predicting fuel consumption. In this study, 10 independent variables were used to predict fuel consumption. All numeric variables have outliers and skewed data. The results of the correlation test showed that there was a correlation between variables. However, the height and front track variables do not have a linear relationship with the fuel consumption variable.

Based on the linear regression model, it was concluded that there are 9 independent variables that affect fuel consumption. The variable that does not affect fuel consumption is the horsepower variable. However, the resulting linear regression model is not good enough because it does not meet the assumptions. However, based on the results of feature importance in the decision tree model, it is concluded that all important variables in the model. However, there are two variables that have a low influence on fuel consumption, namely the height variable and rear brakes.

Besides that, the results from the Neural Networks model show that there are 8 variables that are better used as inputs to predict fuel consumption. However, the NN model has the highest RMSE, MAE, and MAPE values among other models. This shows that the NN model is still not optimal and can be further developed. In addition, you can also use Deep Learning to get even better results. The best model for predicting fuel consumption is the Bayesian Neural Networks model. Because this model produces the smallest RMSE, MAE, and MAPE values compared to other models.

7. ACKNOWLEDGEMENT

This research is fully funded by Bina Nusantara University. This paper is an output of Penelitian International Binus with contract number 085/VRRTT/V/2025. We would like to express our gratitude to Bina Nusantara University for the support of our research.

8. REFERENCES

[1] EIA, "Gasoline explained: Gasoline price fluctuations," U.S. Energy Information Administration, 6 September 2024. [Online]. Available: https://www.eia.gov/energyexplained/gasoline/price-fluctuations.php.

[2] E. Akcura and F. C. Bloch, "The high price of cheap fuel: What new global data reveals about liquid fuel subsidies," World Bank, 19 May 2025. [Online]. Available: https://blogs.worldbank.org/en/opendata/the-high-price-of-cheap-fuel--what-new-global-datareveals-about.

[3] Z. Yang, "Passenger vehicle greenhouse gas emissions and fuel consumption," International Council on Clean Transportation, November 2024. [Online]. Available: https://theicct.org/pv-fuel-economy.

[4] H. Ritchie and M. Roser, "CO_2 emissions," Our World in Data, June 2020. [Online]. Available: https://ourworldindata.org/co2-emissions.

[5] S. C. Karmaker, S. Hosan, M. M. Rahman, K. K. Sen and B. B. Saha, "Dynamic Linkage Between Biomass Energy Consumption and Ecological Footprint: A Panel Analysis for BRICS Countries," in *Proceeding of International Exchange and Innovation Conference on Engineering & Sciences (IEICES)*, Fukuoka, 2021.

[6] J. Rahman, G. I. Wahid, S. Ahmed, F. M. N. Islam, B. B. Saha and S. Ghosh, "Nexus Between CO2 Emission, Economic Development and Energy Consumption in Bangladesh," in *Proceeding of International Exchange and Innovation Conference on Engineering & Sciences (IEICES)*, Fukuoka, 2023.

[7] B. A. Asiedu, A. A. Hassan and M. A. Bein, "Renewable energy, non-renewable energy, and economic growth: evidence from 26 European countries," *Environmental Science and Pollution Research,* vol. 28, no. 9, pp. 11119 - 11128, 2021.

[8] J. Galupino and J. Dungca, "Estimation of Permeability of Soil-Fly Ash Mix using Machine Learning Algorithms," in *Proceeding of International Exchange and Innovation Conference on Engineering & Sciences (IEICES)*, Fukuoka, 2023.

[9] M. A. Vamsi, B. R. Rishita, M. A. Qurishi, N. S. Sandeep and K. R. Ram, "Fuel Efficiency Prediction," *Journal of Emerging Technologies and Innovative Research,* vol. 9, no. 4, 2022.

[10] T. K. Manjunath and K. P. Ashok , "Fuel Prediction Model for Driving Patterns Using Machine Learning Techniques," *Journal of Computer Science,* vol. 20, no. 3, pp. 291 - 302, 2023.

[11] M. A. Hamed, M. H. Khafagy and R. M. Badry, "Fuel Consumption Prediction Model using Machine Learning," *International Journal of Advanced Computer Science and Applications (IJACSA),* vol. 12, no. 11, pp. 406 - 414, 2021.

[12] D. Zhao, H. Li, J. Hou, P. Gong, Y. Zhong, W. He and Z. Fu, "A Review of the Data-Driven Prediction Method of Vehicle Fuel Consumption," *Energies,* vol. 16, no. 5258, pp. 1 - 20, 2023.

[13] J. Islam, "Car Specification Dataset 1945-2020," Kaggle, 25February2023.[Online].Available: https://www.kaggle.com/datasets/jahaidulislam/car-specification-dataset-1945-2020