USED CARS PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES

B. POOJITHA, MCA, DCA, DVR & Dr. Hima Shekar MIC College of Technology, A.P., India.

B. MURALI KRISHNA, Associate Professor, Dept. Of AI & IT, DVR & Dr. Hima Shekar MIC college of

Technology, A.P., India.

Abstract— in this paper discuss various algorithms and the required dataset that were implemented to build this module. Cars of a particular make, model, year, and set of features start out with a price set by the manufacturer. As they age and are resold as used, they are subject to supply and demand pricing for their particular set of features, in addition to their unique history. The more this sets them apart from comparable cars, the harder they become to evaluate with traditional methods. Using Machine L earning algorithms to better utilize data on all the less common features of a car can more accurately assess the value of a vehicle. This study compares the performance of Linear Regression, Ridge Regression, Lasso Regression, and Random Forest Regression ML algorithms in predicting the price of used cars. An important qualification of a price prediction tool is that depreciation can be represented to better utilize past data for current price prediction. The study has been conducted with a large public dataset of used cars. The results show that Random Forest Regression demonstrates the highest price prediction performance e across all metrics used. It was also able to represent average depreciation much more closely than the other algorithms, at 13.7% predicted annual geometric depreciation for the dataset independent of vehicle age.

INTRODUCTION

The world is growing day by day and also expectations of every people are also growing up. Out of all the expectation one of them is to buy a car. But all are not able to buy always a new car, so they will buy used one. But new person don't know about the market price for his or her dream car for old one. That is where we have need a platform which helps new people for car price prediction. The goal of this research is to create machine learning models that can properly forecast the price of a used car based on its attributes so that buyers can make educated decisions. On a dataset

containing the sale prices of various brands and models, we build and analyse several learning approaches. We'll examine the results of numerous machine learning algorithms, such as Linear Regression, Ridge Regression, Lasso Regression, Elastic Net, and Decision Tree Regressor, and pick the best one. The car's pricing will be determined based on a number of factors. Regression Algorithms are employed because they offer us with a continuous number as an output rather than a categorized value, allowing us to anticipate the real price of a car rather than its price range. A user interface has also been created that takes input from any user and displays the price of a car based on their inputs. There are three types of fuel data sets here. They are Diesel , Petrol and LPG are used here.

LITERATURE REVIEW

[1] Gegic, E. et al. (2019) demonstrate the need to create a model to forecast the cost of second hand cars in Bosnia and Herzegovina. They used machine learning techniques such as artificial neural networks, support vector machines, and random forests. However, the aforementioned methods were used in concert. The web scraper, which was created using the PHP programming language, was used to gather the data from the website autopijaca.ba for the forecast. Then, to determine which method best suited the provided data, the respective performances of various algorithms were compared

[2] Kuiper, S. (2008) demonstrated a multivariate regression model that helps in classifying and predicting values in numeric format. It demonstrates how to apply this multivariate regression model to forecast the price of 2005 General Motors (GM) vehicles. The price prediction of cars does not require any special knowledge. So, the data available online is enough to predict prices. The author of the article did the same car price prediction and introduced variable selection techniques that helped in finding which variables were more relevant for inclusion in the model.

[3] In 2019, Pal et al discovered as a methodology for predicting used cars prices using Random Forest. The paper evaluated usedcar price prediction using Kaggle data set which gave an accuracy of 83.62% for test data and 95% for train-data. The most relevant features used for this prediction were price, kilometer, brand, and vehicle type and identified by filtering out outliers and irrelevant features of the data set. Being a sophisticated model, Random Forest provided good accuracy in comparison to prior work using these data sets.

RELATED WORK

New cars of a particular make, model, and year all have the same retail price, excluding optional features. This price is set by the manufacturer. Used car, however, are subject to supply-and-demand pricing. Further, used cars have additional attributes that factor into the price. These include the condition, mileage, and repair history, which sets cars that may have shared are tail price apart. The used car market is generally divided into two categories, retail and wholesale. The retail price is the higher of the two prices and is what an individual should expect when buying a carat dealership. The whole sale price is the lower price which dealers will pay. Whether the dealer has sourced the car from a trade in, auction, or another dealer, this price is considerably lower to ensure that the dealer will make a profit on the vehicle. Prices for peer-to-peer car sales generally lie in-between theretail and wholesale price points. Because there is no "middle-man" inpeer-to-peer transactions, there is only a single price point, rather thantwo. A difficulty in peer-to-peer transactions is for both parties to agree on a fair price. There are many tools which provide an approximation, but do not factor in the particularities of the car into the price. Car markets are to some extent local and therefore location also affects the price. There is therefore a need for a valuation method which can make use of more of the features particular to each car, and extract information from all other previous sales of cars with shared features.

Machine learning (ML) is a subfield of Artificial Intelligence (AI) that works with algorithms and technologies to make useful inferences from data. Machine learning algorithms are well suited to problems entailing large amounts of data which would not be possible to process without such algorithms. ML works algorithmically rather than mathematically and permit a machine to "learn" and adapt its predictions to best fit the data it has trained on.

PROPOSED WORK

In addition to the age of the car and the revaluation of its features, used cars have a unique service history that develops over time. Parts will become worn with time and miles driven (mileage). What is replaced, when it is replaced, and by whom, are all to be considered as it relates to the current working condition of the car and its desirability on the market. The particularities are difficult to account for in traditional price-setting models, as it is a major differentiat or in vehicles. Generally, it is summarized in the "condition" of the car. The value of repairs or custom modifications to the car are recognized only if they noticeably improve the overall condition of the car.

SAMPLE SCREENSHOTS

Actual Prices vs Predicted Prices

MAPE by Algorithm for price percentile
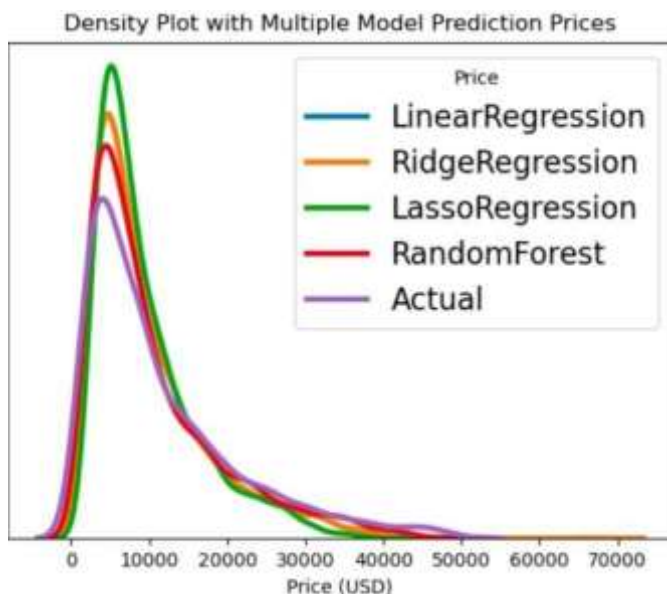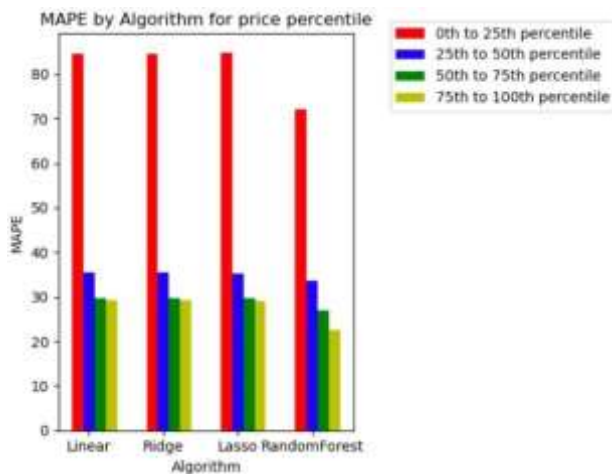
CONCLUSION

In this paper, we propose a Predicting used car prices is a difficult task due to the large number of features and parameters that must be examined in order to get reliable findings. The first and most important phase is data collection and preprocessing. The model was then defined and built in order to implement algorithms and generate results.

References

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

[2] Richardson, M. S. (2009). Determinants of used vehicle resale value

[3] Wu,etal,(2009). An expert system of price forecasting for used vehicles using adaptive neuro-fuzzy inference.

[4] Gegic, E. et al. (2019) 'Car price prediction

Density Plot with Multiple Model Prediction Prices

using machine learning techniques', TEM Journal, 8(1), pp. 113–118. doi: 10.18421/TEM81-16.

[5] Dholiya, M. et al. (2019) 'Automobile Resale System Using Machine Learning', International Research Journal of Engineering and Technology(IRJET), 6(4), pp. 3122–3125.