Fuel Consumption Prediction of Vehicles Using Machine Learning Algorithm

[1]Oladejo Olubusayo, [2]Ozoh Patrick, [3*] Ibrahim Musibau, [4]Adigun Adepeju, [5]Oyinloye Olufunke, [6]Dimple Ariyo, [7]Ojo Oluwafolake and [8]Abanikannda Mutahir

[1]Department of Physics, Osun State University, Osogbo, Nigeria
[2,3,4,5,6,7,8]Faculty of Computer Science and Information Technology, Osun State University, Osogbo, Nigeria

email: [3*]kunle_ibrahim2001@yahoo.com

**Corresponding author*

Received: 15 November 2024 | Accepted: 24 April 2025 | Early access: 19 May 2025

Abstract - *The major objective of this study revolves around accurately estimating the distance a car can travel in kilometres per Litter of fuel consumed. This study develops a precise machine-learning algorithm for predicting vehicle petroleum consumption. This encompasses adapting distinct machine-learning techniques, evaluating their performance, selecting the most optimal model, and validating its real-world applicability. The dataset used for this study includes attributes such as Miles per gallon (Mpg), acceleration, horsepower, displacement, cylinder count, and car model. The implementation strategy entails comprehensive data pre-processing and employing well-established machine learning techniques: Random Forest, Decision Tree, and Linear Regression. The Python programming environment is applied for coding and data manipulation. Model performance assessment uses the Mean Squared Error (MSE) metric. The findings show the performance of the Random Forest algorithm as having the lowest MSE value of 0.008806 among the assessed models. In conclusion, the proficiency of the Random Forest algorithm. in predicting fuel consumption will open avenues for informed decision-making and resource optimization within the automotive sector.*

Keywords: Tracking, estimation, machine learning, model performance, decision making, optimization.

1 Introduction

John McCarthy coined the term machine learning in 1955. He defined it as a branch of engineering. However, it includes experience and quick responses to questions (Gupta et al., 2024). Machine learning can also be defined as that which enables computers to learn (Samuel et al., 2022). This research uses the most effective method for predicting vehicle fuel consumption. It is important to understand the factors that affect fuel consumption and models that can accurately predict fuel consumption (Ashqar et al., 2024). The automobile industry is one of the fastest-growing industries in the world. With the increasing number of vehicles on the road, there is also an increase in fuel consumption. The understanding of factors that affect the models that can accurately predict fuel consumption.

This study consists of five sections. Section one contains the background and contribution to knowledge. Section two includes a review of fundamental concepts and related works. Section three showcases the presentation of the methodology and the techniques, technologies, and tools to be used. Section four presents the results obtained. Section five includes the conclusions and recommendations of this study.

The contributions are

(1) To adapt Random Forest, Decision Tree, and Linear Regression techniques to prediction
(2) Evolve algorithm that accurately predicts fuel consumption in vehicles
(3) Validate the prediction model with the best performance.

2 Literature review

This section provides an in-depth review of the relevant literature about the current study. The review encompasses various aspects, including machine learning techniques in predicting fuel consumption, the methodologies adopted, and an overview of pertinent works in the field. Fuel consumption prediction is vital for optimizing energy use and minimizing environmental impact in transportation and energy systems. It helps us understand consumption patterns, make efficient energy decisions, and develop strategies. Recent progress, driven by data-driven methods and technology, has improved fuel consumption prediction. These predictive models, from single vehicles to entire cities, consider factors like vehicle type, driving conditions, and weather (Su et al., 2023).

Machine learning drives these models, analysing past and real-time data for accurate predictions. This has wide-ranging benefits, reducing costs, and carbon footprints, and enhancing sustainability. The progress in this field also contributes to our understanding of energy dynamics and supports the shift to sustainable energy. By combining machine learning and data analysis, researchers are revolutionizing fuel management for a greener future (Smith, 2022). Chen et al. (2021) introduced machine learning for predicting fuel consumption for spark-ignition engine vehicles based on real-world driving conditions. A set of driving data was collected from an on-board diagnostic (OBD) system, pre-processed data, and the most relevant features using a correlation-based feature selection method. Then, three machine learning algorithms were used to predict fuel consumption: random forest, support vector regression, and multilayer perceptron. The models were evaluated using three metrics: MAE, RMSE, and coefficient of determination. The results showed that the random forest model outperformed the other two models, and the feature importance analysis identified engine load, vehicle speed, and engine coolant temperature as the most critical features for fuel consumption prediction. The authors concluded that their machine learning approach is an effective method for fuel consumption prediction of spark-ignition engine vehicles, which can lead to economic and environmental benefits.

The purpose of Xu et al. (2021) is to develop a general research guideline on structural sciences with the current researchers to have greater knowledge to understand the importance of the basic sciences. Yang et al. (2024) improve the accuracy of model prediction and describe the basis for controlling consumption. Manivannan (2024) applies machine learning to produce a smart energy system for electrical vehicles using optimization methods to reduce costs and time. Katyare et al. (2024) investigate issues in forecasting different levels of systems, with a focus on accuracy. as accurate predictions are important for improving efficiency.

The literature review of the different prediction models is given as follows.

2.1 ANN

Li et al. (2020) demonstrated that machine-learning models predict urban buses. They collected data from a real-world driving cycle and applied four models: decision tree, RF technique, and artificial networks (ANN). They investigated the performance of the accuracy. The three models identified the most critical factors. The study demonstrated that machine learning models can effectively predict urban buses optimize the performance of urban buses, reduce fuel consumption, and minimize carbon emissions.

Tang et al. (2022) developed a prediction model for a heavy-duty truck using artificial neural networks (ANNs). The authors collected data on various driving conditions, including vehicle speed, acceleration, engine load, and other parameters, from an on-board diagnostic (OBD) system installed in the truck. The collected data was pre-processed, and the features were selected using a correlation-based feature selection method. The authors then developed an ANN model for the truck's features. The performance was evaluated using two metrics: mean absolute error (MAE) and coefficient of determination (R^2). The accuracy in predicting fuel consumption has an MAE of 0.3 km/L and R^2 of 0.96. The authors concluded that their ANN-based fuel consumption prediction model can be used to optimize the control strategies of heavy-duty trucks and improve their fuel efficiency, which can have significant economic and environmental benefits.

2.2 Random forest

Wu et al. (2024) state that the random forest method for monitoring malware is an improvement in the application of machine learning. The technique consists of obtaining different sets of datasets. The evaluation for the estimation of accuracy is greater than previous techniques. Noviyanti et al. (2023) provide an increase in the reliability of fast monitoring of diabetes using the Random Forest algorithm model. The study was done with the aid of data collection, data pre-processing, modelling, and evaluation.

2.3 Decision tree

Blockeel et al. (2023) present the importance of decision trees in machine learning. The study investigates changes in the study of decision trees over time. It provides the strong characteristics and shortcomings of decision trees. Costa and Pedreira (2023) present an in-depth of the major progress made in decision tree research, in terms of training data, modeling, and interpretation.

2.4 Linear regression models

To eradicate imbalanced datasets in regression, a regression method is provided together with the Gaussian technique (Wen et al., 2024). These methods involve the gradient-boosting decision tree (GBDT), random forest (RF), and extreme gradient boosting (XGBoost) models. As a result, the GBDT model possesses the highest accuracy with R2 = 0.95, MAPE = 29.8%, and evaluation results (n = 3214, R2 = 0.84, MAPE = 38.8%). Aziz et al. (2024) present a model to monitor information from the data sets. The proposed technique is useful for reducing crime. The methods involve the use of several regression models that are developed depending on the regression techniques, decision tree regression (DTR), simple linear regression (SLR), and support vector regression (SVR). The study was pre-processed using MySQL and R programming.

3 Methodology

This section discusses the study methods used in predicting fuel consumption using machine learning. It provides details of data collection, data pre-processing, and the models used.

3.1 Data collection

This section outlines the methodology adopted to achieve the research goals and objectives. The initial phase involves data collection, with the dataset sourced from Kaggle, providing comprehensive information about fuel consumption. This data set encompasses attributes such as mpg) acceleration, horsepower, displacement, cylinder count, weight, and car model. The data size used consists of 18 vehicles of different models. The dataset contains 9 columns while the target variable is the MPG (miles per gallon) as the target variable in research work represents the primary focus of modeling efforts. For training purposes, the pre-processed dataset is divided into the experimental vehicle and the vehicle of other models. These are converted into numerical values: 0 for experimental cars and 1 for different vehicle models. This study uses 70% of the data: training, and 30%: testing. The fuel consumption dataset is shown in Figure 1.

	A	B	C	D	E	F	G	H	I
1	Mcg	Cylinders	Displacement	Horsepower	Weight	Acceleration	Model year	Origin	Car Model
2	18	8	307	130	3504	12	70	1	Chevrolet chevelle malbu
3	15	8	350	165	3693	11.5	70	1	Buck skylark 320
4	18	8	318	150	3436	11	70	1	Plymouth satelite
5	16	8	304	150	3433	12	70	1	Amc rebel sst
6	17	8	302	140	3449	10.5	70	1	Ford torino
7	15	8	429	198	4341	10	70	1	Ford galaxe 500
8	14	8	454	220	4354	9	70	1	Chevrolet impala
9	14	8	440	215	4312	8.5	70	1	Plymouth fury ii
10	14	8	455	225	4425	10	70	1	Pontiac cataina
11	15	8	390	190	3850	8.5	70	1	Amc ambassador dol
12	15	8	383	170	3563	10	70	1	Dodge challenger se
13	14	8	340	160	3609	8	70	1	Plymouth cuda 340
14	15	8	400	150	3761	9.5	70	1	Chevrolet monte carlo
15	14	8	455	225	3086	10	70	1	Buck estate wagon (sw)
16	24	4	113	95	2372	15	70	3	Toyota corona mark ii
17	22	6	198	95	2833	15.5	70	1	Plymouth duster
18	18	6	199	97	2774	15.5	70	1	Amc hornet
19	21	6	200	85	2587	16	70	1	Ford maverick
20									
21									

Sheet1

Figure 1: Fuel consumption dataset

3.2 Data pre-processing

This step involves cleaning and filtering to achieve further analysis. The technique used for the pre-processing of the data is the One Hot encoding technique. One-hot encoding is a technique used in data pre-processing to convert categorical variables into a binary matrix format. It's commonly applied in machine learning and data analysis when dealing with categorical data that cannot be directly used by algorithms that expect numerical input. In one-hot encoding:

1. Each unique category in the categorical variable.

2. For each observation (row) in the dataset, only one element in the binary vector is 1, indicating the presence of that category, while the other elements are 0, indicating absence.

For the above dataset, one hot encoding was used to convert the cylinder attribute from being a categorical variable to Binary format (0s and 1s). One-hot encoding enables machine learning algorithms to effectively work with categorical data, which are often nominal or ordinal. It prevents the algorithm from assuming any ordinal relationship between categories and treats each category as a distinct and independent feature to capture categorical information without introducing unintended biases. The data pre-processing techniques performed on the cylinder attributes are shown in Figure 2.

```
In [21]: new_df["cylinders"]=new_df["cylinders"].astype(str)
         new_df["origin"]=new_df["origin"].astype(str)
         new_df=pd.get_dummies(new_df)
```

Figure 2: A snippet of the encoding technique performed on the cylinder attribute

3.3 Model description

The models used in this study are described in this section. This involves the process of choosing the most appropriate machine learning model architecture it involves selecting a model that is well-suited to the data and provides accurate and reliable predictions.

3.3.1 Random forest

Random Forest is a powerful supervised learning technique for regression and classification tasks. The classifier is an ensemble algorithm that builds multiple decision trees and combines them to produce a more effective classifier. The Random Forest algorithm pseudocode is shown in the following steps.

Step 1: Select randomly M features from the feature set

Step 2: For each x in M

a. Calculate the Information Gain
   Gain (t, x) = E (t) – E (t, x)
   $E(t) = \sum c \ - p\$ \log 2 \ p\$$
   $E\ (t, x) = \sum c \in K\ P(c)E(c)$
   Where E(t) is the entropy of the two classes, E(t,x) is the entropy of feature x
b. Select the node d
c. Split the node into sub-nodes
d. Repeat steps a, b, and c to construct the tree until it reaches the minimum number of samples to split.

Step 3: Repeat steps 1 and 2 N for times to build a forest of N trees

The structure of the vehicle fuel consumption prediction based on the random forest is given in Figure 3.

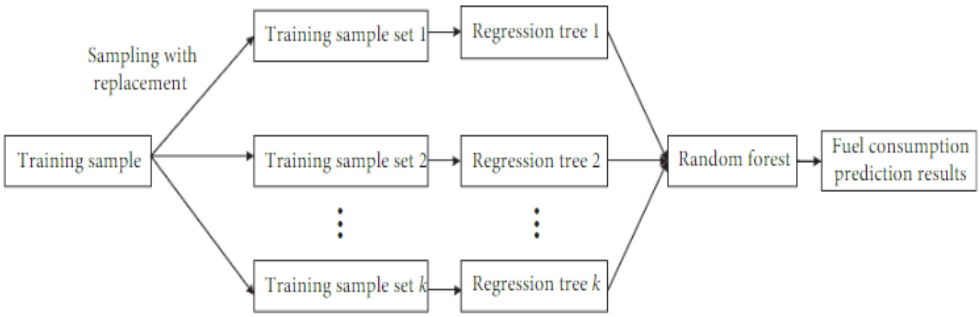


Figure 3: The structure of a vehicle fuel consumption prediction based on the random forest

3.3.2 Linear regression

Linear regression, a form of supervised machine learning, establishes a linear connection between a dependent variable and one or more independent features as shown below:

Given

$y_i \in Y\ (i = 1, 2, \cdots, n)$ are labels to data (Supervised learning)
$x_i \in X\ (i = 1, 2, \cdots, n)$ are the input independent training data (univariate – one input variable/parameter)
$\hat{y}_i \in \hat{Y}\ (i = 1, 2, \cdots, n)$ are the predicted values, with $\hat{Y} = \theta_1 + \theta_2 X$ where $\hat{y}_i = \theta_1 + \theta_2 x_i$ and θ_1, θ_2 are the intercept and slope, respectively.

3.3.3 The DT algorithm

The steps for a decision tree algorithm are given as follows.

1. Start.
2. For every iteration, compute the Entropy(H) and Information gain (IG) of this attribute.
3. Select the attribute with the smallest Entropy or Largest Information gain.
4. Get S.
5. The algorithm begins processing.

3.4 Model evaluation

This section delves into the techniques employed to assess the performance of the algorithms utilized for the prediction model. In the context of this project, the evaluation method is the Mean Square Error (MSE). This is given as follows.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(mpg_i - \widehat{mpg}_i)^2$$

where n represents the total number of instances, mpg_i is the actual "mpg" value for the i-th instance and $\widehat{mpg}_i$ is the predicted "mpg" value for the i-th instance.

4 Results and discussion

This section presents the results of the implementation performed in this study and its evaluation metrics employed to validate the result, and performance of the model. It also provides a detailed discussion of the results and findings of the model.

4.1 Model implementation

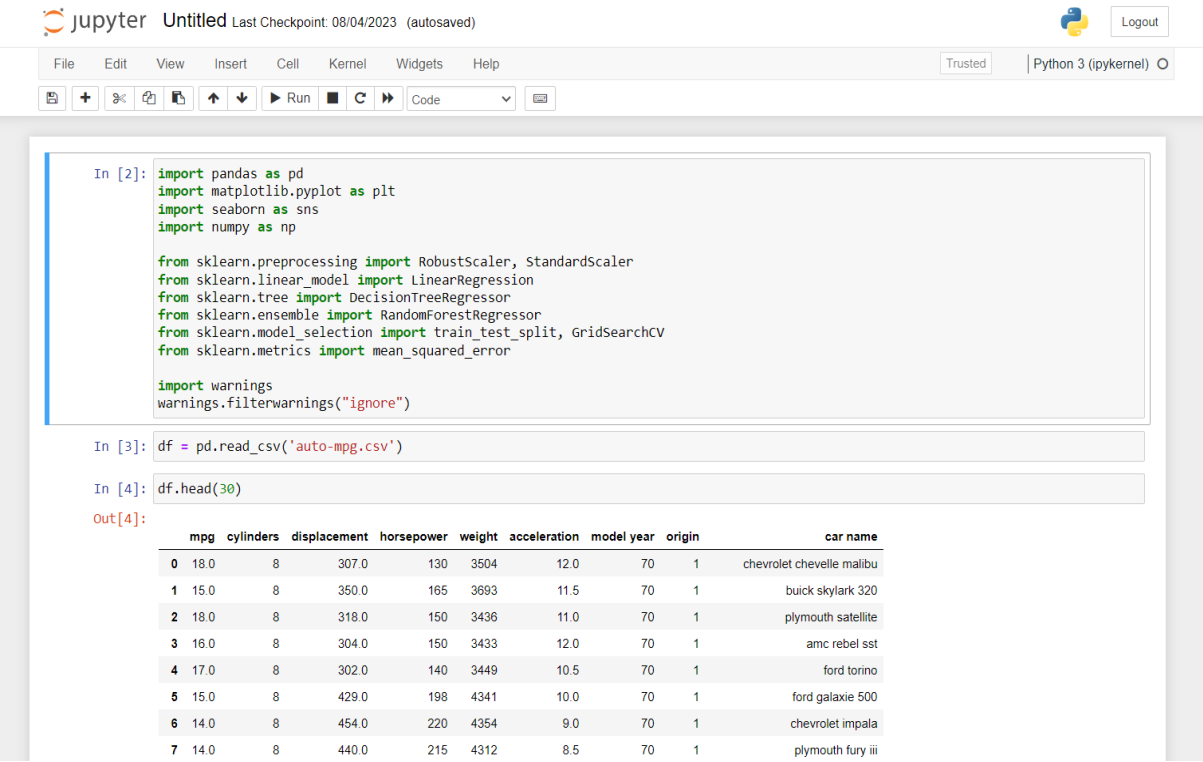
In this study, the model was implemented using hardware and software tools.

4.1.1 Hardware tools

The hardware tool used for this study is a computer system running Windows operating system it serves as the host for the other hardware tools this research requires one computer system.

4.1.2 Implementation of the study

The software used for implementing this study is the Python programming language. Python is good for data mining, analysis, and Machine Learning. The Jupyter Notebook is the version of the Python programming language used. It is an online intelligent computational environment-based software for creating notebook documents. This software allows users to create and share documents with live code, equations, visuals, and narrative text with the open-source Jupyter Notebook program. Figure 4 shows the comparison between the sample data and the corresponding predicted values.



	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin	car name
0	18.0	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165	3693	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150	3436	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150	3433	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140	3449	10.5	70	1	ford torino
5	15.0	8	429.0	198	4341	10.0	70	1	ford galaxie 500
6	14.0	8	454.0	220	4354	9.0	70	1	chevrolet impala
7	14.0	8	440.0	215	4312	8.5	70	1	plymouth fury iii

Figure 4: Results using Jupyter Notebook environment

For the three models that were used for this study, Random Forest has the least MSE (mean square error) value. Figure 5 shows the MSE value for the three models. Random Forest has the least (MSE) value 0.008806, Linear Regression: 0.010937, and Decision Tree: 0.015844 respectively, and the lower value indicates a better fit, hence making the Random Forest the most efficient model. Furthermore, during the development of the model, a comparison was made between a sample with the corresponding predicted data. This comparison served as a validation technique to assess the performance of the trained model. The agreement between the predicted and actual data provided insights into the accuracy and reliability of the model's prediction. This is shown in Figure 5. Figure 5 shows the visualizations for linear regression, decision trees, and the Random Forest technique. The figure shows that the Random Forest technique has the lowest matric value when the three methods are compared. This indicates that Random Forest is the most efficient method.

Out[107]:

	Regression Model	Metrics Result
0	Linear Regression	0.010937
1	Decision Tree	0.015844
2	Random Forest	0.008806

In [103]:
```
import matplotlib.pyplot as plt
```

In [126]:
```
plt.bar(pt['Regression Model'],pt['Metrics Result'], color= ["r",'g','b'])
plt.title("Result Visualization",size=22)
plt.xlabel("Model", size=18)
plt.ylabel("mse value", size=18)

plt.show()
```

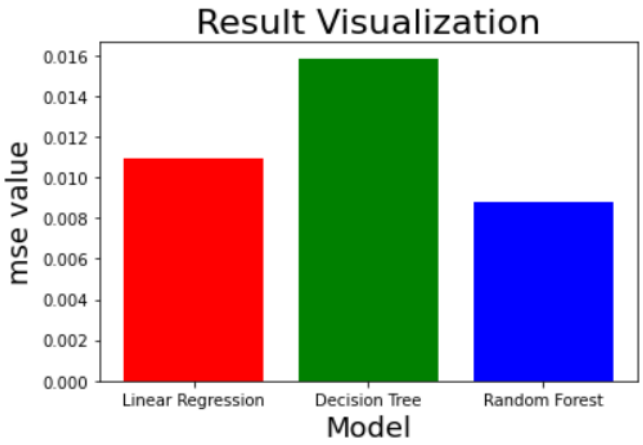


Figure 5: Result visualization of the metric value for the three model

4.2 Evaluation of results

According to Figures 4 and 5, the Random Forest technique is the best-performing technique. To determine its accuracy, it is compared with Eyring et al. (2024) and Asnicar et al. (2024). The three methods were tested on 23 datasets. Table 1 gives the actual data and their respective estimates using Eyring et al. (2024), Asnicar et al. (2024), and the Random Forest technique.

Table 1: Actual data and estimates for techniques

Eyring et al. (2024)		Asnicar et al. (2024)		Random Forest	
Actual data	**Estimates**	**Actual data**	**Estimates**	**Actual data**	**Estimates**
39.59	46.11	39.59	43.56	39.59	44.18
43.60	48.43	43.60	46.77	43.60	47.91
40.74	44.67	40.74	43.14	40.74	43.94
42.75	48.95	42.75	46.32	42.75	47.01
39.18	46.21	39.18	43.34	39.18	44.22
34.79	40.66	34.79	37.21	34.79	38.54
39.68	45.79	39.68	43.11	39.68	44.32
40.84	44.41	40.84	42.32	40.84	42.99
42.03	45.76	42.03	43.01	42.03	43.87
41.78	46.46	41.78	43.32	41.78	44.58
42.94	46.42	42.94	43.11	42.94	44.36
38.12	43.68	38.12	39.92	38.12	40.93
35.91	40.43	35.91	37.88	35.91	38.03

40.95	45.74	40.95	42.43	40.95	43.46
41.56	47.32	41.56	44.47	41.56	45.84
42.34	45.46	42.34	43.13	42.34	43.79
42.64	46.34	42.64	43.73	42.64	44.74
42.15	46.85	42.15	43.58	42.15	44.36
36.90	40.85	36.90	37.41	36.90	38.58
35.76	39.47	35.76	36.74	35.76	37.74
41.58	45.27	41.58	42.82	41.58	43.43
42.15	46.11	42.15	43.15	42.15	44.85
40.92	44.32	40.92	41.89		

The results of the comparisons are shown in Table 2. Table 2 indicates that RMSE for Eyring et al. (2024), Random Forest, and Asnicar et al. (2024) are 0.873, 0.596, and 0.704, respectively. The MAPE for the Random Forest is 0.921%, which is the smallest compared with Eyring et al. (2024), and Asnicar et al. (2024) having MAPE values of 1.957%, and 1.199%, respectively.

Table 2: Evaluation of techniques

	Eyring et al. (2024)	**Random Forest**	**Asnicar et al. (2024)**
RMSE	0.873	0.596	0.704
MAPE (%)	1.957	0.921	1.199

The output for the RMSE and MAPE of Eyring et al. (2024), Random Forest, and Asnicar et al. (2024) indicate that the Random Forest is the most accurate for estimating data. The RMSE and MAPE values of the Random Forest are the smallest evaluated compared to Eyring et al. (2024) and Asnicar et al. (2024). A critical study of past studies depicted in Section 2 shows that the Random Forest technique is the most widely used machine learning technique. Section 2 is organized to show the strengths and weaknesses of each method.

5 Conclusion

The study aims to develop a robust and dependable model capable of accurately predicting fuel consumption. The models' performance was assessed using the Mean Squared Error (MSE) metric. This metric was chosen as it effectively measures the accuracy of predictions, with lower values indicating better performance. Upon evaluating the models, it was observed that the Random Forest algorithm exhibited the least MSE value (0.008806), signifying superior predictive capabilities. The Linear Regression algorithm had an MSE of 0.010937, while the Decision Tree algorithm had an MSE of 0.015844. This outcome underscored the Random Forest algorithm's dominance in accurately predicting fuel consumption. These findings underscore the Random Forest's potential to revolutionize fuel consumption prediction, a crucial aspect of vehicle management and resource allocation. The model developed here offers a reliable tool for vehicle fuel consumption. The future research for this study will consider having diverse samples to validate the model's robustness and universality.

References

Ashqar, H. I., Obaid, M., Jaber, A., Ashqar, R., Khanfar, N. O., & Elhenawy, M. (2024). Incorporating driving behavior into vehicle fuel consumption prediction: methodology development and testing. *Discover Sustainability*, 5(1), 344.

Asnicar, F., Thomas, A. M., Passerini, A., Waldron, L., & Segata, N. (2024). Machine learning for microbiologists. *Nature Reviews Microbiology*, 22(4), 191-205.

Aziz, R. M., Sharma, P., & Hussain, A. (2024). Machine learning algorithms for crime prediction under Indian penal code. *Annals of data Science*, 11(1), 379-410.

Blockeel, H., Devos, L., Frénay, B., Nanfack, G., & Nijssen, S. (2023). Decision trees: from efficient prediction to responsible AI. *Frontiers in Artificial Intelligence*, 6, 1124553.

Chen, J., Li, K., Zhang, Z., Li, K., & Yu, P. S. (2021). A survey on applications of artificial intelligence in fighting against COVID-19. *ACM Computing Surveys (CSUR),* 54(8), 1-32.

Costa, V. G., & Pedreira, C. E. (2023). Recent advances in decision trees: An updated survey. Artificial Intelligence Review, 56(5), 4765-4800.

Eyring, V., Collins, W. D., Gentine, P., Barnes, E. A., Barreiro, M., Beucler, T., & Zanna, L. (2024). Pushing the frontiers in climate modelling and analysis with machine learning. *Nature Climate Change*, 14(9), 916-928.

Gupta, A. K., Pal, G. K., Rajput, K., & Bhatnagar, S. (2024, March). Analysis of Machine Learning Techniques for Fault Detection in 3D Printing. In 2024 2nd International Conference on Disruptive Technologies (ICDT) (pp. 1032-1037). IEEE.

Katyare, P., Joshi, S., & Kulkarni, M. (2024). Utilizing Machine Learning Approach to Forecast Fuel Consumption of Backhoe Loader Equipment. *International Journal of Advanced Computer Science & Applications*, 15(5).

Liu, T., Lin, L., Bi, X., Tian, L., Yang, K., Liu, J., & Pan, F. (2019). In situ quantification of interphasial chemistry in Li-ion battery. *Nature nanotechnology*, 14(1), 50-56.

Manivannan, R. (2024). Research on IoT-based hybrid electrical vehicles energy management systems using machine learning-based algorithm. *Sustainable Computing: Informatics and Systems*, 41, 100943.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. AI magazine, 27(4), 12-12.

Noviyanti, C. N., & Alamsyah, A. (2024). Early Detection of Diabetes Using Random Forest Algorithm. *Journal of Information System Exploration and Research*, 2(1).

Samuel, J., Kashyap, R., Samuel, Y., & Pelaez, A. (2022). Adaptive cognitive fit: Artificial intelligence augmented management of information facets and representations. *International journal of information management*, 65, 102505.

Su, M., Su, Z., Cao, S., Park, K. S., & Bae, S. H. (2023). Fuel Consumption Prediction and Optimization Model for Pure Car/Truck Transport Ships. *Journal of Marine Science and Engineering*, 11(6), 1231.

Tang, X., Zhou, H., Wang, F., Wang, W., & Lin, X. (2022). Longevity-conscious energy management strategy of fuel cell hybrid electric Vehicle Based on deep reinforcement learning. Energy, 238, 121593.

Wen, Z., Wang, Q., Ma, Y., Jacinthe, P. A., Liu, G., Li, S., & Song, K. (2024). Remote estimates of suspended particulate matter in global lakes using machine learning models. *International Soil and Water Conservation Research*, 12(1), 200-216.

Wu, Y. C., & Chang, Y. L. (2024). Ransomware detection on Linux using machine learning with random forest algorithm. Authorea Preprints.

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., & Zhang, J. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4).

Yang, H., Sun, Z., Han, P., & Ma, M. (2024). Data-driven prediction of ship fuel oil consumption based on machine learning models considering meteorological factors. Proceedings of the Institution of Mechanical Engineers, Part M: *Journal of Engineering for the Maritime Environment*, 238(3), 483-502.