e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 7, July 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

6381 907 438 6381 907 438 ijmrset@gmail.com www.ijmrset.com

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |

Prediction of Fuel Efficiency using Machine Learning

Mrs.P.Vanitha, Mr.N.Karthik

Assistant Professor, Department of Computer Applications (UG), Hindusthan College of Arts & Science, Coimbatore, Tamil Nadu, India

Assistant Professor, Department of IT & Cognitive Systems, Sri Krishna Arts and Science College, Coimbatore, Tamil Nadu, India

ABSTRACT: To enhance the accuracy of the fuel consumption prediction model with Machine Learning to minimize Fuel Consumption. This will lead to an economic improvement for the business and satisfy the domain needs. To propose a machine learning model to predict vehicle fuel consumption. The proposed model is based on the Support Vector Machine algorithm. The Fuel Consumption estimation is given as a function of Mass Air Flow, Vehicle Speed, Revolutions per Minute, and Throttle Position Sensor features. The proposed model is applied and tested on a vehicle's On-Board Diagnostics Dataset. The observations were conducted on 18 features. Results achieved a higher accuracy with an R-Squared metric value of 0.97 than other related work using the same Support VectorMachineregressionalgorithm.WeconcludedthattheSupportVectorMachinehasa great effect when used for fuel consumption prediction purposes. Our model can compete with other Machine Learning algorithms for the same purpose which will help manufacturers find more choices for successful Fuel Consumption Prediction models.

KEYWORDS: Machine Learning; Ship Fuel Consumption Prediction; Black-Box Model; White-Box Model; Convolutional Neural Networks.

I. INTRODUCTION

An accurate fuel consumption prediction model is the basis for ship navigation status analysis, energy conservation, and emission reduction. In this study, we develop a black-box model based on machine learning and a white-box model based on mathematical methods to predict ship fuel consumption rates. We also apply the Kwon formula as a data preprocessing cleaning method for the black-box model that can eliminate the data generated during the acceleration and deceleration process. The ship model test data and the regression methods are employed to evaluate the accuracy of the models. Furthermore, we use the predicted correlation between fuel consumption rates and speed under simulated conditions for model performance validation. We also discuss applying the data-cleaning method in the preprocessing of the black-box model. The results demonstrate that this method is feasible and can support the performance of the fuel consumption model in a broad and dense distribution of noise data in data collected from real ships. We improved the error to 4% of the white- box model and the $R2$ to 0.9977 and 0.9922 of the XG Boost and RF models, respectively. After applying the Kwon cleaning method, the value of $R22$ also can reach 0.9954, which can provide decision support for the operation of shipping companies.

1.1 OBJECTIVES OF THE PROJECT

The objectives of the project are as stated below:

1. To develop a system to provide a more effective way of short-listing
2. To determine the key skill characteristic by defining each expert's preferences and ranking.
3. To automate the process of requirement specifications and applicant's
4. To conduct online aptitude and personality
5. To produce ranking decisions that would have relatively higher consistency than those of human

1.2 MODULE DESCRIPTION

This study will be valuable to the management of commercial banks as they will be able to understand the relevance of Curriculum vitae for the main reason being to help the banks in retaining its employers for a considerable period. Consequently exert greater effort on the client personalities by the incumbent resulting in increased client personalities

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |

performance and better attendance. It's a study that will be useful to the management of commercial banks as they will understand the relationship between Curriculum vitae and employer turnover and hence ensure that the employers are informed everything regarding the client personalities they are to take. Further the study is important to all organizations in Nigeria especially the many who have not yet embraced RJP to try it so as to reduce employer turnover especially early employer turnovers. It is also a wakeup call to the human resource practitioners to optimize the use of RJP so as to cut on turnover and hence improve the recruitment process. Further, of it being a new concept, not many workforce are aware of it, thus it will be a means of not only educating the many human resource professionals about this crucial technique and apply it as required but also create awareness to all labor force so that in case the RJP is not used, they know what they are missing out.

The study will benefit the government especially the Ministry of Finance for making policy decisions whose overall objectives are to accelerate the rate of employer performance and reduce the turnover rate thus improving service delivery. The human resource specialists wishing to implement RJPs as part of their recruitment and staffing practices would also be able to maximize the gains attributable to RJP interventions by matching RJP methods to the organizationaloutcomestheyseektoaffect.Theformandtimingthatwillmaximizepositive

Outcomes such as client personalities performance and satisfaction and minimize outcomes such as turnover remain unknown. In addition, the human resource governing bodies such is IHRM (Institute of Human Resource Management) also will come to the realization of using it (RJP) and emphasize its use to the large human resource fraternity.

The scholars will gain knowledge on the role of Curriculum vitae on employers. They will acquire an insight into the influence of Curriculum vitae on employers‟ turnover. They will also be able to use the findings of this study to prove various theories and to use thestudyasabasisforfurtherresearchonothervariablesnotincludedinthisstudy.The findings will also increase the stock of theoretical and empirical knowledge especially in the African context and also form the basis for further research.

II. SYSTEM ANALYSIS

2.1 EXISTING SYSTEM

There are several existing ideas and approaches for predicting fuel efficiency using machine learning. Here are some:

1. **Feature Engineering:** One common approach is to engineer features from vehicle specifications such as engine size, weight, horsepower, torque, aerodynamics, transmission type, fuel type, and tire pressure. These features can serve as input variables for machine learning models.
2. **Historical Data Analysis:** Analyzing historical fuel efficiency data for different vehicles can provide insights into the factors that influence fuel efficiency. Machine learning models can then be trained on this data to predict fuel efficiency for new vehicles.
3. **Real-Time Sensor Data:** Using sensors installed in vehicles, real-time data such as vehicle speed, engine RPM, throttle position, and fuel consumption rate can be collected. Machine learning models can analyses this data to predict fuel efficiency and provide recommendations for fuel-efficient driving behavior.
4. **Weather and Road Conditions:** Weather conditions (e.g., temperature, humidity, wind speed) and road conditions (e.g., traffic congestion, road gradient) can affect fuel efficiency. Integrating weather and road condition data into machine learning models can improve the accuracy of fuel efficiency predictions.
5. **Vehicle Telematics Data**: Vehicle telematics systems can provide rich data on vehicle performance, driving behavior, and maintenance status. Machine learning models can leverage this data to predict fuel efficiency and identify opportunities for fuel savings through vehicle maintenance and driver training.
6. **Hybrid and Electric Vehicles:** For hybrid and electric vehicles, machine learning models can predict energy consumption and battery usage based on driving patterns, route characteristics, and vehicle specifications. These predictions can help optimize energy management strategies to improve fuel efficiency and extend battery life.
7. **Customer Segmentation:** Analyzing customer data such as demographics, purchasing behavior, and vehicle preferences can help identify segments of customers with similar fuel efficiency requirements. Machine learning models can then be tailored to each segment to provide personalized fuel efficiency predictions and recommendations.
8. **Anomaly Detection**: Detecting anomalies in fuel consumption patterns can indicate potential issues such as fuel leaks, engine malfunctions, or inefficient driving behavior. Machine learning models trained on historical data can flag anomalies in real-time data streams and alert vehicle owners or fleet managers to take corrective actions, how machine learning can be applied to predict fuel efficiency and optimize vehicle performance. The key is to leverage relevant data sources, engineer informative features, and select appropriate machine learning algorithms

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |

to build accurate and robust predictive models.

2.2 PROPOSED SYSTEM

DATA COLLECTION PROCESS

1. EXPERIMENT DESIGN

It suggests that driving behavior is affected by various factors such as street design, traffic management methods, traffic conditions, weather conditions and the driver's mental and physical condition. In order to evaluate the effect of the driver's condition on vehicle fuel consumption and simplify the verification process, in this study we fixed the vehicle type, trip route and weather conditions used in our experiment. The only variable factors are the drivers (i.e., their driving behavior) and the traffic conditions. If more than one route were used in the experiment, it would be difficult to determine which factors were primarily responsible for variation in fuel consumption. Therefore, all of the data for our experiment was collected using a fixed route which included some variation in road types. Examples of the two types of roads used in our study are shown in Fig. 2. The total distance of all of the road segments was about 15.2 km, which consisted of a 5.3 km expressway loop with two lanes in each direction and 9.9 km of ordinary road with one lane in each direction.

Our data was collected using 30 normal passenger cars with a 1.2T (85kw) gasoline engine and a six-speed automatic transmission (6AT). Fuel consumption increases by 0.38± 0.079% each time the air temperature decreases by 1°C[36]. Therefore, in order to avoid the possibility of variations in air temperature obscuring the relationship between driving behavior and fuel consumption, the data collection was conducted in the autumn from September to November. 202 drivers are selected to join the experiment, the information of the drivers is shown in Fig. 4.As the supervised and unsupervised learning method need lots of samples, so we try out best to find the experiment participants as much as possible. We choose these 202 drivers from our university's students and the cooperator's staffs. All the participants drove in the experimental route for10 circuits a day and the whole experiment of single drivers last a week. When processing our experiment, we did not give time limitation or some special driving tasks to the participants in order to avoid extra mental pressure. We just tell them the research goal, experimental route and drive as they usually do. Most of the experiment participants are in normal emotion and will be paid after the experiment.

2. DATA COLLECTION AND REDUNDANT DATA PRUNING

The data collection system (DCS) is divided into three parts: a vehicle- mounted data collection system (VMDCS), a wireless transmission system (WTS) and a data center (DC). The VMDCS uses On-Board Diagnostics (OBD) to obtain the vehicle's operating information from the ECU, and uses GPS to track the vehicle's position. The WTS uses a wireless transmission unit (WTU) installed on the vehicle which communicates with the base station via 4G broadband to upload the collected data. Messages from the WTS include a receiving module IP address so that the data can be transmitted to the DC via the internet. The DC server shows the vehicle's position and real-time vehicle information on the Web. The collected data is stored in an SQL database.

In order to improve calculation efficiency, we selected vehicle operation data with a strong relationship to driving behavior, and used the Pearson correlation coefficient (PCC)[37]to determine the relevance of each parameter to vehicle fuel consumption. We treated positive and negative acceleration as different parameters because their effects on fuel economy differ. For example, when calculating fuel cost, if negative acceleration is less than zero, instantaneous

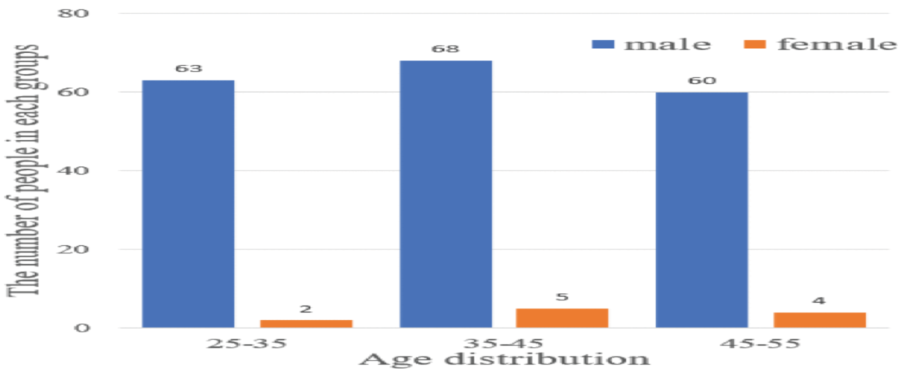
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |

fuel consumption is zero. The calculated correlation coefficients for various features are listed, where PCC value ρ is represented by different color bars according to the following standard guidelines; when |ρ|>0.5=strong correlation, when0.5>|ρ|>0.3=moderate correlation, when |ρ|<0.3=weak correlation [38]. In 'Negative acc' and 'Negative acc variance' have a negative correlation with fuel consumption, so in fact, the PCC of these two parameters are negative values. Then, before using an unsupervised clustering method to abstract the data distribution features, we first pruned the weakly correlated data parameters.



3. FUEL CONSUMPTION CALCULATION

To calculate fuel consumption, we integrated instant fuel consumption information from the ECU to obtain accumulated fuel consumption data. In order to verify the results of our calculations, we compared our calculated results with the results from a fuel consumption analyzer under various traffic conditions. The differences between these two fuel consumption measurement approaches are shown in Table 1.

Vehicle load \ Road type	Urban road	Expressway	Rural road
No-load	4.85%	1.28%	2.18%
Full-load	5.94%	0.81%	3.65%

From the data in Table 1, we can conclude that the difference between our calculation method and actual fuel consumption is less than 6%. As the route used in our experiment is only15 km in length and the goal of the study is to evaluate the effect of driving behavior on fuel consumption, this difference can be ignored.

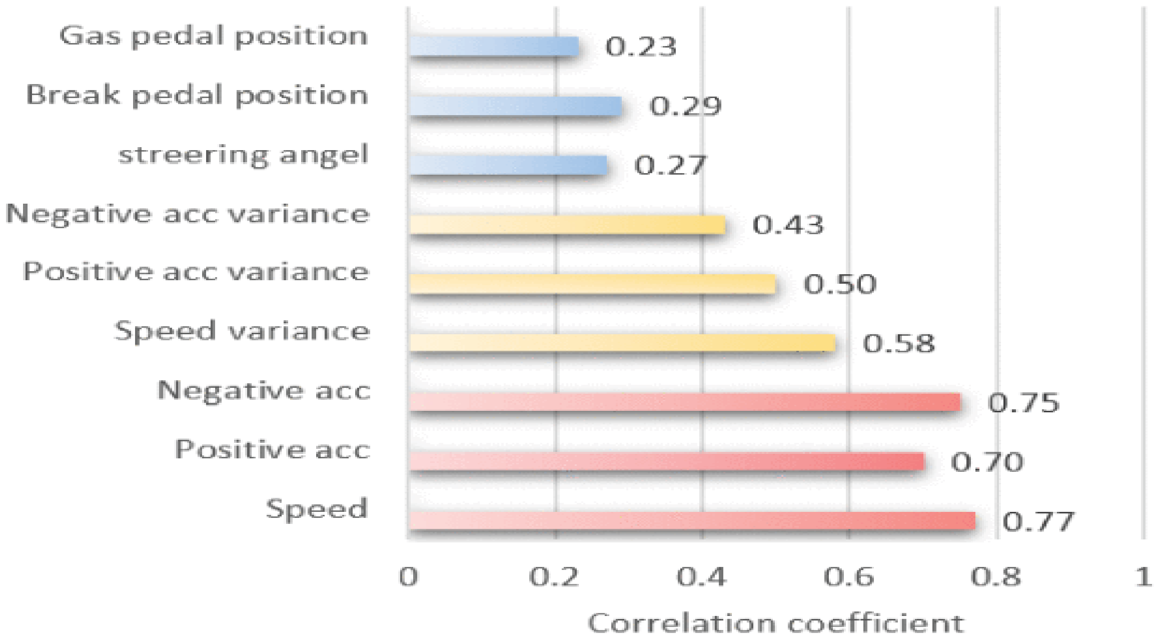
4. DATA SEGMENT CONSTRUCTION

As our research goal is to analyze and predict the impact of driving behavior on fuel consumption within a limited time frame (25 to 35 minutes), in this section we describe the spectral clustering method we used to compare inner similarity within the data set, so as to cluster data with similar features into the same cluster. Our spectral clustering method can only handle data sets of the same size. The data collection rate was 10Hz and we collected 15,000-21,000 data points per circuit of the driving route (we treated each circuit of the driving route as an independent data set). Since the amount of data collected in each data set varied, we needed to compress each data set to the same size.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |



Firstly partitioned the raw data set into several subsets. The driving route was divided into 50 road segments according to their location distribution. And then the whole data will be divided according to their belonging road segment (each data points contain the GPS position). As each road segment contains a different number of data points, we needed to calculate each segment's minimum data size Sn. For example, S1istheminimumdata size of the first road segment (calculated from the entire dataset associated with the first road segment).Each dataset allocated to road segment 1 is then compressed to sizeS1. After data compression, each data set will have the same data size.

In contrast to using maximum information entropy to select the size limit of the data, as in our previous study [39], the data compression method adopted in this paper allows us to retain most of the data points.

5. PREDICTION OF SHORT-TERM FUEL CONSUMPTION USING LSTM

The clustering-based method proposed in Section II above can only provide relatively long-term (25 to 35 minutes) assessment of the impact of a driver's behavior on fuel consumption. When attempting to perform relatively short-term prediction (30 seconds to 5 minutes), the clustering-based method does not work well for classifying driving behavior according to fuel efficiency. Besides, our clustering method is, in fact, a kind of classifier, so it has no prediction ability. Therefore, in this section we propose the Use of a time series learning method (an LSTM network) to model the relationship between driving behavior and fuel consumption, allowing us to predict the short-term fuel consumption state of a driver's behavior. As a driving behavior pattern represents the driver's interaction with a dynamic driving environment, and fuel consumption can be treated as the cost result of this process, in this section we add dynamic driving environment information to our learning data. In the series data construction process described in this section, we first explain how we coded driving environment factors into a digital form. Then the environmental feature data and the behavior data are integrated into time-series data using a sliding window. Fuel consumption state will be the label for the constructed time-series data set. The LSTM-based model is then trained using the time-series data. The model's classification performance and prediction accuracy will be discussed at the end of this section.

6. TIME-SERIES DATA CONSTRUCTION

A. CODING OF ENVIRONMENTAL FACTORS:

As explained in our previous study [49], we divided the environmental factors into two categories, dynamic environmental features (other vehicles, brake lights of leading vehicles, pedestrians, etc.) and static environmental features (features which remain in variable for relatively long periods of time, including road structures such as intersections and curves).Some of the dynamic features are captured by a camera mounted on the vehicle. As shown inFigs. 2and16, two types of roads were used in this study. In Fig. 16, the gray car is the experimental vehicle, the red vehicle is the leading vehicle or leading vehicle in the right lane, the blue vehicle is a parked vehicle, the green vehicle

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |

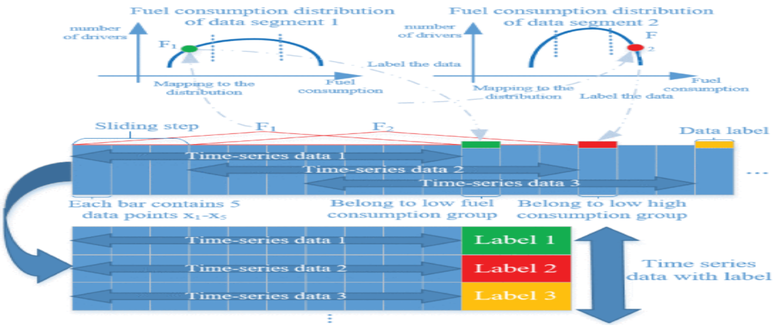
is the first on-coming vehicle in the opposite lane and the yellow vehicle is the second on-coming vehicle in the opposite lane.

$$S_{all} = \sum_{i=1}^{50} S_i$$

In ordinary-road scenes (one lane in each direction), the motorcycle or motorbike and the pedestrian are also considered to be environmental factors which can affect the driver's behavior. Thanks to the development of object detection technology, we can easily extract these traffic environment factors. In this study we used YOLOv3 [50], a deep learning-based, real-time object detection method, to obtain the relative positions of these traffic factors. Using this position information, we can code the traffic factors into a digital form.

7. FUEL CONSUMPTION FEATURE LABELING AND TIME SERIES DATA CONSTRUCTION:
In BT represents the driving behavior data set from one trip along the fixed driving route, while S represents the size of the data (the number of behavior data points) collected during the time period it took to complete the route. S is calculated by applying the method.

The only difference in compressing process used in this section is that here, we divide the experimental road into 150segments instead of 50 in order to obtain much more detailed data features. N in (16) represents the driving behavior categories strongly and moderately correlated with fuel consumption (N = 6).



8. METHODOLOGY

The system has two modules, one being candidate oriented and the other module is organization oriented. In the first case the system would enable the candidate to give the test for a particular company and also view the results of their previous tests which would help them to improve their performance. In the second scenario, the specifications and requirements of available job positions would be posted by the recruiter and the candidates can apply for the same by appearing for the required test.

III. SYSTEM DESIGN

The data in the system has to be stored and retrieved from database. Designing the database is part of system design. Data elements and data structures to be stored have been identified at analysis stage. They are structured and put together to design the data storage and retrieval system.

A database is a collection of interrelated data stored with minimum redundancy to serve many users quickly and efficiently. The general objective is to make database access easy, quick, inexpensive and flexible for the user. Relationships are established between the data items and unnecessary data items are removed.

Normalization is done to get an internal consistency of data and to have minimum redundancy and maximum stability. This ensures minimizing data storage required, minimizing chances of data inconsistencies and optimizing for updates. The MySQL database has been chosen for developing the relevant databases.

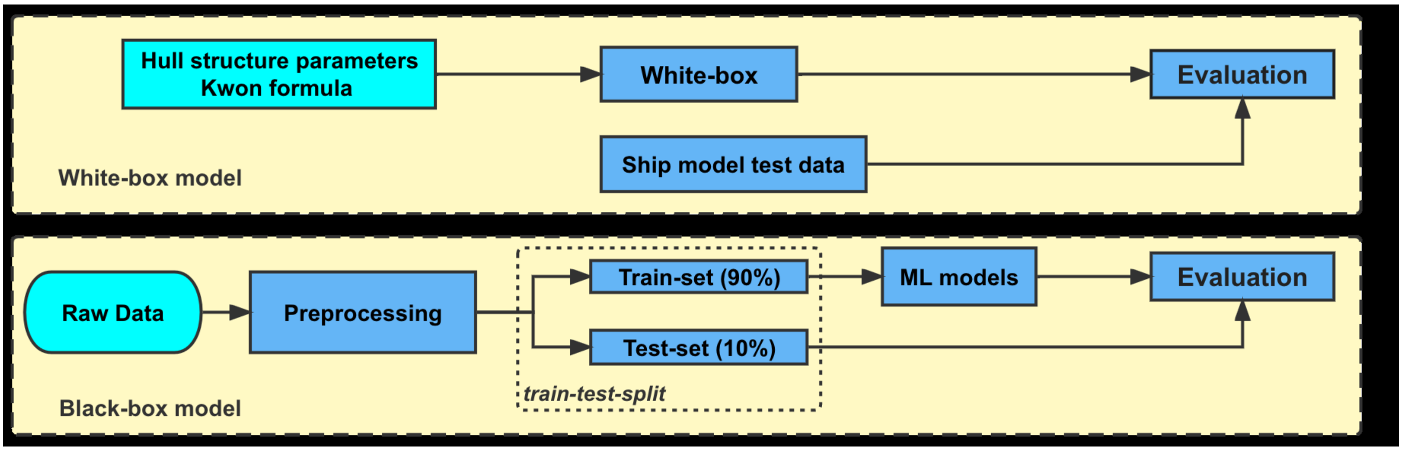
FUEL EFFICIENCY PREDICTION MODELS
We establish two types of models to predict the FCR: white-box and black-box (ML, machine learning) models. The

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |

flow chart of the model building.



IV. SYSTEM TESTING AND IMPLEMETATION

Software testing is a critical element of software quality assurance and represents the ultimate review of specification, design and coding. In fact, testing is the one step in the software engineering process that could be viewed as destructive rather than constructive.

A strategy for software testing integrates software test case design methods into a well-planned series of steps that result in the successful construction of software. Testing is the set of activities that can be planned in advance and conducted systematically. The underlying motivation of program testing is to affirm software quality with methods that can economically and effectively apply to both strategic to both large and small-scale systems.

V. FUTURE ENHANCEMENT

In the future ,we can improve the dataset so that the model will give more appropriate results .We can also integrate an SMS service into our system so that the users can get up dates on the latest job openings. Recruiters can also use our system as a job recruitment platform. Our system can also be use full for college placements .To improve the accuracy of our model, we can add more aptitude tests on-site, to increase the efficiency of our system. Our system can also be of use in matrimonial sites in the future.

VI. CONCLUSION

In concluded unsupervised machine learning method of spectral clustering to classify drivers into three groups using six driving behavior-based fuel consumption features. We then analyzed the macro-behavior of each group, focusing on power demand (speed and acceleration) and control stability (variation in speed and acceleration). Our results showed that the proposed spectral clustering-based method could accurately identify drivers with different fuel consumption profiles, and clearly modelled the relationship between the real- world driving data and the corresponding fuel consumption features. In addition to the estimation of fuel consumption using vehicle operation data, we also performed a qualitative analyses of driving behavior, Speed and acceleration information reveal the amount of power demanded by a driver, while variance in speed and acceleration represent the range of dynamic control exercised by drivers [25],[26]. The results of our analysis showed that high fuel consumption drivers (those in the red cluster) tend to maintain a relatively steady, high demand for power, while their dynamic control of the vehicle is less stable. Their acceleration rates are higher, and their pedal control behavior is less stable compared to drivers in the low fuel consumption cluster. Drivers in the median yellow cluster showed the lowest speed distribution, but their gas and brake pedal operation characteristics were similar to those of the low efficiency drivers in the red cluster. Drivers in the blue cluster had the lowest fuel consumption, since they tended to maintain a consistent speed, and their dynamic control of the vehicle was the most stable among the three groups. We also compared the spectral cluster method with other state of art clustering method such as k- means and KFCM. As show in Table.3, spectral cluster method can achieve the best clustering performance of the two methods.

However, there were drawbacks to our proposed method, in that the spectral clustering-based method requires

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.524| Monthly, Peer Reviewed & Referred Journal|

| Volume 7, Issue 7, July 2024 |

| DOI:10.15680/IJMRSET.2024.0707283 |

relatively long-term data to produce accurate classification results. So, for real-time and short-term fuel consumption feature prediction, this unsupervised method is not appropriate. Furthermore, the results of data mining can only show the impact of a driver's behavior on fuel consumption on a macro-level.

REFERENCES

1. A.Samuel,"Design and implementation of web based Recruitment Portal (A case study of Nigeria Civil Service Commission Enugu) BY NWAMAGHINNA BLESSING," 2016 10 [Online].
2. http://www.Academia.Edu14695018/Design And Implementation of Web Based_ Recruitment_ Portal _A case study of Nigeria Civil Service Commission Enugu BY NWAMAGHINNA BLESSING. [Accessed 20 January 2019].
3. S. Khmer, "Designing a Database for an Online Job Portal," 15 November2016.[Online]Available:http://www.vertabelo.com/blog/technical-articles/designing-a-database-for-anonline-job-portal. [Accessed 3 February 2019].
4. Bettors"bettors,"2015.[Online].Available:http://www.bdtutors.com/.[Accessed5 February2019].
5. W. &. WHATWG, "Hypertext Mark up Language (HTML)," 14 December 2017. [Online] Available: https://en.wikipedia.org/wiki/HTML. [Accessed 9 February 2019].
6. Tutorial Points "Cascading Style Sheet", 6 February 2013.[Online]. Available: https://www.tutorialspoint.com/css/. [Accessed 28 august 2018].
7. Tutorial Points, "JavaScript," 4 June 2017. [Online].Available: https://www.tutorialspoint. com/javascript/ [Accessed 9 February 2018].
8. TutorialPoints,"PHP,"4June2017.[Online].Available:https://www.tutorialspoint.com /php/[Accessed 9 February2018].

INTERNATIONAL JOURNAL OF

MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com