



# PREDICTING CAR PRICES USING MACHINE LEARNING



**Session 2019-2023**

**Project Advisors  
Dr. Muhammad  
Saeed**

**Submitted By**

M. SAMEER SIDDIQUI	BS Seat # B-19102092
MIRZA VIRAAD BAIG	BS Seat # B-19102055
UMAIR IQBAL	BS Seat # B-19102103
MUHAMMAD USMAN	BS Seat # B-19102105

Department of Computer Science  
University of Karachi  
Karachi

BS

Price Prediction using ML

Jan 10, 2024

# **STATEMENT OF SUBMISSION**

A report submitted to the  
Department of Computer Science  
in partial fulfillment of the requirements for the degree  
Bachelor of Science in  
Computer Science

By

Muhammad Sameer Siddiqui  
Mirza Viraad Baig  
Umair Iqbal  
Muhammad Usman

University of Karachi  
January 2024

## **Acknowledgements**

We truly acknowledge the cooperation and help make by Dr. Muhammad Saeed, He has been a constant source of guidance throughout the course of this project. We are also thankful to our friends and families whose silent support led us to complete our project.

(Signed)

Muhammad Sameer Siddiqui B-19102092

Mirza Viraad Baig B-19102055

Umair Iqbal B-19102103

Muhammad Usman B-19102105

Date

January 10, 2024

## **Abstract**

The necessity of personal vehicles for daily commuting and leisure travel is well understood. Whether opting for a new or used vehicle, the decision-making process involved in this purchase often consumes a significant amount of time. Moreover, selling an existing vehicle presents its own challenges, determining the optimal price and method for a smooth transaction.

Imagine a scenario where the process of selling a car doesn't necessitate intermediary intervention. Instead, envision a system capable of estimating the selling price of a car, leveraging a repository of historical selling prices. This eliminates the need for a middleman, offering a direct interface for sellers and buyers alike. Our proposed solution employs machine learning (ML) models to predict and analyze car selling prices, streamlining the vehicle selling and purchasing experience.

The resulting web application offers users an estimated selling price for their car based on factors such as fuel type, years of service, mileage, and transmission type (manual/automatic). Additionally, the platform features a review section where users can share their experiences with owned cars, fostering a community-driven exchange of insights. Furthermore, an analytics section provides comprehensive market trend analysis, offering valuable insights to both buyers and sellers.

In essence, our deployed ML-powered web application serves as a valuable tool, benefiting users by simplifying the selling process, empowering informed decision-making, and providing a reliable estimate of a car's market value.

# Table of Contents

STATEMENT OF SUBMISSION .....	i
Acknowledgements .....	ii
Abstract.....	iii
Table Of Contents .....	iv
Chapter 1 Introduction.....	1
1.1 Introduction.....	2
Chapter 2 Problem Statement .....	3
2.1 Problem Statement .....	4
Chapter 3 Background Research Done .....	5
3.1 Background research done .....	6
Chapter 4 Final Problem Statement .....	7
4.1 Final Problem Statement .....	8
Chapter 5 Requirements and Analysis on this Problem Statement.....	9
5.1 Requirements .....	10
5.1.1 Data Sets .....	10
5.1.2 Programming Requirements.....	10
5.2 Analysis .....	11
Chapter 6 Prototype .....	12
6.1 Design of Prototype .....	13
6.1.0 Login.....	13
6.1.1 Home Page .....	13
6.1.2 Used Car .....	15
6.1.3 New Car .....	16
6.1.4 Imported Car .....	17
6.1.5 Reviews .....	17
6.1.6 Analytics.....	17
6.2 Implementation of Prototype .....	23
6.2.1 Used Car Model .....	23
6.2.2 New Car Model.....	26
6.2.3 Imported Car Model.....	28
6.2.4 Backend Code.....	31
Chapter 7 Results Of The Prototype.....	35

7.1 Results Of Prototype.....	36
7.1.1 Used Car Model Results.....	36
7.1.2 Imported Car Model Results .....	37
7.1.3 New Car Models Results .....	38
Chapter 8 Conclusion of Research.....	40
8.1 Conclusion.....	41
Chapter 9 Future Work in this Field .....	42
9.1 Future Work.....	43
Chapter 10 References .....	44
10.1 Reference .....	45

# **Chapter 1**

## **Introduction**

## 1.1 Introduction

DRIVE AI Is a web app which is designed for the purpose of predicting the value of used cars and new cars. Vehicle price prediction especially when the vehicle is used and not coming direct from the factory, is both a critical and important task. With increase in demand for used cars and up to 10 percent decrease in demand for the new cars in 2020 Due to Covid outbreak more and more vehicle buyers are finding alternatives of buying new cars outright. So by keeping in mind the latest trends of the market and the government policy of not providing any more financing for the buying of imported cars which are more than 1000CC ,and heavy duty taxes on import so the people are now moving towards the purchasing of used cars and whenever someone try to buy or sell used cars the main issue that come into existence is that what will be the correct asking price by the seller for its used car and also the correct price and correct choice of car for the buyer and that's where Drive AI came into play our web app not only predicts the price of the car that user wants to sell by taking the important parameters like transmission, fuel type, Millage, Company and Year of model but it also predict the value for a customer who wants to buy a used car as well and not only buyer and seller can use this app ,it also Provides Users to share their reviews about cars they owned and current market analysis as well. But also car brokers can use our app as well to get the idea and correct price of a car they want to predict. The Drive AI uses Machine Learning and Data Science to predict the value and prices of both used and new cars. The Machine Learning models are trained on the real World data which is gathered from cars selling websites and attested from the market as well which result in the models which predict the price of the car accurately and similar to market. These Machine Learning models are formed using Linear Regression, Ridge and Lasso Regressions and other Machine Learning techniques. After these models are trained and accuracy of greater than 78 percent is achieved then they are deployed on a web app which is created by using flask as backend and then the frontend is created using popular designing languages like HTML, CSS, Bootstrap and JavaScript.



## **Chapter 2**

### **Problem Statement**

## **2.1 Problem Statement**

If we talk about selling and buying of used cars there are many online web and app on internet, but there is no such platform that is providing first-hand knowledge about price of car and guiding people at what price they should buy. As we all know in market there are agents who manipulate the prices and buyer have no other option but depending on them, so we came up with an idea that if a buyer knows approximate value of cars so the agents can't make them fool. For getting approximate price of car we have developed a website which will predict value of buyer's desired car that will give some sort of idea to the user about their desired car. For prediction of cars we have used machine learning models.

## **Chapter 3**

### **Background Research Done**

### **3.1 Background research done**

We have done a lot of background research for Drive AI project. First of all, we started looking for research papers related to this project and found many research papers, some of them were mentioned above. With the help of research paper, we've got an idea how to lead our project step by step. In this project we are predicting car price of used cars, new cars and imported cars so we started collecting real world datasets from multiple places like kaggle, pakwheels, olx so we could train our model on this data and predict the price for our desired cars. As we know that our project is predicting car prices based on machine learning models so we started learning machine learning models from YouTube. We found out that this is a supervised learning problem so we learned many supervised learning models and use some of them in our project, after using all the machine learning models we came to the conclusion that linear regression fits best for our project because it's accuracy is more accurate than any other model.

## **Chapter 4**

### **Final Problem Statement**

## 4.1 Final Problem Statement

Car is one of the most important part of today's lifestyle with increase in demand for used cars and up to 10 percent decrease in demand for the new cars in 2021, more and more vehicle buyers are finding alternatives of buying new cars outright because of the latest trends of the market and the government policy of not providing any more financing for the buying of imported cars which are more than 1000CC ,so the people are now moving towards the purchasing of used cars and whenever someone try to buy or sell used cars the main issue that exist is that what should be the correct price, that's where Drive AI came into play our web app not only predicts the price of the car that user wants to sell but it also predict the value for a customer who wants to buy a used car as well and not only buyer and seller can use this app, But also car brokers can use our app as well to get the idea and correct price of a car they want to predict. The Drive AI uses Machine Learning and Data Science to predict the value and prices of both used and new cars.

**Chapter 5**  
**Requirements and Analysis on this Problem**  
**Statement**

## 5.1 Requirements

According to the above problem statement the requirements which needed to build the desired solution that is in the form of web app are following.

### 5.1.1 Data Sets

The DRIVE AI predicts three types of car prices including used car, new and imported cars and for that purpose we needed three types of Data Sets which can solve and provide us correct data about each type. The Data Sets are gathered from real world car websites like Pakwheels, Olx and sometimes extracted from the official websites of companies like Pak Suzuki and Indus Toyota. The Data Sets are mentioned below:

#### 5.1.1.a Used Car Data Set

The data set provide us the information about the used cars like their company, year, model and other attributes, The Data Set is taken from kaggle and is a combination of both Olx and Pakwheels data. It have more than 2250 records of different used cars.

#### 5.1.1.b Imported Car Data Set

The data set provide us the information about the Imported cars like their company, year, model, fuel type and other attributes, The Data Set is taken from Github and it have more than 650 records of different Imported cars of different brands.

#### 5.1.1.c New Car Data Set

The data set provide us the information about the New cars like their company, year, model, fuel type, suspension and price, The Data Set is taken from no of sources like Pak Suzuki, Toyota Indus and other websites and then it is combined together. It have more than 150 records of different New cars of different brands.

### 5.1.2 Programming Requirements

The programming tools and languages which we need to build the web app according the above problem statement are:

#### 5.1.2.a Python

Used for building Machine Learning models by using Machine Learning libraries and packages include Pandas, Numpy, Seaborn (these libraries are used to clean, preprocess the data and perform EDA on it) and Pickle which is used to dump ML model into pickle file so that it can be used at backend integration. For the Backend integration we used FLASK a web architecture which is used to build web apps.



### **5.1.2.b JavaScript, Html and CSS**

The HTML and CSS are used for designing of the web app Frontend while the JavaScript is used for providing the functionality in the Frontend

### **5.1.2.c Software Development Environment and Interface**

For training of Machine Learning models we need Jupyter Notebook which is used worldwide for Machine Learning, For Frontend we used VS Code which is a code editor and for Backend we used Pycharm one the famous code editor and Integrated Development Environment for Python related programming.

## **5.2 Analysis**

Analysis of the Problem statement tell us that the with increase in demand for used cars and up to 10 percent decrease in demand for the new cars in 2021, more and more vehicle buyers are finding alternatives of buying new cars outright because of the latest trends of the market and the government policy of not providing any more financing for the buying of imported cars which are more than 1000CC ,so the people are now moving towards the purchasing of used cars and whenever someone try to buy or sell used cars the main issue that exist is that what should be the correct price, that's where Drive AI came into play our web app not only predicts the price of the car that user wants to sell by taking the important parameters like Transmission, Fuel type, Mileage, Company and Year of model but it also predict the value for a customer who wants to buy a used car as well and not only buyer and seller can use this app, But also car brokers can use our app as well to get the idea and correct price of a car they want to predict. The Drive AI uses Machine Learning and Data Science to predict the value and prices of both used and new cars. The Machine Learning models are trained on the real World data which is gathered from cars selling websites and attested from the market as well which result in the models which predict the price of the car accurately and similar to market. These Machine Learning models are formed using Linear Regression, Ridge and Lasso Regressions and other Machine Learning techniques. After these models are trained and accuracy of greater than 78 percent is achieved then they are deployed on a web app which is created by using flask as backend and then the frontend is created using popular designing languages like HTML, CSS, Bootstrap and JavaScript.

## **Chapter 6**

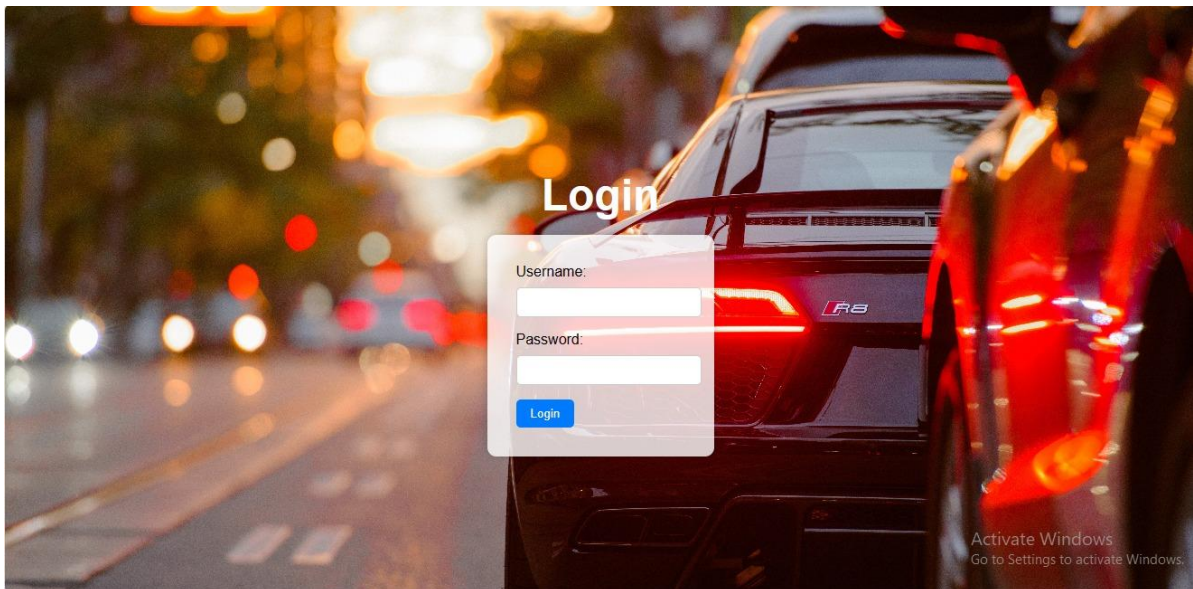
### **Prototype**

## 6.1 Design of Prototype

The prototype is the web app predicting the accurate prices of the cars. For designing the prototype, we used Frontend designing tools like CSS and HTML. The overall prototype consists of five pages namely Home, used car, New car, Imported Car, Help pages, Reviews and Analytics. The frontend images of these pages are given below:

### 6.1.0 Login

The Drive AI app's initial component is the Login Page, meticulously designed to provide users with a secure and intuitive gateway into the application's ecosystem. Upon launching the app, users are presented with a sleek and user-friendly interface, inviting them to log in to their personalized accounts. The Login Page offers a seamless authentication process, allowing users to enter their credentials, including usernames and passwords, ensuring data security and privacy.



### 6.1.1 Home Page

The home page is designed using HTML, CSS, JavaScript and bootstrap and below is the image of home page:

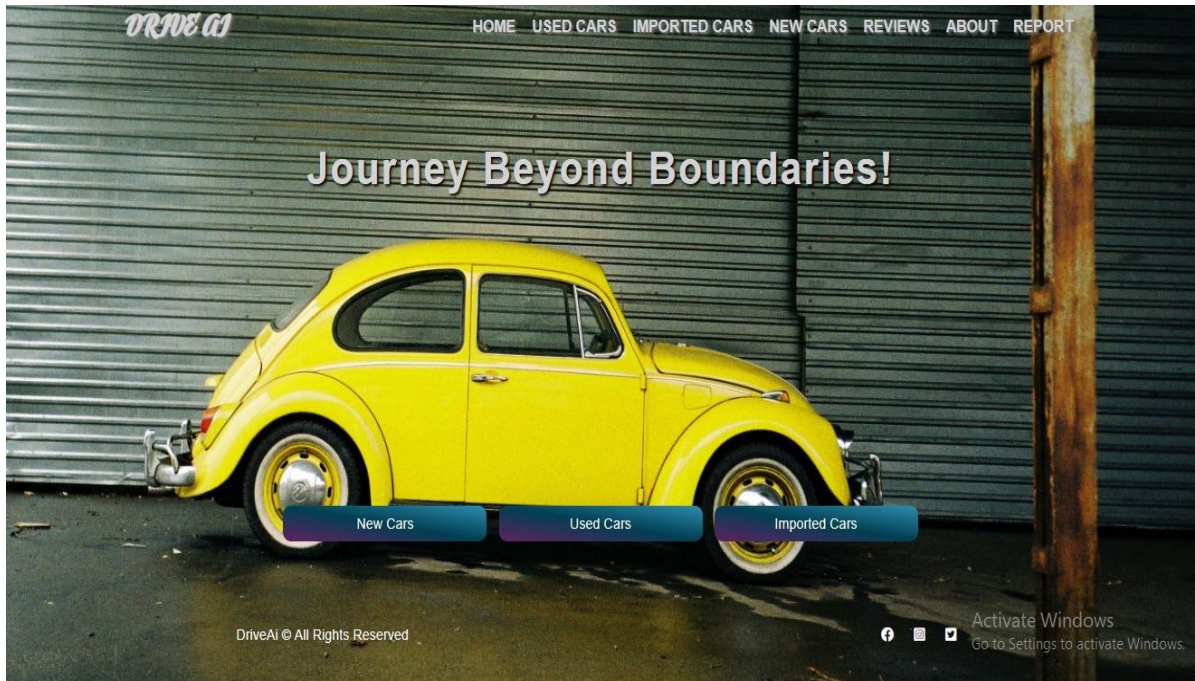


Image 6.1.1 Home Page

The home page is further divided into 3 parts;

#### 6.1.1.a The NavBar

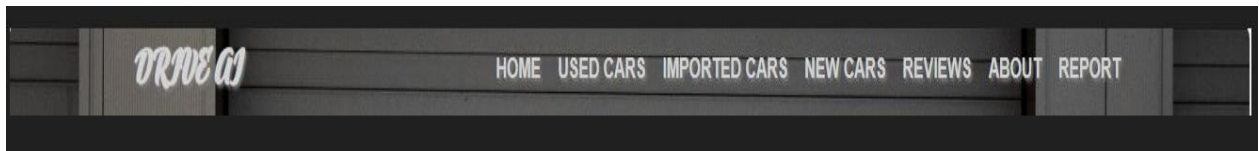


Image 6.1.2 The NavBar

The NavBar is consist of seven parts. The leftmost is the name of our web app “DRIVE AI” and following it is the home button which lead us to home page and following it is Used Car button which have a purpose of navigating the user to Used Car prediction page and following it is the Imported Car which will navigate the user to the Imported Car prediction page and after it we have New Car which will move the user to New Car predict page, Reviews page where users can write reviews, Report where Users can see latest auto industry Trends and lastly we have help which will move user to help page in case if user have any query.



### 6.1.1.b The Buttons

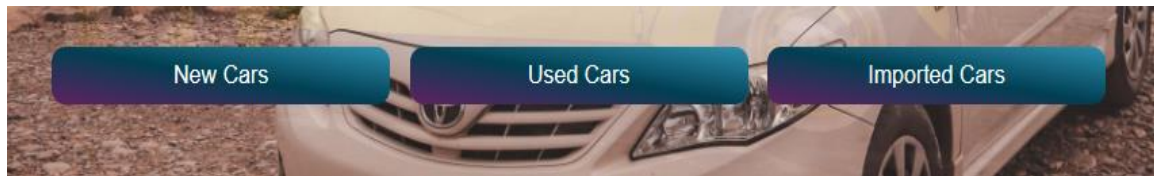


Image 6.1.3 The Buttons

Our home page have three main buttons on center of it. The purpose of each button is to navigate the user to the desired car price predict page. The New Car button as it name suggest it moves the user to New Car price page and the Used Car button move user to use car page and the Imported Car button move the user to Imported Car price page.

### 6.1.1.c The Footer



Image 6.1.4 The Footer

The footer contains the link to our social media channels including Instagram and Facebook and on left hand side it contains the Drive AI all right reserved logo.

### 6.1.2 Used Car

The Used Car page is designed using HTML, CSS, JavaScript and bootstrap and below is the image of Used Car page:

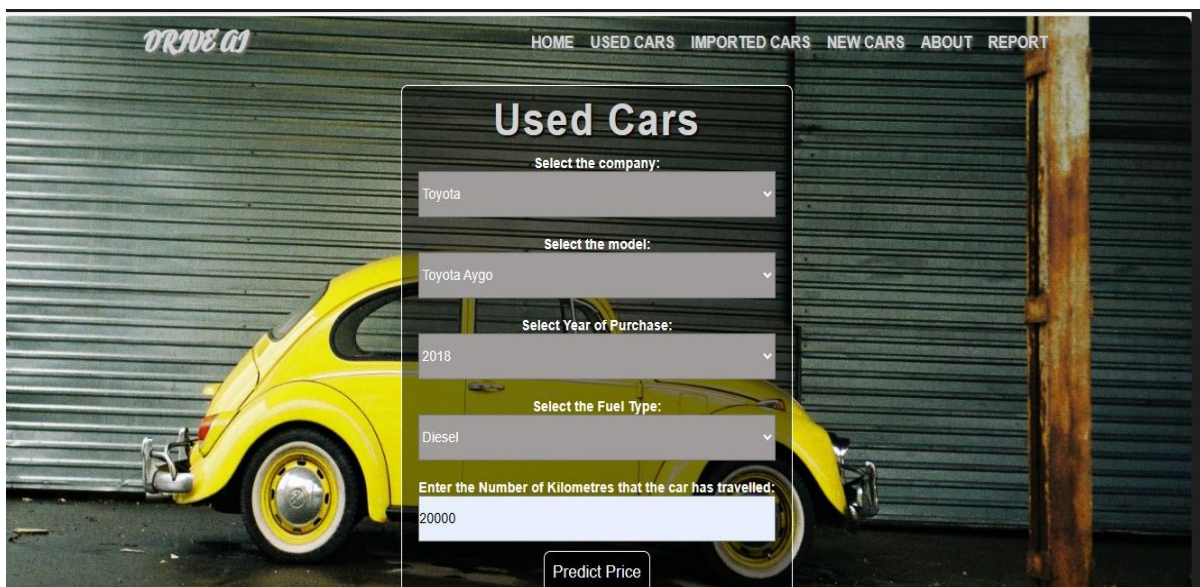
A screenshot of a web application's 'Used Cars' page. The page has a dark background with a yellow Volkswagen Beetle on the left. A white form is centered on the page. The form has a title 'Used Cars' and several dropdown menus: 'Select the company:' with 'Toyota' selected, 'Select the model:' with 'Toyota Aygo' selected, 'Select Year of Purchase:' with '2018' selected, and 'Select the Fuel Type:' with 'Diesel' selected. Below these is a text input field labeled 'Enter the Number of Kilometres that the car has travelled:' with '20000' entered. At the bottom of the form is a button labeled 'Predict Price'. The top of the page has a navigation bar with links: 'HOME', 'USED CARS', 'IMPORTED CARS', 'NEW CARS', 'ABOUT', and 'REPORT'.

Image 6.1.5 Used Car Page

The Used Car page is further divided into 3 parts;

### **6.1.2.a The Input Form**

The image shows a web form titled "Used Cars" overlaid on a background image of a car. The form contains five dropdown menus and one text input field. The first dropdown is labeled "Select the company:" and has "Select Company" as the placeholder. The second is labeled "Select the model:" with an empty placeholder. The third is labeled "Select Year of Purchase:" and has "2018" selected. The fourth is labeled "Select the Fuel Type:" and has "Diesel" selected. The fifth is a text input field labeled "Enter the Number of Kilometres that the car has travelled:" with the placeholder text "Enter the kilometres driven". Below the input fields is a black button with white text that says "Predict Price".

Image 6.1.6 The Input Form

The Input form consist of five input fields namely Company which represent the company of the car, Model which represent the model of the car, Year which represent the model year of car, Fuel type which is used to represent the type of fuel the car uses and lastly it contains the Number of Kilo meters driven by Car. After the user fill all the input fields he then clicks the Predict Button which will give the predicted value.

### **6.1.3 New Car**

The New Car page is designed using HTML, CSS, JavaScript and bootstrap and below is the image of Used Car page:

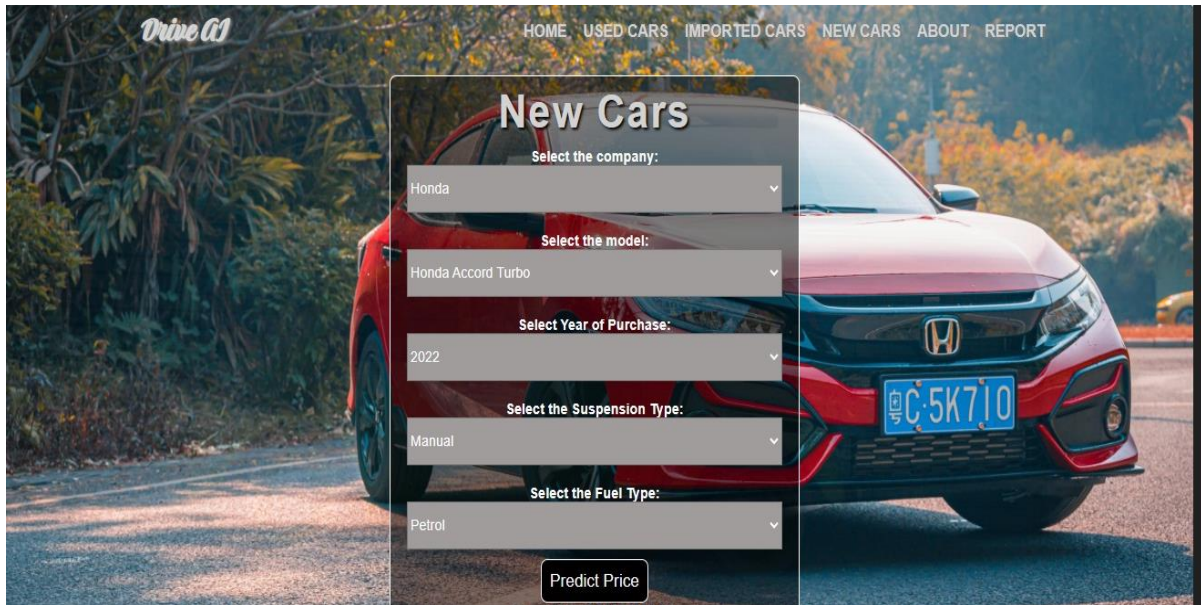


Image 6.1.7 New Car Page

The New Car page is further divided into 3 parts;

#### 6.1.3.a The Input Form



Image 6.1.8 The Input Form

The Input form consist of five input fields namely Company which represent the company of the car, Model which represent the model of the car, Year which represent the model year of car, Fuel type which is used to represent the type of fuel the car uses and lastly it contains the type of Suspension of the Car. After the user fill all the input fields he then clicks the Predict Button which will give the predicted value.

#### 6.1.4 Imported Car

The Imported Car page is designed using HTML, CSS, JavaScript and bootstrap and below is



the image of Used Car page:

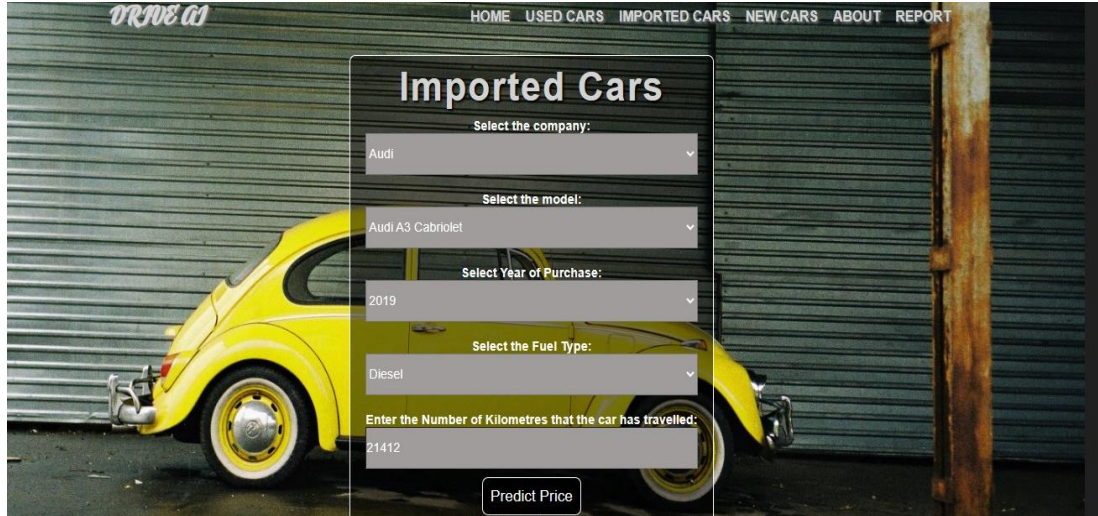


Image 6.1.9 Imported Car Page

The Imported Car page is further divided into 3 parts;

#### **6.1.4.a The Input Form**



Image 6.1.10 The Input Form

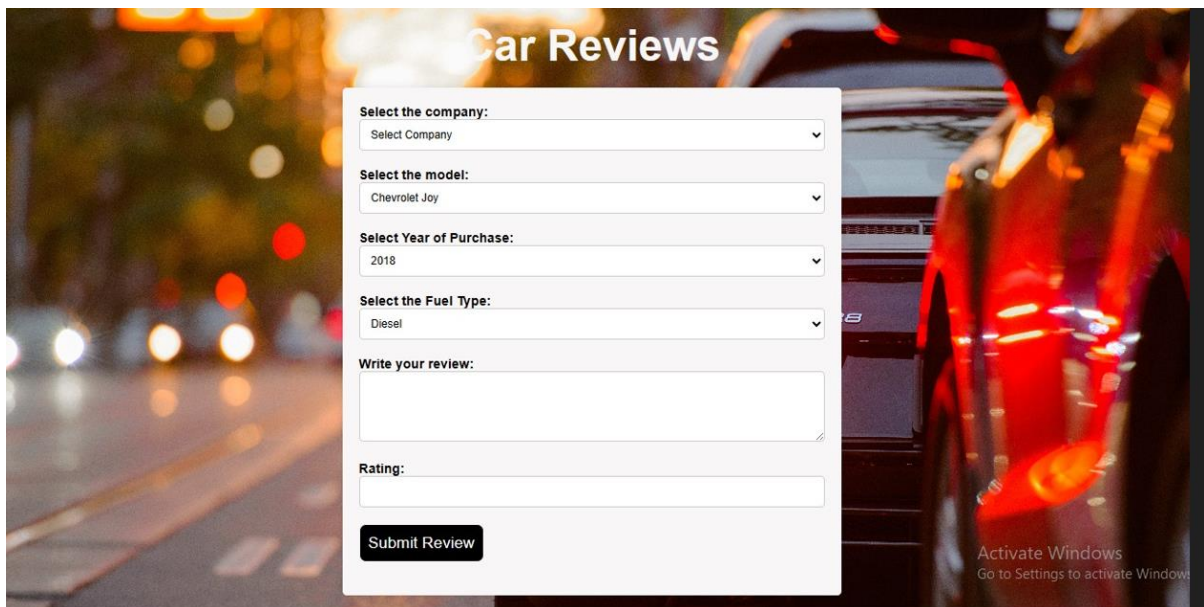
The Input form consist of five input fields namely Company which represent the company of the car, Model which represent the model of the car, Year which represent the model year of car, Fuel type which is used to represent the type of fuel the car uses and lastly it contains the Number of Kilo meters driven by Car. After the user fill all the input fields he then clicks the Predict Button which will give the predicted value.



## 6.1.5 Reviews

The Review Page within the Drive AI app is a dedicated space designed to enable users to share their valuable experiences and insights about their owned vehicles. With a user-friendly interface, this section allows individuals to select the specific details of their car, including the manufacturer/company, model, year of purchase, and fuel type.

Users are prompted to input their thoughts and experiences about their vehicle, providing a platform to share reviews, feedback, and opinions in a structured manner.



**Car Reviews**

Select the company:  
Select Company

Select the model:  
Chevrolet Joy

Select Year of Purchase:  
2018

Select the Fuel Type:  
Diesel

Write your review:

Rating:

Submit Review

Activate Windows  
Go to Settings to activate Windows

## 6.1.6 Report

The Power BI dashboard designed for car sales analysis encapsulates a comprehensive array of insightful features. Delving into sales trends, it meticulously dissects the data to uncover patterns, fluctuations, and growth trajectories within the automotive market. By highlighting top performers, the dashboard offers a clear representation of the most successful vehicles or models, empowering decision-makers to discern influential factors behind their success. The integration of a map feature provides a geographical understanding of sales distribution, unveiling regional hotspots and potential growth areas. Leveraging forecasting capabilities, the dashboard offers a glimpse into the future, enabling proactive strategies based on predictive trends. Additionally, the identification of key influencers in car sales trends elucidates the pivotal drivers affecting purchase behaviors, providing invaluable insights for strategic planning and marketing initiatives. Overall, this Power BI dashboard serves as a dynamic tool, synthesizing diverse analytics facets to empower stakeholders with actionable intelligence in the competitive automotive market landscape.

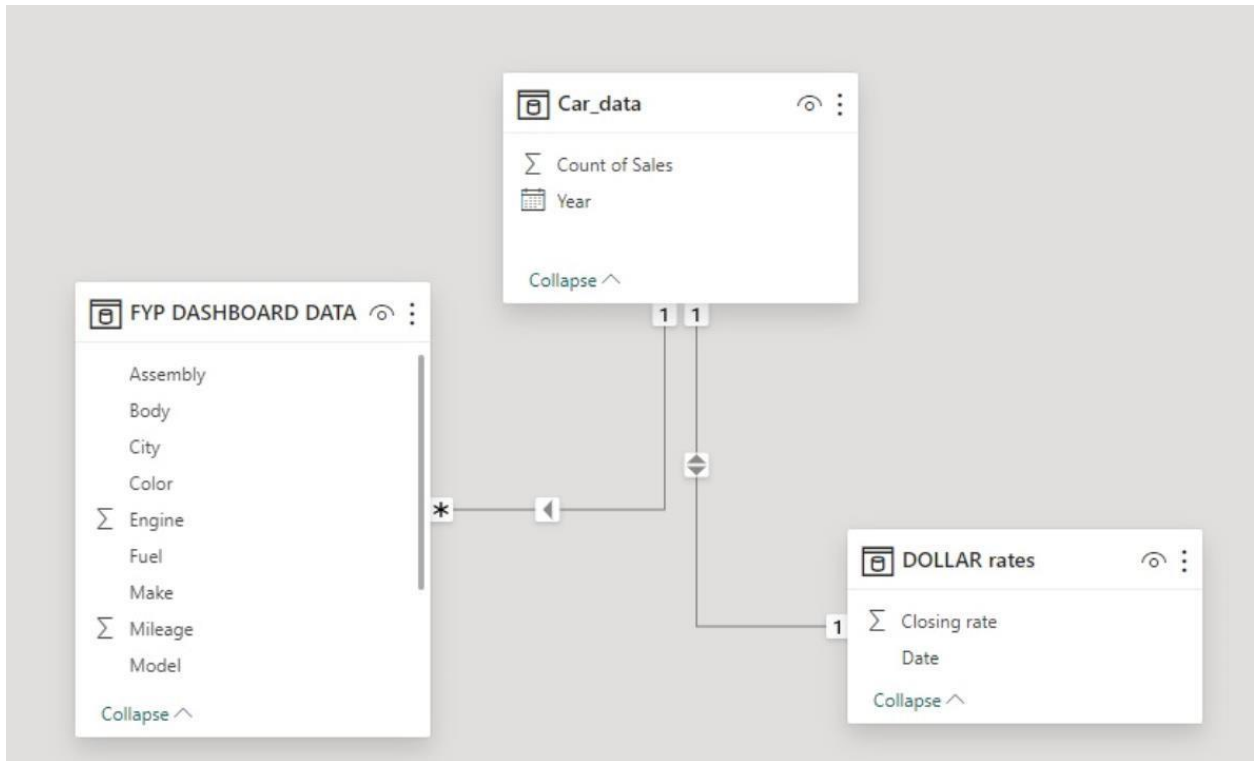
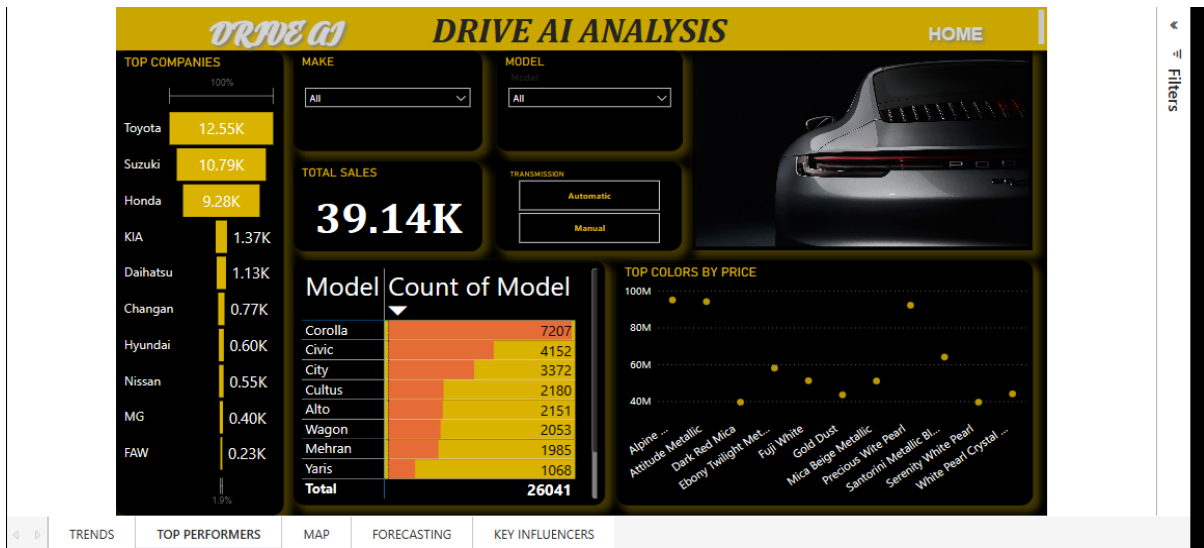
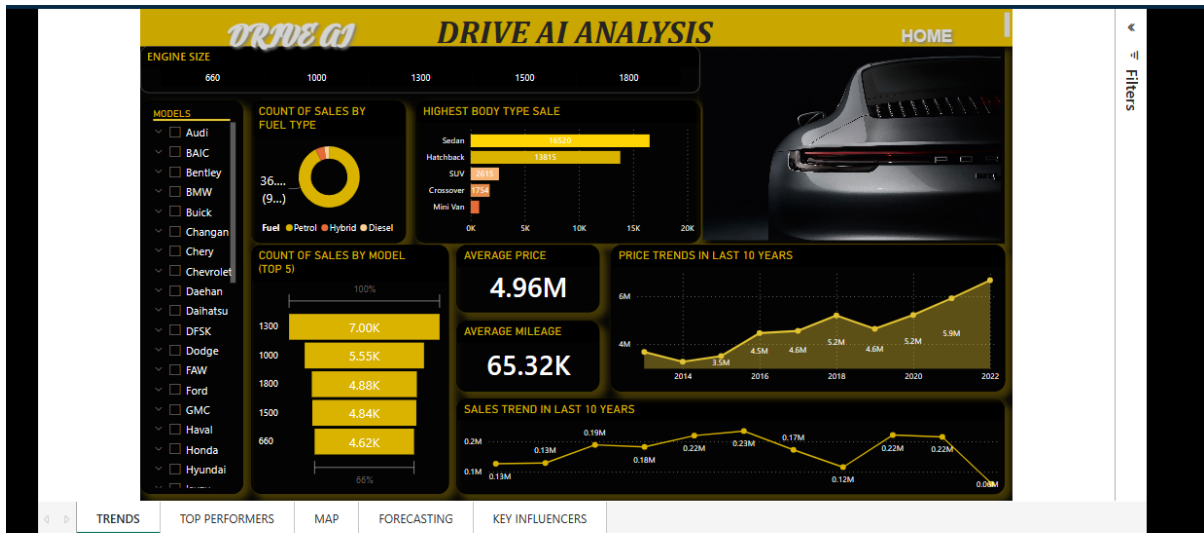
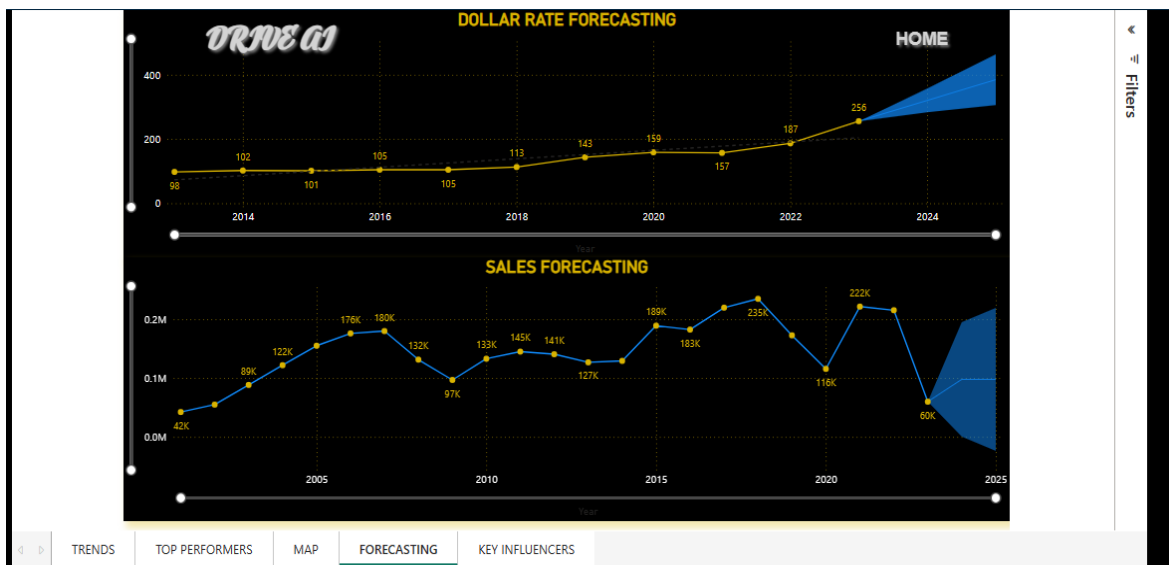
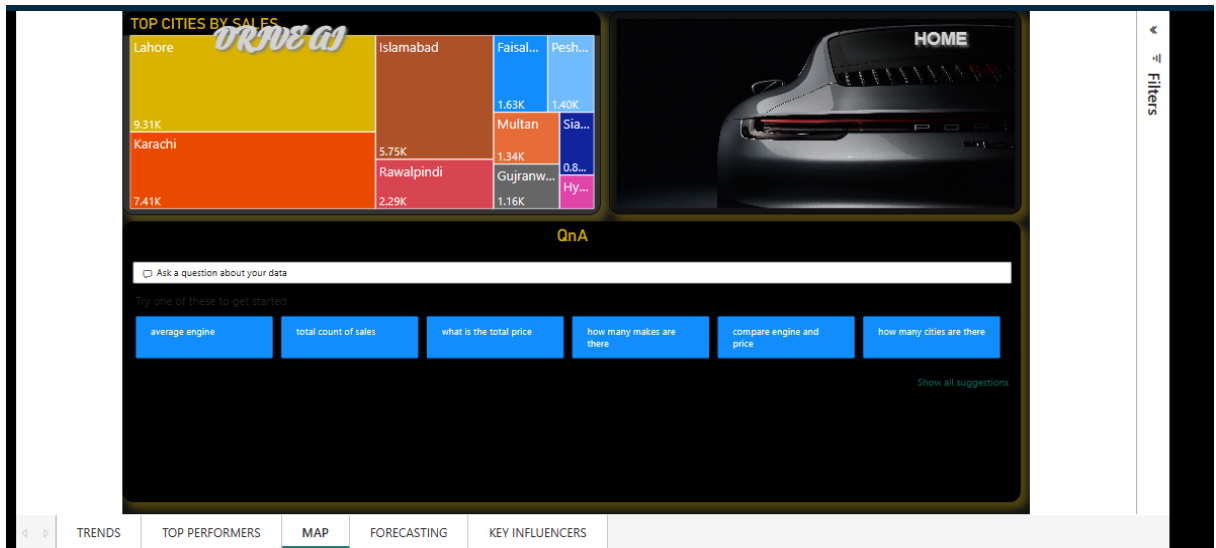
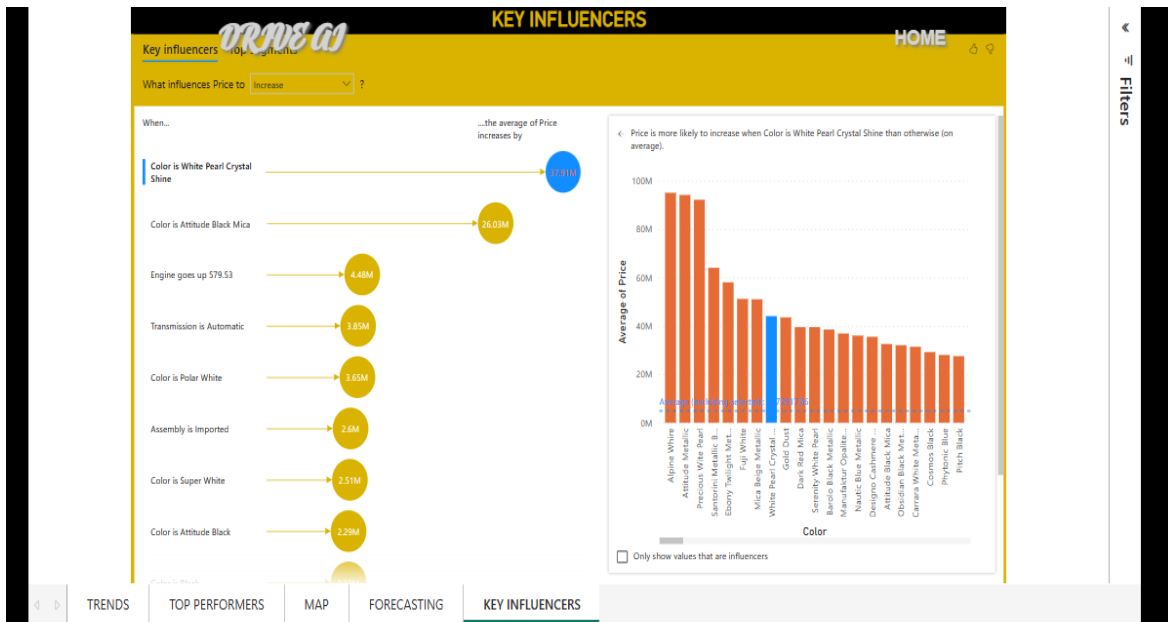


Image 6.1.6 Dashboard Model View







## 6.2 Implementation of Prototype

The Implementation of the prototype is done using Flask for Backend and using python for building Machine Learning models. Since our web app have three Machine Learning Models each of which is predicting the price of specific type of car. The details about each model is given below:

### 6.2.1 Used Car Model

The used car model is used to predict the price of the Used Cars and it is trained on the data set obtained from kaggle the image of the data set is given below:

	name	company	year	Price	kms_driven	fuel_type
0	Toyota Prado	Toyota	1997	2100000	100000	Diesel
1	Suzuki Bolan	Suzuki	2006	380000	100000	Petrol
2	Suzuki Bolan	Suzuki	1998	340000	12345	CNG
3	Suzuki Alto	Suzuki	2010	535000	94000	Petrol
4	Toyota Corolla XLI	Toyota	2013	1430000	100000	Petrol
5	Toyota Corrolla Altis	Toyota	2012	1620000	80000	Petrol
6	Suzuki Cultus VXL	Suzuki	2006	450000	65000	CNG
7	Toyota Corrolla Altis	Toyota	2017	2900000	10241	Petrol
8	Suzuki Alto	Suzuki	2009	490000	83000	CNG
9	Honda Civic VTi	Honda	1997	480000	50000	Petrol
10	Suzuki Khyber	Suzuki	1994	230000	1230000	CNG
11	Suzuki Liana	Suzuki	2006	535000	73000	CNG
12	Toyota Passo	Toyota	2006	770000	78000	Petrol
13	Honda Civic Prosmetic	Honda	1997	490000	100000	Petrol
14	Honda Civic EXi	Honda	1984	160000	170000	CNG
15	Suzuki Alto	Suzuki	2005	425000	200000	Petrol
16	Daihatsu Charade	Daihatsu	1988	180000	100000	Petrol
17	MitsubishiPajero Mini	Mitsubishi	1995	520000	82000	Petrol
18	Suzuki Margalla	Suzuki	1990	400000	10000	CNG
19	Honda City IVTEC	Honda	2014	1400000	69500	Petrol

Image 6.2.1 Used Car Data Set

The data set shown in above image have 6 columns and a total of 2250 rows. The columns name, company, year, fuel type and km's driven represents the input and the price will be the output.

During the model training we will also clean the data from missing and in consistent value and we also try to find the correlation between the different attributes of data set some of these correlation is given below

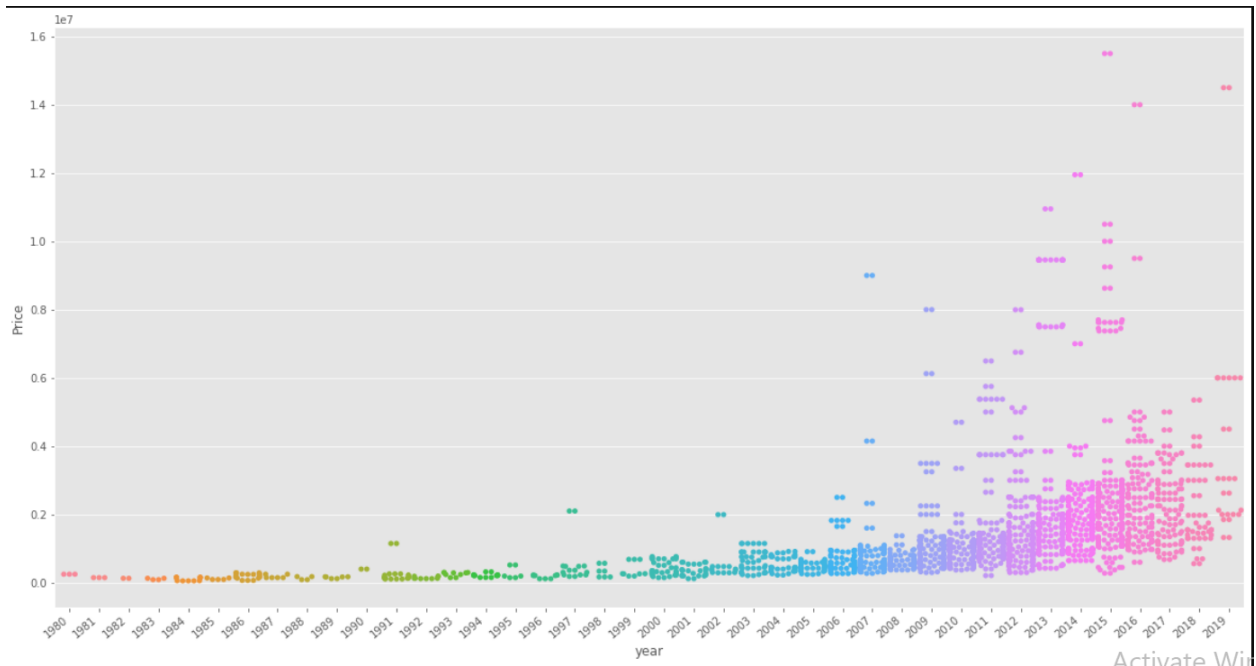


Image 6.2.2 Correlation between Price and Year

Above image represents the correlation between the price of the car with respect to the year of car and it is seen that the older the year of car the lesser its price will be

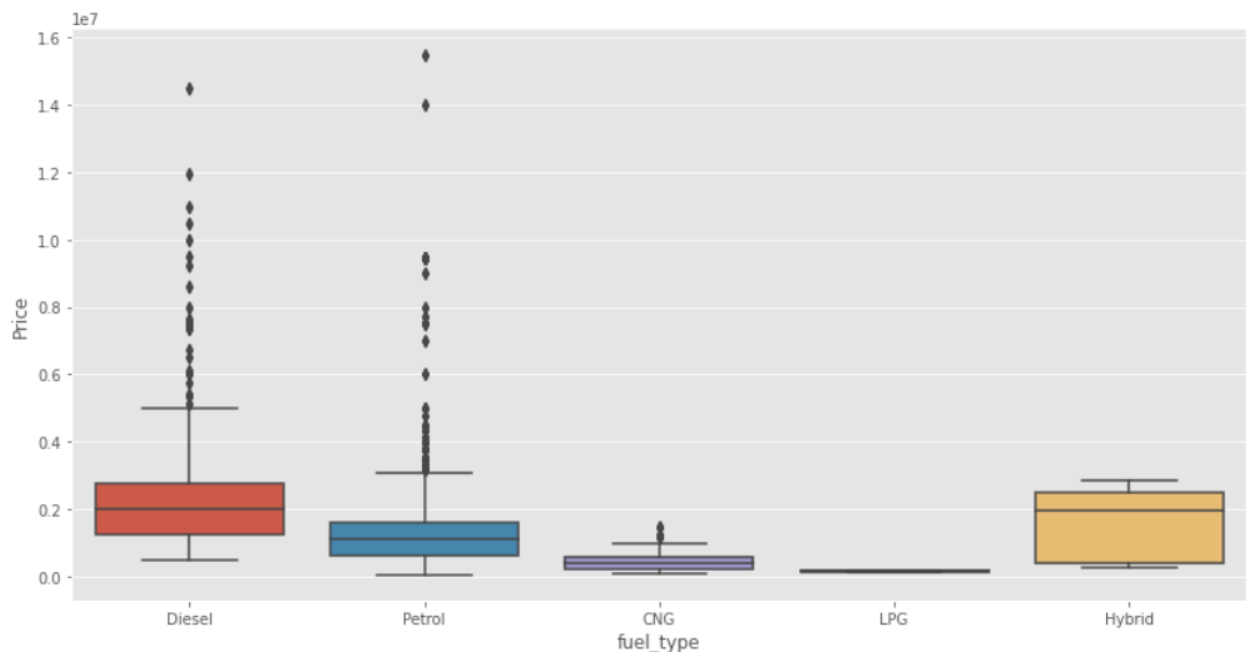


Image 6.2.3 Correlation between Price and Fuel Type

Above image represents the correlation between the price of the car with respect to the fuel type of car and it is seen that the Diesel have the most expensive cars while the LPG have the cheapest of cars.

The model for Used Cars is trained using Linear Regression model and it achieves the test accuracy of 88 percent which is very good in terms of machine learning.

### 6.2.2 New Car Model

The New car model is used to predict the price of the New Cars and it is trained on the data set obtained from websites like Pak Suzuki and Indus Toyota the image of the data set is given below:

	name	company	year	Price	suspension	fuel_type
0	Kia Picanto	Kia	2022	2186000	Manual	Petrol
1	Kia Picanto	Kia	2022	2292000	Automatic	Petrol
2	Kia Stonic EX	Kia	2022	3750000	Automatic	Petrol
3	Kia Stonic EX+	Kia	2022	3975000	Automatic	Petrol
4	Kia Sportage Alpha	Kia	2022	4764000	Automatic	Petrol
5	Kia Sportage AWD	Kia	2022	5276000	Automatic	Petrol
6	Kia Sportage FWD	Kia	2022	5800000	Automatic	Petrol
7	Kia Carnival GLS	Kia	2022	9199000	Automatic	Petrol
8	Kia Carnival GLS+	Kia	2022	9999000	Automatic	Diesel
9	Kia Sorento FWD	Kia	2021	6836000	Automatic	Petrol
10	Kia Sorento AWD	Kia	2022	7499000	Automatic	Petrol
11	Suzuki Alto VX	Suzuki	2022	1306000	Automatic	Petrol
12	Suzuki Alto VXR	Suzuki	2022	1546000	Automatic	Petrol
13	Suzuki Alto VXL	Suzuki	2022	1747000	Automatic	Petrol
14	Suzuki Wagon VXR	Suzuki	2022	1877000	Automatic	Petrol

Image 6.2.4 New Car Data Set

The data set shown in above image have 6 columns and a total of 150 rows. The columns name, company, year, fuel type and suspension represents the input and the price will be the output.

During the model training we will also clean the data from missing and in consistent value and we also try to find the correlation between the different attributes of data set some of these correlation are given below;



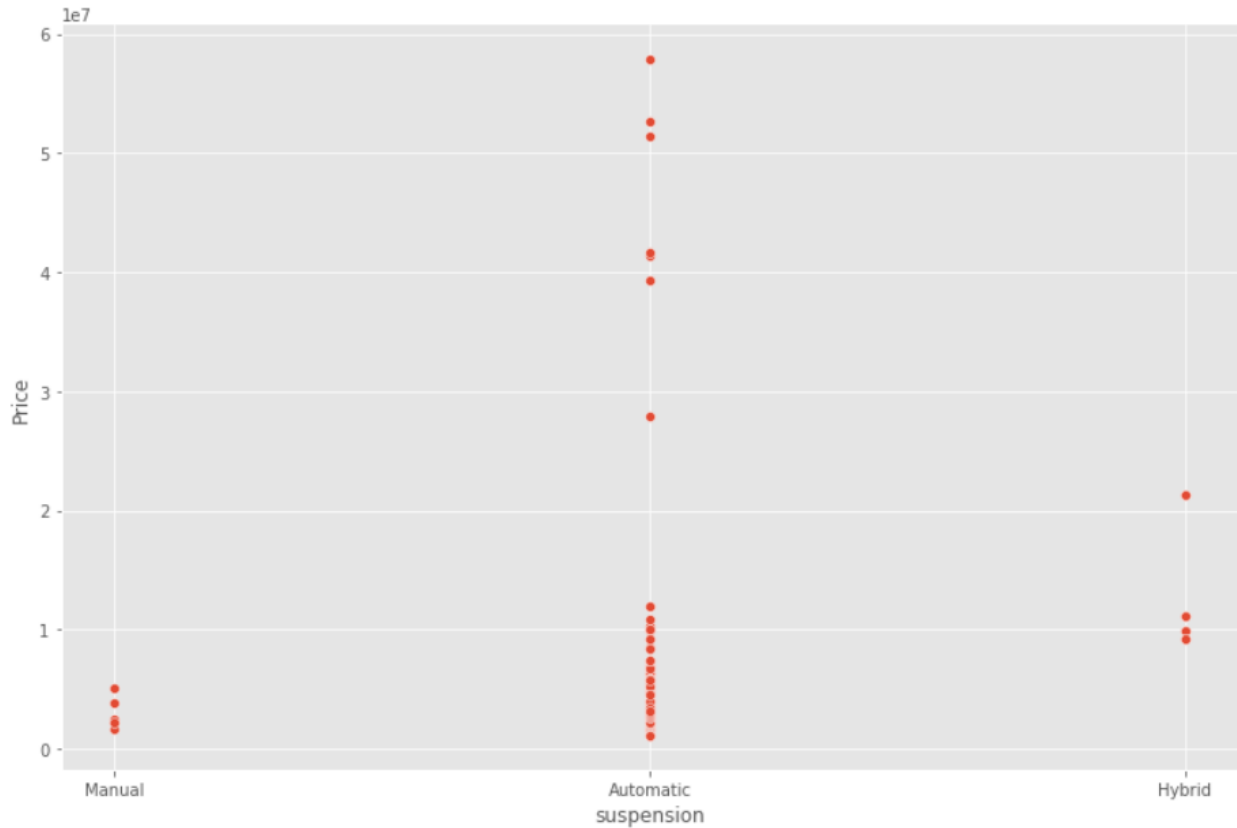


Image 6.2.5 Correlation between Price and Transmission

Above image represents the correlation between the price of the car with respect to the suspension type of car and it is seen that the Automatic suspension have the highest prices while the Manual suspension have the lowest prices and the Hybrid suspension have prices in modest range.

Another correlation is given below:

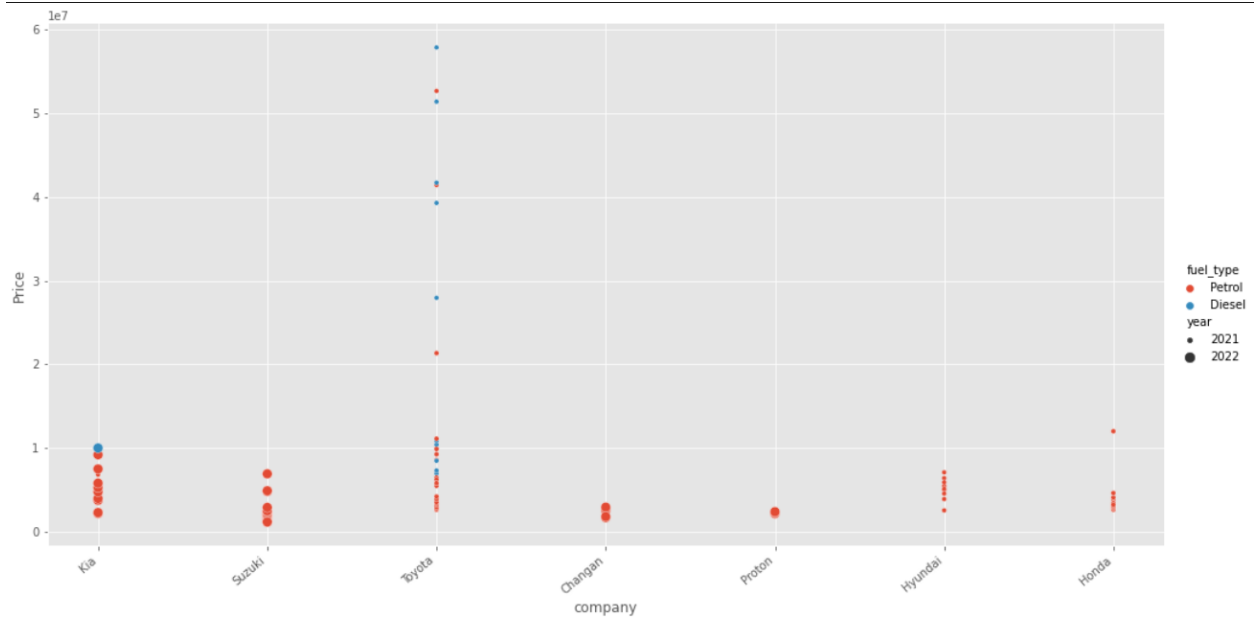


Image 6.2.6 Correlation between Price and Company

Above image represents the correlation between the price of the car with respect to the Company of car and it is seen that the Toyota have the most expensive cars while the Proton have the cheapest of cars and all the other companies lies between.

The model for New Cars is trained using Linear Regression model and it achieves the test accuracy of 98 percent which is very good in terms of machine learning.

### 6.2.3 Imported Car Model

The Imported car model is used to predict the price of the Imported Cars and it is trained on the data set obtained from Github the image of the data set is given below:

	0	name	company	year	Price	kms_driven	fuel_type
0	4	Ford EcoSport Titanium	Ford	2014	2875000	36000	Diesel
1	5	Ford Figo	Ford	2012	875000	41000	Diesel
2	7	Ford EcoSport Ambiente	Ford	2016	4150000	24530	Diesel
3	9	Skoda Fabia Classic	Skoda	2010	910000	60000	Petrol
4	15	Audi A8	Audi	2017	5000000	4000	Petrol
5	16	Audi Q7	Audi	2014	2500000	16934	Diesel
6	17	Ford Figo	Ford	2012	875000	41000	Diesel
7	35	Renault Lodgy 85	Renault	2018	3449995	20000	Diesel
8	36	Skoda Yeti Ambition	Skoda	2012	2240000	68000	Diesel
9	38	Renault Duster 110	Renault	2012	2505000	38000	Diesel
10	39	Renault Duster 85	Renault	2013	2449995	27000	Diesel
11	59	Ford Figo	Ford	2012	875000	41000	Diesel
12	62	Ford EcoSport Ambiente	Ford	2016	4150000	24530	Diesel
13	64	Chevrolet Spark LS	Chevrolet	2010	550000	41000	Petrol
14	65	Volkswagen Polo Highline	Volkswagen	2014	1575000	60000	Petrol

Image 6.2.7 Imported Car Data Set

The data set shown in above image have 6 columns and a total of 650 rows. The columns name, company, year, fuel type and Kms driven represents the input and the price will be the output.

During the model training we will also clean the data from missing and inconsistent value and we also try to find the correlation between the different attributes of data set some of these correlation are given below;

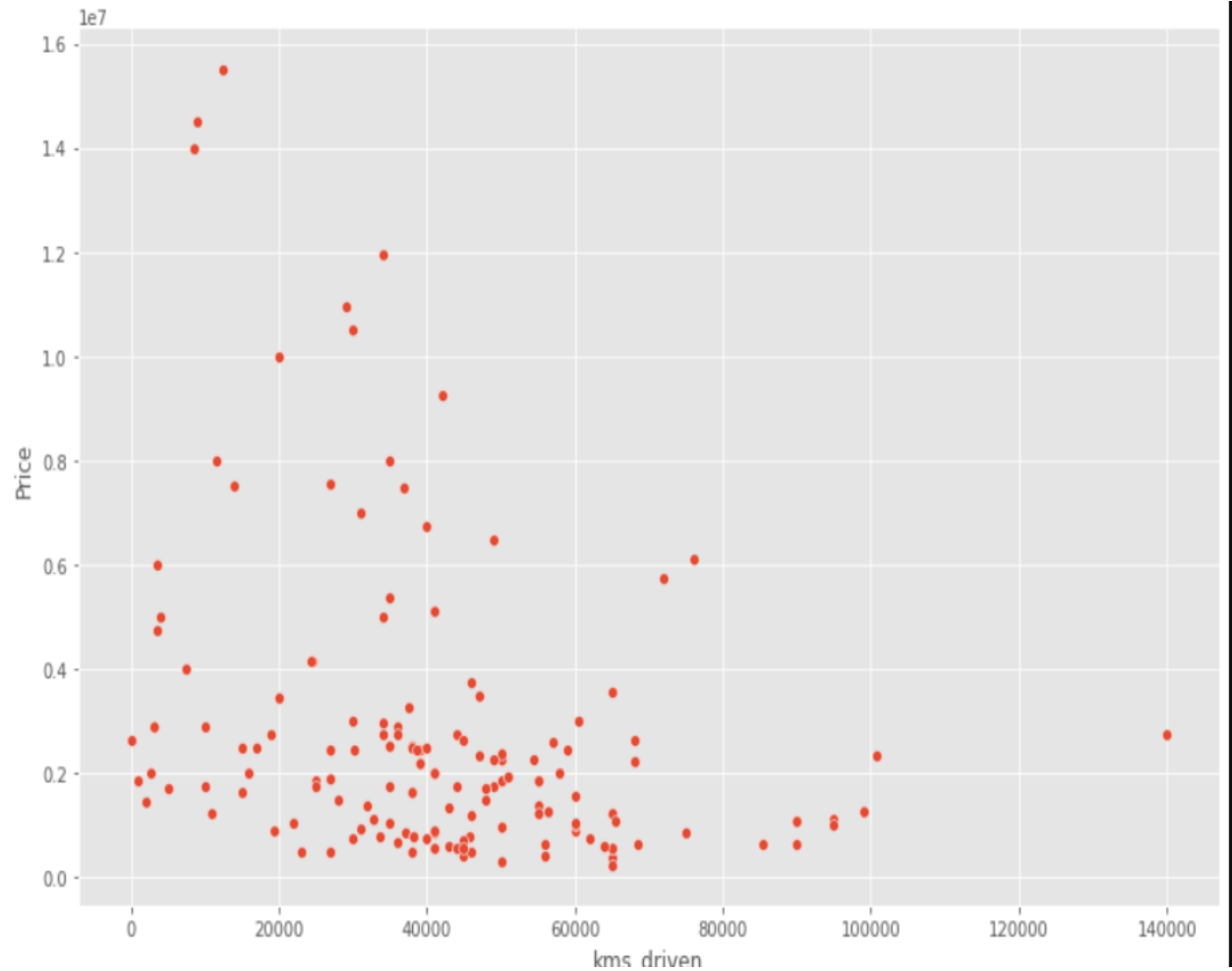


Image 6.2.8 Correlation between Price and Kms\_Driven

Above image represents the correlation between the price of the car with respect to the Kilo meters driven by car and it is seen that the more the car is driven the lesser its price will be.

Another correlation is given below:

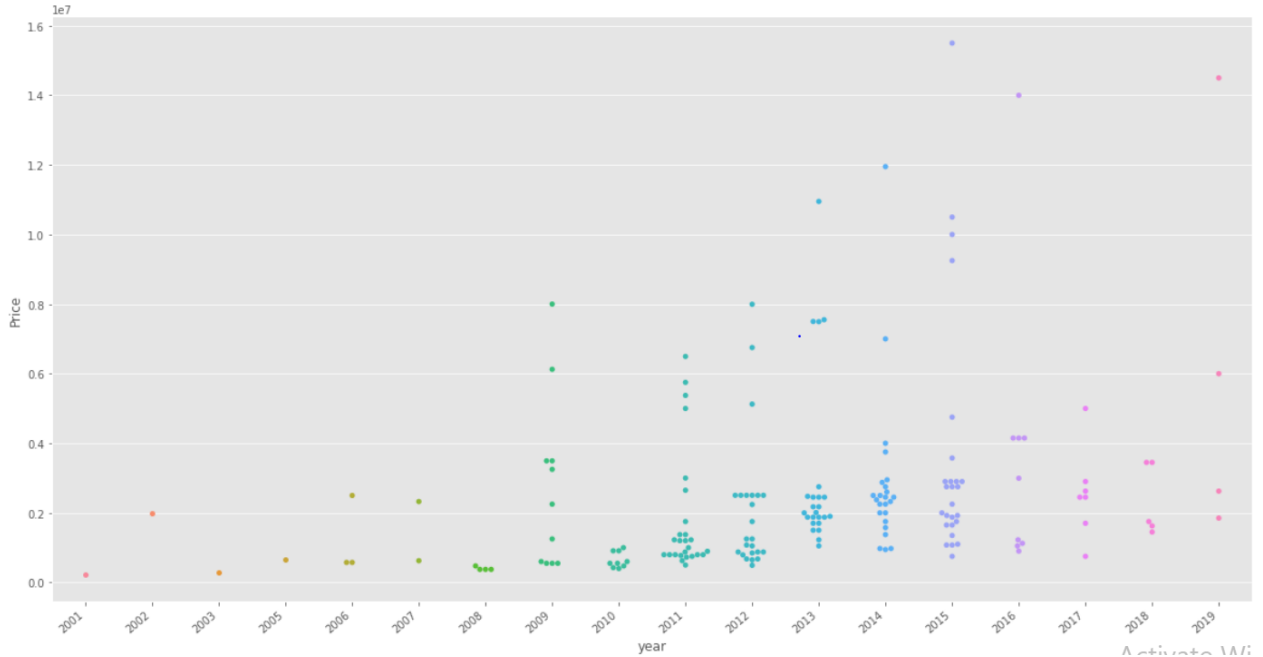


Image 6.2.9 Correlation between Price and Year

Above image represents the correlation between the price of the car with respect to the year of car and it is seen that the car of the latest years has generally higher prices than the cars of the previous or old years.

The model for Imported Cars is trained using Linear Regression model and it achieves the test accuracy of 78 percent which is very good in terms of machine learning.

#### 6.2.4 Backend Code

Now after the models are trained they are deployed on the web app using flask the snip of the code of flask is given below which is used to deploy the models.

```

from flask import Flask, render_template, request, redirect
from flask_cors import CORS, cross_origin
import pickle
import pandas as pd
import numpy as np

app=Flask(__name__)
cors=CORS(app)
model=pickle.load(open('LinearRegressionModel321.pkl','rb'))
car=pd.read_csv('Book111.csv')
car2=pd.read_csv('Book11.csv')
car3=pd.read_csv('newbook11.csv')

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

@app.route('/model1', methods=['GET', 'POST'])
def model1():
    companies=sorted(car['company'].unique())
    car_models=sorted(car['name'].unique())
    year=sorted(car['year'].unique(), reverse=True)
    fuel_type=car['fuel_type'].unique()

    companies.insert(0, 'Select Company')

```

Image 6.2.10 Backend Code Image

```

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

@app.route('/predict',methods=['POST'])
@cross_origin()
def predict():

    company=request.form.get('company')

    car_model=request.form.get('car_models')
    year=request.form.get('year')
    fuel_type=request.form.get('fuel_type')
    driven=request.form.get('kilo_driven')

    prediction=model.predict(pd.DataFrame(columns=['name', 'company', 'year', 'kms_driven', 'fuel_type'],
    data=np.array([car_model,company,year,driven,fuel_type]).reshape(1, 5)))
    print(prediction)

    return str(np.round(prediction[0],2))

model222=pickle.load(open('LinearRegressionModel333.pkl','rb')) #New Car

@app.route('/predict2',methods=['POST2'])
def predict2():
    company = request.form.get('company')

    car_model = request.form.get('car_models')
    year = request.form.get('year')
    suspension = request.form.get('suspension')
    fuel_type = request.form.get('fuel_type')

    prediction2=model222.predict(pd.DataFrame(columns=['name', 'company', 'year', 'suspension', 'fuel_type'],
    data=np.array([car_model,company,year,suspension,fuel_type]).reshape(1, 5)))
    print(prediction2)
    return str(np.round(prediction2[0],2))

if __name__=='__main__':
    app.run()

```

Image 6.2.11 Backend Code Image

In above images it is seen that a pickle file is first loaded for each model and then the html template is rendered for each page of web app and the user after he enter the desired information than it goes to the predict function whose purpose is to take the input provided by user via input form and then send it to the pickle file for each model so that it can predict the desired output value and then take it back from the pickle file and send it to the prediction page to show the user his desired output

```

app = Flask(__name__)
app.config['SQLALCHEMY_DATABASE_URI'] = 'postgresql://postgres:password@localhost/ml-app'
db = SQLAlchemy(app)
with app.app_context():
    db.create_all()
class User(db.Model):
    __tablename__ = "user1"
    id = db.Column(db.Integer, primary_key=True)
    username = db.Column(db.String(80), unique=True, nullable=False)
    password = db.Column(db.String(120), nullable=False)

@app.route('/login', methods=['GET', 'POST'])
def login():
    if request.method == 'POST':
        username = request.form['username']
        password = request.form['password']

        # Query the database for the user
        user = User.query.filter_by(username=username).first()
        passw = User.query.filter_by(password=password).first()
        # Check if the user exists and the password is correct
        if user and passw:
            # Redirect to the index page on successful login
            return redirect(url_for('index'))
        else:
            # Handle invalid credentials
            return 'Invalid username or password'
    else:

```

Image 6.2.12 Login Code snippet

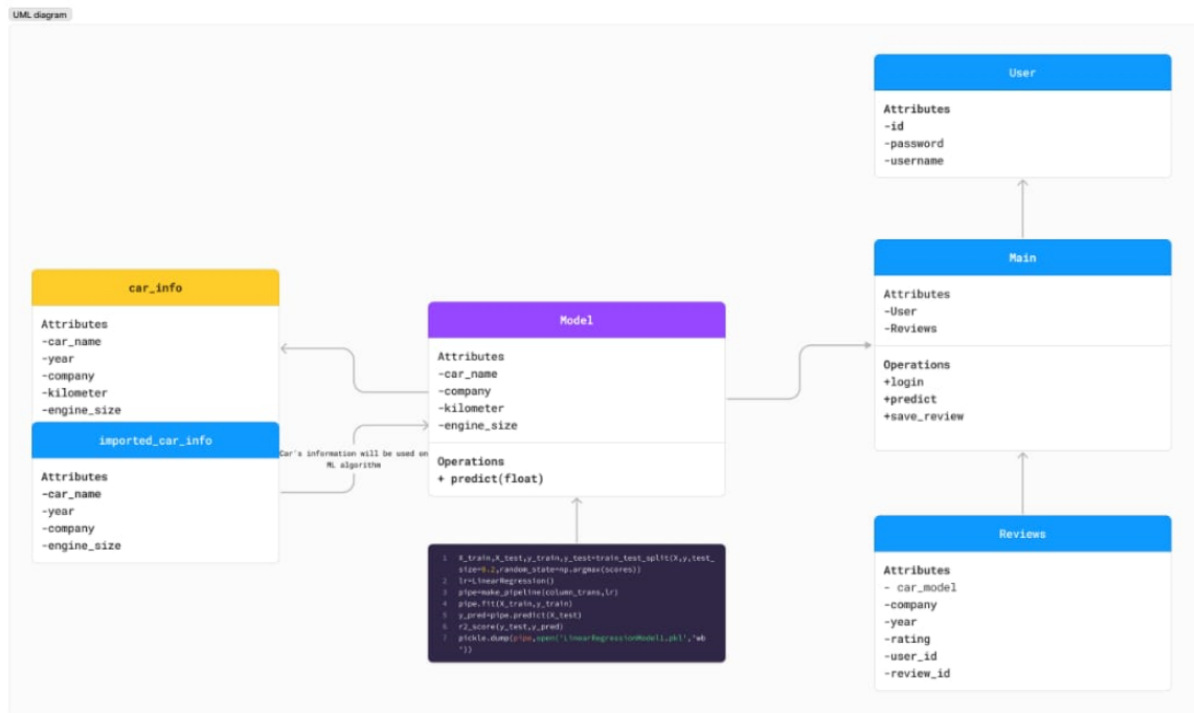


Image 6.2.13 UML Diagram



## **Chapter 7**

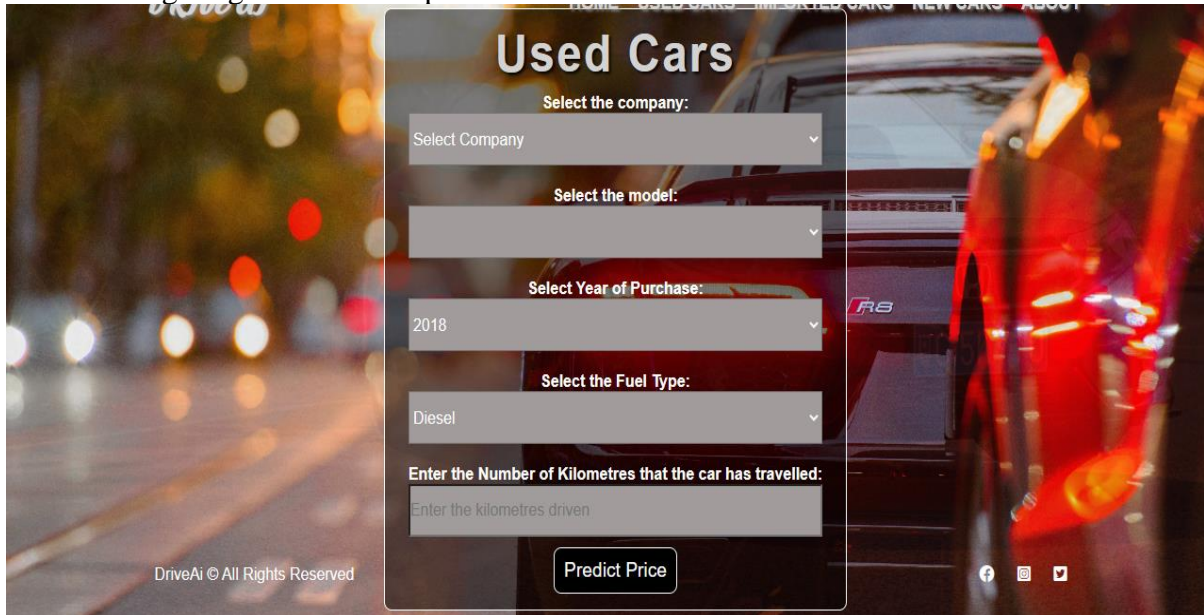
### **Results of The Prototype**

## 7.1 Results of Prototype

The main issue and problem that we try to solve by using our prototype is to predict the price of different type of cars. Since we have three Machine Learning models deployed on our web app with each model predicting the specific type of car price the prediction result of each model are given below:

### 7.1.1 Used Car Model Results

Down below is the image of model before user input and after the user pressing the predict button after giving the desired input.



The screenshot displays a web application titled "Used Cars" overlaid on a background image of a car at night. The form includes the following elements:

- Select the company:** A dropdown menu with "Select Company" as the placeholder text.
- Select the model:** A dropdown menu with an empty field and a downward arrow.
- Select Year of Purchase:** A dropdown menu with "2018" selected.
- Select the Fuel Type:** A dropdown menu with "Diesel" selected.
- Enter the Number of Kilometres that the car has travelled:** A text input field with the placeholder "Enter the kilometres driven".
- Predict Price:** A button located at the bottom right of the form.

At the bottom left of the overlay, the text "DriveAi © All Rights Reserved" is visible. At the bottom right, there are social media icons for Facebook, Instagram, and Twitter.

Image 7.1.1 Used Car Model Before Input

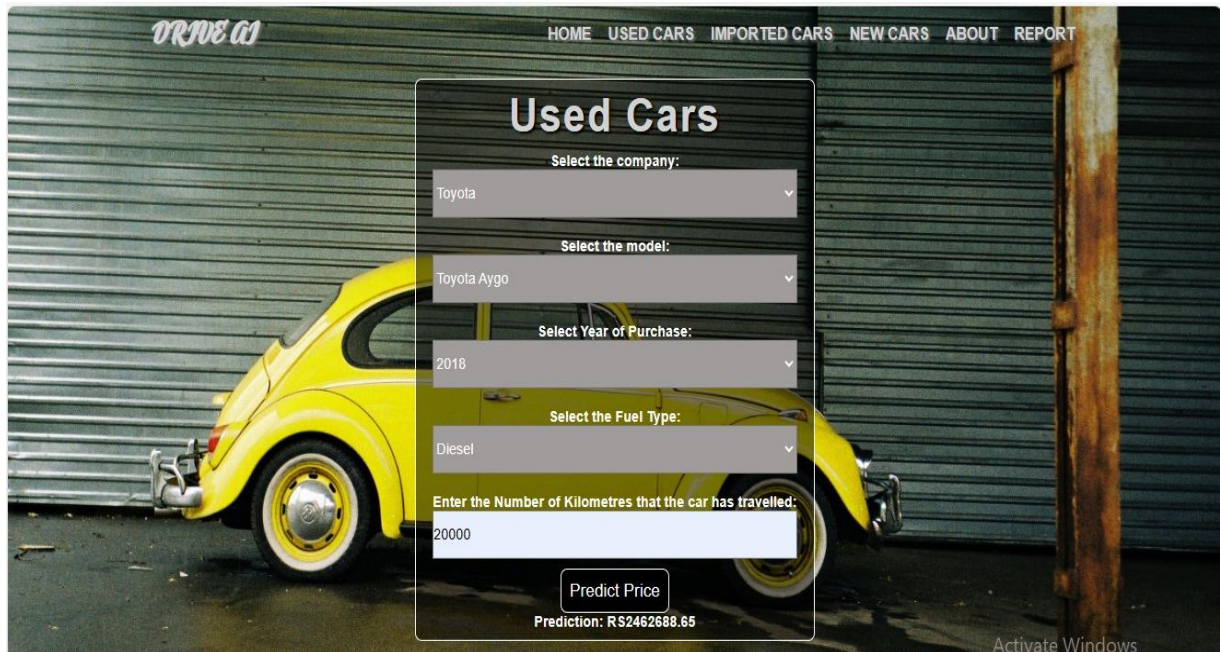


Image 7.1.2 Used Car Model After Input

As it is seen after the user provide the input our model successfully predicts the price.

### 7.1.2 Imported Car Model Results

Down below is the image of model before user input and after the user pressing the predict button after giving the desired input.

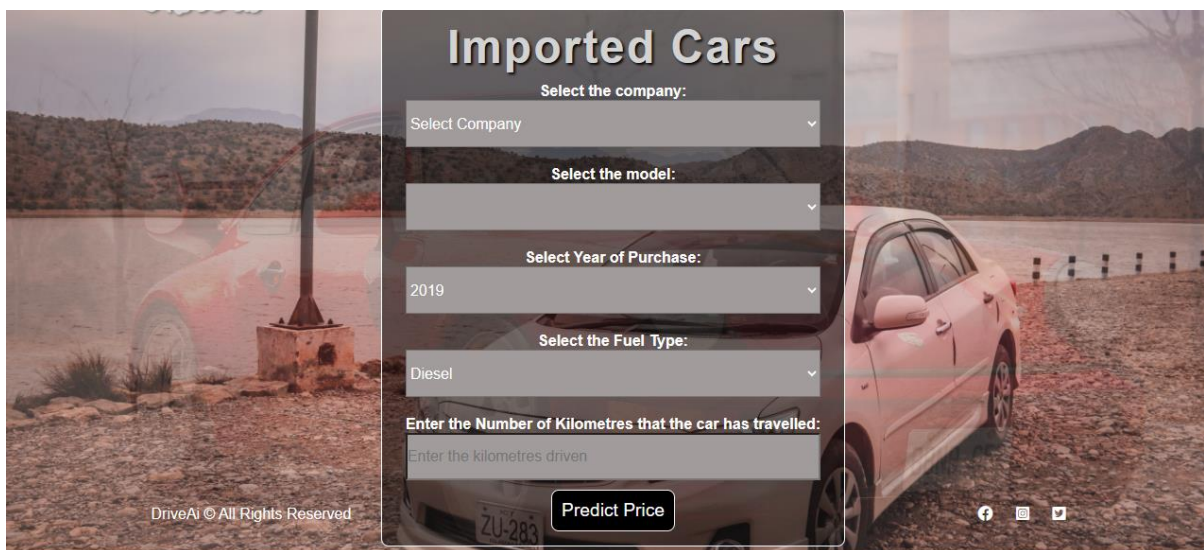


Image 7.2.1 Imported Car Model Before Input



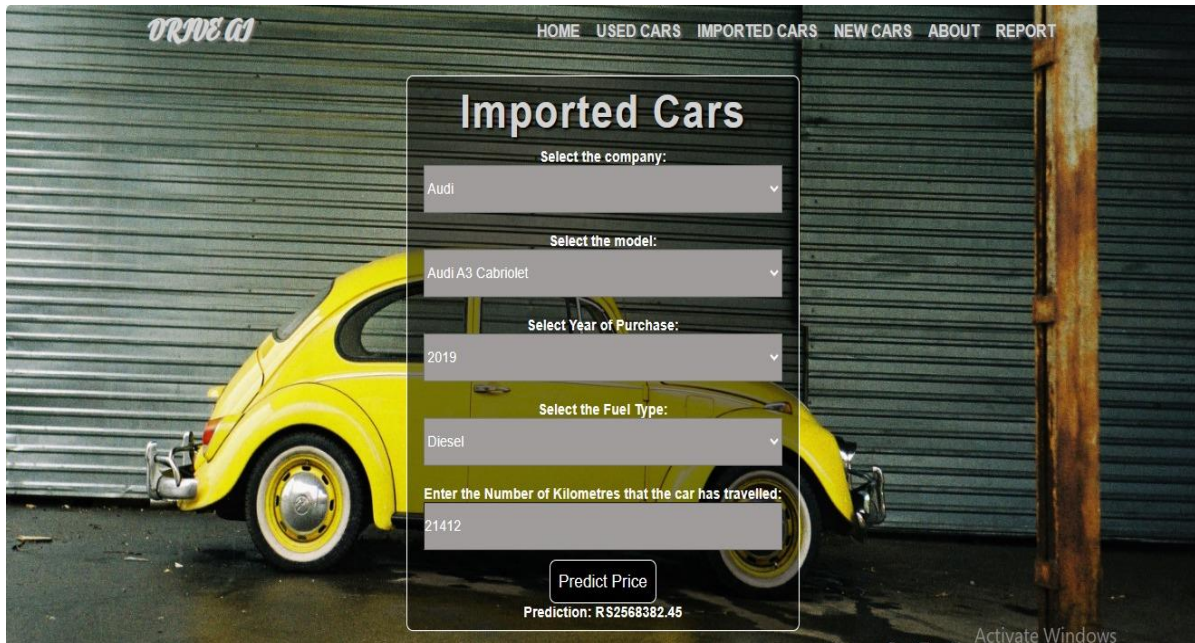


Image 7.2.2 Imported Car Model After Input

As it is seen after the user provide the input our model successfully predict the price.

### 7.1.3 New Car Models Results

Down below is the image of model before user input and after the user pressing the predict button after giving the desired input.

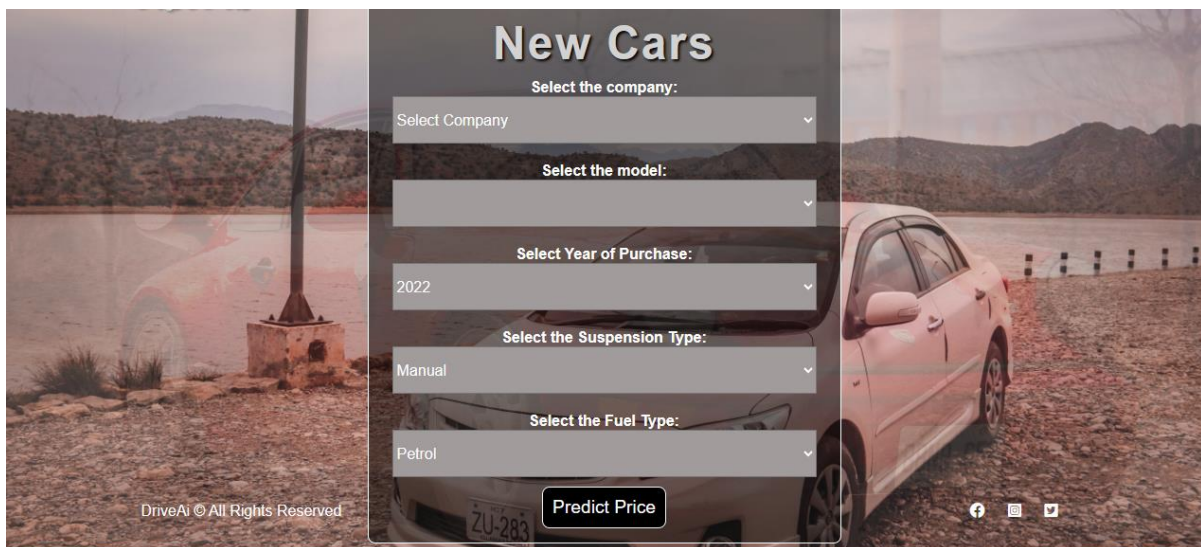


Image 7.3.1 New Car Model Before Input

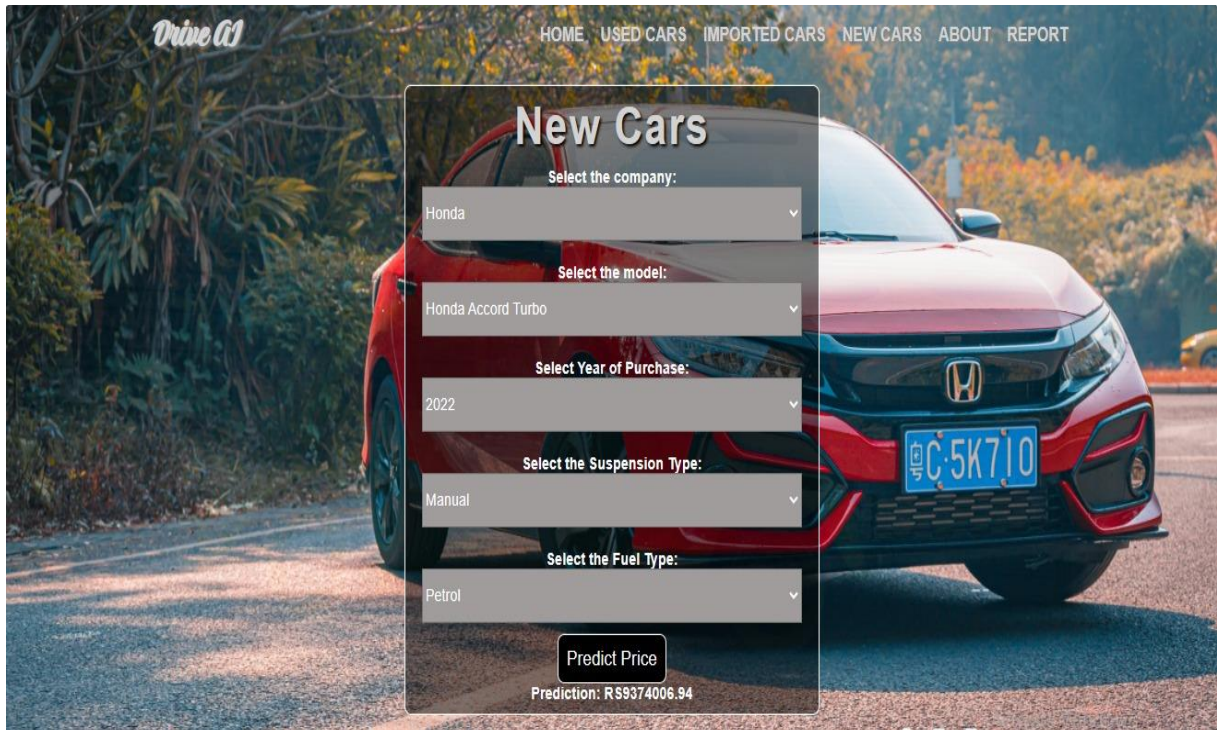


Image 7.3.2 New Car Model After Input

As it is seen after the user provide the input our model successfully predicts the price.

## **Chapter 8**

### **Conclusion of Research**

## 8.1 Conclusion

The data sets used in this web app can be very valuable in conducting similar research using different prediction techniques. The prices of vehicles can be predicted using this web app for more than one type of car. The data obtained under this research facilitated in prediction of prices of used cars, imported cars and new cars through linear regression method with an accuracy of more than 78 percent. Many assumptions were made on the basis of the data set. The proposed system evaluated variables and selected the most relevant variables out of the dataset and reduced the complexity of model by eliminating unrelated variables during processing and analysis phase. All these models are deployed on Flask which is the technology used for backend integration and used to link models with front end which is developed by using frontend design technologies using HTML, CSS and JavaScript.

## **Chapter 9**

### **Future Work in this Field**



## 9.1 Future Work

Our web app uses Machine Learning to predict the prices of the different types of cars. This feature can be further used to predict the prices of other type of automobiles like trucks, busses and motor bikes. The machine learning technique can be used to predict the prices of other commodities as well. Our web app can be further used and nourished into a complete car buying and selling web sites like Carfirst and Pakwheels and using filters for prices instead they can use our web app idea and use Machine Learning for that which is of higher accuracy Further more we can extend our app as a one stop solution for buying selling of cars ,getting information through blogs and news feed on our web site and analyzing market trends by just one click and we can also integrate AI chat bots to assist customers to increase Customer relationships.

## **Chapter 10**

### **References**

## 10.1 Reference

- 1- Pudaruth,S. 2014. “Predicting the Price of Used Cars Using Machine Learning Techniques”, International Journal of information & Computation Technology,4(7), p.753-764.
- 2- <https://towardsdatascience.com/predicting-car-price-using-machine-learning-8d2df3898f16>
- 3- <https://www.analyticsvidhya.com/blog/2021/07/car-price-prediction-machine-learning-vs-deep-learning/>
- 4- <https://medium.com/geekculture/used-carprice-prediction-complete-machine-learning-project-d25559cf2d2a>
- 5- Listiani M. 2009. Support Vector Regression Analysis for Price Prediction in a Car Leasing Application. Master Thesis. Hamburg University of Technology.