NAALAIYA THIRAN PROJECT - 2022
19ECI01-PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP

MACHINE LEARNING BASED VEHICLE PERFORMANCE ANALYZER

A PROJECT REPORT

Submitted by

SWETHA.S	**1904060**
VEDHESH.A.S	**1904061**
DHINESH.R	**1904075**
JERIN.A	**1904083**

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

COIMBATORE INSTITUTE OF TECHNOLOGY,

COIMBATORE – 641 014

(Government Aided Autonomous Institution affiliated to Anna University)

ANNA UNIVERSITY: CHENNAI 600025

NOVEMBER 2022

COIMBATORE INSTITUTE OF TECHNOLOGY

(Government aided Autonomous Institution Affiliated to Anna University)

COIMBATORE – 641014

ANNA UNIVERSITY: CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this report **"MACHINE LEARNING BASED VEHICLE PERFORMANCE ANALYZER"** is the Bonafide work of **SWETHA.S (1904060), VEDHESH.A.S (1904061), DHINESH.R (1904075), JERIN.A (1904083)** who carried out the **19ECL77 Professional Readiness for Innovation, Employability and Entrepreneurship** project offered by IBM and Anna University, Chennai.

SIGNATURE	**SIGNATURE**	**SIGNATURE**
Ms.T.LOGANAYAKI , M.E.,	**Dr.S.UMA,M.E.,Ph.D.,**	**Dr. A. RAJESWARI, M.E.Ph.D.,**
FACULTY MENTOR	**FACULTY EVALUATOR**	**PRINCIPAL AND HEAD**
Assistant Professor	Professor	Professor
Department of Electronics and Communication Engineering	Department of Electronics and Communication Engineering	Department of Electronics and Communication Engineering

TABLE OF CONTENTS

`

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 PROJECT OVERVIEW

The automotive industry is extremely competitive. With increasing fuel prices and picky consumers. Automobile makers are constantly optimising their processes to increase fuel efficiency. The performance analysis of the car is based on the various parameters.

To develop new car models with advanced features, it is first necessary to analyse its current performance. However, for an effective prediction of a vehicle's performance, it is important to consider its engine type, number of engine cylinders, fuel type, horsepower, etc. Hence, the proposed model will predict the quality of a car based on these factors which can then be improved by decreasing its fuel consumption and thus increase the overall efficiency. A machine learning model is proposed to help one understand the influence of these parameters on the vehicle's performance. By applying different algorithms according to the dataset and visualisation, the best algorithm for this problem statement is found. And the final estimation will be of much use in the car industry.

1.2 PURPOSE

Predicting the performance level of cars is an important and interesting problem. The main goal is to predict the performance of the car to improve certain behaviours of the vehicle. This can significantly help to improve the system's fuel consumption and increase efficiency.

2.IDEATION PHASE

2.1 LITERATURE SURVEY

2.1.1 EXISTING SOLUTIONS

● Enhancing Performance Prediction Robustness by Combining Analytical Modeling and Machine Learning
● Design of a Performance Analyzer for Electric Vehicle Taxi Systems
● SPRINTA: SPRINT PERFORMANCE ANALYZER BASED ON ACCELEROMETER DATA
● Innovative Analytic Test Vehicle Integrated into Automated Indoor Braking Analyzer
●ElectricVehiclePerformanceAnalyzer: https://www.pestingers.net/pdfs/other-computers/circuit-cellar/1999/circuit-cellar-110.pdf#page=14

2.1.2 INFERENCE FROM THE EXISTING SOLUTIONS

1. From a Taxi performance analyser we went through : An event tracker, a stream handler, object interfaces, and strategy integrator, the analysis procedure can measure the performance of a dispatch and relocation strategy in terms of dispatch latency, customer waiting time, and the number of daily fast charging operations
2. From an Enhancing Performance Prediction : we explore several hybrid/Gray box techniques that exploit AM and ML in synergy in order to get the best of the two worlds. We evaluate the proposed techniques in case studies targeting two complex and widely adopted middleware systems: a NoSQL distributed key-value store and a Total Order Broadcast (TOB) service.
3. From an EV Performance analyser: Design of the analyser and cover both the microcontroller and PC software (both in C). The micro's software was developed with a Hitachi C compiler, and the software for the PC was developed with CVI (C for Virtual Instrument) from National Instrument

2.2 EMPATHY MAP:

An empathy map is a visualization tool used to articulate what a product team knows about a user. This tool helps product teams build a broader understanding of the "why" aspect behind user needs and wants. This tool forces product teams to practice empathic design, which shifts the focus from the product they want to build to the people who will use this product. As a team identifies what they know about the user and places this information on a chart, they gain a more holistic view of the user's world and his or her problems, or opportunity space.

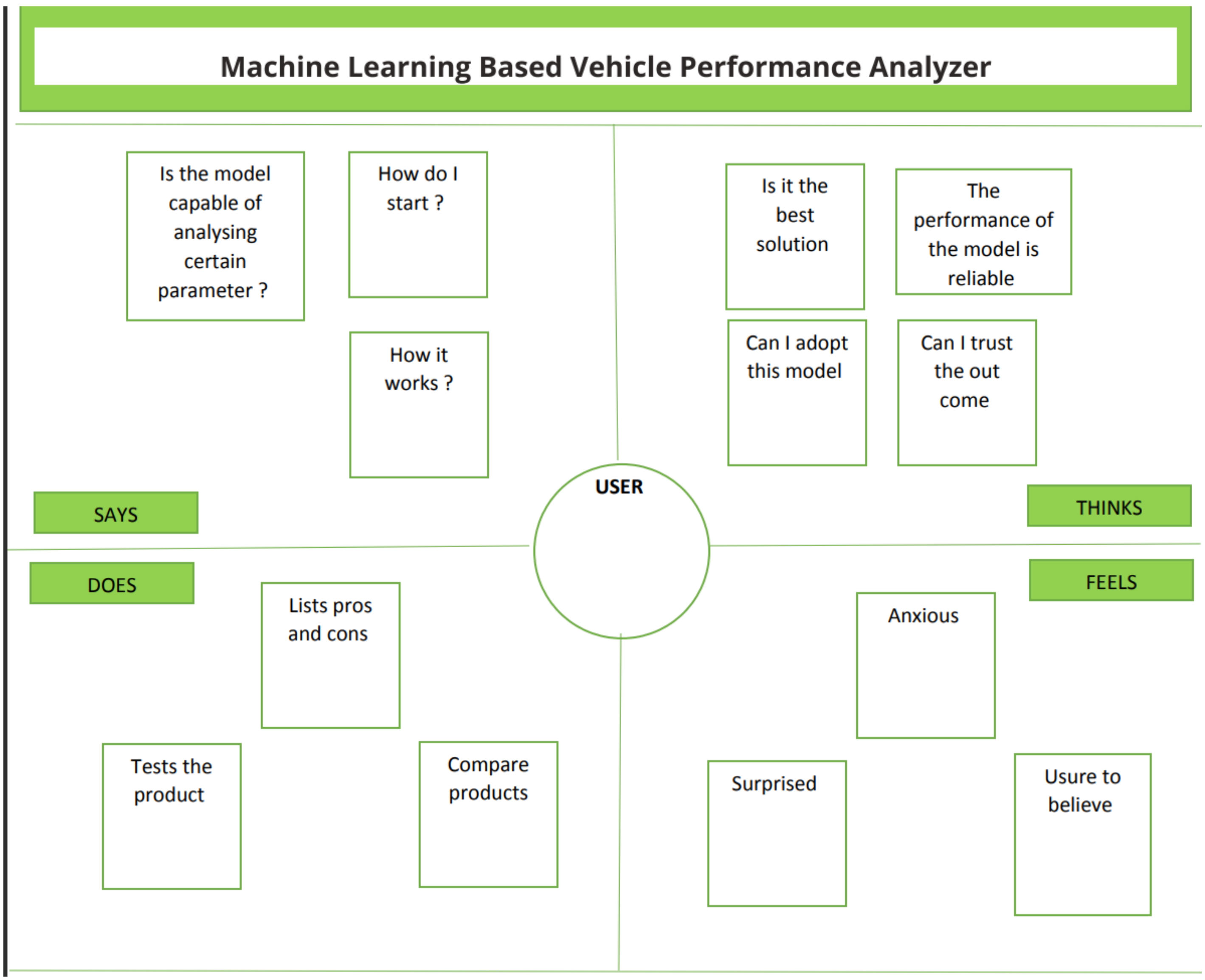


Fig.2.2 Empathy map

2.3 PROBLEM STATEMENT

Predicting the performance level of cars is an important and interesting problem. The main goal is to predict the performance of the car to improve certain behaviours of the vehicle. This can significantly help to improve the system's fuel consumption and increase efficiency. The performance analysis of the car is based on the engine type, no of engine cylinders, fuel type, horsepower, etc. These are the factors on which the health of the car can be predicted. It is an ongoing process of obtaining, researching, analyzing, and recording health based on the above three factors. The performance objectives like mileage, dependability, flexibility and cost can be grouped together to play a vital role in the prediction engine and engine management system. This approach is a very important step towards understanding the vehicle's performance.

2.4 IDEATION:

Ideation is the process where you generate ideas and solutions through sessions such as Sketching, Prototyping, Brainstorming, Brainwriting, Worst Possible Idea, and a wealth of other ideation techniques. Ideation is also the third stage in the Design Thinking process.

Fig.2.4.1 Ideation

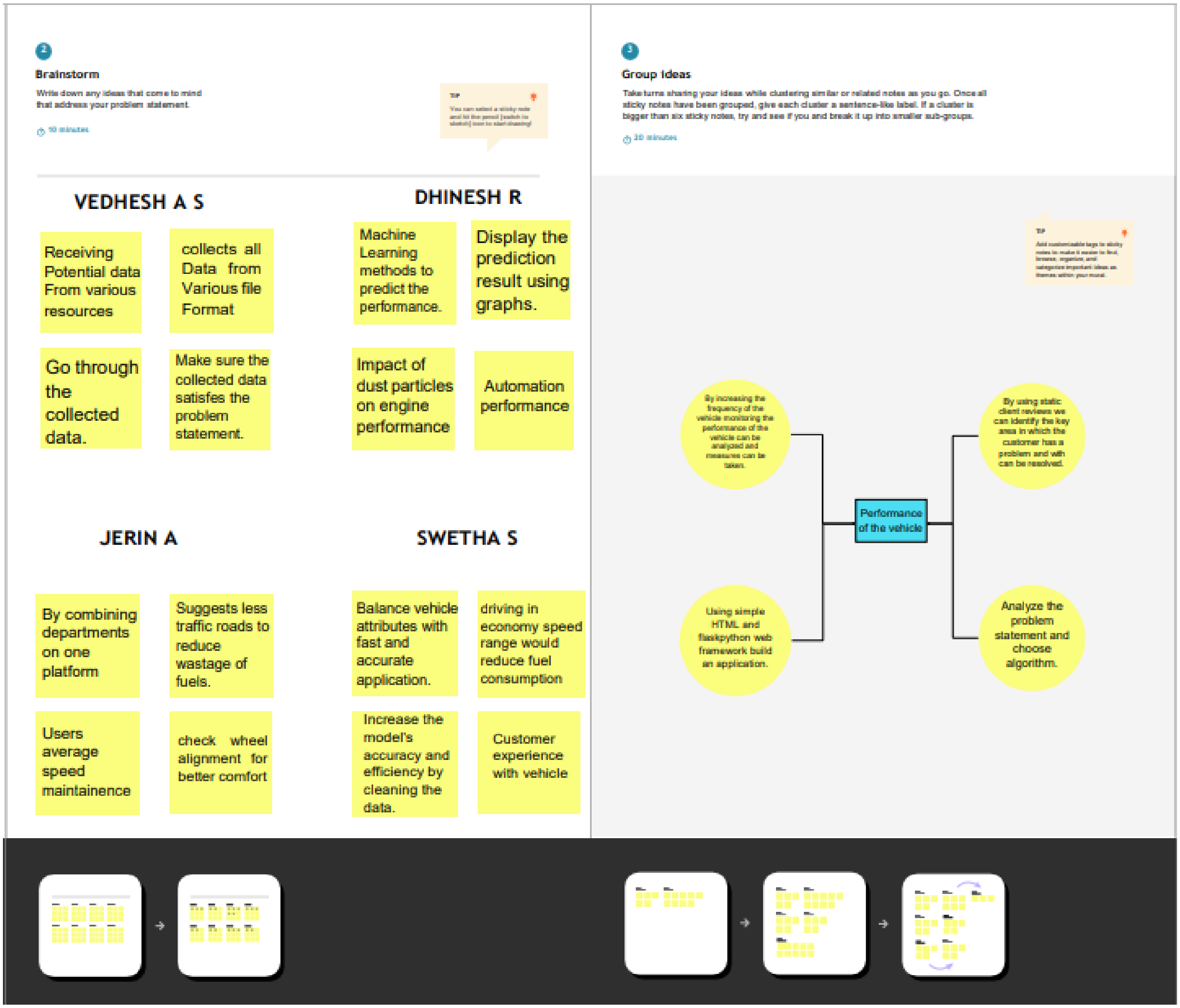


Fig.2.4.2 Ideation continued

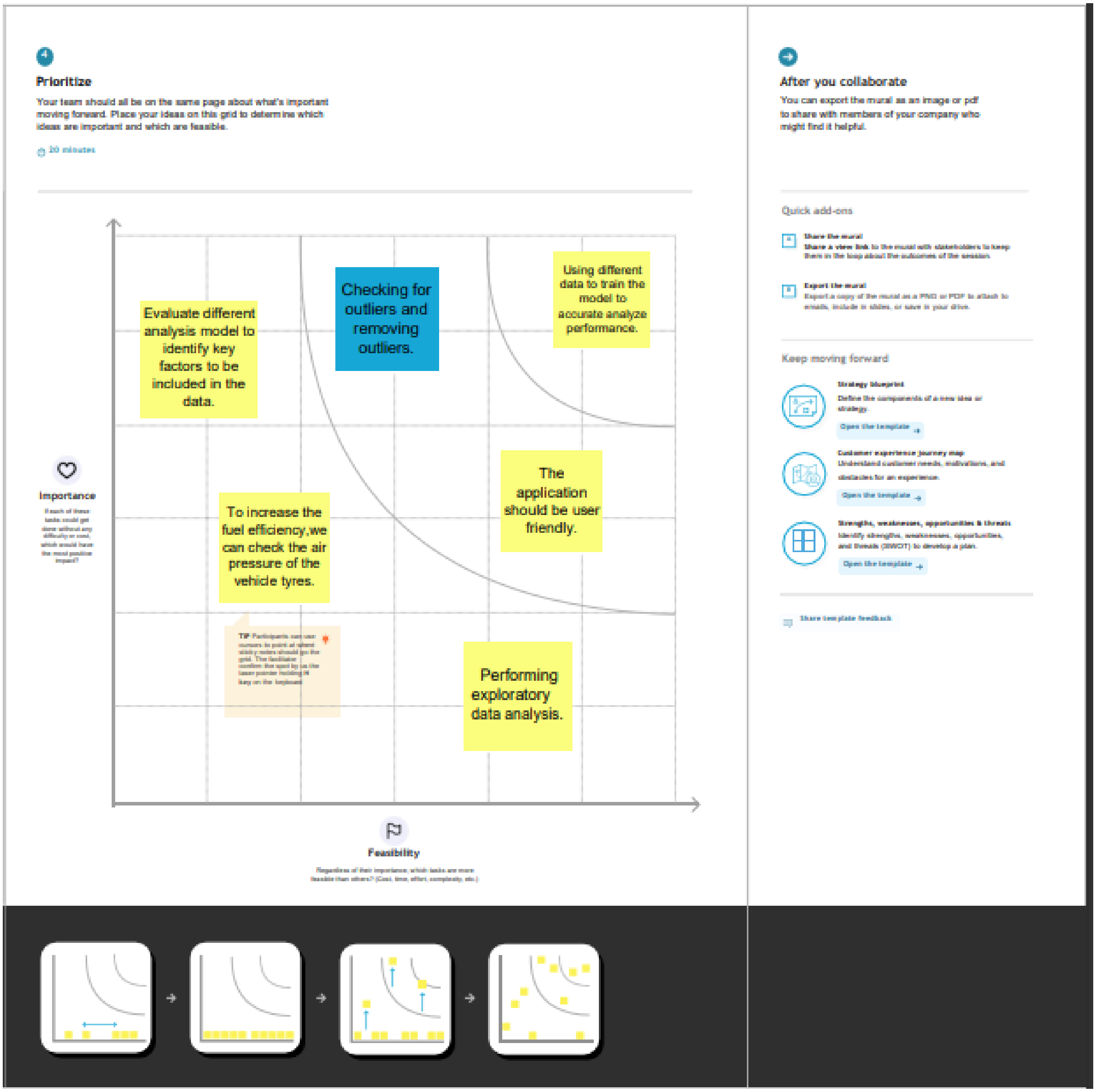


Fig.2.4.3 Ideation continued

3. PROJECT DESIGN PHASE-I

3.1 PROPOSED SOLUTION

The proposed Machine Learning Based Vehicle Performance Analyser system is shown in Figure 3.1

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Predicting the performance level of a vehicle has some problems based on complexities with data because it needs about a million relevant records to train an ML model.
2.	Idea / Solution description	Using a supervised learning algorithm to know the target value for the problem. To train such a model, which can be identified as the vehicle parameters preferable with the variety of configurations, they are required as input variables.
3.	Novelty / Uniqueness	In machine learning, the dataset used in the training phase is a significant factor in building successful predictions.
4.	Social Impact / Customer Satisfaction	Perfection may include and extend beyond driving safety performance, estimation of the vehicle’s life, fuel efficiency, and long-distance driving efficiency.

5.	Business Model (Revenue Model)	A vehicle’s fuel consumption is influenced by external and internal factors, although the engine and vehicle type minimise fuel consumption.
6.	Scalability of the Solution	From the study's conclusion, it is inferred that fuel consumption rate and vehicle driver index (VDI), a measure of driving behaviour, are deeply related.

Table 3.1 Proposed solution

3.2 PROBLEM SOLUTION FIT

The problem solution for Machine Learning Based Vehicle Performance Analyser is shown in fig 3.2

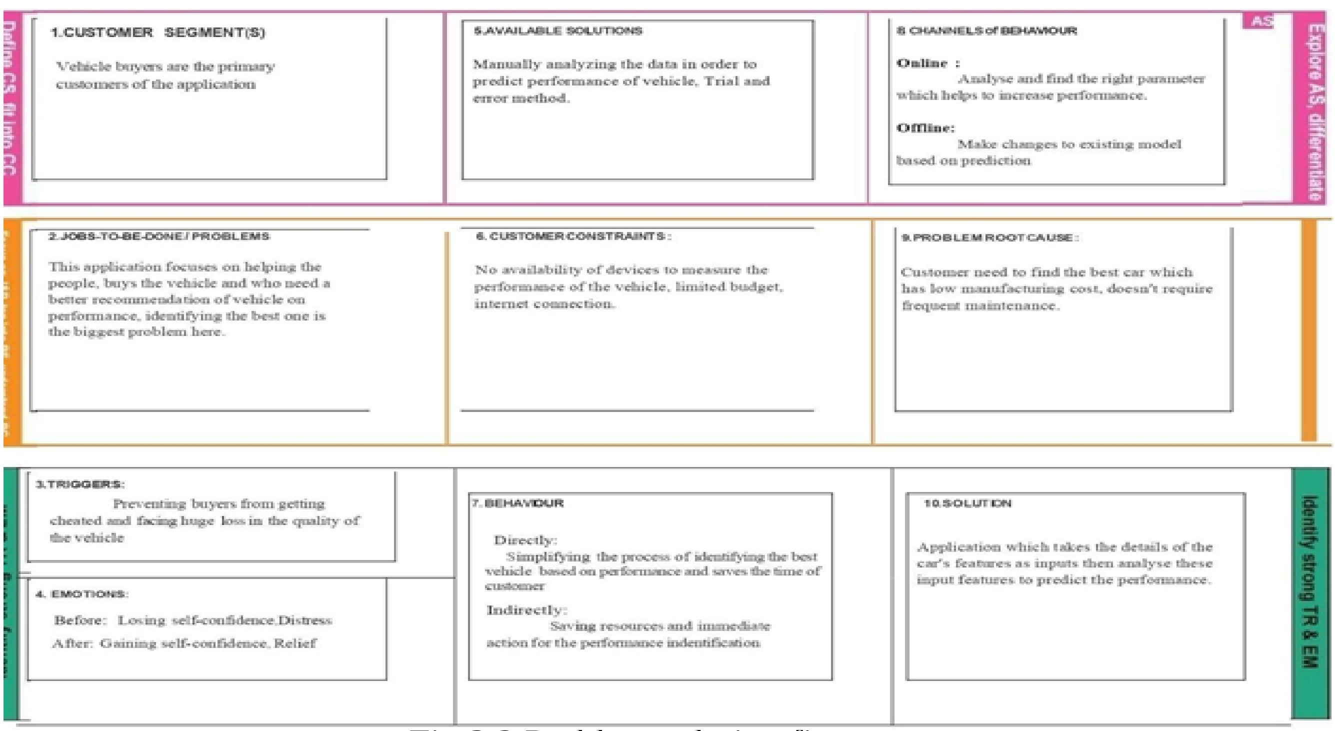


Fig 3.2 Problem solution fit

3.3 SOLUTION ARCHITECTURE DIAGRAM

The Solution Architecture for Machine Learning Based Vehicle Performance Analyser is shown in fig 3.3

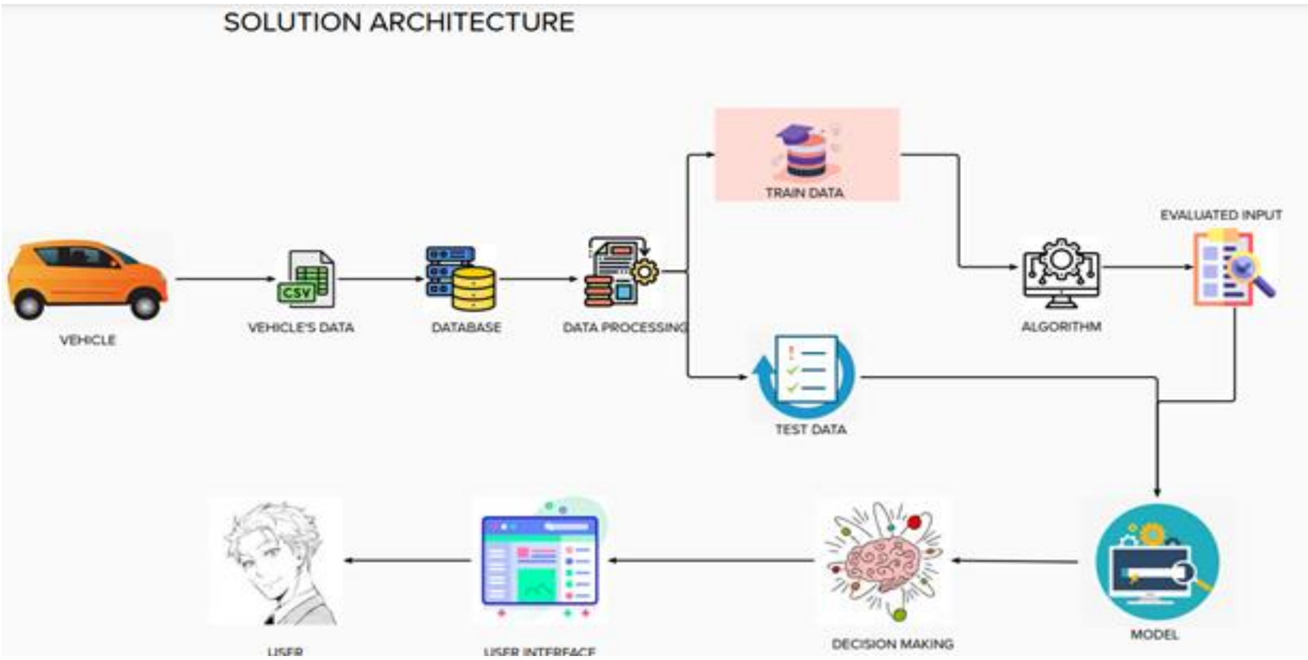


Fig 3.3 Solution Architecture

4.PROJECT DESIGN PHASE-II

4.1 REQUIREMENT ANALYSIS

The Requirement Analysis for Machine Learning Based Vehicle Performance Analyser is shown in fig 4.1

FR.No	FUNCTIONAL REQUIREMENTS (Epic)	NON-FUNCTIONAL REQUIREMENTS
FR-1	Enter the Inputs	Get Inputs through a form
FR-2	User Essential	Predict the performance of the vehicle
FR-3	Data Prepossessing	Sample Dataset for training purpose
FR-4	User input Evaluation	Evaluating the given user values
FR-5	Prediction	Fuel consumption and efficiency of the vehicle

Table 4.1 Requirement analysis

4.2 CUSTOMER JOURNEY

Customer journey maps are used to map the relationship between a customer and an organization over time and across all channels on which they interact with the business. Design teams use customer journey maps to see how customer experiences meet customers' expectations and find areas where they need to improve designs.

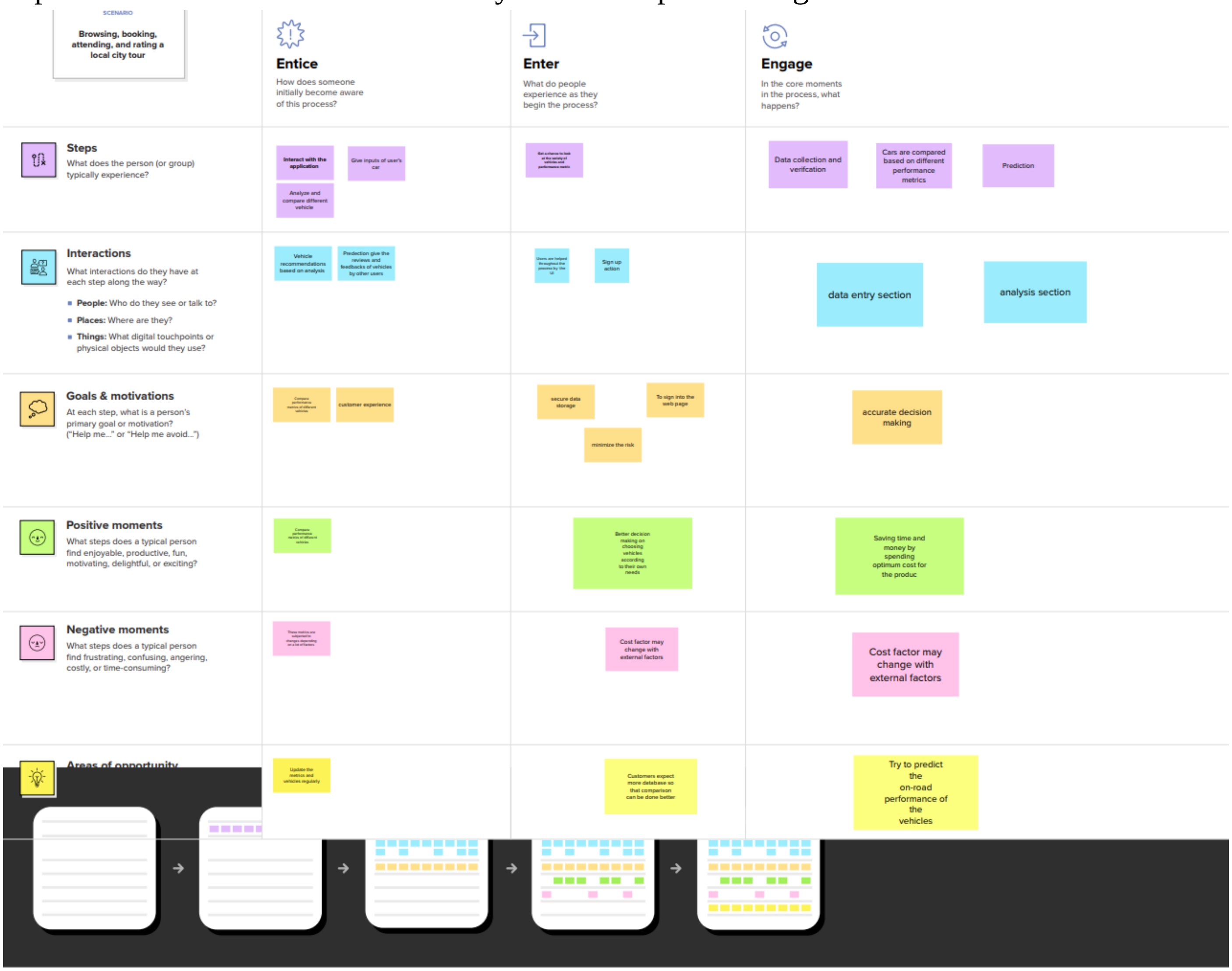


Fig. 4.2 Customer Journey

Fig. 4.2 Customer Journey

4.3 DATA FLOW DIAGRAMS

A data-flow diagram is a way of representing a flow of data through a process or a system. The DFD also provides information about the outputs and inputs of each entity and the process itself. A data-flow diagram has no control flow — there are no decision rules and no loops.

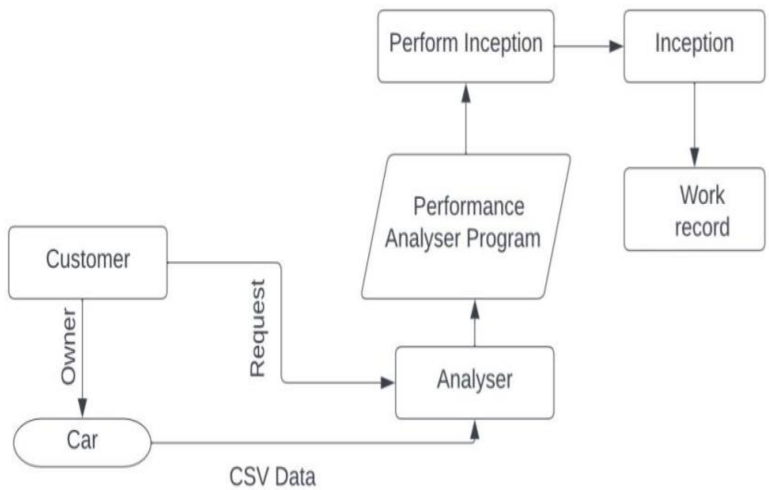


Fig. 4.3 Data flow diagram

User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer	Access the webpage	USN-1	Anyone can access the webpage to check the specifications of the vehicle.	I can access my webpage online at any time.	High	Sprint-1
Customer	Performance of the vehicle	USN-2	As per the usage of the user, the performance of the vehicle should be predictable.	Prediction can be done in an easy way.	High	Sprint-2
Customer	Accuracy to check the performance and health of the car	USN-3	By using our prediction, it helps to check the health of the car.	The efficiency of the car can be predicted.	High	Sprint-1

Table. 4.3 Data flow

4.4 TECHNOLOGY ARCHITECTURE:

A technical architecture diagram provides an overview of the various components of your system and how they work together. They are beneficial when planning and managing large-scale technology projects, as they facilitate better decision-making and understanding.

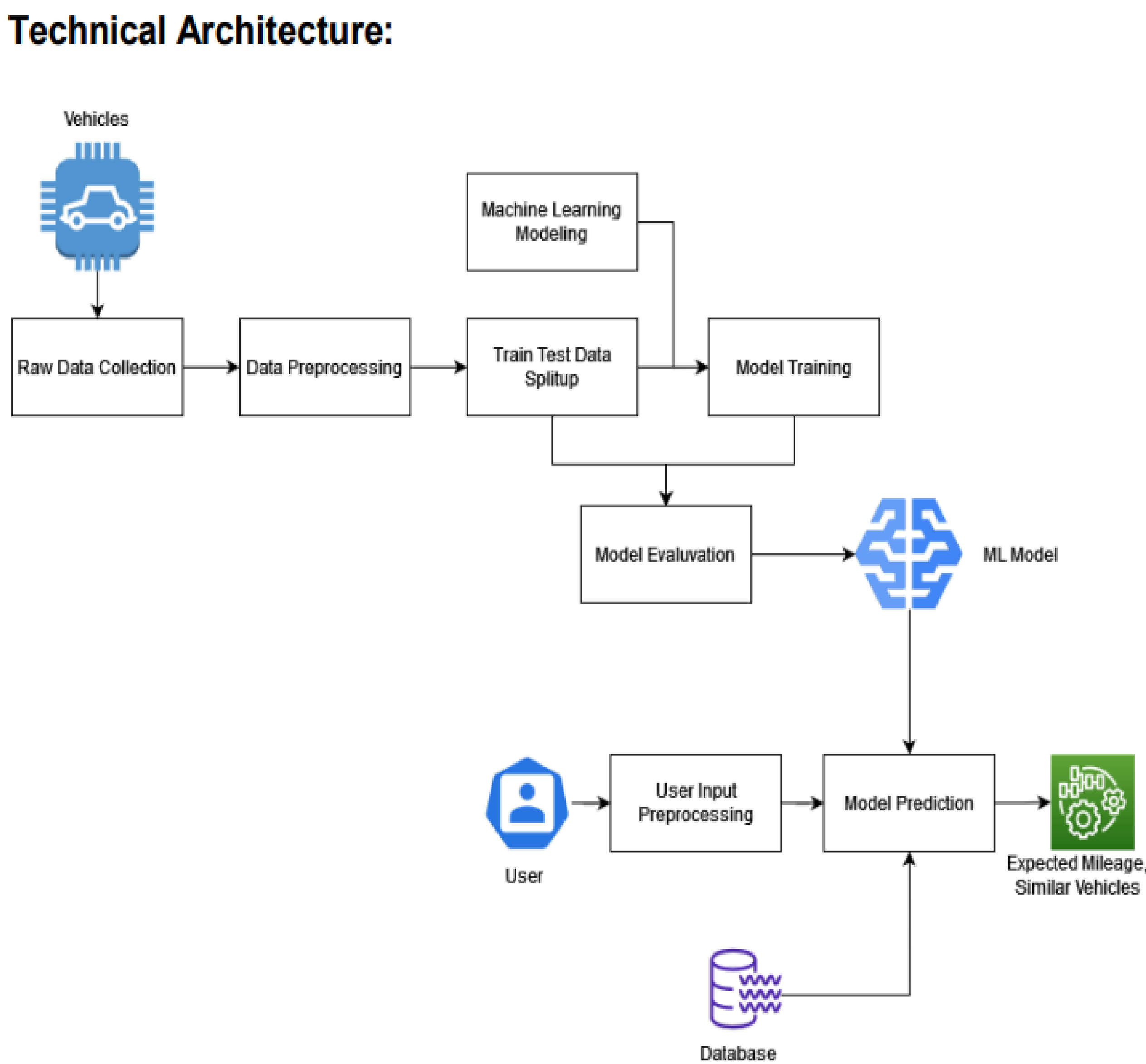


Fig 4.4 Technology architecture

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	The user interacts with the application through a Web Application that is responsive to the device that is being used.	React Js
2.	Get User Data	The process collects the user input data that is collected via a form to the server as a JSON Object	REST API
3.	Model Prediction	Use the data collected from the user to make predictions on the mileage expected.	IBM Watson ML
4.	Send User Report	Send the predictions along with suggestions to the user as JSON Object	REST API
5.	Database	Database contain user information such as name, email, vehicle basic information, mileage predicted over time.	MySQL
6.	Cloud Database	Database Service on Cloud	IBM DB2
7.	External API-1	Vehicle Details Database	https://api.auto-data.net/
8.	Machine Learning Model	The machine learning model is used to predict mileage from the user inputs	Regression Modelling.
9.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Core i5, 8GB RAM Cloud Server Configuration :	Local, Docker

Table. 4.4.1 Components & Technologies

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	React Js, Flask, Sci-kit Learn	Javascript, Python
2.	Security Implementations	Identity and Access Management, OAUTH, WAF	IBM Cloud
3.	Scalable Architecture	3 Tier Architecture, Model-View-Controller implementation.	Model - SQL DB, View - ReactJS, Controller - Flask Server
4.	Availability	Proxy servers, Load Balancers to help balance traffic among servers to help improve uptime	IBM Cloud load balancers
5.	Performance	The frontend is detached from the Business logic server reducing requests sent to the server.	Nginx proxy

Table. 4.4.2 Application characteristics

5. PROJECT PLANNING PHASE

5.1. MILESTONE AND ACTIVITY LIST

The Milestones on the project workings has been set and the activity levels were updated

S.No	Milestone	Description	Duration	Working status
1	Project Objectives	Project objectives are what you plan to achieve by the end of your project. This might include deliverables and assets or more intangible objectives like increasing productivity or motivation	1 WEEK	Completed
2	Project Flow	It is a visual aid to understand the methodology you're using to manage the project. The diagram shows the interdependent and parallel processes over the project's life cycle	1 WEEK	Completed
3	Pre-Requisites	Pre-requisites are all the needs at the required level needed for the execution	1 WEEK	Completed

		of the different phases of a project.		
4	Prior Knowledge	Prior knowledge is defined as all the knowledge one has before learning about a particular topic	1 WEEK	Completed
5	Data Collection	It is the process of gathering and measuring information on targeted variables in an established system, which then enables one to answer relevant questions and evaluate outcomes	1 WEEK	Completed
6	Data Pre-processing	Data pre-processing is a process of cleaning the raw data i.e. the data is collected in the real world and is converted to a clean data set. In other words, whenever the data is gathered from different sources it is collected in a raw format and this data isn't feasible for the analysis.	1 WEEK	Completed

7	Model Building	It is the process of developing a probabilistic model that best describes the relationship between the dependent and independent variables	1 WEEK	In Progress
8	Application Building	Application is the process of creating a computer program. In this phase build our Flask application which will be running in our local browser with a user interface.	2 WEEKS	In Progress
9	Train the Model on IBM	Training the built model on the IBM Cloud.	2 WEEKS	In Progress
10	Ideation Phase	Ideation is the process where you generate ideas and solutions through sessions such as sketching, Prototyping, Brainstorming, Brainwriting, Worst Possible Idea, and a wealth of other ideation techniques.	1 WEEK	Completed

11.	Project Design Phase-I	Project design is an early phase of a project where the project’s key features, structure, criteria for success, and major deliverables are planned out. The aim is to develop one or more designs that can be used to achieve the desired project goals.	1 WEEK	Completed
12	Project Design Phase-II	Project design is an early phase of a project where the project’s key features, structure, criteria for success, and major deliverables are planned out. The aim is to develop one or more designs that can be used to achieve the desired project goals.	1 WEEK	Completed
13	Project Planning Phase	In the Planning phase, the project manager works with the project team to create the technical design, task list, resource plan, Communications plan budget, and initial schedule for the project and establishes the roles and responsibilities of the project team and its stakeholders.	1 WEEK	Completed

14	Project Development Phase	Project development is the process of planning and allocating resources to planning and allocating resources to fully develop a project or product from concept to go-live	1 WEEK	In Progress

Table. 5.1 Milestone and activity list

6.SPRINT DELIVERY PLAN

6.1.Product Backlog, Sprint Schedule, and Estimation:

The purpose of sprint planning is to define what can be delivered in the sprint and how that work will be achieved. Sprint planning is done in collaboration with the whole scrum team.

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Visiting Webpage	USN-1	As a user, I can able to view the website.	10	Low	Team leader
Sprint-1	Design	USN-2	As a user, I can Enter the data of the vehicle.	20	High	Team member1
Sprint-2	Result	USN-3	As a user, I can get the predicted performance of the vehicle using the given data.	20	High	Team member2

Sprint-3	Design	USN-4	As a user, I want a good user experience.	10	Low	Team member 3
Sprint-3	Result	USN-5	As a user, I want the website to work fast and predict performance quickly.	10	Low	Team leader
Sprint-4	Result	USN-6	As a user, I expect the prediction is highly accurate	20	High	Team leader

Table. 6.1 Product Backlog, Sprint Schedule, and Estimation

6.2.Project Tracker, Velocity & Burndown Chart:

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint -1	30	6 Days	24 Oct 2022	29 Oct 2022	30	29 Oct 2022
Sprint -2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint -3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint -4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

Table. 6.2 Project Tracker, Velocity & Burndown Chart

7.RESULTS AND DISCUSSIONS

Importing libraries and datasets is the first step that was done after which files were accessed using the IBM Cloud Object Storage including credentials. Multiple features were taken and they were correlated with the available parameters related to the vehicle. This existing dataset was analysed and visualised using pre-existing tag values of different parameters and mpg.
Random forest regressor was used to identify the best possible relationship that can be formed with the car parameters and mpg. The ultimate result provided an outlook toward highlighting the best possible strength between the parameters and mpg was visualised using a heatmap from an existing widely used visualisation library, seaborn.

Out[4]:

	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

Fig 7.1 Importing dataset

```
Out[3]: mpg             False
        cylinders       False
        displacement    False
        horsepower      False
        weight          False
        acceleration    False
        model year      False
        origin          False
        car name        False
        dtype: bool
```

There are no null characters in the columns but there is a special character '?' in the 'horsepower' column. So we we replaced '?' with nan and replaced nan values with mean of the column.

Fig 7.2 Finding missing data

Out[10]:

	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin
count	398.000000	398.000000	398.000000	398.000000	398.000000	398.000000	398.000000	398.000000
mean	23.514573	5.454774	193.425879	104.469388	2970.424623	15.568090	76.010050	1.572864
std	7.815984	1.701004	104.269838	38.199187	846.841774	2.757689	3.697627	0.802055
min	9.000000	3.000000	68.000000	46.000000	1613.000000	8.000000	70.000000	1.000000
25%	17.500000	4.000000	104.250000	76.000000	2223.750000	13.825000	73.000000	1.000000
50%	23.000000	4.000000	148.500000	95.000000	2803.500000	15.500000	76.000000	1.000000
75%	29.000000	8.000000	262.000000	125.000000	3608.000000	17.175000	79.000000	2.000000
max	46.600000	8.000000	455.000000	230.000000	5140.000000	24.800000	82.000000	3.000000

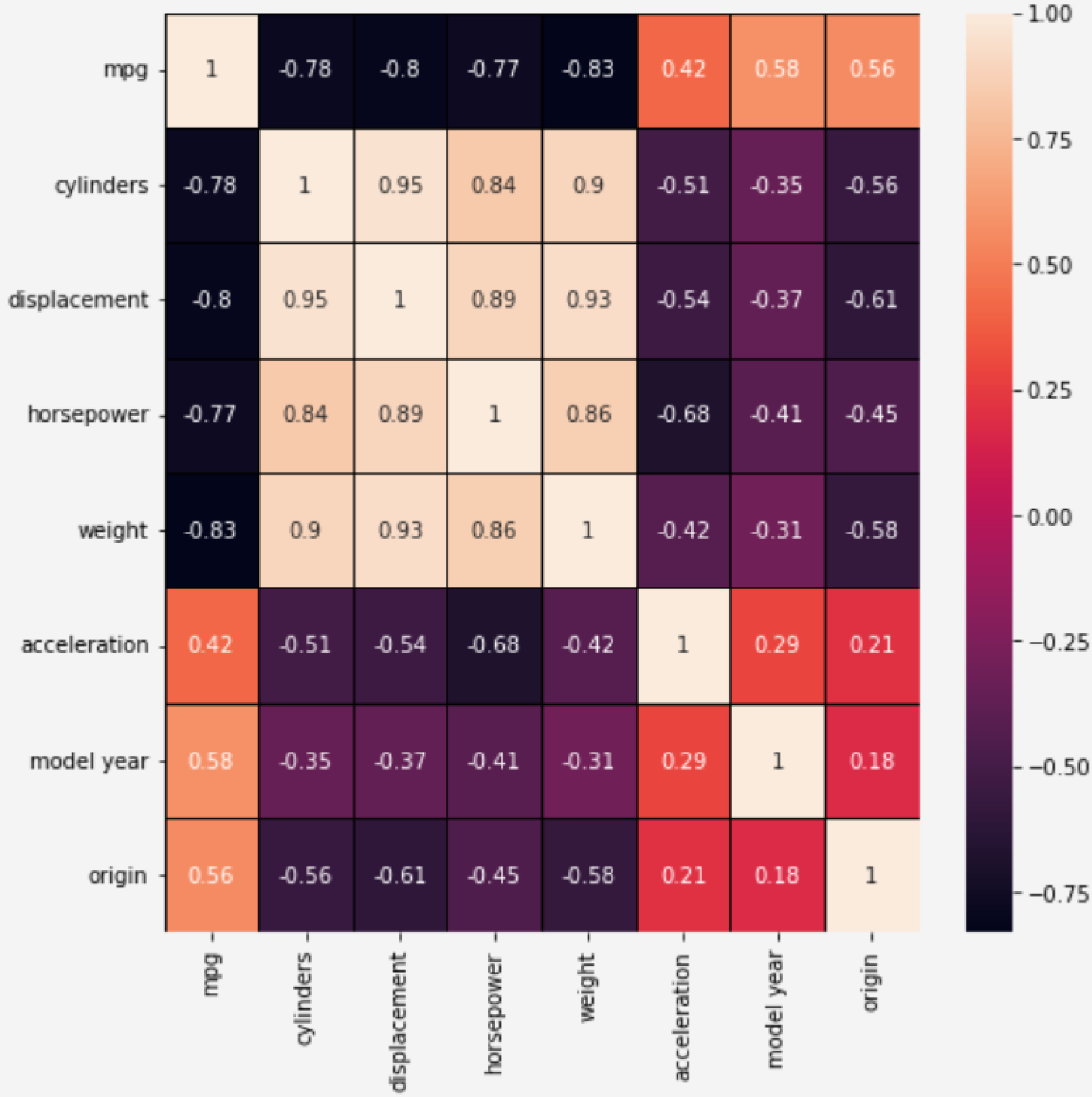
There is no use with car name attribute so drop it

Fig 7.3 Finding missing data

Out[12]:

	mpg	cylinders	displacement	horsepower	weight	acceleration	model year	origin
mpg	1.000000	-0.775396	-0.804203	-0.771437	-0.831741	0.420289	0.579267	0.563450
cylinders	-0.775396	1.000000	0.950721	0.838939	0.896017	-0.505419	-0.348746	-0.562543
displacement	-0.804203	0.950721	1.000000	0.893646	0.932824	-0.543684	-0.370164	-0.609409
horsepower	-0.771437	0.838939	0.893646	1.000000	0.860574	-0.684259	-0.411651	-0.453669
weight	-0.831741	0.896017	0.932824	0.860574	1.000000	-0.417457	-0.306564	-0.581024
acceleration	0.420289	-0.505419	-0.543684	-0.684259	-0.417457	1.000000	0.288137	0.205873
model year	0.579267	-0.348746	-0.370164	-0.411651	-0.306564	0.288137	1.000000	0.180662
origin	0.563450	-0.562543	-0.609409	-0.453669	-0.581024	0.205873	0.180662	1.000000

Fig 7.4 Finding missing data



Visualizations of each attributes w.r.t rest of all attributes

Fig 7.5 Heatmap-Data visualization

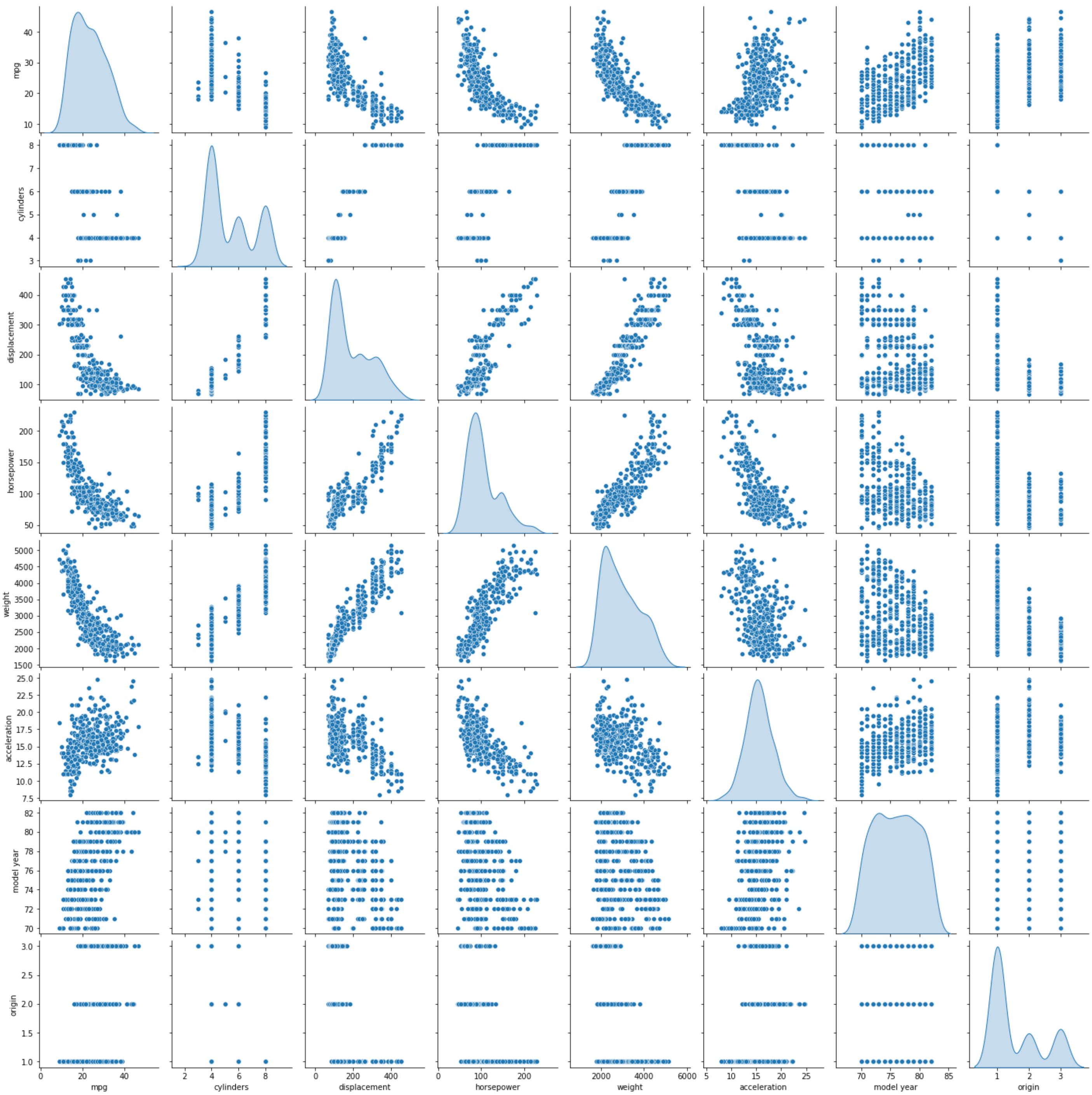


Fig 7.6 Regression plot-Data visualization

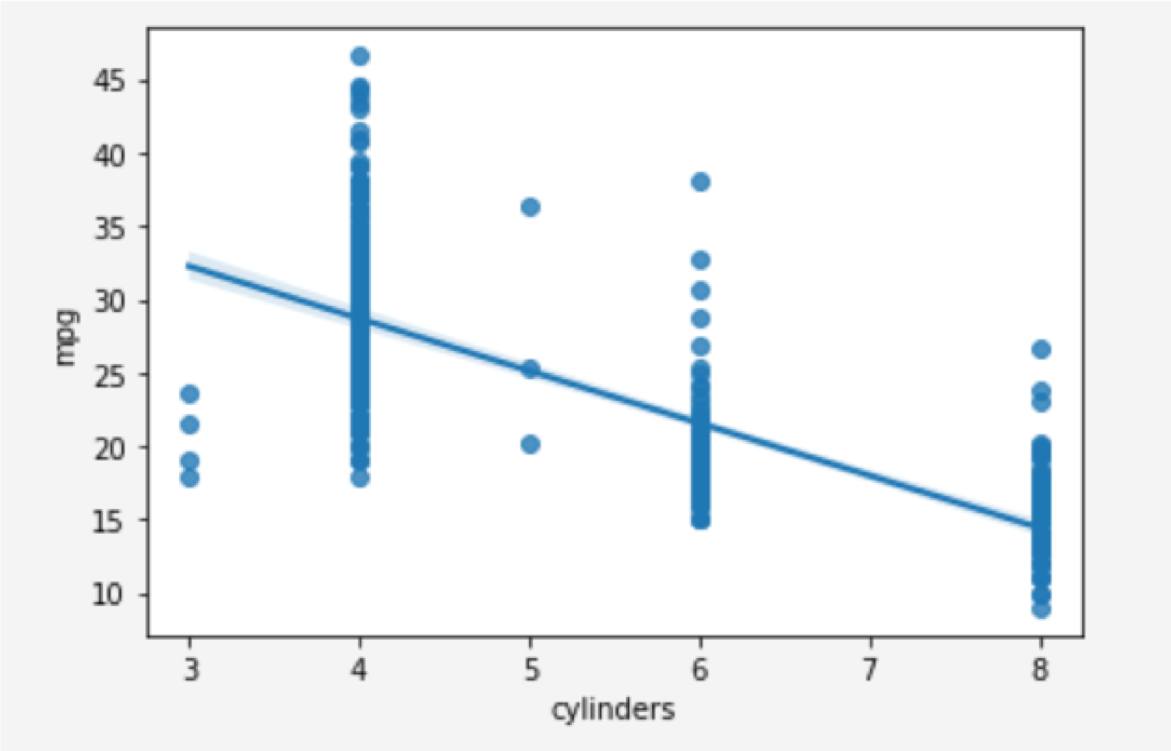


Fig 7.7 Displacement vs mpg

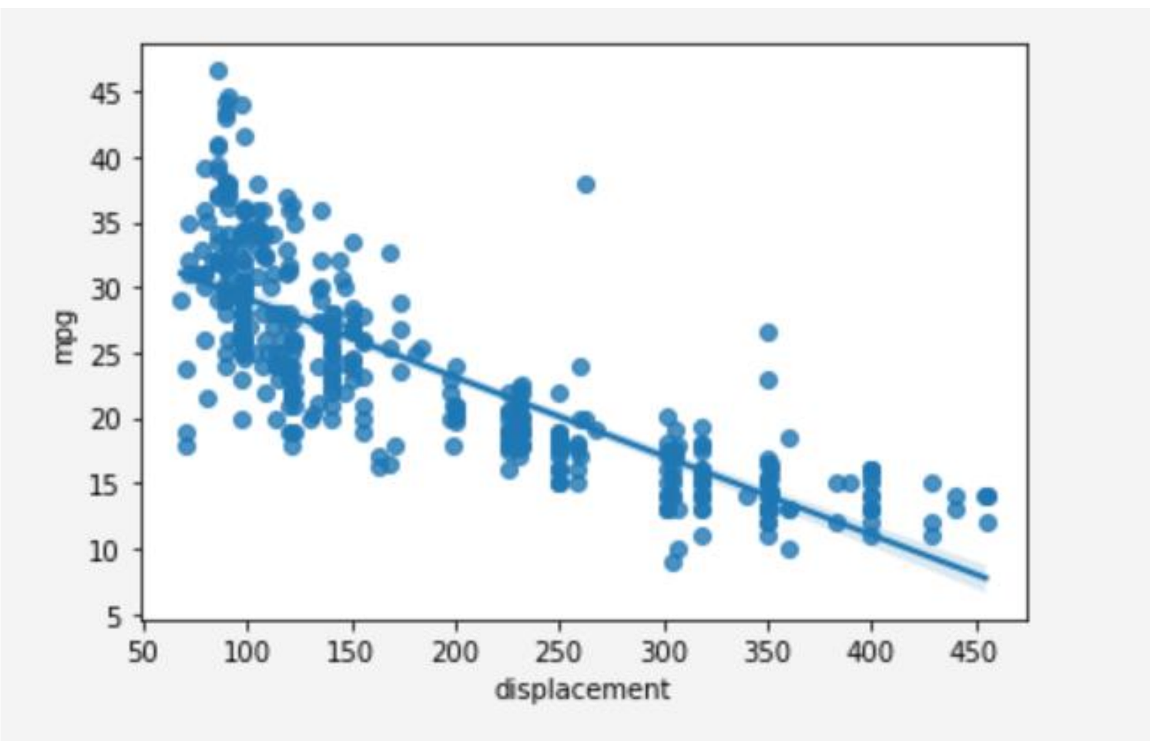


Fig 7.8 Horsepower vs mpg

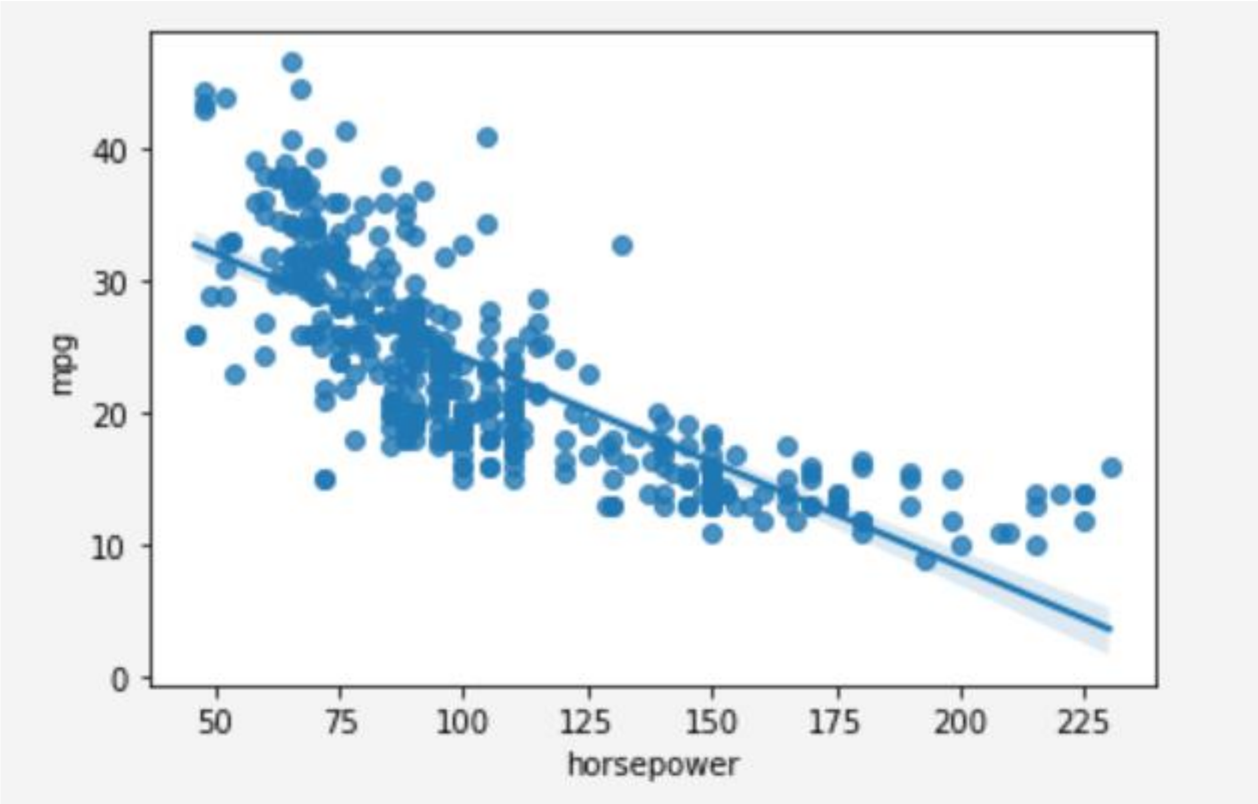


Fig 7.9 Weight vs mpg

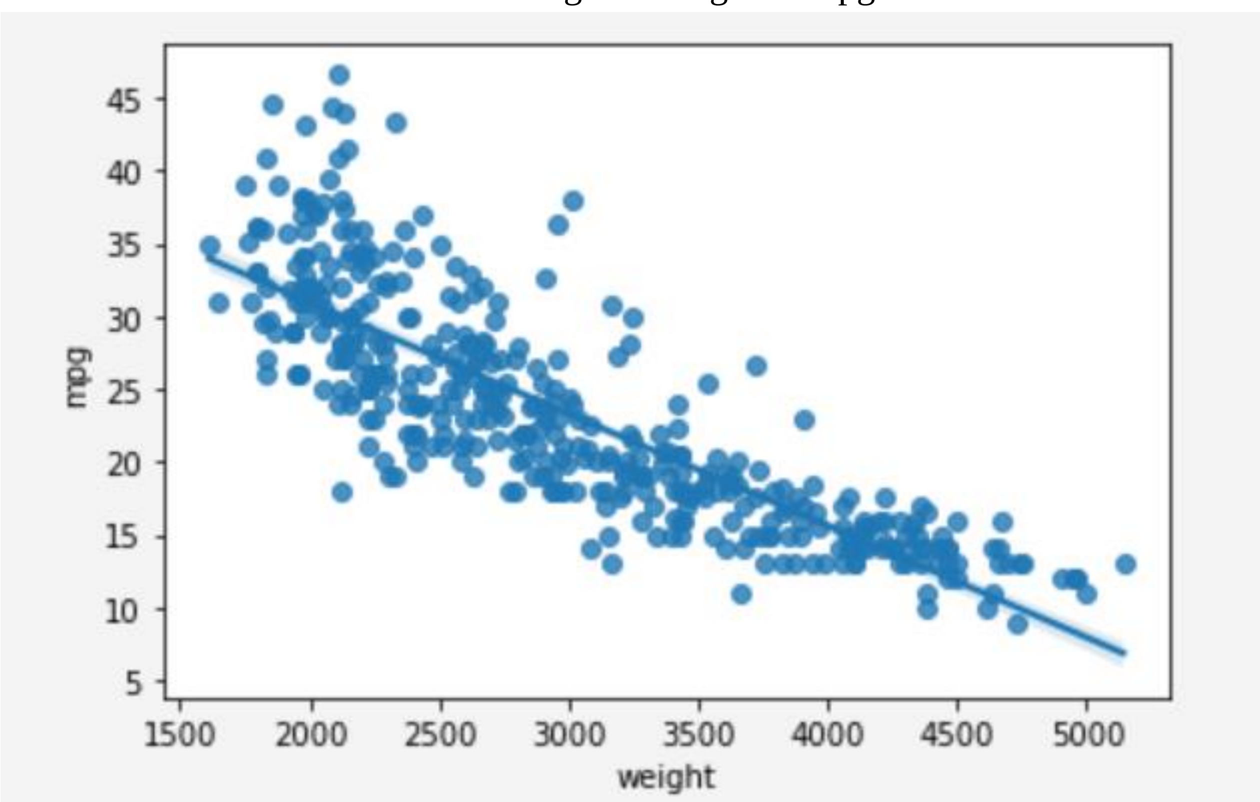


Fig 7.10 Acceleration vs mpg

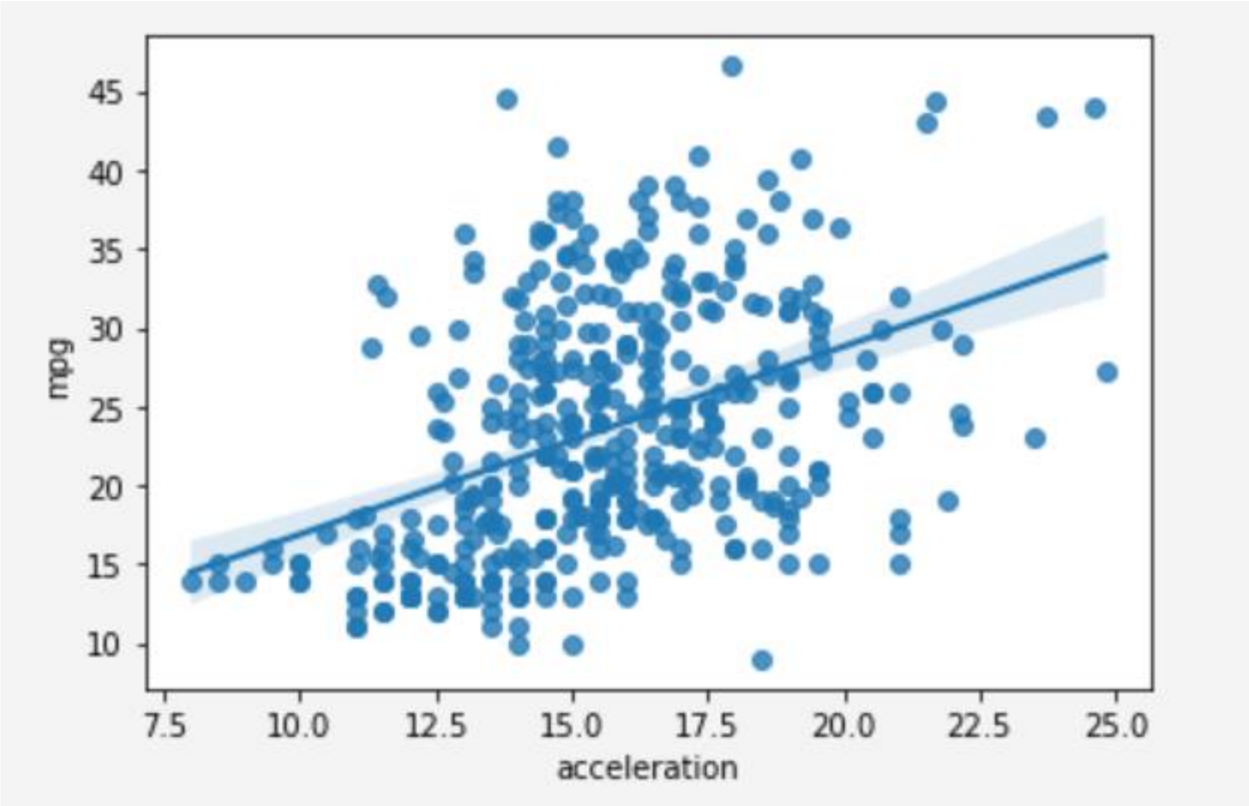


Fig 7.11 Acceleration vs mpg

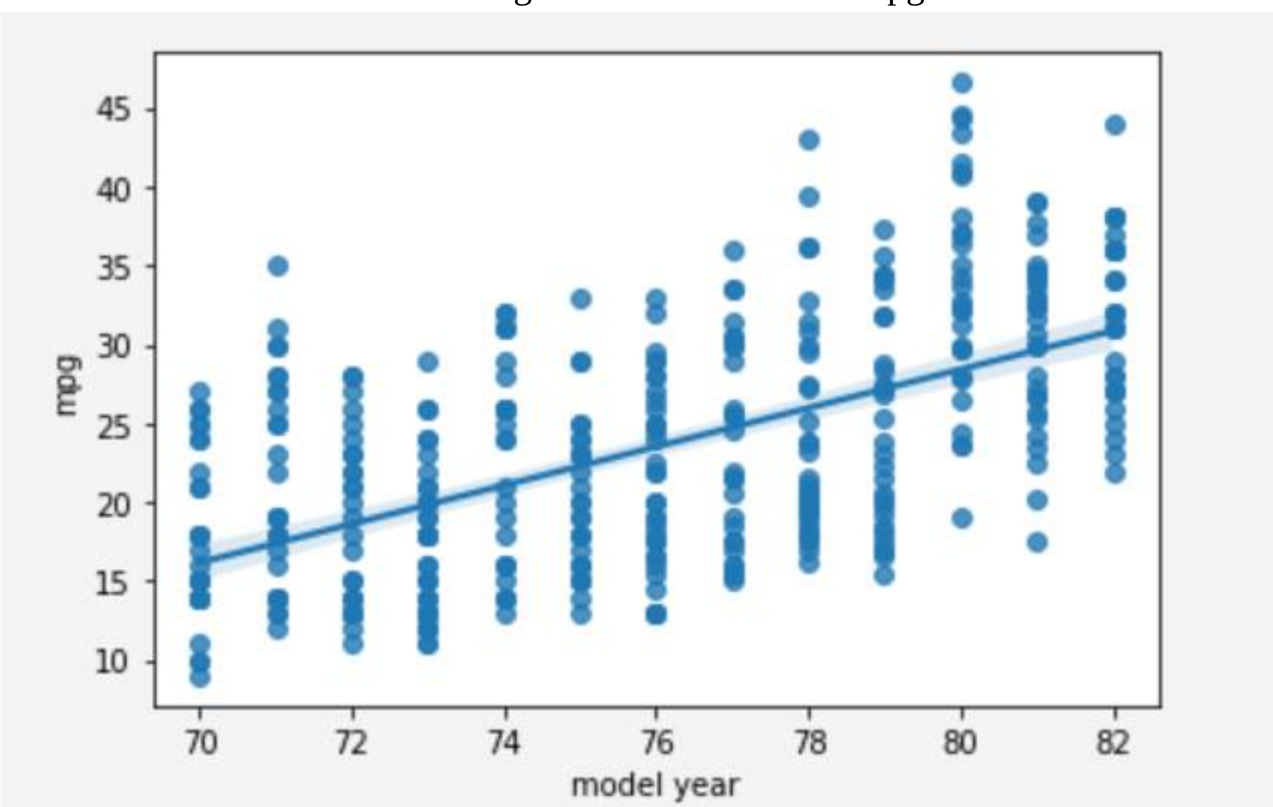


Fig 7.12 Model year vs mpg

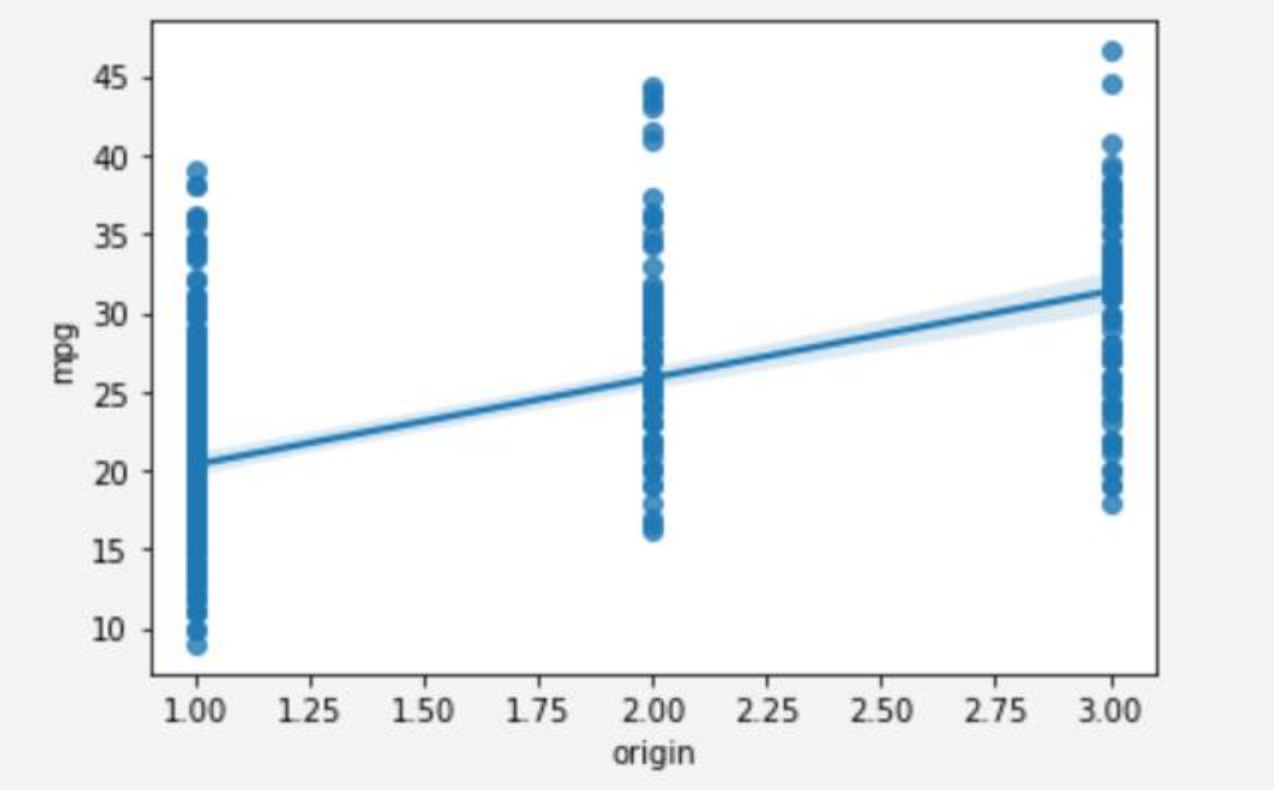


Fig 7.13 Origin vs mpg

Out[31]:

OLS Regression Results

Dep. Variable:	mpg	**R-squared:**	0.717
Model:	OLS	**Adj. R-squared:**	0.713
Method:	Least Squares	**F-statistic:**	165.5
Date:	Sun, 13 Nov 2022	**Prob (F-statistic):**	4.84e-104
Time:	15:17:17	**Log-Likelihood:**	-1131.1
No. Observations:	398	**AIC:**	2276.
Df Residuals:	391	**BIC:**	2304.
Df Model:	6		
Covariance Type:	nonrobust		

Fig 7.14 OLS Regression results

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	42.7111	2.693	15.861	0.000	37.417	48.005
cylinders	-0.5256	0.404	-1.302	0.194	-1.320	0.268
displacement	0.0106	0.009	1.133	0.258	-0.008	0.029
horsepower	-0.0529	0.016	-3.277	0.001	-0.085	-0.021
weight	-0.0051	0.001	-6.441	0.000	-0.007	-0.004
acceleration	0.0043	0.120	0.036	0.972	-0.232	0.241
origin	1.4269	0.345	4.136	0.000	0.749	2.105

Omnibus:	32.659	**Durbin-Watson:**	0.886
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	43.338
Skew:	0.624	**Prob(JB):**	3.88e-10
Kurtosis:	4.028	**Cond. No.**	3.99e+04

Fig 7.15 OLS Regression results

```
array([[8.000e+00, 3.070e+02, 1.300e+02, 3.504e+03, 7.000e+01, 1.000e+00],
       [8.000e+00, 3.500e+02, 1.650e+02, 3.693e+03, 7.000e+01, 1.000e+00],
       [8.000e+00, 3.180e+02, 1.500e+02, 3.436e+03, 7.000e+01, 1.000e+00],
       ...,
       [4.000e+00, 1.350e+02, 8.400e+01, 2.295e+03, 8.200e+01, 1.000e+00],
       [4.000e+00, 1.200e+02, 7.900e+01, 2.625e+03, 8.200e+01, 1.000e+00],
       [4.000e+00, 1.190e+02, 8.200e+01, 2.720e+03, 8.200e+01, 1.000e+00]])
```

Fig 7.15 Independent variables

Dependent variables

```
In [33]: y=dataset.iloc[:,0:1].values
         y
```

```
Out[33]: array([[18. ],
                [15. ],
                [18. ],
                [16. ],
                [17. ],
                [15. ],
                [14. ],
                [14. ],
                [14. ],
                [15. ],
                [15. ],
                [14. ],
                [15. ],
                [14. ],
                [24. ],
                [22. ],
                [18. ],
                [21. ],
                [27. ],
                [26. ],
                [25. ],
                [24. ],
                [25. ],
                [26. ],
                [21. ],
                [10. ],
                [10. ],
                [11. ],
                [ 9. ],
```

Fig 7.16 Dependent variables

```
In [46]: client.set.default_space(space_uid)

Out[46]: 'SUCCESS'
```

Fig 7.17 Random forest regressor

```
NAME                           ASSET_ID                              TYPE
default_py3.6                  0062b8c9-8b7d-44a0-a9b9-46c416adcbd9  base
kernel-spark3.2-scala2.12      020d69ce-7ac1-5e68-ac1a-31189867356a  base
pytorch-onnx_1.3-py3.7-edt     069ea134-3346-5748-b513-49120e15d288  base
scikit-learn_0.20-py3.6        09c5a1d0-9c1e-4473-a344-eb7b665ff687  base
spark-mllib_3.0-scala_2.12     09f4cff0-90a7-5899-b9ed-1ef348aebdee  base
pytorch-onnx_rt22.1-py3.9      0b848dd4-e681-5599-be41-b5f6fccc6471  base
ai-function_0.1-py3.6          0cdb0f1e-5376-4f4d-92dd-da3b69aa9bda  base
shiny-r3.6                     0e6e79df-875e-4f24-8ae9-62dcc2148306  base
tensorflow_2.4-py3.7-horovod   1092590a-307d-563d-9b62-4eb7d64b3f22  base
pytorch_1.1-py3.6              10ac12d6-6b30-4ccd-8392-3e922c096a92  base
tensorflow_1.15-py3.6-ddl      111e41b3-de2d-5422-a4d6-bf776828c4b7  base
autoai-kb_rt22.2-py3.10        125b6d9a-5b1f-5e8d-972a-b251688ccf40  base
runtime-22.1-py3.9             12b83a17-24d8-5082-900f-0ab31fbfd3cb  base
scikit-learn_0.22-py3.6        154010fa-5b3b-4ac1-82af-4d5ee5abbc85  base
default_r3.6                   1b70aec3-ab34-4b87-8aa0-a4a3c8296a36  base
pytorch-onnx_1.3-py3.6         1bc6029a-cc97-56da-b8e0-39c3880dbbe7  base
kernel-spark3.3-r3.6           1c9e5454-f216-59dd-a20e-474a5cdf5988  base
pytorch-onnx_rt22.1-py3.9-edt  1d362186-7ad5-5b59-8b6c-9d0880bde37f  base
tensorflow_2.1-py3.6           1eb25b84-d6ed-5dde-b6a5-3fbdf1665666  base
spark-mllib_3.2                20047f72-0a98-58c7-9ff5-a77b012eb8f5  base
tensorflow_2.4-py3.8-horovod   217c16f6-178f-56bf-824a-b19f20564c49  base
runtime-22.1-py3.9-cuda        26215f05-08c3-5a41-a1b0-da66306ce658  base
do_py3.8                       295addb5-9ef9-547e-9bf4-92ae3563e720  base
autoai-ts_3.8-py3.8            2aa0c932-798f-5ae9-abd6-15e0c2402fb5  base
tensorflow_1.15-py3.6          2b73a275-7cbf-420b-a912-eae7f436e0bc  base
kernel-spark3.3-py3.9          2b7961e2-e3b1-5a8c-a491-482c8368839a  base
pytorch_1.2-py3.6              2c8ef57d-2687-4b7d-acce-01f94976dac1  base
spark-mllib_2.3                2e51f700-bca0-4b0d-88dc-5c6791338875  base
pytorch-onnx_1.1-py3.6-edt     32983cea-3f32-4400-8965-dde874a8d67e  base
spark-mllib_3.0-py37           36507ebe-8770-55ba-ab2a-eafe787600e9  base
spark-mllib_2.4                390d21f8-e58b-4fac-9c55-d7ceda621326  base
autoai-ts_rt22.2-py3.10        396b2e83-0953-5b86-9a55-7ce1628a406f  base
xgboost_0.82-py3.6             39e31acd-5f30-41dc-ae44-60233c80306e  base
pytorch-onnx_1.2-py3.6-edt     40589d0e-7019-4e28-8daa-fb03b6f4fe12  base
pytorch-onnx_rt22.2-py3.10     40e73f55-783a-5535-b3fa-0c8b94291431  base
```

Fig 7.18 Client software specifications list

```
In [75]: rf.predict([[4.000e+00, 1.210e+02, 7.600e+01, 2.511e+03, 7.200e+01, 2.000e+00]])

Out[75]: array([23.2])
```

Fig 7.19 Prediction output

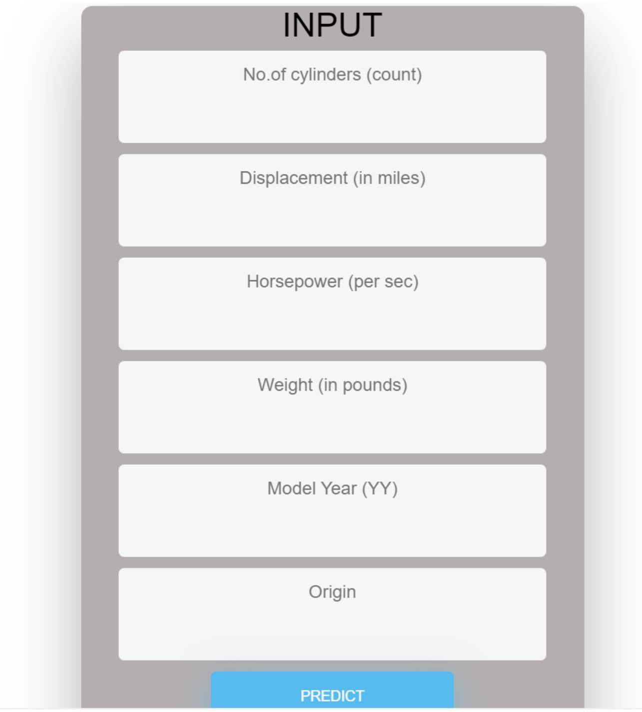


Fig 7.20 APPLICATION OUTPUT

8.CONCLUSION

As discussed in the previous section, the model developed has promising results in predicting the vehicle performance with the Regression model by constantly predicting with different train and test split ratio. The xxxxx model's performance is low only when there is low fuel efficiency repeatedly but in comparison with other models developed xxxx model's performance is exceptional and the values obtained for RMSExxx, MAExxx and R2xxx is also acceptable. Although this model was run on the data collected from small passenger cars, the model is not limited only to that class and can be generalised for any vehicle with the driving data and vehicle characteristics available. 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.

The xxxxx model underperforms only when fuel efficiency is consistently low, but when compared with other models currently being made, the xxxx model continues to perform exceptionally well, and the values observed for RMSExxx, MAExxx, and R2xxx are also acceptable. Although this model was tested using data from small passenger cars, it can be implemented to any vehicle with the assistance of available driving information and vehicle characteristics. Future research and analysis of fuel economy has more possibilities if it takes into account extra components like the state of the roadways and current traffic using Google Maps; this will enable for a more thorough examination.

APPENDIX I
SOURCE CODE

HTML:

```
<link href="//maxcdn.bootstrapcdn.com/bootstrap/4.0.0/css/bootstrap.min.css"
rel="stylesheet" id="bootstrap-css">
<link href="https://fonts.googleapis.com/css2?family=Girassol&display=swap"
rel="stylesheet">
<script src="//maxcdn.bootstrapcdn.com/bootstrap/4.0.0/js/bootstrap.min.js"></script>
<script src="//cdnjs.cloudflare.com/ajax/libs/jquery/3.2.1/jquery.min.js"></script>
<link rel="stylesheet" href="D:\IBM\css.css">
<link rel="shortcut icon" href="{{ "C:\Users\vedhe\Documents\GitHub\final
goat\Application building\Website\static\css\favicon.ico "}}>
<div class="navbar">

   <h1>PREDICT YOUR VEHICLE'S PERFORMANCE</h1>

</div>

<div class="wrapper fadeInDown">
 <div id="formContent">
  <!-- Tabs Titles -->
  <section class="date">
  <!-- Icon -->

  <div class="fadeInDown">
  <form action="{{ url_for('y_predict')}}"method="post">
    <label style="font-size:30px;">INPUT</label>
    <br>
   <input type="text" name="Cylinders" placeholder="No.of cylinders (count)"
required="required" />
```

```
        <input type="text" name="Displacement" placeholder="Displacement (in miles)"
required="required" />
        <input type="text" name="Horsepower" placeholder="Horsepower (per sec)"
required="required" />
        <input type="text" name="Weight" placeholder="Weight (in pounds)"
required="required" />
        <input type="text" name="Model Year" placeholder="Model Year (YY)"
required="required" />
        <input type="text" name="Origin" placeholder="Origin" required="required" />
        <br>
       <input type="submit" class="fadeIn fourth" value="Predict">
    </form>
    </div>
    </section>

    <div id="formFooter">
      <a class="underlineHover" href="#">
         <strong>{{ prediction_text }}</strong></a>
    </div>
   </div>
  </div>
</div>
```

<u>CSS:</u>

```
/* BASIC */

html,body {
    width: 100%;
    height: 100%;
```

```
    display: table;
}

body {
  font-family: "Poppins", sans-serif;
  display: table-cell;
  min-height: 100%;
  background-image: url('background.jpg');
  background-position: center;
  background-repeat: no-repeat;
  background-attachment: fixed;
  background-size: cover;
  margin: 0;
}

a {
  color: #92badd;
  display:inline-block;
  text-decoration: none;
  font-weight: 400;
}

h2 {
  text-align: center;
  font-size: 16px;
  font-weight: 600;
  text-transform: uppercase;
  font-family: cursive;
  display:inline-block;
  margin: 40px 8px 10px 8px;
```

```css
  color: #cccccc;
}

/* STRUCTURE */
.navbar {
  display: flex;
  font-size: 15px;
  font-family: "Muli", sans-serif;
  color: white;
  height: 20%;
  width: 100vw;
  box-shadow: 2px 2px 5px grey;
  background-color: #0a090c;
}

.date {
  display: flex;
  align-items: center;
  justify-content: center;
}

.title {
  display: flex;
  align-items: center;
  padding-left: 10px;
}

.wrapper {
  display: flex;
  align-items: center;
```

```
  flex-direction: column;
  justify-content: center;
  width: 100%;
  min-height: 100%;
  padding: 20px;
}

#formContent {
  -webkit-border-radius: 10px 10px 10px 10px;
  border-radius: 10px 10px 10px 10px;
  background: #B3AFAE ;
  padding: 30px;
  width: 90%;
  max-width: 450px;
  position: relative;
  padding: 0px;
  -webkit-box-shadow: 0 30px 60px 0 rgba(0,0,0,0.3);
  box-shadow: 0 30px 60px 0 rgba(0,0,0,0.3);
  text-align: center;
}

#formFooter {
  background-color: #2C4150;
  padding: 25px;
  text-align: center;
  -webkit-border-radius: 0 0 10px 10px;
  border-radius: 0 0 10px 10px;
  -webkit-box-shadow: 0 30px 60px 0 rgba(0,0,0,0.3);
  box-shadow: 0 30px 60px 0 rgba(0,0,0,0.3);
}
```

```
/* TABS */

h2.inactive {
  color: #cccccc;
}

h2.active {
  color: #0d0d0d;
  border-bottom: 2px solid #5fbae9;
}

/* FORM TYPOGRAPHY*/

input[type=button], input[type=submit], input[type=reset]  {
  background-color: #56baed;
  border: none;
  color: white;
  padding: 15px 80px;
  text-align: center;
  text-decoration: none;
  display: inline-block;
  text-transform: uppercase;
  font-size: 13px;
  -webkit-box-shadow: 0 10px 30px 0 rgba(95,186,233,0.6);
  box-shadow: 0 10px 30px 0 rgba(95,186,233,0.6);
  -webkit-border-radius: 5px 5px 5px 5px;
  border-radius: 5px 5px 5px 5px;
```

```
  margin: 5px 20px 40px 20px;
  -webkit-transition: all 0.3s ease-in-out;
  -moz-transition: all 0.3s ease-in-out;
  -ms-transition: all 0.3s ease-in-out;
  -o-transition: all 0.3s ease-in-out;
  transition: all 0.3s ease-in-out;
}

input[type=button]:hover, input[type=submit]:hover, input[type=reset]:hover  {
  background-color: #39ace7;
}

input[type=button]:active, input[type=submit]:active, input[type=reset]:active  {
  -moz-transform: scale(0.95);
  -webkit-transform: scale(0.95);
  -o-transform: scale(0.95);
  -ms-transform: scale(0.95);
  transform: scale(0.95);
}

input[type=text] {
  background-color: #f6f6f6;
  border: none;
  color: #0d0d0d;
  padding: 10px 32px;
  padding-bottom: 50px;
  text-align: center;
  text-decoration: none;
  display: inline-block;
  font-size: 16px;
```

```
  margin: 5px;
  width: 85%;
  border: 2px solid #f6f6f6;
  -webkit-transition: all 0.5s ease-in-out;
  -moz-transition: all 0.5s ease-in-out;
  -ms-transition: all 0.5s ease-in-out;
  -o-transition: all 0.5s ease-in-out;
  transition: all 0.5s ease-in-out;
  -webkit-border-radius: 5px 5px 5px 5px;
  border-radius: 5px 5px 5px 5px;
}

input[type=text]:focus {
  background-color: #fff;
  border-bottom: 2px solid #5fbae9;
}

input[type=text]:placeholder {
  color: #cccccc;
}

/* ANIMATIONS */

/* Simple CSS3 Fade-in-down Animation */
.fadeInDown {
  -webkit-animation-name: fadeInDown;
  animation-name: fadeInDown;
  -webkit-animation-duration: 1s;
  animation-duration: 1s;
  -webkit-animation-fill-mode: both;
```

```
  animation-fill-mode: both;
}

@-webkit-keyframes fadeInDown {
  0% {
    opacity: 0;
    -webkit-transform: translate3d(0, -100%, 0);
    transform: translate3d(0, -100%, 0);
  }
  100% {
    opacity: 1;
    -webkit-transform: none;
    transform: none;
  }
}
.col-md-12 .inputDefault
{
    background-color:#DCDCDC;
    height:40px;
    width: 50%;
    display: inline-block;
}
.col-md-12 .form-group {
margin-bottom: 0;
}
.col-md-12 .form-group:nth-child(1n)
{
    display: inline-block;
    width: 50%;
    padding-top: 10px;
```

```
}
.col-md-12 .form-group:nth-child(1n)::before
{
    content: "";
    border-top: 1px solid #333;
    width: 50%;
    display: inline-block;
    padding-bottom: 10px;
}
.col-md-12 .form-group:nth-child(1)::before {
display: none;
}

@keyframes fadeInDown {
  0% {
    opacity: 0;
    -webkit-transform: translate3d(0, -100%, 0);
    transform: translate3d(0, -100%, 0);
  }
  100% {
    opacity: 1;
    -webkit-transform: none;
    transform: none;
  }
}

/* Simple CSS3 Fade-in Animation */
@-webkit-keyframes fadeIn { from { opacity:0; } to { opacity:1; } }
@-moz-keyframes fadeIn { from { opacity:0; } to { opacity:1; } }
@keyframes fadeIn { from { opacity:0; } to { opacity:1; } }
```

```
.fadeIn {
  opacity:0;
  -webkit-animation:fadeIn ease-in 1;
  -moz-animation:fadeIn ease-in 1;
  animation:fadeIn ease-in 1;

  -webkit-animation-fill-mode:forwards;
  -moz-animation-fill-mode:forwards;
  animation-fill-mode:forwards;

  -webkit-animation-duration:1s;
  -moz-animation-duration:1s;
  animation-duration:1s;
}

.fadeIn.first {
  -webkit-animation-delay: 0.4s;
  -moz-animation-delay: 0.4s;
  animation-delay: 0.4s;
}

.fadeIn.second {
  -webkit-animation-delay: 0.6s;
  -moz-animation-delay: 0.6s;
  animation-delay: 0.6s;
}

.fadeIn.third {
  -webkit-animation-delay: 0.8s;
```

```
  -moz-animation-delay: 0.8s;
  animation-delay: 0.8s;
}

.fadeIn.fourth {
  -webkit-animation-delay: 1s;
  -moz-animation-delay: 1s;
  animation-delay: 1s;
}

/* Simple CSS3 Fade-in Animation */
.underlineHover:after {
  display: block;
  left: 0;
  bottom: -10px;
  width: 0;
  height: 2px;
  background-color: #56baed;
  content: "";
  transition: width 0.2s;
}

.underlineHover:hover {
  color: #0d0d0d;
}

.underlineHover:hover:after{
  width: 100%;
}
/* OTHERS */
```

```
*:focus {
   outline: none;
}

#icon {
  width:60%;
}
```

APPLICATION :

```
from flask import Flask, request, render_template
import requests

app = Flask(__name__)



@app.route('/')
def home():
    return render_template('index.html')



@app.route('/y_predict', methods=['POST'])
def y_predict():
    '''
    For rendering results on HTML GUI
    '''
    x_test = [[int(x) for x in request.form.values()]]
    print("xtest= ", x_test)
    # sc = load('scalar.save')
```

```
# NOTE: you must manually set API_KEY below using information retrieved from your
IBM Cloud account.
API_KEY = "k0ToNjB4fREMsVxEr0C3pjHT0bNJzgZvVt1S0SikVpMJ"
token_response = requests.post('https://iam.cloud.ibm.com/identity/token', data={"apikey":
                                                         API_KEY,
                                                        "grant_type":
'urn:ibm:params:oauth:grant-type:apikey'})
mltoken = token_response.json()["access_token"]

header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}

# NOTE: manually define and pass the array(s) of values to be scored in the next line
payload_scoring = {
    "input_data": [{"field": [['cylinders', 'displacement', 'horsepower', 'weight', 'model year',
'origin']],
            "values": x_test}]}

response_scoring = requests.post(
    'https://us-south.ml.cloud.ibm.com/ml/v4/deployments/3950d430-efb8-43ea-b408-
28233df071d7/predictions?version=2022-11-13',
    json=payload_scoring,
    headers={'Authorization': 'Bearer ' + mltoken})
print("Scoring response")
prediction = response_scoring.json()
print(prediction)
output = prediction['predictions'][0]['values'][0]
output = output[0]
print(output)
if (output <= 9):
```

```
        pred = "Worst performance with mileage " + str(output)
    if (output > 9 and output <= 17.5):
        pred = "Low performance with mileage " + str(output)
    if (output > 17.5 and output <= 29):
        pred = "Medium performance with mileage " + str(output)
    if (output > 29 and output <= 46):
        pred = "High performance with mileage " + str(output)
    if (output > 46):
        pred = "Very high performance with mileage " + str(output)

    return render_template('index.html', prediction_text='{}'.format(pred))


if __name__ == "__main__":
    app.run(debug=True)
```

MODEL :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import statsmodels.formula.api as smf
dataset.isnull().any()
dataset['horsepower']=dataset['horsepower'].replace('?',np.nan)
dataset['horsepower'].isnull().sum()
dataset['horsepower']=dataset['horsepower'].astype('float64')
dataset['horsepower'].fillna((dataset['horsepower'].mean()),inplace=True)
dataset.isnull().any()
dataset.info() #Pandas dataframe.info() function is used to get a quick overview of the
dataset.
dataset.describe() #Pandas describe() is used to view some basic statistical details of a data
```

frame or a series of numeric values.

```
dataset=dataset.drop('car name',axis=1) #dropping the unwanted column.
corr_table=dataset.corr()#Pandas dataframe.corr() is used to find the pairwise correlation of all columns in the dataframe.
corr_table
sns.heatmap(dataset.corr(),annot=True,linecolor ='black', linewidths = 1)#Heatmap is a way to show some sort of matrix plot,annot is used for correlation.
fig=plt.gcf()
fig.set_size_inches(8,8)
sns.pairplot(dataset,diag_kind='kde') #pairplot represents pairwise relation across the entire dataframe.
plt.show()
sns.regplot(x="cylinders", y="mpg", data=dataset)
sns.regplot(x="displacement", y="mpg", data=dataset)
sns.regplot(x="horsepower", y="mpg", data=dataset)
sns.regplot(x="weight", y="mpg", data=dataset)
sns.regplot(x="acceleration", y="mpg", data=dataset)
sns.regplot(x="model year", y="mpg", data=dataset)
sns.regplot(x="origin", y="mpg", data=dataset)
sns.set(style="whitegrid")
sns.boxplot(x=dataset["mpg"])
from scipy import stats
pearson_coef, p_value = stats.pearsonr(dataset['cylinders'], dataset['mpg'])
print("The Pearson Correlation Coefficient is", pearson_coef, " with a P-value of P =", p_value)
pearson_coef, p_value = stats.pearsonr(dataset['displacement'], dataset['mpg'])
print("The Pearson Correlation Coefficient is", pearson_coef, " with a P-value of P =", p_value) pearson_coef, p_value = stats.pearsonr(dataset['horsepower'], dataset['mpg'])
print("The Pearson Correlation Coefficient is", pearson_coef, " with a P-value of P =", p_value)
pearson_coef, p_value = stats.pearsonr(dataset['weight'], dataset['mpg'])
print("The Pearson Correlation Coefficient is", pearson_coef, " with a P-value of P =", p_value)
pearson_coef, p_value = stats.pearsonr(dataset['acceleration'], dataset['mpg'])
print("The Pearson Correlation Coefficient is", pearson_coef, " with a P-value of P =",
```

```
p_value)
pearson_coef, p_value = stats.pearsonr(dataset['model year'], dataset['mpg'])
print("The Pearson Correlation Coefficient is", pearson_coef, " with a P-value of P =",
p_value)
pearson_coef, p_value = stats.pearsonr(dataset['origin'], dataset['mpg'])
print("The Pearson Correlation Coefficient is", pearson_coef, " with a P-value of P =",
p_value)
test=smf.ols('mpg~cylinders+displacement+horsepower+weight+acceleration+origin',dataset)
.fit()
test.summary()
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)
```

APPENDIX II :