International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 03 | March - 2023 Impact Factor: 7.185 ISSN: 2582-3930

PREDICTION OF CAR RESELLING PRICES USING MACHINE LEARNING

V. Jeevan Babu
Computer Science department
(Asst. Professor)
Hyderabad Institute of Technology And Management
Hyderabad, India
jeevanbabu.cse@hitam.org

Sharkhane Bhargav Raj
Computer Science department
Hyderabad Institute of Technology And Management
Hyderabad, India
rajsharkhana@gmail.com

Thipparapu Shivaji
Computer Science department
Hyderabad Institute of Technology And Management
Hyderabad, India
luckkys178@gmail.com

Shanam Mukund
Computer Science department
Hyderabad Institute of Technology And Management
Hyderabad, India
shanammukund98@gmail.com

V. Jai Harshith Kumar
Computer Science department
Hyderabad Institute of Technology And Management
Hyderabad, India
jaiharshith.veerla@gmail.com

ABSTRACT

It is generally known that, taking wise and challenging decisions is really a crucial task in every business. To propose a novel solution for this challenge, this research work majorly focuses on one of the retail businesses i.e., used car sales business. The proposed research work shows that, the predictive analytical models will be a great add-on to business mainly for assisting the decision making process. The major objective of our paper is to build a prediction model i.e., a fair price mechanism to predict the cars selling price based on their features like the car model, the number of years that a car is old, the type of fuel it uses, the type of seller, the type of transmission and the number of kilometers that the car has driven so far. This paper will help to get an approximation about selling price of a used car based on its features and reduces the seller and consumer risk in business. The proposed model utilizes the machine learning algorithms and regression techniques of statistics like linear, decision tree and random forest regressions to achieve this task.

KEYWORDS

Used car sales business, decision making process, machine learning algorithms, regression techniques, linear, decision tree and random forest regressions.

INTRODUCTION

Predicting the price of used cars in both an important and interesting problem. According to data obtained from the National Transport Authority, the number of cars registered between 2003 and 2013 has witnessed a spectacular increase of 234%. From 68, 524 cars registered in 2003, this number has now reached 160, 701. With difficult economic conditions, it is likely that sales of secondhand imported (reconditioned) cars and used cars will increase. It is reported in the sales of new cars has registered a decrease of 8% in 2013. In many developed countries, it is common to lease a car rather than buying it outright. A lease is a binding contract between a buyer and a seller (or a third party – usually a bank, insurance firm or other financial institutions) in which the buyer must pay fixed instalments for a pre-defined number of months/years to the seller/financer. After the lease period is over, the buyer has the possibility to buy the car at its residual value, i.e. its expected resale value. Thus, it is of commercial interest to seller/financers to be able to predict the salvage value (residual value) of cars with accuracy. If the residual value is under-estimated by the seller/financer at the beginning, the instalments will be higher for the clients who will certainly then opt for another seller/financer. If the residual value is over-estimated, the instalments will be lower for the clients but then the seller/financer may have

much difficulty at selling these high-priced used cars at this overestimated residual value. Thus, we can see that estimating the price of used cars is of very high commercial importance as well. Manufacturers' from Germany made a loss of 1 billion Euros in their USA market because of mis-calculating the residual value of leased cars . Most individuals in Mauritius who buy new cars are also very apprehensive about the resale value of their cars after certain number of years when they will possibly sell it in the used cars market.

AIMS AND OBJECTIVE

Aim:

The aim of predicting used car prices using machine learning is to create a model that can accurately estimate the market value of a used car based on various factors, such as its make, model, year of manufacture, mileage, and condition.

Objective:

- Data collection: Compile a sizable database of previously used vehicles containing information on their make, model, year, mileage, condition, and location, among other factors that may influence their pricing.
- Data preparation: Preparation of data Cleaning, translating, and encoding the data into a machine-readable format that can be utilised for model training and testing constitutes preprocessing.
- Feature selection: Choose the most crucial attributes that affect the cost of the car and use them to train the model.
- Model selection: Choose the suitable machine learning technique, such as regression or decision trees, that can precisely forecast the car's pricing based on the chosen features.
- Model training: To make sure the model is reliable and effectively generalises to new data, train the model on the prepared dataset using appropriate methods, such as cross-validation.
- Analyze the model's performance on a test dataset and adjust its hyperparameters to get the most accurate results.

LITERATURE SURVEY

Predicting Used Car Prices Using Machine Learning Techniques" by S. Bhatnagar et al. (2019) - This study uses machine learning methods including decision trees, random forests, and linear regression to forecast the values of secondhand automobiles. The data for this study was gathered from a variety of places, including automobile dealerships and internet auto marketplaces.

"Predicting Used Car Prices with Machine Learning Techniques" by H. Asghar and N. Ahmad (2018) - In order to anticipate used automobile values, this study investigates the application of various machine learning methods, including

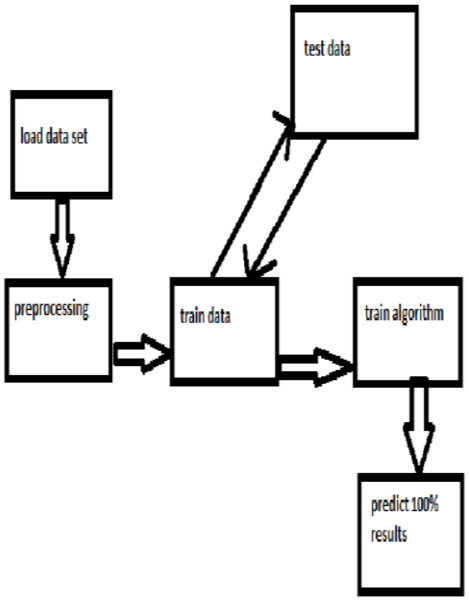
random forest, support vector machine, and artificial neural networks. The market in Pakistan provided the data set for this investigation.

"A Comparative Study of Machine Learning Techniques for Predicting Used Car Prices" by M. R. Alam et al. (2020) - The performance of different machine learning techniques, including gradient boosting, random forest, and decision trees, in forecasting used automobile prices is compared in this study. The market in Bangladesh provided the data for this study.

PROPOSED SYSTEM

The proposed research work shows that, the predictive analytical models will be a great add-on to business mainly for assisting the decision making process. Predictive Analytics is a process, where the businesses use statistical methods and technologies to analyze their historical data for delivering new insights and plan the future accordingly. The major objective of our paper is to build a prediction model i.e., a fair price mechanism to predict the cars selling price based on their features like the car model, the number of years that a car is old, the type of fuel it uses, the type of seller, the type of transmission and the number of kilometers that the car has driven so far. This paper will help to get an approximation about selling price of a used car based on its features and reduces the seller and consumer risk in business. The proposed model utilizes the machine learning algorithms and regression techniques of statistics like linear, decision tree and random forest regressions to achieve this task.

SYSTEM ARCHITECTURE



CONCLUSION

The increased prices of new cars and the financial incapability of the customers to buy them, Used Car sales are on a global increase. Therefore, there is an urgent need for a Used Car Price Prediction system which effectively determines the worthiness of the car using a variety of features. The proposed system will help to determine the accurate price of used car price prediction. This paper compares 3 different algorithms for machine learning : Random forest Regression, Lasso Regression and Ridge Regression.

FUTURE ENHANCEMENT

There are various potential future applications for the fast expanding subject of used-car price prediction using artificial neural networks and machine learning.

Integration of more data sources: Adding new data sources to the study is one technique to increase the precision of used car price forecast. For instance, utilising information from online auto markets like Autotrader or CarGurus could offer insightful knowledge into current market patterns and enhance the precision of the forecasts.

Expansion of the model to include more features: The existing model for estimating the price of used cars may only incorporate a small number of features, like the brand, model, and year of the vehicle. The precision of the forecasts can be increased any more by extending the model to incorporate extra information like the location, mileage, and state of the vehicle.

REFERENCES

[1] S. Pudaruth, "Predicting the Price of Used Cars using Machine Learning Techniques," International Journal of Information & Computation Technology, vol. 4, no. 7, pp. 753–764, 2014.

[2] N. Kanwal and J. Sadaqat, "Vehicle Price Prediction System using Machine Learning Techniques," International Jounal of Computer Applications, vol. 167, no. 9, pp. 27–31, 2017.

[3] S. Peerun, N. H. Chummun, and S. Pudaruth, "Predicting the Price of Second-hand Cars using Artificial Neural Networks," The Second International Conference on Data Mining, Internet Computing, and Big Data, no. August, pp. 17–21, 2015.

[4] N.Sun, H. Bai, Y. Geng, and H. Shi, "Price evaluation model in second-hand car system based on BP neural network theory," in 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), jun 2017, pp. 431–436.

[5] G.Rossum, "Python Reference Manual," Amsterdam, The Netherlands, The Netherlands, Tech. Rep., 1995.

[6] A. K. Elmagarmid, P. G. Ipeirotis, and V. S. Verykios, "Duplicate Record Detection: A Survey," IEEE Transactions on Knowledge and Data Engineering, vol. 19, no. 1, pp. 1–16, jan 2007.

[7] G.Chandrashekar and F. Sahin, "A survey on featureselection methods," Computers & Electrical Engineering, vol. 40, no. 1, pp. 16–28, 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0045790613003066

[8] M.C.Newman,"Regression analysis of log-transformed data: Statistical bias and its correction," Environmental Toxicology and

Chemistry, vol. 12, no. 6, pp. 1129–1133, 1993. [Online]. Available: http://dx.doi.org/10.1002/etc.5620120618

[9] R.Taylor, "Interpretation of the Correlation Coefficient: A Basic Review," Journal of Diagnostic Medical Sonography, vol. 6, no. 1, pp. 35–39, 1990.

[10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vander-(5)plas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duches-nay, "Scikit-learn: Machine Learning in fPgython," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.

[11] J. Morgan, "Classification and Regression Tree Analy-sis," Bu.Edu, no. 1, p. 16, 2014. [Online]. Available: http://www.bu.edu/sph/files/2014/05/MorganCART.pdf