www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

Examination of Machine Learning for the Purpose of Predicting the Value of Used Vehicles

[1] Manjusha Nambair, [2] Ch. Archana,

[1]Assistant Professor, Megha Institute of Engineering & Technology for Women, Ghatkesar.
[2] MCA Student, Megha Institute of Engineering & Technology for Women, Ghatkesar.

Abstract—

At the outset, there is a manufacturerset price for vehicles that are particular to brand, model, year, and features. As they get older and are sold as used, their value is affected by market forces related to their unique characteristics, as well as their personal background. Traditional techniques of determining their worth grow more difficult the more unique they are in comparison to other vehicles of a similar kind. A more precise evaluation of a vehicle's value is possible with the use of Machine Learning (ML) algorithms that make better use of data on all of the less common automotive qualities. This study examines the effectiveness of several Machine Learning (ML) algorithms in predicting the values of used vehicles. These algorithms include Linear Regression, Ridge Regression, Lasso Regression, and Random Forest Regression. To better use previous data in projecting current prices, an effective price prediction algorithm must be able to integrate depreciation. The examination of a large public dataset consisting of pre-owned automobiles was important to this investigation.

Keywords— Used car prices, regression analysis, supervised machine learning techniques, and depreciation

I. INTRODUCTION

An advanced online tool, the Car Price Predictor, is at the forefront of technology that aims to provide consumers with an exact estimate of the worth of used cars. The program is based on state-of-the-art machine learning methods and aims to use the linear regression algorithm to its full potential. The comprehensive dataset that this tool runs on has been carefully selected to include a wide variety of factors that are important in calculating the market value of an automobile. Included in this set of characteristics are important facts including the vehicle's make, model, year of purchase, fuel type, and total kilometers driven. The Car Price Predictor's data handling is quite careful, which shows their devotion to accuracy. In order to ensure that the tool's predictive model is as accurate as possible, it performs thorough cleansing and preprocessing on the dataset before making any predictions [1]. This meticulousness ensures that the model is trained on a high-quality dataset, which improves its capacity to provide dependable predictions in the complex and ever-changing used automobile market. Machine learning, and more especially a linear regression model trained on the carefully selected dataset, is the backbone of the Car Price Predictor. In this scenario, the characteristics included in the automobile affect its price, and linear regression, a basic statistical analysis tool, skillfully fits a linear equation to the observed data points to enable the model to predict the numeric value of the target variable. While training, the suggested model takes into account a large number of input features—the year of purchase, model, and manufacturer—and outputs a single variable—the price of the vehicle. The model can extend its knowledge from past data and produce accurate predictions for new occurrences by learning from previous data. When dealing with market variables like gasoline price swings, economic uncertainty, and other factors that impact vehicle costs, the capacity to quickly adjust to new circumstances is invaluable. The Car Price Predictor does a great job of turning a trained and verified model into an intuitive online app. The app's user-friendly layout allows users to quickly and easily enter their desired car's specifications. Included in

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

this are not just the standard pieces of data like make, model, and purchase year, but also more specific pieces like fuel type and mileage. Everyone, even those with little technical knowledge, will be able to utilize the tool because to its intuitive design. The Car Price Predictor goes into gear as soon as it gets these inputs. It takes the supplied data and runs it through the trained linear regression model to provide an accurate estimate of the car's pricing [2]. Afterwards, the customer is provided with an easily digestible forecast along with any accompanying explanations or visuals that may shed light on the variables impacting the expected price. Because of this openness, consumers are able to make educated judgments about what to buy or sell in a market that could be difficult to understand at times. Beyond its intuitive UI, the Car Price Predictor takes a holistic approach. The tool's dedication to provide precise and dependable estimations in the dynamic used vehicle price market is evident in its meticulous data pretreatment and implementation of state-of-the-art machine learning models. Users may confidently traverse the complexity of the used automobile market thanks to the fusion of innovative technology and diligent attention to data quality. Designed to improve decision-making and promote a more informed and efficient automobile industry for all players, the tool goes beyond just a prediction engine.

II. LITERATURE REVIEW

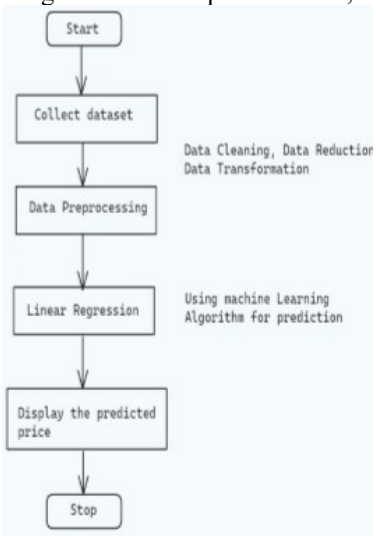
Predicting the pricing of used cars in Mauritius using supervised machine learning algorithms is the subject of an existing research effort. Daily newspapers provide the historical data used to make the projections. Multiple linear regression analysis, k-nearest neighbors, naive Bayes, and decision trees are among the methods that have been used to create these forecasts [1]. Using machine learning methods to forecast future vehicle costs is the subject of yet another research paper [2]. In order to guarantee the accuracy and dependability of the forecasts, a wide variety of distinct characteristics are examined. The researchers have utilized three separate machine learning techniques—ANN (Artificial Neural Network), Random Forest, and others—to simulate the prediction of used automobile pricing [3]. The used car industry's price assessment methodology is the subject of the research suggested in [4]. Back propagation (BP) neural networks are the main tools used in this investigation. This methodology stands out because it uses big data analysis to combine a large amount of dispersed vehicle data with a massive database of transaction data. Proposing a strong price assessment model that makes use of the abundance of accessible data is the main goal of this research. Using the improved BP neural network technique, the model is designed to examine pricing data for different car kinds. To make sure the model can accurately capture the intricacies of price dynamics in the used automobile market, this algorithm is fine-tuned to improve the analysis' efficiency and accuracy. This research aims to fill gaps in our knowledge of the variables that influence the price of used automobiles by using big data and cutting-edge neural network methods. Considerations such as the car's type and year, its performance record, current market conditions, and other pertinent details fall under this category. One way to get useful and reliable price forecasts is with the use of the suggested price assessment model. To function, it sifts through the massive information, finds trends and patterns, and applies this knowledge to forecast prices that are most in line with the features of each individual vehicle type [5]. Our goal is to develop a model that can adjust to changing market circumstances and emerging trends throughout time. The overarching goal of this study is to add to the literature on used-car pricing by presenting a complex model that makes use of neural network methods and huge data. The goal is to help people who are buying and selling used cars make better, more informed decisions by offering a more dynamic and accurate way to evaluate costs.

III. PROPOSED METHODOLOGY

There are essentially two stages to how the mechanism works: In the first stage, known as "training," the system learns how to use the information to fit a model, which may be a line or a curve, according to the method used. B. The testing step involves feeding the system data and seeing how it responds, with an eye on making sure it's accurate. Hence, appropriate data must be used for both training and evaluating the model [6]. The goal of the system is to find used vehicle prices and make predictions about them, thus it needs algorithms that are good at both of those things. A number of

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

methods were evaluated for their accuracy before being chosen for implementation, and the



one deemed most suitable for the task was chosen.

Fig. 1: Proposed System Flowchart

Figure 1 shows that the procedure starts with gathering the dataset. Data Preprocessing follows, which includes activities like data cleansing, data minimization, and data transformation, among others. The next step in making price predictions is to use machine learning techniques, such as Linear Regression. The most accurate model in terms of price prediction is then selected as the best one. The user is then shown the estimated pricing depending on their inputs once the best model has been chosen. The machine learning model can estimate used automobile prices with the help of user inputs provided via an easy-to-use online interface [7]. To forecast the price of an automobile given the brand, model, year, mileage, engine size, and a variety of other features of a car. In C. Linear Regression (LR), one variable is designated as the dependent factor and a linear equation is applied to the data in order to establish a link between the other variables. Consider a data analyst who, using a linear regression model, is trying to determine the relationship between people's weight and their height. Here, the model is used to investigate and measure the possible connection between these two crucial factors [8].

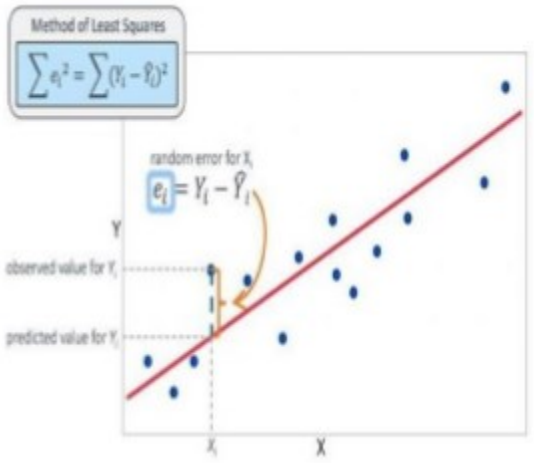


Fig. 2: General Representation of Linear Regression Process

Whether there are a number of independent variables or only one, linear regression is useful for discovering the correlations between continuous variables [9]. It is a useful tool for discovering relationships and interdependencies in a dataset, which may help you understand how changing one variable might affect the conclusion you're after.

a = c1D1 + c2D2 + c3D3 + …….. + e

mc1, c2, c3 → Slope

e → a Intercept

D1, D2, D3 ……. → independent variables

a → dependent variables

The suggested model may also have an API built using frameworks like Flask or Django [10]. This research keeps tabs on the model's accuracy by retraining it with fresh data at regular intervals and seeing how well it can identify new data. Model reliability is defined as the capacity to provide robust and consistent predictions over a wide range of situations and datasets [11].

www.ijbar.org
ISSN 2249-3352 (P) 2278-0505 (E)
Cosmos Impact Factor-5.86

IV. RESULTS & DISCUSSION

Table I shows the sample snippet of the dataset used in this research study

TABLE 1: DATASET USED

ID	Make	Model	Year	Price	Mileage	Engine Size	Fuel Type	Transmission	Horsepower
1	Toyota	Corolla	2018	15000	30000	1.8L	Petrol	Automatic	132
2	Ford	Mustang	2016	25000	40000	2.3L	Petrol	Manual	310
3	BMW	X5	2019	40000	15000	3.0L	Diesel	Automatic	335

Figure 3 shows the software generated output for the given dataset.

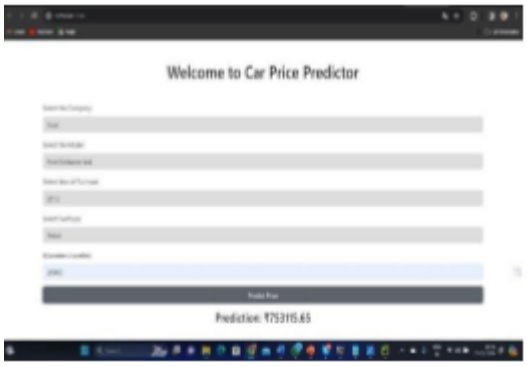


Fig. 3: OUTPUT

V. CONCLUSION & FUTURE SCOPE

The importance of the worldwide market for used automobiles is growing, and this research concludes that a reliable method for predicting their prices is necessary. Using the Linear Regression (LR) approach from supervised machine learning, this research created a strong model that can evaluate a car's worth depending on many different factors. This method trains the model to find correlations and trends that affect vehicle pricing by using extensive datasets that include a variety of parameters including year, condition, make, model, and miles. The suggested approach shows great promise in delivering accurate price predictions, which may help buyers and sellers make well-informed decisions. Our system's goal is to create a more fair and efficient used automobile market by increasing market transparency and efficiency, which will help close the gap between consumer needs and available technology. The accuracy of the model may be improved in future work by adding more variables and honing the algorithm using sophisticated machine learning methods. The system's stability and usefulness may be enhanced even more by adding a wider variety of vehicles and market circumstances to the dataset. Taken together, the findings of this study provide a useful tool for both buyers and sellers in the used automobile market by demonstrating the potential of machine learning to solve real-world problems in this sector.

REFERENCES

[1] Sameerchand Pudaruth, “Predicting the Price of Used Cars using Machine Learning Techniques”;(IJICT 2014).

[2] Enis gegic, Becir Isakovic, Dino Keco, Zerina Masetic, Jasmin Kevric, ”Car Price Prediction Using Machine Learning”; (TEM Journal 2019).

[3] 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).

[4] Nitis Monburinon, Prajak Chertchom, Thongchai Kaewkiriya, Suwat Rungpheung, Sabir Buya, Pitchayakit Boonpou, “Prediction of Prices for Used Car by using Regression Models” (ICBIR 2018).

[5] 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)

[6] LISTIANI, M., 2009. Support Vector Regression Analysis for Price Prediction in a Car

LeasingApplication. Thesis (MSc). Hamburg University of Technology.

[7] Oprea, C, Making the decision on buying second-hand car market using data mining techniques (Special, 2010), pp.17-26.

[8] Used Car Price Prediction Based on the Iterative Framework of XGBoost+LightGBM:This study, published in Electronics in 2023

[9] Vehicle Price Classification and Prediction Using Machine Learning in the IoT Smart Manufacturing Era:Published in Sustainability in 2023.

[10] Used Cars Price Prediction Using Machine Learning:This research, available on SSRN from 2023.

[11] "Price Prediction and Classification of Used-Vehicles Using Supervised Machine Learning" by Jasmina Pašagić Škrinjar, Tomislav Fratrović, and Borna Abramović.