**A PROJECT REPORT ON CAR RESALE VALUE PREDICTION**

**DOMAIN :** Applied Data Science

**TEAM :**IBM-Project-35390-1660284416

**COLLEGE NAME** : Government College of Engineering-Bargur

**TEAM MEMBERS**

**TEAM ID:** **PNT2022TMID06615**

|  |
| --- |
| TEAM LEADER : GAJALAKSHMI R(61071911114) |
| TEAM MEMBER 1: SAVITHA R(61071911144) |
| TEAM MEMBER 2: RAMYA S (61071911136) |
| TEAM MEMBER 3: SOWFIYA BEGAM S(61071911146) |

**TABLE OF CONTENT**

1. **INTRODUCTION**

1.1Project Overview

1.2Purpose

1. **LITERATURE SURVEY**

2.1Existing problem

2.2References

2.3Problem Statement Definition

1. **IDEATION & PROPOSED SOLUTION**

3.1Empathy Map Canvas

3.2Ideation & Brainstorming

3.3Proposed Solution

3.4Problem Solution fit

1. **REQUIREMENT ANALYSIS**

4.1Functional requirement

4.2Non-Functional requirements

1. **PROJECT DESIGN**

5.1Data Flow Diagrams

5.2Solution & Technical Architecture

5.3User Stories

1. **PROJECT PLANNING & SCHEDULING**

6.1Sprint Planning & Estimation

6.2Sprint Delivery Schedule

6.3Reports from JIRA

1. **CODING & SOLUTIONING (Explain the features added in the project along with code)**

7.1Feature 1

7.2 Feature 2

7.3Database Schema (if Applicable)

1. **TESTING**

8.1Test Cases

8.2User Acceptance Testing

1. **RESULTS**

9.1Performance Metrics

1. **ADVANTAGES & DISADVANTAGES**
2. **CONCLUSION**
3. **FUTURE SCOPE**
4. **APPENDIX**

Source Code

GitHub & Project Demo Link

**CHAPTER 1**

**INTRODUCTION**

**1.1 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 kilometre driven, fuel type, etc. Need for the System 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 loss. Car 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 biased valuation.

**BACKGROUND**

In this project, we mainly focus on the analysis of the Vehicle Resale Predict and then predict the results through them using training data. AI calculations can be utilized to anticipate the retail worth of a vehicle, in light of a specific arrangement of highlights. Various sites have various calculations to create the retail cost of the trade-in vehicles, and subsequently there is certainly not a brought together calculation for deciding the cost. Via preparing measurable models at foreseeing the costs, one can undoubtedly get a good guess of the cost without really entering the subtleties into the ideal site.

**1.2PROJECT PURPOSE**

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.The 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, kilometres driven, fuel of car, year of purchase.Car 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 biased valuation. The new system developed by us consists of two parts - Data gathering and Prediction using **Machine Learning based algorithms**.

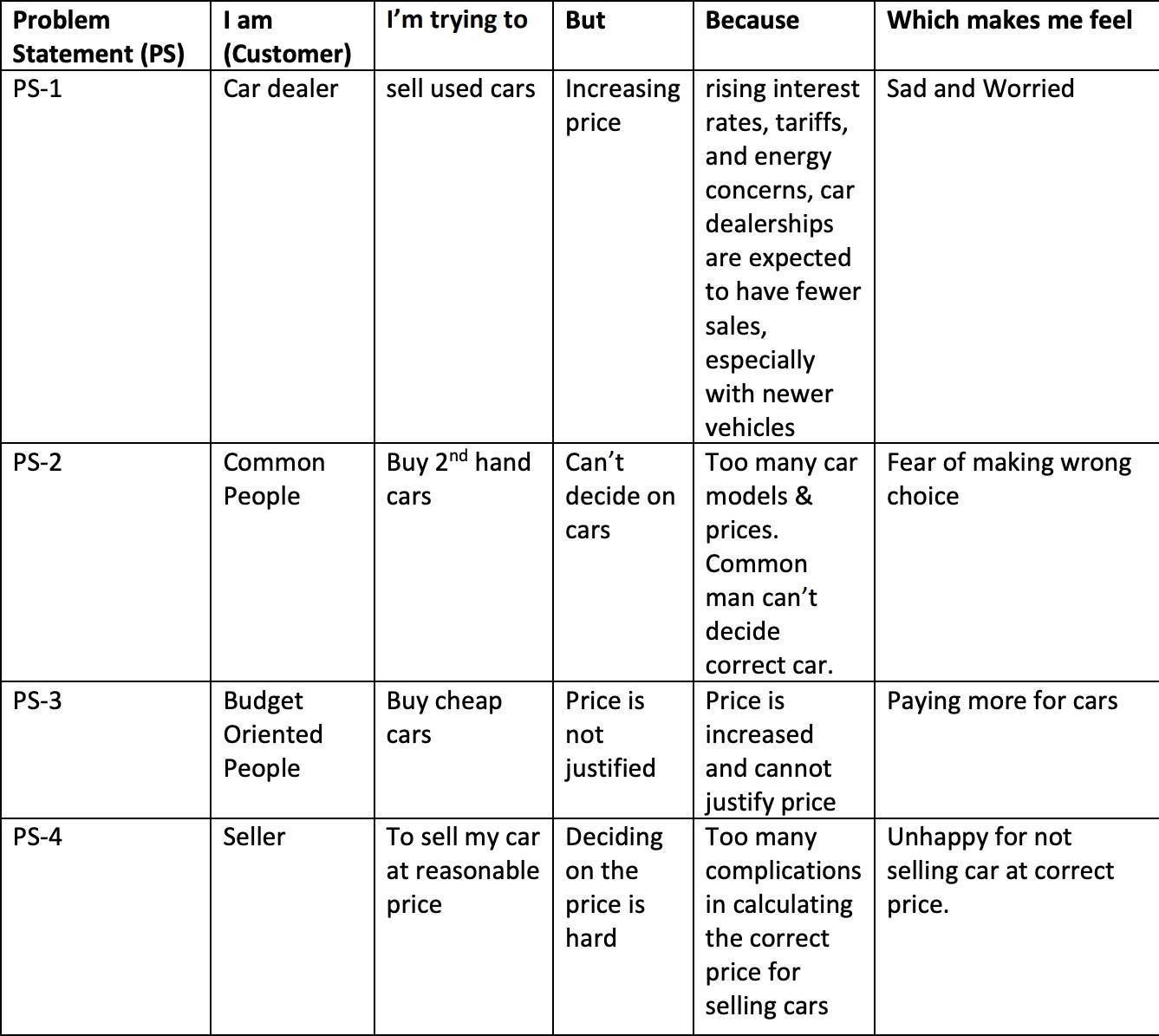
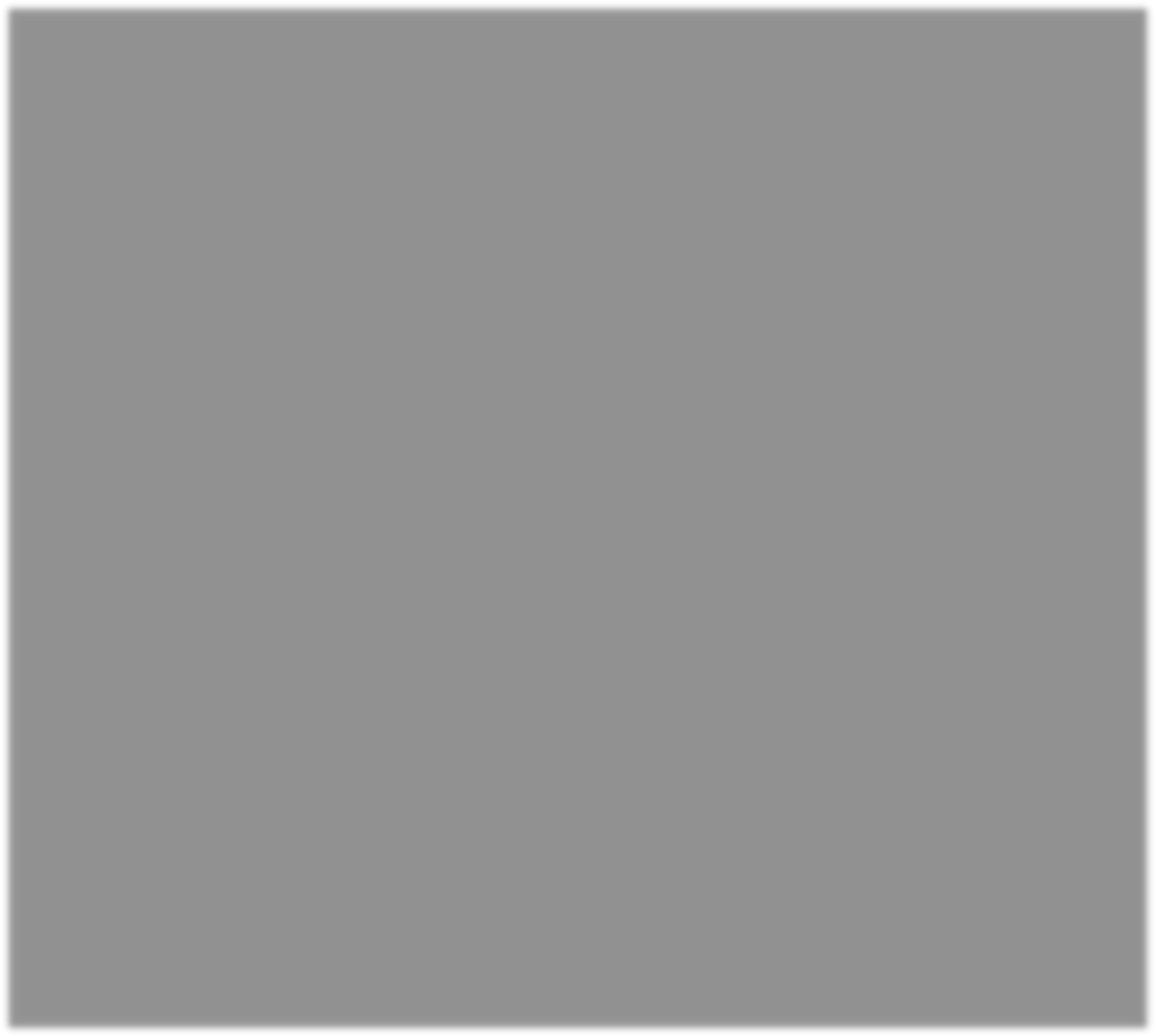
**CHAPTER 2**

**LITERATURE SURVEY**

The first paper is Predicting the price of Used Car Using Machine Learning Techniques. In this paper, they investigate the application of supervised machine learning techniques to predict the price of used cars in Mauritius. The predictions are based on historical data collected from daily newspapers. Different techniques like multiple linear regression analysis, k-nearest neighbours, naïve bayes and decision trees have been used to make the predictions. The Second paper is Car Price Prediction Using Machine Learning Techniques. Considerable number of distinct attributes are examined for the reliable and accurate prediction

**2.1 EXISTING PROBLEM**

.



**2.2 REFERENCES**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.NO** | **AUTHORS** | **PAPER** | **YEAR** |
| **1** | Sameerchand Pudaruth | Predicting the Price of Used Cars using Machine Learning Techniques | IJICT 2014 |
| **2** | Enis gegic, Becir Isakovic  Dino Keco  Zerina Masetic  Jasmin Kevric, | Car Price Prediction Using Machine Learning | TEM Journal 2019 |
| **3** | Ning sun  Hongxi Baily  uxia Geng  Huizhu Shi | Price Evaluation Model In Second Hand Car System Based On BP Neural Network Theory | Hohai University Changzhou, China |
| **4** | Nitis Monburinon  Prajak Chertchom  Thongchai Kaewkiriya  Suwat Rungpheung  Sabir Buya  Pitchayakit Boonpou, | Prediction of Prices for Used Car by using Regression Models | ICBIR 15 |
| **5** | Doan Van Thai  Luong Ngoc Son  Pham Vu Tien  Nguyen Nhat Anh  Nguyen Thi Ngoc Anh | Prediction car prices car prices using qualify qualitative data and knowledge-based system | Hanoi National University |

**2.3 PROBLEM STATEMENT**

**DEFINITION**

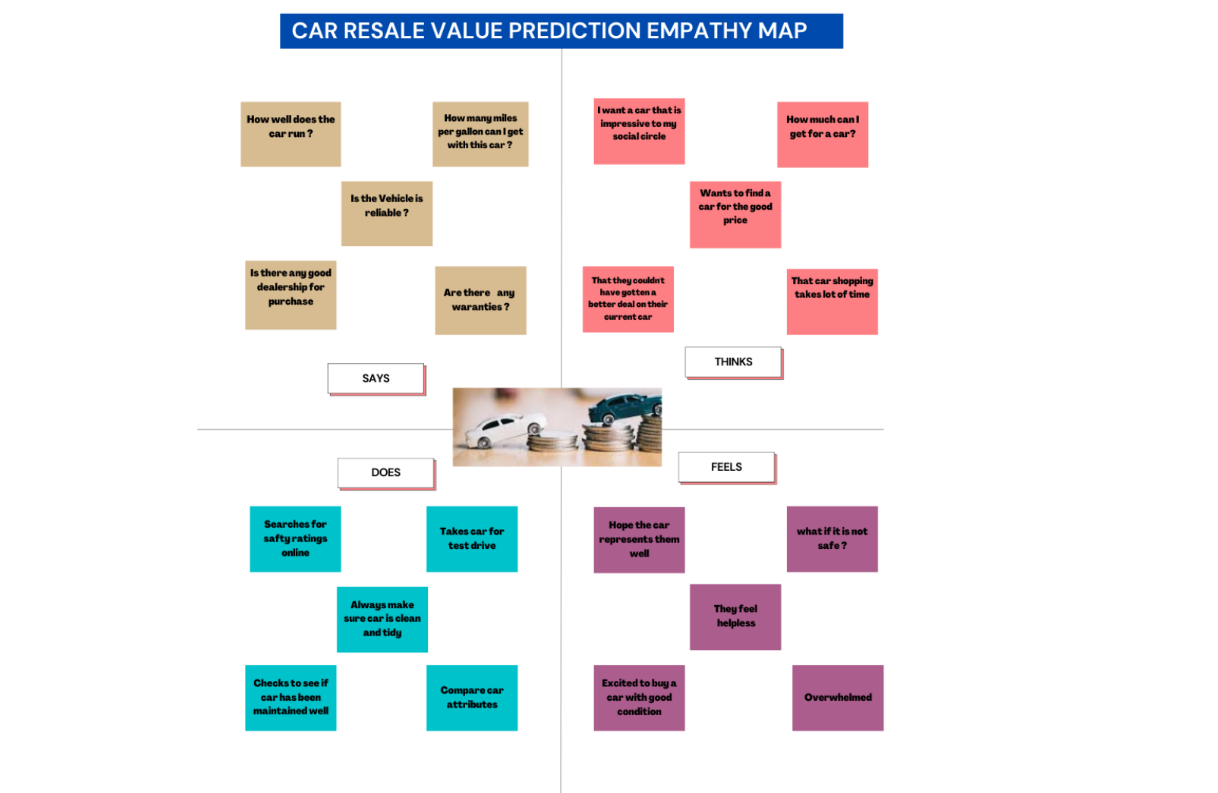
The main aim of this project is to predict the price of used cars using the various Machine Learning (ML) models. This can enable the customers to make decisions based on different inputs or factors namely Brand or Type of the car one prefers like Ford, Hyundai. Model of the car namely Ford Figo, Hyundai Creta. Location like Delhi, Chennai, Mumbai. It based on year of manufacturing like 2020, 2022, type of fuel namely Petrol, Diesel. Price range or Budget and type of transmission which the customer prefers like Automatic or Manual, Mileage to name a few characteristic features required by the customer. The project Car Price Prediction deals with providing the solution to these problems. Car makers face several challenges in the second-hand market. The depth crisis in the European Union, the general problem of overcapacity, increasing competition from Asian manufacturers, and the trend toward more eco-friendly cars are only a few factors that add to the difficulty of selling used vehicles in the second-hand market and decrease sales margins. Therefore, car makers require sophisticated decision support systems to sustain the profitability of the used car business. A core component of such systems is a prediction model that estimates resale prices on the basis of car attributes and other factors.

**CHAPTER 3**

**IDEATION & PROPOSED SOLUTION**

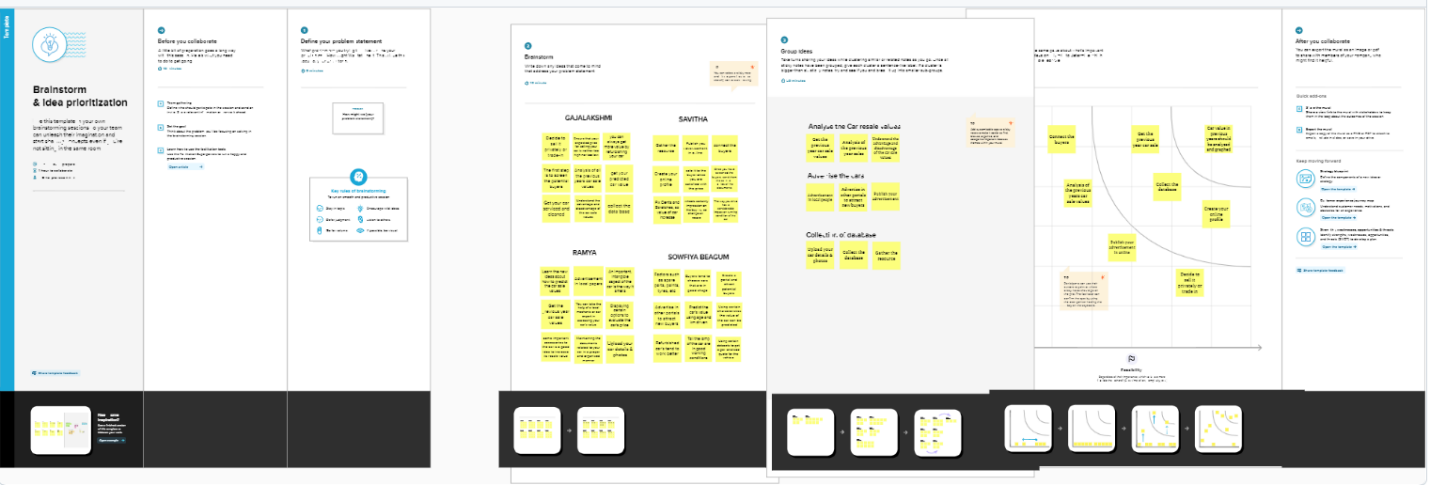
**3.1 EMPATHY MAP**

An empathy map is a visualization tool used to articulate what a product team knows about a user. This tool helps product teams build a broader understanding of the “why” aspect behind user needs and wants. This tool forces product teams to practice empathic design, which shifts the focus from the product they want to build to the people who will use this product. The four empathy map quadrants look at what the user **says**, **thinks**, **feels**, and **does**.

****

**3.2 IDEATION AND BRAINSTROM**

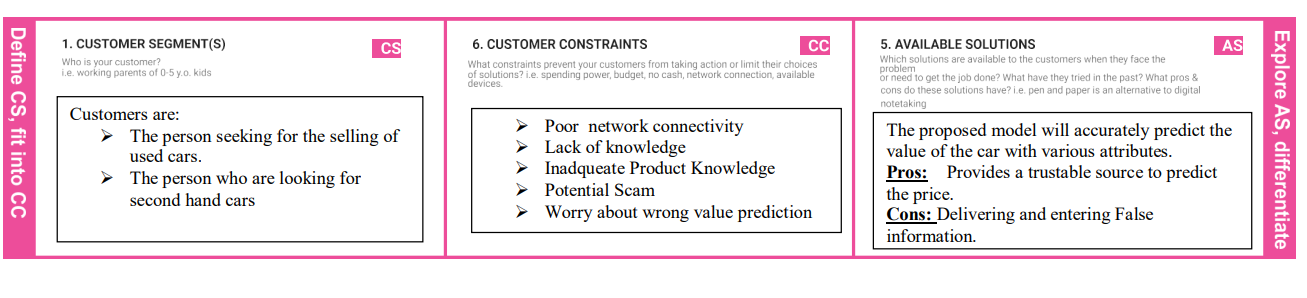
The term brainstorming has now become the accepted way of referring to group attempts to solve specific problems or develop new ideas by amassing spontaneous, unrestrained contributions by members. Brainstorming can be used to generate possible solutions for simple problems, but it is unrealistic to expect it to accomplish most problem-solving or planning tasks. The technique is of value as part of a larger effort that includes individual generation of information and ideas and subsequent compilation, evaluation, and selection. Brainstorming can be used to generate components of a plan, process, solution, or approach and to produce checklists.

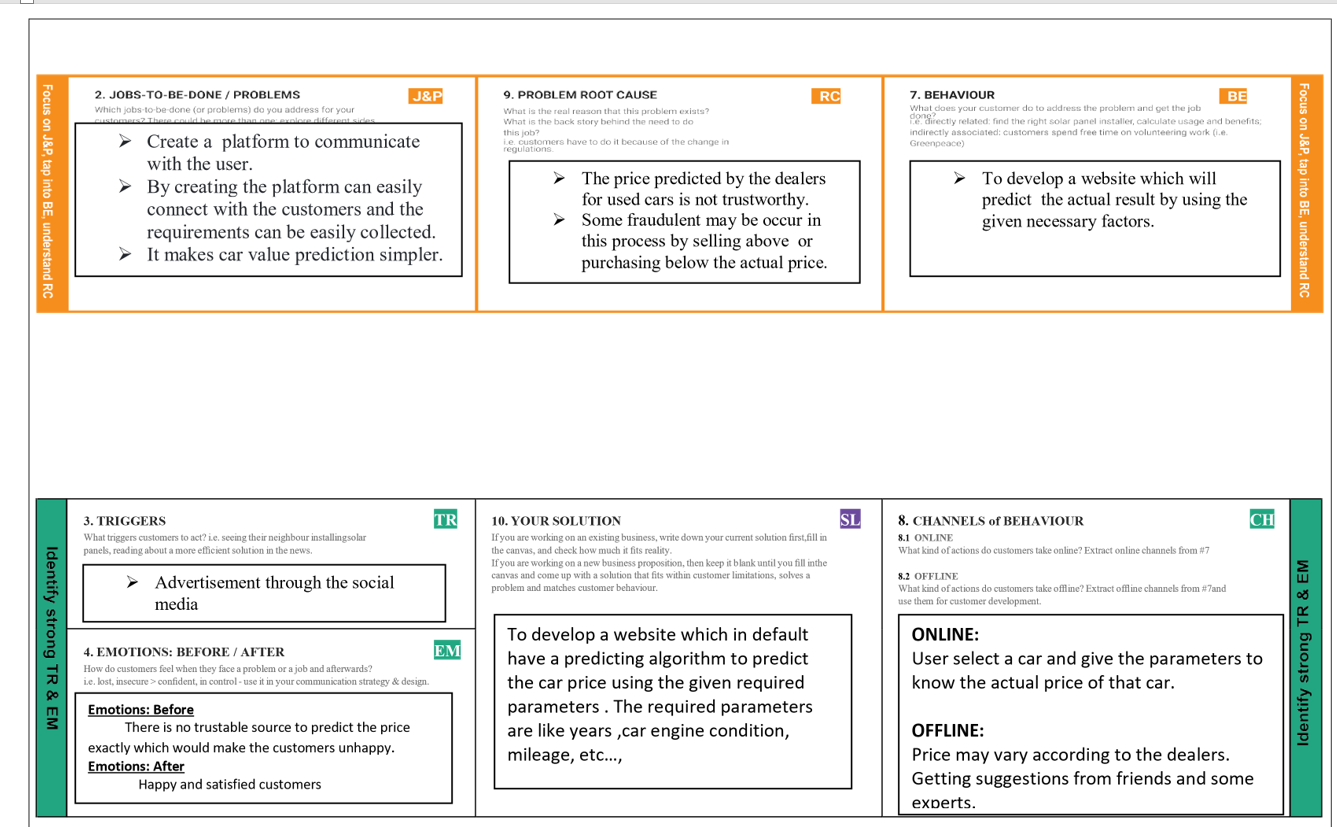


**3.3 PROPOSED SOLUTION**

|  |  |  |
| --- | --- | --- |
| **S.NO** | **Parameter** | **Description** |
| **1** | Problem statement | The main aim of project is to predict the resale value of a used car using linear regression algorithms. |
| **2** | Idea/ solution description | The resale value of car depends on km, fuel etc  The data is then pre-processed to handle missing values and outliers and split into depend variable and independent variable.  After that the model is developed using regression algorithms to predict the resale price to the car. |
| **3** | Novelty/Uniqueness | This is a real time problem which can benefit both customer and seller.  The novelty of this proposal is to predict the resale value as near as possible to actual value. |
| **4** | Social impact | Provided the current economic times it is more likely that the user age of second-hand cars will increase. This mutual commercial interest to both customer and sellers can increase profit. |
| **5** | Business model | The proposed solution could be sold to reseller so that they could use to find perfect price for bidding.  It could be developed into an application and get revenue from it if more on of users started to using it to find the best value to second hand cars. |

**3.4 PROBLEM SOLUTION FIT**

****

****

**CHAPTER 4**

**REQIUREMENT ANALYSIS**

**4.1 FUNCTIONAL REQUIREMENT**

Followingarethefunctionalrequirementsoftheproposedsolution.

|  |  |  |
| --- | --- | --- |
| **FRNo.** | **FunctionalRequirement(Epic)** | **SubRequirement(Story/Sub-Task)** |
| FR-1 | UserRegistration | RegistrationthroughWebsite |
| FR-2 | UserConfirmation | Confirmationviawebsite |
| FR-3 | CarRegistration | Registeringthecardetails |
| FR-4 | ValuePrediction | Predictingthecarresalevalue |

**4.2 NON-FUNCTIONAL ANALYSIS**

Followingarethenon-functionalrequirementsoftheproposedsolution.

|  |  |  |
| --- | --- | --- |
| **FRNo.** | **Non-FunctionalRequirement** | **Description** |
| NFR-1 | Usability | Predictingtheresalevalue |
| NFR-2 | Security | Providingsecuritytothewebsite |
| NFR-3 | Reliability | Providinghighreliabilitybypredictingvaluesfordifferenttypesofcars |
| NFR-4 | Performance | Providinghighperformance by usingsomemachinelearningtechniques |
| NFR-5 | Availability | Itisusedforalltypesofcars |
| NFR-6 | Scalability | Predictingvaluesfordifferenttypesofcars |

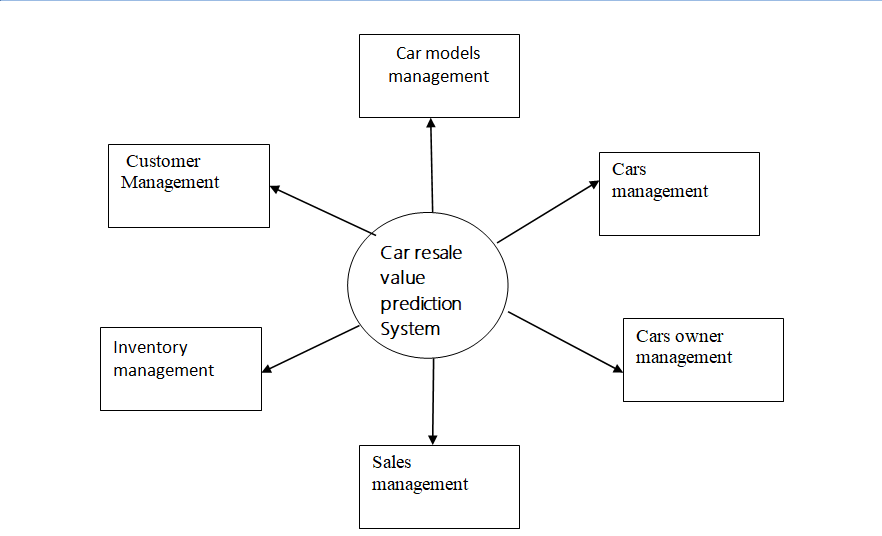
**CHAPTER 5**

**PROJECT DESIGN**

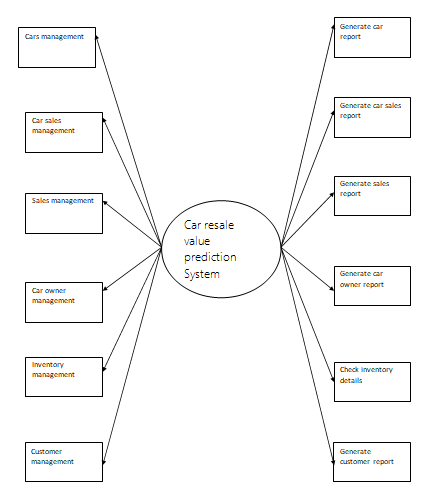
**5.1 DATA FLOW DIAGRAM**

ADataFlowDiagram(DFD)isatraditionalvisualrepresentationofthe informationflowswithina system. A neat and clear DFD can depict the right amount of the system requirementgraphically. It shows how data enters and leaves the system, what changes the information,andwheredatais stored.

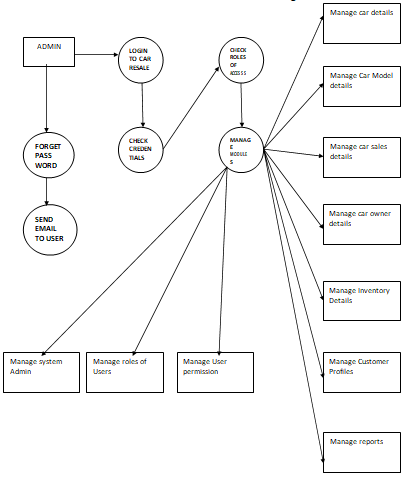
**ZeroLevelDataFlowDiagram**



**First Level Data Flow Diagram**

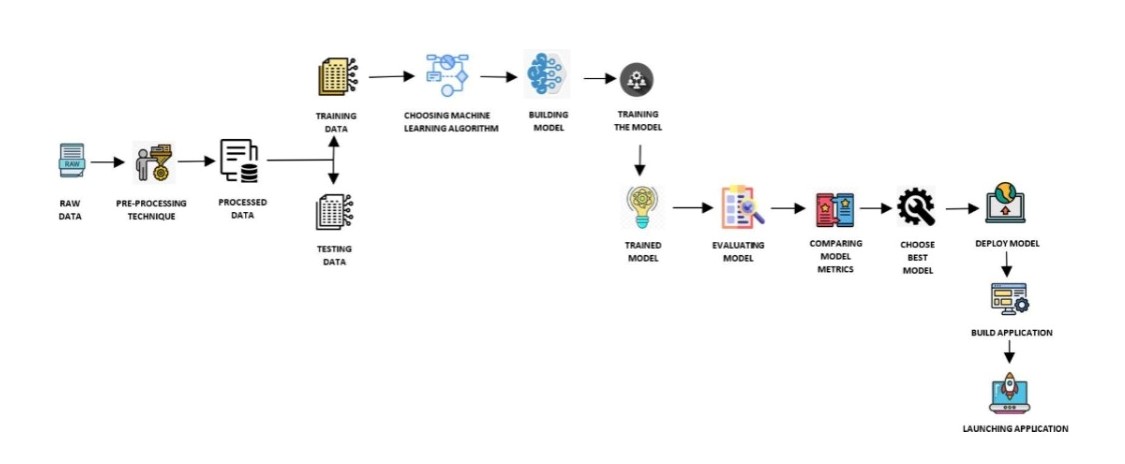
****

# **Second LevelDataFlowDiagram**

****

**5.2SOLUTION & TECHNICAL**

**ARCHITECTURE**

****

**CHAPTER 6**

**PROJECT PLANNING & SCHEDULING**

**6.1 SPRINT PLANNING & ESTIMATION**

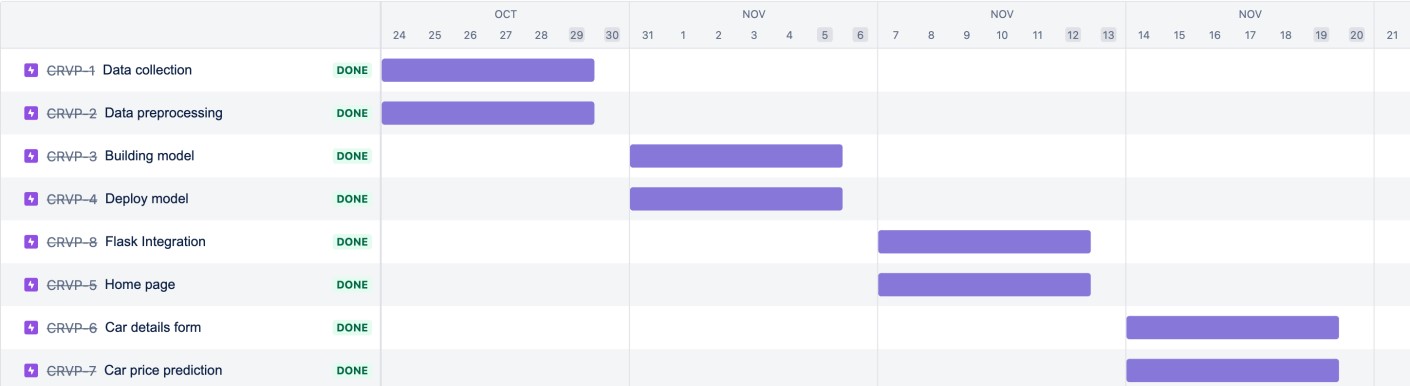
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **USER** | **SPRINT** | **FUNCTIONAL REQUIREMENT** | **USER STORY NUMBER** | **USER STORY/TASK** | **PRIORITY** | **TEAM MEMBER** |
| **Admin** | Sprint 1 | Dataset collection | USN-1 | Collect the required data for the car resale prediction | High | Sowfiya Begam S |
|  | Sprint 1 |  | USN-2 | Perform data cleaning to optimize the dataset | Medium | Sowfiya Begam S |
|  | Sprint 2 | Model building | USN-3 | Build the model using regression algorithms classify the data | High | Ramya S |
|  | Sprint 2 |  | USN-4 | Deployment of ML model using IBM cloud | High | Ramya S |
|  | Sprint 3 | Train the model | USN-5 | Integrate the web app developed using flask with IBM model | High | Savitha R |
| **Customer** | Sprint 3 |  | USN-6 | Details about the application and the car resale process | Low | Savitha R |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **User** | **Sprint** | **FUNCTIONAL REQUIREMENT** | **USER STORY NUMBER** | **USER STORY/TASK** | **PRIORITY** | **TEAM MEMBERS** |
|  | Sprint 4 | Car details | USN-7 | As a user I should give the car details like car model engine and fuel type. | Medium | Gajalakshmi R |
|  | Sprint 4 | Car details | USN-8 | As a user I can view the current rate of the used car price | High | Gajalakshmi R |

**6.2 SPRINT DELIVERY SCHEDULE**



**6.3 REPORTS FROM JIRA**



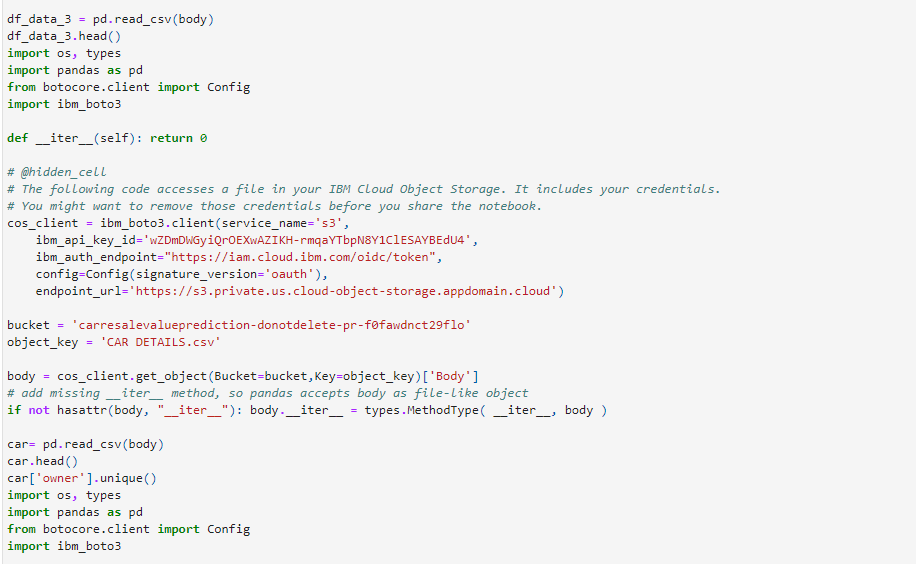
**CHAPTER 7**

**CODING & SOLUTIONING**

**7.1 LINEAR REGRESSION**

****

**7.2DEPLOYED THE MODEL IN IBM CLOUD**

****

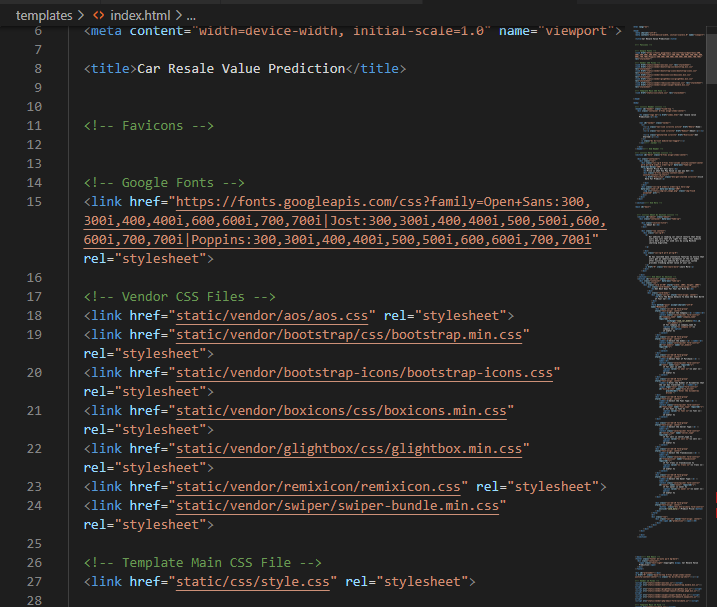
**7.3 USING DEPLOYED MODEL TO PREDICT THE RESULT**



Used IBM Watson to deploy the model, instead of storing the large modelfile in the local which is not feasible for practical use. This code sends the API request to the deployed model along with the data that the user had entered for which we want to predict the result. After successful prediction the result comes in the form of json which is later parsed and the resale value is obtained.

****

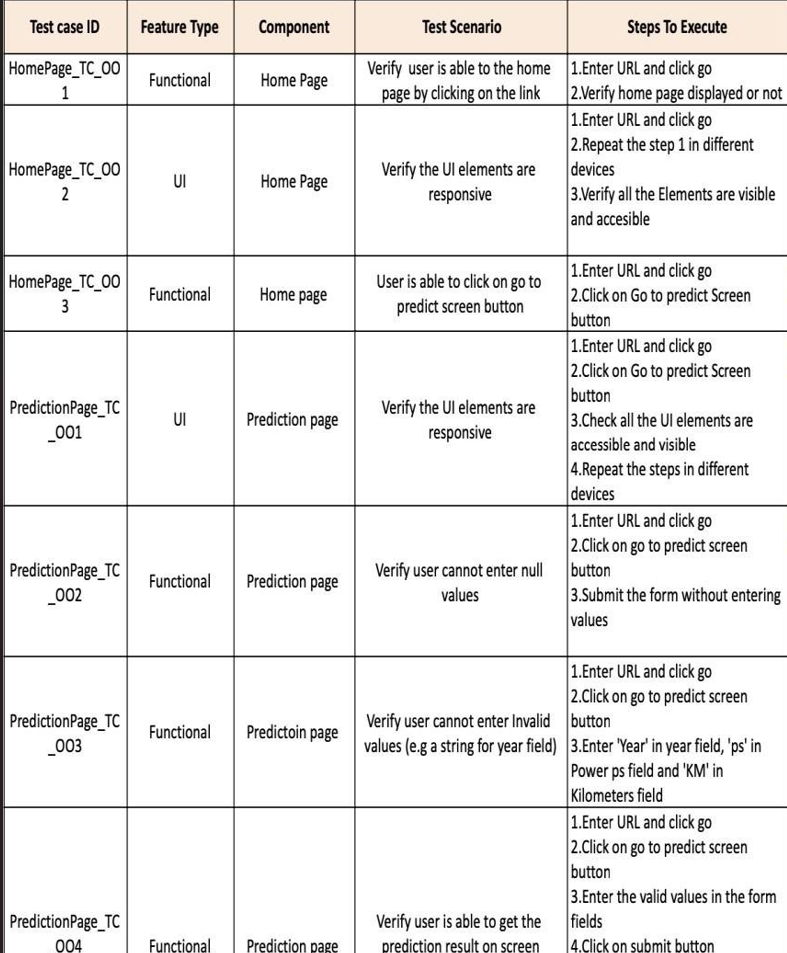
**7.5 BOOTSTRAP**

****

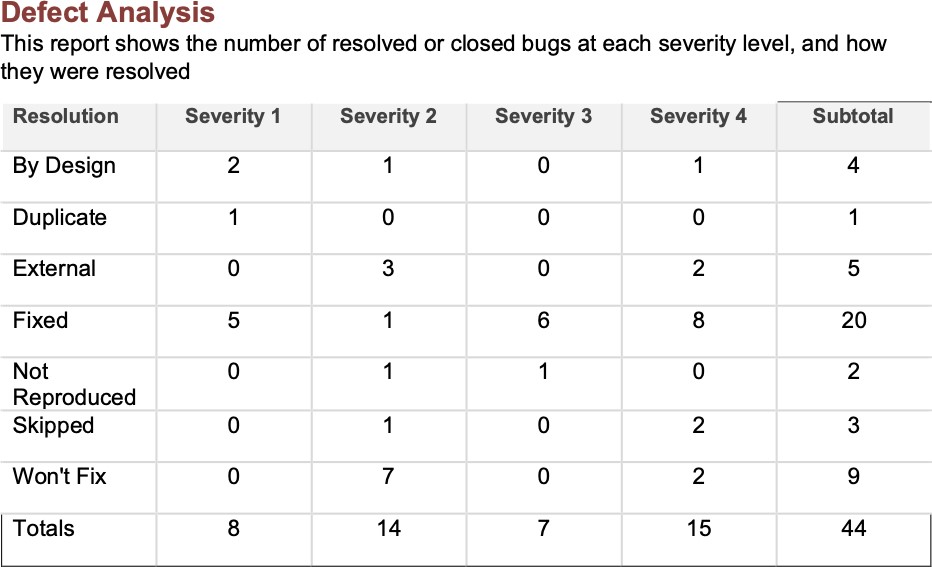
**CHAPTER 8**

**TESTING**

* 1. **TEST CASES**

****

* 1. **USER ACCEPTANCE TESTING**



**CHAPTER 9**

**RESULTS**

**PERFORMANCE METRICS**

****

**CHAPTER 10**

**ADVANTAGES**

* Very Fast prediction .
* Simple UI, just click the link and predict .
* Better accuracy without overfitting .
* High availability .

**DISADVANTAGES**

* If the actual value is higher then, the data points may have higher deviation in the `predicted value.
* User entered data are not stored in database .

**CONCLUSION**

Thus, by using Random Forest regressor, the model can predict the resale value of the car with maximum accuracy without overfitting. It also consumes low memory and faster than other regressors for this dataset. Also, deploying the model in IBM cloud allows us to use the model from the hosted website. In depth analysis and powerful computers can make this model more accurate.

**FUTURE SCOPE**

This project will be more useful in future, as renting and reselling of a car is becoming more common, there are some services such as cars24.com where we can buy and sell second hand cars, the customers of these services may need to estimate or predict the resale value of the cars based on its configuration.

**APPENDIX**

Source Code :

<https://github.com/IBM-EPBL/IBM-Project-35390-1660284416/tree/main/Final%20Deliverables/PROJECT>

GitHub:

<https://github.com/IBM-EPBL/IBM-Project-35390-1660284416>

Project Demo Link:

<https://youtu.be/yHxr-zQD0ds>