



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 6, June 2022

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.118

9940 572 462

6381 907 438

[ijirset@gmail.com](mailto:ijirset@gmail.com)

[www.ijirset.com](http://www.ijirset.com)

# A Review on Car Fare Forecast

Rashmi S Manvi, Ms.Ashwini.C

MCA Student, UBTD College of Engineering, Davanagere, Karnataka, India

Dept. of Master of Computer Applications, UBTD College of Engineering, Davanagere, Karnataka, India

**ABSTRACT:** The demand for used cars has increased significantly in past decade, and it has doubled its value in last few years. This has given rise to the used car market, which on its own has become a profitable industry. This paper aims to build a model to predict used cars at reasonable prices based on multiple features, including year of manufacture, vehicle mileage, fuel consumption, fuel type, engine size, transmission and road tax. Due to the increased prices of new cars and financial incapability of the customers to buy them, used car sales are gradually increasing. The recent online portal has ease the need for both customer and the seller to be better informed about the current and patterns that determine the value of used in the market. Using machine learning algorithms such as Multiple Regression, Lasso Regression and Regression Trees, we will try to develop a statistical model which will help to forecast the used car price based on earlier customer data and a given set of reviews. We will also be comparing the prediction efficiency of these models to determine the best one. User interface has also been developed which receives the input from any customer and displays the price of a car according to customer's input.

## I. INTRODUCTION

The used car market is a booming industry, which has enlarged its market value in last few years. According to the recent report on India's pre-owned car market by Indian Blue Book, nearly 4 million used cars were purchased and sold in 2018-19. The focus of this project is developing machine learning models that accurately predict the price of a used car based on its features, in order to make informed purchases. The arrival of online portals such as Quikr, CarWale, OLX and many other has facilitated the need for customer and the seller to be better informed about the trends and patterns that forecast the value of used car in the market.

An Accurate estimation of car prices requires specialized expertise, because price usually depends on different features and variables. Most people prefer to buy the used cars because of Cost-Effective, and they can resell that again after some years of usage which may get some profit. Distinct features like door number, type of transmission, exterior color, safety, dimensions, Interior whether it has navigation or not also influence the car price.

Different websites have distinct algorithms to generate the market price of the used car and hence there isn't a combined algorithm for determining the price. In this paper we have used different methods and techniques in order to obtain higher fineness of used car price prediction.

### 1.1 Objectives:

The objective of our system is to achieve good accuracy and to develop an effective and efficient model which forecasts the price of used car according to the user input.

Our main objective here is to develop a user Interface which is user-friendly and takes the input from the customer and predicts the price.

## II. LITERATURE SURVEY

Over fitting and Under fitting come into picture when we design our statistical models. The models might be too partial to the training data and might not perform best on the test data. This is known as over fitting. Similarly, the models might not take into consideration all the variance present in the population and perform poorly on test data set. This is known as under fitting. A perfect balance need to be executed between these two, which leads to the concept of Bias-Variance trade-off. In this paper, the price evaluation model based on the big data analysis is proposed, which takes advantage of widely circulated vehicle data and numerous of vehicle transaction data for each type of vehicle by using the optimized BP neural network algorithm. One can list their used car for sale along with details like image and



year of purchase and kilometres so that it is displayed to lakhs of interested car buyers in the web sites. It aims to establish a second-hand car price evaluation model to get the price that best matches the car

Considerable number of different attributes are examined for the decisive and accurate prediction. To construct a model for predicting the price of second –hand cars in Herzegovina and Bosnia ,they have used three machine learning techniques(Artificial Neural Network, Random Forest and Support Vector Machine).

### III. SYSTEM IMPLEMENTATION

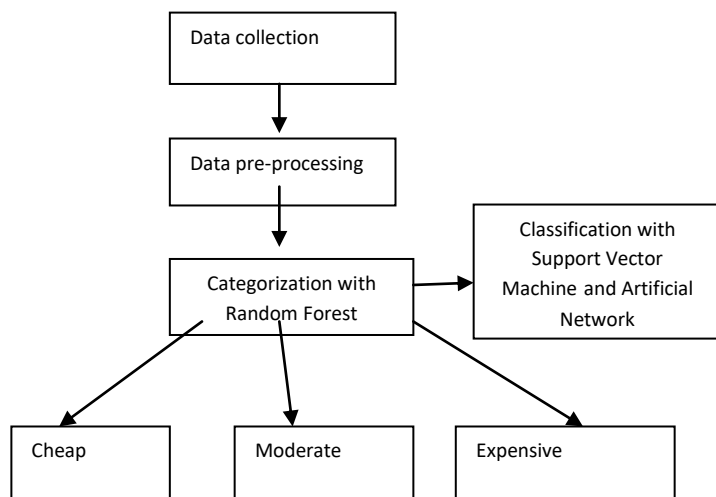
Single machine learning classifier approach that has been used in all previous researches was also tested in this research .The collected data is split into testing and training subsets. They are Support Vector Machine and Random Forest Classifier and Artificial Neural Network models were build.

#### 3.1 Existing System

The price of the car should be known for sale or purchase of old car, we need to meet the agent to know the car in existing system .We need to pay for the agent in the existing system to know the market price of the car. It may take long time to predict the price. There may be chances of miscalculation in the agent guessing

#### 3.2. Proposed System

Based on the different features and factors ,and also with the help of specialized expertise, prediction of price is done accurately. The most important element for prediction is model and brand, vehicle usage period ,vehicle mileage. The fuel consumption per mile and fuel type used in vehicle affect the price of vehicle due to a constant change in the price of fuel. In order to achieve higher precision we applied different methods and techniques of used car price prediction.



**Figure:** Block diagram of the overall classification process

Data collection is made from local web portal for buying and selling cars.

The following attributes were used for each car: Model, Brand, Fuel ,Year of manufacturing, number of doors, transmission type, mileage, navigation(yes/no),damaged(yes/no),leather(yes/no),parking sensor(yes/no), alarm(yes/no),digital air condition(yes/no)

### IV. METHODOLOGY

There are two primary phases in the system, they are Training phase and Testing phase.

Training phase: The data in data set is used in system training and fits a model based on algorithm that is chosen

accordingly.

Testing phase: The inputs will be provided in the system and is tested for working.

The accuracy is examined. The data that has to train the model or test, has to be appropriate. Appropriate algorithms have to be used in the system to detect and predict the price of the used car. Different algorithms were compared for its accuracy, before the algorithms are selected for further use.

#### 4.1. The Client

To be ready to predict the value of used cars can help both buyers and sellers.

**Used car sellers:** They are one in every of the most important target group which will have an interest in results of this study. If used car sellers better understand what makes a car desirable, what the significant features are used for a car, then they will consider this information and offer an improved service.

**Online pricing services:** There are websites that give an estimate value of a car. They may have good prediction model. However, having a second model may help them to relinquish a more robust prediction to their users. Therefore, the model developed during this study may help online web services that tell a second user car's market price.

**Individuals:** There are many individuals who have an interest within the used car markets at some points in their life because they wanted to sell their car or buy a second hand car during this process, it's a giant corner to pay an excessive amount of or sell less than it's market price

## V. RELATED WORK

Predicting price of a second hand cars has been studied extensively in various researches. Listian discussed, in her paper written for Master thesis, that regression model that was built using Support vector Machine(SVM) can predict the value of a car that has been leased with better precision than multivariate regression or some simple multivariate correlation this might get on the grounds that Support Vector Machine(SVM) is best in handling datasets with more dimension it's vulnerable to over fitting and under fitting. The weakness of this research is that a change of easy regression with more advanced SVM regression wasn't shown in basic indicators like mean, variance.

Another approach was given by Richardson applied multiple correlation analysis and demonstrated that hybrid cars retain their value for extended time than traditional cars. This has the roots in environmental concerns about the climate and it gives higher fuel efficiency. Conducted car price prediction study, by using neuro-fuzzy knowledge system. They look into consideration the subsequent attribute: year of production, brand and sort of engine. Their prediction model produced similar results because the regression model. Moreover, they made an expert system named ODAV(optimal distribution of Auction Vehicles) as there's a high demand for selling the car at the top of the leasing year by car dealers.

Regression model supported k-nearest neighbour algorithms was accustomed predict the worth of a car. This technique features a tendency to be exceptionally successful since quite two million vehicles were exchanged through it.

Gonggie proposed a model that's built using ANN(Artificial Neural Network) for the worth prediction of a second hand car. He considered several attributes miles passed, estimated car life and brand. The proposed model was built, so it could pander to nonlinear relations in data which wasn't the case with previous models that were utilizing the straightforward statistical regression techniques.

Furthermore, Pudaruth applied various machine learning algorithms, namely k-nearest neighbor, multiple simple regression analysis, decision trees and naive bayes for car prediction in the Mauritius. The data set went to create a prediction model was collected manually from the local newspaper in period but one month, as time can have a clear impact on price of car.

He studied the subsequent attributes: brand, model, cubic capacity, mileage in kilometres, production year, exterior color, transmission type and price. However, the author revealed that Naive Bayes and Decision Tree were unable to predict and classify numeric values. However, it's noticeable that single machine learning algorithm approach failed to give remarkable results and will be enhanced by assembling various machine learning methods in an ensemble.

## VI. DATASETS



We used an open dataset to coach our model, to predict the value of used car accurately. The dataset contains the costs and attributes of over 450,000 used cars sold on the web site across 40 brands. Our data set contains 20 unique attributes car being sold, out of which some columns will be removed that have little to no impact on car's price from our analysis, like as postal code, pictures and name of advertisement.

The following attributes are considered:

**Price:** The calculated retail price of cars. The data set selected for car were all but a year old and were considered to be in best condition.

**Mileage:** Total number of miles travelled by the car.

**Model:** The unique models for every car.

**Doors :**Total number of doors in car

**Color :** The outside color of the car

**Cylinder:** Number of cylinder present within the car

**Sound :**A categorical variable ,that represents whether upgraded speakers are present within the car

## VII. FUTURE SCOPE

In future Machine learning model may attach the various website which can provide real time data for prediction of price . Android app can be build as user interface ,which makes easy for interacting with user. To improve the accuracy of machine learning model we may add large historical data of car price .Rather than using old data set, we use adaptive learning rates and train on cluster of data and judiciously design deep learning network structures for better performance.

## VIII. CONCLUSION

Prediction of car price can be difficult task due to higher number of attributes that should be considered for the accurate prediction. Data collection and pre-processing is the major step in prediction. We have used different types of machine learning algorithms to obtain more accuracy ,and it is achieved 93 percent of efficiency. This ability of single and multiple groups of machine learning algorithm is notable.

To increase the prediction performance Data cleaning processes is used ,it is insufficient for the cases of complex data sets as the one in this research. Single machine algorithm is applied on the data set accuracy was less than 50%. Therefore, the ensemble of multiple machine learning algorithms has been proposed and machine learning methods of this combination gains accuracy of 92.38%.

## REFERENCES

1. Pudaruth ,S.(2014).Predicting the price of used car using machine learning techniques .Int.J.Inf. comput.Technology,4(7),753-764
2. Used car database.(n.d.)Retrieved from: <https://www.kaggle.com/orgesleka/used-cars-database>.
3. Enis gegic,Becir Isakovic,Dino Keco,Zerina Masetic,Jasmin Kevric,"Car Price Prediction Using Machine Learning"Agencija za statistiku BiH.(n.d.) ,retrieved from :<http://www.bhas.ba>
4. OLX.(n.d.) , Retrieved from: <http://olx.ba>
5. <https://ieeexplore.ieee.org/Xplore/home.jsp>
6. <https://machinelearningmystery.com/k-fold-cross-validation/>



**INNO SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 8.118**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details