# 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 biased valuation.

# 2.LITERATURE SURVEY:

With the recent arrival of internet portals, buyers and sellers may obtain an appropriate status of the factors that ascertain the market price of a used automobile. Lasso Regression, Multiple Regression, and Regression Trees are examples of machine learning algorithms. We will try to develop a statistical model that can forecast the value of a pre-owned automobile based on prior customer details and different parameters of the vehicle. [2] This paper aims to compare the efficiency of different models' predictions to find the appropriate one. On the subject of used automobile price prediction, several previous studies have been conducted. To anticipate the value of pre-owned automobiles in Mauritius, Pudaruth employed naive Bayes, k-nearest neighbours, multiple linear regression, and decision trees. However, because there were fewer carsobserved, their results were not good for prediction. In his article, Pudaruth concluded that decision trees and naive Bayes are ineffective for continuous-valued variables.[4] To anticipate the price of a vehicle, Noor and Jan employed multiple linear regression. They used a variable selection methodology to determine the variables that had the highest influence and then eliminated the remainder. Only a few variables are included in the data, which were utilised to create the linear regression model. With an R-square of 98 percent, the outcome was outstanding. [4] Peerun et al. conducted study to assess the neural network's performance in predicting used automobile prices. However, especially on higher-priced cars, the estimated value is not very close to the real price. In forecasting the price of a used car, they found that support vector machine regression outperformed neural networks and linear regression by a little margin. [4] To accurately anticipate the price of a car, many different approaches have been used in the digital world, ranging from machine learning approaches like multiple linear regression, k-nearest neighbor, and naive bayes to random forest and decision tree to the SAS enterprise miner. In [7], [8], [9], [10] and [11] all of these solutions took into account distinct sets of attributes when making predictions based on the historical data used to train the model. We attempted to construct a web application where a user may verify the effective market price of their automobiles using a model for prediction based on the factors that have the greatest impact on vehicle prices.

# 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 leasing application

[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.

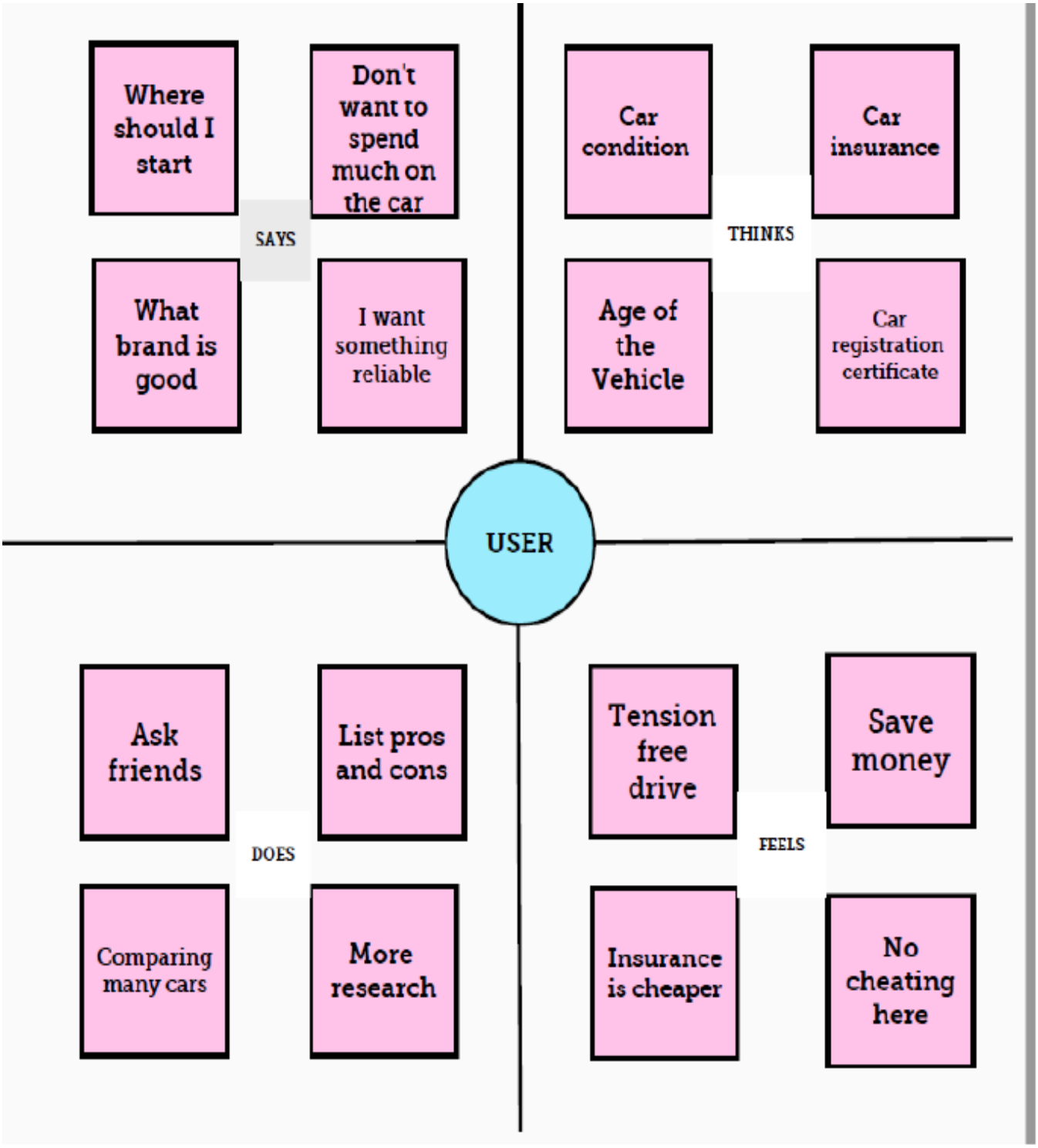
In order to predict the resale value of the car, we proposed an intelligent, flexible, and effective system that is based on using regression algorithms. Considering the main factors which would affect the resale value of a vehicle a regression model is to be built that would give the nearest resale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

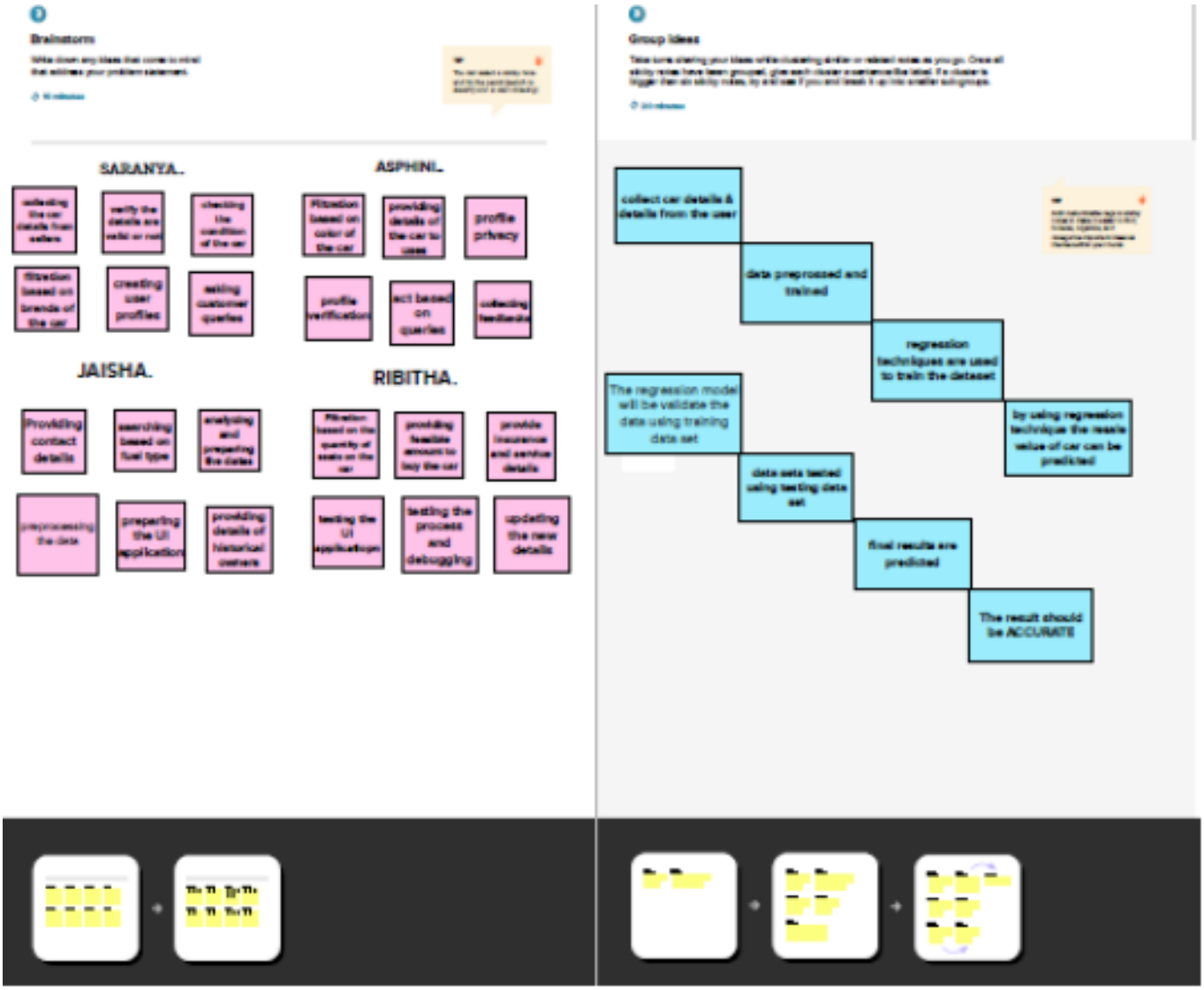
An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudesIt is a useful tool to helps teams better understand their users.

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.


Where should I start
Don't want to spend much on the car
SAYS
What brand is good
I want something reliable
Car condition
Car insurance
THINKS
Age of the Vehicle
Car registration certificate
USER
Ask friends
List pros and cons
DOES
Comparing many cars
More research
Tension free drive
Save money
FEELS
Insurance is cheaper
No cheating here

# 3.2. IDEATION & BRAINSTORMING:

Brainstorming provides a free and open environment that encourages everyone within a team to 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:

In view of the differing highlights and factors, and furthermore with the assistance of master information the vehicle value forecast has been done precisely. The most essential elements for forecast are brand and model, period use of vehicle, mileage of vehicle, gear type and fuel type utilized in the vehicle just as fuel utilization per mile profoundly influence cost of a vehicle because of continuous changes in the cost of a fuel. Various highlights like (discretionary) outside shading, entryway number, sort of transmission, measurements, security, cool, inside, if it has route will likewise impact the vehicle cost. In this, we applied distinctive methods(like relapse, grouping, bunching and so forth) and techniques(like regulated, solo, semi managed) to accomplish higher accuracy of the pre-owned car value expectation

A well-articulated customer problem statement allows you and your team to find the ideal solution for the challenges your customers face. Throughout the process, you'll also be able to empathize with your customers, which helps you better understand how they perceive your product or service.

| | | |
|---|---|---|
| **I am** | Describe customer with 3-4 key characteristics - *who are they?* | Describe the customer and their attributes here |
| **I'm trying to** | *List their outcome or "job" the care about - what are they trying to achieve?* | List the thing they are trying to achieve here |
| **but** | Describe what problems or barriers stand in the way – *what bothers them most?* | Describe the problems or barriers that get in the way here |
| **because** | Enter the "root cause" of why the problem or barrier exists – *what needs to be solved?* | Describe the reason the problems or barriers exist |
| **which makes me feel** | Describe the emotions from the customer's point of view – *how does it impact them emotionally?* | Describe the emotions the result from experiencing the problems or barriers |

## 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.

Define CS, fit into CC

## 1. CUSTOMER SEGMENT(S)

CS

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

- Dealeis
- Avid Buyeis ovei the age of 18

## 6. CUSTOMER CONSTRAINTS

CC

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

Customeis aie hesitant due to stigma of computei piedicted values might not be accuiate.

## 5. AVAILABLE SOLUTIONS

AS

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

Visit online websites to see how much othei people with similai cais aie selling theii cais foi.

By visiting dealeiships and getting estimates.

Explore AS, differentiate

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

## 2. JOBS-TO-BE-DONE / PROBLEMS

J&P

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

1'o build a supeivised machine leaining model that utilizes iegiession methods to accuiately piedict/anticipate the value of a Used cai based on the following factois:

- Condition of the cai
- Kilometeis diiven
- Life Span
- Damages
- No. of owneis

## 9. PROBLEM ROOT CAUSE

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

1'he value pioposed by dealeis and othei paities foi a cai may be untiustwoithy and extiemely low.

Useis aie unsuie how much theii can actually sell foi oi at a piice which theycan bid foi.

## 7. BEHAVIOUR

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

Pioviding false claims on damages inand on the cai.

1'o oveisell non-existent featuies.

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

Identify strong TR & EM

## 3. I'RIGGERS

TR

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

Useis may othei sites to make a compaiison whichcateis the decision piocess.

## 4. EMOTIONS: BEIORE / AIIER

EM

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

Befoie:
1'he usei might be conceined about the inaccuiate piediction based on human assessment.
Afiei:
without usei inteivention, the usei may decide the attiibutes of the cai on theii own

## 10. YOUR SOLUI'ION

SL

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

A machine leaining model can be utilized to develop this system which can accuiately piedict the iesale value of the cai given a set of attiibutes of the cai.

## 8. CHANNELS of BEHAVIOUR

CH

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

8.2 OIILINE
What kind of actions do customeis take offline? Extract offline channels from 7and usetr them foi customer development.

Online:
Customeis don't just look at the infoimation piovided by cai biand websites but they also makea compaiison study on piicings on vaiious websites.

Offline:
If an usei is inteiested in buying a cai. 1'hey would visit a lot of dealeiships to get a quotation and do a compaiison study.

Identify strong TR & EM

# 4. REQUIREMENT ANALYSIS:

## 4.1.FUNCTIONAL REQUIREMENT:

**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<br>Registration through Gmail<br>Registration through Application |
| FR-2 | User Confirmation | Confirmation via Email<br>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 |

## 4.2.NON-FUNCTIONAL REQUIREMENTS:

**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<br>Simple and easy to Understand |
| NFR-2 | **Security** | Aware of scams |
| NFR-3 | **Reliability** | The system must perform without failure |
| NFR-4 | **Performance** | The landing page must support several users must provide 5 second or less response time |
| NFR-5 | **Availability** | Uninterrupted services must be available all time except the time of server updation. |
| NFR-6 | **Scalability** | That can handle any amount of data and perform many computations in a cost-effective and time-saving way to instantly serve millions of users residing at global locations. |

# 5. PROJECT DESIGN:

# 5.1 DATA FLOW DIAGRAMS:

A data flow diagram (DFD) maps out the flow of information for any process or system. It uses defined symbols like rectangles, circles and arrows, plus short text labels, to show data inputs, outputs, storage points and the routes between each destination. Data flowcharts can range from simple, even hand-drawn process overviews, to in-depth, multi-level DFDs that dig progressively deeper into how the data is handled. They can be used to analyze an existing system or model a new one. Like all the best diagrams and charts, a DFD can often visually "say" things that would be hard to explain in words, and they work for both technical and nontechnical audiences, from developer to CEO. That's why DFDs remain so popular after all these years. While they work well for data flow software and systems, they are less applicable nowadays to visualizing interactive, real-time or database-oriented software or systems.

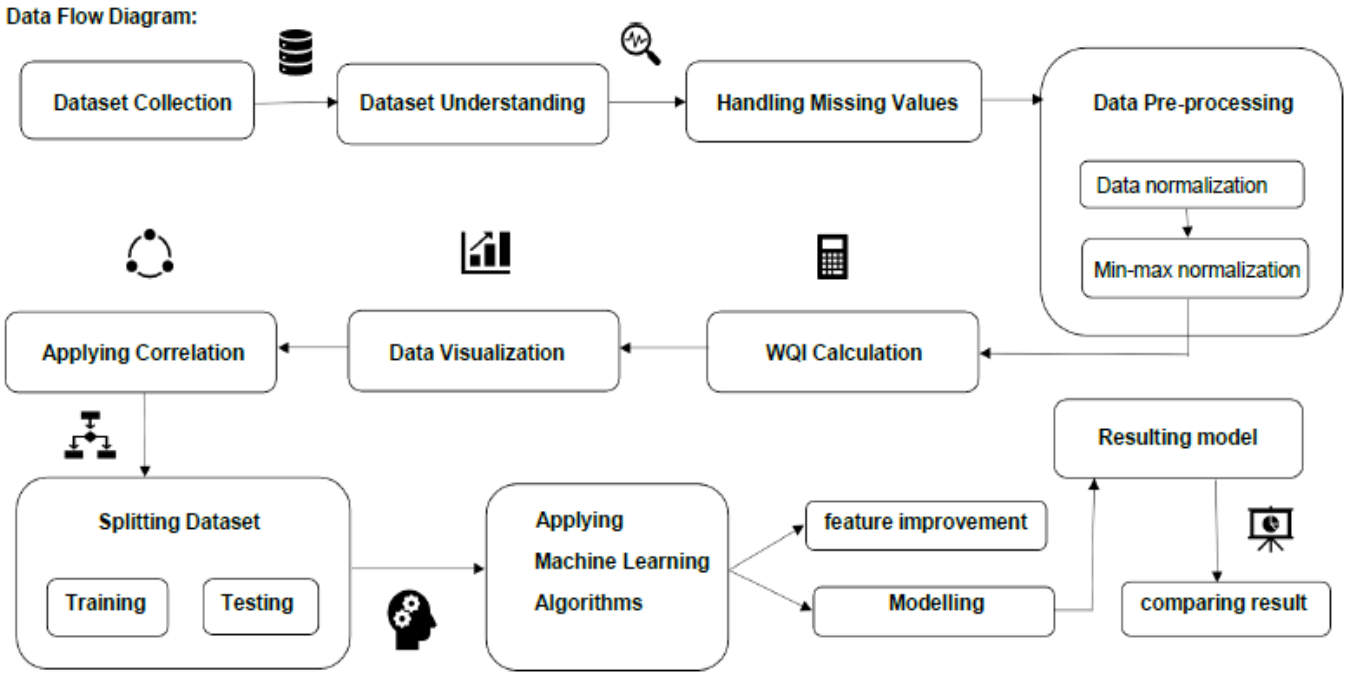
Data flow diagrams were popularized in the late 1970s, arising from the book Structured Design, by computing pioneers Ed Yourdon and Larry Constantine. They based it on the "data flow graph" computation models by David Martin and Gerald Estrin. The structured design concept took off in the software engineering field, and the DFD method took off with it. It became more popular in business circles, as it was applied to business analysis, than in academic circles.

Also contributing were two related concepts:

Object Oriented Analysis and Design (OOAD), put forth by Yourdon and Peter Coad to analyze and design an application or system.

Structured Systems Analysis and Design Method (SSADM), a waterfall method to analyze and design information systems. This rigorous documentation approach contrasts with modern agile approaches such as

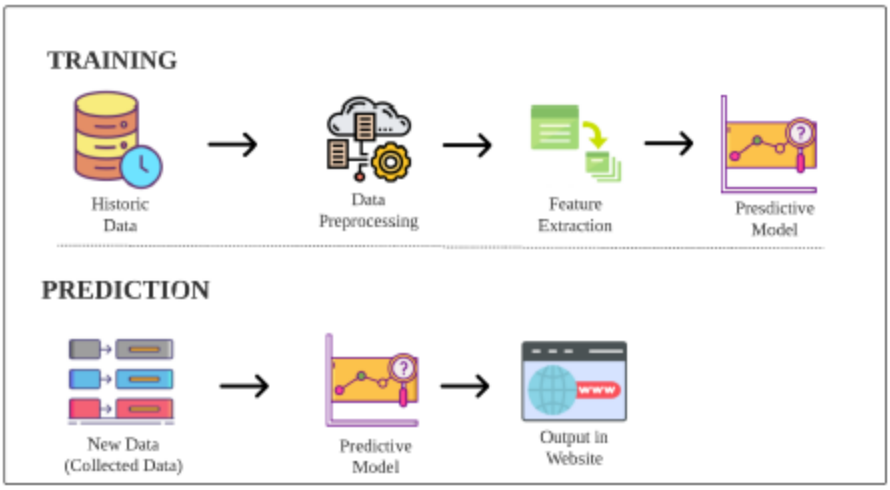
Scrum and Dynamic Systems Development Method (DSDM.)
Three other experts contributing to this rise in DFD methodology were Tom DeMarco, Chris Gane and Trish Sarson. They teamed up in different combinations to be the main definers of the symbols and notations used for a data flow diagram.




# 5.2. SOLUTION & TECHNICAL ARCHIECTURE:

## solution architecture:

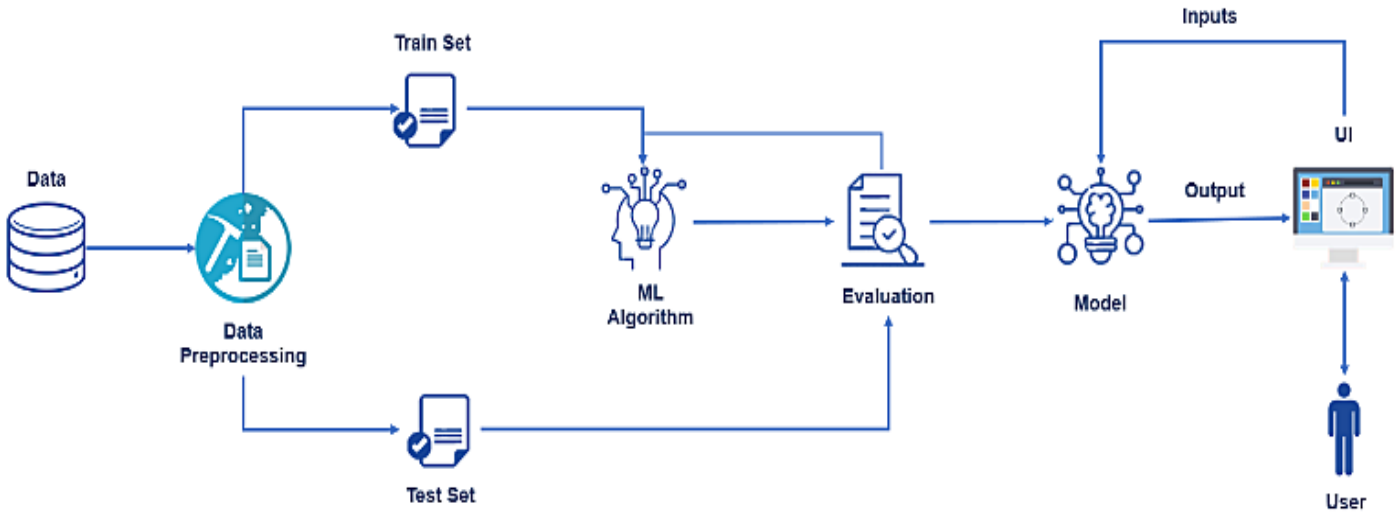
Solution architecture is the process of developing solutions based on predefined processes, guidelines and best practices with the objective that the developed solution fits within the enterprise architecture in terms of information architecture, system portfolios, integration requirements and many more.
It can then be viewed as a combination of roles, processes and documentation that are intended to address specific business needs, requirements or problems through the design and development of applications and information systems.Solution architecture is the initial step taken when an organization aims to create a set of enterprise solutions, applications and processes that integrate with each other in order to address specific needs and requirements and that often lead to software architecture and technical architecture work.The solution architecture is described in a document that specifies a certain level of vision for all current and future solutions, applications and processes that the organization has. Design and development of solutions and applications then follow the guidelines specified in the solution architecture document to ensure that they conform to set standards that make integration and communication easier, and make the tracking of problems and inconsistencies between solutions easier as well.




## Technical architecture:

Technical architecture—which is also often referred to as application architecture, IT architecture, business architecture, etc.—refers to creating a structured software solution that will meet the business needs and expectations while providing a strong technical plan for the growth of the software application through its lifetime. IT architecture is equally important to the business team and the information technology team.

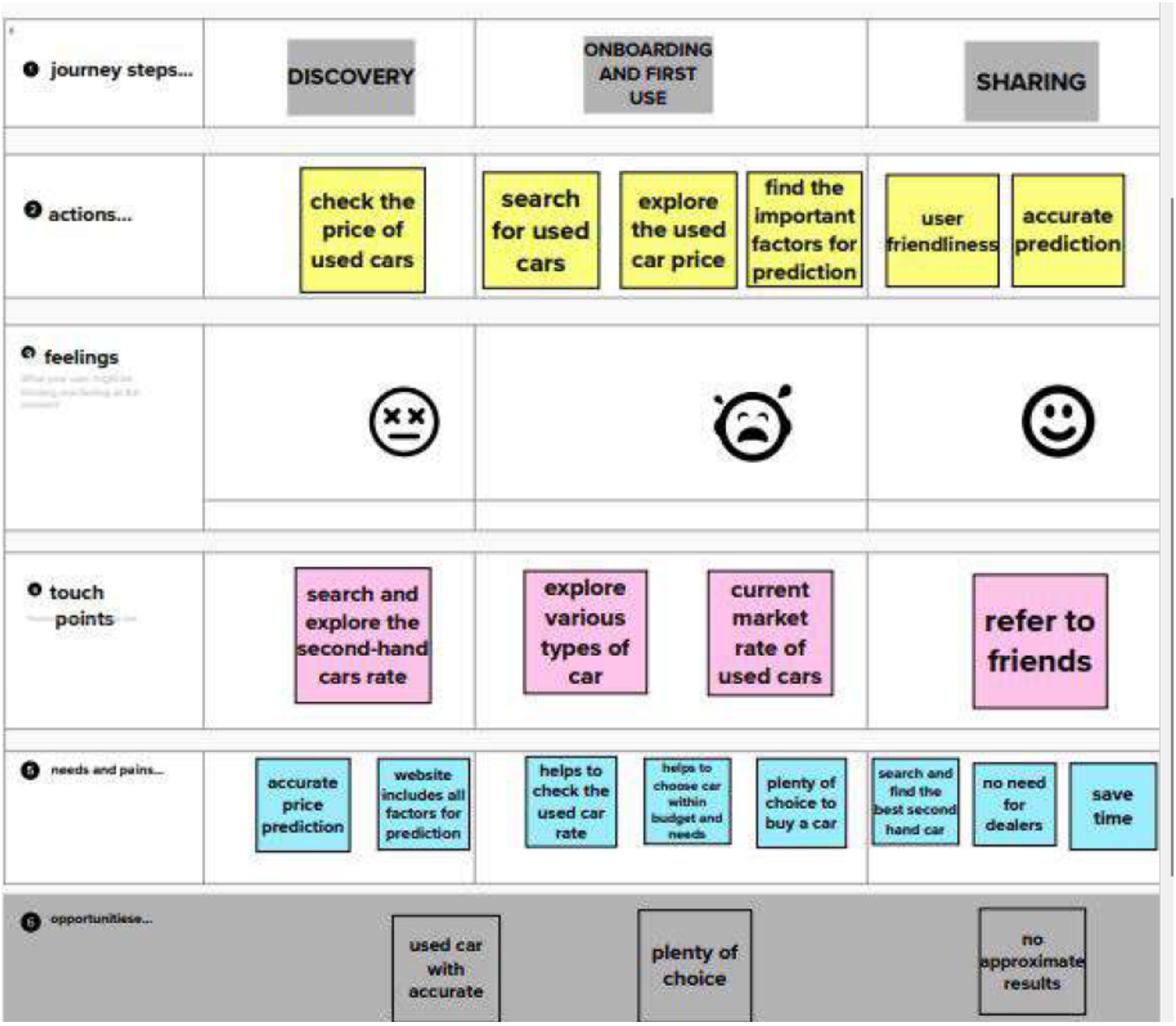
Technical architecture includes the major components of the system, their relationships, and the contracts that define the interactions between the components. The goal of technical architects is to achieve all the business needs with an application that is optimized for both performance and security.



## 5.3. USER STORY:

A user story is an informal, general explanation of a software feature written from the perspective of the end user. Its purpose is to articulate how a software feature will provide value to the customer. A user story is the smallest unit of work in an agile framework. It's an end goal, not a feature, expressed from the software user's perspective.

A user story is an informal, general explanation of a software feature written from the perspective of the end user or customer. The purpose of a user story is to articulate how a piece of work will deliver a particular value back to the customer. The user story for the project is as follows:



# 6.PROJECT PLANNING & SCHEDULING:

## 6.1. Sprint Planning and Estimation:

Sprint planning is an event in scrum that kicks off the sprint. The purpose of sprint planning is to define what can be delivered in the sprint and how that work will be achieved. Sprint planning is done in collaboration with the whole scrum team. In scrum, the sprint is a set period of time where all the work is done. However, before you can leap into action you have to set up the sprint. You need to decide on how long the time box is going to be, the sprint goal, and where you're going to start. The sprint planning session kicks off the sprint by setting the agenda and focus. If done correctly, it also creates an environment where the team is motivated, challenged, and can be successful. Bad sprint plans can derail the team by setting unrealistic expectations. The following is the sprint planning and estimation for the project.

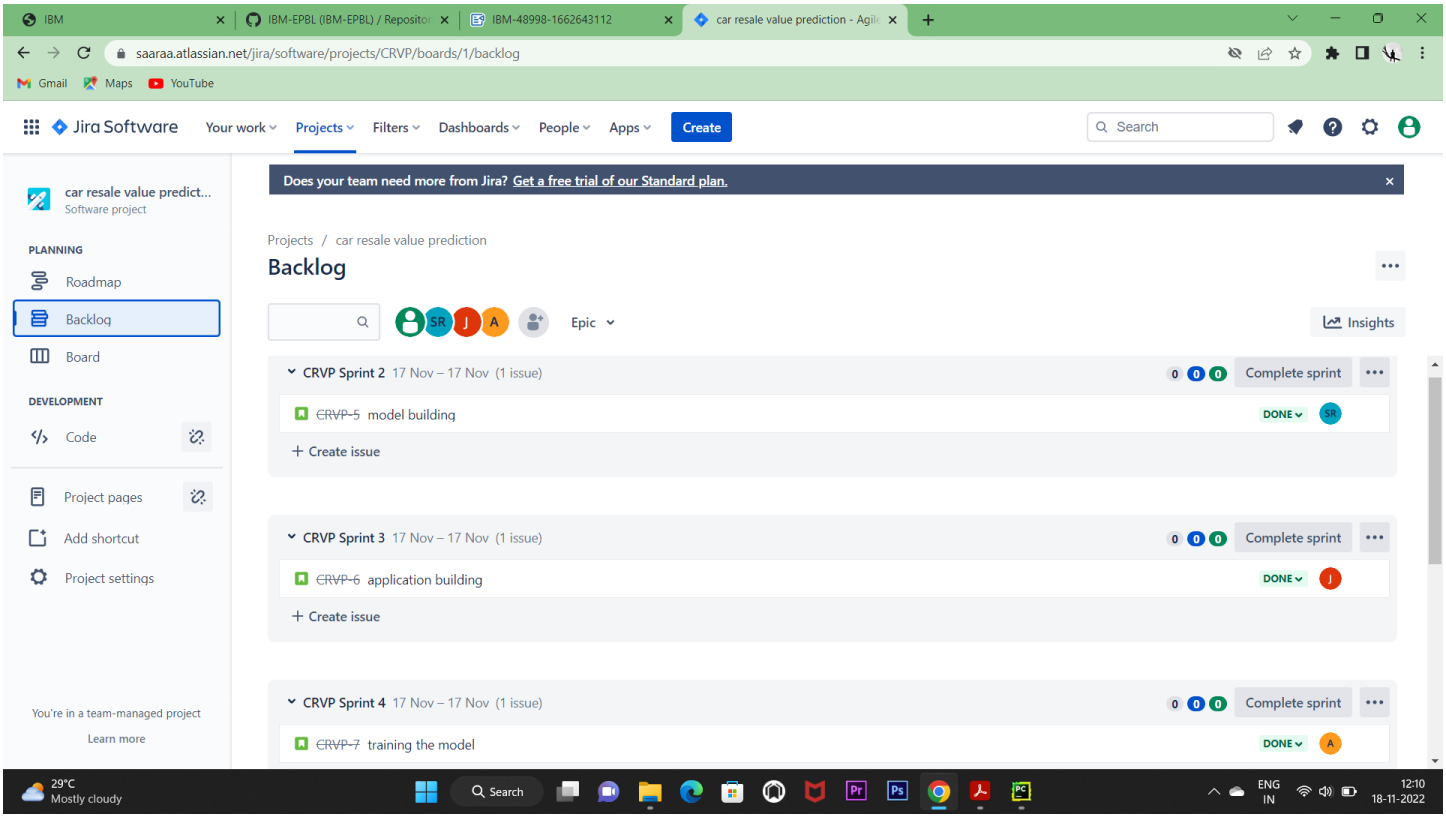
| Sprint | Functional Requirement (Epic) | User Story Number | User Story / Task | Story Points | Priority |
|---|---|---|---|---|---|
| Sprint-1 | Pre-process data | USN-1 | Collect Dataset | 1 | Low |
| Sprint-1 | | USN-2 | Import required libraries | 1 | Low |
| Sprint-1 | | USN-3 | Read and clean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

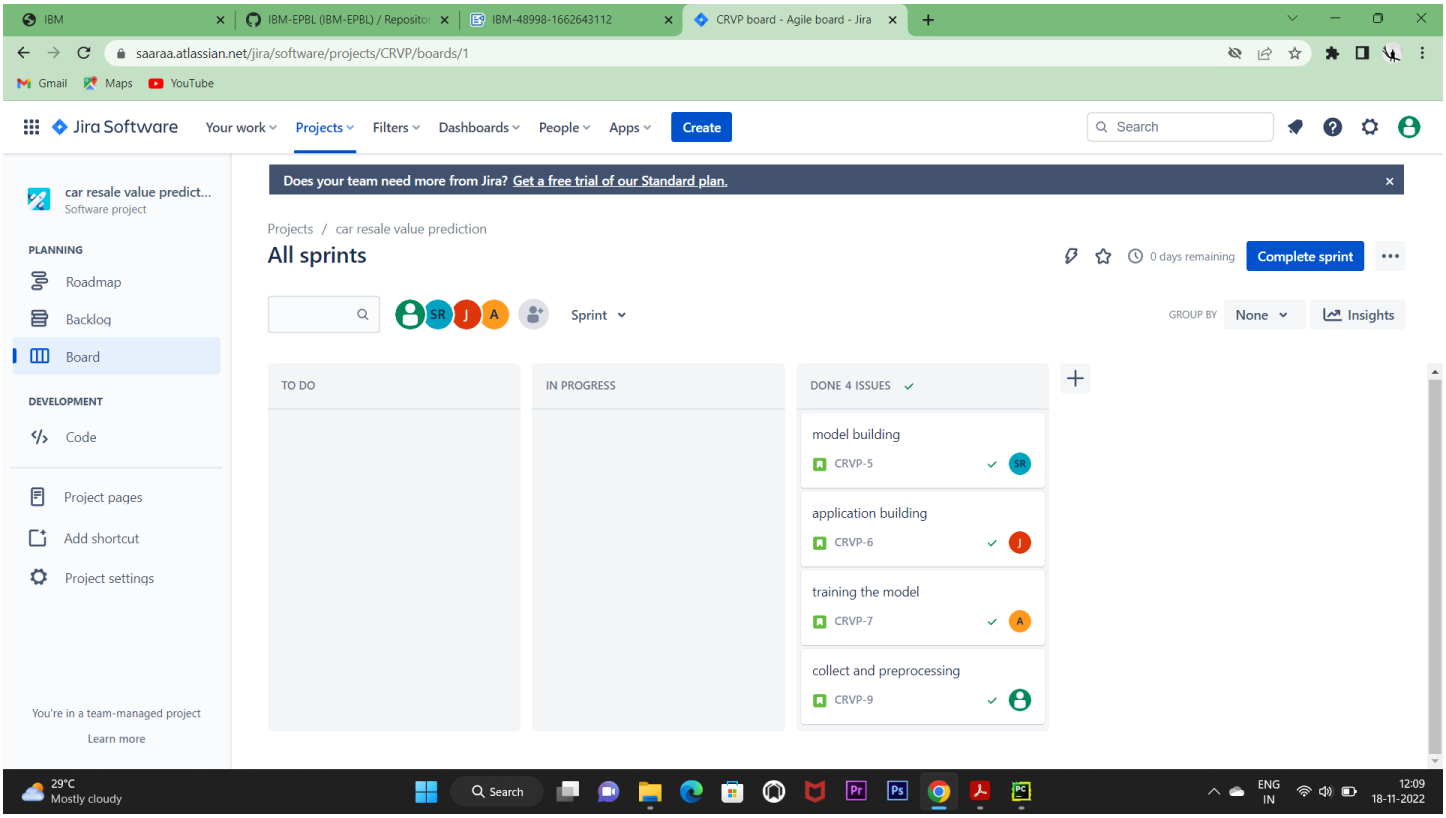
## 6.2 Sprint Delivery Schedule:

Since sprints take place over a fixed period of time, it's critical to avoid wasting time during planning and development. And this is precisely where sprint scheduling enters the equation.In case you're unfamiliar, a sprint schedule is a document that outlines sprint planning from end to end. It's one of the first steps in the agile sprint planning process—and something that requires adequate research, planning, and communication.

| 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 | 6 Days | 24 Oct 2022 | 29 Oct 2022 | 20 | 29 Oct 2022 |
| Sprint-2 | 20 | 6 Days | 31 Oct 2022 | 05 Nov 2022 | 20 | 05 Nov 2022 |
| Sprint-3 | 20 | 6 Days | 07 Nov 2022 | 12 Nov 2022 | 20 | 12 Nov 2022 |
| Sprint-4 | 20 | 6 Days | 14 Nov 2022 | 19 Nov 2022 | 20 | 19 Nov 2022 |

# 6.3. Reports from JIRA:



# 7. CODING & SOLUTIONING:

## 7.1. 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

## 7.2. 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).

# 1. 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 | 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 |

## 2.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:

# 9.1. 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:

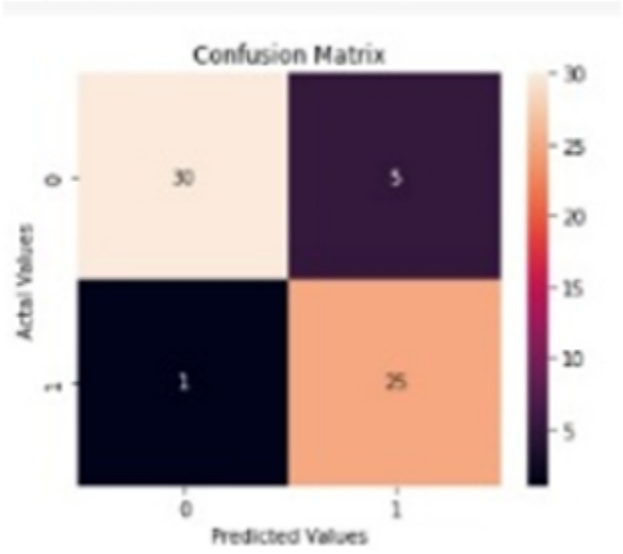
## Regression Model:

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 the dataset.

```
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 the resale 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 inRandom Forest, different selections of features and number of trees will be tested to check for change in performance.

# 13. APPENDIX:

## Source code

## app.py

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

app = Flask(__name__)
filename = 'resale_model (1).sav'
model_rand = pickle.load(open(filename, 'rb'))


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


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


@app.route('/predict')
def predict():
    return render_template('booking.html')


@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 = request.form.get('vehicletype')
    new_row = {'yearOfRegistration': regyear, 'powerPS': powerps, 'kilometer': kms,
'monthOfRegistration': regmonth,
               '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

```
<!DOCTYPE html>
<html>
    <head>
        <style>



            }
        </style>
    </head>
    <body>
        <h1>Car Resale Value Prediction</h1>
```

```
        <h2>welcome</h2>
        <form action='predict'>
            <p>Click here to predict the price</p>
            <input type="submit" value="Click here">
         </form>

    </body>
</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;
         }

         body {
            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";>Predict the Price!</h1>
            </center>

            <center>
               <h3 style="font-family:verdana; color: white;">{{ypred}}</h3>
            </center>
            <div class="row">
               <h4>Registration Details</h4>
               <div class="input-group"><input type="number" name="regyear" id="regyear"
                     placeholder="Registration Year" required/>
               </div>
               <div>
                  <h4>Registration Month</h4>
                  <!-- <div class="input-group"><input type="number" name="regmonth"
id="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>
                     </select>
                  </div>

                  </div>

            </div>
               <h4>Power PS</h4>
               <div class="input-group"><input type="number" name="powerps"
placeholder="Power of the Car in PS" required/>
               </div>
               <h4>Kilometers Driven</h4>
               <div class="input-group"><input type="number" name="kms"
placeholder="Kilometers the car has driven" required/>
               </div>

               <div>
                  <h4>Gear Box Type</h4>
                  <div class="input-group">
                     <input id="gear-manual" type="radio" name="gearbox" value="manual" />
                     <label for="gear-manual">Manual</label>
                     <input id="gear-automatic" type="radio" name="gearbox"
value="automatic" />
                     <label for="gear-automatic">Automatic</label>
                     <input id="gear-notdeclared" type="radio" name="gearbox" value="not
declared" />
                     <label for="gear-notdeclared">Not declared</label>

                  </div>
               </div>

               <div>
                  <h4>Is Car Damaged</h4>
                  <div class="input-group">
                     <input id="yes" type="radio" name="damaged" value="Yes" />
                     <label for="yes">Yes</label>
                     <input id="damaged-no" type="radio" name="damaged" value="No" />
                     <label for="damaged-no">No</label>
                     <input id="damaged-notdeclared" type="radio" name="damaged" value="Not
declared" />
                     <label for="damaged-notdeclared">Not declared</label>
```

```
            </div>
        </div>

        <div>
            <h4>Fuel Type</h4>
            <div class="input-group">
                <select name="fuel" id="fuel">
                    <option value="not-declared">not-declared</option>
                    <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 Car</h4>
            <div class="input-group">
                <select name="brand" id="brand">
                    <option value="audi">audi</option>
                    <option value="jeep">jeep</option>
                    <option value="Mahindra">Mahindra</option>
                    <option value="volkswagen">volkswagen</option>
                    <option value="skoda">skoda</option>
                    <option value="bmw">bmw</option>
                    <option value="peugeot">peugeot</option>
                    <option value="ford">ford</option>
                    <option value="mazda">mazda</option>
                    <option value="nissan">nissan</option>
                    <option value="renault">renault</option>
                    <option value="mercedes_benz">mercedes_benz</option>
                    <option value="honda">honda</option>
                    <option value="fiat">fiat</option>
                    <option value="opel">opel</option>
                    <option value="mini">mini</option>
                    <option value="smart">smart</option>
                    <option value="hyundai">hyundai</option>
                    <option value="alfa_romeo">alfa_romeo</option>
                    <option value="subaru">subaru</option>
                    <option value="volvo">volvo</option>
                    <option value="mitsubishi">mitsubishi</option>
                    <option value="kia">kia</option>
```

```
                <option value="seat">seat</option>
                <option value="lancia">lancia</option>
                <option value="porsche">porsche</option>
                <option value="citroen">citroen</option>
                <option value="toyota">toyota</option>
                <option value="chevrolet">chevrolet</option>
                <option value="dacia">dacia</option>
                <option value="suzuki">suzuki</option>
                <option value="daihatsu">daihatsu</option>
                <option value="chrysler">chrysler</option>
                <option value="sonstige_autos">sonstige_autos</option>
                <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</option>
                <option value="lada">lada</option>
                <option value="trabant">trabant</option>

            </select>
        </div>
    </div>


    <div>
        <h4>Model Type</h4>
        <div class="input-group">
            <select name="model_type" id="model_type">
                <option value="not-declared">not-declared</option>
                <option value="grand">grand</option>
                <option value="scorpio s11">scorpio s11</option>
                <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</option>
                <option value="scirocco">scirocco</option>
                <option value="5er">5er</option>
                <option value="andere">andere</option>
                <option value="civic">civic</option>
                <option value="punto">punto</option>
```

```
                        <option value="e_klasse">e_klasse</option>
                        <option value="clio">clio</option>
                        <option value="kadett">kadett</option>
                        <option value="one">one</option>
                        <option value="fortwo">fortwo</option>
                        <option value="1er">1er</option>
                        <option value="b_klasse">b_klasse</option>
                        <option value="a8">a8</option>
                        <option value="jetta">jetta</option>
                        <option value="c_klasse">c_klasse</option>
                        <option value="micra">micra</option>
                        <option value="vito">vito</option>
                        <option value="sprinter">sprinter</option>
                        <option value="astra">astra</option>
                        <option value="156">156</option>
                        <option value="escort">escort</option>
                        <option value="forester">forester</option>
                        <option value="xc_reihe">xc_reihe</option>
                        <option value="fiesta">fiesta</option>
                        <option value="scenic">scenic</option>
                        <option value="ka">ka</option>
                        <option value="a1">a1</option>
                        <option value="transporter">transporter</option>
                        <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</option>
                        <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="caddy">caddy</option>
                        <option value="mondeo">mondeo</option>
                        <option value="cordoba">cordoba</option>
                        <option value="colt">colt</option>
                        <option value="impreza">impreza</option>
                        <option value="vectra">vectra</option>
                        <option value="lupo">lupo</option>
                        <option value="berlingo">berlingo</option>
                        <option value="m_klasse">m_klasse</option>
                        <option value="tiguan">tiguan</option>
                        <option value="6_reihe">6_reihe</option>
                        <option value="c4">c4</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</option>
                        <option value="6er">6er</option>
                        <option value="modus">modus</option>
                        <option value="fox">fox</option>
                        <option value="matiz">matiz</option>
                        <option value="beetle">beetle</option>
                        <option value="rio">rio</option>
                        <option value="touareg">touareg</option>
                        <option value="logan">logan</option>
                        <option value="spider">spider</option>
                        <option value="cuore">cuore</option>
                        <option value="s_max">s_max</option>
                        <option value="a2">a2</option>
                        <option value="x_reihe">x_reihe</option>
                        <option value="a5">a5</option>
                        <option value="galaxy">galaxy</option>
                        <option value="c3">c3</option>
                        <option value="viano">viano</option>
                        <option value="s_klasse">s_klasse</option>
                        <option value="1_reihe">1_reihe</option>
                        <option value="sharan">sharan</option>
                        <option value="avensis">avensis</option>
                        <option value="sl">sl</option>
                        <option value="roomster">roomster</option>
```

```
                        <option value="q5">q5</option>
                        <option value="santa">santa</option>
                        <option value="leon">leon</option>
                        <option value="cooper">cooper</option>
                        <option value="4_reihe">4_reihe</option>
                        <option value="sportage">sportage</option>
                        <option value="laguna">laguna</option>
                        <option value="ptcruiser">ptcruiser</option>
                        <option value="clk">clk</option>
                        <option value="primera">primera</option>
                        <option value="espace">espace</option>
                        <option value="exeo">exeo</option>
                        <option value="159">159</option>
                        <option value="transit">transit</option>
                        <option value="juke">juke</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</option>
                        <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</option>
                        <option value="agila">agila</option>
                        <option value="duster">duster</option>
                        <option value="v50">v50</option>
                        <option value="mx_reihe">mx_reihe</option>
                        <option value="meriva">meriva</option>
                        <option value="discovery">discovery</option>
```

```
                        <option value="c_reihe">c_reihe</option>
                        <option value="v_klasse">v_klasse</option>
                        <option value="yaris">yaris</option>
                        <option value="c5">c5</option>
                        <option value="aygo">aygo</option>
                        <option value="seicento">seicento</option>
                        <option value="cc">cc</option>
                        <option value="carnival">carnival</option>
                        <option value="fusion">fusion</option>
                        <option value="bora">bora</option>
                        <option value="cl">cl</option>
                        <option value="tigra">tigra</option>
                        <option value="300c">300c</option>
                        <option value="500">500</option>
                        <option value="100">100</option>
                        <option value="q3">q3</option>
                        <option value="cr_reihe">cr_reihe</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>
                        <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</option>
                        <option value="cayenne">cayenne</option>
                        <option value="wrangler">wrangler</option>
                        <option value="lybra">lybra</option>
                        <option value="range_rover_sport">range_rover_sport</option>
                        <option value="lancer">lancer</option>
                        <option value="freelander">freelander</option>
                        <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</option>
                    <option value="cherokee">cherokee</option>
                    <option value="clubman">clubman</option>
                    <option value="arosa">arosa</option>
                    <option value="legacy">legacy</option>
                    <option value="pajero">pajero</option>
                    <option value="auris">auris</option>
                    <option value="c2">c2</option>
                    <option value="niva">niva</option>
                    <option value="s60">s60</option>
                    <option value="nubira">nubira</option>
                    <option value="vivaro">vivaro</option>
                    <option value="g_klasse">g_klasse</option>
                    <option value="lodgy">lodgy</option>
                    <option value="850">850</option>
                    <option value="serie_2">serie_2</option>
                    <option value="charade">charade</option>
                    <option value="croma">croma</option>
                    <option value="outlander">outlander</option>
                    <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</option>
                    <option value="move">move</option>
                    <option value="materia">materia</option>
                    <option value="terios">terios</option>
                    <option value="kalina">kalina</option>
                    <option value="elefantino">elefantino</option>
                    <option value="i3">i3</option>
                    <option value="samara">samara</option>
                    <option value="kappa">kappa</option>
                    <option value="serie_3">serie_3</option>
                    <option value="discovery_sport">discovery_sport</option>
                </select>
            </div>
        </div>
```

```
            <div>
               <h4>Vehicle Type</h4>
               <div class="custom">
                  <select name="vehicletype" id="wgtmsr">
                     <option value="coupe">coupe</option>
                     <option value="suv">suv</option>
                     <option value="small car">small car</option>
                     <option value="limousine">limousine</option>
                     <option value="convertible">convertible</option>
                     <option value="bus">bus</option>
                     <option value="combination">combination</option>
                     <option value="not-declared">not-declared</option>
                     <option value="others">others</option>
                  </select>
               </div>
            </div>


            <div class="row">
               <center><button class="button button2" href="/"
type="submit">Predict</button></center>
            </div>

            <br>
            <center> <a href="/home">Click here to go back Home!</a></center>
         </form>

      </div>
   </section>
   </html>
```

## Github link

https://github.com/IBM-EPBL/IBM-Project-48998-1660814897

## Project demo link

https://drive.google.com/file/d/1xYc5KzijlZCITJtvx_sspC1XAhfpCybB/view?usp=drivesdk