USED CAR PRICE PREDICTION USING RANDOM FOREST ALGORITHM

A Mini Project Report Submitted
In partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology
in
Computer Science and Engineering

by

KALALI MADHUSUDHAN GOUD	**20N31A0597**
KASALA SHREYA REDDY	**20N31A05A7**
M.SHIVAPRIYA REDDY	**20N31A05B9**

Under the esteemed guidance of

MRS.P.Honey Diana
Assistant Professor

Department of Computer Science and Engineering

Malla Reddy College of Engineering & Technology
(Autonomous Institution- UGC, Govt. of India)

(Autonomous Institution- UGC, Govt. of India)

(Affiliated to JNTUH, Hyderabad, Approved by AICTE, NBA &NAAC with ‘A’ Grade)
Maisammaguda, Kompally, Dhulapally, Secunderabad – 500100
website: www.mrcet.ac.in

2020-2024

Malla Reddy College of Engineering & Technology

(Autonomous Institution- UGC, Govt. of India)
(Affiliated to JNTUH, Hyderabad, Approved by AICTE, NBA &NAAC with 'A' Grade)
Maisammaguda, Kompally, Dhulapally, Secunderabad – 500100
website: www.mrcet.ac.in

CERTIFICATE

This is to certify that this is the Bonafide record of the project entitled "**USED CAR PRICE PREDICTION USING RANDOM FOREST ALGORITHM**", submitted by **KALALI MADHUSUDHAN GOUD(20N31A0597), KASALA SHREYA REDDY (20N31A05A7) and M.SHIVAPRIYA REDDY (20N31A05B9)** of B.Tech in the partial fulfillment of the requirements for the degree of Bachelor of Technology in Computer Science and Engineering, Department of CSE during the year 2023-2024. The results embodied in this project report have not been submitted to any other university or institute for the award of any degree or diploma.

Internal Guide	Head of the Department
Mrs.P.Honey Diana	**Dr. S. Shanthi**
Assistant Professor	**Professor**

External Examiner

DECLARATION

We hereby declare that the project titled "**USED CAR PRICE PREDICTION USING RANDOM FOREST ALGORITHM** " submitted to Malla Reddy College of Engineering and Technology (UGC Autonomous), affiliated to Jawaharlal Nehru Technological University Hyderabad (JNTUH) for the award of the degree of Bachelor of Technology in Computer Science and Engineering is a result of original research carried-out in this thesis. It is further declared that the project report or any part thereof has not been previously submitted to any University or Institute for the award of degree or diploma.

KALALI MADHUSUDHAN GOUD - (20N31A0597)

KASALA SHREYA REDDY - (20N31A05A7)

M.SHIVAPRIYA REDDY- (20N31A05B9)

ACKNOWLEDGEMENT

We feel honored to place our warm salutation to our college Malla Reddy College of Engineering and Technology (UGC-Autonomous) for giving us an opportunity to do this Project as part of our B.Tech Program. We are ever grateful to our **Director Dr. V.S.K Reddy** and **Principal Dr. S. Srinivasa Rao** who enabled us to have experience in engineering and gain profound technical knowledge. We express our heartiest thanks to our **HOD, Dr. S. Shanthi** for encouraging us in every aspect of our course and helping us realize our full potential.

We would like to thank our Project Guide **Mrs. P. Honey Diana** for her regular guidance, suggestions and constant encouragement. We are extremely grateful to our Project Coordinator **Mr. Sandeep Agarwalla** for his continuous monitoring and unflinching co-operation throughout project work. We would like to thank our Class Incharge **Mr. M. Sandeep** who in spite of being busy with his academic duties took time to guide and keep us on the correct path. We would also like to thank all the faculty members and supporting staff of the Department of CSE and all other departments who have been helpful directly or indirectly in making our project a success.

We are extremely grateful to our parents for their blessings and prayers for the completion of our project that gave us strength to do our project.

With regards and gratitude

KALALI MADHUSUDHAN GOUD - (20N31A0597)

KASALA SHREYA REDDY - (20N31A05A7)

M.SHIVAPRIYA REDDY - (20N31A05B9)

ABSTRACT

A car price prediction has been a high interest research area, as it requires noticeable effort and knowledge of the field expert. Considerable number of distinct attributes are examined for the reliable and accurate prediction. To build a model for predicting the price of used cars the applied three machine learning techniques are Random forest regression and linear regression. Respective performances of different algorithms were then compared to find one that best suits the available data set. The final prediction model was integrated into Java application. Furthermore, the model was evaluated using test data and the accuracy of 82% was obtained. Aims to build a model to predict used cars' reasonable prices based on multiple aspects, including vehicle mileage, year of manufacturing, fuel consumption, transmission, road tax, fuel type, and engine size. This model can benefit sellers, buyers, and car manufacturers in the used cars market. The model building process involves machine learning and data science. Various regression methods, including linear regression, decision tree regression, and random forest regression, were applied in the research to achieve the highest accuracy. The dataset was divided and modified to fit the regression, thus ensure the performance of the regression. To evaluate the performance of each regression, R-square was calculated. Among all regressions in this project, Linear Regression achieved the highest R-square of 0.90416. Compared to previous research, the resulting model includes more aspects of used cars while also having a higher prediction accuracy.

Keywords: Price, Car, Regression.

TABLE OF CONTENTS

S.NO	TITLE	PG.NO
1	**INTRODUCTION**	**01**
	1.1 PURPOSE AND OBJECTIVES	01
	1.2 EXISTING AND PROPOSED SYSTEM	03
	1.3 SCOPE OF PROJECT	04
2	**SYSTEM ANALYSIS**	**05**
	2.1 HARDWARE AND SOFTWARE REQUIREMENTS	05
	2.2 SOFTWARE REQUIREMENTS SPECIFICATION	06
3	**SYSTEM DESIGN**	**07**
	3.1 ARCHITECTURE	07
	3.2 UML DIAGRAMS	08
4	**METHODOLOGY**	**12**
	4.1 MODULE DESCRIPTION	12
	4.2 PROCESS/ALGORITHM	13
5	**IMPLEMENTATION**	**17**
	5.1 SAMPLE CODE	17
	5.2 OUTPUT SCREENS	23
6	**CONCLUSION & FUTURE SCOPE**	**25**
7	**BIBLIOGRAPHY**	**26**

LIST OF FIGURES

FIGURE.NO	NAME	PG.NO
3.1	SYSTEM ARCHITECTURE	07
3.2.1	CLASS DIAGRAM	09
3.2.2	USE CASE DIAGRAM	10
3.2.3	SEQUENCE DIAGRAM	10
3.2.4	ACTIVITY DIAGRAM	11
4.2.2	SUPERVISED LEARNING	14
4.2.3	UNSUPERVISED LEARNING	15
4.2.4	REINFORCEMENT LEARNING	16
4.2.5	ARCHITECTING THE MACHINE LEARNING PROCESS	16

LIST OF OUTPUTS

FIGURE.NO	NAME	PG.NO
5.2.1	DATA SET	23
5.2.2	TEST SET RESULT	23
5.2.3	HOME SCREEN	24
5.2.4	OUTPUT SCREEN	24

1. INTRODUCTION

Our Aim is to build a model to predict used cars' reasonable prices based on multiple aspects, including vehicle mileage, year of manufacturing, fuel consumption, transmission, road tax, fuel type, and engine size. This model can benefit sellers, buyers, and car manufacturers in the used cars market. Upon completion, it can output a relatively accurate price prediction based on the information that users input. The model building process involves machine learning and data science.

1.1 PURPOSE, AIM AND OBJECTIVES:

1.1.1 PURPOSE AND AIM:

Why car price prediction is important?

predicting the price of cars accurately is of great importance for buyers, sellers. For example, it can create relationship between seller and buyers for future products also it will definitely help understand current situations in an easy way.

And it is helpful for the people who are willing to buy car

learning is used for analyzing information based on past experience and predicting future performance.

1.1.2 OBJECTIVES:

knowing the cost of car based on past data

Helping to sellers and buyers for estimating the cost of cars

1.1.3 BACKGROUND OF PROJECT

The project background is a one-page section of your project proposal that explains the problem that your project will solve. You should explain when this issue started, its current state and how your project will be the ideal solution.

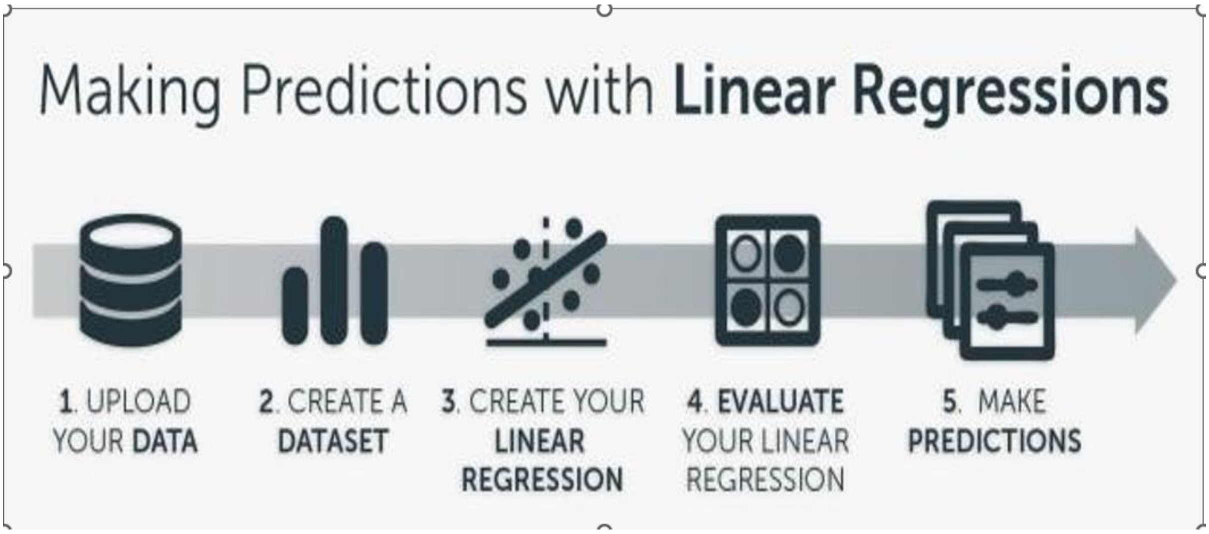


Fig : Background of Project

This is the project in which we can predict the future car prices based on the past data usingpython as the language and in that we have algorithms like:

Random Forest Regression

Linear Regression than processing and analysis data and train test split then

finally Evaluation.

1.2 EXISTING AND PROPOSED SYSTEM:

1.2.1 EXISTING SYSTEM

The existing system includes a basic car price prediction using the inputted data and displays its output on simple static website and there is no use of different kinds of regression algorithms. There is only use of simple linear regression algorithm in the project. The existing gives less accuracy in result. The below mentioned flowchart includes the behaviour of the existing system.

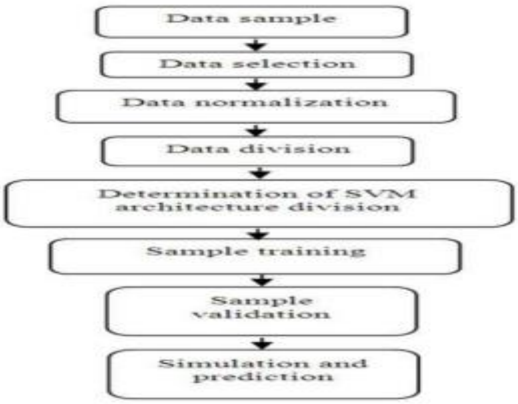


Fig:1.2.1 Existing System

1.2.3 PROPOSED SYSTEM:

We have used dataset and machine learning algorithms like multiple linear regression, random forest and gradient boosting for analyzing the relation between the variables and predict the gold price. We have used these machine learning algorithms to improve accuracy. Random forest machine learning algorithm gives the highest accuracy among all. algorithm. In the proposed System, we have tried to solve the issues in existing system. We have developed a GUI. Also we have used three machine learning algorithms, linear and random regression respectively in our system to predict the output more correctly with higher accuracy.

1.3 SCOPE OF PROJECT:

A dataset, as the name suggests, is a collection of data. In Machine Learning projects, we always need a dataset. Firstly, we need the training dataset to train our model, to help it predict. Then, we use testing datasets to predict and check how accurate is our model. ′ In this project, I have used the dataset available on Kaggle. One can find various such sites to download from.The dataset that I’ve used in my code was the data available on Kaggle. It

Ultimately,the success of the project is determined by the accuracy of predictions, the functionality of the user interface, and the positive impact on student support and outcomes.

2. SYSTEM ANALYSIS

2.1 HARDWARE AND SOFTWARE REQUIREMENTS:

2.1.1 HARDWARE REQUIREMENTS:

System	:	Intel 13 2.4 GHz. Or Advanced
Hard disk	:	512 MB
RAM	:	256 MB

2.1.2 SOFTWARE REQUIREMENTS:

Operating System	:	Windows/Linux
Coding Language	:	Python/Anaconda
Tools Used	:	Jupyter Notebook

2.2 SOFTWARE REQUIREMENT SPECIFICATION:

2.2.1 SRS:

Software Requirement Specification (SRS) is the starting point of the software developing activity. As system grew more complex it became evident that the goal of the entire system cannot be easily comprehended. Hence the need for the requirement phase arose. The software project is initiated by the client needs. The SRS is the means of translating the ideas of the minds of clients (the input) into a formal document (the output of the requirement phase.)

The SRS phase consists of two basic activities:

1)Problem/Requirement Analysis:

The process is order and more nebulous of the two, deals with understand the

problem, the goal and constraints.

2)Requirement Specification:

Here, the focus is on specifying what has been found giving analysis such as representation, specification languages and tools, and checking the specifications are addressed during this activity.

The Requirement phase terminates with the production of the validate SRS document. Producing the SRS document is the basic goal of this phase.

2.2.2 SCOPE:

This document is the only one that describes the requirements of the system. It is meant for the use by the developers, and will also be the basis for validating the final delivered system. Any changes made to the requirements in the future will have to go through a formal change approval process. The developer is responsible for asking for clarifications, where necessary, and will not make any alterations without the permission of the client.

3. SYSTEM DESIGN

System design is transition from a user oriented document to programmers or data base personnel. The design is a solution, how to approach to the creation of a new system. This is composed of several steps. It provides the understanding and procedural details necessary for implementing the system recommended in the feasibility study. Designing goes through logical and physical stages of development, logical design reviews the present physical system, prepare input and output specification, details of implementation plan and prepare a logical design walkthrough.

3.1 ARCHITECTURE:

Architecture diagram is a diagram of a system, in which the principal parts or functions are represented by blocks connected by lines that show the relationships of the blocks. The block diagram is typically used for a higher level, less detailed description aimed more at understanding the overall concepts and less at understanding the details of implementation.

The architecture of the "Medical Emergency App" project is designed to provide a seamless and secure experience for users seeking emergency medical assistance.

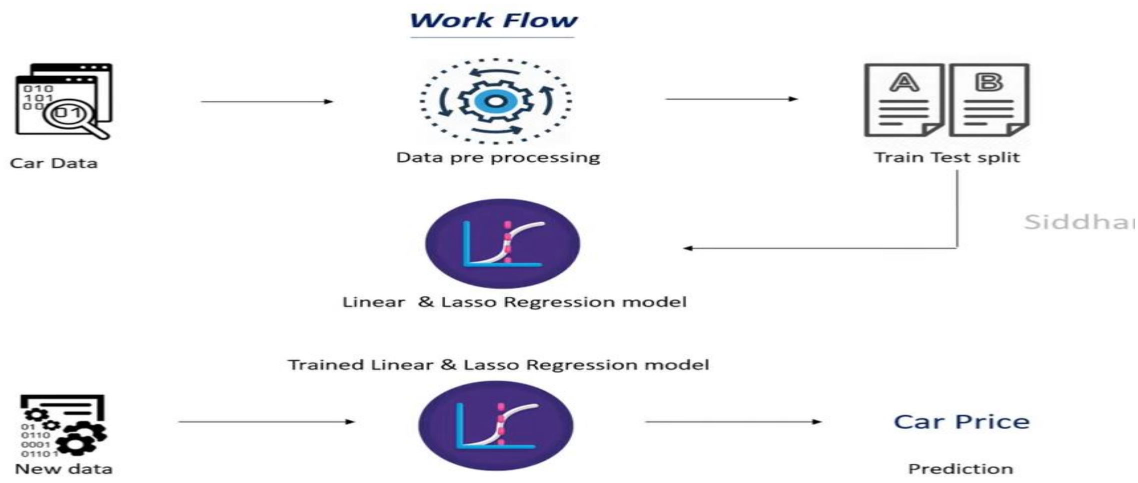


Fig : 3.1 : System Architecture

3.2 UNIFIED MODELING LANGUAGE (UML) :

The unified modeling is a standard language for specifying, visualizing, constructing and documenting the system and its components is a graphical language which provides a vocabulary and set of semantics and rules. The UML focuses on the conceptual and physical representation of the system. It captures the decisions and understandings about systems that must be constructed. It is used to understand, design, configure and control information about the systems.

Depending on the development culture, some of these artifacts are treated more or less formally than others. Such artifacts are not only the deliverables of a project; they are also critical in controlling, measuring, and communicating about a system during its development and after its deployment.

The UML addresses the documentation of a system's architecture and all of its details. The UML also provides a language for expressing requirements and for tests. Finally, the UML provides a language for modeling the activities of project planning and release management.

UML DIAGRAMS:

A use case is a description of how a person who actually uses that process or system will accomplish a goal. It's typically associated with software systems, but can be used in reference to any process. For example, imagine you're a cook who has a goal of preparing a grilled cheese sandwich. The use case would describe through a series of written steps how the cook would go about preparing that sandwich.

3.2.1 CLASS DIAGRAM:

A class is a representation of an object and, in many ways; it is simply a template from which objects are created. Classes form the main building blocks of an object-oriented application. Although thousands of students attend the university, you would only model one class, called Student, which would represent the represent the entire collection of students.

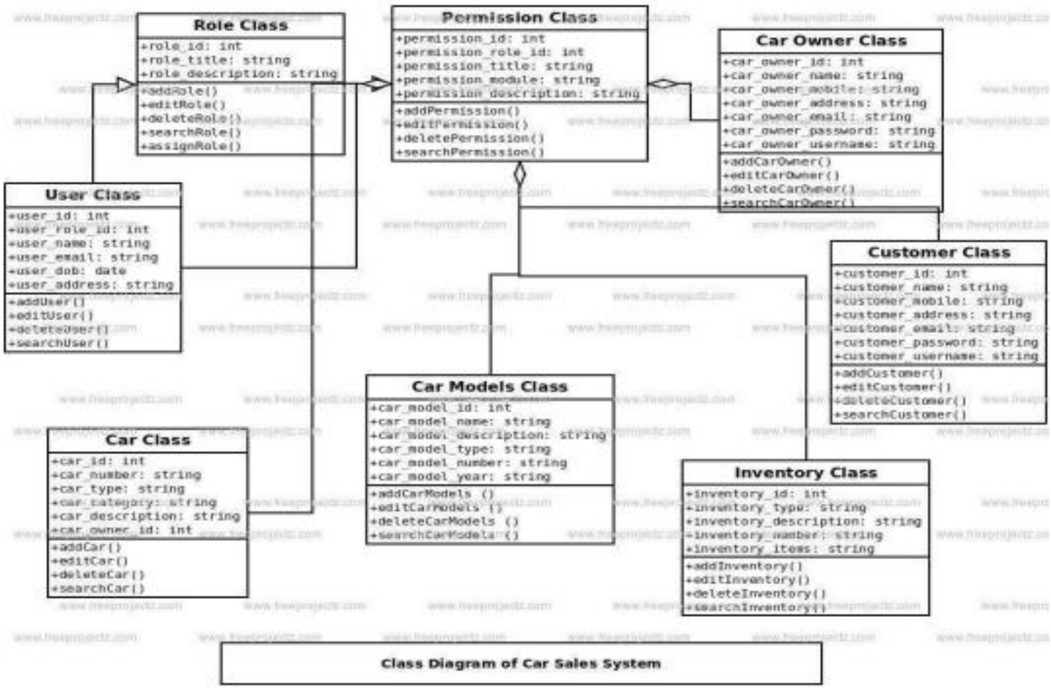


Fig 3.2.1 Class Diagram

3.2.2 USE CASE DIAGRAM:

A use case diagram is a graph of actors set of use cases enclosed by a system boundary, communication associations between the actors and users and generalization among use cases.

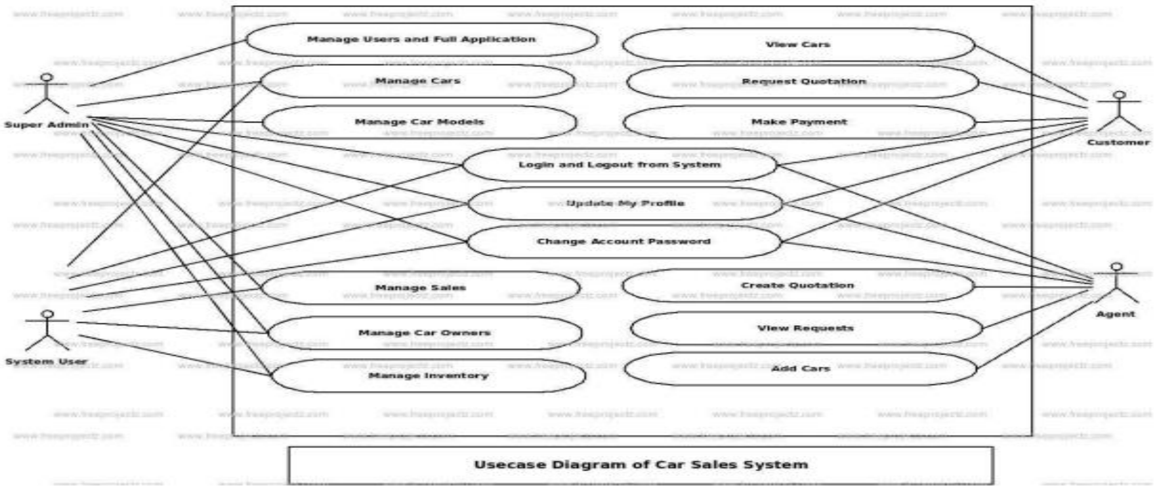


Fig 3.2.2 : Use Case Diagram

3.2.3 SEQUENCE DIAGRAM:

Sequence diagram are used to represent the flow of messages, events and actions between the objects or components of a system. Time is represented in the vertical direction showing the sequence of interactions of the header elements, which are displayed horizontally at the top of the diagram.

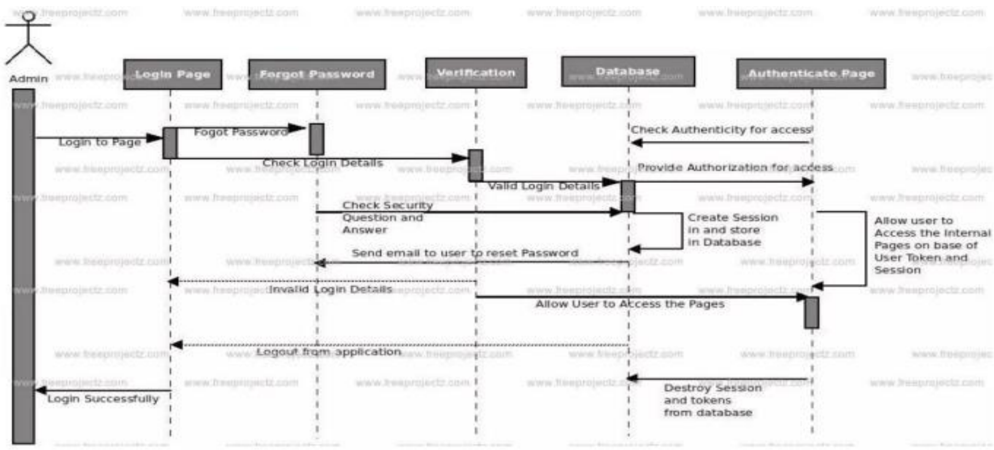


Fig 3.2.3 : Sequence Diagram

3.2.4 ACTIVITY DIAGRAM:

Activity diagram represent the business and operational workflows of a system. An Activity diagram is a dynamic diagram that shows the activity and the event that causes the object to be in the particular

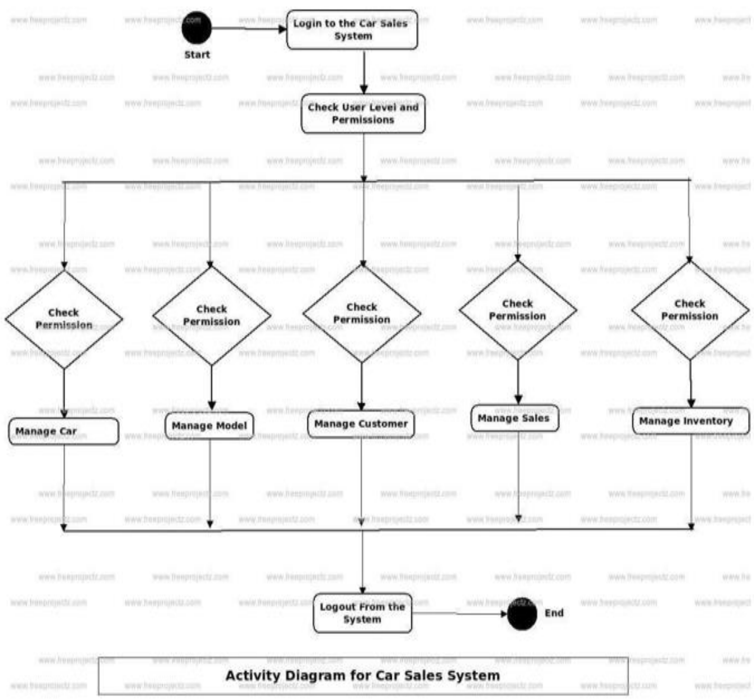


Fig 3.2.4: Activity Diagram

4. METHODOLOGY

4.1 MODULE DESCRIPTION:

1.Numpy: NumPy is a library for the Python programming language, adding support for large, multi- dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays.

2.Pandas: Pandas is a software library written for the Python programming language for data manipulation and analysis. It is built on top of another package named Numpy, which provides support for multi- dimensional arrays

3.Pickle : Pickle in Python is primarily used in serializing and deserializing a Python object structure. 22 In other words, it's the process of converting a Python object into a byte stream to store it in a file/database, maintain program state across sessions, or transport data over the network.

4.Scipy : SciPy is a scientific computation library that uses NumPy underneath.SciPy stands for Scientific Python. It provides more utility functions for optimization, stats and signal processing. Like NumPy, SciPy is open source so we can use it freely. SciPy was created by NumPy's creator Travis Olliphant. SciPy has optimized and added functions that are frequently used in NumPy and Data Science

5.Scikit : Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering

4.2 PROCESS/ALGORITHM:

4.2.1 MACHINE LEARNING :

Machine learning is a subfield of artificial intelligence that involves the development of algorithms and statistical models that enable computers to improve their performance in tasks through experience. These algorithms and models are designed to learn from data and make predictions or decisions without explicit instructions. There are several types of machine learning, including supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training a model on labeled data, while unsupervised learning involves training a model on unlabeled data. Reinforcement learning involves training a model through trial and error. Machine learning is used in a wide variety of applications, including image and speech recognition, natural language processing, and recommender systems.

Definition of learning: A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks T, as measured by P , improves with experience E.

Classification of Machine Learning

Machine learning implementations are classified into four major categories, depending on the nature of the learning “signal” or “response” available to a learning system which are as follows

4.2.2 Supervised learning:

Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs. The given data is labeled. Both classification and regression problems are supervised learning problems.

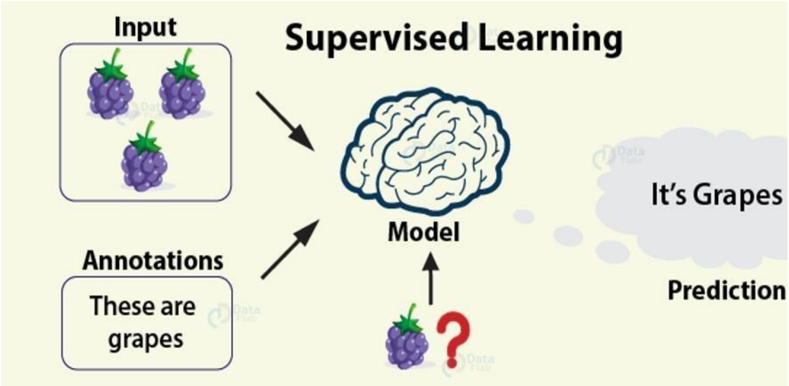


Fig 4.2.2 Supervised Learning

Linear Regression:

In linear regression, we measure the linear relationship between two or more than two variables. Based on this relationship, we perform predictions that follow this linear pattern.

Random Forest:

Random Forests are an ensemble learning method that is for performing classification, regression as well as other tasks through the construction of decision trees and providing the output as a class which is the mode or mean of the underlying individual trees.

Gradient Boosting

Gradient Boosting is an ensemble learning method that is a collection of several weak decision trees which results in a powerful classifier.

Support Vector Machine

SVMs are powerful classifiers that are used for classifying the binary dataset into two classes with the help of hyperplanes.

Logistic Regression

It makes use of a bell-shaped S curve that is generated with the help of logit function to categorize the data into their respective classes.

4.2.3 Unsupervised Learning

Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses. In unsupervised learning algorithms, classification or categorization is not included in the observations. Example: Consider the following data regarding patients entering a clinic. The data consists of the gender and age of the patients.

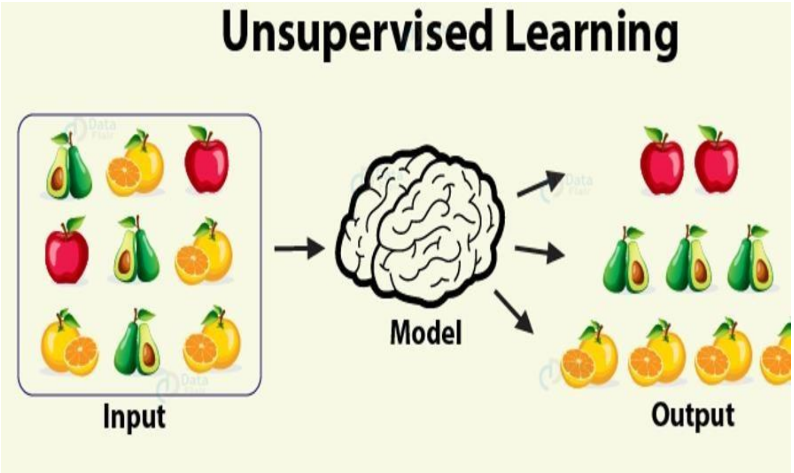


Fig 4.2.3: Unsupervised Learning

Clustering

Clustering, also known as cluster analysis, is a technique of grouping similar sets of objects in the same group that is different from the objects in other group. Some of the essential clustering techniques are as follows –

K-means : The aim of the k-means clustering algorithm is to partition the n observations in the data into k clusters such that each observation belongs to the cluster with the nearest mean. This serves as the prototype of the cluster.

4.2.4 Reinforcement Learning

Reinforcement learning is the problem of getting an agent to act in the world so as to maximize its rewards. A learner is not told what actions to take as in most forms of machine learning but instead must discover which actions yield the most reward by trying them. For example — Consider teaching a dog a new trick: we cannot tell him what to do, what not to do, but we can reward/punish it if it does the right/wrong thing.

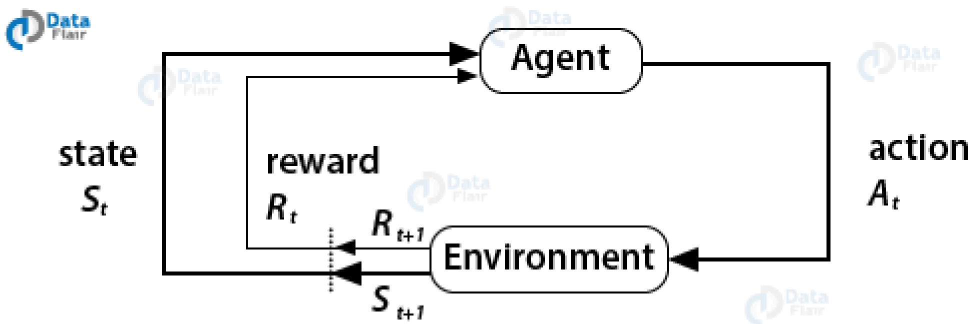


Fig 4.2.4: Reinforcement Learning

4.2.5 Architecture Of Machine Learning

Machine Learning architecture is defined as the subject that has evolved from the concept of fantasy to the proof of reality. As earlier machine learning approach for pattern recognitions has lead foundation for the upcoming major artificial intelligence program.

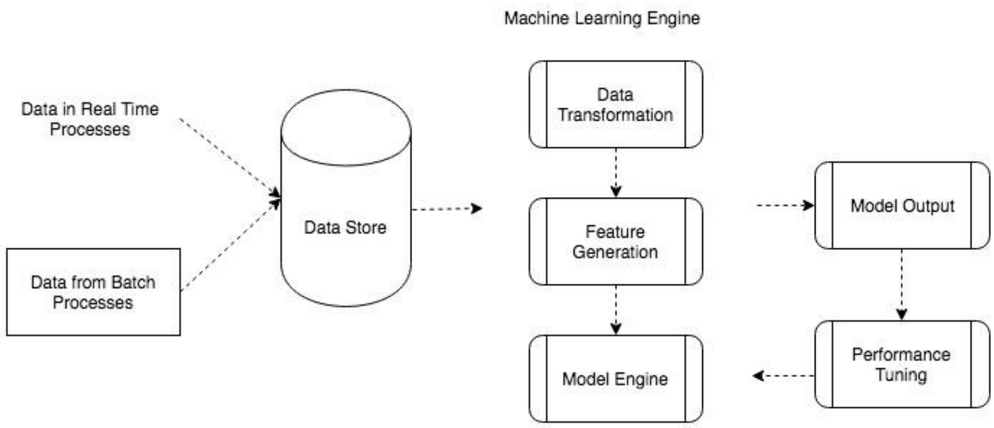


Fig 4.2.5 Architecting the Machine Learning Process

5. IMPLEMENTATION

5.1 SAMPLE CODE:

import pandas as pd df=pd.read_csv('car data.csv') df.shape print(df['Seller_Type'].unique()) print(df['Fuel_Type'].unique()) print(df['Transmission'].unique()) print(df['Owner'].unique()) ##check missing values df.isnull().sum() df.describe() final_dataset=df[['Year','Selling_Price','Present_Price','Kms_Driven','Fuel_Type','Seller_Type','Trans mission','Owner']] final_dataset.head() final_dataset['Current Year']=2020 final_dataset.head() final_dataset['no_year']=final_dataset['Current Year']-final_dataset['Year'] final_dataset.head() final_dataset.drop(['Year'],axis=1,inplace=True) final_dataset.head() final_dataset=pd.get_dummies(final_dataset,drop_first=True) final_dataset.head() final_dataset.head() final_dataset=final_dataset.drop(['Current Year'],axis=1) final_dataset.head() final_dataset.corr() import seaborn as sns sns.pairplot(final_dataset)

import seaborn as sns

#get correlations of each features in dataset corrmat = df.corr()

top_corr_features = corrmat.index plt.figure(figsize=(20,20))

#plot heat map

g=sns.heatmap(df[top_corr_features].corr(),annot=True,cmap="RdYlGn")X=final_dataset.iloc[:,1:]

y=final_dataset.iloc[:,0]

X['Owner'].unique() X.head() y.head()

Feature Importance from sklearn.ensemble import ExtraTreesRegressor import matplotlib.pyplot as plt

model = ExtraTreesRegressor() model.fit(X,y) print(model.feature_importances_)

#plot graph of feature importances for better visualization feat_importances = pd.Series(model.feature_importances_, index=X.columns) feat_importances.nlargest(5).plot(kind='barh')

plt.show()

from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=0) from sklearn.ensemble import RandomForestRegressor regressor=RandomForestRegressor() n_estimators = [int(x) for x in np.linspace(start = 100, stop = 1200, num = 12)]print(n_estimators)

from sklearn.model_selection import RandomizedSearchCV

#Randomized Search CV 32 # Number of trees in random forest n_estimators = [int(x) for x in np.linspace(start = 100, stop = 1200, num = 12)]

Number of features to consider at every split max_features = ['auto', 'sqrt'] # Maximum number of levels in tree max_depth = [int(x) for x in np.linspace(5, 30, num = 6)] # max_depth.append(None)

Minimum number of samples required to split a node min_samples_split = [2, 5, 10, 15, 100]

Minimum number of samples required at each leaf node min_samples_leaf = [1, 2, 5, 10]

Create the random grid random_grid = {'n_estimators': n_estimators,

'max_features': max_features,

'max_depth': max_depth,

'min_samples_split': min_samples_split,

'min_samples_leaf': min_samples_leaf}

print(random_grid)

Use the random grid to search for best hyperparameters

First create the base model to tune rf = RandomForestRegressor()

Random search of parameters, using 3 fold cross validation,

search across 100 different combinations rf_random = RandomizedSearchCV(estimator = rf, param_distributions = random_grid,scoring='neg_mean_squared_error', n_iter = 10, cv = 5, verbose=2, random_state=42, n_jobs = 1) rf_random.fit(X_train,y_train)

rf_random.best_params_ rf_random.best_score_ predictions=rf_random.predict(X_test)

sns.distplot(y_test-predictions)

plt.scatter(y_test,predictions) from sklearn import metrics print('MAE:', metrics.mean_absolute_error(y_test, predictions)) print('MSE:', metrics.mean_squared_error(y_test, predictions)) print('RMSE:', np.sqrt(metrics.mean_squared_error(y_test, predictions))) import pickle

open a file, where you ant to store the data file = open('random_forest_regression_model.pkl', 'wb')

dump information to that file pickle.dump(rf_random, file)

app.py from flask import Flask, render_template, request import jsonify import requests import pickle import numpy as np import sklearn from sklearn.preprocessing import StandardScaler app = Flask(__name__) model = pickle.load(open('random_forest_regression_model.pkl', 'rb'))

@app.route('/',methods=['GET']) def Home():

return render_template('index.html') standard_to = StandardScaler() @app.route("/predict", methods=['POST']) def predict(): 33

Fuel_Type_Diesel=0 if request.method == 'POST': Year = int(request.form['Year']) Present_Price=float(request.form['Present_Price'])

Kms_Driven=int(request.form['Kms_Driven']) Kms_Driven2=np.log(Kms_Driven)

```
Owner=int(request.form['Owner'])       Fuel_Type_Petrol=request.form['Fuel_Type_Petrol']
if(Fuel_Type_Petrol=='Petrol'):
Fuel_Type_Petrol=1 Fuel_Type_Diesel=0 else:
Fuel_Type_Petrol=0
Fuel_Type_Diesel=1                                                          Year=2020-Year
Seller_Type_Individual=request.form['Seller_Type_Individual']
if(Seller_Type_Individual=='Individual'):          Seller_Type_Individual=1          else:
Seller_Type_Individual=0   Transmission_Mannual=request.form['Transmission_Mannual']
if(Transmission_Mannual=='Mannual'): Transmission_Mannual=1
 else:
return    render_template('index.html',prediction_text="You    Can    Sell    The    Car    at
{}".format(output))
else:
return render_template('index.html') if __name__=="__main__":
app.run(debug=True)
```

Source code of html:

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Document</title>
</head>
```

```
<body>
 <div style="color:blue">
        <form action="{{ url_for('predict')}}" method="post">
            <h2>Predictive analysis</h2>
            <h3>Year</h3>
<input id="first" name="Year" type="number ">
            <h3>What is the Showroom Price?(In lakhs)</h3><br><input id="second"
name="Present_Price" required="required">
            <h3>How Many Kilometers Drived?</h3><input id="third" name="Kms_Driven"
required="required">
            <h3>How much owners previously had the car(0 or 1 or 3) ?</h3><br><input
id="fourth" name="Owner" required="required">
            <h3>What Is the Fuel type?</h3><br><select name="Fuel_Type_Petrol" id="fuel"
required="required">
                <option value="Petrol">Petrol</option>
                <option value="Diesel">Diesel</option>
                <option value="Diesel">CNG</option>
            </select>
            <h3>Are you A Dealer or Individual</h3><br><select
name="Seller_Type_Individual" id="resea" required="required">
```

```
            <option value="Dealer">Dealer</option>
            <option value="Individual">Individual</option>
          </select>
          <h3>Transmission type</h3><br><select name="Transmission_Mannual" id="research" required="required">
            <option value="Mannual">Manual Car</option>
            <option value="Automatic">Automatic Car</option>
          </select>
          <br><br><button id="sub" type="submit ">Calculate the Selling Price</button>
          <br>
        </form>
        <br><br><h3>{{ prediction_text }}<h3>
    </div>    <style>       body {           background-color: lightslategray;           text-align: center;           padding: 0px;
        }

        #research {           font-size: 18px;
    </style>
</body>
</html>
```

5.2 OUTPUT SCREENS

5.2.1 Data Set

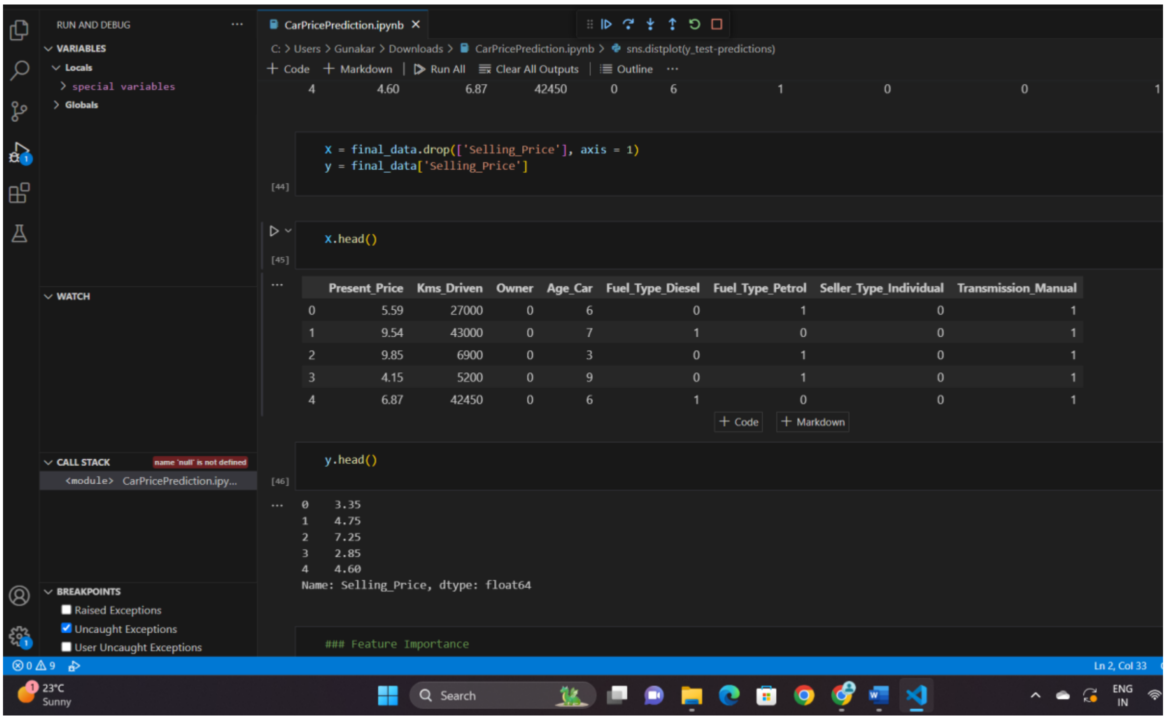
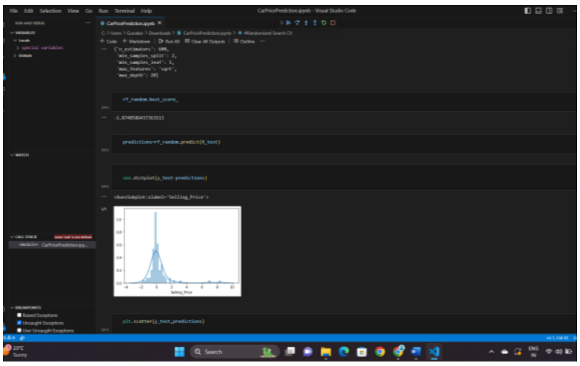


Fig 5.2.1: Data set

5.2.2 Test ResultSet:



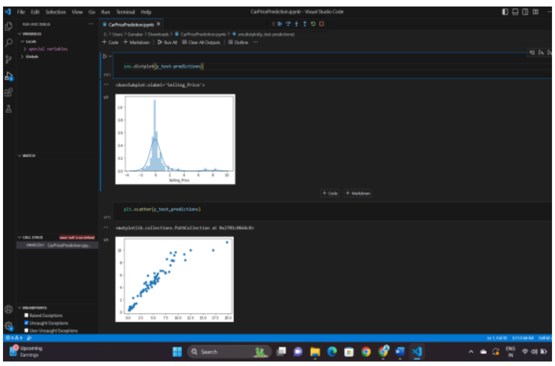


Fig:5.2.2.1 y_predict

Fig 5.2.2.2 : X_predict

5.2.3 Home Screen:

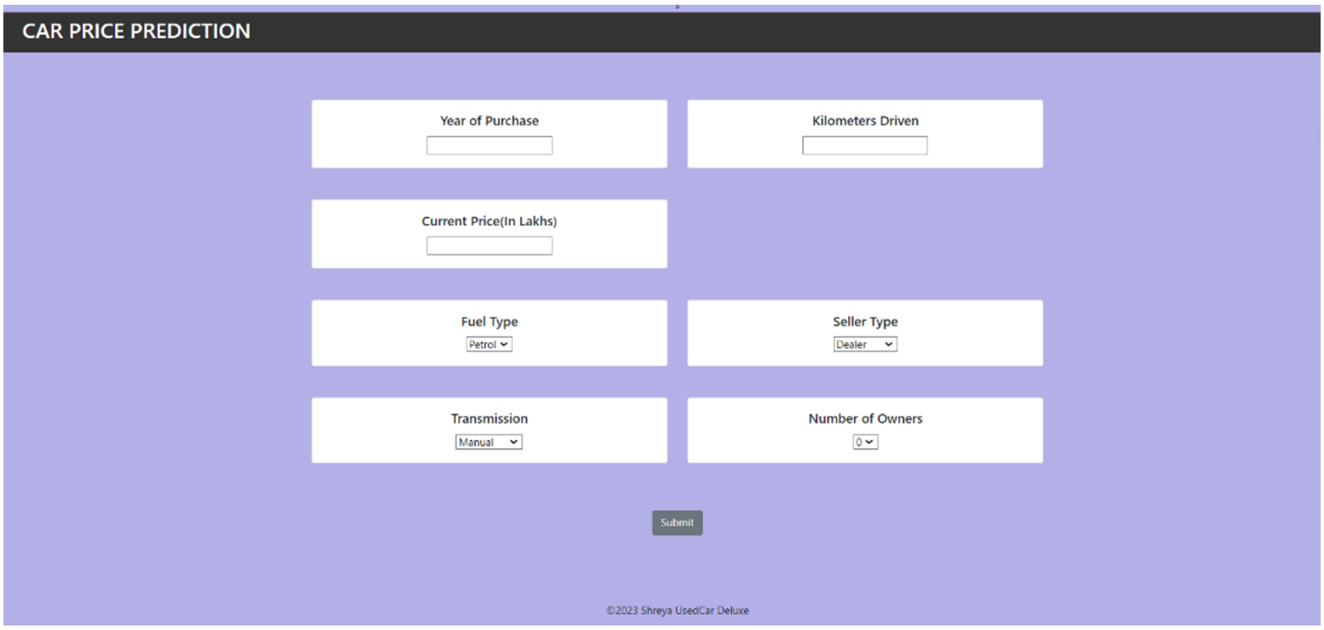


Fig 5.2.3: Home Screen

5.2.4 Output Screen:

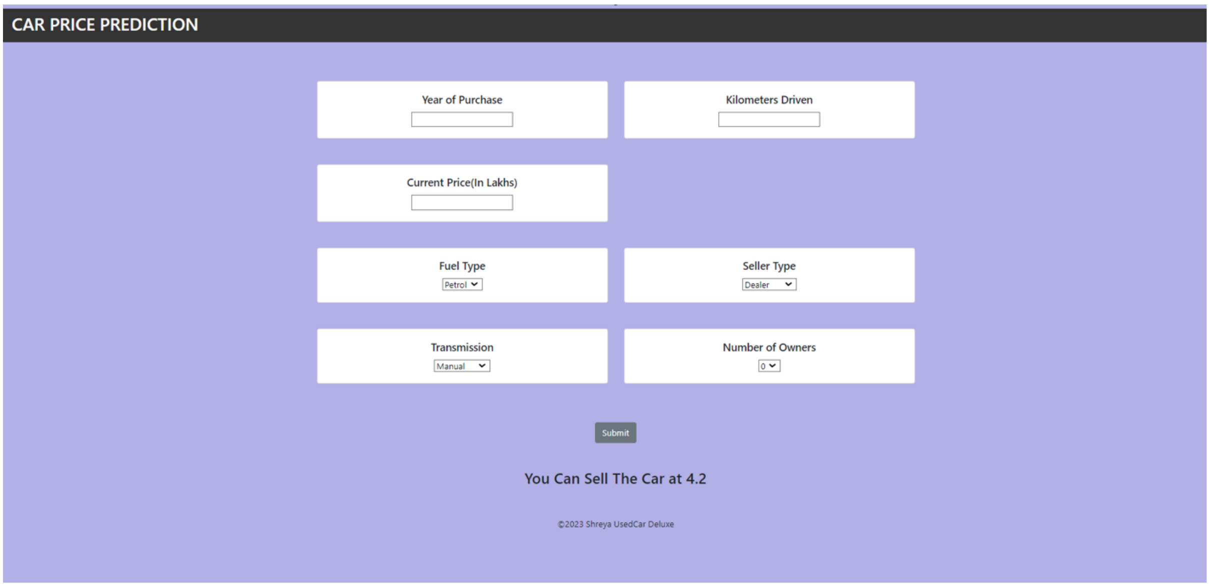


Fig 5.2.4 Output Screen

6. CONCLUSION & FUTURE SCOPE

6.1 CONCLUSION:

Used Car price prediction can be a challenging task due to the high number of attributes that should be considered for the accurate prediction. The major step in the prediction process is collection and preprocessing of the data. The increased prices of new cars and the financial incapability of the customers to buy them, Used Car sales are on a global increase. Therefore, there is an urgent need for a Used Car Price Prediction system which effectively determines the worthiness of the car using a variety of features. The proposed system will help to determine the accurate price of used car price prediction. This paper compares 2 different algorithms for machine learning : Random forest,Linear regression .

6.2 FUTURE SCOPE:

In future this machine learning model may bind with various website which can provide real time data for price prediction. Also we may add large historical data of car price which can help to improve accuracy of the machine learning model. We can build an android app as user interface for interacting with user. 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

7. BIBLIOGRAPHY

- Sameerchand Pudaruth, “Predicting the Price of Used Cars using Machine Learning Techniques”;(IJICT 2014)
- Enis gegic, Becir Isakovic, Dino Keco, Zerina Masetic, Jasmin Kevric, ”Car Price Prediction Using Machine Learning”; (TEM Journal 2019)
- Ning sun, Hongxi Bai, Yuxia Geng, Huizhu Shi, “Price Evaluation Model In Second Hand Car System Based On BP Neural Network Theory”; (Hohai University Changzhou, China) ′ Nitis Monburinon, Prajak Chertchom, Thongchai Kaewkiriya, Suwat Rungpheung, Sabir Buya, Pitchayakit Boonpou, “Prediction of Prices for Used Car by using Regression Models” (ICBIR 2018)
- Doan Van Thai, Luong Ngoc Son, Pham Vu Tien, Nguyen Nhat Anh, Nguyen Thi Ngoc Anh, “Prediction car prices using qualify qualitative data and knowledge-based system” (Hanoi National University)

- Building A Car Price Prediction Model Using Machine Learning - (analyticsvidhya.com) Basic Concepts in Machine Learning - Javatpoint Machine Learning - GeeksforGeeks Python Machine Learning (w3schools.com)