International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 03, 2021

Forecasting Fuel Depletion Of Automobiles Through Machine Learning Algorithms

[1] D. Sasikala, [2]k. Venkatesh Sharma

[1]Department of CSE Nalla Malla Reddy Engineering College, Hyderabad
[2]Department of CSE , CVR College of Engineering, Hyderabad

Email: [1]godnnature@gmail.com, [2]venkateshsharma.cse@gmail.com

Abstract: Forecasting fuel depletion of automobiles - project targets at predicting the intake of fuel by automobiles which is vital in improving fuel economy of automobiles and preventing fraudulent activities. In the globe currently fuel plays a major role in transportation domain. ***Distance, capacity, automobile features, and motorist performance are the internal factors influencing the fuel depletion and pathway conditions, traffic flow, and climate shows a dynamic part of external factors. The foremost task is to model and predict the fuel depletion only with the available data with the stimulus of internal and external factors. However, a few of these factors are measured or available for the fuel depletion analysis. That is, a case is considered where only a subset of the above-mentioned factors is available as a multi-variate time series from a long distance, for different vehicles say, public and private bus, cabs etc.,. The recommended system using Machine Learning (ML) algorithms - Linear Regression Method, Multi Variate Method, and Random Forest Method performs a significant part in prediction. Random Forest Method overtakes the other two ML algorithms in its accomplishment.*** ***These predictions also assist in realizing how they can serve in progress of the ecosystem.***

Keywords: Prediction, Forecasting, Fuel depletion analysis, multi-variate time series, Machine Learning (ML) algorithms - Linear Regression Method, Multi Variate Method, and Random Forest Method.

1. INTRODUCTION

Fuel is a substance that causes reaction with new substances releasing energy used for work. They are reliable with other constituents or tools stock up potential energy, for example those that precisely remit electrical energy (for instance batteries and capacitors) or mechanical energy (such as regulating wheels, springs, dense air, or water in a cenote). Liquid fuels are combustible or energy-initiating particles that can be sourced to spawn mechanical energy largely turn out kinetic energy; they also require taking the form of their holder. It is the emissions of liquefied fuels that are igneous instead of the solution. Fuel prediction uses the data of fuel consumed, visualizes the data and predicts the fuel depletion. Not all the factors can be measured, as considering a case where only a subset of the above-mentioned factors is available as a multi-variate time series from a long distance, then in such cases the model should be trained with the available factors. Such analysis is powered by Machine Learning (ML), as the improvement of model is done by learning the patterns from the data and finally tested too.

Sources for Fuel and the growth of population are opposite to each other i.e., sources are decreasing and the latter is increasing which turn out to be a confronting task to preserve the

fuel. Hence, organized steps are not sufficient for fuel preservation, but insightful analysis should be done that confines the usage of fuel which wholly helps in fuel preservation. In the overseeing environment there is a lot of rising demand for fuel and it is known that fuel in a non-renewable resource and over time it can extinguish, so this work supports even common people to predict the amount of fuel being consumed by their automobiles so that over time they can lessen it and preserve the fuel for the future generation to come that is also known as the liquid gold. It would also help some local cab service owners to know the amount of fuel to be filled in their vehicles. As they have many vehicles, keeping an accurate track of fuel being consumed overall becomes complex and this research would help them to predict the total fuel being used by them in a month as well as to predict for the upcoming months. This work would help the government transport agencies, office transport services and many more.

2. LITERATURE SURVEY

This research [1] concentrates on assessing the extent of fuel ingested by heavy duty trucks in the European Union and thus limits the emissions being churned out. Neural Network was the utmost powerful algorithm amongst all the other learning techniques. The work performed in this paper [2] is with ML techniques to predict fuel depletion centered on independent variables.

An innovative structure was recommended [3] to uphold fuel-efficient driving performances over real-time spontaneous perception and motorist response. Using a dataset from a long-distance bus, this model generated an accuracy of 85.2% also raising the fuel efficiency up to 16.4%. This document [4] put on a data summarization approach established on distance instead of the customary time period while building adapted ML patterns for fuel ingesting. Article [5] aimed on Driver/Driving Behavior with big data.

Ability of this investigation [6] is to model and predict the fuel intake in improving fuel finances of automobiles and put off unlawful actions in fleet administration. Also by analyzing the three techniques, random forest prediction peaks the gradient boosting and neural networks. The drive of this manuscript [7] is to develop fuel ingestion reviewing databases centered on mobile phone data. The fuel ingesting prediction models are constructed using back propagation (BP) neural network, support vector regression (SVR), and random forests. The random forest model is evidenced to ensure the utmost accuracy and executes more rapidly, being appropriate for extensive uses and puts a basis for appraising databases and fine controlling of city transportation fuel ingestion. In script [8] neural networks offered precise approximations among the models that aid in ascertaining associations among energy usage and aspects as road and weather conditions. The intent of this paper [9] was to offer the real usage of niche technologies in cracking utmost acute complications of Fleet management. The suggested approaches in [10] lead to thorough report for the identification of aircraft fuel ingestion. Purposes of this investigation study [11] are modeling fuel ingestion, recognizing the correlation of motorist performance to fuel usage and spotting fuel fraud.

3. EXISTING SYSTEM

In the existing systems the vehicle itself is able to predict the fuel consumption and the distance the vehicle will be able to travel based on the way the vehicle is driven and the internal attributes of the vehicle. The problems of the existing system are given below: • Not very accurate.

- Difficult to access every time.

International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 03, 2021

4. PROPOSED SYSTEM

The proposed system using machine learning helps to overcome these shortcomings of the existing system. As the machine is intelligent enough to learn on its own over time the accuracy also increases. These are some of the major merits that lead to develop this Forecasting fuel depletion of automobiles - project. Merits of the proposed system: • Helps to manage one's money being spent on fuel. • Fuel consumption prediction to common users is easy.

The System after vigilant analysis has been identified to be present with the following modules. **1. Machine Learning Model Module 2. User Module 3. System Module**.

Software Requirements: **Operating System**: Microsoft Windows 7 or Above. **Technology**: Machine Learning and **Programming Language**: Python.

Hardware Requirements: **Processor** : Intel core5 or above and **RAM** : 4GB DDR4

1. Machine Learning Model Module:

Here, the model is used for predicting the fuel and it is a Supervised model which is trained based on the information of dataset which consists of tuples which were observed in Canada. It consists of 6000 tuples, after constructing the model it is saved for further predictions. Random Forest Algorithm was used for predicting the values.

2. User Module

This Module is the user who interacts with the model through an Interface. The Operations performed are: 1. Enters Engine Size 2. Enters emission of CO2 in gm/cc 3. Enters no. of Cylinders 4. Enters No. of km's to be travelled. This module maintains the user record and tracks the data accordingly.

3. System Module

System acts as an Interface between the Machine Learning Model and the User using it. It helps in User understanding by displays certain values and taking certain values.

UML Diagrams:

1. Use Case Diagrams

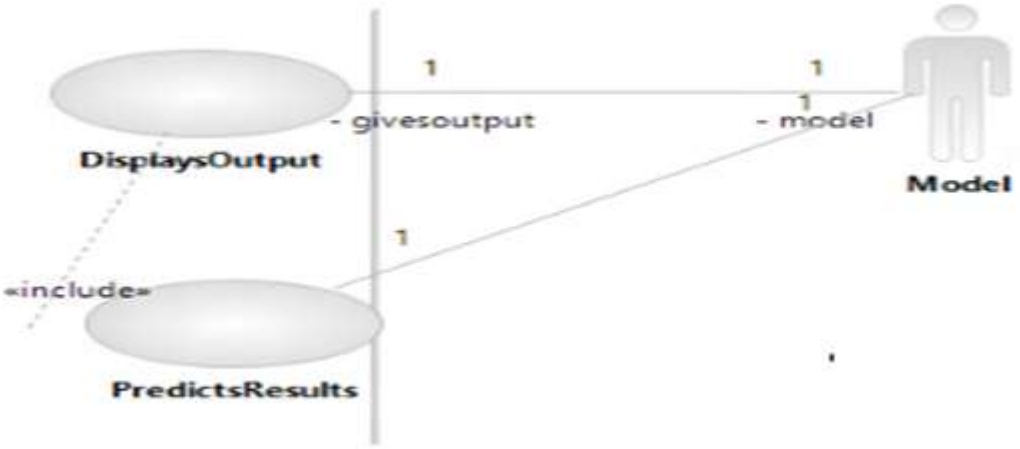

Figure 1. Use case Diagram for Machine Learning Model Module

Figure 1. illustrates the efficacy of a Machine Learning Module, where the model takes the input values processes them using the trained model and sends the values to the system that further displays them.

International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 03, 2021

Figure 2 Use case Diagram for System Module

Figure 2 depicts the functionality of System Module. The System acts an interface between the user and Trained machine learning module. It takes the values from the user requests for the model to process and gets the predicted values from the model and displays it to the user. It plays an important role in the prediction[12].

Figure 3. gives the picture of the role of User module and implements the operations of both the Developer and the User where from the programmer side it develops a model and from the user side it takes the values and processes it and gets to know the fuel to be filled in the vehicle in order to travel certain no. of kms.

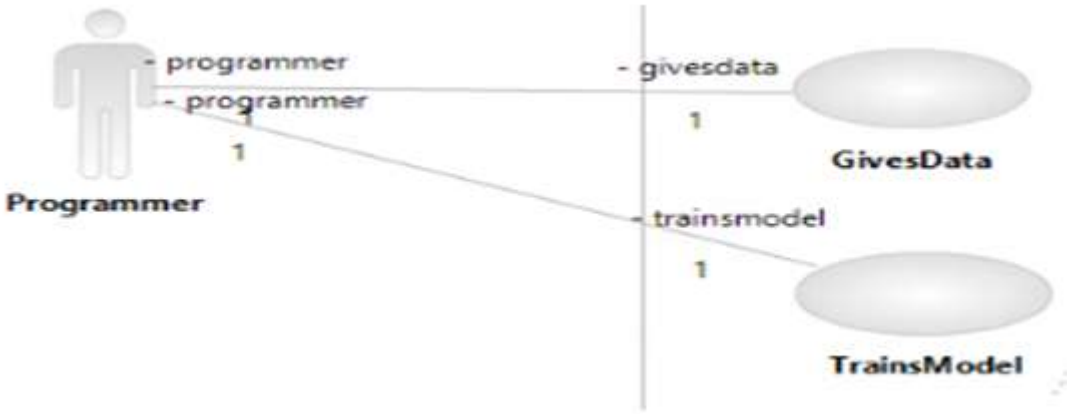

Figure 3. Use case Diagram for User Module

2. Class Diagram

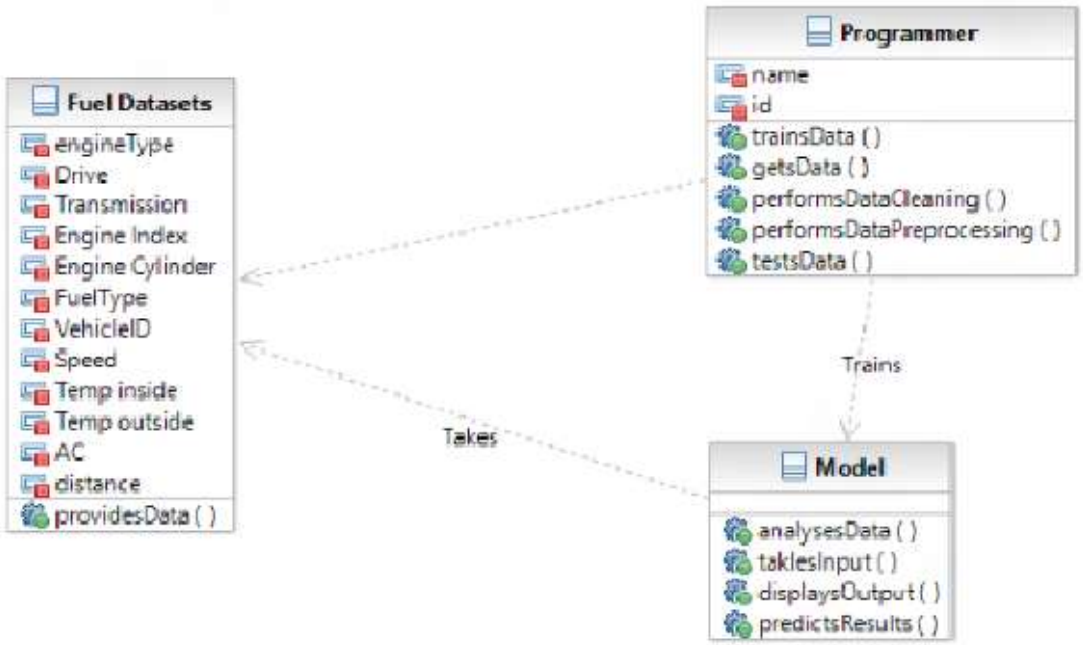

Figure 4. Class Diagram for Prediction of Fuel Consumption of Vehicles

Three classes are identified; they are User, Programmer and Model where they inter-connected to each other in performing the operations. The user and the programmer are considered the same with different operations.

3. Sequence Diagram

Initially the user enters the Interface and enters the Engine size-no of cylinders and CO2 emissions of his vehicle and upon clicking the predict button the system sends the values to the Machine Learning (ML) model for processing and predicting the value of fuel to be predicted [13]. Further, the user tries to know the amount of fuel to be filled in the vehicle for travelling certain no. of km's for that he enters the km's to be travelled and the system sends the values for processing and it calculates the values and displays the values to the user.

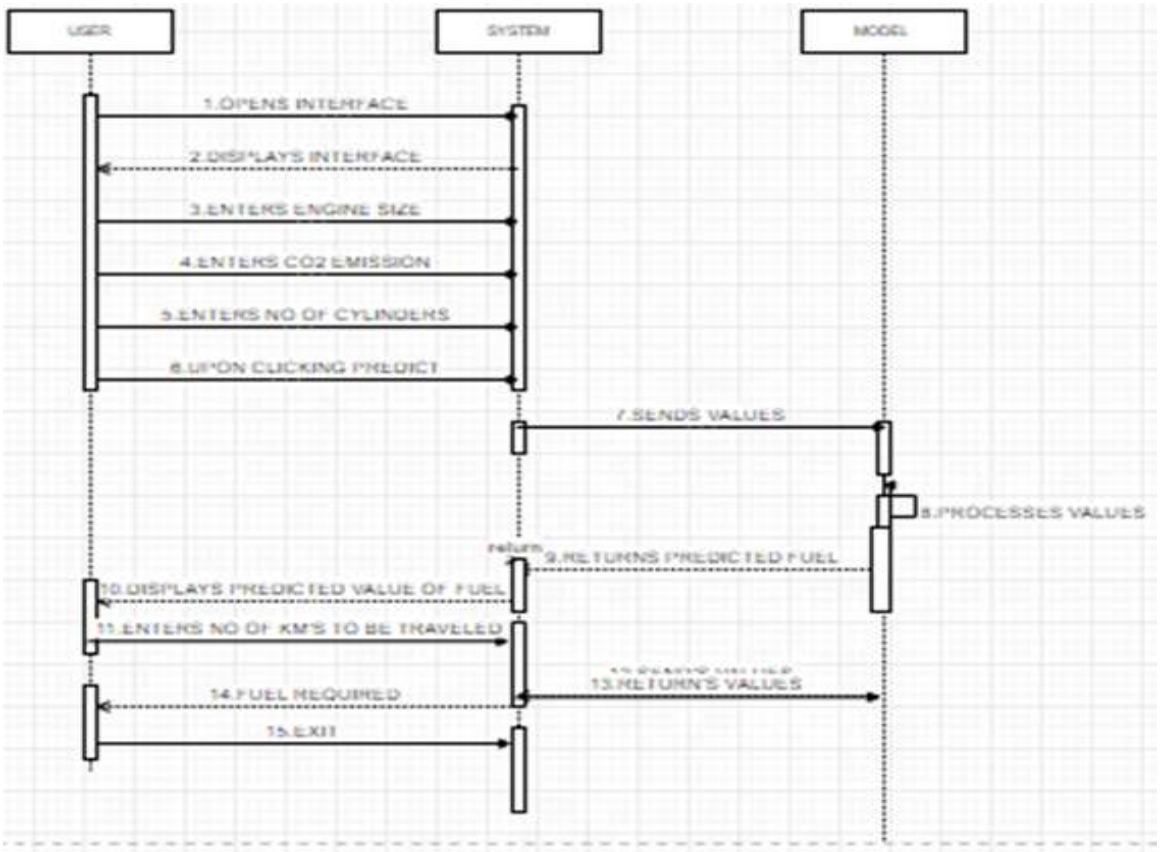

Figure 5. Sequence Diagram for Prediction of Fuel Consumption of Vehicles

4. Activity Diagram

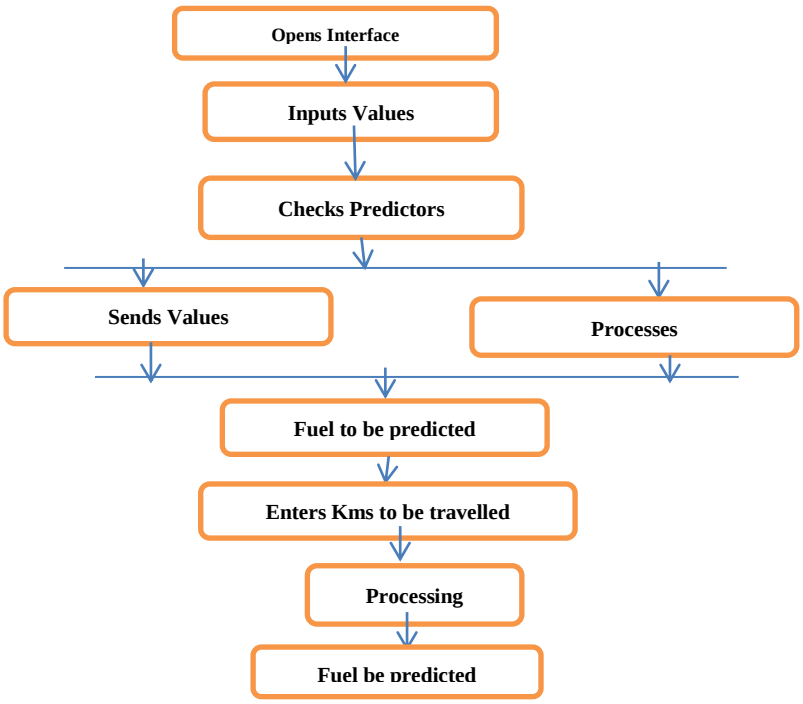

Figure 6 Activity Diagram

Figure 6 describes the flow of the events in the whole system as shown above. The events from starting to the end are represented.
These various ML algorithms, techniques and methods are used to build models for predicting and solving real-life problems in fuel depletion by using fuel consumption data with Python implementation. Then the implementation stage of this project is a factual display of the defining moments that make this project a success with the following: Linear Regression Method, Multi Variate Method, and Random Forest Method [14].

Linear Regression Method
Linear regression is a well-defined arithmetical model that investigates the linear correlation amongst a dependent variable with a particular set of independent variables. Linear regression is of the subsequent two types − • Simple Linear Regression • Multiple Linear Regression

Multi Variate Method
Multivariate refers to analyzing the linear relationship amongst more dependent variables with different variances or distributions.

Random Forest Method
Random forest is a supervised learning algorithm used for both classification and regression. But however it is applied for classification problems. Generally a forest is made up of trees and more trees mean more robust forest. Likewise, random forest algorithm constructs decision trees on data samples and then gets the prediction from each of them and lastly opts for the best result by means of voting. It is an ensemble method that is superior to a single decision tree as it lessens the over-fitting by averaging the result.
Working of Random Forest Algorithm
- Step 1 − First, start with the selection of random samples from a given dataset.
- Step 2 − Then, this process will build a decision tree for every single sample getting the prediction outcome from each decision tree.
- Step 3 − Here, voting will be accomplished for all predicted results.
- Step 4 – Finally, choose the most voted prediction result as the ultimate prediction result.

5. RESULT ANALYSIS

Output Screens of various functionalities in our interface are shown over here along with the description

	A	B	C	D	E	F	G	H
1	MODEL_YE	MAKE	ENGINE_SI	CO2_EMISS	CYLINDERS	CITY	HWY	AC
2	2014	ACURA	2	196	4	9.9	6.7	1
3	2014	ACURA	2.4	221	4	11.2	7.7	0
4	2014	ACURA	1.5	136	4	6	5.8	1
5	2014	ACURA	3.5	255	6	12.7	9.1	0
6	2014	ACURA	3.5	244	6	12.1	8.7	0
7	2014	ACURA	3.5	230	6	11.9	7.7	1
8	2014	ACURA	3.5	232	6	11.8	8.1	0
9	2014	ACURA	3.7	255	6	12.8	9	1
10	2014	ACURA	3.7	267	6	13.4	9.5	1
11	2014	ACURA	2.4	212	4	10.6	7.5	0
12	2014	ACURA	2.4	225	4	11.2	8.1	0
13	2014	ACURA	3.5	239	6	12.1	8.3	0
14	2014	ALFA ROMI	1.8	193	4	9.7	6.9	1
15	2014	ASTON MA	5.9	359	12	18	12.6	1
16	2014	ASTON MA	5.9	359	12	18	12.6	1
17	2014	ASTON MA	4.7	338	8	17.4	11.3	0
18	2014	ASTON MA	4.7	354	8	18.1	12.2	0
19	2014	ASTON MA	4.7	338	8	17.4	11.3	0
20	2014	ASTON MA	4.7	354	8	18.1	12.2	0
21	2014	ASTON MA	5.9	359	12	18	12.6	1
22	2014	AUDI	2	202	4	9.9	7.4	1
23	2014	AUDI	2	230	4	11.5	8.1	1
24	2014	AUDI	2	214	4	10.8	7.5	1
25	2014	AUDI	2	230	4	11.5	8.1	1
26	2014	AUDI	2	230	4	11.5	8.1	1
27	2014	AUDI	2	214	4	10.8	7.5	0
28	2014	AUDI	2	235	4	12	8.1	1
29	2014	AUDI	3	251	6	12.8	8.6	1
30	2014	AUDI	3	224	6	9.8	6.4	1
31	2014	AUDI	3	258	6	13.1	8.8	1
32	2014	AUDI	3	224	6	9.8	6.4	1

Figure 7 Dataset

The dataset which consists of different attributes such as CO2 emission, km's travelled in city, km's travelled in highway, Engine-Size, No. of cylinders of Vehicle are used. The values of these attributes are practically recorded on the Canadian roads. For building this model CO2 emissions, Engine Size and No. of cylinders are used as the independent variables.

Machine Learning is used for building this model, in ML Supervised Learning is used. Three different algorithms for building this model are used and they are:

Analysis of the dataset using linear regression:

First, a model using Linear Regression algorithm is built for predicting the fuel by considering the independent attribute as CO2 emission, but a conclusion that the fuel couldn't be predicted using only one attribute i.e. CO2 emission is drawn. The relation between CO2 emission and Fuel is shown in Figure 8:

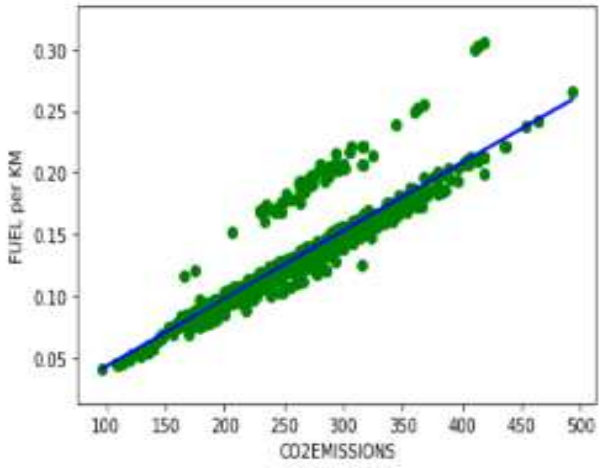

Figure 8. Graph between Fuel and CO2 -Linear relation

The proper relation between the fuel and CO2 emission couldn't be found. Upon analysis, thereby, it can be told that prediction of fuel by only using CO2 as the attribute cannot be found.

Analysis of dataset using multivariate regression:

International Journal of Aquatic Science
ISSN: 2008-8019
Vol 12, Issue 03, 2021

Multi-Variate Regression is considered for further considering different attributes for predicting the FUEL. Here, No. of cylinders and Engine Size as the Independent variables are considered. Upon building the model it is known that the values vary differently in accordance with the Fuel.

```
Intercept    58578718172.66581
coefficient  [-5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10  2.98023224e-08 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -5.85787182e+10 -5.85787182e+10
              -5.85787182e+10 -5.85787182e+10 -4.08172607e-04 -2.51083374e-02]
```

Figure 9. Error rate of Multi-variate

As, it is proved that fuel can't be predicted using Multi-Variate as the intercept value is high and coefficients are not related. Therefore, Multi-variate for building the model can't be used.

Random Forest:

This Algorithm is used for Final Model as it could be predicted that the values of fuel are with high efficiency.

```
Mean Absolute Error: 0.7297180957914039
Mean Squared Error: 2.1927461443635132
Root Mean Squared Error: 1.4807924042091496
```

Figure 10. Error outputs of random forest

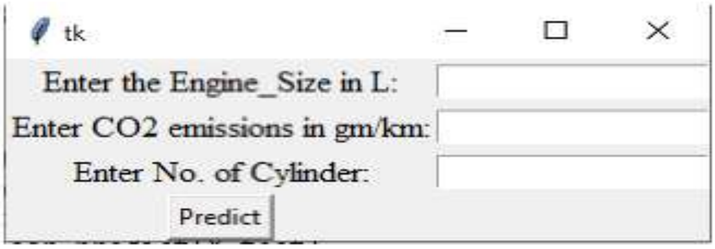

Figure 11. Interface of Prediction of Fuel Consumption

This is the basic Interface of Prediction of Fuel Consumption of Fuel. Here the Attributes that are taken for prediction are Engine Size, CO_2 and no. of cylinders of the particular vehicle is taken into consideration and the fuel values is predicted. The Engine size is taken in L,CO_2 emissions are taken as gm/km and the no. of cylinders are taken.

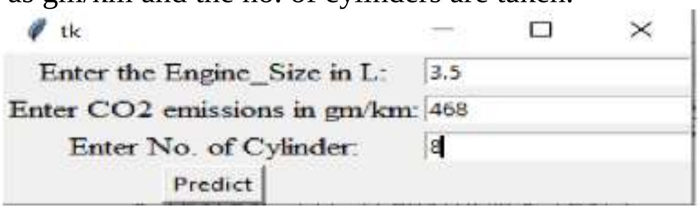

Figure 12. Values Input

Based on the available knowledge the user enters the values of the Engine size, no. of cylinders and the CO_2 emitted up to a certain period of time recorded. The values entered are

measured in specific unit's for different attributes, they are used for predicting the fuel to be filled the user for travelling. The values taken are from the mechanical units of the vehicle i.e. L,gm/km etc. which are truly based on the design of the engine. The values vary from Vehicle to vehicle and vehicle type to vehicle type.

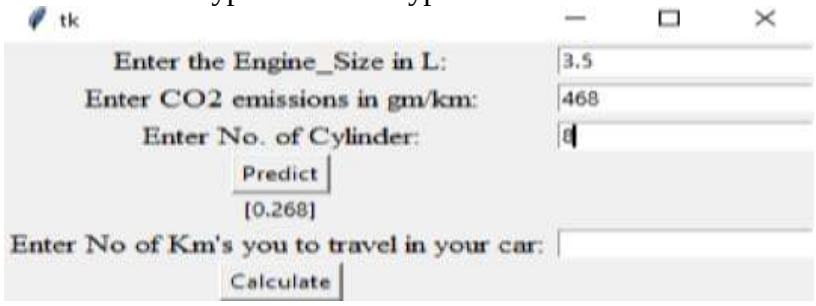

Figure 13. Predicting the Fuel

When the User clicks the Predict Button the amount of fuel to be used by the user of that particular vehicle for 1km is predicted by the MODEL and displayed by the System. If, further the user wants to know the amount of Fuel needed to travel certain km's he can enter the value in the Entry box given based on which the amount of fuel to be used is predicted.

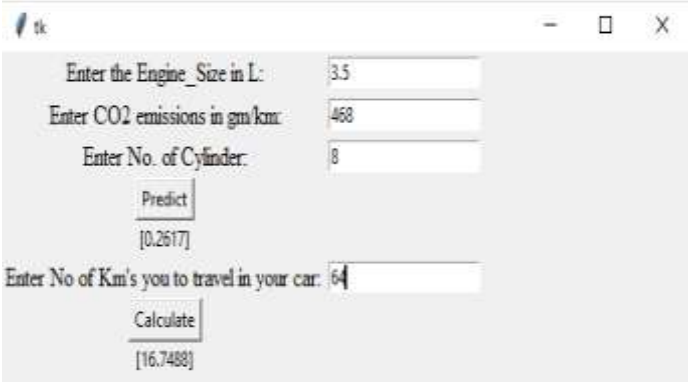

Figure 14. Final Output

Upon Clicking Calculate button the amount of fuel to be filled by the user in the vehicle is predicted. Based on which the vehicle user can fill the fuel for travelling. The value is predicted based on the Model which we have trained the model using the dataset which consist of 6000 records. The input value is taken the system and considered for predicting the fuel by takes the input values as the dependent variables and dependent variable as Fuel.

6. FUTURE WORK AND CONCLUSION

Certain ways how this project can be extended are:

- Building an application to help people keep track of their ingestion and predict their monthly consumption.
- Helps to know the overall depletion of fuel in the month or year by government or other bodies.
- By predicting fuel and the amount of CO2 that is being released by the consumption it can help to reduce pollution.

The fuel depletion prediction helps to identify the most suitable algorithm for the fuel prediction. This prediction would assist in knowing the amount of fuel being consumed by a person that would help to realize how they can benefit in betterment of the environment. This project also helps at predicting the real time fuel depletion of different vehicles.

It is very complex to create a system that composes all the necessities of the user. User requirements keep changing as the system is being used so with time the working is to be changed which could make it more reliable and efficient.

7. REFERENCES

[1] Akshay Bhoraskar, "Prediction of Fuel Consumption of Long Haul Heavy Duty Vehicles using Machine Learning and Comparison of the Performance of Various Learning Techniques.", Corpus ID: 203688706, Semantic Scholar, 2019.

[2] Thomas Bousonville, Martin Dirichs, Thilo Krüger, "Estimating truck fuel consumption with machine learning using telematics, topology and weather data", International Conference on Industrial Engineering and Systems Management (IESM), 25－27 September 2019, Shanghai, China.

[3]Sandareka Wickramanayake, Dilum Bandara, and Nishal Samarasekera,"Real-Time Monitoring and Driver Feedback to Promote Fuel Efficient Driving", https://www.researchgate.net/publication/342733294, July 2020.

[4] Alexander Schoen, Andy Byerly, Brent Hendrix, Zina Ben-Miled, "A Machine Learning Model for Average Fuel Consumption in Heavy Vehicles", IEEE Transactions on Vehicular Technology, May 2019.

[5] Ramazan Terzi, Seref Sagiroglu and Mustafa Umut Demirezen, "Big Data Perspective for Driver/Driving Behavior", IEEE Intelligent Transportation Systems Magazine, November 2018.

[6]Sandareka Wickramanayake, H. D. Dilum Bandara, "Fuel consumption prediction of fleet vehicles using Machine Learning: A comparative study", Moratuwa Engineering Research Conference (MERCon), 5-6 April 2016, Moratuwa, Sri Lanka.

[7] Ying Yao, Xiaohua Zhao, Chang Liu, and Yuelong Su, "Vehicle Fuel Consumption Prediction Method Based on Driving Behavior Data Collected from Smartphones", Journal of Advanced Transportation, March 2020.

[8] Afiya Ayman, Michael Wilbur, Amutheezan Sivagnanam, Philip Pugliese, Abhishek Dubey, and Aron Laszka, "Data-Driven Prediction of Route-Level Energy Use for Mixed-Vehicle Transit Fleets", 6th IEEE International Conference on Smart Computing(SMARTCOMP 2020), Bologna, Italy, 14^{th} – 17^{th} September 2020.

[9] Priya Singh, Milind Sukram Suryawanshi and Darshana Tak, "Smart Fleet Management System Using IoT, Computer Vision, Cloud Computing and Machine Learning Technologies", IEEE 5th International Conference for Convergence in Technology (I2CT), 29-31 March 2019, Pune, India.

[10] S. Baumann and U. Klingauf, "Modeling of aircraft fuel consumption using machine learning algorithms", CEAS Aeronautical Journal, September 2019.

[11] Sandareka Wickramanayake, Dilum Bandara, Shashika Muramudalige, Sasikala Kottegoda, Asiri Ishanka Liyana Arachchi, Afkham Azeez, Malintha Amarasinghe, Biman D Hettiarachchi and Chami Keerthisinghe, "Modeling and Enhancing Fuel Economy of Fleet Vehicles Based on Vehicular Data Analytics", Project 2016.

[12] Sujatha, K & Shalini Punithavathani, D, 'Optimized ensemble decision-based multi-focus image fusion using binary genetic Grey-wolf optimizer in camera sensor networks", Spinger, DOI: 10.1007/s11042-016-4312-3, Multimedia Tools and Applications .

[13] K Venkatachalam, S Balakrishnan, R Prabha, SP Premnath, Effective Feature Set Selection And Centroid Classifier Algorithm For Web Services Discovery, International Journal of Pure and Applied Mathematics,Vol 119, issues-12,pp 1157-1172,2018

[14]https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_random_forest.htm

Authors Profile

Dr D.Sasikala[1] ME, PhD, MBA, MPhil, MSc. MISTE, Professor, Department of CSE, Nalla Malla Reddy Engineering College, Area of Specialization: Digital Image Processing, Artificial Intelligence, Deep Learning & Machine Learning, Software Engineering, Soft Computing & Optimization Techniques, Data Mining & Data Warehousing, Data Science & Big Data Analytics.

Dr K.Venkatesh Sharma[2] MTech, AMIETE, PhD, Professor, Department of CSE, CVR College of Engineering, Area of Specialization: Artificial Intelligence and Machine Learning, Software Engineering, Soft Computing & Optimization Techniques, Data Mining & Data Warehousing, Big Data Analytics & Data Science.