MACHINE LEARNING BASED VECHICLE PERFORMANCE ANALYZER

NALAIYA THIRAN PROJECT BASED LEARNING
On

PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP

A PROJECT REPORT

TEAM ID: PNT2022TMID10381

KANNAN M	720819104050
ANNAMALAI R	720819104008
BHARATHI JR	720819104017
KARTHICK RAJA P	720819104051

BACHELOR OF ENGINEERING

IN
COMPUTER SCIENCE ENGINEERING

HINDUSTHAN INSTITUTE OF TECHOLOGY

Approved by AICTE, New Delhi, Accredited with 'A' Grade by NAAC

(Affiliated to Anna University, Chennai)

COIMBATORE – 641 032

November 2022

TABLE OF CONTENTS

1.INTRODUCTION

1.1 Project Overview

Forest fires are a major environmental issue, creating economic and ecological damage while endangering human lives. There are typically about 100,000 wildfires in the United States every year. Over 9 million acres of land have been destroyed due to treacherous wildfires. It is difficult to predict and detect Forest Fire in a sparsely populated forest area and it is more difficult if the prediction is done using ground-based methods like Camera or Video-Based approach. Satellites can be an important source of data prior to and also during the Fire due to its reliability and efficiency. The various real-time forest fire detection and prediction approaches, with the goal of informing the local fire authorities.

Automotive Technologies are providing improvised services to the driver's safety and vehicle security under the umbrella of Intelligent Transportation System (ITS). In the development of ITS, advanced Automotive Technologies shall play a crucial role in determining the overall experience of users by making it much at ease in terms of reducing the risk of road accidents, risk of cybercrime in the vehicle, buying a used car etc. It is often noted that judging the driver's driving skill is subjective and is difficult to set a standard for driver's skills. The modern approach to transportation system is focusing on rapidly evolving with the intelligent vehicles. High rise in recorded traffic density, road accidents and crisis faced in regulating the effective management of traffic control in urban and rural

areas have concerned us to develop a smart solution in context to ITS. The automotive industry has great expectations from these futuristic solutions to improve the safety of people and security of vehicles. It is observed that the users are shifting from individualistic approach to the data-centric approach based on OBD-II scanner to avail the augmented driving experience. In spite of the modern command, control, communication, computers and intelligent systems, we are still facing numerous calamities in which thousands of precious human lives are lost in accidents. Therefore, it should be an immediate need to tackle the small scale yet serious issues using the state-of-the-art techniques. We are mainly focusing on analyzing the data which is collected from the vehicle using the OBD-II scanner and eventually providing the driver's safety solutions. We aim to obtain the solutions by observing the blind-spots accurately and efficiently using pattern recognition techniques from supervised learning.

2. LITERATURE SURVEY

Singh D, Singh M., "Internet of Vehicles for Smart and Safe Driving",International Conference on Connected Vehicles and Expo (ICCVE),Shenzhen, 19 -23 Oct.,2015. (This paper has discussed about smarttransportation services in cloud (Cloud-STS) for safety and convenience. STSprovide driver centric board services in the cloud networks. STS composed ofVehicle to WiFi networks (VtoWiFi), Vehicle to Cloud Network (VtoCN),Vehicle to Vehicle (VtoV), and Cloud Network to service provider (CNtoSP).The idea is to utilize the (WiFi enabled) Smart Highways and 3Dcameraenabled dash board

navigation device to enhance accident prevention /monitoring and control.) Zhang, Y., Lin, W., and Chin, Y., "Data -Driven Driving Skill Characterization:Algorithm Comparison and Decision Fusion," SAE Technical Paper2009 -01 -1286, 2009, https://doi.org/10.4271/2009 -01 -1286.Azevedo, C. LCardoso. (By adapting vehicle control systems to the skill level of the driver,the overall vehicle active safety provided to the driver can be furtherenhanced for the existing active vehicle controls, such as ABS, Traction Control, Vehicle Stability Enhancement Systems. As a follow-up to thefeasibility study in, this paper provides some recent results on data-drivendriving skill characterization. In particular, the paper presents anenhancement of discriminant features, the comparison of three different learning algorithms for recognizer design, and the performanceenhancement with decision fusion. The paper concludes with the discussionsof the experimental results and some of the future work.)

- J. E. Meseguer, C. T. Calafate, J. C. Cano and P. Manzoni, "DrivingStyles: Asmartphone application to assess driver behavior," 2013 IEEE Symposium onComputers and Communications (ISCC), Split,2013, pp.000535 -000540. oi:10.1109/ISCC.2013.6755001.(The DrivingStyles architecture integrates bothdata mining techniques and neural networks to generate a classification ofdriving styles by analyzing the driver behavior along each route. In particular,based on parameters such as speed, acceleration, and revolutions per minuteof the engine (rpm), we have implemented a neural network based algorithmthat is able to characterize the type of road on which the vehicle is moving,as well as the degree of aggressiveness of each

driver. The final goal is toassist drivers at correcting the bad habits in their driving behavior, whileoffering helpful tips to improve fuel economy. In this work we takeadvantage of two key-points: the evolution of mobile terminals and theavailability of a standard interface to access car data.)

Kenneth L. Clarkson. 1985. Algorithms for Closest -Point Problems(Computational Geometry). Ph.D. Dissertation. Stanford University, Palo Alto,CA. UMI Order Number: AAT 8506171.(This dissertation reports a variety ofnew algorithms for solving closest-point problems. The input to thesealgorithms is a set or sets of points in d-dimensional space, with an associated L (, p) metric. The problems considered are:(1) The all nearestneighbors problem. For point set A, find the nearest neighbors in A of eachpoint in A.(2) The nearest foreign neighbor problem. For point sets A and B,find the closest point in B to each point in A. The geometric inimumspanning tree problem.)Goszczynska H., Kowalczyk L., Kuraszkiewicz B. (2014) Correlation Matrices asa ool to Analyze the Variability of EEG Maps. In: Piętka E., Kawa J., WieclawekW. (eds) InformationTechnologies in Biomedicine, Volume 4. Advances inIntelligent Systems and Computing, vol 284. Springer.(The aim of this paperis to present the selected examples of possible applications of image ofcorrelation coefficients matrix of EEG map series in the analysis of variation ofthe topography of the isopotential areas in EEG maps, and thus in the assessment of stationarity, spatio-temporal variability and trends of changesof bioelectric activity of the brain. The image of correlation coefficients matrix shows similiarity of all pairs of maps in a series.

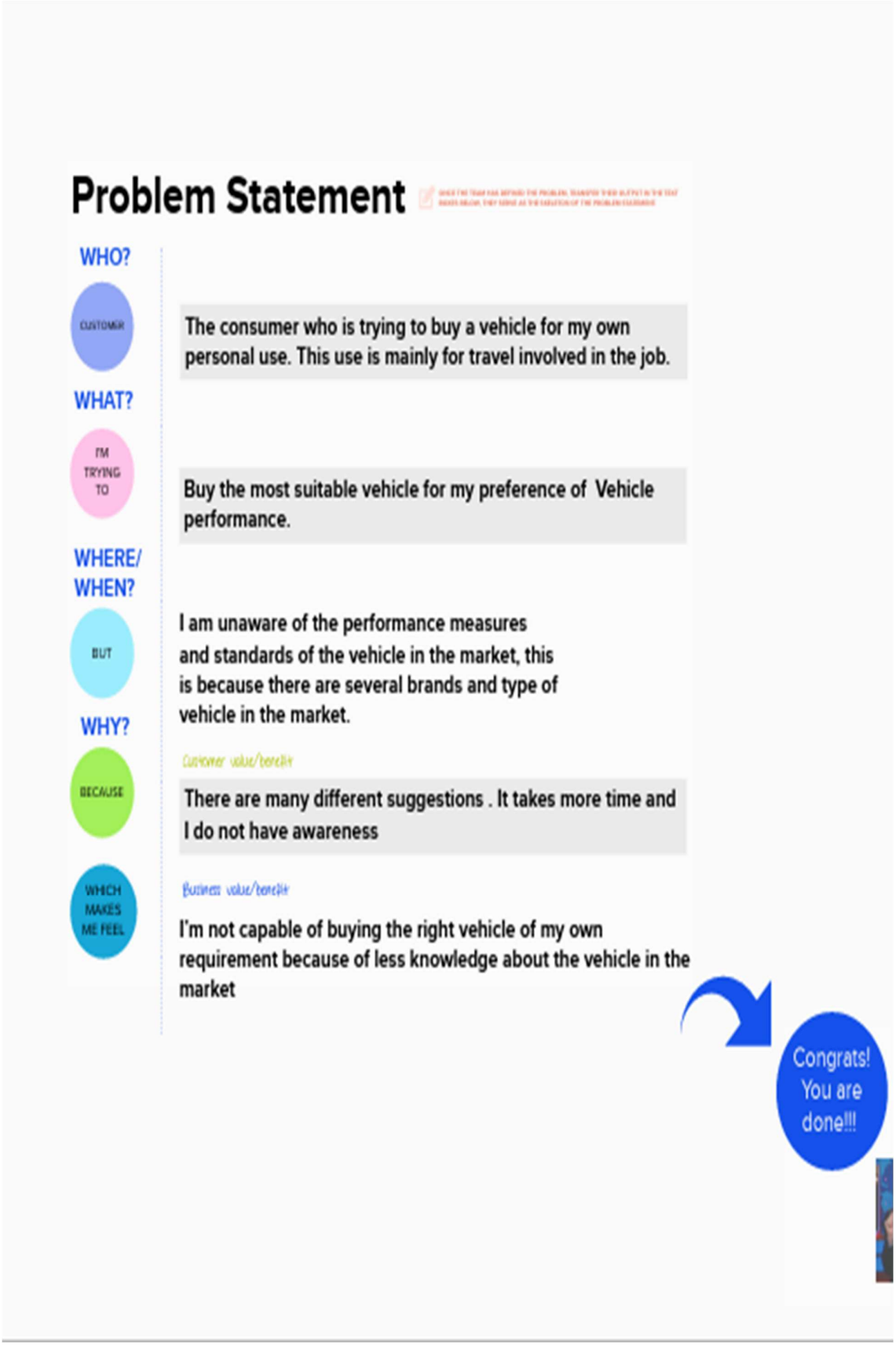
The choice of segmentationthreshold of characteristic areas in images of the correlation coefficientsmatrix of EEG map series corresponding to the sequence similarityrelationships in a series of maps was based on the results of researchconducted on test series.)

2.2 References

- Singh D, Singh M., "Internet of Vehicles for Smart and Safe Driving", International Conference on Connected Vehicles and Expo (ICCVE), Shenzhen, 19-23 Oct., 2015.
- Zhang, Y., Lin, W., and Chin, Y., "Data-Driven Driving Skill Characterization:
Algorithm Comparison and Decision Fusion," SAE Technical Paper 2009 - 01-1286, 2009, https://doi.org/10.4271/2009-01-1286.Azevedo, C. L Cardoso.
- J. E. Meseguer, C. T. Calafate, J. C. Cano and P. Manzoni, "DrivingStyles: Asmartphone application to assess driver behavior,"2013 IEEE Symposium on Computers and Communications (ISCC), Split, 2013, pp.000535- 000540. doi: 10.1109/ISCC.2013.6755001.
- Schneider, A., Hommel, G., & Blettner, M. (2010). Linear Regression Analysis:

Part 14 of a Series on Evaluation of Scientific Publications. Deutsches Ärzteblatt International, 107(44), pp. 776–782.

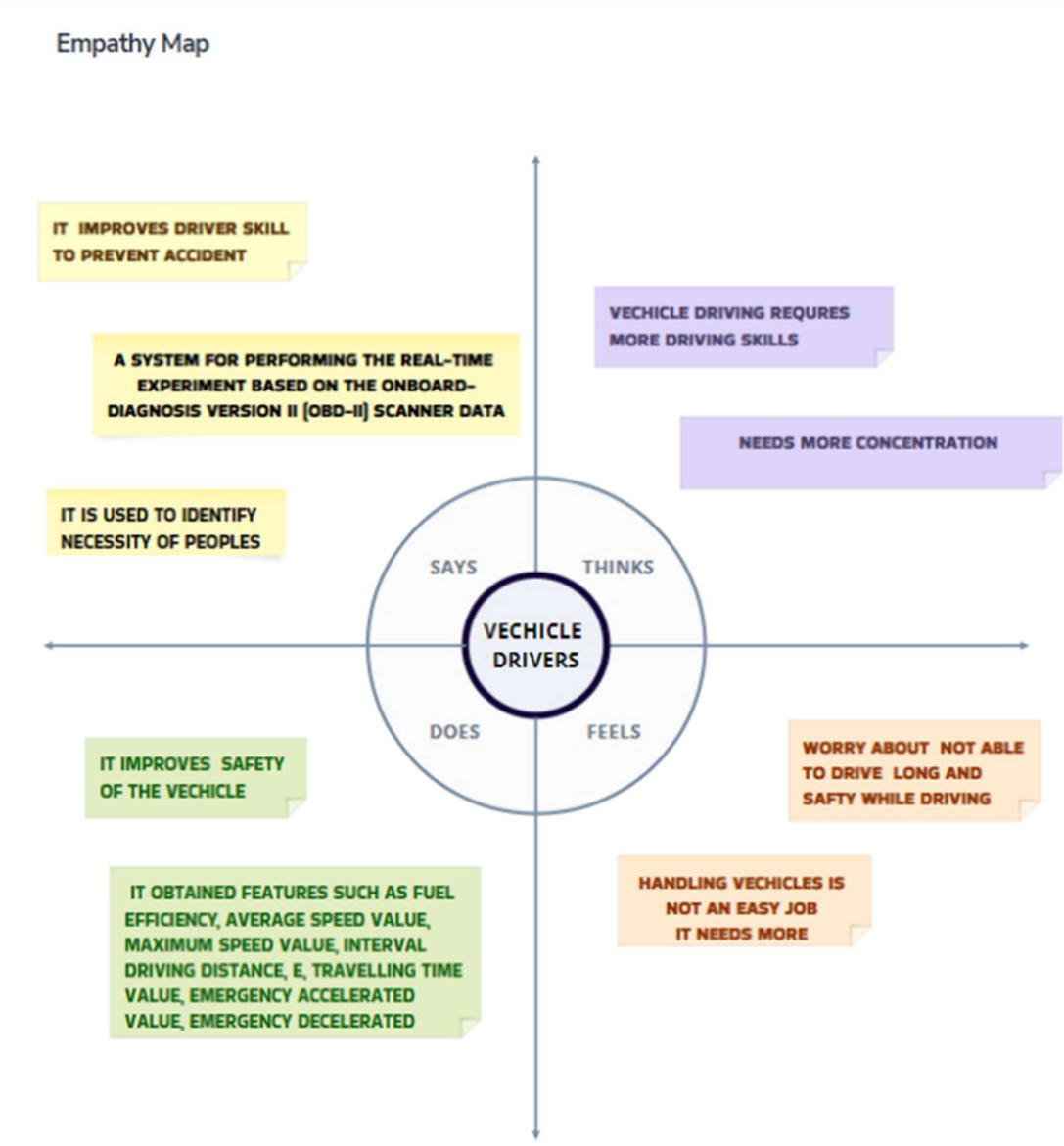
2.3 Problem Statement Definition



3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

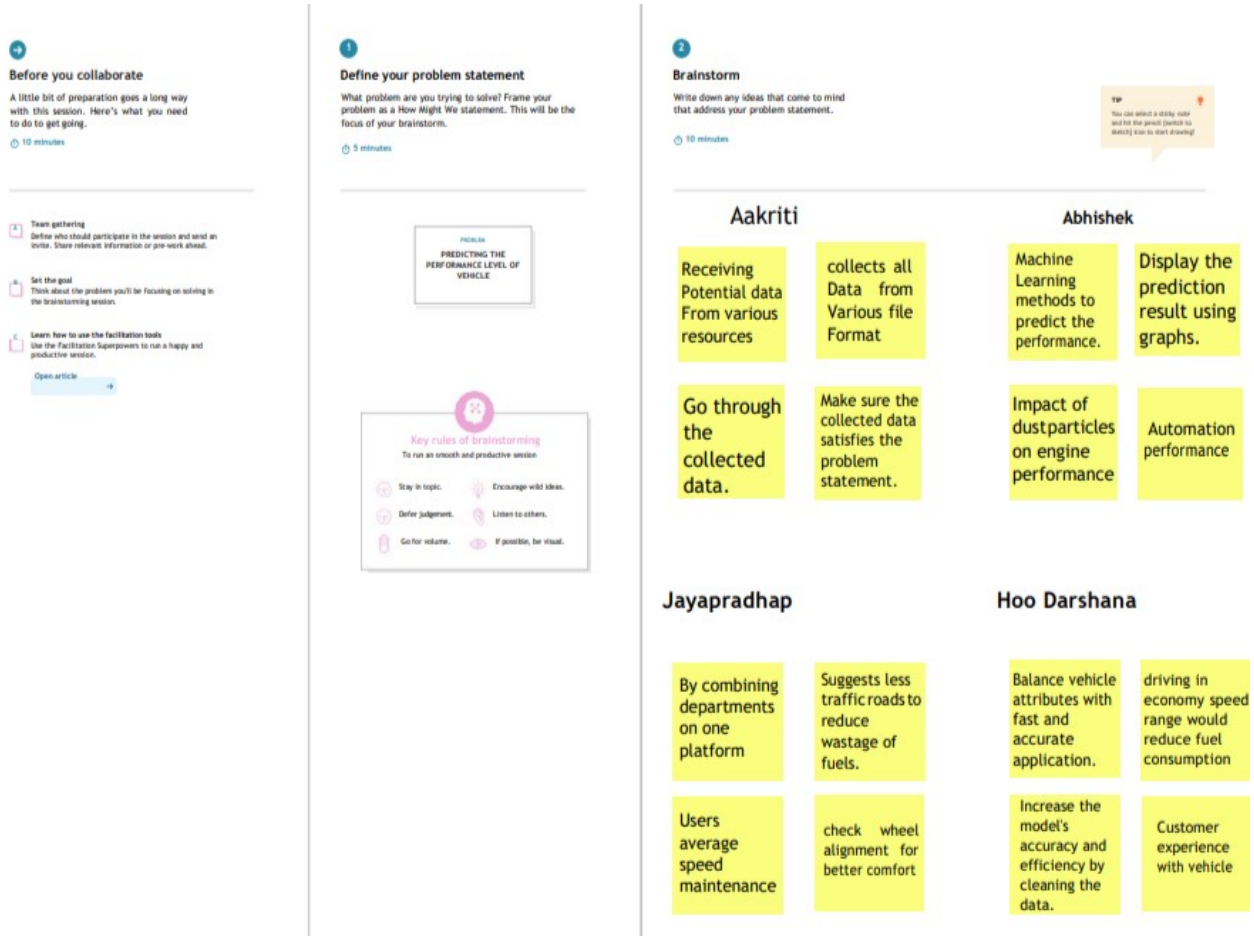
An empathy map is a collaborative tool teams can use to gain a deeper insight into their customers. Much like a user persona, an empathy map can represent a group of users, such as a customer segment.



3.2 Ideation & Brainstorming

Brain storming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Use this template in your own brain storming sessions so your team can unleash the irimagination and start shaping concepts even if you're not sitting in the same room.

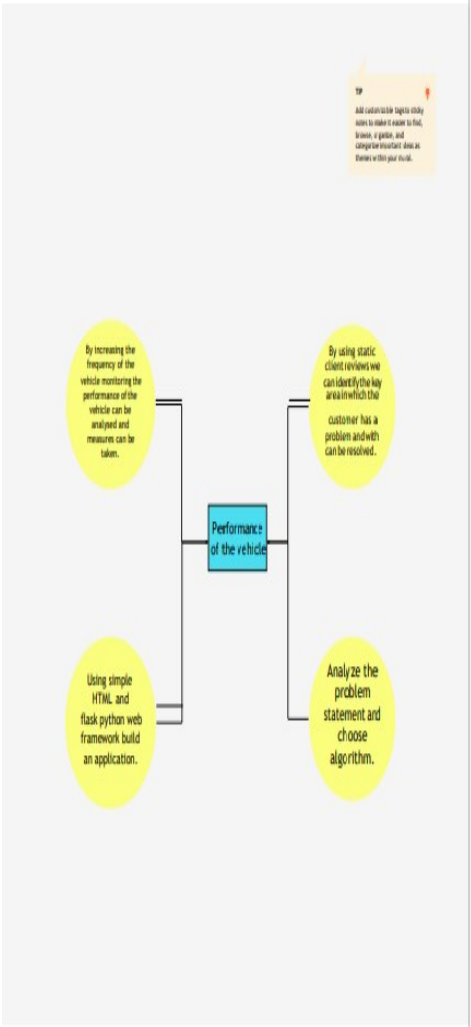


3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes

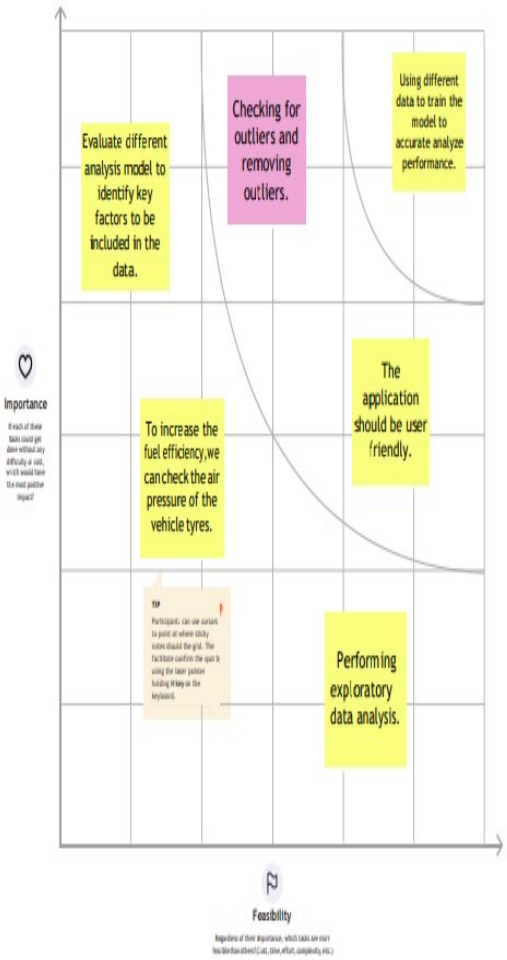


4

Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

20 minutes



After you collaborate

You can export the mural as an image or pdf to share with members of your company who might find it helpful.

Quick add-ons

Share the mural
Share a view link to the mural with stakeholders to keep them in the loop about the outcomes of the session.

Export the mural
Export a copy of the mural as a PNG or PDF to attach to emails, include in slides, or save in your drive.

Keep moving forward

Strategy blueprint
Define the components of a new idea or strategy.
Open the template

Customer experience journey map
Understand customer needs, motivations, and obstacles for an experience.
Open the template

Strengths, weaknesses, opportunities & threats
Identify strengths, weaknesses, opportunities, and threats (SWOT) to develop a plan.
Open the template

Share template feedback

3.3 Proposed Solution

S.No.	Parameter	Description
1.	Problem Statement	The user is an Automobile manufacturer which is Looking for new technologies to analyse the performance of its vehicles.
2.	Idea / Solution description	The variables that are measured across different parts of the car and its performance can be used to for ecast the condition of the vehicle. This continual process of collecting, studying, analysing, and documenting vehicle's performance can be used to Gain competitive advantage.
3.	Novelty /Uniqueness	Giving the public and the manufacturer the feature to analyse their vehicle's performance.
4.	Social Impact/Customer Satisfaction	The product improves vehicles performance, increases mileage, efficiency, lifetime, etc. These Properties will have a positive response.
5.	Business Model	Selling the product as an application and/or selling through collaboration with vehicle manufacturers. Product canal so be sold to Racing companies for Them to dominate in their field.

3.4 Problem Solution Fit:

4.REQUIREMENT ANALYSIS

4.1 Functional requirement

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement(Epic)	Sub Requirement (Story/Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Reset Password	ResetpasswordthroughGmail Reset password through Mobile number
FR-4	Feedback	The user can submit the feedback through a contact form in the website or through Gmail.

4.2 Non-Functional requirements

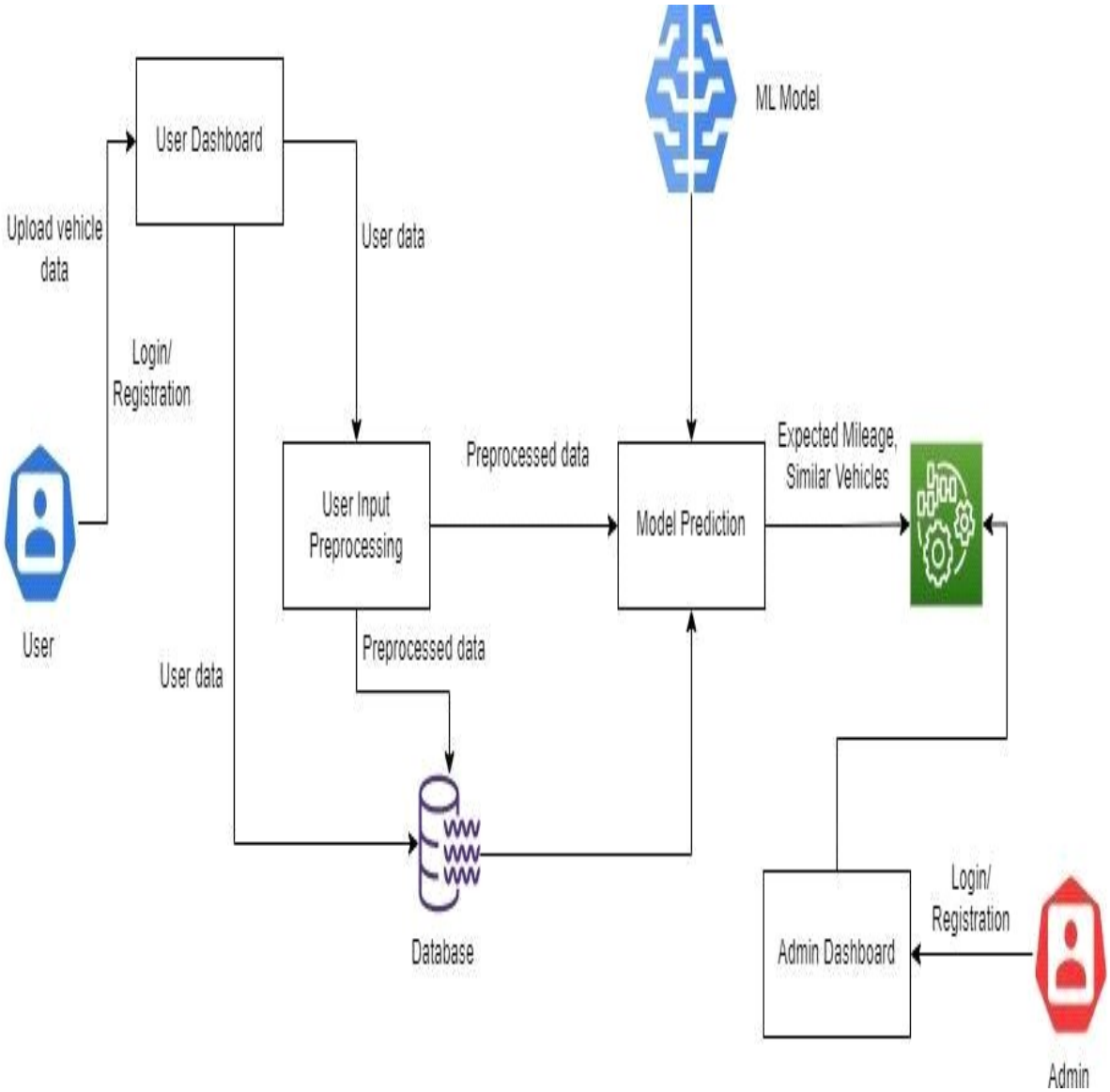
Following are the non-functional requirements of the proposed solution.

FR No.	Non-FunctionalRequirement	Description
NFR-1	Usability	The analyser allows the user to improve performance based on the results provided. It is easy to use with just the data required.
NFR-2	Security	The security is improved by using vehicle alarm, wheel lock, vehicle lock and also GPS tracker.
NFR-3	Reliability	The reliability rating is good due to best performance, less frequency of problem occurrence and cost for repairing is low.
NFR-4	Performance	Thevehicleisupgradedintheirqualityandinfrastructure to provide better performance likegoodmileage, smooth travel.
NFR-5	Availability	The data required is collected by research personsandthisdata canbe usedtoprovidebetterresults.

5. PROJECT DESIGN

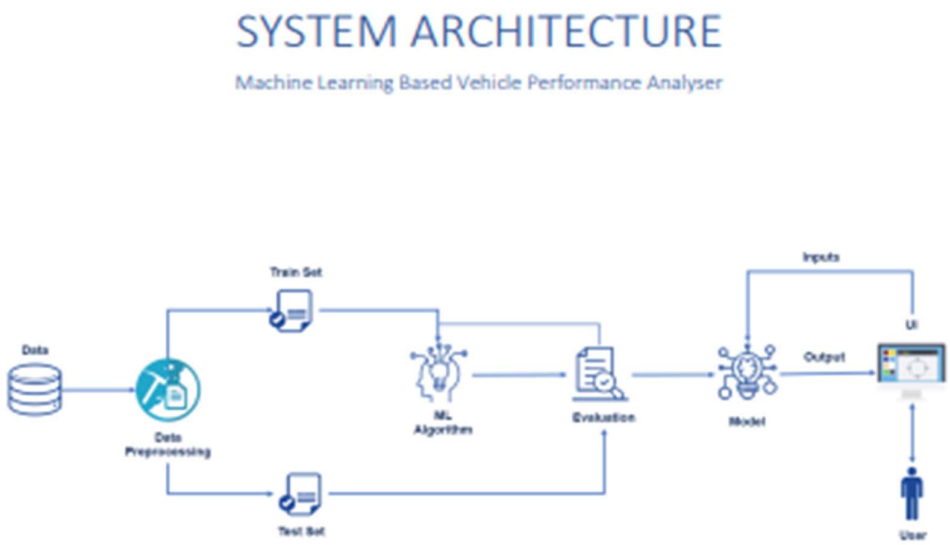
5.1 Data Flow Diagrams

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.



5.2 Solution & Technical Architecture

A solution architecture (SA) is an architectural description of a specific solution. Solution Architectures combine guidance from different enterprise architecture viewpoints (business, information and technical), as well as from the enterprise solution architecture (ESA).



5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Google OAUTH		Medium	Sprint-1
	Login	USN-4	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-5	As a user, I can upload vehicle data and get an analysis of the data		Medium	Sprint-2
Administrator	Login	USN-6	As an admin, I can log into the application by entering email & password	I can access my admin dashboard	High	Sprint-1

6.PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Product Backlog, Sprint Schedule, and Estimation

Use the below template to create product backlog and sprint schedule

Project tracker:

Sprint	Total story point	Duration	Sprint start date	Sprint end date	Story points completed	Sprint released ate
Sprint-1	20	6days	27-oct-2022	28-oct-2022	20	04-nov-2022
Sprint-2	20	6days	02-nov-2022	05-nov-2022	20	07-nov-2022
Sprint-3	20	6days	08-nov-2022	12-nov-2022	20	12-nov-2022
Sprint-4	20	6days	14-nov-2022	19-nov-2022	20	19-nov-2022

Velocity:

average velocity=80/20=4 story points per day

7.CODING & SOLUTIONING

7.1 Feature 1

In [5]:
```
import numpy as np
import pandas as pd
```

In [6]:
```
#Reading The Dataset

datas = pd.read_csv(r"C:\Users\sunda\Desktop\IBM\car performance.csv")
datas.head()
```

Out[6]:

	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin	car name
0	18.0	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165	3693	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150	3436	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150	3433	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140	3449	10.5	70	1	ford torino

In [7]:
```
# Handling Missing Values

datas.isnull().any()
```

Out[7]:
```
mpg             False
cylinders       False
displacement    False
horsepower      False
weight          False
acceleration    False
model year      False
origin          False
car name        False
dtype: bool
```

In [8]:
```
# Splitting The Dataset Into Dependent And Independent Variable.

x = datas.iloc[:,1:8].values
```

In [9]:
```
y = datas.iloc[:,0].values
```

In [10]:
```
# Split The Dataset Into Train Set And Test Set

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.2,random_state=0)
```

In [11]:
```
# Normalizing

from sklearn.preprocessing import StandardScaler
sd = StandardScaler()
x_train = sd.fit_transform(x_train)
x_test = sd.fit_transform(x_test)
```

In [12]:
```
x_train
```

In [12]:
```
x_train
```

Out[12]:
```
array([[ 1.49526939,  1.22961301,  1.24359144, ..., -0.79520768,
        -1.13752513, -0.73301171],
       [-0.85285735, -0.92367663, -1.16092059, ...,  1.24411524,
        -1.41177304,  0.5068698 ],
       [-0.85285735, -0.92367663, -0.68001818, ...,  0.05760009,
         1.05645814,  0.5068698 ],
       ...,
       [-0.85285735, -1.206235  , -1.45480539, ...,  1.42950823,
        -0.86327722,  0.5068698 ],
       [ 0.32120602,  0.56706235, -0.09224857, ..., -0.2390287 ,
        -1.41177304, -0.73301171],
       [-0.85285735, -0.99188037, -0.86703579, ...,  0.31715028,
        -0.31478141,  0.5068698 ]])
```

In [13]:
```
# Build The Model With The Random Forest Regressor

from sklearn.ensemble import RandomForestRegressor
d = RandomForestRegressor (n_estimators=30,random_state = 0)
d.fit(x_train,y_train)
```

Out[13]:
```
RandomForestRegressor(n_estimators=30, random_state=0)
```

In [14]:
```
# prediction

y_pred = d.predict(x_test)
y_pred
```

Out[14]:
```
array([14.38333333, 24.25666667, 14.21666667, 20.56666667, 18.47333333,
       30.21666667, 34.63333333, 21.15      , 16.30333333, 25.76      ,
       36.60333333, 36.27      , 19.53666667, 27.32333333, 16.54333333,
       32.99333333, 28.32333333, 27.49666667, 17.03      , 35.82      ,
       16.47333333, 23.54      , 23.16666667, 20.7       , 33.69666667,
       26.45      , 33.79666667, 30.37333333, 31.93666667, 16.57333333,
       20.26666667, 32.99      , 19.79666667, 34.08333333, 20.85666667,
       25.02      , 19.65333333, 17.14      , 34.78333333, 12.76666667,
       13.73333333, 15.2       , 28.32      , 32.76666667, 28.74333333,
       22.68666667, 20.54333333, 16.50666667, 23.38      , 29.88333333,
       34.31666667, 26.5       , 17.63      , 27.78333333, 15.96666667,
       12.96666667, 18.86666667, 26.91666667, 31.95666667, 15.68      ,
       20.81      , 25.97      , 19.84666667, 21.6       , 13.46666667,
       15.33333333, 14.2       , 18.90333333, 24.72666667, 14.21666667,
       34.87666667, 13.25      , 22.96666667, 18.77666667, 23.83333333,
       32.16666667, 28.17666667, 31.23666667, 31.94      , 14.35      ])
```

In [19]:
```
from sklearn.metrics import r2_score
accuracy = r2_score(y_test,r)
accuracy
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
Input In [19], in ()
      1 from sklearn.metrics import r2_score
```

7.2 Feature 2

In [4]:
```
import numpy as np
import pandas as pd
```

In [5]:
```
#Reading The Dataset

datas = pd.read_csv(r"C:\Users\sunda\Desktop\IBM\car performance.csv")
datas.head()
```

Out[5]:

	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin	car name
0	18.0	8	307.0	130	3504	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165	3693	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150	3436	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150	3433	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140	3449	10.5	70	1	ford torino

In [6]:
```
# Handling Missing Values

datas.isnull().any()
```

Out[6]:
```
mpg             False
cylinders       False
displacement    False
horsepower      False
weight          False
acceleration    False
model year      False
origin          False
car name        False
dtype: bool
```

In [7]:
```
# Splitting The Dataset Into Dependent And Independent Variable.

x = datas.iloc[:,1:8].values
```

In [8]:
```
y = datas.iloc[:,0].values
```

In [9]:
```
# Split The Dataset Into Train Set And Test Set

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.2,random_state=0)
```

In [10]:
```
# Normalizing

from sklearn.preprocessing import StandardScaler
sd = StandardScaler()
```

In [11]:
```
x_train
```

Out[11]:
```
array([[ 1.49526939,  1.22961301,  1.24359144, ..., -0.79520768,
        -1.13752513, -0.73301171],
       [-0.85285735, -0.92367663, -1.16092059, ...,  1.24411524,
        -1.41177304,  0.5068698 ],
       [-0.85285735, -0.92367663, -0.68001818, ...,  0.05760009,
         1.05645814,  0.5068698 ],
       ...,
       [-0.85285735, -1.206235  , -1.45480539, ...,  1.42950823,
        -0.86327722,  0.5068698 ],
       [ 0.32120602,  0.56706235, -0.09224857, ..., -0.2390287 ,
        -1.41177304, -0.73301171],
       [-0.85285735, -0.99188037, -0.86703579, ...,  0.31715028,
        -0.31478141,  0.5068698 ]])
```

In [12]:
```
# Build The Model With The Random Forest Regressor

from sklearn.ensemble import RandomForestRegressor
d = RandomForestRegressor (n_estimators=30,random_state = 0)
d.fit(x_train,y_train)
```

Out[12]:
```
RandomForestRegressor(n_estimators=30, random_state=0)
```

In [13]:
```
# prediction

y_pred = d.predict(x_test)
y_pred
```

Out[13]:
```
array([14.38333333, 24.25666667, 14.21666667, 20.56666667, 18.47333333,
       30.21666667, 34.63333333, 21.15      , 16.30333333, 25.76      ,
       36.60333333, 36.27      , 19.53666667, 27.32333333, 16.54333333,
       32.99333333, 28.32333333, 27.49666667, 17.03      , 35.82      ,
       16.47333333, 23.54      , 23.16666667, 20.7       , 33.69666667,
       26.45      , 33.79666667, 30.37333333, 31.93666667, 16.57333333,
       20.26666667, 32.99      , 19.79666667, 34.08333333, 20.85666667,
       25.02      , 19.65333333, 17.14      , 34.78333333, 12.76666667,
       13.73333333, 15.2       , 28.32      , 32.76666667, 28.74333333,
       22.68666667, 20.54333333, 16.50666667, 23.38      , 29.88333333,
       34.31666667, 26.5       , 17.63      , 27.78333333, 15.96666667,
       12.96666667, 18.86666667, 26.91666667, 31.95666667, 15.68      ,
       20.81      , 25.97      , 19.84666667, 21.6       , 13.46666667,
       15.33333333, 14.2       , 18.90333333, 24.72666667, 14.21666667,
       34.87666667, 13.25      , 22.96666667, 18.77666667, 23.83333333,
       32.16666667, 28.17666667, 31.23666667, 31.94      , 14.35      ])
```

In [14]:
```
# Model Evaluation

from sklearn.metrics import r2_score
accuracy = r2_score(y_test,y_pred)
accuracy
```

Out[14]:
```
0.8914224071232417
```

In [17]:
```
# save the model

import pickle
pickle.dump(d,open('regression.pkl','wb'))
```

In [ ]:

In [ ]:

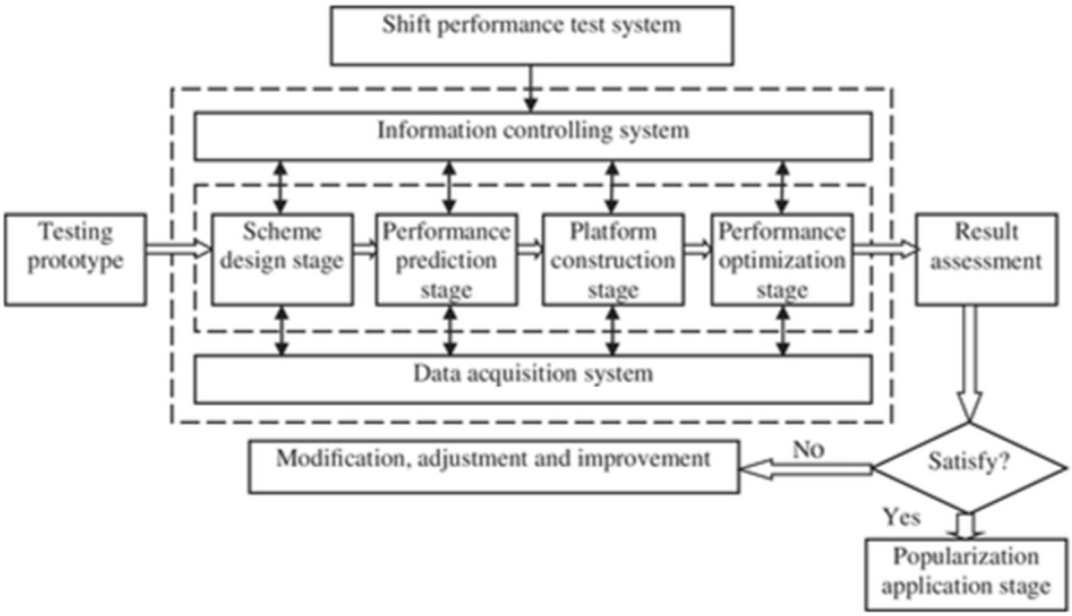
7.3 Database Schema (if Applicable)

8.Testing

8.1 Test Cases

GIF Gesellschaft für Industrieforschung is an automotive company based out of Germany, covering the vehicle transmissions and the design, testing, and development of powertrain system. GSA system, which was developed by GIF, has proven in many years of testing to be a valuable tool, mainly by complex transmission and gear development projects. GSA system is not only applied for objectively assessing the change of gears in a vehicle, but also permits measuring and analyzing the influencing factors of gear shifting quality. Therefore, during the new development and integration of transmissions, transmission component testing, and quality assurance, GSA system is indispensable as the right tool. Based on GSA shifting performance test evaluation system, this project completes the control strategy research of improving the shifting quality of automobile transmission. The technical route to be adopted. The development process of transmission shifting performance is mainly divided into scheme design stage, performance prediction stage, platform construction stage, performance optimization stage, and popularization application stage. First, the research idea is clear, the research scheme is determined, and the theoretical research framework of shifting test and evaluation is constructed. Second, relevant materials are collected, the performance parameters of transmission and shifting control mechanism are summarized, and the shifting performance of the whole system is predicted. Third, the shifting control performance bench test and GSA test system are built to carry out the subjective and objective test and research of shifting performance, and the real-time acquisition of shifting performance is carried out. The performance indicators in the process are analyzed and evaluated to provide data support for improving the shifting performance of the system. Then the optimization improvement measures are specified and the shifting performance is re-evaluated to obtain the best matching relationship of the shifting performance and verify the effectiveness of the

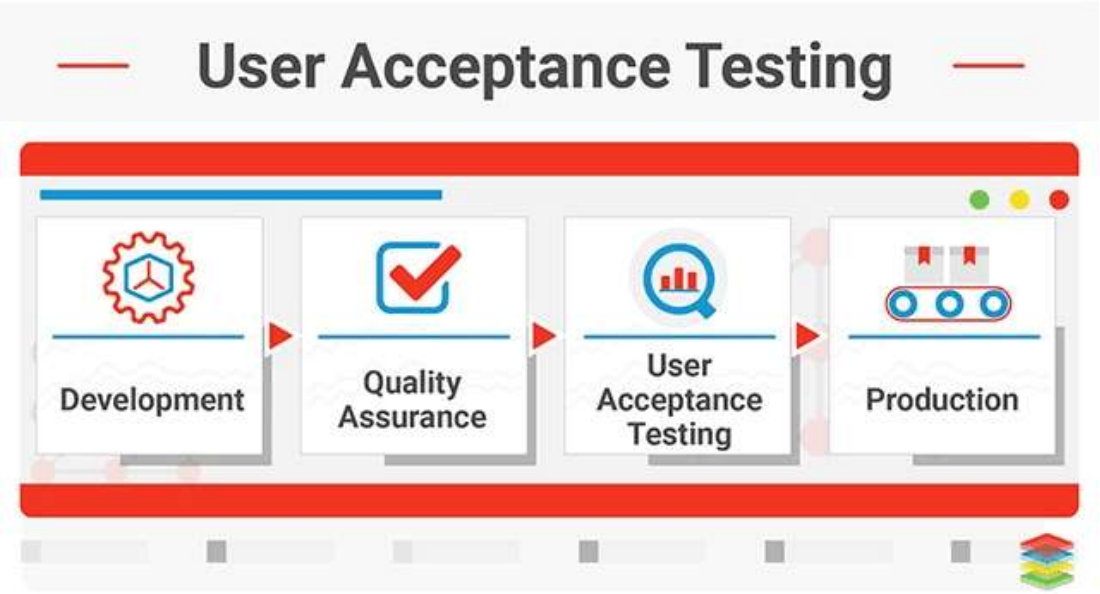
control strategy research scheme. Finally, the shifting optimization test and analysis path based on GSA test technology is formed.



GSA test analysis means that the whole test system includes hardware acquisition equipment and software processing analysis tools. Through real-time measurement and analysis of the force, travel, acceleration, and other important parameters of the shifting lever knob, the performance of the whole vehicle is systematically evaluated from the aspects of the shifting and selection force (travel), system stiffness, free play, and dynamic impact.

8.2 User Acceptance Testing (UAT)

It is a process to check that system accepts requirements of a user or not. It's performed at a time when the system used by actual users. User acceptance testing comes after →Unit Testing →Integration Testing →System Testing →Acceptance Testing in the process of testing



Alpha Testing

It is the type of testing, executed to identify all possible issues/bugs before releasing a product to every data users or public. It simulates real users by using Black Box and White Box testing techniques. The primary function is to carry out tasks executed by a typical user. Alpha testing carried out in a developed environment or in a lab environment where a product developed and the user of a product are internal employees of an organization.

Beta Testing

It is the type of testing in which users of software or application are real users. In this application is tested in a real environment and considered as a form of an external User Acceptance testing. Beta version of a software released to a limited number of a user when tested in a Real-Time environment with the help of real users, to obtain feedback on product quality. Beta testing reduces failures, risks and provides increased quality of a product through customer validation. It is a final test before shipping a product to the customers. In this type of testing, getting direct feedback from users is a significant benefit. It is required to test a product finally in a Real-Time environment.

9.RESULTS

9.1 Performance Metrics

The result analysis consisted of the collection of data from the OBD-II scanner through the app, which was then processed into the machine learning model and finally trained as shown in Trained Output values were used as the benchmark for testing against the gathered data. To do so, we performed a k-fold cross-validation technique with k=10, to train the model. We performed several experiments on the parameters which are essential for the testing of vehicle's safety and economic efficiency. In our first experiment, a relationship between Maximum speed value and the travel time (red zone) is obtained. This relationship describes the total distances travelled while crossing the road given the signal was red.We clearly observed, as shown in Fig. 4, that given the speed of the car was high, it was more likely for the driver to cross the road at red signal and in turn this implies the increasing likelihood of meeting with an accident We performed a k-fold cross-validation technique [6] with k=10, to train the model. In the Hypothesis-1, ECN_DRVG_INDX, we found that majority of the data was congested in the lower left part of the graph suggesting an inverse logarithmic growth of the trend based on the training data. This showed a positive growth of the ECN_DRVG_INDX based on the hypothesis value. However, the data scatters as the value on x-axis increases, hinting at a somewhat lesser correlation for predicted value based on hypothesis value. Therefore, the ECN_DRVG_INDX is found to be an inverse
logarithmic function of the features

10. ADVANTAGES & DISADVANTAGES

ADVANTAGES

- Easy Implementation
- Low cost
- Can know and maintain the driver and car performance

DISADVANTAGE

- It is very difficult to find the place for placing temperature sensor and pressure sensor.

11.CONCLUSION

In this paper we have obtained some newer insights about the car data analysis such as economic driving index (ECN_DRVG_INDX) and safety driving index (SFTY_DRVG_INDX.) The results have proven to be approximately 80% fitting the given features and are very helpful to be used in different use cases such as a parameter in finding the driver's driving performance in a driving school, as a good estimate for finding an optimal price for a used car that can be based on several factors which we have analyzed in this paper etc. We also found that the model used to train the data can be improved further by finding better hyper parameter values for the features. It is also possible that different features can be considered for improving the hypothesis.

12. FUTURE SCOPE

There is more scope in future for research and analysis of fuel efficiency by including other factors like the road condition and real-time traffic with the help of google maps, this would help in analysing much deeper. The knowledge discovered from the research and future work can be used by the car manufacturing companies to improve the fuel economy by considering the characteristics that substantially influence the fuel efficiency.

13. APPENDIX

Source code

GitHub : **https://github.com/IBM-EPBL/IBM-Project-4446-1658732517**

Project Demo link:

https://drive.google.com/file/d/1kMKEiHhT59kEpT4AgvQF_4uhka1NQFlC/view?usp=share_link