CAR RESALE VALUE PREDICTION

A PROJECT REPORT

Submitted by

MOHAMMED NISHAR S	**(721819104014)**
AVINASHLINGAM S	**(721819104003)**
MOHAMMED JAMEAL J	**(721819104013)**
ROHITH R V	**(721819104021)**

in the partial fulfilment of the requirements for the award of a degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

RATHINAM TECHNICAL CAMPUS, COIMBATORE.

ANNA UNIVERSITY : CHENNAI 600 025

NOVEMBER 2022

INDEX

Project Report

1. INTRODUCTION

1.1 Project Overview

In this project we have used different algorithms with different techniques for developing Car resale value prediction systems considering different features of the car. In a nutshell, car resale value prediction helps the user to predict the resale value of the car depending upon various features like kilometres driven, fuel type, etc.

This resale value prediction system is made for general purpose to just predict the amount that can be roughly acquired by the user.

1.2 Purpose

Used car resale market in India was marked at 24.2 billion US dollars in 2019. Due to the huge requirement of used cars and lack of experts who can determine the correct valuation, there is an utmost need of bridging this gap between sellers and buyers. This project focuses on building a system that can accurately predict a resale value of the car based on minimal features like kms driven, year of purchase etc. without manual or human interference and hence it remains unbiased.

The main idea of making a car resale value prediction system is to get hands-on practice for python using Data Science. Car resale value prediction is the system to predict the amount of resale value based on the parameters provided by the user. User enters the details of the car into the form given and accordingly the car resale value is predicted.

Determining whether the listed price of a used car is a challenging task, due to the many factors that drive a used vehicle’s price on the market. The focus of this project is developing machine learning models that can accurately predict the price of a used car based on its features, in order to make informed purchases.

2. LITERATURE SURVEY

2.1 Existing problem

The rise of online websites and other tools like it have made it easier for both buyers and sellers to get a better understanding of the factors that determine the market value of a used car. Based on a set of factors, Machine Learning algorithms may be used to forecast the price of any automobile. The cost is calculated using the amount of characteristics. They used linear regression and lasso regression to develop a price model for used automobiles in comparative research. The main goal of this study is to discover the best predictive model for estimating the price of a used car.

In existing model, it was to estimate the cost of the used cars using the K nearest neighbour algorithm which is simple and suitable for small data set. Here, they have collected a used cars dataset and analysed the same. The data was trained by the model and examined the accuracy of the model among different ratios of trained and test set. The same model is cross-validated for assessing the performance of the model using the K- Fold method which is easy to understand and implement. They have used the K nearest Neighbour algorithm and got accuracy 85% where the accuracy of linear regression is 71%. The proposed model is also validated with 5 and 10 folds by using K Fold Method.

2.2 References

1. CAR PRICE PREDICTION USING MACHINE LEARNING [KETAN AGRAHARI, AYUSH CHAUBEY ET AL, 2021]

2. VEHICLE RESALE PRICE PREDICTION USING MACHINE LEARNING": [B.LAVANYA, SK.RESHMA, N.NIKITHA, M.NAMITHA , L.KANYA KUMARI, S.KISHORE BABU, 2021]

3. AN EXPERT SYSTEM OF PRICE FORECASTING FOR USED VEHICLES USING ADAPTIVE NEURO-FUZZY INFERENCE [WU, ET AL, 2009]

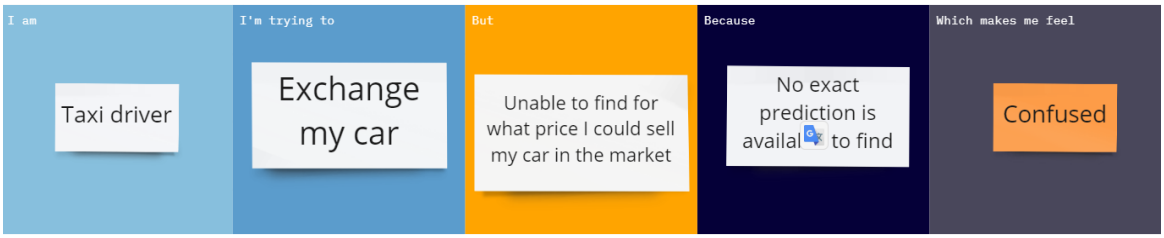
4. USED CAR PRICE PREDICTION USING K-NEAREST NEIGHBOR BASED MODEL [SAMRUDDHI, ASHOK KUMAR, 2020]

2.3 Problem Statement Definition

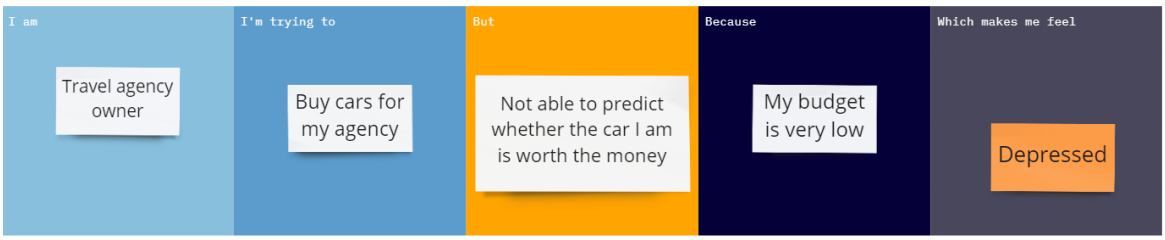
The research objective of this study is to predict used cars prices, using data mining techniques, by scraping data from websites that sell used cars, and analysing the different aspects and factors that lead to the actual used car price valuation. To enable consumers to know the actual worth of their car or desired car, by simply providing the program with a set of attributes from the desired car to predict the car price. The purpose of this study is to understand and evaluate used car prices and to develop a strategy that utilizes data mining techniques to predict used car prices.

PROBLEM STATEMENT 1

PROBLEM STATEMENT 2

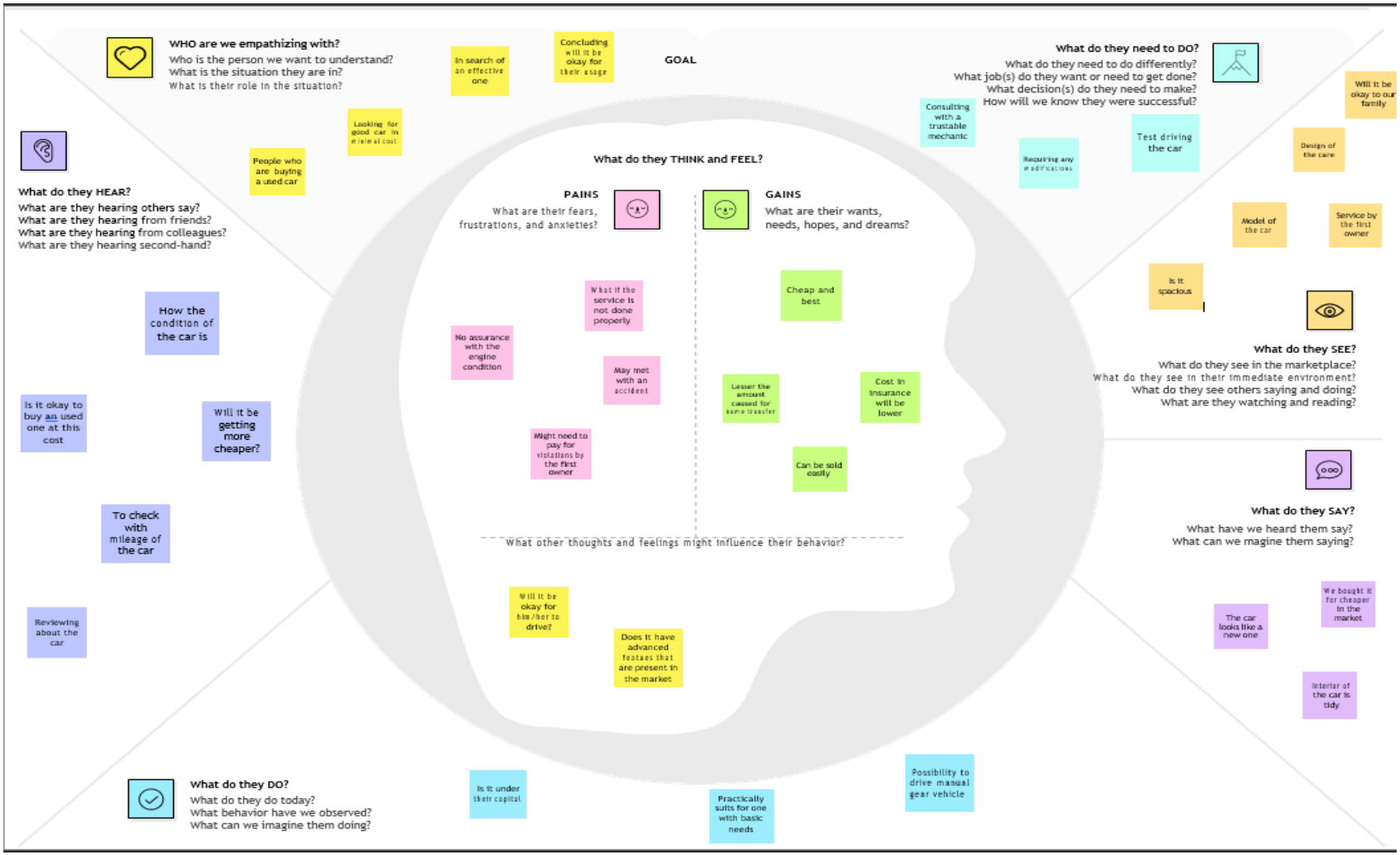


PROBLEM STATEMENT 3

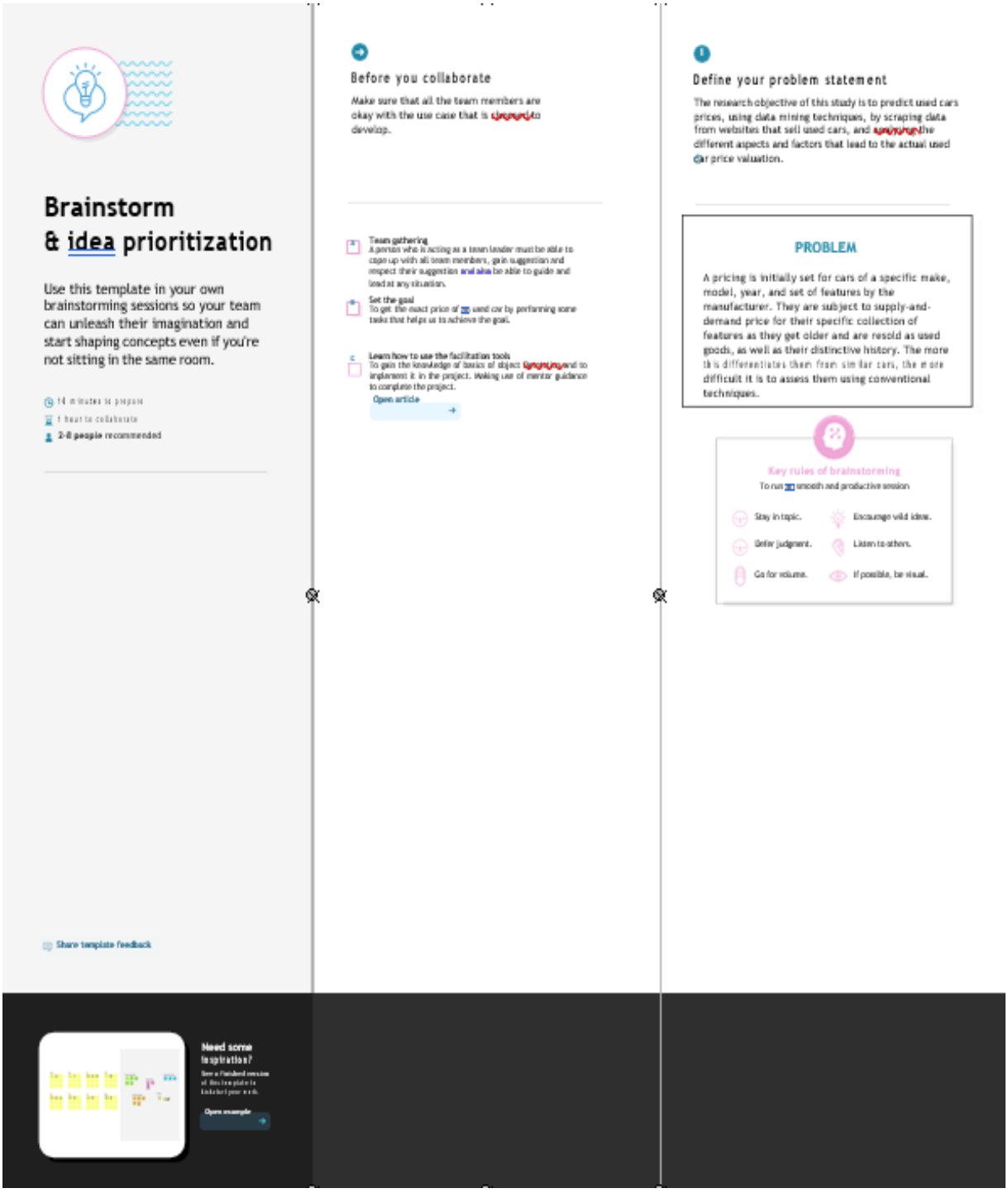


3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



3.2 Ideation & Brainstorming



2

Brainstorm

Write down any ideas that come to mind that address your problem statement.

MOHAMMED NISHAR

ROHITH

To build a tidy code

AVINASHLINGAM

MOHAMMED JAMEEL

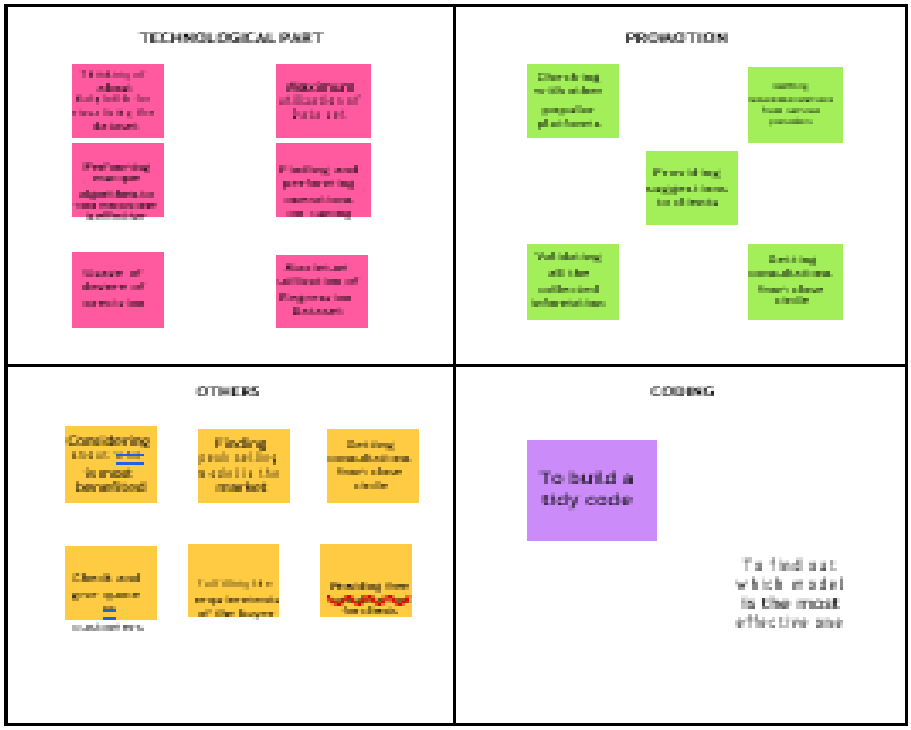
Finding post selling details the market

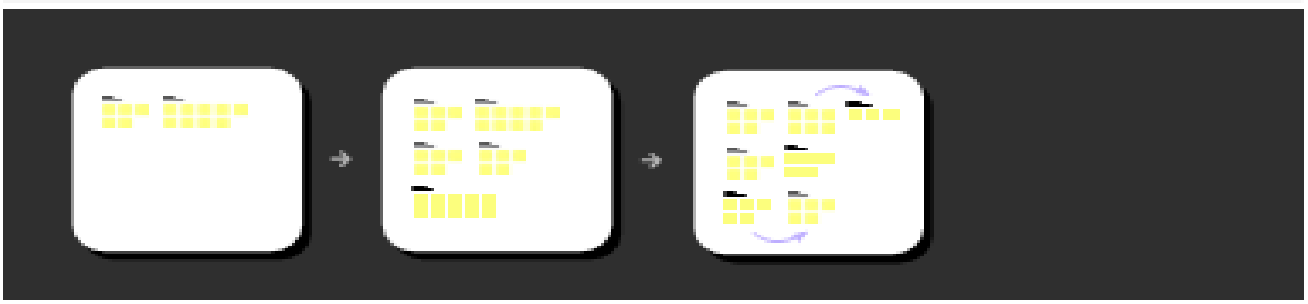
3

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.

GROUPING OF IDEAS

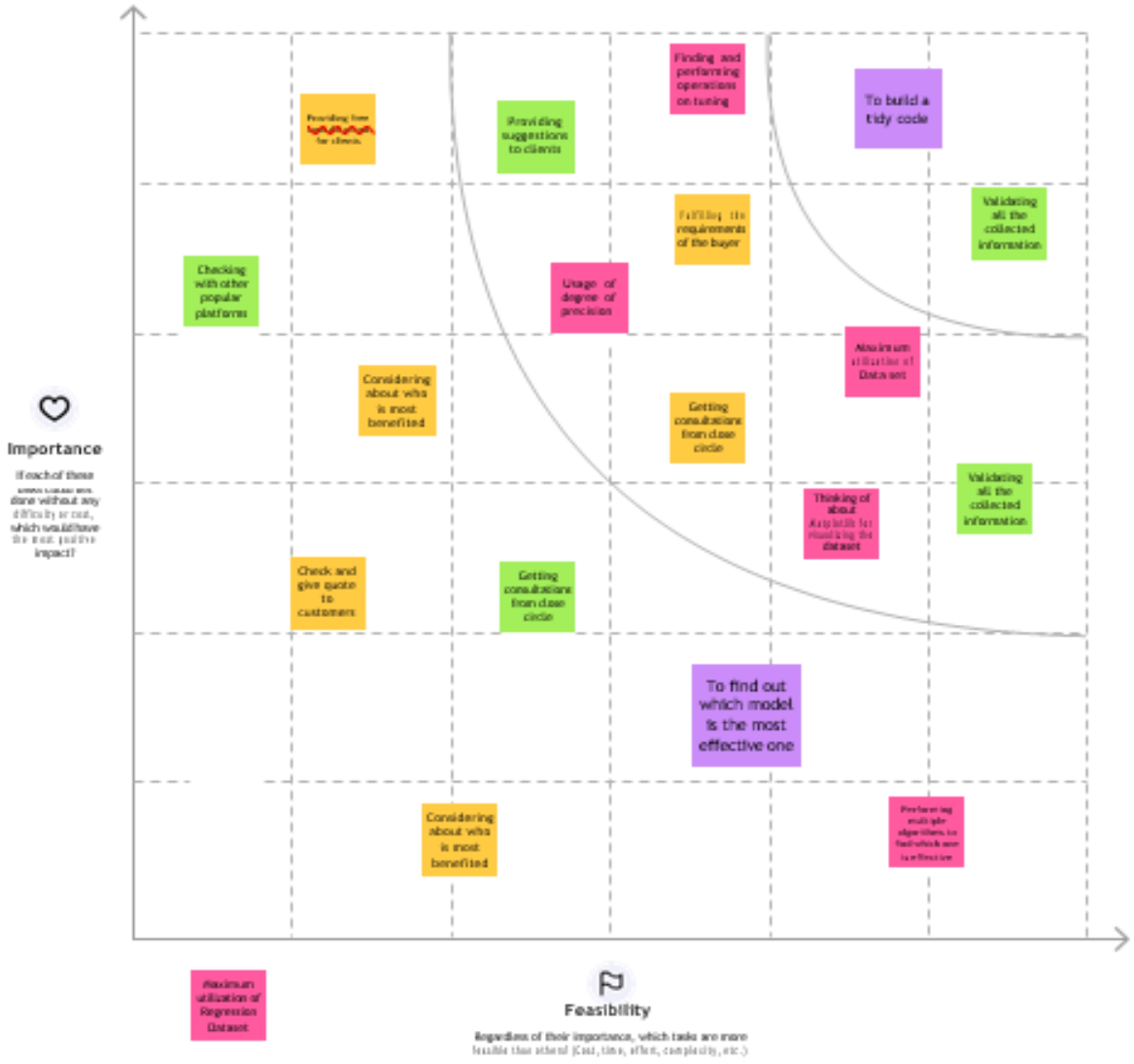




4

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.



3.3 Proposed Solution

S.No	Parameter	Description
1.	Problem Statement (Problem to be solved)	Any person who is going to buy a car for the first time in his/her life would have not much idea about the car. If a person is going to buy a used car, then the person should conduct some tests on that car for its effectiveness. Manual tests may not be effective. This is where the proposed solution is implemented, conducting a set of tasks which are trained using machine learning concepts, The project should take parameters related to used car as inputs and enable the customers to make decisions by their own
2.	Idea / Solution description	The retail price of a new car, excluding extra equipment, is the same for all vehicles of the same make, model, and year. The producer determines this cost. However, the price of used cars is determined by supply and demand. Additionally, there are other characteristics of second-hand cars that affect the price. These distinguish between cars that may have had a similar retail price by taking into account their condition, mileage, and repair history.
3.	Novelty / Uniqueness	Utilizing the built programme, used vehicle price may quickly and accurately assess a car's value on its own. Users can enter the car model, miles travelled, type of car, and the number of submitted car photographs. Based on the data, the car price will be forecasted.
4.	Social Impact / Customer Satisfaction	Once this idea is implemented, let it be anyone might be normal person looking for his household or a dealer who is a reseller of cars or a cab driver could get accurate quote of the vehicle he/she is going to buy just by passing all the relevant information that are needed to analyse the value of a car. Just by Machine learning's understanding as it relates to

		automobile valuations and other comparable price prediction issues will grow because of this. The purpose of this effort is to address the research questions. They all involve contrasting several machine learning (ML) price prediction methods. This will be done by collecting and generating a dataset that will allow for fair training and comparison of all the algorithms. For the same dataset to be utilised for all the methods, they must be similar enough. Additionally, if the adjustments do not improve the performance of the other models, no significant optimizations of the dataset will be made. It is outside the purview of this work to maximise the price prediction performance of any one algorithm in ways that do not provide improved comparisons.
5.	Business Model (Revenue Model)	This implementation helps in customers save a lot of money as the customer directly deals with the seller and is able to get the exact quote without the help of any third person like dealer or platforms like CARS24. A Business model is a plan that outlines how a new company will make money from its regular business operations and how it will cover its operational costs and expenses in the desired format.
6.	Scalability of the Solution	As the model developed is able to predict the resale value of an used by collecting some information about the car. It gives a 90% accurate result. The scalability of the model is very high because the developed model could give the quote for any type of car. The best parameters were chosen by trial and error as the models were being developed, and the cache model was then implemented with the settings that produced the best performance. All of the models approximated geometric appreciation, which means that regardless of the age of the vehicle, a fixed percentage of value is lost year.

3.4 Problem Solution fit

The problem solution Fit simply means that you have identified a customer-related issue and that the resolution you have developed genuinely addresses the issue. It assists business owners, marketers, and corporate innovators in seeing behavioural trends and understanding what would be successful.

Purpose:

- Solve complex problems in a way that fits the state of your customers.
- Succeed faster and increase your solution adoption by tapping into existing mediums and channels of behaviour.
- Sharpen your communication and marketing strategy with the right triggers and messaging.
- Increase touch-points with your company by finding the right problem-behaviour fit and building trust by solving frequent annoyances, or urgent or costly problems.
- Understand the existing situation in order to improve it for your target group.

PROJECT TITLE: CAR RESALE VALUE PREDICTION | PROJECT DESIGN PHASE – I | TEAM ID: PNT2022TMID43743

Problem-Solution fit canvas 2.0

Purpose / Vision

PROBLEM SOLUTION FIT

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) — CS

Who is your customer?
i.e. working parents of 0-5 y.o. kids

- A normal person trying to assess the worth of his asset
- A taxi driver trying to buy a used car
- A car reseller looking for a car for his client
- Anyone trying to know the value of a used car

6. CUSTOMER CONSTRAINTS — CC

What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices.

- To estimate a car's value on your own, without assistance of someone.
- To prevent human error in prediction. Its not possible to get the exact worth manually, but Machine Learning Techniques can.
- Reduction in amount spent to dealers

5. AVAILABLE SOLUTIONS — AS

Which solutions are available to the customers when they face the problem or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking

- We would have chosen the random forest approach over the linear regression algorithm.
- The goal of the system for predicting automobile resale value is to use data mining techniques on a set of vehicle data to help with that forecast.

Explore AS, differentiate

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

2. JOBS-TO-BE-DONE / PROBLEMS — J&P

Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.

- When using ML approaches to test the autos, any damages must be appropriately assessed.
- With the available methods, only rough predictions can be made; an exact outcome cannot be provided.
- Not every factor is taken into account when testing.

9. PROBLEM ROOT CAUSE — RC

What is the real reason that this problem exists?
What is the back story behind the need to do this job?
i.e. customers have to do it because of the change in regulations.

- The fear of the car's state, the engine's condition, the gasoline type, the mileage of the vehicle, and physical damage are the main risk factors for estimating values and believing anonymous vendors. The best course of action is to avoid buying from strangers, shop around for the best deal, inspect the vehicle, and use predictive analysis to make predictions.

7. BEHAVIOUR — BE

What does your customer do to address the problem and get the job done?
i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)

- They presented paperwork and a record of your car's condition that is questionable. The model was just created to get rid of human-predicted anonymous value.
- Create or enhance the strategy vision. Identify customer segments via vehicle customization.
- Having trouble forecasting second-hand car values and relying on unidentified brokers

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

Identify strong TR & EM

3. TRIGGERS — TR

What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.

- Create or enhance the strategy vision. Segment buyers and users can use vehicle

4. EMOTIONS: BEFORE / AFTER — EM

How do customers feel when they face a problem or a job and afterwards?
i.e. lost, insecure > confident, in control - use it in your communication strategy & design.

BEFORE SELLING: Depressed, Frustrated **AFTER SELLING:** Happy
BEFORE BUYING: Confused
AFTER BUYING: Feels Joyous if cheaper

10. YOUR SOLUTION — SL

If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality.
If you are working on a new business proposition, then keep it blank until you fill in the canvas and come up with a solution that fits within customer limitations, solves a problem and matches customer behaviour.

- We employ supervised learning algorithms, which can be resolved by random forset regression. Even with the enormous dataset, it operates effectively and predicts the outcome with a high degree of accuracy. When a significant amount of the data is absent, accuracy can still be maintained.

8. CHANNELS of BEHAVIOUR — CH

8.1 ONLINE
What kind of actions do customers take online? Extract online channels from #7

- Utilizing the many parameters provided by the owner, the customer should be able to estimate the value of the vehicle, Customers can compare with multiple sources

8.2 OFFLINE
What kind of actions do customers take offline? Extract offline channels from #7 and use them for customer development.

- Offline service providers like Dealers, Resellers, would face some difficulties.

Extract online & offline CH of BE

AMALTAMA

4. REQUIREMENT ANALYSIS

4.1 Functional requirements

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 New use portal Registration through gmail
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Libraries & Modules	**NumPy** for scientific computing **Pandas** for using open-source data analysis **Sklearn** for using machine learning techniques
FR-4	Frameworks	Flask framework for building Web applications
FR-5	IDE	VS code Anaconda Navigator for using Machine learning related techniques

4.2 Non-Functional requirements

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	Estimating the value at resale
NFR-2	**Security**	Offering the website with security policy
NFR-3	**Reliability**	Because it could estimate the worth of any type of car, it would be highly reliable.
NFR-4	**Performance**	Implementing a website with heavier performance, that does not lag on type pf input data
NFR-5	**Availability**	The website would be easily to anyone, since it is a website its accessible form any part of the world

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.

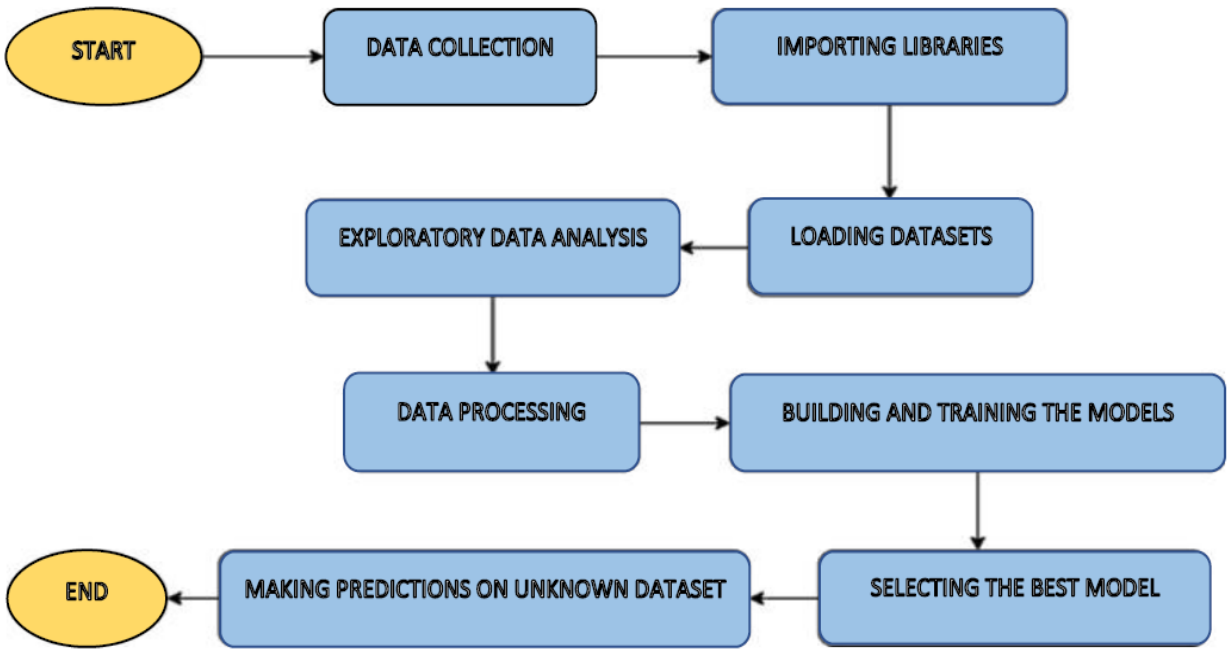
These diagrams are used as visualization tool to help the audience get a better idea of what exactly is going on in the system. The DFDs are use to :

(i) discuss with the user a diagrammatic interpretation of the process in the system and clarify what is currently being performed.

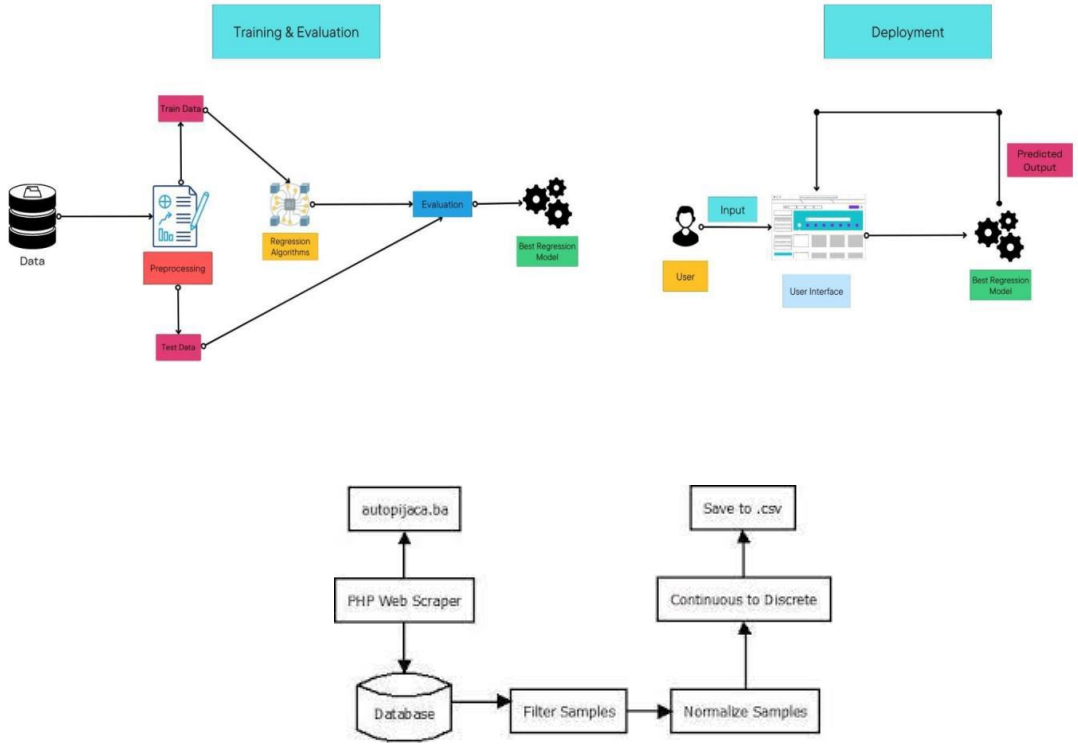
(ii) determine what the new system should be able to do and what information is required for each different process the should be carried out.

(iii) Check that the completed system conforms to its intended design.

(iv) provides easy presentation and communication between technical and non technical staff.



5.2 Solution & Technical Architecture



S.No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / React Js etc.
2.	Application Logic-1	Logic for a process in the application	Python
3.	Database	Data Type, Configurations etc.	NoSQL
4.	Cloud Database	Database Service on Cloud	IBM DB2
5.	File Storage	File storage requirements	IBM Block Storage
6.	Machine Learning Model	Purpose of Machine Learning Model	Regression model
7.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

5.3 User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	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 Facebook	I can register & access the dashboard with Facebook Login	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail		Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard					
Customer (Web user)	Registration	USN-1	As a user, I can register and enter my car details	I can enter the details of my asset	Medium	Sprint-1

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

TITLE	DESCRIPTION	DATE
LITERATURE SURVEY	Analysis of the use case chosen by referring various journals and research papers	20 SEPTEMBER 2022
INFORMATION GATHERING	Collecting data by consulting technical documents, research articles, etc.	24 SEPTEMBER 2022
PREPARATION OF EMPATHY MAP	Prepare a list of problem statements in publications, a canvas for an empathy map to capture the user's gains and pains, etc.	25 SEPTEMBER 2022
IDEATION	List the ideas generated during the brainstorming session and rank the top three according to relevance and viability.	30 SEPTEMBER 2022
PROPOSED SOLUTION	Create a proposal for a solution that details its innovation, viability as a business idea, social impact, scalability, and other factors.	01 OCTOBER 2022
PROBLEM SOLUTION FIT	Get a problem-solution-fit document ready.	06 OCTOBER 2022
SOLUTION ARCHITECTURE	Document the solution architecture.	09 OCTOBER 2022
CUSTOMER JOURNEY	Create customer journey maps to comprehend how users engage with and use the application from entry to exit.	14 OCTOBER 2022
FUNCTIONAL REQUIREMENT	Document the functional requirements.	16 OCTOBER 2022
DATA FLOW DIAGRAMS	Create the data flow diagrams, then submit them for evaluation.	16 OCTOBER 2022
TECHNOLOGY ARCHITECTURE	Create the diagram of the technological architecture.	20 OCTOBER 2022
PREPARE MILESTONE & ACTIVITY LIST	Create a list of the project's milestones and activities.	28 OCTOBER 2022
PROJECT DEVELOPMENT - DELIVERY OF SPRINT-1, 2, 3 & 4	Create the code, develop it and submit it after testing it.	IN PROGRESS

6.2 Sprint Delivery Schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Pre-processing of data	USN-1	Collect Dataset	5	High	Mohammed Nishar
Sprint-1		USN-2	Import required libraries	5	High	Rohith
Sprint-1		USN-3	Read and clean data sets	2	Medium	Mohammed Jameal
Sprint-2	Building the Model	USN-1	Split data into independent and dependent variables	4	High	Avinashlingam
Sprint-2		USN-2	Apply using regression model	2	Medium	Rohith
Sprint-3	Application building	USN-1	Build python flask application and HTML page	5	High	Avinashlingam
Sprint-3		USN-2	Execute and test the application	1	Low	Mohammed Nishar
Sprint-4	Training the model	USN-1	Train machine learning model that was built	5	High	Mohammed Jameal
Sprint-4		USN-2	Integrate flask	5	High	Avinashlingam

7. CODING & SOLUTIONING

(Explain the features added in the project along with code)

```
# Import Libraries
import pandas as pd
import numpy as np
from flask import Flask, render_template, Response, request
import pickle
from sklearn.preprocessing import LabelEncoder
import requests

# NOTE: you must manually set API_KEY below using information retrieved from your IBM
Cloud account.
API_KEY = "iGmBU8ohAyWS6fHk34WZVDB2xFYfOP3y09AFaAwbnf87"
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}



app = Flask(__name__)#initiate flask app

def load_model(file='../Result/resale_model.sav'):#load the saved model
        return pickle.load(open(file, 'rb'))

@app.route('/')
def index():#main page
        return render_template('car.html')

@app.route('/predict_page')
def predict_page():#predicting page
        return render_template('value.html')

@app.route('/predict', methods=['GET','POST'])
def predict():
        reg_year = int(request.args.get('regyear'))
        powerps = float(request.args.get('powerps'))
        kms= float(request.args.get('kms'))
        reg_month = int(request.args.get('regmonth'))

        gearbox = request.args.get('geartype')
```

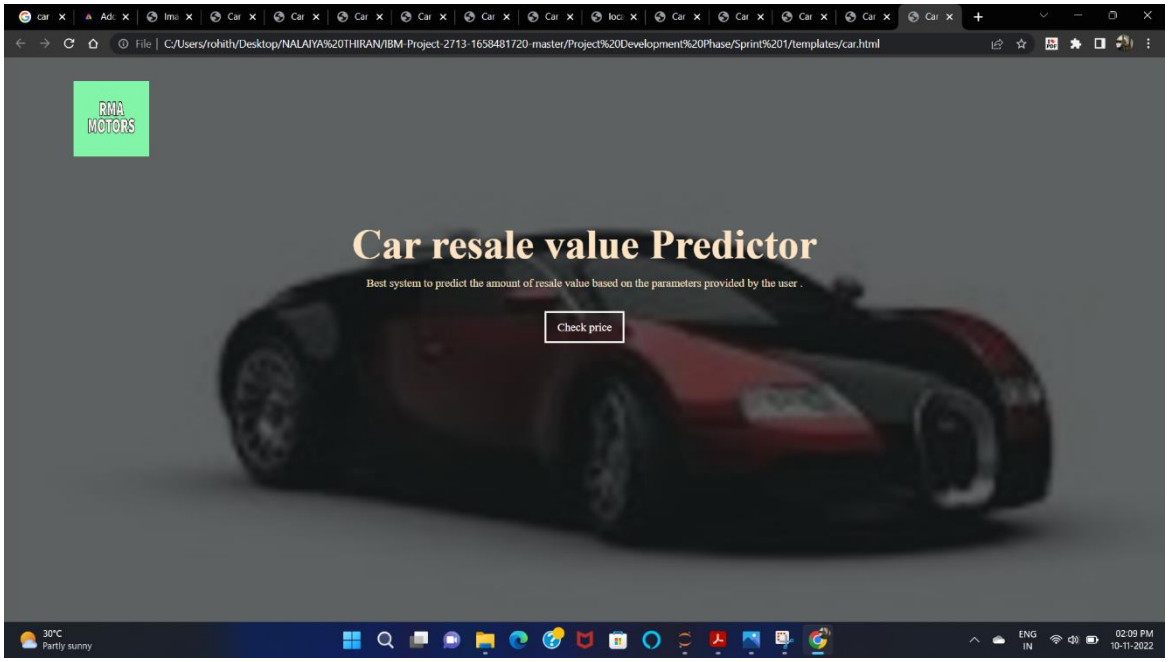
```
        damage = request.args.get('damage')
        model = request.args.get('model')
        brand = request.args.get('brand')
        fuel_type = request.args.get('fuelType')
        veh_type = request.args.get('vehicletype')

        new_row = {'yearOfReg':reg_year, 'powerPS':powerps, 'kilometer':kms,
                                'monthOfRegistration':reg_month, 'gearbox':gearbox,
                                'notRepairedDamage':damage,
                                'model':model, 'brand':brand, 'fuelType':fuel_type,
                                'vehicletype':veh_type}

        print(new_row)

        new_df = pd.DataFrame(columns=['vehicletype','yearOfReg','gearbox',
                'powerPS','model','kilometer','monthOfRegistration','fuelType',
                'brand','notRepairedDamage'])
        new_df = new_df.append(new_row, ignore_index=True)
        labels = ['gearbox','notRepairedDamage','model','brand','fuelType','vehicletype']
        mapper = {}

        for i in labels:
                mapper[i] = LabelEncoder()
                mapper[i].classes = np.load('../Result/'+str('classes'+i+'.npy'), allow_pickle=True)
                transform = mapper[i].fit_transform(new_df[i])
                new_df.loc[:,i+'_labels'] = pd.Series(transform, index=new_df.index)

        labeled = new_df[['yearOfReg','powerPS','kilometer','monthOfRegistration'] + [x+'_labels'
for x in labels]]

        X = labeled.values.tolist()
        print('\n\n', X)
        #predict = reg_model.predict(X)

        # NOTE: manually define and pass the array(s) of values to be scored in the next line
        payload_scoring = {"input_data": [{"fields": [['yearOfReg', 'powerPS', 'kilometer',
'monthOfRegistration','gearbox_labels', 'notRepairedDamage_labels',
'model_labels','brand_labels', 'fuelType_labels', 'vehicletype_labels']], "values": X}]}

        response_scoring = requests.post('https://us-
south.ml.cloud.ibm.com/ml/v4/deployments/7f67cbed-6222-413b-9901-
```

```
b2a72807ac82/predictions?version=2022-10-30', json=payload_scoring, headers={'Authorization':
'Bearer ' + mltoken})
        predictions = response_scoring.json()
        print(response_scoring.json())
        predict = predictions['predictions'][0]['values'][0][0]
        print("Final prediction :",predict)

        return render_template('predict.html',predict=predict)

if __name__=='__main__':
        reg_model = load_model()#load the saved model
        app.run(host='localhost', debug=True, threaded=False)
```

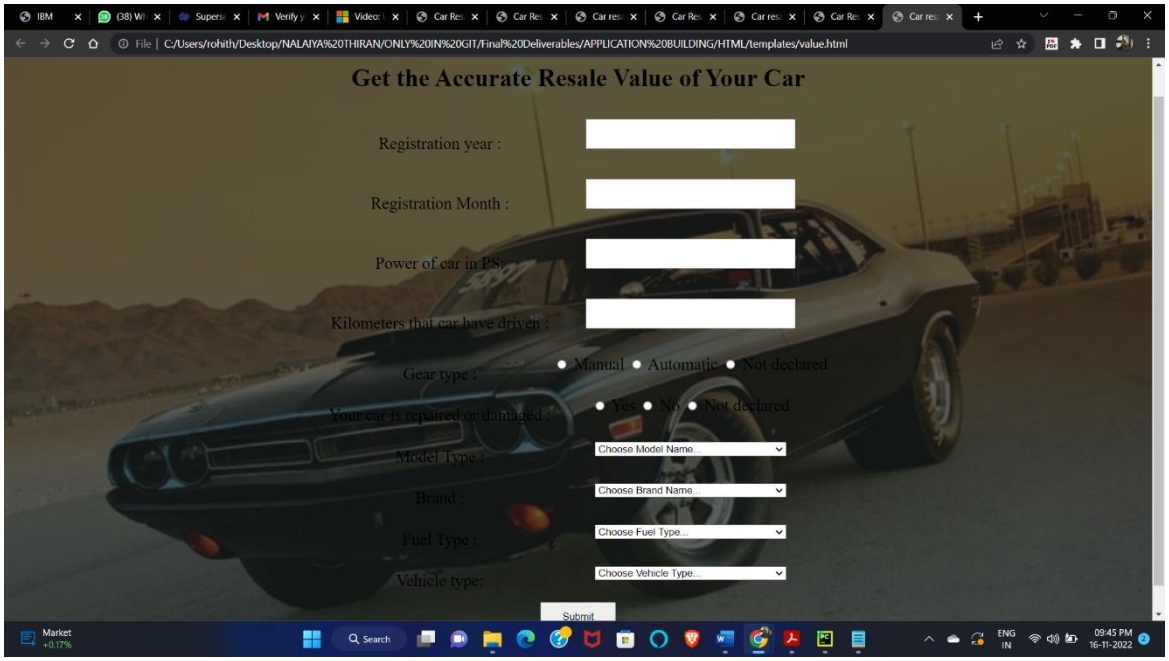
8. TESTING

8.1 Test Cases

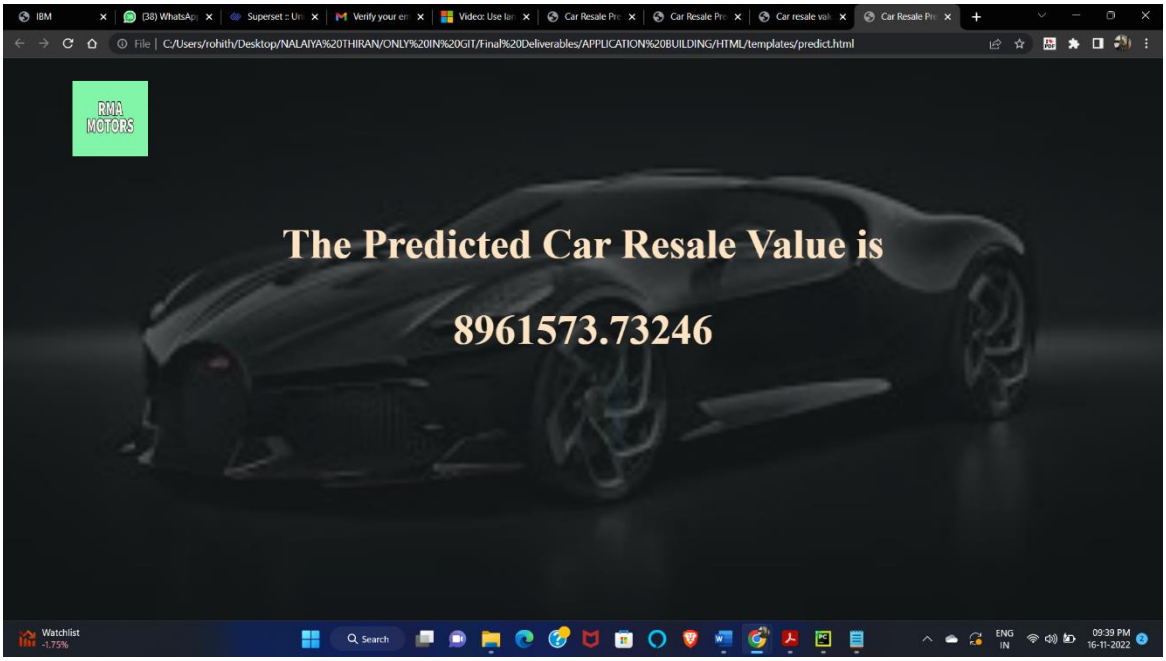
1. LOGIN PAGE



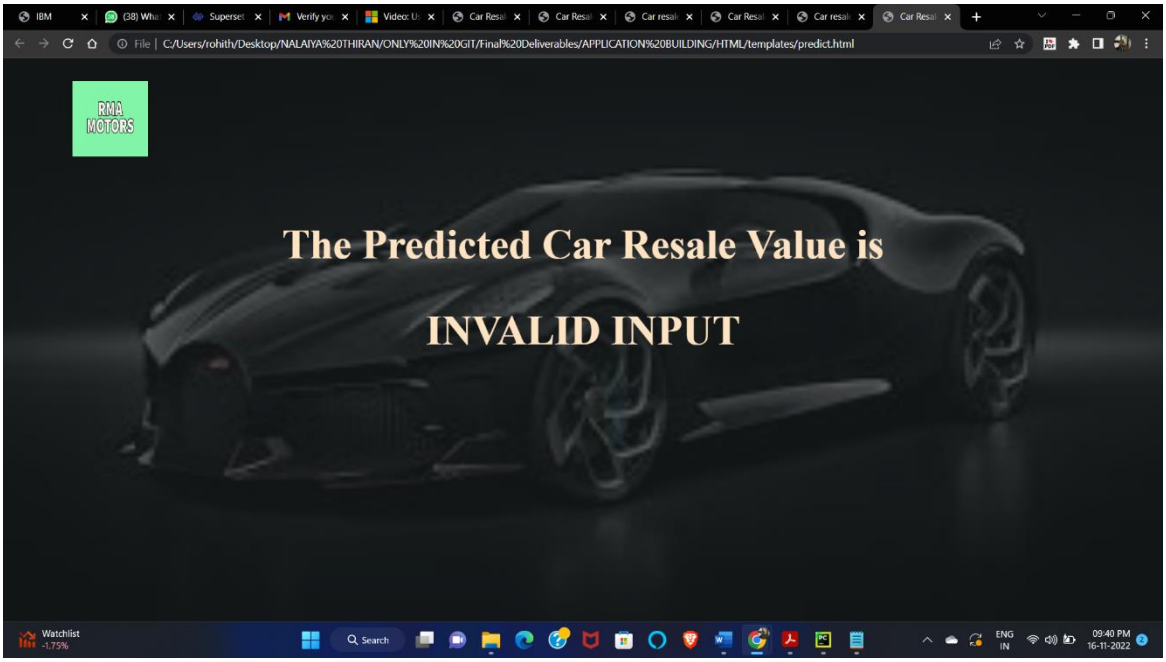
2. USER INPUT PAGE



3. CORRECT INPUT CASE



2. INVALID INPUT CASE



8.2 User Acceptance Testing

User Acceptance Testing (UAT) is a type of testing performed by the end user or the client to verify/accept the software system before moving the software application to the production environment. UAT is done in the final phase of testing after functional, integration and system testing is done.

The main **Purpose of UAT** is to validate end to end business flow. It does not focus on cosmetic errors, spelling mistakes or system testing. User Acceptance Testing is carried out in a separate testing environment with production-like data setup. It is kind of black box testing where two or more end-users will be involved.

UAT is performed by –

- Client
- End users

On Performing User Acceptance Testing the trained model gives the expected result which in turn meets the business requirements and satisfies all the conditions

UAT is the last chance to identify and rectify defects. Businesses may suffer losses if UAT is not performed properly (or at all). The losses that may occur (by fixing system issues after production), are much more expensive than fixing before production.

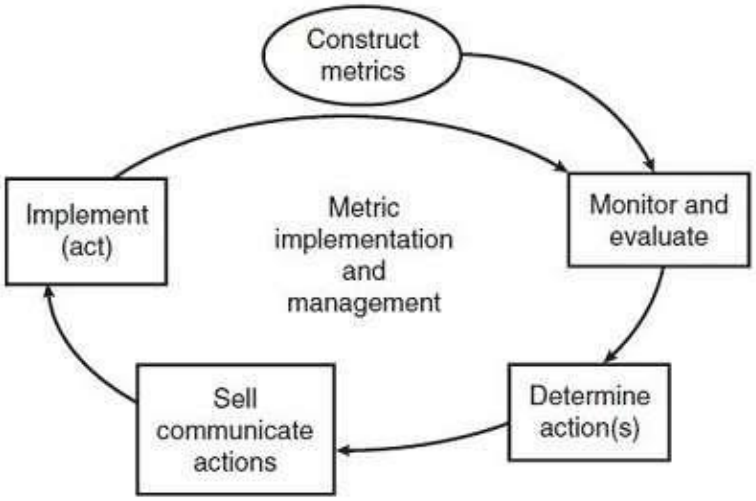
The organization may also lose some reputation as a result of moving defective software to production. Therefore, an UAT is vital. Software testing is a broad field in computing.

9. RESULTS

9.1 Performance Metrics

The following performance measurement necessities are the same whether you're measuring business, service, process, or laboratory variables. Together, they constitute a measurement plan.

- **Definition of purpose:** Why is a measurement being made? What process or variable is being measured? For what will the resulting data be used?
- **Statement of the required measurement performance indicators (accuracy, precision, resolution):** These may be determined by organizational policy, adherence to a published standard or an analysis of the requirements based on use, ability to measure, or more.
- **The unit or variable being measured** and a statement as to why measuring that particular variable supports the purpose of the measurement.
- **An operational definition:** A detailed, yet easily understood, description of the measurement process.
- **An analysis plan:** A typical example is a monthly report that makes comparisons to the previous month, year over year, and year to date. The different time frames provide greater context and allow the data to be presented graphically.
    - A control chart is a simple analysis plan template. It provides a graphical context that shows the continuity of changes over time, plus some analysis (control limits) that enables the viewer to differentiate among common causes, special causes, and random variation.



10. ADVANTAGES

1. Easily identifies trends and patterns

Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans. For instance, for an e-commerce website like Amazon, it serves to understand the browsing behaviours and purchase histories of its users to help cater to the right products, deals, and reminders relevant to them. It uses the results to reveal relevant advertisements to them.

2. No human intervention needed (automation)

With ML, you don't need to babysit your project every step of the way. Since it means giving machines the ability to learn, it lets them make predictions and also improve the algorithms on their own. A common example of this is anti-virus softwares; they learn to filter new threats as they are recognized. ML is also good at recognizing spam.

3. Handling multi-dimensional and multi-variety data

Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.

11. DISADVANTAGES

1. Data Acquisition

Machine Learning requires massive data sets to train on, and these should be inclusive/unbiased, and of good quality. There can also be times where they must wait for new data to be generated

2. Time and Resources

ML needs enough time to let the algorithms learn and develop enough to fulfil their purpose with a considerable amount of accuracy and relevancy. It also needs massive resources to function. This can mean additional requirements of computer power for you

3. Interpretation of Results

Another major challenge is the ability to accurately interpret results generated by the algorithms. You must also carefully choose the algorithms for your purpose.

12.CONCLUSION

By performing different ML models, we aim to get a better result or less error with max accuracy. Our purpose was to predict the price of the used cars having 25 predictors and 509577 data entries.

Initially, data cleaning is performed to remove the null values and outliers from the dataset then ML models are implemented to predict the price of cars.

Next, with the help of data visualization features were explored deeply. The relation between the features is examined.

From the below table, it can be concluded that XGBoost is the best model for the prediction for used car prices. XGBoost as a regression model gave the best MSLE and RMSLE values.

Result of Models:

Model	MSLE	RMSLE	Accuracy
Linear regression	0.00243399	0.04933557	59.3051%
Ridge regression:	0.00243399	0.04933553	59.3051%
Lasso regression	0.00243400	0.04933566	59.305%
KNN	0.00144004	0.03794796	76.4681%
Random Forest	0.00077811	0.00077811	87.5979%
Bagging Regressor	0.00143192	0.03784080	76.809%
Adaboost Regressor	0.00084475	0.02906475	86.4084%
XGBoost Regressor	0.00065047	0.02550431	89.6623%

13.FUTURE SCOPE

- ✓ Thus, the Random Forest Regressor Regression model trained by us using IBM Watson Studio with the dataset provided by the mentor gives 95% exact resale value of the car.
- ✓ Need to collect more data and develop the dataset and train more in the model for best results and to consider all the models that are present in the real world.
- ✓ In future this machine learning model may bind with various website which can provide real time data for price prediction. Also we may add large historical data of car price which can help to improve accuracy of the machine learning model.
- ✓ We can build an android app as user interface for interacting with user. For better performance, we plan to judiciously design deep learning network structures, use adaptive learning rates and train on clusters of data rather than the whole dataset.

14.APPENDIX

14.1 Source Code

Car.html

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
  <head>
    <meta charset="utf-8">
    <title>Car resale value </title>
    <link rel="stylesheet" href="../static/css/style.css">
    <link rel="stylesheet" href="https://cdnjs.cloudflare.com/ajax/libs/font-awesome/4.7.0/css/font-awesome.min.css">
  </head>
  <body>
    <section class="header">
      <nav>
        <a href="/"><img src="../static/Images/roh.png" width="100" height="100"></a>

      </nav>
        <div class="text-box">
          <h1>Car resale value Predictor</h1>
          <p>Best system to predict the amount of resale value based on the parameters provided by the user .</p>
          <a href="./predict_page" class="visit-btn ">Check price</a>
        </div>
    </section>

  </body>
</html>
```

Predict.html

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <link rel="stylesheet" href="../static/css/predict.css">
    <title>Car Resale Predicted Value</title>
</head>
<body>
```

```
        <section class="header">
    <nav>
      <a href="/"><img src="../static/Images/roh.png" width="100" height="100"></a>
    </nav>
      <div class="text-box">
        <h1>The Predicted Car Resale Value is </h1>
                    <h1>{{predict}}</h1>
      </div>
  </section>
</body>
</html>
```

Value.html

```
<!DOCTYPE html>
<html lang="en" dir="ltr">
<head>
<link rel="stylesheet" href="../static/css/value.css">
<title>Car resale value</title>

</head>
<body>
        <section class="form">
        <form action="http://localhost:5000/predict" method="GET">
  <table border="0" align="center">
  <tbody>
                <h1>Get the Accurate Resale Value of Your Car</h1>
  <tr>
        <td><label for="year" padding:10px>Registration year : </label></td>
        <td><input id="year" maxlength="50" name="regyear" type="text" />
        <br>
        <br>
        </td>
        </tr>
        <tr>
        <td><label for="month">Registration Month : </label></td>
        <td><input id="month" maxlength="50" name="regmonth" type="text" />
        <br>
        <br>
        </td>
        </tr>

        <tr>
```

```
<td><label for="power">Power of car in PS: </label></td>
<td><input id="power" maxlength="50" name="powerps" type="text" />
<br>
<br>
</td>
</tr>
<tr>
<td><label for="kilometer">Kilometers that car have driven : </label></td>
<td><input id="kilometer" maxlength="50" name="kms" type="text" />
<br>
<br>
</td>
</tr>
<tr>
<td><label for="geartype">Gear type : </label></td>
<td><input type="radio" name="geartype" value="manual"/> Manual
<input type="radio" name="geartype" value="automatic"/> Automatic
<input type="radio" name="geartype" value="not-declared"/> Not declared
<br>
<br>
</td>
</tr>
<tr>
<td><label for="damage">Your car is repaired or damaged : </label></td>
<td><input type="radio" name="damage" value="yes"/> Yes
<input type="radio" name="damage" value="no"/> No
<input type="radio" name="damage" value="not-declared"/> Not declared
<br>
<br>
</td>
</tr>
<tr>
<td><label for="model">Model Type : </label></td>
<td>
<select name="model" id="model">
<option value="" disabled selected hidden>Choose Model Name...</option>
<option value="golf">Golf </option>
<option value="grand">Grand </option>
<option value="fabia">Fabia </option>
<option value="3er">3er </option>
<option value="2_reihe">2 Reihe </option>
<option value="andere">Andere </option>
<option value="c_max">C Max </option>
```

<option value="3_reihe">3 Reihe </option>
<option value="passat">Passat </option>
<option value="navara">Navara </option>
<option value="ka">Ka </option>
<option value="polo">Polo </option>
<option value="twingo">Twingo </option>
<option value="a_klasse">A klasse </option>
<option value="scirocco">Scirocco </option>
<option value="5er">5er </option>
<option value="meriva">Meriva </option>
<option value="arosa">Arosa </option>
<option value="c4">C4 </option>
<option value="civic">Civic </option>
<option value="transporter">Transporter </option>
<option value="punto">Punto </option>
<option value="e_klasse">E Klasse </option>
<option value="clio">Clio </option>
<option value="kadett">Kadett </option>
<option value="kangoo">Kangoo </option>
<option value="corsa">Corsa </option>
<option value="one">One </option>
<option value="fortwo">Fortwo </option>
<option value="1er">1er </option>
<option value="b_klasse">B Klasse </option>
<option value="signum">Signum </option>
<option value="astra">Astra </option>
<option value="a8">A8 </option>
<option value="jetta">Jetta </option>
<option value="fiesta">Fiesta </option>
<option value="c_klasse">C Klasse </option>
<option value="micra">Micra </option>
<option value="vito">Vito </option>
<option value="sprinter">Sprinter </option>
<option value="156">156 </option>
<option value="escort">Escort </option>
<option value="forester">Forester </option>
<option value="xc_reihe">Xc Reihe </option>
<option value="scenic">Scenic </option>
<option value="a4">A4 </option>
<option value="a1">A1 </option>
<option value="insignia">Insignia </option>
<option value="combo">Combo </option>
<option value="focus">Focus </option>

<option value="tt">Tt </option>
<option value="a6">A6 </option>
<option value="jazz">Jazz </option>
<option value="omega">Omega </option>
<option value="slk">Slk </option>
<option value="7er">7er </option>
<option value="80">80 </option>
<option value="147">147 </option>
<option value="glk">Glk </option>
<option value="100">100 </option>
<option value="z_reihe">Z Reihe </option>
<option value="sportage">Sportage </option>
<option value="sorento">Sorento </option>
<option value="v40">V40 </option>
<option value="5er">5er </option>
<option value="ibiza">Ibiza </option>
<option value="3er">3er </option>
<option value="mustang">Mustang </option>
<option value="eos">Eos </option>
<option value="touran">Touran </option>
<option value="getz">Getz </option>
<option value="a3">A3 </option>
<option value="almera">Almera </option>
<option value="megane">Megane </option>
<option value="7er">7er </option>
<option value="1er">1er </option>
<option value="lupo">Lupo </option>
<option value="r19">R19 </option>
<option value="zafira">Zafira </option>
<option value="caddy">Caddy </option>
<option value="2_reihe">2 Reihe </option>
<option value="mondeo">Mondeo </option>
<option value="cordoba">Cordoba </option>
<option value="colt">Colt </option>
<option value="impreza">Impreza </option>
<option value="vectra">Vectra </option>
<option value="berlingo">Berlingo </option>
<option value="80">80 </option>
<option value="m_klasse">M Klasse </option>
<option value="tiguan">Tiguan </option>
<option value="i_reihe">I Reihe </option>
<option value="espace">Espace </option>
<option value="sharan">Sharan </option>

<option value="6_reihe">6 Reihe </option>
<option value="panda">Panda </option>
<option value="up">Up </option>
<option value="seicento">Seicento </option>
<option value="ceed">Ceed </option>
<option value="5_reihe">5 Reihe </option>
<option value="yeti">Yeti </option>
<option value="octavia">Octavia </option>
<option value="mii">Mii </option>
<option value="rx_reihe">Rx Reihe </option>
<option value="6er">6er </option>
<option value="modus">Modus </option>
<option value="fox">Fox </option>
<option value="matiz">Matiz </option>
<option value="beetle">Beetle </option>
<option value="c1">C1 </option>
<option value="rio">Rio </option>
<option value="touareg">Touareg </option>
<option value="logan">Logan </option>
<option value="spider">Spider </option>
<option value="cuore">Cuore </option>
<option value="s_max">S Max </option>
<option value="a2">A2 </option>
<option value="x_reihe">X Reihe </option>
<option value="a5">A5 </option>
<option value="galaxy">Galaxy </option>
<option value="c3">C3 </option>
<option value="viano">Viano </option>
<option value="s_klasse">S Klasse </option>
<option value="1_reihe">1 Reihe </option>
<option value="avensis">Avensis </option>
<option value="sl">Sl </option>
<option value="roomster">Roomster </option>
<option value="q5">Q5 </option>
<option value="kaefer">Kaefer </option>
<option value="santa">Santa </option>
<option value="cooper">Cooper </option>
<option value="leon">Leon </option>
<option value="4_reihe">4 Reihe </option>
<option value="500">500 </option>
<option value="laguna">Laguna </option>
<option value="ptcruiser">Ptcruiser </option>
<option value="clk">Clk </option>

<option value="primera">Primera </option>
<option value="exeo">Exeo </option>
<option value="159">159 </option>
<option value="transit">Transit </option>
<option value="juke">Juke </option>
<option value="qashqai">Qashqai </option>
<option value="carisma">Carisma </option>
<option value="accord">Accord </option>
<option value="corolla">Corolla </option>
<option value="lanos">Lanos </option>
<option value="phaeton">Phaeton </option>
<option value="boxster">Boxster </option>
<option value="verso">Verso </option>
<option value="swift">Swift </option>
<option value="rav">Rav </option>
<option value="kuga">Kuga </option>
<option value="picanto">Picanto </option>
<option value="kalos">Kalos </option>
<option value="superb">Superb </option>
<option value="stilo">Stilo </option>
<option value="alhambra">Alhambra </option>
<option value="911">911 </option>
<option value="mx_reihe">Mx Reihe </option>
<option value="m_reihe">M Reihe </option>
<option value="roadster">Roadster </option>
<option value="ypsilon">Ypsilon </option>
<option value="cayenne">Cayenne </option>
<option value="galant">Galant </option>
<option value="justy">Justy </option>
<option value="90">90 </option>
<option value="sirion">Sirion </option>
<option value="crossfire">Crossfire </option>
<option value="6_reihe">6 Reihe </option>
<option value="agila">Agila </option>
<option value="duster">Duster </option>
<option value="cr_reihe">Cr Reihe </option>
<option value="v50">V50 </option>
<option value="discovery">Discovery </option>
<option value="c_reihe">C Reihe </option>
<option value="v_klasse">V Klasse </option>
<option value="yaris">Yaris </option>
<option value="c5">C5 </option>
<option value="aygo">Aygo </option>

<option value="cc">Cc </option>
<option value="carnival">Carnival </option>
<option value="fusion">Fusion </option>
<option value="bora">Bora </option>
<option value="forfour">Forfour </option>
<option value="100">100 </option>
<option value="cl">Cl </option>
<option value="tigra">Tigra </option>
<option value="156">156 </option>
<option value="300c">300c </option>
<option value="100">100 </option>
<option value="147">147 </option>
<option value="q3">Q3 </option>
<option value="spark">Spark </option>
<option value="v70">V70 </option>
<option value="x_type">X Type </option>
<option value="5_reihe">5 Reihe </option>
<option value="ducato">Ducato </option>
<option value="s_type">S Type </option>
<option value="x_trail">X Trail </option>
<option value="toledo">Toledo </option>
<option value="altea">Altea </option>
<option value="7er">7er </option>
<option value="voyager">Voyager </option>
<option value="calibra">Calibra </option>
<option value="bravo">Bravo </option>
<option value="range_rover">Range Rover </option>
<option value="antara">Antara </option>
<option value="tucson">Tucson </option>
<option value="q7">Q7 </option>
<option value="citigo">Citigo </option>
<option value="jimny">Jimny </option>
<option value="cx_reihe">Cx Reihe </option>
<option value="wrangler">Wrangler </option>
<option value="lybra">Lybra </option>
<option value="range_rover_sport">Range Rover Sport </option>
<option value="lancer">Lancer </option>
<option value="159">159 </option>
<option value="freelander">Freelander </option>
<option value="captiva">Captiva </option>
<option value="c2">C2 </option>
<option value="500">500 </option>
<option value="range_rover_evoque">Range Rover Evoque </option>

<option value="sandero">Sandero </option>
<option value="note">Note </option>
<option value="900">900 </option>
<option value="147">147 </option>
<option value="defender">Defender </option>
<option value="cherokee">Cherokee </option>
<option value="clubman">Clubman </option>
<option value="samara">Samara </option>
<option value="2_reihe">2 Reihe </option>
<option value="1er">1er </option>
<option value="3er">3er </option>
<option value="601">601 </option>
<option value="3_reihe">3 Reihe </option>
<option value="4_reihe">4 Reihe </option>
<option value="5er">5er </option>
<option value="6_reihe">6 Reihe </option>
<option value="legacy">Legacy </option>
<option value="pajero">Pajero </option>
<option value="auris">Auris </option>
<option value="niva">Niva </option>
<option value="5_reihe">5 Reihe </option>
<option value="s60">S60 </option>
<option value="nubira">Nubira </option>
<option value="vivaro">Vivaro </option>
<option value="g_klasse">G Klasse </option>
<option value="lodgy">Lodgy </option>
<option value="850">850 </option>
<option value="serie_2">Serie 2 </option>
<option value="6er">6er </option>
<option value="charade">Charade </option>
<option value="croma">Croma </option>
<option value="outlander">Outlander </option>
<option value="gl">Gl </option>
<option value="doblo">Doblo </option>
<option value="musa">Musa </option>
<option value="amarok">Amarok </option>
<option value="156">156 </option>
<option value="move">Move </option>
<option value="9000">9000 </option>
<option value="v60">V60 </option>
<option value="145">145 </option>
<option value="aveo">Aveo </option>
<option value="200">200 </option>

```
<option value="300c">300c </option>
<option value="b_max">B Max </option>
<option value="delta">Delta </option>
<option value="terios">Terios </option>
<option value="rangerover">RangeRover </option>
<option value="90">90 </option>
<option value="materia">Materia </option>
<option value="kalina">Kalina </option>
<option value="elefantino">Elefantino </option>
<option value="i3">I3 </option>
<option value="kappa">Kappa </option>
<option value="serie_3">Serie 3 </option>
<option value="48429">48429 </option>
<option value="serie_1">Serie 1 </option>
<option value="discovery_sport">Discovery Sport </option>

</select>
<br>
<br>
</td>
</tr>
<tr>
<td><label for="brand">Brand :</label></td>
<td>
<select name="brand" id="brand">
<option value="" disabled selected hidden>Choose Brand Name...</option>
<option value="volkswagen">Volkswagen </option>
<option value="audi">Audi </option>
<option value="jeep">Jeep </option>
<option value="skoda">Skoda </option>
<option value="bmw">Bmw </option>
<option value="peugeot">Peugeot </option>
<option value="ford">Ford </option>
<option value="mazda">Mazda </option>
<option value="nissan">Nissan </option>
<option value="renault">Renault </option>
<option value="mercedes_benz">Mercedes Benz </option>
<option value="opel">Opel </option>
<option value="seat">Seat </option>
<option value="citroen">Citroen </option>
<option value="honda">Honda </option>
<option value="fiat">Fiat </option>
<option value="mini">Mini </option>
```

```
<option value="smart">Smart </option>
<option value="hyundai">Hyundai </option>
<option value="sonstige_autos">Sonstige Autos </option>
<option value="alfa_romeo">Alfa Romeo </option>
<option value="subaru">Subaru </option>
<option value="volvo">Volvo </option>
<option value="mitsubishi">Mitsubishi </option>
<option value="kia">Kia </option>
<option value="suzuki">Suzuki </option>
<option value="lancia">Lancia </option>
<option value="porsche">Porsche </option>
<option value="toyota">Toyota </option>
<option value="chevrolet">Chevrolet </option>
<option value="dacia">Dacia </option>
<option value="daihatsu">Daihatsu </option>
<option value="trabant">Trabant </option>
<option value="saab">Saab </option>
<option value="chrysler">Chrysler </option>
<option value="jaguar">Jaguar </option>
<option value="daewoo">Daewoo </option>
<option value="rover">Rover </option>
<option value="land_rover">Land Rover </option>
<option value="lada">Lada </option>
</select>
<br>
<br>
</td>
</tr>
<tr>
<td><label for="fuelType">Fuel Type :</label></td>
<td>
<select name="fuelType" id="brand">
<option value="" disabled selected hidden>Choose Fuel Type...</option>
<option value="petrol"> Petrol </option>
<option value="diesel"> Diesel </option>
<option value="not-declared"> Not Declared </option>
<option value="lpg">LPG </option>
<option value="cng">CNG </option>
<option value="hybrid">Hybrid </option>
<option value="others">Others </option>
<option value="electric">Electric </option>
</select>
<br>
```

```
        <br>
        </td>
        </tr>
        <tr>
        <td><label for="vehicletype">Vehicle type:</label></td>
        <td>
        <select name="vehicletype" id="vehicle" >
        <option value="" disabled selected hidden>Choose Vehicle Type...</option>
        <option value="coupe">Coupe </option>
  <option value="suv">SUV </option>
  <option value="kleinwagen">Kleinwagen </option>
  <option value="limousine">Limousine </option>
        <option value="cabrio">Cabrio </option>
        <option value="bus">Bus </option>
        <option value="kombi">Kombi </option>
        <option value="andere">Andere </option>
        <option value="volkswagen">Volkswagen </option>
        </select>
        <br>
        <br>
        </td>
        </tr>
        </tbody>
        </table>
        <input name="Submit" type="Submit" value="Submit" id="button"/>
        </form>
 </section>

</body
</html>
```

15. GITHUB LINK

https://github.com/IBM-EPBL/IBM-Project-48089-1660804377