SYNOPSYS

ON

**“USED CAR PRICE PREDCTION”**

For

PG Diploma in Big Data Analytics

(PG-DBDA)

**Submitted by**

Prajakta Waykos (2220310125014)

Balaji Kamble (220310125001)

Mansi Patil (2203101225030)

**Under the guidance of**

Sh. Pappu Dindayal Kapgate



ACTS-CDAC, NEW DELHI

March 2022

**Abstract**

As the automobile industry is booming, the demand for new cars is increasing. With the demand for new cars, used cars have also gained tremendous value. Estimating the price of these used cars has become crucial by which we can provide the actual value of the car to the buyers and also to the sellers for transparency.

**Table of contents**

|  |  |
| --- | --- |
| **Title** | **Page no** |
| Abstract |  |
| 1. Background |  |
| 2. Problem Statement |  |
| 3. Aim and Objectives |  |
| 4. Significance of the Study |  |
| 5. Scope of the Study |  |
| 6. Research Methodology |  |
| 7. Requirements Resources |  |
| 8. Research Plan |  |
| Reference |  |

**List of figures**

|  |  |
| --- | --- |
| **Sr. No.** | **Name of figure** |
| 6.1 | Engine parts |
| 6.2 | Data Transformation |

**List of tables**

|  |  |
| --- | --- |
| **Sr. No.** | **Name of Table** |
| 7.1 | Software tools and python libraries |

**1. Introduction**

The used car market in 2019 is valued at $ 828.24 billion and is expected to be worth $ 1.355 billion by 2027. During the 2013-2018 evaluation period, India's used car market experienced double-digit CAGR revenue growth. The market was in the early stages of expansion due to higher vehicle replacement rates, faster time to market for new cars, growth in the middle class, higher average ticket sizes, and lower import taxes on new cars. Older cars are preferred in India because new cars are expensive and popular among people in the socio-economic group below the middle class. Also, renting a car is becoming a widespread practice, and rather than purchasing a car outright, you can do so by making fixed monthly payments for a predetermined number of months. You have the option to purchase the vehicle at the residual value, or the anticipated selling price, at the end of the lease. To make it easier for people to buy a car, we have developed a predictive model that predicts the price of used cars as demand increases.

**2. Problem Statement OR Related Research OR Related Work**

The 17th century industrial revolution started from the British colony. Automobile industry started developing and increasing demand. As more manufacturing competition increased according to demand. But as technology gets developed, old transportation gets avoided by the consumers.

Used car prices are based on Brand, Manufacturer and Transmission type. This process done

By professionals who understand the condition and market demand of cars. Condition of cars depends on the engine, body of cars, average and condition of cars.

But due to the increase in price of new cars and the inability of customers to buy new cars due to lack of funds, used car sales are on a global increase. There is a need for a used car price prediction system to effectively determine the worthiness of the car using a variety of features.

**3. Aim and Objectives**

The main aim of this project is to determine the price of used cars using professionals’ evaluation. Predict the price of a used car using various machine Learning Models. This helps the customer to make decisions based on different input or factors such as Price, brand, Condition of vehicles.

The objectives of the research are outlined as follows:

* To investigate the modern approach to determine the used car price.
* To determine the most favourable technique to estimate the parameters for learning models.
* To predict human capitalism with high accuracy using machine learning models.
* To evaluate performance of various supervised models
* To improve automobile economy

**4. Significance of the Study**

We have developed a system to accurately predict the price of used cars. The problem of predicting used car prices is valuable because various studies show that the used car market will continue to grow in the short term. The manufacturer sets the price of a new car in the industry, with the government incurring some additional costs in the form of taxes. As a result, customers purchasing a new car can be confident that the money they invest will be worthwhile. Due to rising new car prices and customers' financial inability to purchase them, used car sales are on the rise globally. As a result, there is an urgent need for a Used Car Price Prediction system that accurately determines the worthiness of the vehicle based on a variety of features.

**5. Scope of the Study**

The main aim of this project is to determine the price of used cars using professional evaluation. Predict the price of a used car using various machine Learning Models. This helps the customer to make decisions based on different input or factors such as Price, brand, and condition of vehicles.

We offer a Machine Learning-based methodology for predicting the prices of secondhand cars based on their characteristics. The cost is calculated using the amount of characteristics. Then, to illustrate our findings, we construct a responsive website that includes all of the countless used car listings. Our efforts culminated in this deployed service, which integrates data, machine learning, and features. This methodology can assist consumers looking to purchase a used car in making more informed judgments. Customers can now look for all automobiles in a region without physical efforts, anytime and from any location.

We used linear regression and lasso regression to develop a price model for used automobiles in a comparative research. The main goal of this study is to discover the best predictive model for estimating the price of a used car.

**6. Research Methodology**

In this section, the research has been conducted with several setups to determine the price of used cars by employing the supervised learning model.

**6.1 Introduction**

The used automobile market is a growing business with a market value that has nearly doubled itself in previous years. The rise of online websites and other tools like it have made it easier for both buyers and sellers to get a better understanding of the factors that determine the market value of a used car. Based on a set of factors, Machine Learning algorithms may be used to forecast the price of any automobile. The data set will include information on a variety of automobiles. There will be information regarding the vehicle's technical elements, such as the engine type, fuel type, the kilometres per litre, and more, for each car.

[REF: 4Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology, AKTU]

**6.2 Dataset description**

In this project data of used cars will be used. Dataset contains car name, fuel type, previous car price, door numbers, car body and location of engine. Another criteria of data are wheelbase of the car then length and width of car and cylinder numbers are considered.

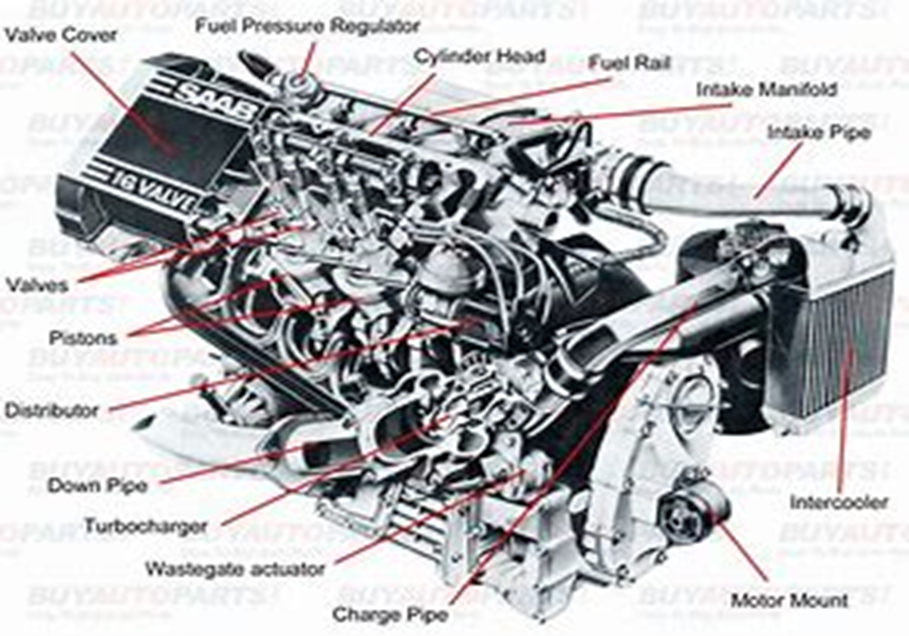


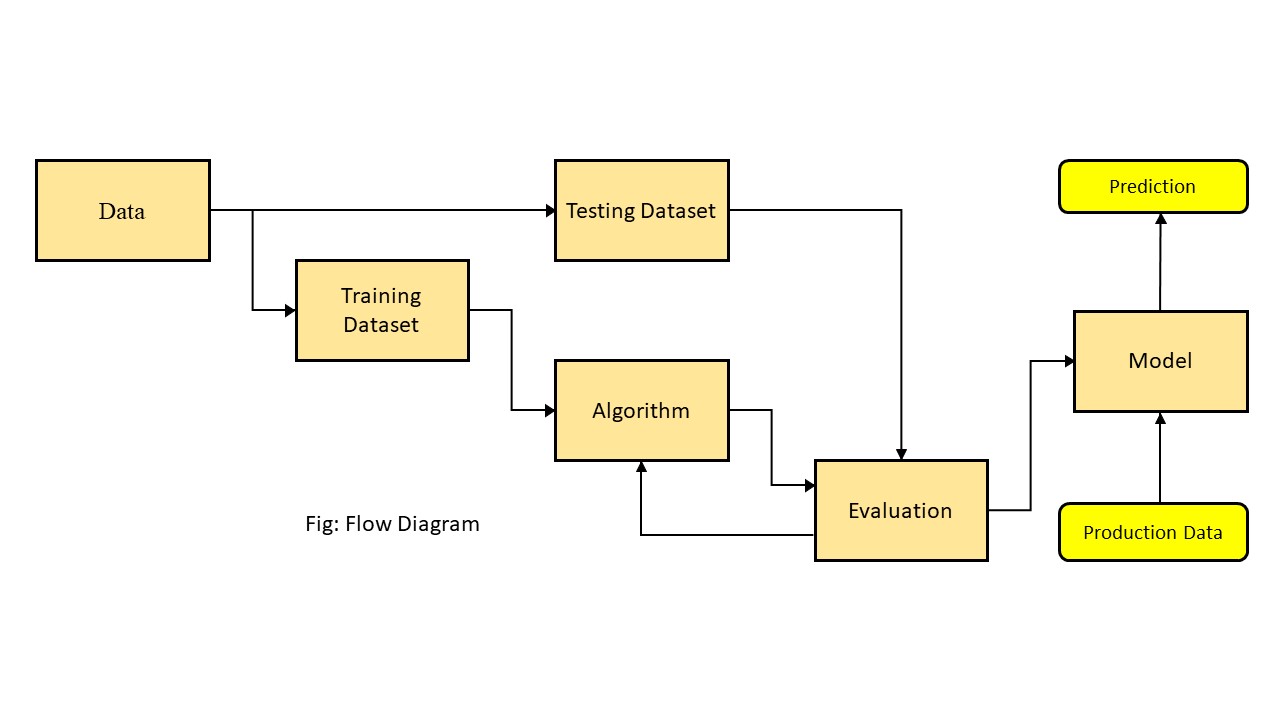
Fig 6.1. Engine parts

**6.3 Data Pre-processing:**

The given dataset is already available in the pre-processed form. The pre-processing has been done within excel, NumPy, panda’s applications. In our case, we will use the pre-process dataset which is in .csv format for the project.

**6.4 Transformation**

In this section, to obtain the predicted car price for a used car can be determined using machine learning models. A detailed explanation of each step required in the transformation process has been explained.



**7. Requirements Resources**

In this section, the required hardware and software tools have been mentioned. To conduct this study, the available hardware and software are also mentioned.

**7.1 Hardware Requirement**

In this study, we need the following hardware specifications to perform the task.

• CPU: 2

• CPU Model Name: AMD EPYC 7742 64-core processor

• Memory: 64 GB (min)

• Storage: 200 GB (min)

**7.2 Software Requirement**

The below table describes the required software and python library requirement to complete this study.

Table 7.1 Software tools and python libraries

|  |  |  |
| --- | --- | --- |
| Sr. No. | Package Name | Version |
| 1 | Operating System | 20.0.4.2 LTE |
| 2 | Power-bi | 2.106.582.0 |
| 3 | Mysql | 8.0.29 |
| 4 | Python | 3.10. 0 |
| 5 | Anaconda | 1.10. 0. |
| 6 | Jupyter lab | v6. 4.11 |
| 7 | Numpy | numpy 1.23.0 |
| 8 | Pandas | 1.4.3 |
| 9 | Scipy | 1.8.0 |
| 10 | Sklearn | 0.14.5 |
| 11 | Matplotlib | 3.5.2 |
| 12 | Seaborn | 0.11. 0 |