# Trip Based Modeling of Fuel Consumption in Modern Fleet Vehicles

IBM–DOCUMENTATION

**UNDER THE GUIDANCE OF**

Industry Mentor Name : Prof Swetha

Faculty Mentor Name : Dr. S. RAMESH

**TEAM ID : PNT2022TMID38910**

**SUBMITTED BY:**


| | |
|---|---|
| Nandhini S | 421319104019 |
| Jayashree V | 421319104009 |
| Kalaiselvi K | 421319104011 |
| Sharmila M | 421319104026 |

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**KRISHNASAMY COLLEGE OF ENGINEERING AND TECHNOLOGY**

**ANNA UNIVERSITY :: 2019-2023**

# 1. INTRODUCTION

## 1.1 Project Overview

Heavy-duty trucks contribute approximately 20% of fuel consumption in the United States of America(USA).The fuel economy of heavy- (HDV) duty vehicles worldparameterslikeroadparameters,driverbehaviour,weatherconditions,andvehicleparameters ,etc.Althoughmodernvehiclescomplywithemissionsregulations,potentialmalfunctionoftheengi ne,regular wear and tear, or other factors could affect vehicle performance. Predicting fuel consumptionper trip based on dynamic on-road data can help the automotive roadtesting.Datamodellingcaneasilyhelptodiagnosethereasonbehindfuelconsumptionwith a knowledge of input parameters. In this paper, an artificial neural network (ANN) was implemented to model fuel consumption in modern heavy-duty trucks for predicting the total and instantaneous fuel consumption of a trip based on very few key parameters, such as engine load (%),engine speed (rpm), and vehicle speed (km/h). Instantaneous fuel cantopredictpatternsinfuelconsumptionforoptimizedfleetoperations.Inthiswork,thedatausedfor modelling was collected at a frequency of 1Hz during on-road testing of modern heavy-duty vehicles(HDV) at the West Virginia University Centre for Alternative Fuels Engines and Emissions (WVUCAFEE) using the portable emissions monitoring system (PEMS). The performance of the artificialneural network was evaluated using mean absolute error (MAE) and root mean square error (RMSE).The model was further evaluated with data collected from a vehicle on-road trip. The study showsthat artificial neural networks performed slightly better than other machine learning techniques suchas linear regression (LR), and random forest (RF), with high R-squared ($R2$) and lower root meansquareerror.

## 1.2 Purpose

The fuel efficiency of heavy-duty trucks can be beneficial not only for the automotive the and transportation industry but also for a country’s economy and the global environment. The cost of fuel consumed contributes to approximately 30% of a heavy-duty truck's life cycle.

# 2. LITERATURE SURVEY

## 2.1 Existing Problem

In the various nations, different financial, technological and transportation policies adopted to significantly reduce the enormous amount of gasoline consumption, which can be seen in Fig.1. The brief description collected from Silva et al. (2009), Jonkoping et al. (2010), Javari and Baratimalayerei (2008), Verdinejad (2009), Arête and Zimmerman's (2004), Bates (2000), Baht (2004), Foumunung et al. (1999) can be seen in Fig.1. All the above mentioned facts have turned gasoline consumption into a crisis in Iran. Unfortunately, the present energy conservation strategies, including standards, developing alternative fuel (CNG/LPG), technical inspection, and excluding old vehicles have not reduced consumption efficiently. The proposed solutions, which are categorized into short-term (such as increasing buses, restriction of trips to central regions of major cities, registration policy-making, etc.), mid-term (such as increasing production capacity of refineries, justifying and goal-orienting subsidies, increasing fuel prices to international levels, etc.) and long-term (developing rail-transportation systems, etc.) solutions, are essential steps for transportation energy consumption providence. The purpose of this research is to use fuel consumption data, and other related records, to model travel demand conditions. The vision behind this research can be understood by this statement; “how to establish four steps classic transportation models by records from transport fleet consumption’s data." Absolute results or answers for this question also solve more problems in different areas as economic aspects, environmental aspects, decision making, fuel demand, future forecasting, etc. Also, these actions help planners to forecast and plan with real and factual situation data. Efforts for answering this question can be effective to conquer the present difficulties in different economic, environmental, decision making, fuel demand, future forecasting aspects and etc. to reach the suitable sustainable development in transportation. Furthermore, these efforts can be helpful for planners to identify reliable data and exploring the alternative methods for travel demand modelling.

## 2.2 References

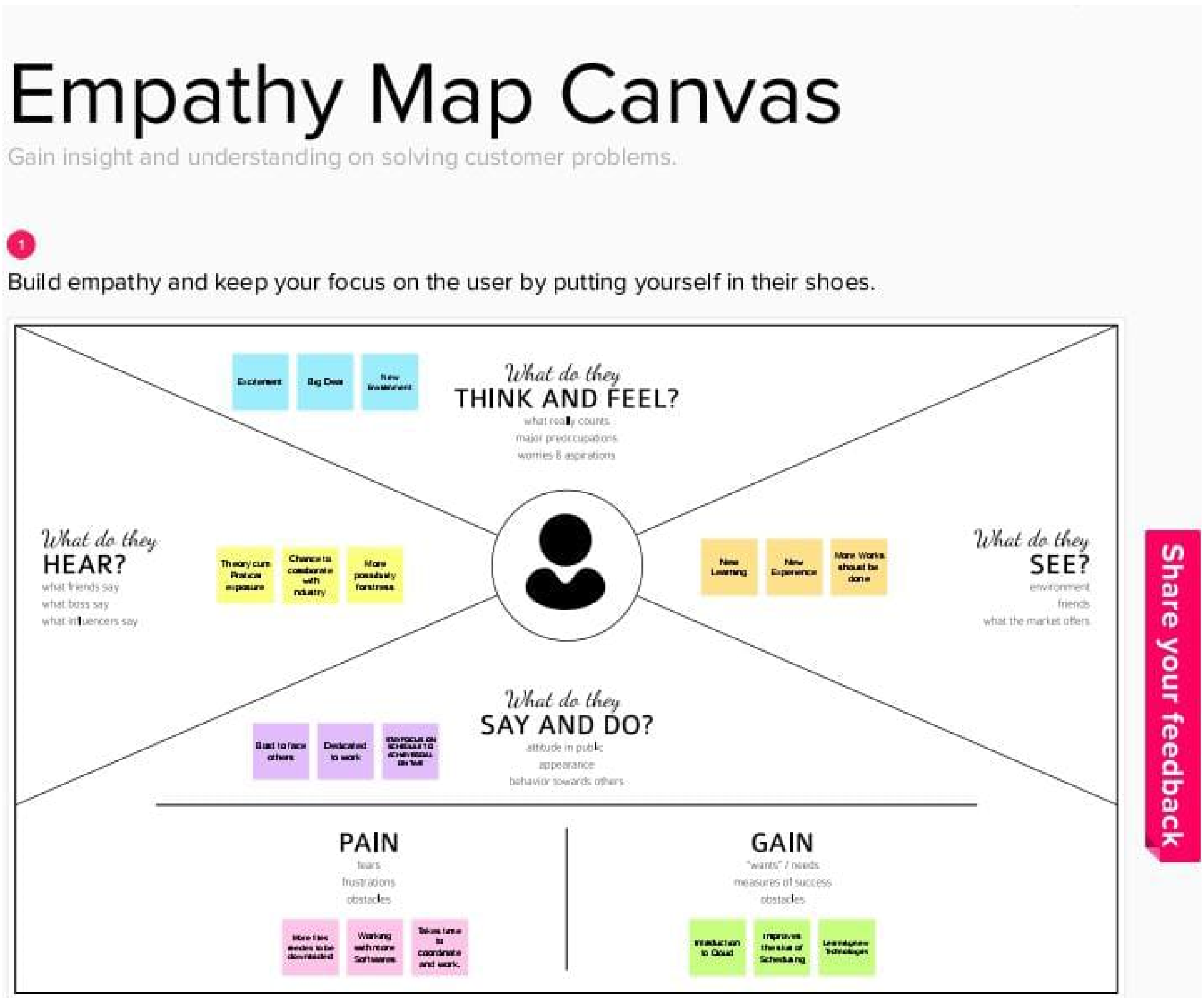

- Agency, E.P. Sources of Greenhouse Gas Emissions.Available online:https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions(accessedon5October2021).
- NHTSA's Corporate Average Fuel Economy (CAFE). Available online: https://www.nhtsa.gov/laws-regulations/corporate-average-fuel-economy(accessedon5October2021).


## 2.3 Problem Statement Definition

- A new asymmetric twin-scroll turbocharged engine with two EGR circuits is first presented.
- Experiment and simulation are combined on the diesel engine with asymmetric turbocharger.
- Effect laws of turbine critical parameters and EGR valves control strategy are explored.
- The new system has the maximum EGR rate and fuel economy improvements of 8.59% and 1.98%.
- The research collects bus fuel consumption data for diesel buses.
- Models are developed to compute the fuel consumption levels of buses.
- The optimum bus fuel economy cruising speeds range between 40 and 50 km/h.

## 3. IDEATION AND PROPOSED SOLUTION

### 3.1 Empathy Map Canvas




### 3.2 Ideation And Brain Storming

Hardware Requirements

1. 8GB RAM
2. Intel Core i3
3. Laptop/Desktop

4. Windows/MAC/Linux OS.

Software Requirements

1. Python
2. numpy
3. Pandas
4. Seaborn
5. Matplotlib.pyplot

## 3.3 Proposed solution

| **S. No.** | **Parameter** | **Description** |
|---|---|---|
| 1. | Problem Statement (Problemtobe solved) | A new asymmetric twin-scroll turbo charged engine with two EGRcircuits is first presented. Experiment and simulation a recombined on the diesel engine with asymmetric turbo charger. Effect laws of turbine critical parameters and EGR valves control strategy are explored. |

| | | |
|---|---|---|
| 2. | Idea/Solution description | Capture and prevent fuel theft and leakage.<br>Fuel monitoring and anti-siphoning devices can update you on fuel amounts in trucks and on-sitetanks and send alerts about fuellevels.<br><br>**Calculate and report fuel taxes**. Integrated with your vehicle's GPS, a fuel management system can automatically calculate travelled distance and purchased fuel to help file your IFTA tax reports.<br><br>The ultimate benefit of fuel management systems is automation operations that used to be done manually in spread sheets can happen automatically in the background, providing analytics for you to base your decisions on. Now, let's cover the main opportunities and how they work. |

| | | |
|---|---|---|
| 3. | Novelty/Uniqueness | Use the intuitive charts and reportsprovided on the Onboard Cloud toanalysessafetytrendsfor yourfleet.<br><br>Deep dive into vehicle and driver-level analytics to pinpoint safetyissues and identify actions.<br><br>Onboard Cloud also combinescollision avoidance alerts with othertelemetric parameters to provide apowerful picture of how your fleetsafetyisimproving. |

# 3.4 Proposed Solution Fit

**Project Title: Trip Based Modeling of Fuel Consumption in Modern Fleet Vehicles**

**Project Design Phase-I - Solution Fit Template**

**Team ID:** PNT2022TMID38910

Define CS, fit into CC

**1. CUSTOMER SEGMENT(S)** CS

Transport Vehicles Driver/owner

Transport Companies

**6. CUSTOMER CONSTRAINTS** CC

Experiment and simulation are combined on the diesel engine with asymmetric turbocharger

**5. AVAILABLE SOLUTIONS** AS

Ideations

Providing data and research analysis of fuel economy potential by country and region

Creating database with data collected using PEMS devices

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

**2. JOBS-TO-BE-DONE / PROBLEMS** J&P

Models are developed to compute the fuel consumption levels of vehicles

The research collects bus fuel consumption data for diesel vehicles

**9. PROBLEM ROOT CAUSE** RC

The model is more consistent with empirical observations compared to the MOVES and CMEM models

**7. BEHAVIOUR** BE

A new asymmetric twin-scroll turbocharged engine with two EGR circuits is first presented

The new system has the maximum EGR rate and fuel economy improvements of 8.59% and 1.98%

Focus on J&P, tap into BE, understand RC

**3. TRIGGERS**

TR

Performing correlation analysis on the input parameters selected to eliminate multi-colinear variables.

**10. YOUR SOLUTION**

SL

Developing the neural networks and identifying the network with best-performing hyper parameters

**8. CHANNELS of BEHAVIOUR**

CH

8.1 ONLINE

Check current ongoing fuel consumption

8.2 OFFLINE

**4. EMOTIONS: BEFORE / AFTER**

EM

before customer can only monitor the fuel consumption

**after developing the model customer can anti-siphoning devices update on fuel amounts in trucks**

**They incorporate data about fuel transactions into analytics**

The hyper parameters include the number of hidden layers learning rate and optimization function.

Customer can view previous fuel monitoring status

# 4. REQUIREMENT ANALYSIS

## 4.1 FUNCTIONAL REQUIREMENTS

Followingarethefunctionalrequirementsoftheproposedsolution.

| **FR No.** | **Functional Requirement** | **Sub Requirement(Story/Sub-Task)** |
|---|---|---|
| FR-1 | User Registration | Registration through<br>Form Registration<br>through Gmail<br>Registration through LinkedIN |
| FR-2 | User Confirmation | Confirmation via Email<br>Confirmation via OTP |
| FR-3 | User Signup | Signup through register mail id<br>Signup through password |
| FR-4 | User Signout | Sign out the register mail id |
| FR-5 | User Forgot Password | Change password  through mail id |

## 4.2 NON-FUNCTIONAL REQUIREMENTS

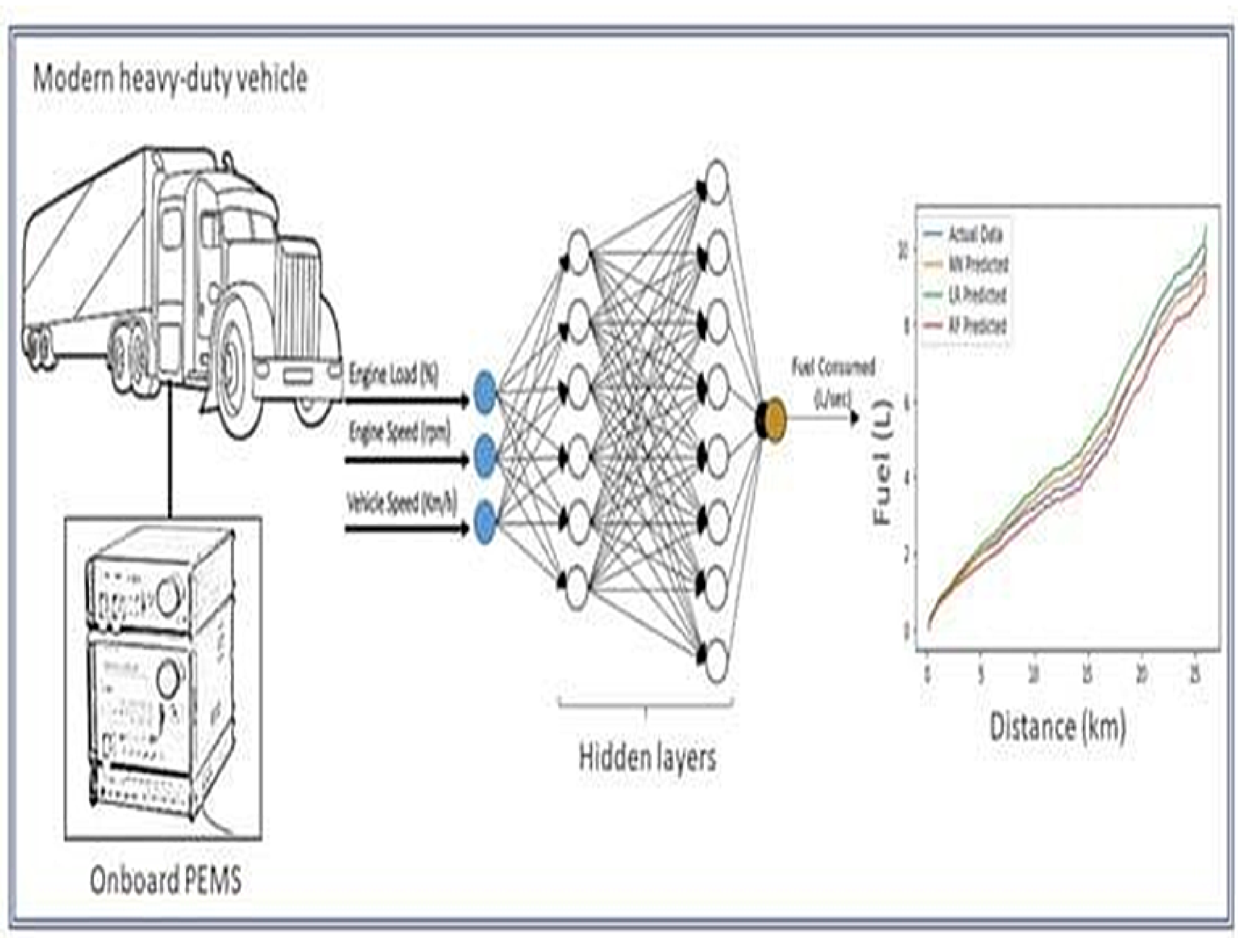
Following are the non-functional requirements of the proposed solution.

| **FR No.** | **Non-Functional Requirement** | **Description** |
|---|---|---|
| NFR-1 | **Usability** | Automation is a significant future of fleet management software. Automator reminders and e-mail notifications enable customer to have a systematic plan for vehicle |

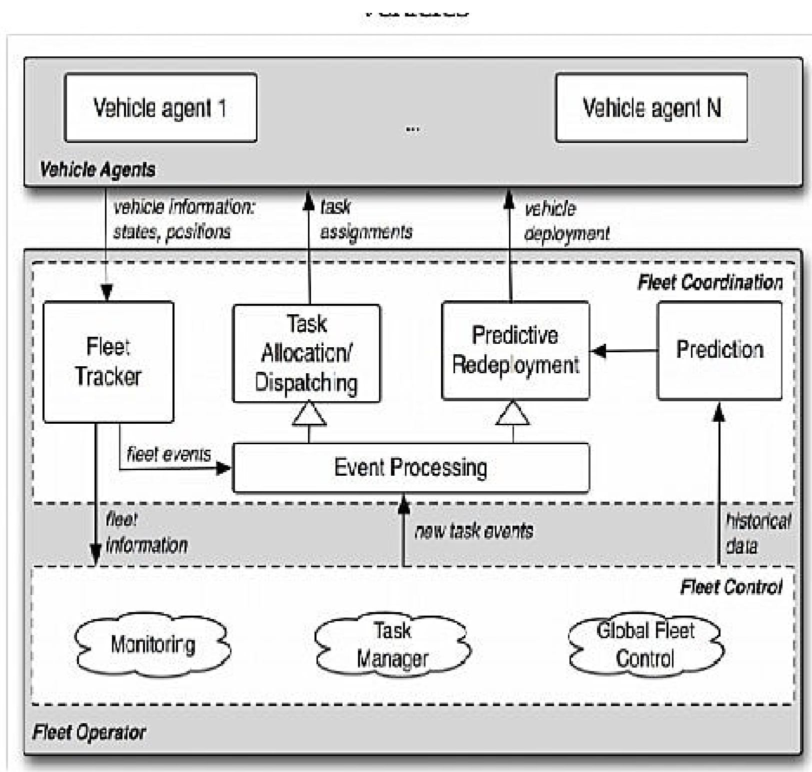
| | | |
|---|---|---|
| | | maintenance, repairs Timely fleet audits. |
| NFR-2 | **Security** | As sustainability rises in social and commercial importance, fleet managers will need to integrate strategies to track and reduce emissions, monitor green performance, as well as managing the transition from traditional vehicle to the nextgeneration. |
| NFR-3 | **Reliability** | The fleet management software not only collectsvaluabledatabutalso provides insights.Withthe<br>help of information collected through GPS trackingsystem. |
| NFR-4 | **Performance** | Fleet managers can access more insights than ever |

# 5. PROJECT DESIGN

## 5.1 DATA FLOW DIAGRAM



5.2 SOLUTION AND TECHNICAL ARCHITECTURE



## 5.3 USER STORIES

- Most customers discover fuel consumption analysis.
- A customer navigates to the fuel monitoring section of our website or app.
- Fuel monitoring section of the website iOS app or android app.
- Fuel prediction of the website iOS app or android app.
- Fuel statistics analysis of the website iOS app.
- Complete modern vehicles analysis in fuel consumption.
- Direct monitoring with the people, and potentially other group vehicles.

# 6. PROJECT PLANNING AND SCHEDULING

## 6.1 SPRINT PLANNING AND ESTIMATION

ProductBacklog,SprintSchedule,andEstimation

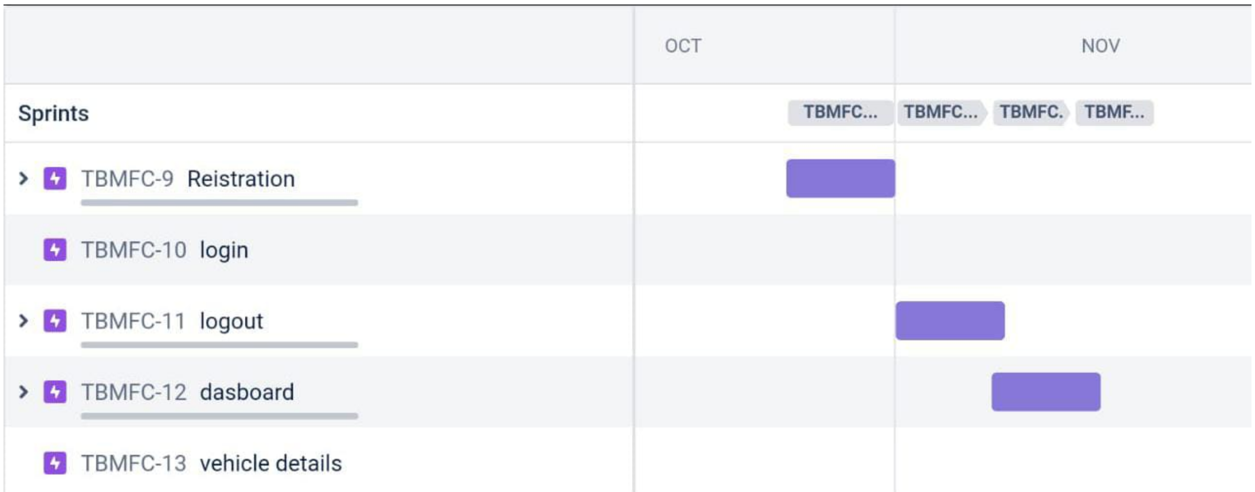
| **Sprint** | **FunctionalRequirement(Epic)** | **UserStoryNumber** | **UserStory/Task** |
|---|---|---|---|
| Sprint-1 | Registration | USN-1 | Asauser,Icanregisterfortheapplicationbyentering mypassword. |
| Sprint-1 | | USN-2 | Asauser, IwillreceiveconfirmationemailonceI |
| Sprint-2 | | USN-3 | Asauser,Icanregisterfortheapplicationthrough |
| Sprint-1 | | USN-4 | Asauser,Icanregisterfortheapplicationthrough |
| Sprint-2 | Login | USN-5 | Asauser,Icanlog intotheapplicationbyentering |
| Sprint-2 | logout | USN-6 | Asauser, Icanlogoutmy application |
| Sprint-3 | Dashboard | USN-7 | Asauser,Ican monitor my vehiclefuelconsum |
| Sprint-4 | Vehicledetails | USN-8 | Asauser, Ican enter myvehicledetails |

## 6.2 SPRINT DELIVERY SCHEDULE

ProductBacklog,SprintSchedule,andEstimation(4Marks)

| **Sprint** | **FunctionalRequirement(Epic)** | **UserStoryNumber** | **UserStory/Task** |
|---|---|---|---|
| Sprint-1 | Registration | USN-1 | Asauser,Icanregisterfortheapplicationbyen ng mypassword. |
| Sprint-1 | | USN-2 | Asauser, Iwillreceiveconfirmationemailon |
| Sprint-2 | | USN-3 | Asauser,Icanregisterfortheapplicationthrou |
| Sprint-1 | | USN-4 | Asauser,Icanregisterfortheapplicationthrou |
| Sprint-2 | Login | USN-5 | Asauser,Icanlog intotheapplicationbyenter |
| Sprint-2 | logout | USN-6 | Asauser, Icanlogoutmy application |
| Sprint-3 | Dashboard | USN-7 | Asauser,Ican monitor my vehiclefuelconsu |
| Sprint-4 | Vehicledetails | USN-8 | Asauser, Ican enter myvehicledetails |

## 6.3 REPORTS FROM JIRA



# 7. CODING AND SOLUTIONING

## 7.1 FEATURE 1

Data collection method such as onboard mission measurement , laboratory measurement, and tunnel study been used in past An on road data collection method using PEMS is increasingly being used, which makes it possible to collected world fuel consumption and emission data and has proved to be reliable.The data used in the current study was collected using a PEMS device mounted on the vehicle during on-road testing data frequency of 1Hz.PEMS software outputs for the sensor ports were used to process second by second data into comma-separated values(CSV) file for each trip. Over 100 parameters such as fuel rate(L/h),engine speed(rpm), speed(km/h), gas temperature, $CO_2$, Nix, GPS altitude, GPS longitude, GPS latitude, etc. were collected for each trip based on data loggers settings. Data were collected from two modern heavy-duty trucks with the same make/model of diesel engine manufactured in Detroit in 2016 were used in this study. The trucks were Cascade models manufactured by Freight liner with DD13engines and used good movement trucks. The on-road tests were performed in California, the fuel consumption during the on-road tests were recorded, and the cumulative fuel consumed per trip was calculated by summing the values. The vehicles were tested for multiple on-road trips with different routes, drivers, and conditions. However, modelling with to many parameters might over fit the ANN model resulting in poor performance. Hence, as ubset of 10 features based on previous studies and domain knowledge was selected. These features included trip number, engine speed(rpm), trip distance(km), vehicle speed(km/h), fuel temperature(◦C), fuel rate(L/s), accelerator pedal position(%), actual torque(Nm), power(kW), and engine load(%). For better modelling of the neural network, the data collected must be representative.

## 7.2 FEATURE 2

–

The last four rows and columns of the correlation matrix indicate that the independentfeatures accelerator pedal position (%), actual torque (ftlb), power (bhp), and engine load(%) are highly correlated to each other with a correlation coefficient of 0.85 and higher.Hence, to prevent over fitting of the model, only engine load (%) of the four parameters was used in modelling. Feature dimension can further be reduced by identifying the highlycorrelated features with the target variable fuel rate (L/s). The recursive feature elimination(RFE) method was used to identify and plot the feature importance scores. RFE determines the features based on the desired withallfeaturesandrecursivelyremovingfeatures.Featureimportanceoftheremainingfeatures,includi ngengineload(%),acceleratorpedalposition(%),fueltemperature(degC), vehicle speed (km/h), trip distance (km), and engine speed (rpm) concerning fuelrate (L/s), was determined with the RFE technique, and the top three features with thehighest scores were selected for modelling. Based on the feature analysis, three independent features, namely engine load (%), vehicle speed (km/h), and engine speed (rpm), will be calculated.

## 7.3 DATABASE SCHEMA

| distance | consume | speed | temp_inside | temp_outside | specials | gas_type | AC |
|---|---|---|---|---|---|---|---|
| 28 | 5 | 26 | 21,5 | 12 | | E10 | 0 |
| 12 | 4,2 | 30 | 21,5 | 13 | | E10 | 0 |
| 11,2 | 5,5 | 38 | 21,5 | 15 | | E10 | 0 |
| 12,9 | 3,9 | 36 | 21,5 | 14 | | E10 | 0 |
| 18,5 | 4,5 | 46 | 21,5 | 15 | | E10 | 0 |
| 8,3 | 6,4 | 50 | 21,5 | 10 | | E10 | 0 |
| 7,8 | 4,4 | 43 | 21,5 | 11 | | E10 | 0 |
| 12,3 | 5 | 40 | 21,5 | 6 | | E10 | 0 |
| 4,9 | 6,4 | 26 | 21,5 | 4 | | E10 | 0 |
| 11,9 | 5,3 | 30 | 21,5 | 9 | | E10 | 0 |
| 12,4 | 5,6 | 42 | 21,5 | 4 | | E10 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 11,8 | 4,6 | 38 | 21,5 | 0 | E10 | 0 |
| 12,3 | 5,9 | 59 | 21,5 | 10 | E10 | 0 |
| 24,7 | 5,1 | 58 | 21,5 | 12 | E10 | 0 |
| 12,4 | 4,7 | 46 | 21,5 | 11 | E10 | 0 |
| 17,3 | 5,1 | 24 | 21,5 | 5 | E10 | 0 |
| 33,4 | 5,6 | 36 | 21,5 | 3 | E10 | 0 |
| 11,8 | 5,1 | 32 | 21,5 | 3 | E10 | 0 |
| 25,9 | 4,9 | 39 | 21,5 | 8 | E10 | 0 |
| 11,8 | 4,7 | 40 | 21,5 | 4 | E10 | 0 |
| 25,3 | 5,5 | 32 | 21,5 | 3 | E10 | 0 |
| 14,2 | 5,9 | 38 | 21,5 | 1 | E10 | 0 |
| 17,9 | 5,7 | 37 | 21,5 | 1 | E10 | 0 |
| 11,8 | 4,7 | 36 | 21,5 | 1 | E10 | 0 |
| 12,3 | 5,9 | 62 | 21,5 | 6 | E10 | 0 |
| 12,4 | 4,1 | 57 | 21,5 | 9 | E10 | 0 |
| 18,4 | 5,7 | 21 | 22,5 | 2 | E10 | 0 |
| 18,4 | 5,8 | 28 | 21,5 | 3 | E10 | 0 |
| 18,3 | 5,5 | 29 | 21,5 | 1 | E10 | 0 |
| 18,4 | 5,7 | 35 | 21,5 | 4 | E10 | 0 |
| 12,3 | 5,3 | 51 | 21,5 | 11 | E10 | 0 |
| 28 | 5 | 26 | 21,5 | 12 | E10 | 0 |
| 12 | 4,2 | 30 | 21,5 | 13 | E10 | 0 |
| 11,2 | 5,5 | 38 | 21,5 | 15 | E10 | 0 |
| 12,9 | 3,9 | 36 | 21,5 | 14 | E10 | 0 |
| 18,5 | 4,5 | 46 | 21,5 | 15 | E10 | 0 |
| 8,3 | 6,4 | 50 | 21,5 | 10 | E10 | 0 |
| 7,8 | 4,4 | 43 | 21,5 | 11 | E10 | 0 |
| 12,3 | 5 | 40 | 21,5 | 6 | E10 | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4,9 | 6,4 | 26 | 21,5 | 4 | E10 | 0 |
| 11,9 | 5,3 | 30 | 21,5 | 9 | E10 | 0 |
| 12,4 | 5,6 | 42 | 21,5 | 4 | E10 | 0 |
| 11,8 | 4,6 | 38 | 21,5 | 0 | E10 | 0 |
| 12,3 | 5,9 | 59 | 21,5 | 10 | E10 | 0 |
| 24,7 | 5,1 | 58 | 21,5 | 12 | E10 | 0 |
| 12,4 | 4,7 | 46 | 21,5 | 11 | E10 | 0 |
| 17,3 | 5,1 | 24 | 21,5 | 5 | E10 | 0 |
| 33,4 | 5,6 | 36 | 21,5 | 3 | E10 | 0 |
| 11,8 | 5,1 | 32 | 21,5 | 3 | E10 | 0 |
| 25,9 | 4,9 | 39 | 21,5 | 8 | E10 | 0 |
| 11,8 | 4,7 | 40 | 21,5 | 4 | E10 | 0 |
| 25,3 | 5,5 | 32 | 21,5 | 3 | E10 | 0 |
| 14,2 | 5,9 | 38 | 21,5 | 1 | E10 | 0 |
| 17,9 | 5,7 | 37 | 21,5 | 1 | E10 | 0 |
| 11,8 | 4,7 | 36 | 21,5 | 1 | E10 | 0 |
| 12,3 | 5,9 | 62 | 21,5 | 6 | E10 | 0 |
| 12,4 | 4,1 | 57 | 21,5 | 9 | E10 | 0 |
| 18,4 | 5,7 | 21 | 22,5 | 2 | E10 | 0 |
| 18,4 | 5,8 | 28 | 21,5 | 3 | E10 | 0 |

# 8. TESTING

## 8.1 TEST CASE

In machine learning systems, however, data and desired behaviour are the inputs and the models learn the logic as the outcome of the training and optimization processes. In this case, testing involves validating the consistency of the model’s logic and our desired behaviour. Due to the process of models learning the logic, there are some notable obstacles in the way of testing Machine Learning systems. They are:

- **Indeterminate outcomes**: on retraining, it’s highly possible that the model parameters vary significantly.
- **Generalization**: it’s a huge task for Machine Learning models to predict sensible outcomes for data not encountered in their training.
- **Coverage**: there is no set method of determining test coverage for a Machine Learning model.
- **Interpretability**: most ML models are black boxes and don’t have a comprehensible logic for a certain decision made during prediction.

## 8.2 USER ACCEPTANCE TESTING

The modelling structure will be similar to the traditional four-stage Travel Demand Forecast Model (TDFM) for the requirements of this study area. Traffic Analysis Zones (TAZs) contains 15 internal zones and 8 external zones, but Traffic Analysis Districts (TADs) contain 157 internal zones. Demography observation and trip's survey were estimated with data from the Shiraz city; centre of Fars province, Iran country, 1999 Sharif university data, and travel observation data has been collected (2009). Each individual TAZ is divided to about 10 districts based on road's boundaries, which has the close equivalency factor for contained districts. All the statistics and demographic data controlled and tested as well as both zonal layers. The main advantage of this method is that it can provide the opportunity to consider more concentrated centroids compared to just 15 centroids in TAZs, furthermore, implementing this type of model, precise identification of different zonal applications (residential, industrial, public and private) can be easily feasible. On the other hand, one of the evident disadvantages of this type of modelling is that considering the excessive number of attributes can cause increasing the time, effort and money. This model is essentially a conventional four-stage model with a number of additional sub models. Fuel price and socio-economic conditions affect trip generation. Fuel consumption shows the passing distance to reach a location and intensity of the activity, and socio-economic factors determine the magnitude and extent of population activity. Author developed transportation networks via ArcGIS9.2 & ArcGIS9.3 software and travel demand models are developed by using CUBE5 and TransCAD4.5 transportation planning software packages in GIS platform. All these software packages are implemented at the Transportation System Engineering (TSE) laboratory with the license purchased by the Indian Institute of Technology Bombay. The methodology used for planning travel demand model and its application for evaluating appropriate land use and sustainability of the model can be explained in following steps: 1. Generation and Creation of network for case study and define all necessary attributes. 2. Define all characteristics and attribute of network for travel demand model generation. 3. Generation and Validation of base year OD Matrices. 4. Development and Calibration of Travel Demand Model. 5. Model Application. 6. Fuel Indices and Evaluation of Alternatives. 2.1. Base year OD matrix generation This process starts with the last 15 years Home Interview Survey (HIS) data as input. HIS obtained sample size magnitude equal to 4.3% of the number of households, i.e. to select one sample out of 23 households(literature standards mentioned that for more than 1 million population cities, it could be more than 4%sample size magnitude).

## 9. RESULTS

### 9.1 PERFORMANCE METRICS

This study presents the fuel consumption modelling for modern heavy-duty vehicles using PEMS data under various driving conditions, different routes, and external factors. Engine Load (%), Engine Speed (rpm), and Vehicle Speed (km/h) were used as inputs for the ANN. Based on the hyper-parameter tuning, the neural network was trained for100 epochs with a learning rate of 0.00001.During each epoch, the loss for each dataitem/batch in the training dataset and validation dataset was calculated. The loss plots show indicate the mean absolute error(MA) The performance of the machine learning model for the regression problem was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R- Squared ($R2$).The research object of Wickramanayake and Bandera is the fuel consumption prediction of the bus, and this study focuses on the fuel consumption of the taxicabs. At the same time, the driving behaviuor data of this study are collected from mobile phones with higher flexibility and complexity rather than a fixed GPS device. This method could predict vehicle fuel consumption with high accuracy and efficiency based on cell phone data and provide strong support for traffic management departments to monitor the ecological levels of driving behaviuor of taxi drivers.

## 10. ADVANTAGES

It can save you money

Driving a fuel-efficient car reduces the running costs you will have to pay throughout the lifetime of your vehicle. There are many different savings associated with fuel efficiency, and together they add up to a significant amount of money.

Spend less on fuel

If your car is fuel efficient and goes further per litre of petrol or diesel, your fuel costs will be significantly reduced. Diesel engines in particular are more frugal than ever and can cover distances of more than 80 miles on one gallon of fuel.

As well as saving you money, this also means that you do not have to stop so often to fill up the tank. As the driver of a fuel-efficient model you can spend less time standing on chilly fuel station forecourts and more time enjoying your car’s performance from behind the wheel.

## DISADVANTAGES

### Gasoline use contributes to air pollution

The vapours given off when gasoline evaporates and the substances produced when gasoline is burned (carbon monoxide, nitrogen oxides, particulate matter, and unburned hydrocarbons) contribute to air pollution. Burning gasoline also produces carbon dioxide, a greenhouse gas.

## 11. CONCLUSION

In conclusion, the study demonstrates the modelling of fuel consumption in modern heavy-duty vehicles with an artificial neural network using very few technical parameters . An attempt was made to develop a model using very few parameters collected under different conditions. Data from modern heavy-duty trucks with the same make and model, driven by different persons on various routes under different external conditions, were used for training the artificial neural network. The model relies on very few parameters that couldbeobtainedquicklyandeasilyfromavehicleduringatrip,unlikeotherparameterssuch as road grade, latitude, longitude, traffic information, etc. Moreover, the three parameters used were able tocaptureaminimumof78%ofthevarianceinthe fuel rate compared to other studies where many parameters are used. Adding more input parameters would improve the performance of ANN, but collecting such data might require additional equipment setup. The performance measures MAE, RMSE, and $R2$indicate thataccurate prediction can be obtained with the model. The data modelling can help to identifythe trend in instantaneous fuel consumption and to calculate the total fuel consumed by the vehicle for each trip, which can further help in diagnosing vehicle performance in the case of abnormalities. Models that are accurate, fast, and able to predict in real-time willenable the optimization of fuel model more complex data from other vehicles with different makes and models that do not have the amount on-road data needed to train a network. This work can be extended to include other factors such as time, traffic information, road information, GPS data, etc. that affect fuel consumption, and to estimate vehicle exhausted missions.

## 12. FUTURE SCOPE

In this study, driving behaviuor data and fuel consumption data of taxi drivers collected from OBD and mobile phone terminals, respectively, were matched. The correlation between driving behaviuor and fuel consumption was analyzed, and relevant driving behaviuor indicators affecting fuel consumption were extracted through the filter-based feature selection method. Using the seven selected driving behaviuor indicators (namely, average speed, ASEI, average acceleration, average deceleration, acceleration time percentage, deceleration time percentage, and cruising time percentage), three fuel consumption prediction models based on a BP neural network, SVR, and a random forest were constructed. The results of model error and the run time comparison analysis show that the three models could predict fuel consumption accurately, and the random forest model had the highest accuracy and efficiency, with an RMSE of 0.783 L/100 km, mean absolute percentage error (K) of 6.9%, and model running time of 0.14 s. This finding is consistent with the research of Wickramanayake and Banderawhich also shows that random forest models are most effective in predicting fuel consumption based on driving behaviuor data. The research object of Wickramanayake and Bandera is the fuel consumption prediction of the bus, and this study focuses on the fuel consumption of the taxicabs. At the same time, the driving behaviuor data of this study are collected from mobile phones with higher flexibility and complexity rather than a fixed GPS device. This method could predict vehicle fuel consumption with high accuracy and efficiency based on cell phone data and provide strong support for traffic management departments to monitor the ecological levels of driving behaviuor of taxi drivers.

# 13. APPENDIX

jupyter model building Last Checkpoint: 11/03/2022 (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3 (ipykernel)

Run Code

## build our model

In [15]:
```
x=pd1
```

In [21]:
```
y=pd2
```

In [22]:
```
from sklearn.model_selection import train_test_split
```

In [26]:
```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2)
```

## knn model

In [27]:
```
from sklearn.neighbors import KNeighborsClassifier
```

In [29]:
```
knn=KNeighborsClassifier(n_neighbors=5)
```

In [ ]:
```
pred=knn.predict(x_test)
```

In [ ]:
```
from sklearn.metrices import accuracy_score,confusion_matrix
```

In [ ]:
```
accuracy_score(y_test,pred)
```

In [ ]:
```
confusion_matrix(y_test,pred)
```

jupyter fuel consumption prediction Last Checkpoint: a few seconds ago (autosaved)

Logout

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3 (ipykernel)

Run Markdown

In [30]:
```
data_main.isnull().sum()
```

Out[30]:
```
distance                   0
consume                    0
speed                      0
temp_inside                0
temp_outside               0
gas_type                   0
AC                         0
rain                       0
sun                        0
refill liters              0
refill gas                 0
specials_AC                0
specials_AC Sun            0
specials_AC rain           0
specials_AC snow           0
specials_AC sun            0
specials_ac                0
specials_ac rain           0
specials_half rain half sun 0
specials_rain              0
specials_snow              0
specials_sun               0
specials_sun ac            0
dtype: int64
```

In [31]:
```
data_main.isna().sum().sum()
```

Out[31]:
```
0
```

## preprocessing the data done

SOURCE CODE

```
import pandas as pd
import numpy as np
import seaborn as sns
data=pd.read_csv("fuel consumption.csv")
data.info()
data.head()
data.tail()
data.isna().sum()
data.isna().sum().sum()
data.isna()
data[data['temp_inside'].isna()]
data.shape
remove=data.dropna()
data.mean()
data['gas_type'].replace({'E10':10,'SP98':98},inplace=True)
data.head()
data['refill gas'].replace({'E10':10,'SP98':98},inplace=True)
data.head()
data_main=pd.get_dummies(data,columns=['specials'])
data_main['temp_inside'].mode()
data_main['temp_inside'].fillna(data_main['temp_inside'].mode(),inplace=Tru
data_main['refill liters'].fillna(data_main['refill liters'].mean(),inplace=True)
data_main['refill gas'].fillna(data_main['refill gas'].mean(),inplace=True)
data_main.head(10)
data_main.isnull().sum()
data_main.isna().sum().sum()
```

GITHUB AND PROJECT DEMO LINK

GitHub link:

https://github.com/IBM-EPBL/IBM-Project-40937-166037537

Demo video link:

https://youtu.be/dMMaFCDzyVg