PROJECT REPORT

TEAM ID: PNT2022TMID37526

Project Name: Car Resale Value Prediction

Team lead: B.Manikandaan
Team members:
K.Manikandan
J.Fennjoy
S.Praveen Kumar
M.Karuppasamy

1. INTRODUCTION

1.1 Project Overview

- To understand the problem to classify if it is a regression or a classification kind of problem.
- To pre-process/clean the data using different data pre-processing techniques.
- Applying different algorithms according to the dataset

1.2 Purpose

To build a working web application using the Python Flask Framework and deploy our built model on it to meet user satisfaction.

2. LITERATURE SURVEY

2.1 Existing problem

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.

2.2 References

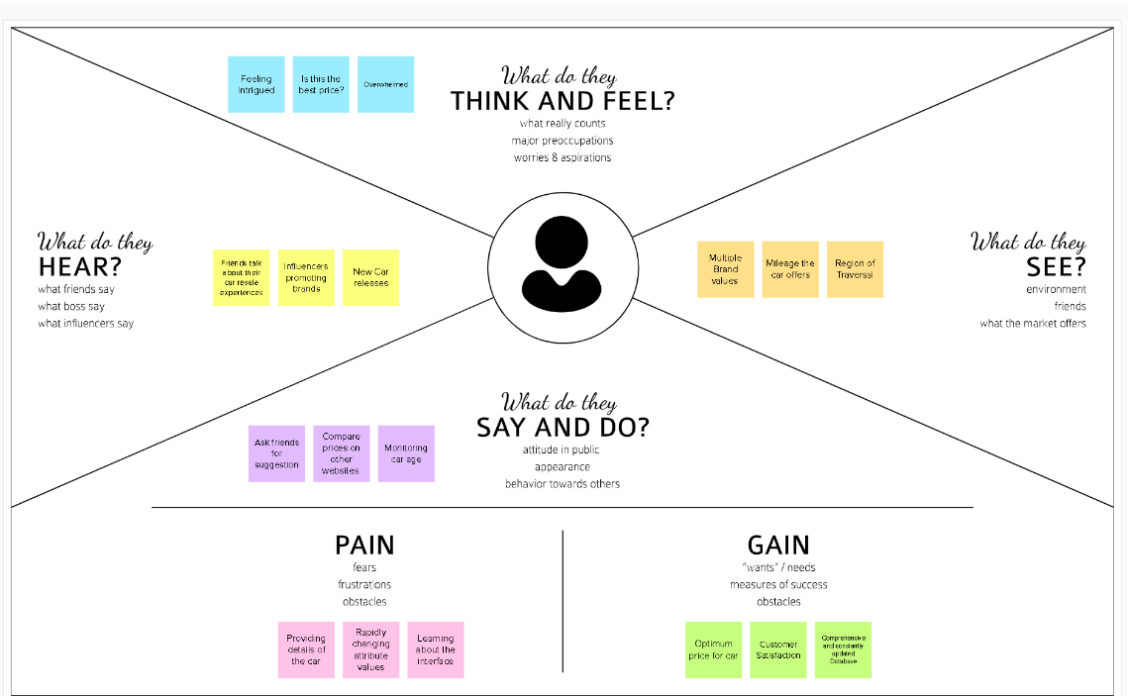
- Voß, S. (2013). Resale Price Prediction in the Used Car Market.
- Kiran, S., 2020. Prediction of resale value of the car using linear regression algorithm. *Int. J. Innov. Sci. Res. Technol*, *6*(7), pp.382-386.\
- Gegic, E., Isakovic, B., Keco, D., Masetic, Z. and Kevric, J., 2019. Car price prediction using machine learning techniques. *TEM Journal*, *8*(1), p.113.
- Ganesh, Mukkesh & Venkatasubbu, Pattabiraman. (2019). Used Cars Price Prediction using Supervised Learning Techniques. International Journal of Engineering and Advanced Technology. 9. 216-223. 10.35940/ijeat.A1042.1291S319.
- Pudaruth, S., 2014. Predicting the price of used cars using machine learning techniques. *Int. J. Inf. Comput. Technol*, *4*(7), pp.753-764.

2.3 Problem Statement Definition

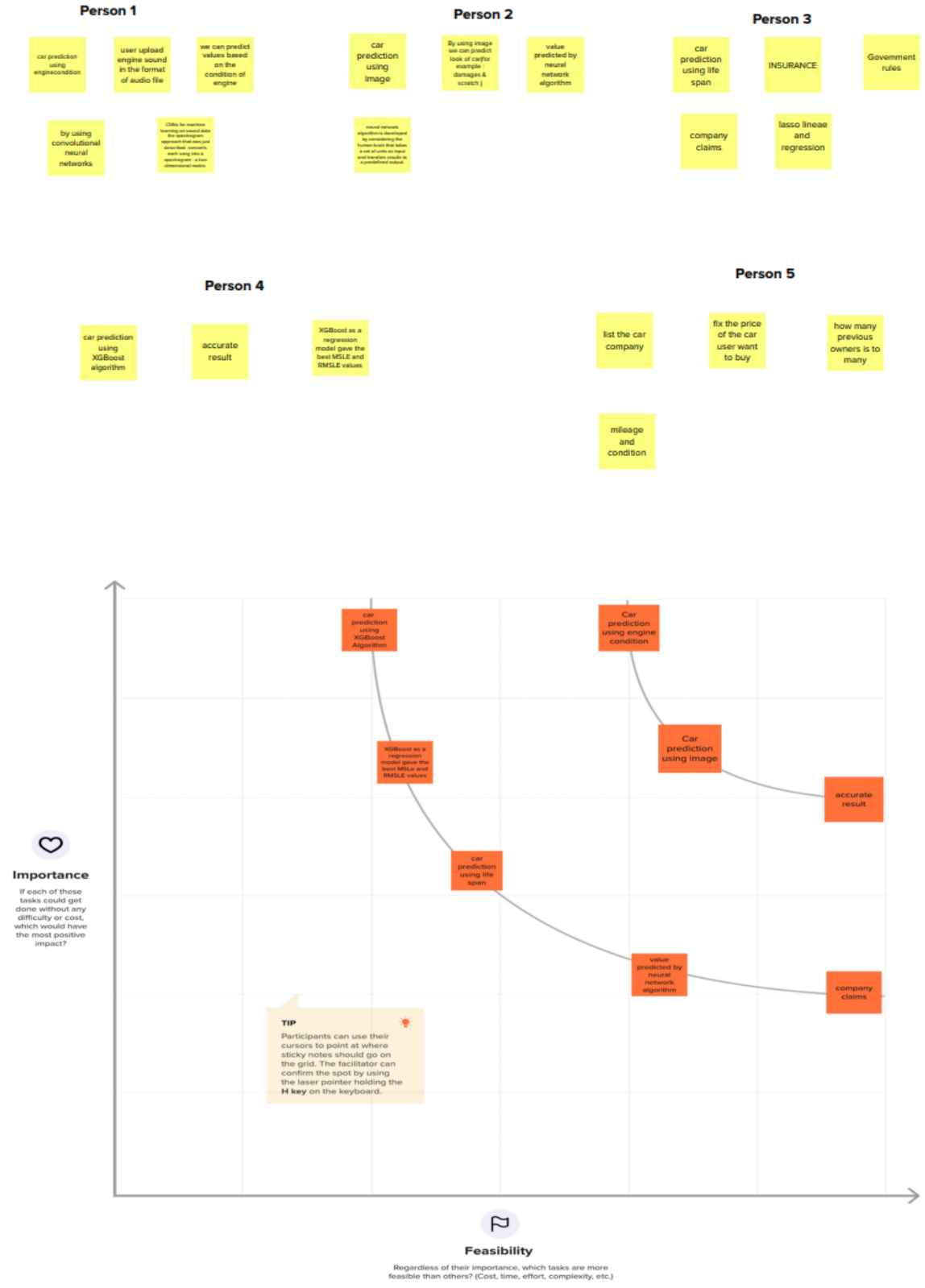
To predict the resale value of the car, we proposed an intelligent, flexible, and effective system that is based on using regression algorithms. Considering the main factors which would affect the resale value of a vehicle a regression model is to be built that would give the nearest resale value of the vehicle.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



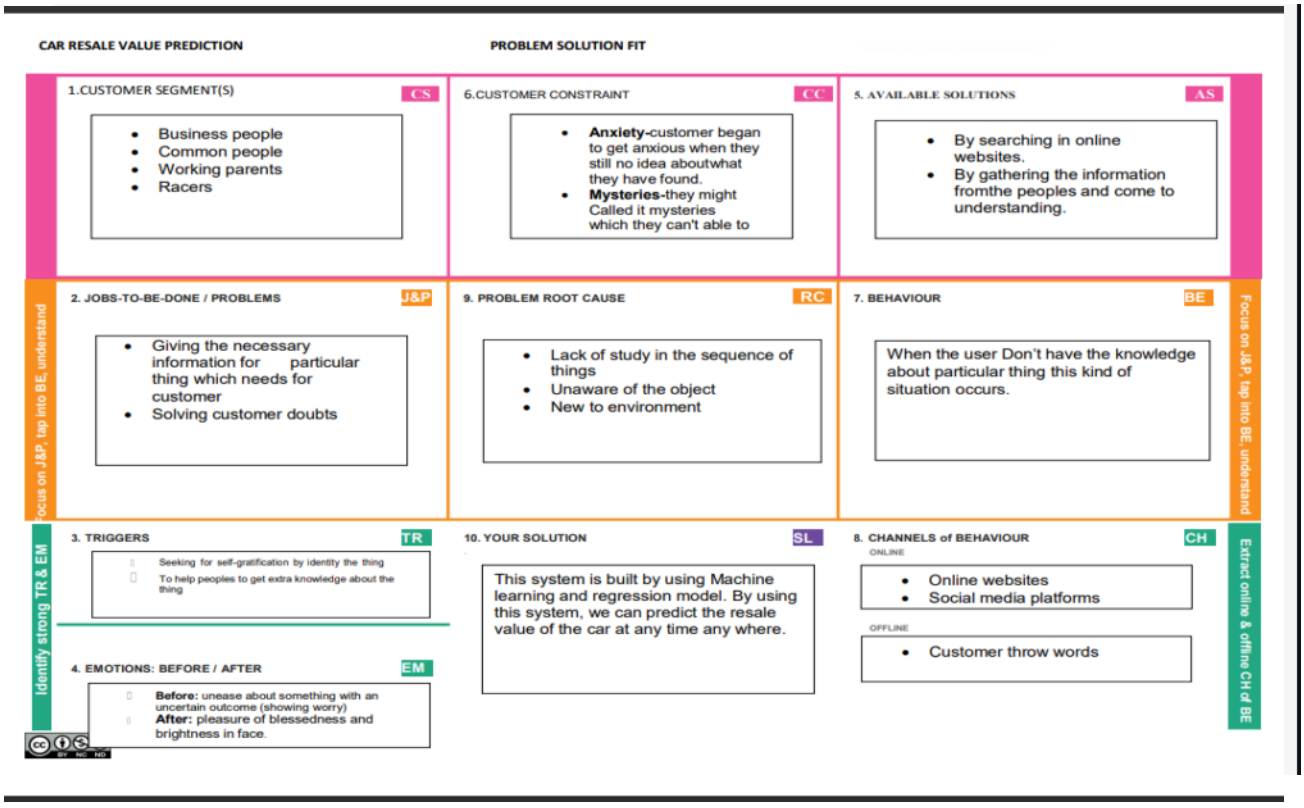
3.2 Ideation & Brainstorming



3.3 Proposed Solution

The dataset has to be pre-processed in an appropriate way prior to Machine Learning model implementation. The stacking algorithm is employed to create an ensemble of strong learners. These learning algorithms can be finalized through experimentation. The model deployment is executed via the Flask platform and the model will be deployed via “pikl” files and on the IBM Cloud platform.

3.4 Problem Solution fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration and Login	Registration through Form Registration through Email Login through Email
FR-2	User should be able to input car details	Car information like date of purchase, price, damages incurred, etc are entered by the user
FR-3	User should be able to view past predictions	User can view the previous predictions the model has made on different cars and categorise according to the brand, type of car, date of purchase, etc.

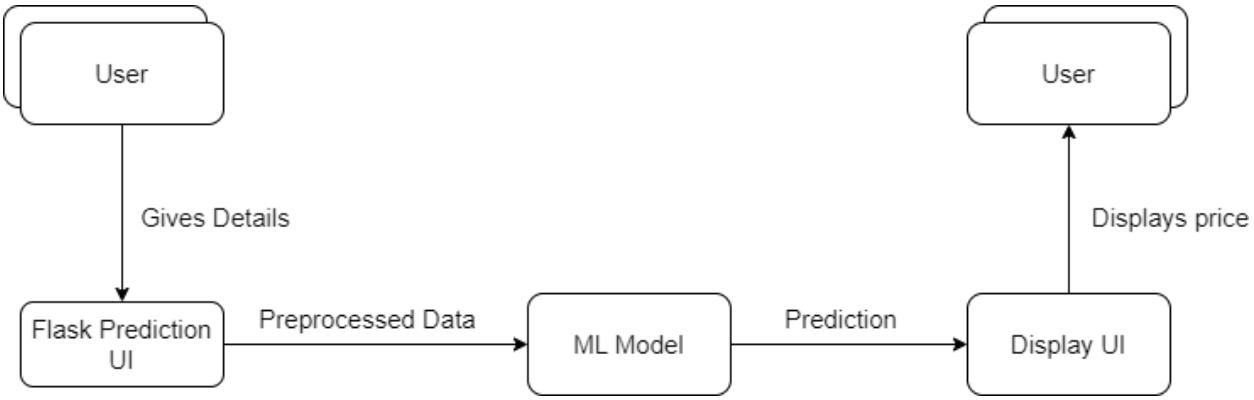
FR-4	System predicts car resale vale	Taking input features given by user, system should be able to predict car price by forwarding the prediction request to the ML model.
FR-5	Admin should be notified of any errors in the system	Any error that occurs like the model taking a long time to evaluate resale price should be notified to the admin so that the problem might be fixed.

4.2 Non-Functional requirements

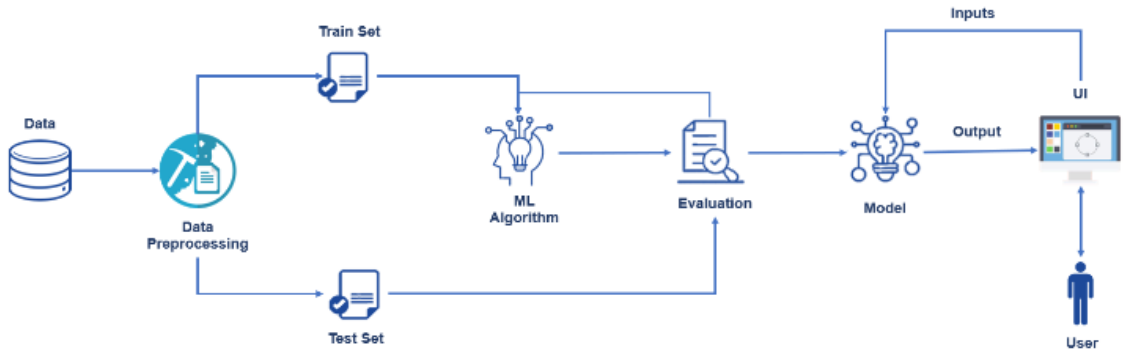
NFR-1	**Usability**	Effective User Interface with descriptions for each feature and proper layout that ensures each user finds it easy to access and interact with the system.
NFR-2	**Security**	Account creation for each user with a mandatory password strength check while creating the account.
NFR-3	**Reliability**	Chance of critical failure should be less than or equal to 2%.
NFR-4	**Performance**	The system must provide a webpage rendering images and texts upon receiving a request within a time of 8 seconds over a standard internet connection.
NFR-5	**Availability**	The website should be available to users 24x7. Any issues or errors will be addressed within the next 24 hours.
NFR-6	**Scalability**	The system must be scalable enough to support 1,00,000 requests at the same time without crashing.

5. PROJECT DESIGN

5.1 Data Flow Diagrams



5.2 Solution & Technical Architecture



5.3 User Stories

User Type	Functional Requirements	User Story Number	User Story	Acceptance Criteria	Priority	Release
Customer	Entering Car Details	USN-1	The user enters the necessary car details	All the mandatory fields are filled	High	Sprint 1
Customer	Viewing valuation	USN-2	The user's car's value is predicted by model	The value is displayed	High	Sprint 1
Admin	Updating model	USN-3	The admin can update the ML model after modifying it.	The ML model is properly loaded into the app.	High	Sprint 2

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.
Sprint-1	Login	USN-2	As a user, I can log into the application by entering email & password
Sprint-2	Model Development		Gather dataset and perform pre-processing and cleaning
Sprint-2			Training and testing of model using data at hand

Sprint-2			Integrating model in application
Sprint-3	Input for prediction	USN-3	As a user, I should be able to give inputs to the application that will be sent to the model for prediction
Sprint-3	Prediction	USN-4	As a user, I will be able to view the predicted valuation for given vehicle
Sprint-4	Previous predictions	USN-5	As a user, I will be able to view previous predictions made by the model
Sprint-4	Update model	USN-6	As an admin, I will be able to update the model according to the fluctuating environment

6.2 Sprint Delivery Schedule

Sprint	**Duration**	**Sprint Start Date**	**Sprint End Date (Planned)**
Sprint-1	6 Days	24 Oct 2022	29 Oct 2022
Sprint-2	6 Days	31 Oct 2022	05 Nov 2022
Sprint-3	6 Days	07 Nov 2022	12 Nov 2022
Sprint-4	6 Days	14 Nov 2022	19 Nov 2022

6.3 Reports from JIRA

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

7.1 Features added

7.1.1 Login,Register

Users can log into the application like a fully deployed website logging in or creating an account using the register feature.

7.1.2 Update

Users can update their password or account details.

7.2 Database Schema

Database made using inbuilt sqllite3 library available in python

User
id INTEGER PRIMARY KEY email TEXT NOT NULL username TEXT NOT NULL roll_number INTEGER NOT NULL pass_word TEXT NOT NULL

8. TESTING

8.1 Test Cases

TestcasesIBM.xlsx

8.2 User Acceptance Testing

1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the Car resale value prediction project at the time of the release to User Acceptance Testing (UAT).

2. Defect Analysis

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

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	10	4	2	3	20

Duplicate	0	0	0	0	0
External	1	0	0	0	1
Fixed	5	0	0	20	25
Not Reproduced	0	0	0	0	0
Skipped	2	0	0	0	2
Won't Fix	0	0	0	0	0
Totals	18	4	2	23	51

3. 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	10	0	0	10
Client Application	5	0	0	5

Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

9. RESULTS

9.1 Performance Metrics

RMSE: Root Mean Squared Error is the metric used by us to evaluate our model. It is most commonly used to evaluate regression models. Root Mean Square Error as the name suggests is calculated as the root of the mean squared errors of the predicted values. The formula is given below:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} \|y(i) - \hat{y}(i)\|^2}{N}},$$

Where zfi is the predicted value, zoi is the actual value and N is the sample size

We have used the sklearn.metrics package to calculate the RMSE.

Our ensembled model had a RMSE value of 3545.68 and the individual model error rates are given below:

Decision Tree: 4136.24

Random Forest: 3167.32

XGBoost: 3333.48

10. ADVANTAGES & DISADVANTAGES

Advantages:

- **The model is fairly accurate and is able to give a good prediction of what the actual resale value might be.**
- **The model is very quick in calculating the predictions.**
- **Errors of one model will be reduced by the ensembling with other models.**
- **It is easier for us to upload a better trained version of the model onto the cloud.**

Disadvantages:

- **The datasets available and the dataset the model has trained on do not give sufficient information to the model as it does not have very useful information.**
- **Attributes such as Fuel Mileage and Popularity of model has not been taken into account which can give a better idea about the resale value**

11. CONCLUSION

Thus with our project we can help both working resale dealers in the market and general users who want to sell their used cars by providing them with accurate resale value prediction.

12. FUTURE SCOPE

Can help with shaping the future of car resale market as our project opens up an insight into how the factors can be taken into for predicting the value of used cars.

13. APPENDIX

Source Code:

HTML FILES:

Home page : resalepredict.html

```
<!DOCTYPE html>
<html lang="en">

    <head>
        <meta charset="UTF-8">
        <meta name="viewport" content="width=device-width, initial-scale=1.0">
        <link rel="stylesheet" href="{{url_for('static',filename='css/resaleintro.css
        <title>Document</title>
    </head>

    <body>

        <div class="main">
            <div id="bg"></div>
            <div id="head">
                <h1>Car Resale Value</h1>
            </div>

            <form action="/predict" method="post">
                <div id="row">
                    <h3>Brand</h3>
                    <select name="brand" required="required">
                        <option value="" selected disabled hidden>choose</option>
                        <option value="audi">Audi</option>
                        <option value="bmw">BMW</option>
                        <option value="jaguar">Jaguar</option>
                        <option value="jeep">Jeep</option>
                        <option value="honda">Honda</option>
                        <option value="hyundai">Hyundai</option>
                        <option value="chevrolet">Chevrolet</option>
                        <option value="land_rover">Land_rover</option>
                        <option value="nissan">Nissan</option>
                        <option value="toyota">Toyota</option>
                        <option value="volvo">Volvo</option>
                        <option value="suzuki">Suzuki</option>
                    </select>
                </div>
                <div id="row">
                    <h3>Vehicle Type</h3>
                    <select name="vehicletype" required="required">
                        <option value="" selected disabled hidden>choose</option>
                        <option value="bus">Bus</option>
                        <option value="combination">Combination</option>
                        <option value="compact">Compact</option>
                        <option value="coupe">Coupe</option>
                        <option value="limousine">Limousine</option>
                        <option value="suv">SUV</option>
                    </select>
```

```
            </div>

            <div id="row">
                <h3>Year</h3>
                <input name="regyear" type="number" placeholder="eg: 2001" required>
            </div>
            <div id="row">
                <h3>Kilometers Drived</h3>
                <input name="kms" type="number" placeholder="eg: 9999" required="required">
            </div>

            <div id="row">
                <h3> Fuel type</h3>
                <select name="fuel" required="required">
                    <option value="" selected disabled hidden>choose</option>
                    <option value="petrol">Petrol</option>
                    <option value="diesel">Diesel</option>
                    <option value="cng">CNG</option>
                    <option value="lgp">LPG</option>
                    <option value="electric">Electric</option>
                </select>
            </div>

            <div id="row">
                <h3>Transmission Type</h3>
                <select name="gearbox" required="required">
                    <option value="" selected disabled hidden>choose</option>
                    <option value="manual">Manual</option>
                    <option value="automatic">Automatic</option>
                </select>
            </div>

            <div id="row">
                <h3>Engine (cc)</h3>
                <input name="powerps" type="number" placeholder="eg: 1000" required>
            </div>

            <div id="row">
                <h3>Damage</h3>
                <input name="damage" placeholder="Yes or No" required>
            </div>


            <button id="sub" type="submit ">Calculate the Selling Price</button>
        </form>
    </div>
</body>
</html>
```

Output.html

```html
<html>
    <head>
        <title>predicted value</title>
        <style>
            * {
                padding: 0;
                margin: 0;
            }

            body {
                height: 100vh;
                width: 100vw;
            }

            #bg {
                height: 100%;
                width: 100%;
                background-color: white;
                background-image: url('static/image/car.jpg');
                background-size: cover;
                opacity: 0.7;
            }

            h1 {
                font-size: 2rem;
                background-color: rgb(11, 218, 255);

                color: rgb(0, 0, 0);
                position: absolute;
                top: 80%;
                left: 30%;
                padding: 5px 20px;
                border-radius: 8px;
            }

        </style>

    </head>
    <body>
        <div id="bg"></div>
        <h1>{{ pred }}</h1>
    </body>
</html>
```

resaleredict.css

```css
* {
  padding: 0;
  margin: 0;
  font-family: Georgia;
}

.main {
  height: 100vh;
  width: 100vw;
  display: flex;
  justify-content: center;
  align-items: center;
  flex-direction: column;
}

.main #bg{
  width: 100%;
  height: 100vh;
  position:absolute;
  background-color: white;
  background-image: url("../image/car2.jpg");
  background-repeat: no-repeat;
  background-size: cover;
  opacity: 0.5;
}

.main #head {
  position: absolute;
  top: 0;
  height: 3rem;
  width: 100%;
  background-color: rgba(0, 0, 0, 0.6);

}

.main form {
  z-index: 100;
}
.main #head h1 {
  color: aqua;
  position: absolute;
  right: 0;
  padding: 7px 25px;
  font-size: 2rem;
}
.main h3 {
  width: 25rem;
}
```

```css
.main #row {
  display: flex;
  margin-top: 1.5rem;
}
.main #row input {
  margin-top: -7px;
  padding: 2px 10px;
  font-size: 15px;
  outline: none;
  border: 2px solid;
  height: 1.4rem;
  border-radius: 10px;
}
.main #row input:required {
  border: 2px solid rgb(0, 255, 128);
}
.main #row input:required:invalid {
  border: 2px solid rgb(255, 0, 242);
}
.main #row select {
  padding: 5px;
  border-radius: 10px;
  outline: none;
}
.main #row select option {
  background-color: rgb(77, 242, 248);
  color: rgb(0, 0, 0);
}
.main #row select:required {
  border: 2px solid rgb(0, 255, 128);
}
.main #row select:required:invalid {
  border: 2px solid rgb(225, 0, 255);
}
.main #sub {
  font-size: 1rem;
  border: 1px solid rgb(110, 110, 110);
  padding: 5px;
  position: relative;
  margin-top: 4rem;
  margin-left: 35%;
  border-radius: 5px;
  outline: none;
}
.main #sub:active {
  background-color: rgb(0, 17, 255);
  color: rgb(0, 0, 0);
}
```

Flask code
app.py

```python
import json
from flask import Flask, request, render_template
import pickle
import pickle

app=Flask(__name__)

model = pickle.load(open('model.pkl','rb'))

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

@app.route('/predict', methods=['GET', 'POST'])
def predict():
    if request.method == "POST":
        with open('mapping.json', 'r') as file:
            mapping = json.load(file)

        vehicle_type = mapping['vehicleType'][request.form.get('vehicletype')]
        years_old = 2022 - int(request.form['regyear'])
        gearbox = mapping['gearbox'][request.form.get('gearbox')]
        powerps = float(request.form['powerps'])
        kms = float(request.form['kms'])
        fuelType = mapping['fuelType'][request.form.get('fuel')]
        brand = mapping['brand'][request.form.get('brand')]
        damage = mapping['notRepairedDamage'][request.form.get('damage').lower()]

        data = [[vehicle_type, years_old, gearbox, powerps, kms, fuelType, brand, damage]]
        pred = model.predict(data)
        print(pred)
    return render_template('output.html', pred='The resale value predicted is ${:.2f}'.format(pred[0]))
```