iJRASET
International Journal For Research in Applied Science and Engineering Technology



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** III **Month of publication:** March 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.49832

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue III Mar 2023- Available at www.ijraset.com*

# Price Prediction for Second Hand Cars Using Machine Learning

Meghsha Singh[1], Thaduru Sathwik Reddy[2], Simma Krishna Sameera[3], Sai Srinivas[4], Sathi Surya Srijan Reddy[5]
*Department of Computer Science and Engineering, GITAM*

***Abstract:*** ***The purpose of this endeavour is to create a model which can forecast the cost of a used car depending on a variety of factors, including the vehicle's model used, year of production, fuel type, price, and kilometres driven. This method may help sellers, buyers, and automobile manufacturers in the used car market. Taking into account the data provided by the users, it may then provide a relatively accurate price estimate. In the model-building process, machine learning and data science are applied. The information came from classified advertising for used cars. The researchers employed a number of regression methodologies to achieve maximum accuracy, Regression algorithms are utilized because they generate an output that is a continuous random variable instead of a categorised value. Before beginning the model-building process, this project displayed the data in order to better understand the dataset. The dataset was partitioned and altered to match the regression to ensure its performance. To assess the effectiveness of each regression, R-square method was used as more than just a result, it will be possible to figure out a car's actual price instead of merely its range in pricing. Also, a user interface that accepts input from every user and displays the pricing of a specific vehicle also on user's inputs has been developed..***
***Keywords: Machine Learning, and Analysis Ridge Regression, Lasso Regression, and Linear Regression are all types of regression.***

## I. INTRODUCTION

In order to enable purchasers to make informed choices, Our development's objective is to create models using ML that can reliably forecast this used car's cost point on its attributes or on it's features. Because of the multiple factors that determine the market pricing of a used vehicle, evaluating if the quoted price is correct is a challenging task Overall, a regression model It uses a range of inputs to determine a car's exact price. parameters can be a useful tool for both buyers and sellers in the automobile market. However, it's important to use the model's predictions in conjunction with other sources of information and to be aware of its limitations and potential sources of error.

## II. LITERATURE SURVEY

The first study is about predicting used automobile prices using machine learning approaches. In this work, they investigate way guided methods of machine learning may be applied to foretell the cost of used cars in Mauritius. The forecasts are predicated on historical information from daily newspapers. Several methodologies, The results of multiple linear regression analyses, for example, were used to produce the predictions. Author Sameer Chand bases automobile pricing estimates on historical data gleaned from daily newspapers. They used supervised machine learning algorithms to estimate automobile prices. Multiple linear regression, KNN algorithms, naive-based techniques, and other decision tree algorithmic techniques have also been used. The best prediction method was determined after analyzing all used algorithms, they have encountered several difficulties visualizing the algorithms, but they eventually gained results. The goal of this work, according to authors et al, is collecting data from an internet domain using web scraping. They were then ranked using multiple machine learning algorithms to anticipate automobile costing in a straightforward manner. then they separated the prices into previously determined discrete price categories. To create classifier models on various datasets were used.

In their work, Wu et al. developed an adaptive neuro fuzzy inference experience and understanding system to show how and where to estimate the price of automobiles. They projected a model with comparable results to the basic regression model while accounting for the following characteristics: engine type, brand, and year of manufacturing. Due to the increased demand for leased vehicles among auto dealers at the conclusion of the lease year, they have created an expert system called and known as ODAV. This strategy gives information on the best automobile prices as well as the best location to obtain them. Auto prices were estimated using the K-nearest neighbour machine Learning technique, which utilises regression modelling.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

According to the authors, this study focuses on the interaction between the vendor and the buyer. Additional featuresTo predict the price of four-wheelers, factors such the previously mentioned cost, kilometers, brand, brand, trim, type, cylinder, and litre, as well as seats, cruise, sound, and luxury, are required.
These characteristics were used to project the cost of a car using a statistical analytic method for exploratory analysis of data.

## III. METHODOLOGY

This model was developed using a variety of information and methods in this section. A data collection with 93386 records is going to be utilized to train the model. The number of miles driven, year during registration, fuel type, model of car, wealth, auto maker, and gear type are all variables that influence an automobile's value. We created five algorithms—Lasso Regression, Ridge Regression, Linear Regression, and Random Forest Regression—because this is a technique used to predict.

### *A. Lasso Regression*

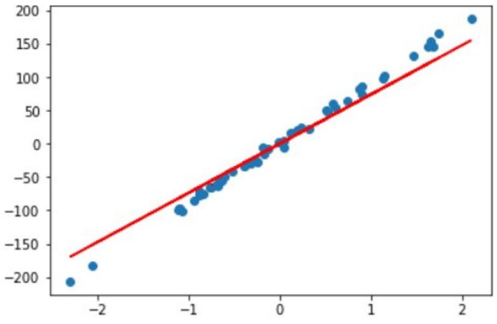
By using lasso regression, it is possible to determine which one of these features is most crucial for estimating the cost of a used car and to lessen the influence of any redundant or unnecessary characteristics that might be in the data set. Thus, lasso regression can boost the used car price forecast model's precision and dependability.
The trained model can then be used to new data to produce predictions, such as the cost of a certain used car given its specifications. Metrics like mean squared error or R-squared can be used to assess the model's accuracy.

$$\sum_{i=1}^{M}(y_i - y'_i)^2 = \sum_{i=1}^{M}\left(y_i - \sum_{j=0}^{n}\beta_j * x_{ij}\right)^2 + \lambda\sum_{j=0}^{n}|\beta_j|$$

Xij = Y or Active Parameter Characteristics
Yi = Dependent Variable
βi2 = A feature's importance is indicated by its weight or magnitude.
λ =cross-validation estimated error rate as little as possible

### *B. Ridge Regression*

Another kind of regression analysis that may be employed to forecast used car prices is ridge regression. Ridge regression uses regularisation, like lasso regression does, but it does so using a separate penalty function that can help to mitigate the effects of cointegration in the dataset.




Graphical Representation of Linear Regression

The training set may be applied to new data to produce predictions, such as the cost of a certain used vehicle given its specifications. Metrics like mean squares of the errors or R-squared can be used to assess the model's accuracy. Overall, when working with data that have a high degree of multicollinearity, regression model can be a beneficial method for predicting the price of used cars.

## IV. OBJECTIVE

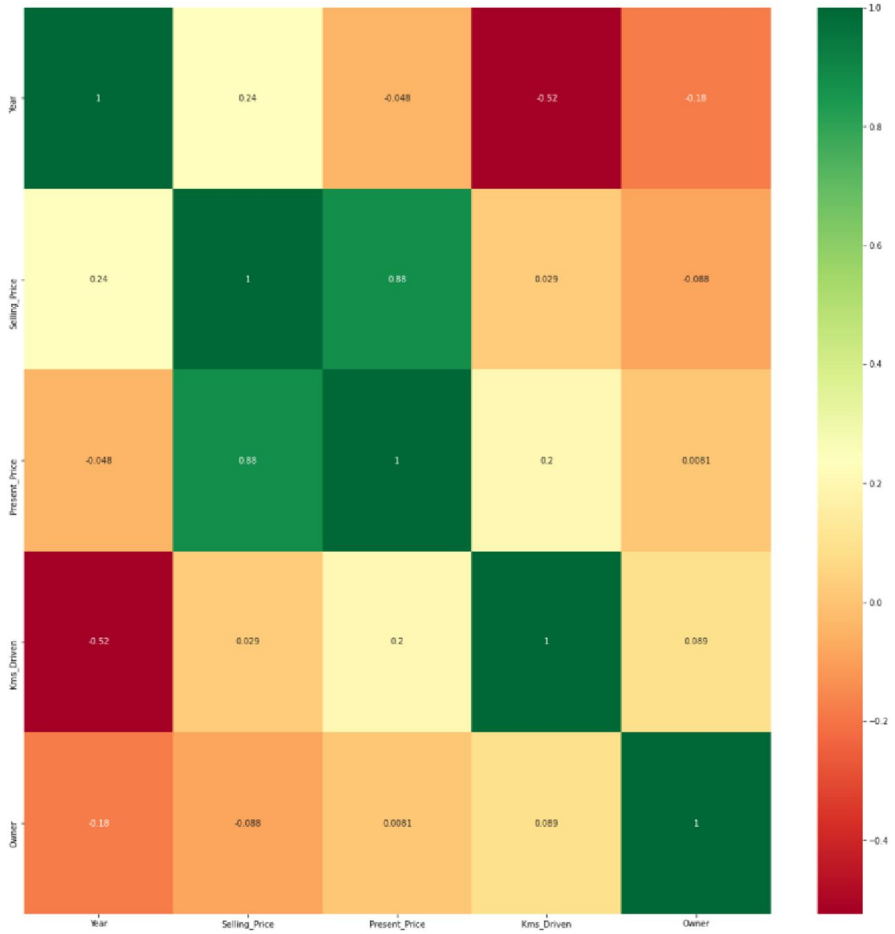
*1)* To develop a model that accurately and efficiently calculates the cost of a used specific car on the user's inputs.
*2)* To design an easy-to- use User Interface that takes user input and predicts pricing

## V. DATASET

The dataset consists of columns the columns are car model, fuel type, year of purchased

| | name | company | year | Price | kms_driven | fuel_type |
|---|---|---|---|---|---|---|
| 0 | Hyundai Santro Xing XO eRLX Euro III | Hyundai | 2007 | 80,000 | 45,000 kms | Petrol |
| 1 | Mahindra Jeep CL550 MDI | Mahindra | 2006 | 4,25,000 | 40 kms | Diesel |
| 2 | Maruti Suzuki Alto 800 Vxi | Maruti | 2018 | Ask For Price | 22,000 kms | Petrol |
| 3 | Hyundai Grand i10 Magna 1.2 Kappa VTVT | Hyundai | 2014 | 3,25,000 | 28,000 kms | Petrol |
| 4 | Ford EcoSport Titanium 1.5L TDCi | Ford | 2014 | 5,75,000 | 36,000 kms | Diesel |

### *A. Heatmap*




Heat map fig

## VI. CONCLUSION

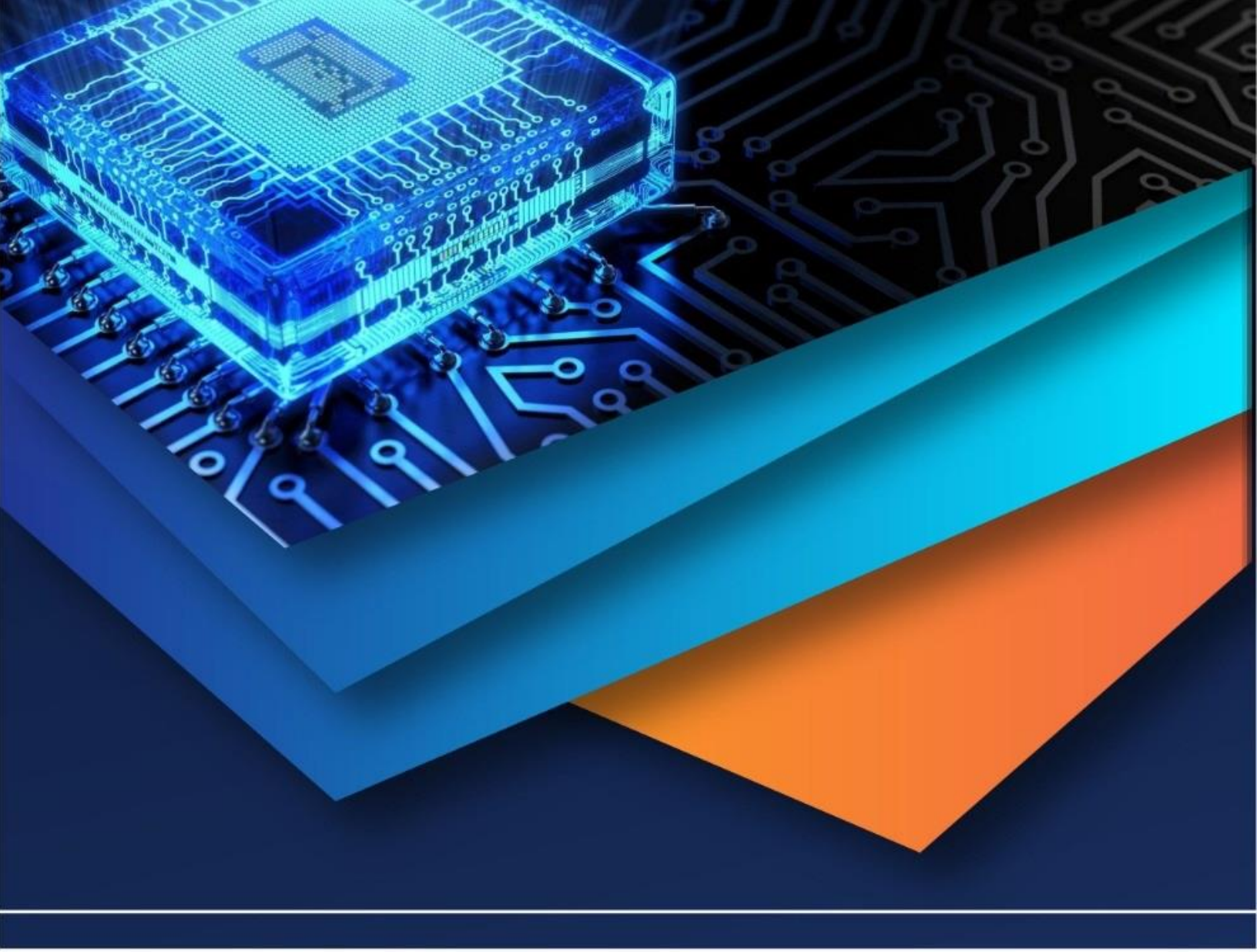
Due to the numerous aspects that must be taken into account for an accurate prediction, predicting car prices is a difficult task.
Data gathering and preparation are the most crucial steps in the prediction process.
In order to build machine learning algorithms, automobile data from kaggle.com is converted into CSV format throughout the study.
Data was split into two distributions, one for training and the other for testing, by the SVM classifier (Support Vector Machine). In other words, 75% of the data was used for machine learning testing, and 25% for training.
The three ML models' accuracy is then compared and rated. This is a significant comparison of machine learning methods from single versus many groups.
As a result, the more accurate forecasting model will assist in determining the actual cost of the car.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

## REFERENCES

[1] 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)
[2] Sameerchand Pudaruth, "Predicting the Price of Used Cars using Machine Learning Techniques"; (IJICT 2014).
[3] Richardson, M. S. (2009). Determinants of used vehicle resale value
[4] Wu,etal,(2009). An expert system of price forecasting for used vehicles using adaptive neuro-fuzzy inference.
[5] Enis gegic, Becir Isakovic, Dino Keco, Zerina Masetic, Jasmin Kevric, "Car Price Prediction Using Machine Learning"; (TEM Journal 2019)
[6] 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)
[7] Nitis Monburinon, Prajak ChertchoSm, Thongchai Kaewkiriya, Suwat Rungpheung, Sabir Buya, Pitchayakit Boonpou, "Prediction of Prices for Used Car by using Regression Models" (ICBIR 2018)

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)