JAYARAM COLLEGE OF ENGINEERING AND TECHNOLOGY

CAR RESALE VALUE PREDICTION

NALAIYA THIRAN PROJECT REPORT

2022

Team ID: IBM-Project-55141-1666074362

College Name: JAYARAM COLLEGE OF ENGINEERING AND TECHNOLOGY

Submitted by

NIRMAL KUMAR V	**(811419104006)**
ARCHANA T	**(811419104003)**
PAVITHRA S	**(811419104007)**
ROSHMABAGA M M	**(811419104009)**

JASEEMA BEGUM J **(811419205002)**

SOWNDHARYA S **(811419205006)**

Project Report Format

CAR RESALE VALUE PREDICTION

1.INTRODUCTION

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. We implement and evaluate various learning methods on a dataset consisting of the sale prices of different makes and model. We will compare the performance of various machine learning algorithms like Linear Regression, Ridge Regression, Lasso Regression, Elastic Net, Decision Tree Regressor and choose the best out of it. Depending on various parameters we will determine the price of the car. Regression Algorithms are used because they provide us with continuous value as an output and not a categorized value because of which it will be possible to predict the actual price a car rather than the price range of a car. User Interface has also been developed which acquires input from any user and displays the Price of a car according to user's inputs.

1.1 PROJECT OVERVIEW:

system is defined in the python language that predicts the amount of resale value based on the given information. The system works on the trained dataset of the machine learning program that evaluates the precise value of the car. User can enter details only of fields like purchase price of car, kilometers driven, fuel of car, year of purchase.Upon form submission, the data is sent to the ML model via Flask API and the model responds with a predicted resale value of the car based on user input. This prediction is displayed on the web page using a render template. Thus, with minimal information and without human intervention or manual examination, a user can predict the resale value of his car.

1.2 PURPOSE

This resale value prediction system is made for general purpose to just predict the amount that can be roughly acquired by the user. We try to predict the amount of resale by best 70% accuracy so the user can get estimated value before he resales the car and doesn't make a deal in lossThe 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 predicteCar resale value prediction system is made with the purpose of predicting the correct valuation of used cars that helps users to sell the car remotely with perfect valuation and without human intervention in the process to eliminate biasedvaluation.

2.LITERATURE SURVEY:

1. **Car's Selling Price Prediction using Random Forest Machine LearningAlgorithm**

Abhishek Pandey, Vanshika Rastogi , Sanika Singh

India has one of the biggest automobile markets all over the globe every day many buyers usually sell their cars after using for the time to another buyer, we call them as 2nd /3rd owner etc. Many platforms such as cars24.com, cardekho.com and OLX.com provides these buyers witha platform where they can sell their used cars, but what should be the price of the car, this is the toughest question ever. Machine Learning algorithms can bring a solution to this problem. Using a history of previously used cars selling data and using machine learning techniques such as Supervised Learning can predict a fair price of the car, here I also used machine learning algorithms such as Random Forest and Extra Tree Regression along with powerful python libraryScikit-Learn to predict the selling price of the used car. The result has shown that these both algorithms are highly accurate

in prediction even the dataset is large or small, irrespective of the size of the dataset they give a precise result.

2. Prediction of Resale Value of the Car Using Linear Regression Algorithm

Kiran S,Computer Science Engineering,SJB Institute of Technology, Bangalore, India In this research the price of the car is considered as dependent variable for target prediction .The data used for prediction was taken from web. The suitability of linear regression algorithm is identified and implemented in this research work for accurately predicting the resale value of the vehicle based on most significant attributes that are beenselected on the basis of highest correlation.The Linear Regression model for prediction ofresale value of the car is providing an accuracy of 90% and an error of 10%. Linear Regression model is better suited for prediction of target attribute that is msrp (car price). Further this work can be implemented using different machine learning algorithms and approaches in order to get higher accuracy rate and lower error percentage.

3. PREDICTIVE ANALYSIS OF USED CAR PRICES USING MACHINE LEARNING

Ashutosh Datt Sharma, Vibhor Sharma, Sahil Mittal, Gautam Jain, Sudha Narang

In this swiftly-moving world, managing our professional as well as personal lives have become quite hectic and if we don't have our own personal vehicle for transportation, life is a lot more hectic. To be on the safe side, one should have a more reliable and easy mode for transportation and a personal vehicle is always the best option. Having a car is very important for people these days as it gives a certain social status and also gives a certain extent of personal control to individual owning it. In some areas with low population, having a car becomes essential as it provides the only option for covering long distances in case of an absence of public transport. Old aged people, who have difficulties in walking or cycling to places, have driving the sole option for moving without being dependent. And for those that don't have enough resources to purchase a brand-new car, buying an old vehicle becomes a necessity and that too at a reasonable

price. The car manufacturing has been increasing swiftly over the years during past decade, with about 92 million cars that were manufactured in 2019. This provides a big boost for the market of old and used cars which is now coming up as a progressively growing industry. The recent entries of various websites and web-portals have fulfilled the requirements of customers up to some extent as they now know the present trends and scenario to get the market value of any old vehicle present in the market. Machine Learning has a lot of applications in real world scenario but one of the most known application is the use of Machine Learning in resolving the prediction problems. The project being discussed here is very much based upon one among such applications.

Employing various Machine Learning Algorithms, we will try and build a statistical model based upon given data and features set to estimate the prices of used cars.

4. **Used Car Price Prediction using Machine Learning: A Case Study**
   Mustapha Hankar; Marouane Birjali; Abderrahim Beni-Hssane
   Published in 2022 11th International Symposium on Signal, Image, Video and Communications (ISIVC)

   Several regression techniques were used based on supervised machine learning to predict the resale price of used cars given many factors such as mileage, fuel type, fiscal power , mark , model, and the production year of the car. In all tested models, gradient boosting regressor showed a high R-squared score and low root mean square error. The results showed that gradient boosting regressor outperformed all tested models with a highest R2 score and a minimized root mean squared error. As a future work, it is intended to increase the performance of the model by scaling the training data and adding more other variables to the feature set.

5. **Car Price Prediction using Machine Learning Techniques**

Enis Gegic, Becir Isakovic, Dino Keco, Zerina Masetic, Jasmin Kevric

A car price prediction has been a high interest research area, as it requires noticeable effort and knowledge of the field expert. Considerable number of distinct attributes are examined for the reliable and accurate prediction. To build a model for predicting the price of used cars in Bosnia and Herzegovina, we applied three machine learning techniques (Artificial Neural Network, Support Vector Machine and Random Forest).The data used for the prediction was collected from the web portal autopijaca.ba using web scraper that was written in PHP programming language. Respective

performances of different algorithms were then compared to find one that best suits the available data set. The final prediction model was integrated into Java application. Furthermore, the model was evaluated using test data and the accuracy of 87.38% was obtained.

6. **An Automated Car Price Prediction System Using Effective Machine LearningTechniques**

Santosh Kumar Satapathy,Rutvikraj Vala,Shiv Virpariya ICT, Pandit Deendayal EnergyUniversity, Gandhinagar, INDIA

This research focuses on Building a mathematical model that could predict the price of a second-hand car based on its current features. Determining the price of a used automobile is a difficult task because several factors like Current Mileage, Current Condition, Make, Year, etc.,can influence the prediction prices of an automobile. And, from the perspective of a person who sells, it becomes a dilemma to predict the price of a second-hand car accurately. Thus, the point of interest of this challenge is in growing gadgets, studying models that can correctly expect the price of a used car primarily based on its capabilities. Due to this, in turn, a consumer can make a much more informed purchase. Therefore, implementing and examining various Machine Learning Techniqueswith Data Analysis will be useful to Provide an Accurate and Easy to use solution.

7. **Used Cars Price Prediction and Valuation using Data Mining Techniques**

AlShared, Abdulla, "Used Cars Price Prediction and Valuation using Data Mining Techniques" (2021). Thesis. Rochester Institute of Technology. Accessed from

A primary objective of this project is to estimate used car prices by using attributes that are highly correlated with a label (Price). To accomplish this, data mining technology has been employed. Null, redundant, and missing values were removed from the dataset during pre-processing. In this supervised learning study, three regressors (Random Forest Regressor, Linear Regression, and Bagging Regressor) have been trained, tested, and compared against a benchmark dataset. The researchers of this project anticipate that in the near future, the most sophisticated algorithm is used for making predictions, and then the model will be integrated into a mobile app or web page for the general public to use.

8. **Used Cars Price Prediction using Supervised Learning Techniques**

P Venkatasubbu, M Ganesh - Int. J. Eng. Adv. Technol.(IJEAT), 2019 - researchgate.net The recent advent of online portals has facilitated the need for both the customer and

the seller to be better informed about the trends and patterns that determine the value of a used car in the market. Using Machine Learning Algorithms such as Lasso Regression, Multiple Regression and Regression trees, they try to develop a statistical model which will be able to predict the price of a used car, based on previous consumer data and a given set of features and also comparing the prediction accuracy of these models to determine the optimal one. To get even more accurate models, we can also choose more advanced machine learning algorithms

such as random forests, an ensemble learning algorithm which creates multiple decision/regression trees, which brings down overfitting massively or Boosting, which tries to bias the overall model by weighing in the favor of good performers.

9. Predicting the Price of Pre-Owned Cars Using Machine Learning and Data Science

G. Kalpana1, Dr. A. Kanaka Durga2, T. Anoop Reddy3, Dr. G.Karuna4 Department of Computer Science And Engineering, Vidya Jyothi Institute of Technology, Hyderabad, India

This model reduces time and cost and is also more user friendly as a result of which there is improvement in business by selling more cars. Here we are also conducting a comparative study on performance of regression based on supervised machine learning models. Each model is trained using data of used car market collected from e-commerce website. As a result, Linear regression gives the best performance with Root mean square error (RMSE) =8902.410 . Followed by ridge, random forest regression algorithms respectively. We can also extend this project by considering more attributes like Resale history, Lic , Accidents history, image etc to the data set for getting clear and accurate analysis.

10. Value Based Pricing meets Data Science: A Concept for Automated Spare Part Valuation

Clemens Wickboldt1, Natalia Kliewer2 Freie Universität Berlin, Professur für Wirtschaftsinformatik, Berlin, Germany,

An early-stage concept for automated spare part valuation which classifies pricing data before applying appropriate valuation methods is presented and hereby combines

methods from multiple disciplines. Information from heterogeneous sources is aggregated, transformed and then supports machine learning methods to automatically determine a Fair Market Value for surplus spare parts. The concept for automated spare part valuation is a promising alternative for value determination and pricing in secondary markets and thus may serve as a foundation for building a generic surplus part trading platform to overcom market transparency issues if the obstacles of validation are overcome. Handling incomplete historical data sets as well as validating the calculated Fair Market Value are some of the challenges which become visible.

2.1 EXISTING PROBLEM:

The forecasts of vehicle cost from the chronicled information that has been gathered from every day papers. They have utilized the administered AI strategies for foreseeing the cost of vehicles. Numerous different calculations like various straight relapse, k-closest neighbor calculations, gullible based, and some choice tree calculations additionally been utilized. Every one of the four calculations are looked at and tracked down the best calculation for forecast. They have confronted a few challenges in looking at the calculations, by one way or another they have overseen. As indicated by creators Pattabiraman, this paper is more focused on the connection among vender and purchaser. To foresee the cost of four wheelers, more highlights are required like previously given value, mileage, make, model, trim, type, chamber, liter, entryways, voyage, sound, cowhide. Utilizing these highlights the cost of vehicle has been anticipated with the assistance of factual investigation framework for exploratory information examination. As per creators EnisGegic et al, in this paper the chiefly focus on gathering different information from web entryway by utilizing web scrap methods. Furthermore, those have been contrasted and the assistance of various AI calculation.

2.2 REFERENCES:

[1] Kanwal Noor, 2017, Vehicle Price Prediction System using Machine Learning Techniques International Journal of Computer Applications. Volume 167 - Number 9

[2] Mariana Lusitania et al, (2009). Support vector regression analysis for price prediction in a vehicle leasingapplication

[3] Richardson, M. S. (2009). Determinants of used vehicle resale value.

[4] Listiani, M. (2009). Support vector regression analysis for price prediction in a car leasing application (Doctoral dissertation, Master thesis, TU Hamburg-Harburg).

[5] Richardson, M. S. (2009). Determinants of used car resale value. Retrieved from:https:/ digitalcc.coloradocollege.edu/islandora/object

[6] Pudaruth, S., 2014. "Predicting the Price of Used Cars using Machine Learning Techniques." Vol 4, Number 7 (2014), pp. 753-76.

[7] Gokce, E. (2020, January 10). "Predicting used car prices with machine learning techniques. "

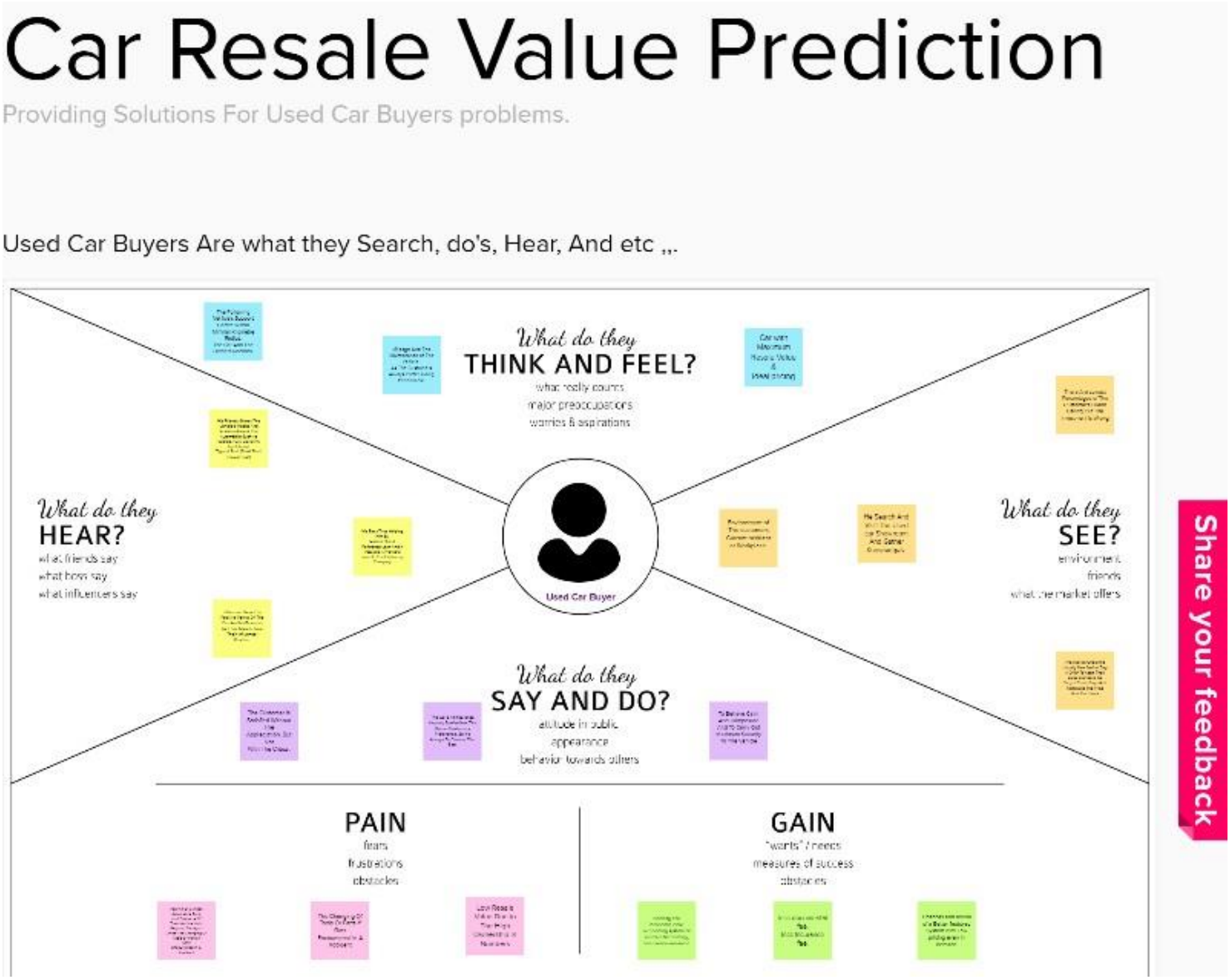
2.3.PROBLEM STATEMENT DEFINITION:

With difficult economic conditions, it is likely that sales of second-hand imported (reconditioned) cars and used cars will increase. In many developed countries, it is common to lease a car rather than buying it outright. After the lease period is over, the buyer has the possibility to buy the car at its residual value, i.e. its expected resale value. Thus, it is of commercial interest to sellers/financers to be able to predict the salvage value (residual value) of cars with accuracy. In order to predict the resale value of the car, we proposed an intelligent, flexible, and effective system that is based on using regression algorithms. Considering the main factors which would affect the resale value of a vehicle a regression model is to be built that would give the nearest resale value of the vehicle. We will be using various regression algorithms and algorithm with the best accuracy will be taken as a solution, then it will be integrated to the web-based application where the user is notified with the status of his product

3.IDEATION & PROPOSED SOLUTION:

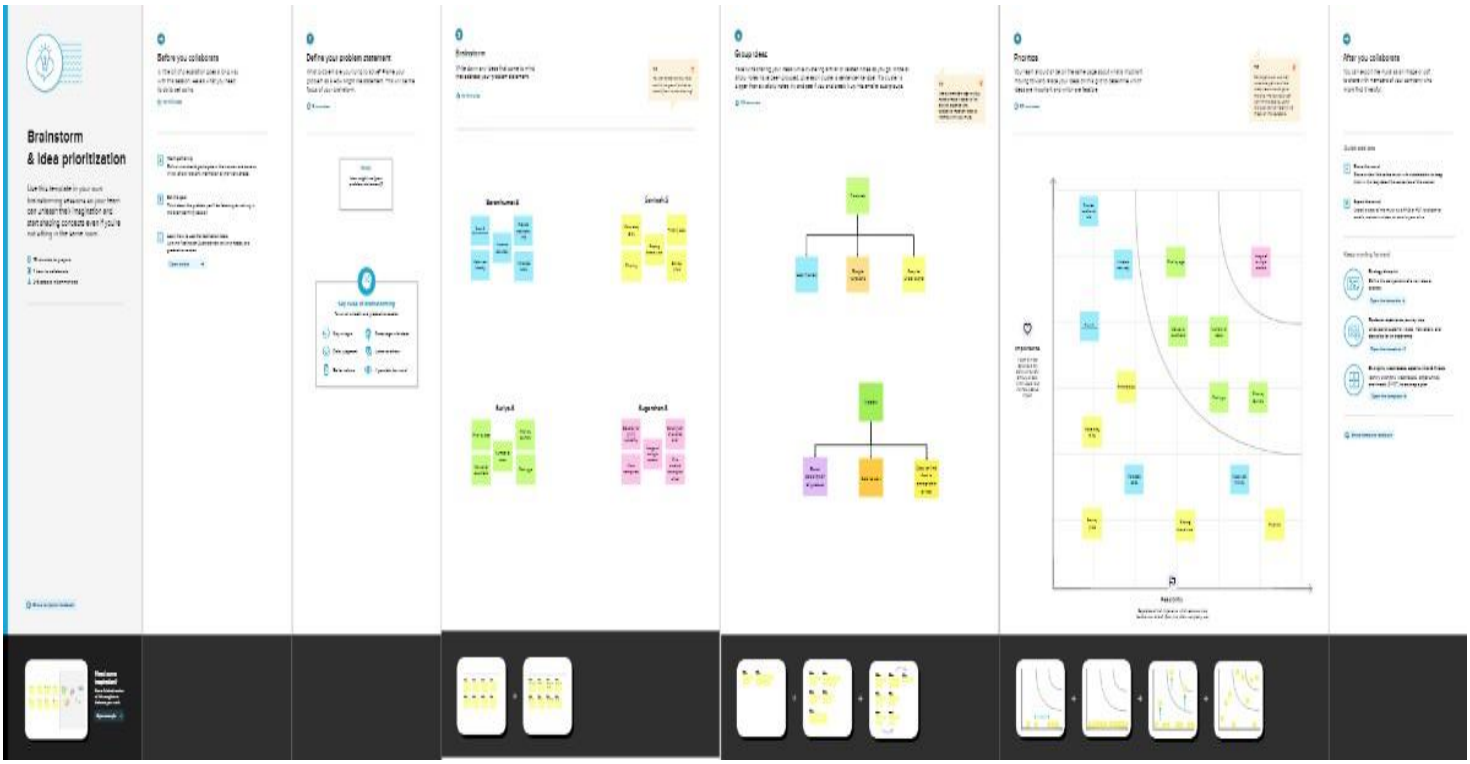
3.1 EMPATHY MAP CANVAS:

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudesIt is a useful tool to helps teams better understand their users. Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.



3.2 IDEATION & BRAINSTORMING:

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



3.3 PROPOSED SOLUTION:

Proposed Solution:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	➢ Sales prediction is the current numerous trend in which all the business companies thrive and it also aids the organization or concern in determining the future goals for it and its plan and procedure to achieve it. ➢ Resales of cars almost occupy a major part in every sales economy. ➢ In that regard various factors like registration year, engine condition, company service record, spare parts condition, tyre condition, car body condition, kilometers covered, Interior look, color, mileage, number of owners, battery condition are taken into consideration before buying it along with engine condition and insurance. ➢ The predication using the factorswould suggest the final productto be brought. ➢ But these data may be inaccurate at times and there is a need of a proper algorithm that will provide a result with good accuracy rate.
2.	Idea / Solution description	➢ The idea starts with login page where the user initially directed to the login page with some authenticated credentials. ➢ Then the user is been taken to the dataset uploading page where the data is read with the integrated flask. ➢ Flask is used as middleware for integrating frontend and backend. ➢ This application had the features of extracting the data , cleaning thedata and perform some EDA for accurate predictions. ➢ Then the processed data is split into train and test the dataset is trained based on any of the machine learning algorithm.The model that we areusing is time series forecasting.

		➢ Then the trained dataset is predicted based on the user needs like model, price, design, kilometres covered, Interior look, colour, mileage, number of owners and many features. ➢ Then the predicted result is shown in the interactive dashboard.
3.	Novelty / Uniqueness	➢ Look at similarities in other features such as pricing, innovation, and seasonality for example. ➢ Consumer behavior changes, it's a fact. So for better accuracy select a more recently added product when possible. ➢ You can use multiple reference products to get the best average and the novelty sales estimates will be based on features from all of them using the average.
4.	Social Impact / Customer Satisfaction	➢ Sales forecasting helps you attain this revenue efficiency by offering insight into the likely behavior of your most valuable customers. ➢ You can predict future sales, as well as improve pricing, advertising, and product development.
5.	Scalability of the Solution	➢ Here we are using time series analysis so, When historical data for a product or product line is available and patterns are obvious, organisations typically employ the time series analysis technique to demand forecasting. ➢ A time series analysis can help you detect seasonal variations in demand, cyclical patterns, and major sales trends. ➢ The time series analysis approach works best for well-established organisations with several years of data to work with and very steady trend patterns.

3.4 PROBLEM SOLUTION FIT:

The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem. A problem-solution-fit occurs if a startup has proved both: 1) that there is a 'problem worth solving' for one or more clearly defined customer groups, and 2) that there is evidence that these customer groups would consider the value proposition of the solution the firm proposes.

4 REQUIREMENT ANALYSIS:

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 to the related websites	Registration through Form Registration through Gmail Registration through Application
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Users Profile	Personal details, Bank account ,Is He/She interested in buying a car
FR-4	Gather information about the vehicle	Through the registered websites they collect information
FR-5	Display the functionality of the vehicle	Details: Fuel type , Manufactured year , Miles Driven , Record

Non-functional Requirements:

Following are the non-functional requirements of the proposed solution

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	User friendly UI Simple and easy to Ur
NFR-2	**Security**	Aware of scams
NFR-3	**Reliability**	The system must perf
NFR-4	**Performance**	The landing page mus provide 5 second or le
NFR-5	**Availability**	Uninterrupted service except the time of ser
NFR-6	**Scalability**	That can handle any a many computations i time-saving way to ins residing at global loca

5.PROJECT DESIGN:

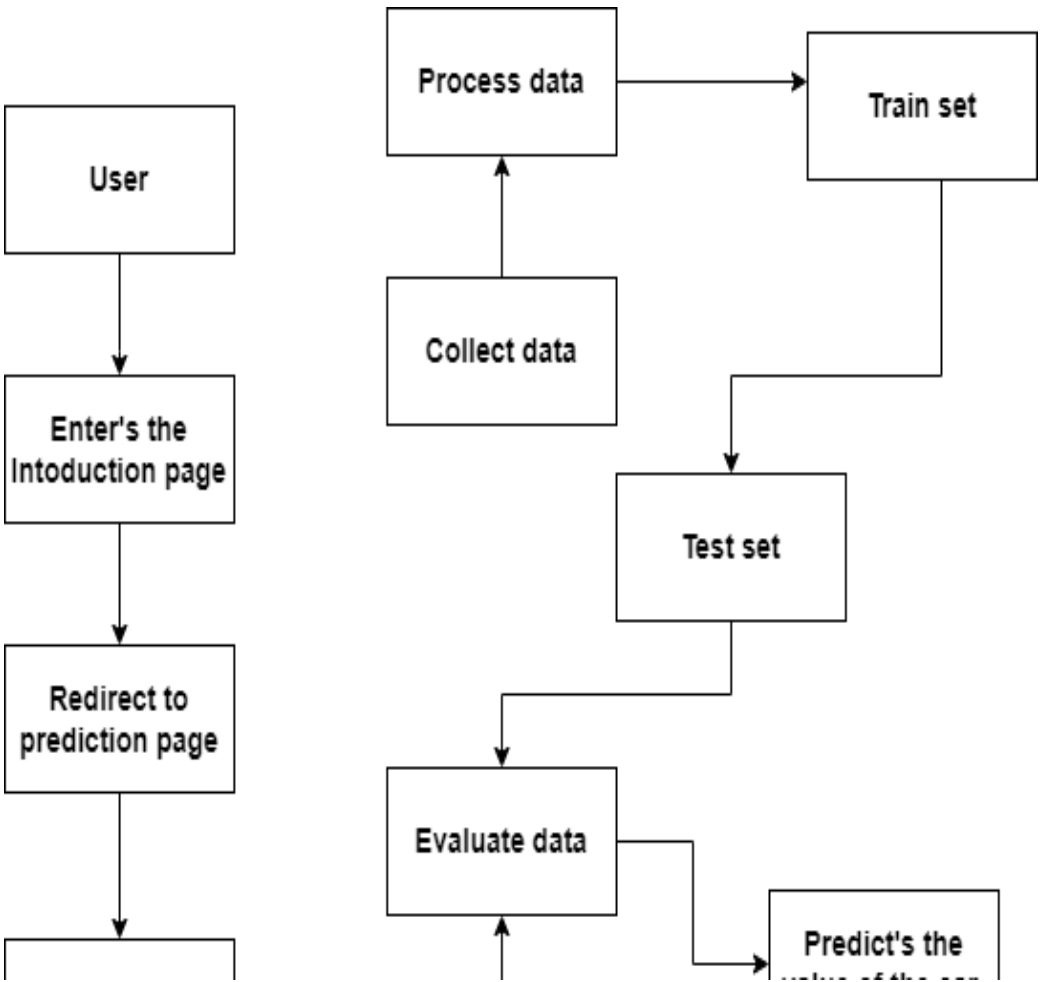
5.1 DATA FLOW DIAGRAMS

A data flow diagram (DFD) m

they work for both technical and nontechnical audiences, fro
That’s why DFDs remain so popular after all these years. Wh
data flow software and systems, they are less applicable nc
interactive, real-time or database-oriented software or system

Data flow diagrams were popularized in the late 1970s,
Structured Design, by computing pioneers Ed Yourdon an
They based it on the “data flow graph” computation models
Gerald Estrin. The structured design concept took off in the
field, and the DFD method took off with it. It became mor
circles, as it was applied to business analysis, than in academ

5.1 Data Flow Diagram



5.2 SOLUTION & TECHNICAL ARCHIECTURE:

Functional Requirements:

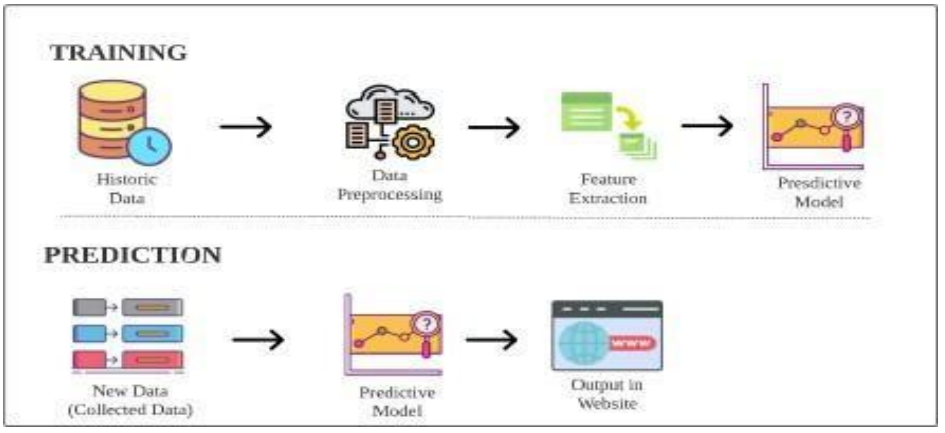
Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIN.
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP.
FR-3	Users Profile	Personal details, Bank account ,If customer interestedto buy a car
FR-4	Gather information about the vehicle	Then get the registered websites by customer collect information.
FR-5	Display the functionality of the vehicle	Through the registered websites they collect information about the vehicles.

Non-functional Requirements:

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	To create an UI makes as a user friendly, it makes a simple way to Understand
NFR-2	**Security**	Aware about fraudulent sites, it gives a fakeinformation about the vehicle.
NFR-3	**Reliability**	Application must perform good and without failure
NFR-4	**Performance**	Website performance measures how quickly the pages of a website load and display in the web browser.
NFR-5	**Availability**	Website availability (also called website uptime) refers to the ability of the users to access and use a website or web service. A website's availability is typically communicated as a percentage for a given span of time.
NFR-6	**Scalability**	Application scalability is the ability of an applicationto handle a growing number of users and load, without compromising on performance and causingdisruptions to user experience. To put it another way, scalability reflects the ability of the software to grow or change with the user's demands..



Technical Architecture:

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2

Example: Car Resale value Prediction

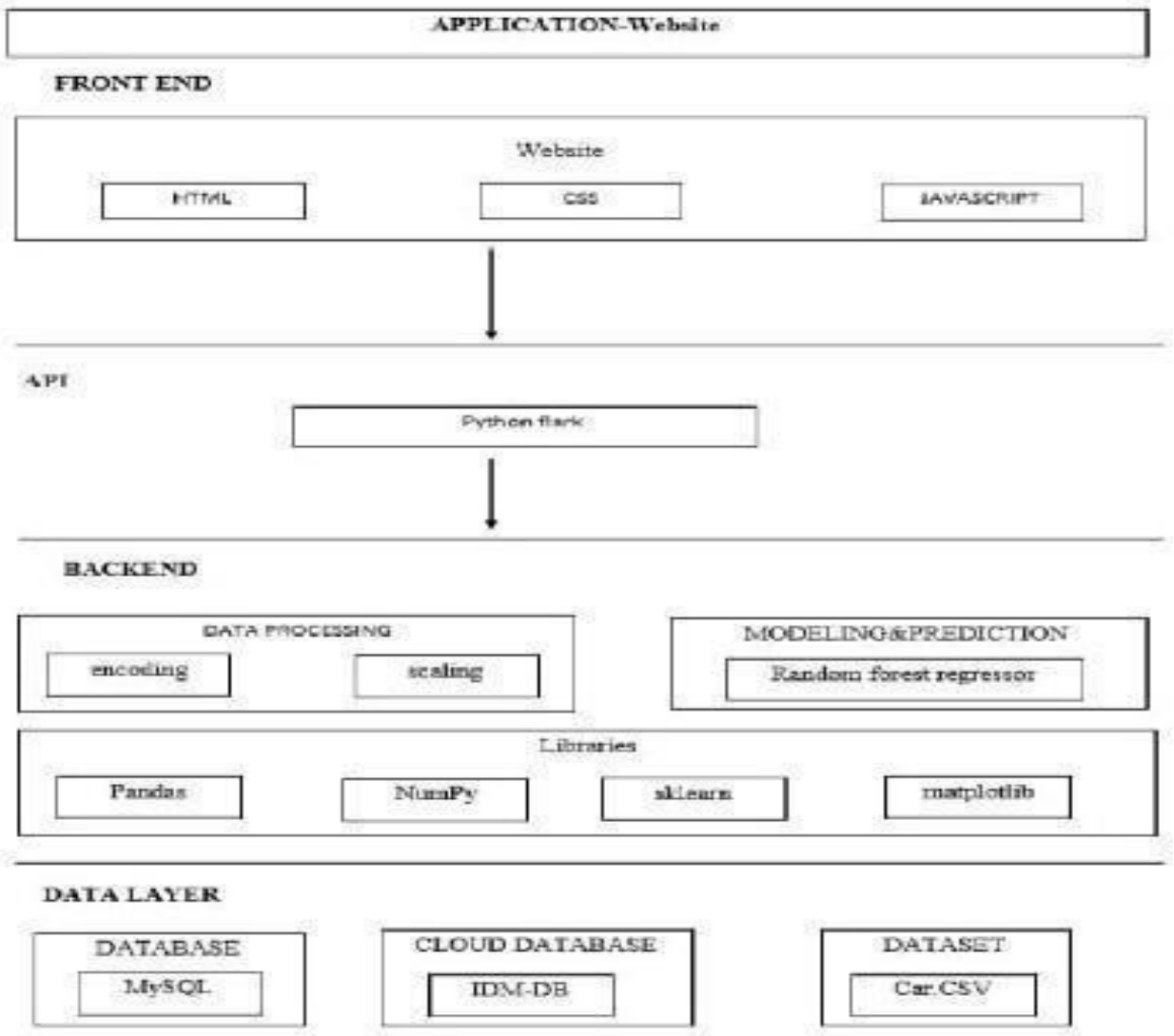


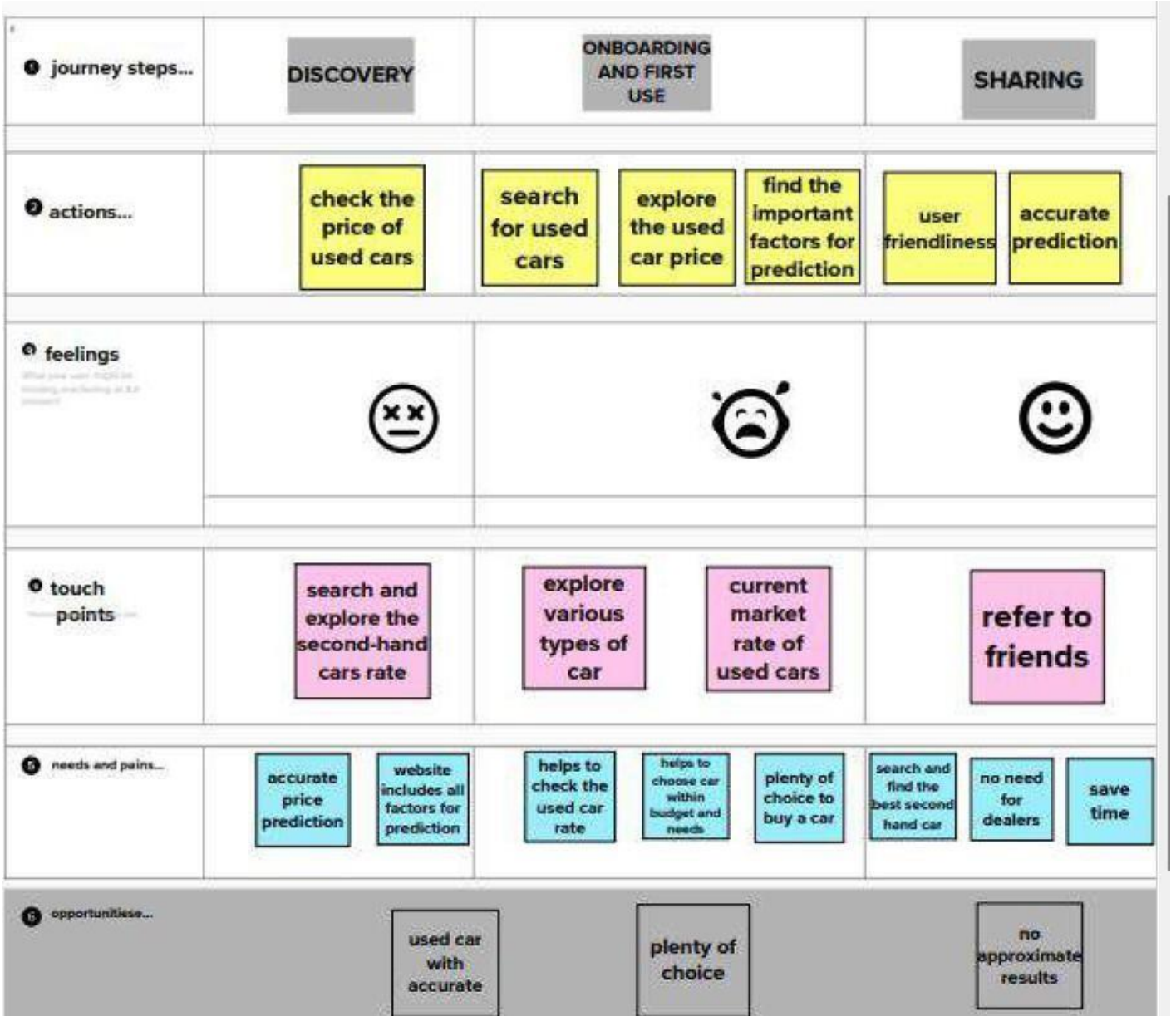
Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js / React Js etc.
2.	Application Logic-1	Logic for a process in the application	Java / Python
3.	Application Logic-2	Logic for a process in the application	IBM Watson STT service
4.	Application Logic-3	Logic for a process in the application	IBM Watson Assistant
5.	Database	Data Type, Configurations etc.	MySQL, NoSQL, etc.
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
8.	External API-1	Purpose of External API used in the application	IBM Weather API, etc.
9.	External API-2	Purpose of External API used in the application	Aadhar API, etc.
10.	Machine Learning Model	Purpose of Regression Model	Regression Model, etc.
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local, Cloud Foundry, Kubernetes, etc.

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	Technology of Opensource framework
2.	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	e.g. SHA-256, Encryptions, IAM Controls, OWASP etc.
3.	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Technology used
4.	Availability	Justify the availability of application (e.g. use of load balancers, distributed servers etc.)	Technology used
5.	Performance	Design consideration for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc.	Technology used

5.3 USER STORY:



6 PROJECT PLANNING AND SCHEDULING

Product Backlog, Sprint Schedule, and Estimation

Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Dataset Reading and Preprocessing	USN-1	Cleaning the dataset end splitting to dependent and independent variables	2	High	Baranikumar.S,
Sprint-2	Building the Model	USN-2	Choosing the appropriate model for building andsavingthe models pickle file	1	High	Santhosh.S
Sprint-3	Application Building	USN-3	Using flask deploying the ML model	2	Medium	Suganthan.S
Sprint-4	Train the Model in ibm	USN-4	Finally train the model on IBM cloud and deploy the application	2	Medium	Suriya.S

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	20	5 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	5 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	5 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	5 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

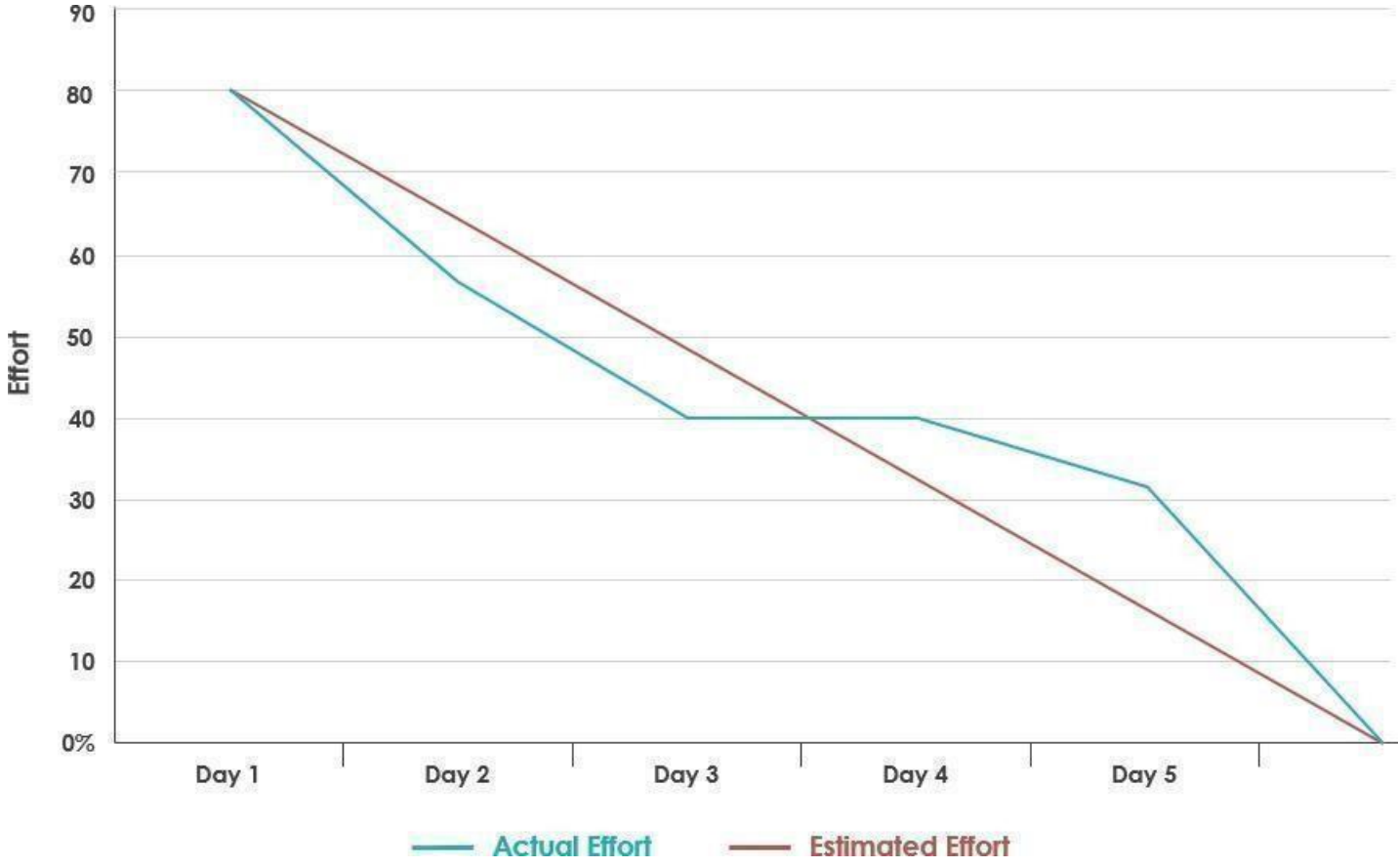
Velocity:

Imagine we have a 5-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) periteration unit (story points per day)

AV =sprint duration\velocity=20/5=4

Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time



7.CODING & SOLUTIONING:

Feature 1:

This is a supervised learning problem and can be solved using regression techniques. We need to predict the selling price of a car based on the given car's features. Supervised Regression problems require labeled data where our target or dependent variable is the selling price of a car. All other features are independent variables.

Linear Models are relatively less complex and explainable, but linear models perform poorly on data containing the outliers. Linear models fail to perform well on non-linear datasets. In such cases, non-linear regression algorithms Random Forest Regressor perform better in fitting the nonlinear data.

This model was hence chosen to account for the large number of features in the dataset and compare a bagging technique with the following gradient boosting methods

Feature 2:

Given the evaluation parameters the Random Forest Regressor outperformed as it has the highest accuracy as well as the lowest error in all three valuation parameters.

As a result of preprocessing and transformation, Random Forest Regressor came out on with 90% accuracy

8 :TESTING:

8.1 TEST CASES:

• Missing values

The trained ML model requires few feature inputs for predicting the output. Failing which, the model throws invalid Input error. All the fields in the html form have been marked required using CSS and thus the user must input all fields.

• Invalid Input

The trained ML model requires only numerical input for all features. Thus, if the user uses symbols such as a comma while inputting, the model may throw an error. To overcome the same, preprocessing script is deployed in the backend which removes all unwanted characters like comma, whitespaces etc. so that model gets required input.

8.2. USER ACCEPTANCE TESTING

Acceptance testing focuses even more on the overall system features and functionality that are visible to the customer. Acceptance testing is often performed by customers to ensure customer usability and satisfaction. The purpose of this is to briefly explain the test coverage and open issues of the [ProductName] project at the time of the release to User Acceptance Testing (UAT).

Defect Analysis

This report shows the number of resolved or closed bugs at each severitylevel, and how they were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	10	4	2	3	19
Duplicate	1	0	1	0	2
External	2	1	0	2	5
Fixed	13	2	2	20	37
Not Reproduced	0	0	1	1	2
Skipped	0	0	1	1	2
Won't Fix	0	2	1	1	4
Totals	26	9	8	28	71

1. Test case analysis

This report shows the number of test cases that have passed, failed, and untested

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	5	0	0	5
Client Application	13	0	0	13
Security	2	0	0	2
Outsource Shipping	4	0	0	4
Exception Reporting	6	0	0	6
Final Report Output	4	0	0	4
Version Control	2	0	0	2

9. RESULTS: Performance Metrics:

Performance metrics are used to track progress. Metrics give some sort of concrete answer which easily can be followed up. There are different types of metrics used for testing.

The regression model can be evaluated on following parameters:

Regression Model:

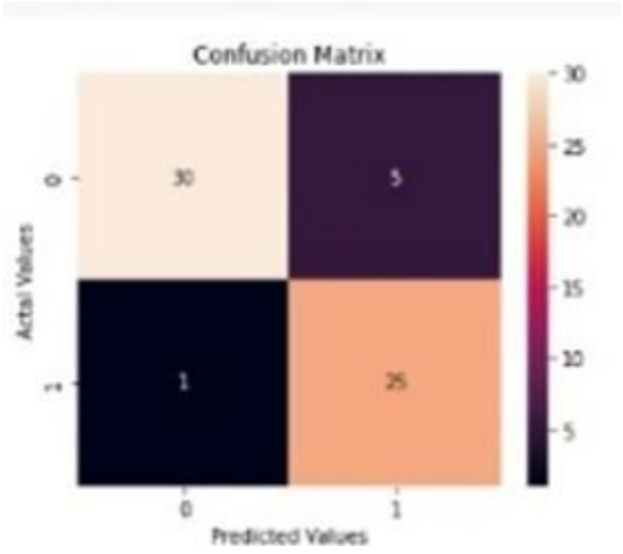
1. Mean Square Error (MSE): MSE is the single value that provides information about goodness of regression line. Smaller the MSE value, better the fit because smaller value implies smaller magnitude of errors.
2. Root Mean Square Error (RMSE): RMSE is the quadratic scoring rule that also measures the average magnitude of the error. It is the square root of average squared difference between prediction and actual observation.
3. Mean Absolute Error (MAE): This measure represents the average absolute difference

between the actual and predicted values in the dataset. It represents the average residual from thedataset.

```
RMSE:
0.31362502409359

MSE:
0.31362502409359

MAE:
0.09836065573770492

R2 SCORCE:
0.5978021978021978
```

Classification Model:

Confusion Matrix, Accuray Score- 0.9016 & Classification Report



Classification report

```
[60] from sklearn.metrics import classification_report
     print(classification_report(original_classes, pred_classes))

                  precision    recall  f1-score   support

              0.0       0.97      0.86      0.91        35
              1.0       0.83      0.96      0.89        26

         accuracy                           0.90        61
        macro avg       0.90      0.91      0.90        61
     weighted avg       0.91      0.90      0.90        61
```

10. ADVANTAGES & DISADVANTAGES:

Advantages:

- Accuracy of our model is 90%.
- Prediction runs for different types of cars.
- Disadvantages:
- Accuracy can be improved.
- Prediction is done using only a few criteria

11. CONCLUSION:

We started with understanding the use case of machine learning in the Automotive industry and how machine learning has transformed the driving experience. We build a Random Forest Regression model to predict the resale value of a used car. Finally, we evaluated the performance of the model using the R squared score and Residual Plot. We could have also used simpler regression algorithms like Linear Regression and Lasso Regression. Still, we need to make sure there are no outliers in the dataset before implementing them. Pair plots and scatter plots help visualize the outliers.

Then we have used a Flask application to display the predicted value to the users based on their corresponding input. This car resale value prediction can be used by the public to estimate theresale value of the car.

12. FUTURE SCOPE:

Currently, only few features are used to predict resale value of the car. This can be extended to more features. One can also implement CNN to determine physical condition of the car from images like identifying dents, scratches etc. and thus predicting more relevant resale value of a car. 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. To correct for overfitting in ce

13. APPENDIX

13.1 Source Code1

```
import pandas as pdimport numpy as np
fromflaskimport Flask, render_template, Response, requestimport pickle
from sklearn.preprocessing import LabelEncoderimport pickle
app = Flask(__name__)
filename ='resale_model (1).sav'
model_rand =pickle.load(open(filename, 'rb'))
@app.route('/')def
index():return
render_template('index.hm
l')
@app.route('/home')def
home():return
render_template('index.html'
)@app.route('/predict')def
predict():return
render_template('booking.ht
ml')
        @app.route('/y_predict', methods=['GET', 'POST'])def y_predict():
        regyear = int(request.form['regyear']) powerps =
        float(request.form['powerps'])kms = float(request.form['kms'])
        regmonth = int(request.form.get('regmonth'))gearbox =
        request.form['gearbox']
        damage = request.form['damaged'] model =
        request.form.get('model_type')brand =
        request.form.get('brand') fuelType =
        request.form.get('fuel')
        ('vehicletype'){'yearOfRegistration': 'powerPS':'kilometer':,'monthOfRegistration': th,
                        'gearbox': gearbox,'notRepairedDamage': damage,'model': model,'brand': brand, 'fuelType':
fuelType,'vehicleType':
vehicletype}
```

```
    print(new_row)
    new_df = pd.DataFrame(
        columns=['vehicleType', 'yearOfRegistration', '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(str('classes' + i + '.npy'), allow_pickle=True
        tr = mapper[i].fit_transform(new_df[i])
        new_df.loc[:, i + '_Labels'] = pd.Series(tr, index=new_df.index)
    labeled = new_df[
        ['yearOfRegistration', 'powerPS', 'kilometer', 'monthOfRegistration'] + [x
"_Labels" for x in labels]]

    X = labeled.values
    print(X)
    y_prediction = model_rand.predict(X)
    print(y_prediction)
    return  render_template('booking.html',
                            ypred="The resale value predicted is ₹ {:.2f}
".format(y_prediction[0]))


if __name__ == '__main__':
    app.run(host='localhost', debug=True, threaded=False)
```

index.html

```
1   <!DOCTYPE html>
2  <html>
3      <head>
4          <link rel="stylesheet" href="styles.css">
5          <link rel="stylesheet" type="text/css" href="{{
   url_for('static',filename='styles/formstyle.css') }}">
6          <link rel="stylesheet" type="text/css" href="{{ url_for('static',filename='styles/font-
   awesome.css') }}"><style>
7     background eat: no-repeat;
8     background-color: black;
9     background-image: url('https://w0.peakpx.com/wallpaper/235/193/HD-wallpaper-black-panther-is-sitting-
   in-car-background-black-panther.jpg');
10    background-size: cover;
11    background-attachment: fixed;14
      }
15                  body{
16                      color: white;
17                  }
18                  h1{
19                      font-weight: 100;
20                  font-size: 50px;21
                    }
22                  h2{23
24                      font-size: 45px;
25                      color: white;
26
27                  }
28                  p{
12
```

ep

```
28                  font-weight: 300;
29                  font-size: 33px;
30                  color:white;
32              }
33              #btn{
34                  padding: 0.5rem;
35                  border-color: black ;
36                  font-weight: 100;
37                  font-size: 24px;
38              background-color: bisque;39     }
40          </style>
41      </head>
42          <body>
43          <center>
44          <h1>Car Resale Value Prediction</h1>
45      <h2>welcome</h2>
46      <form action='predict'>
47          <p>Click here to predict the price</p>
48          <input id="btn" type="submit" value="Click here">
49       </form>
50       </center>
51       </body>
52          <\html
```

booking.html

```
<html>

  <head>
    <link rel="stylesheet" href="styles.css">

    <link rel="stylesheet" type="text/css" href="{{
url_for('static',filename='styles/formstyle.css')  }}">
    <link rel="stylesheet" type="text/css" href="{{
url_for('static',filename='styles/font-awesome.css')  }}">
    <style>
          a{color: white;}
        .button {
          border: none;
          color: white;
          padding: 15px 32px;
          text-align: center;
          margin-top: 20px;
          text-decoration: none;
          display: inline-block;
          font-size: 16px;
          margin: 4px 2px;
          margin-top: 2em;
            cursor:pointer;
        }
            padding:1em;
            font-family:"Open Sans","Helvetica Neue", Helvetica, Arial, sans-serif;font-size:15px;
            background-image: url('https://digitalsynopsis.com/wp-
  content/uploads/2014/06/supercar-wallpapers-bugatti-4.jpg');
            color:#b9b9b9;
            background-color: black;background-repeat: no-repeat;background-size:
                cover;
                background-attachment: fixed;
        }

        /* Green */
        .button2 {
            background-color: white;color:
                black;
```

```
border-color:black solid;}
                    </style>
                </head><section class="banner_main">
                    <div
                class="container">
                    <form action="/y_predict" method="post">
                <center>
                                <h1 style="color:beige; font-family: monospace";>Pre
            Price!</h1>
                            </center>

                            <center>
                                <h3 style="font-family:verdana; color: white;">{{ypred
                            </center>
                            <div class="row">
                                <h4>Registration Details</h4>
                                <div class="input-group"><input type="number" name=
                                id="regmonth"
                                placeholder="Registration Year" required/>
    </div><div>
        <h4>Registration Month</h4>
        <!-- <div class="input-group"><input type="number" name="regmonth"
                                        placeholder="Registration Month" />-->
                            <div class="input-group" >
                                <select name="regmonth" id="regmonth">
                                    <option value=1>January</option>
                                        <option value=2>February</option>
                                        <option value=3>March</option>
                                        <option value=4>April</option>
                                        <option value=5>May</option>
                                        <option value=6>June</option>
                                        <option value=7>July</option>
                                        <option value=8>August</option>
                                        <option value=9>September</option>
                                        <option value=10>October</option>
                                        <option value=11>November</option>
                                        <option value=12>Decemeber</option>
```

```
                                    <div><h4>Gear Box Type</h4><div  class="input-group
 <input id="gear-manual"  type="radio"  name="gearbox"value="manual" />value="automa
  <label  for="gear-manual">Manual</label><input  id="gear-automatic"  type="radio"  nam
<label for="gear-automatic">Automatic</label><input id="gear-notdeclared" type="radio
declared" /><label  for="gear-notdeclared">Not declared</label></div>
                                </div><div><h4>Is Car Damaged</h4><div   class="inp
                                        <inputid="yes"  type="radio"  name="damaged"
                                        <label for="yes">Yes</label>
                                        <input  id="damaged-no" type="radio" name="d
                                        <label  for="damaged-no">No</label>
                                        <input  id="damaged-notdeclared"  type="radio"
            value="Not declared" />
                                        <label  for="damaged-notdeclared">Not  declare
                                    </div></div><div><h4>Fuel  Type</h4>
                                    <div  class="input-group"><select name="fuel" id="
                                            <option   value="not-declared">not-declare
                                            <option   value="diesel">diesel</option>
                                            <option   value="petrol">petrol</option>
                                            <option value="lpg">lpg</option>
                                            <option   value="others">others</option>
                                            <option   value="hybrid">hybrid</option>
                                            <option value="cng">cng</option>
                                            <option   value="electric">electric</option>
                                        </select></div></div><div><h4>Brand  of  the
                                    <div   class="input-group"><selectname="brand" id
                                            <option   value="audi">audi</option>
                                            <option   value="jeep">jeep</option>
                                            <option   value="Mahindra">Mahindra</opt
                                            <option  value="volkswagen">volkswagen
                                            <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">mercede
                                            <option value="honda">honda</option>
```

```
                    <option   value="porsche">porsche</option
                    <option   value="citroen">citroen</option>
                    <option value="toyota">toyota</option>
                    <option   value="chevrolet">chevrolet</opti
                    <option value="dacia">dacia</option>
                    <option value="suzuki">suzuki</option>
                    <option   value="daihatsu">daihatsu</optio
                    <option   value="chrysler">chrysler</optior
                    <option   value="sonstige_autos">sonstige_
                    <option   value="jaguar">jaguar</option>
                    <option   value="daewoo">daewoo</option
                    <option value="rover">rover</option>
                    <option value="saab">saab</option>
                    <option   value="land_rover">land_rover</c
                    <option value="lada">lada</option>
                    <option   value="trabant">trabant</option><
</div></div><div><h4>Model  Type</h4>
<div class="input-group">
        <select name="model_type"  id="model_type">
             <option  value="not-declared">not-declare
            <option value="grand">grand</option>
            <option value="scorpio s11">scorpio s11<
            <option value="golf">golf</option>
            <option value="fabia">fabia</option>
            <option value="3er">3er</option>
            <option   value="2_reihe">2_reihe</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="polo">polo</option>
            <option value="twingo">twingo</option>
            <option   value="a_klasse">a_klasse</optio
            <option   value="scirocco">scirocco</optio
            <option value="5er">5er</option>
            <option value="andere">andere</option>
            <option   value="civic">civic</option>
```

```
<option value="vito">vito</option>
<option  value="sprinter">sprinter</option>
<option value="astra">astra</option>
<option value="156">156</option>
<option value="escort">escort</option>
<option  value="forester">forester</option
<option  value="xc_reihe">xc_reihe</optio
<option  value="fiesta">fiesta</option>
<option  value="scenic">scenic</option>
<option  value="ka">ka</option>
<option  value="a1">a1</option>
<option value="transporter">transporter</
<option value="focus">focus</option>
<option value="a4">a4</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="combo">combo</option>
<option  value="corsa">corsa</option>
<option value="80">80</option>
<option value="147">147</option>
<option value="glk">glk</option>
<option  value="z_reihe">z_reihe</option>
<option  value="sorento">sorento</option>
<option value="ibiza">ibiza</option>
<option  value="mustang">mustang</optio
<option value="eos">eos</option>
<option value="touran">touran</option>
<option value="getz">getz</option>
<option  value="insignia">insignia</option
<option  value="almera">almera</option>
<option  value="megane">megane</option
<option value="a3">a3</option>
<option value="r19">r19</option>
```

```
<option value="panda">panda</option>
<option value="up">up</option>
<option value="i_reihe">i_reihe</option>
<option value="ceed">ceed</option>
<option value="kangoo">kangoo</option>
<option value="5_reihe">5_reihe</option>
<option value="yeti">yeti</option>
<option value="octavia">octavia</option>
<option value="zafira">zafira</option>
<option value="mii">mii</option>
<option value="rx_reihe">rx_reihe</optio
<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="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</optio
<option value="1_reihe">1_reihe</option>
<option value="sharan">sharan</option>
<option value="avensis">avensis</option>
<option value="sl">sl</option>
<option value="roomster">roomster</optio
<option value="q5">q5</option>
<option value="santa">santa</option>
<option value="leon">leon</option>
```

```
<option value="v40">v40</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="rav">rav</option>
<option value="kuga">kuga</option>
<option  value="qashqai">qashqai</option>
<option value="swift">swift</option>
<option  value="picanto">picanto</option>
<option value="superb">superb</option>
<option value="stilo">stilo</option>
<option  value="alhambra">alhambra</opti
<option value="911">911</option>
<option  value="m_reihe">m_reihe</option
<option  value="roadster">roadster</option
<option  value="ypsilon">ypsilon</option>
<option value="galant">galant</option>
<option value="justy">justy</option>
<option value="90">90</option>
<option  value="sirion">sirion</option>
<option  value="signum">signum</option>
<option  value="crossfire">crossfire</optio
<option value="agila">agila</option>
<option value="duster">duster</option>
<option value="v50">v50</option>
<option  value="mx_reihe">mx_reihe</opt
<option value="meriva">meriva</option>
<option  value="discovery">discovery</opt
<option  value="c_reihe">c_reihe</option>
<option  value="v_klasse">v_klasse</optio
<option value="yaris">yaris</option>
<option value="c5">c5</option>
<option value="aygo">aygo</option>
```

```
<option value="spark">spark</option>
<option value="x_type">x_type</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="voyager">voyager</option
<option value="calibra">calibra</option>
<option value="v70">v70</option>
<option value="bravo">bravo</option>
<option value="range_rover">range_rover
<option value="forfour">forfour</option>
<option value="tucson">tucson</option>
<option value="q7">q7</option>
<option value="c1">c1</option>
<option value="citigo">citigo</option>
<option value="jimny">jimny</option>
<option value="cx_reihe">cx_reihe</optio
<option value="cayenne">cayenne</option
<option value="wrangler">wrangler</optio
<option value="lybra">lybra</option>
<option value="range_rover_sport">range
<option value="lancer">lancer</option>
<option value="freelander">freelander</op
<option value="captiva">captiva</option>
<option value="range_rover_evoque">
range_rover_evoque</option>
<option value="sandero">sandero</option
<option value="note">note</option>
<option value="antara">antara</option>
<option value="900">900</option>
<option value="defender">defender</opti
<option value="cherokee">cherokee</opti
<option value="clubman">clubman</optio
<option value="arosa">arosa</option>
<option value="legacy">legacy</option>
```

```
                croma</option>
                <option  value="outlander">outlander</opt
                <option value="gl">gl</option>
                <option value="kaefer">kaefer</option>
                <option value="doblo">doblo</option>
                <option value="musa">musa</option>
                <option  value="amarok">amarok</option>
                <option value="9000">9000</option>
                <option value="kalos">kalos</option>
                <option  value="v60">v60</option>
                <option  value="200">200</option>
                <option  value="145">145</option>
                <option  value="b_max">b_max</option>
                <option  value="delta">delta</option>
                <option value="aveo">aveo</option>
                <option  value="rangerover">rangerover</o
                <option value="move">move</option>
                <option  value="materia">materia</option>
                <option  value="terios">terios</option>
                <option  value="kalina">kalina</option>
                <option  value="elefantino">elefantino</op
                <option  value="i3">i3</option><option  v
                <option value="kappa">kappa</option>
                <option  value="serie_3">serie_3</option>
                <option  value="discovery_sport">discover
            </select>
        </div>
    </div>


    <div>
      <h4>Vehicle  Type</h4>
        <divclass="custom">
     <select  name="vehicletype"  id="wgtmsr">
       <option  value="coupe">coupe</option>
                <option value="suv">suv</option>
                <option  value="small  car">small  car</opti
```

3.2

GitHub: **IBM-Project-49171-1660816515**

ProjectDemoLink:https://drive.google.com/file/d/1-_21JeJtUi_u8ndkazGWcME2dOGlKvec/view?usp=