IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2021
Impact Factor: 4.819

Fuel Economy Prediction using Feature Engineering

Yash Nadkarni, Siddhesh Deo, Aditya Patwardhan and Amey Ponkshe
UG Students, Department of Computer Engineering,
Marathwada Mitra Mandal's College of Engineering, Pune, Maharashtra, India
Savitribai Phule Pune University, Pune, Maharashtra, India

Abstract: *The traditional way to calculate fuel economy is done by using odometer reading and fuel consumed by car to travel that particular distance. This is a very narrow approach as fuel economy is affected by a variety of factors in the real world. Features such as throttle response, engine temperature, coolant temperature, gross weight of vehicle, etc. have a huge influence on the fuel economy. In order to overcome this problem, we have tried to predict fuel economy based on various features extracted from telemetric data in our project. In order to achieve this, we have implemented various feature selection and feature extraction techniques by further analyzing them with the purpose of calculating the effectiveness of those features to achieve high performance of machine learning algorithms that ultimately improves the predictive accuracy of the classifier. This provides us with the information regarding the amount of influence a particular feature has on the overall fuel economy of the vehicle.*

Keywords: Machine Learning, Fuel Economy, Vehicle Telemetry, Data Mining, Feature Engineering, Performance Analysis, Feature Extraction, Machine Learning Algorithms.

I. INTRODUCTION

As compared to other countries, the prices of petrol or diesel or any other form of by-products of petroleum has always been higher in India. Despite being subsidized from the government, the prices of petrol and diesel have always been a big but necessary burden on our income. Due to this there is a vast measure of inter-state inequality with reference to fuel prices with Maharashtra being one of the worst. The daily necessary consumption expenses rise far higher due to this factor. Many world leading automobile industries are currently working on optimizing the consumption of fuel. Hence, we have tried to find a way to reduce those expenses, so that it might become easier on our personal disposable income. Using this fuel economy, we can adapt it to a particular driver using his/her previous data, and according to the terrain, can gauge how the driver will perform. This score can then be compared with how the driver actually performs on the vehicle and helps us in identifying the most suitable, economical and optimal sets for use.

Automated feature Engineering that can be incorporated by Machine Learning software has become available only recently. Our goal, in this paper, is to create a model that automatizes specific features that are extracted according to our needs. Feature Engineering focuses on the extraction of raw features via data mining techniques. Various machine learning algorithms are then applied to the processed dataset and one of them is deployed based on evaluation of the models' individual performance. The first step in creating a model for the fuel economy is feature extraction. Feature Engineering is one of the most important domains in the course of developing projects. It can be defined as the process of extracting features by using raw techniques which belong to Data Mining. Extracting Features is also one of the most challenging tasks that needs to be performed, and without which data cannot be analyzed properly. A feature can have various priorities according to the project in hand. Likewise, a feature can have various levels of relevance. Machine Learning is one of the essential domains of Computer Science, as it provides simplicity and reduces complexity, to both the creator, as well as the user. It is a highly preferred domain, used in the case of vehicles, as the data from the sensors

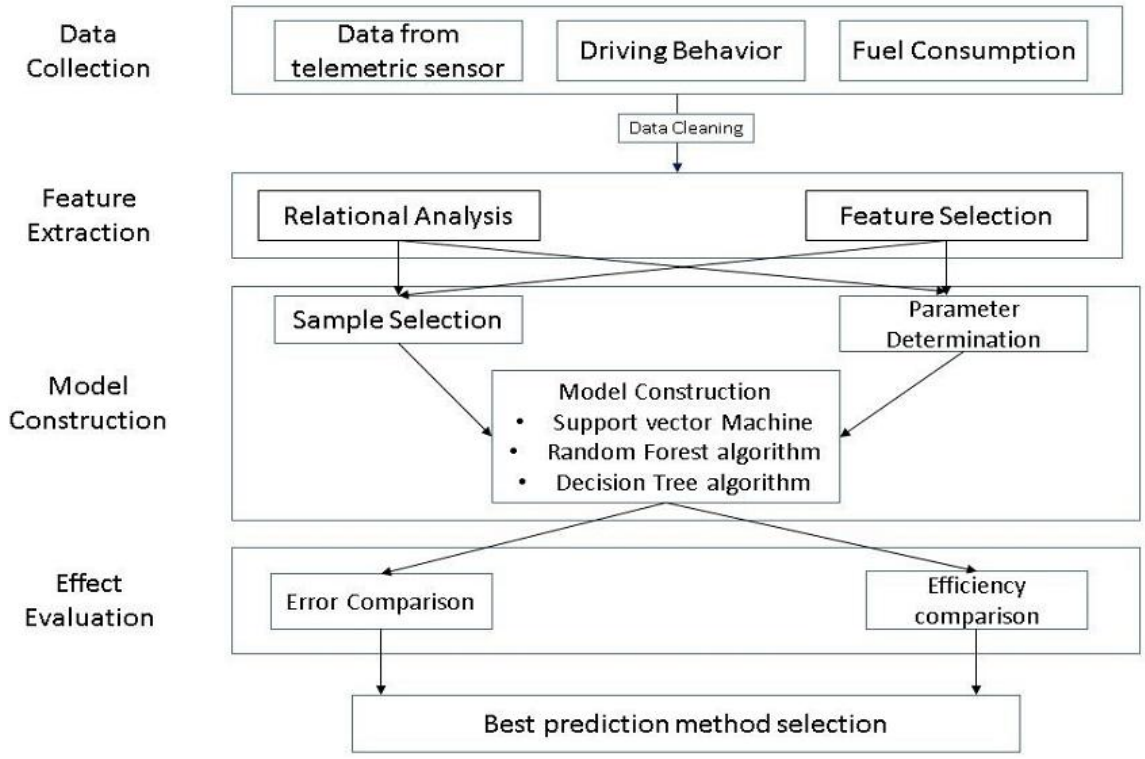
IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2021
Impact Factor: 4.819

can be directly used to train the ML models. This project focuses on predicting the fuel economy of a car based on vehicle telemetric data.

II. Research Methodology

A very important factor that affects energy consumption, driving behaviour and such is transportation [8]. In the traditional sense, fuel economy was always calculated using the odometer readings, amount of fuel that was consumed (in litres/gallons) and some such. But in this paper, we have used various other factors for the calculation of fuel economy. After careful inspection, we have found that these features, although less in quantity, in the strictly technical sense of the word, exert a huge influence in the overall fuel economy of a car. Due to the lack of real driving data for this paper, we have collected a dataset which measures the values of each individual sensor. These sensors were placed on different test vehicles and put through multiple trips, therein deriving the test data.

After each trip, we had a comprehensive collection of data that provided us with the working of almost every sensor that was placed in the test vehicle such as Torque Engine Percentage, Switch pedal break, Temp Intake Manifold, Coolant Temp Engine, Course In Degrees, etc. Using these factors, we used feature engineering to weigh out individual factors which affected the actual fuel consumption of the vehicle. The diagram below depicts the overall research methodology used in this paper.



III. Implementation

In the beginning, we obtained vehicle telemetric data which had information about gears, altitude, vehicle speed, fuel remaining, oil as well as engine temperature, intake temperature, mileage, etc. at 0.5 second time intervals. So, in order to perform feature extraction, we first divided the data into parts where each part consisted of vehicle data which spans a length of road of 4 km. Hence, for this, we implemented a sliding window for the division of our data. Our first window consisted of the first 4 kilometers of the trip. Then, the window slides forward by 2 kilometers, thereby keeping 50% overlapping. This made our next reading. In this way, we kept some part of the trip data constant so that

IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2021
Impact Factor: 4.819

our machine learning model does not encounter completely new data during its training. In order to find out the factors that affect fuel economy, we decided to develop various functions in order to extract and analyze features. The first function helped us to identify the instances of sudden acceleration during the drive. Here, instead of upshifting, if the driver just hammers down the pedal, the RPM(revolutions per minute) of the engine increases tremendously. This is because higher the RPM, higher will be the amount of fuel consumed. Thus, marking such instances of sudden throttle increase helps us to analyze one of the factors affecting fuel economy. This same function can also be implemented for sudden braking. The snippet given below depicts change in RPM per Gear:

```
def rpm_range_gear_1():
    min_val = 1300
    max_val = 2000
    count = 0
    for i in range(len(data)-1,0,-1):
        if data.iloc[i]['RPM'] > max_val and data.iloc[i]['Gear']==1:
            count +=1

    return count

rpm_range_gear_1()
```

In the process of feature extraction we came to know that fuel economy is greatly affected by throttle damping. Combining factors such as throttle pressed, gear and RPM we can calculate torque which directly affects fuel consumption at particular instances. Each gear in a car has a specific predetermined speed range, if the driver over speeds the particular speed range it will be reflected on fuel economy. The function given below depicts sudden Change in Acceleration:

```
def Throttle_bad_or_good():
    fixed_percentage_value=20    #For example
    for i in range(len(data)-1,0,-2): # -2 for each sec of iteration as freq is 2 Hz
        print(data['Time'][i])
        if (data['Throttle_Percentage'][i+2]-data['Throttle_Percentage'][i])>fixed_percentage_value:
            print('BAD')
        else:
            print('GOOD')

```

Fuel economy is greatly affected by driving vehicles in the wrong gear. If the gear is not optimal with the corresponding speed of the vehicle more fuel is consumed and affects the fuel economy vastly. Such instances are marked when speed is not in range of the optimal gear which helps us to analyze one of the factors affecting the fuel economy. The code given below depicts the number of times coolant temperature is out of the range:

```
def coolent_temp():
    min_temp=70
    max_temp=93
    count =0
    for i in range(len(data)-1,0,-1):
        if data.iloc[i]['Coolant_Temp'] < min_temp or data.iloc[i]['Coolant_Temp'] > max_temp:
            count +=1
    return count

coolent_temp()
```

After extracting these features, the data was passed on to the model for training. We implemented various classifiers like decision tree, random forest, KNN and support vector machines and compared their scores. The algorithm given below provides a general idea about the working of the project:

Algorithm:

1. Import the data from car telemetric sensors in json format.
2. Perform data cleaning as telemetric sensor data will consist of many irregularities.
3. Analysis and Feature Extraction:

IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, February 2021
Impact Factor: 4.819

a. Analyse inter-relation between factors in order to find which factors are most dominant in affecting fuel economy. If necessary, combine multiple factors into a single feature.

4. Sample selection and Parameter determination
   a. Perform feature selection and finalize the dataset. Pass this feature extracted dataset to the model
5. Model Construction
   a. Try out different machine learning algorithms like decision tree, random forest, KNN, SVMs, etc. and find out their accuracy and scores.
6. Compare the scores and select the best machine learning algorithm for the model accordingly.
7. Stop

IV. CONCLUSION

Hence, by deploying various feature extraction and selection techniques, we were able to predict the fuel economy of a car during a trip and the amount of influence other factors had on it. As our project was intended for research and analysis of the different components of vehicles, this is deemed useful for fuel consumption analysis.

FUTURE WORK

This software can be used to analyse the performance of the vehicle. Further, the analysis made by the software could be used to grade the driver's performance. Automotive companies or transport companies can use this software to know the vehicle performance and grade the driver's journey. This system can be deployed in not just cars but all sorts of different vehicles. Currently, we are analyzing the data from only a single car. However, this can be extended to multiple cars and other vehicles. Another potential future direction is to set up a dashboard which will provide all the analysis of the vehicle's fuel economy based on the sensor data. A real-time driver scoring a system can be set up which will continuously evaluate the driver thereby showing his score as the vehicle is being driven around.

REFERENCES

[1]. Wingston Harrington. Fuel Economy and Motor Vehicle Emissions (July 1997).
[2]. Karl-Heinz hauer Dipl. Analysis Tool for Fuel Cell Vehicle hardware and Software (Controls) with an Application to Fuel Economy Comparisons of Alternative Systems Designs (1990)
[3]. Bernie Porter, Hugh Blaxill and Noor Jariri. A Study of Potential Fuel Economy Technologies to Achieve CAFE 2025 Regulations using Fleet Simulation Modelling Software (July 2015)
[4]. Serhad Soyhan, Hakan Emre, Kilic, Mehmet Gokalp, Burak Taymaz and Imdat. Performance Comparison of MATLAB and Neuro Solution Software on Estimation of Fuel Economy by Using Artificial Neural Network (2009)
[5]. Peng Ping, Wenhu Qin, Yang Xu , Chiyomi Miyajimaand Kazuya Takeda. Impact of Driver Behaviour on Fuel Consumption: Classification, Evaluation and Prediction Using Machine Learning (June 2019)
[6]. Tushar D. Gaikwad, Zachary D. Asher, Kuan Liu, Mike Huang, Ilya Kolmanovsky. Vehicle Velocity Prediction and Energy Management Strategy Integration of Machine Learning Vehicle Velocity Prediction with Optimal Energy Management to Improve Fuel Economy (Jan 2019)
[7]. Hang-Bong Kang. Various Approaches for Driver and Driving Behavior Monitoring(2013)
[8]. Xiaohua Zhou, Jian Huang, Weifeng Lv, Dapeng Li, "Fuel Consumption Estimates Based on Driving Pattern Recognition", 2013 IEEE International Conference on Green Computing and communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing