2581-4575
International Journal For Recent Developments in Science & Technology
Crossref
A Peer Reviewed Research Journal

PREDICTING THE PRICE OF USED CARS USING MACHINE LEARNING TECHNIQUES

Mr.G.Bhanu Prasad[1],Gooty Sowmya[2], Gundeboina Anjali[3], Kottinti Sowmya[4]

[1]Associate Professor, School of CSE ,Malla Reddy Engineering College For Women(Autonomous Institution), Maisammaguda, Dhulapally,Secunderabad,Telangana-500100

[234]UG Student, Department of CSE,Malla Reddy Engineering College for Women, (Autonomous Institution), Maisammaguda,Dhulapally,Secunderabad,Telangana-500100

Email: bhanu.mrecw@gmail.com

ABSTRACT

A car price prediction has been a high interest research area, as it requires noticeable effort and knowledge of the field expert. Considerable number of distinct attributes are examined for the reliable and accurate prediction. To build a model for predicting the price of used cars. We applied machine learning (linear regression , logistic regression, Support Vector Machine and Random Forest). However, the mentioned techniques were applied to work as an ensemble and regressions. The data used for the prediction was collected Respective performances of different algorithms were then compared to find one that best suits the available data set. With the help of algorithms we have to predict the best percentage and also reveal the output with respective algorithms.

Keywords - car price prediction, support vector machines, classification, machine learning

I. INTRODUCTION

Predicting car prices is an interesting problem, and one that is becoming increasingly difficult, especially as the car market continues to grow. Bureau of Motor Vehicle Statistics show that 84% of vehicles registered in 2014 were for personal use, a figure that increased by 2.7% from 2013. For buyers and sellers, an effective way to estimate the value of a car is vital for making appropriate decisions in this dynamic market. Still, it is not an easy thing to predict car prices correctly. This will help know what factors influence the value of the vehicle. The make and model are the most important, but many factors have an impact on the price. The age and mileage of the vehicle are the most influencing factors; obviously, new vehicles with lower mileage command higher prices. The second significant factor is horsepower; the more horsepower, the greater the performance and thus higher demand for the vehicle. In addition, the type of fuel a vehicle uses and its fuel efficiency play a huge role in determining its price. With the changing cost of fuel and increasing concerns about sustainability, consumers may be concerned about how much a car gets per mile. This shift in priorities will impact the market, with vehicles that are more fuel-efficient or environmentally friendly becoming more expensive. Other than these apparent factors, there are many other factors that contribute to

the price of a car. These include exterior color, number of doors, type of transmission, and vehicle size-all factors that reflect consumer preference and market demand for the car. More expensive cars will be those with advanced safety features, interior comforts like air conditioning and interior trim, and high-tech features like navigation systems. In this paper, we would discuss the complexity of price prediction of used cars using various techniques and methods of machine learning to improve accuracy. By analyzing factors in more depth, we aim to build a more accurate model for car price prediction. The goal here is to take into consideration all the complexities and dynamics that affect the value of a car in the present market. This study not only helps buyers and sellers, but also gives an overall sense of how better one can understand the complex mechanisms behind the pricing of the automobile industry.

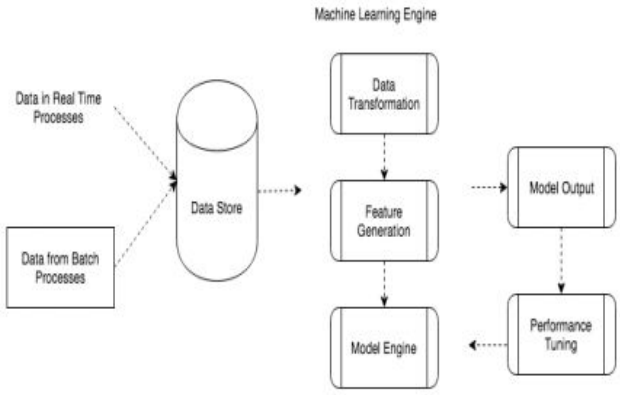

Fig 1: System Architecture

II. RELATED WORK

Machine learning for car price prediction

Author: C Jin(2021)

This paper provides an application of machine learning models for the prediction of used car prices. It gives focus on challenges that arise from different characteristics of the used car market, including make, model, year, mileage, and condition of the car. This paper reviews the various machine learning models used and compares their effectiveness in predicting prices in these markets.

Linear regression is widely used in used car price prediction

Author:MU Sumeyra and K YILDIZ (2023)

This paper presents the main applications of linear regression in used car price prediction. It outlines a step-by-step process for feature selection, model training, and evaluation using linear regression. This paper further analyzes and compares the performance of linear regression with other algorithms, simplifies its effectiveness as a base model for price prediction, and shows its simplicity in action.

Car Price prediction Using Regression: an Application of Multiple Regression Algorithms in KNIME Analytics Platform

Author:BC Arjun and AR Chitra(2023)

This is a paper that focuses on an analysis of multiple regression algorithms with their performance and suitability on implementing in car price prediction on KNIME analytics platform. The authors compare algorithms such as linear regression, decision trees, and support vector machines (SVM) and analyze their performance on metrics such as accuracy and computational efficiency.

Integration of devices and systems: Wireless Bluetooth Time-Lapse Printer for Restaurant and Hospital Management.

Author:R.V, AC, S. B. M, A. Kumari P, V. S. Reddy R me S. Murthy R (2022)

Not directly related to the used car price forecast, this article talks about the integration of hardware and software systems in real-time data applications. Bluetooth-based thermal sensors are used in restaurant and hospital management with the integration of hardware technology and machine learning to achieve efficient data transmission and processing.

In-vehicle radar monitoring system: uses the most significant signals known algorithm to detect and locate people inside the vehicle.

Author:A, Viswanatha, A.C. Ramachandra, P.T.Hegde, M.V. Raghunatha Reddy,V. Hegde and V. Sabhahit (2022)

The article explains how to use the radar monitoring and signal recognition algorithms in detecting and locating people inside a vehicle. Although not directly related to cost prediction, the application of machine learning to vehicle tracking systems shows an interesting connection with future machine learning technologies that can be applied to intelligent vehicle systems or predictive maintenance in the automotive field.

Implement intelligent security systems in agriculture using embedded machine learning

Author: Viswanatha, A. C. Ramachandra, P. T. Hegde, M. V. Raghunatha Reddy, V. Hegde and V. Sabhahit(2023)

In this paper, the authors have applied real-time safety applications in agriculture based on embedded machine learning systems. In this work, sensor data and predictive analytics, akin to automotive safety systems monitoring and maintenance, are discussed. Even though the application focus is on agriculture, the concepts of embedded machine learning can be extended to other applications as well, such as in the automotive industry.

III IMPLEMENTATION

It begins with data collection and preprocessing to obtain relevant automobile attributes, which include its make, model, year of production, mileage, the type of fuel used, and the power rating along with prices from the online sources available as well as existing data. Data gathered is processed for missing values integrated to fill in with continuous attribute values such as mileage and price to enable the consistency within the whole data set. After that, the data is split into a training set and a test set to train the model and evaluate its performance on raw data.

Then,Feature Identification is done to find out which attributes are most relevant to predicting car prices. Techniques such as correlation analysis and feature significance using models such as random forests help in selecting relevant features and reducing bias. Features like mileage, age, and horsepower are very important for the price prediction of a car, but features like color and number of doors are less important.

After feature selection, move on to model selection and training. Several types of machine learning algorithms were tested, including random forests, support vector machines (SVMs), and neural networks.

Random forests are chosen with the consideration that they can handle big complex datasets and produce appropriate predictions, while support vector machines are used after having feature scaling.

Neural networks and deep learning models are used specifically because of their ability to handle large nonlinear datasets. The performance of the models is tested using metrics like mean absolute error (MAE), root mean square error (RMSE), and R-squared to determine the best models. Finally, in the delivery phase, the best model is embedded in a website or mobile application for car price prediction.

The backend, such as created using Flask, serves in running a model so the user can come in with features of any car and view an estimate about its price. This implementation serves users who quickly access car prices through instantaneous forecasts based on the model and can then be updated continually to meet the changing information of price trends of market vehicles.

IV ALGORITHM

The car price prediction system collects data from a variety of sources, including online car markets and historical data sets, which provide key car attributes such as make, model, year, mileage, fuel type, electronics, and price. After collecting the data, it performs imputation to replace missing values using interpolation methods, and continuous features such as mileage and price are compared to be consistent. Features like type of fuel and brand are encoded, and the data is split into training and test sets before training the model and evaluating it.

Following the cleaning of the data, feature selection is performed to identify which features contribute most to predicting car prices.

Techniques like correlation analysis as well as the feature importance using models like the random forest will help determine the key features such as mileage, age, or horsepower among others and eliminate more irrelevant ones. If feature set dimensionality needs further reduction then dimensionality-reduction technique such as PCA can also be used in this kind of problem. Model selection and training:

These involve applying machine learning techniques, including random forests, SVM, and neural networks. The random forests are picked due to their ability in handling complex data sets. There are support vector machine models applied after feature adjustment; these are used for large data sets with the potential of modeling nonlinear relationships in such data sets.

The model is learned and evaluated on the metrics on the training data using MAE, RMSE, or R-squared of prediction accuracy. Once this step has been completed and a choice of the best model achieved, the system transitions into the deployment phase, which involves the placement of the model in the deployed website or mobile application. Backend services are built using frameworks like Flask to offer services for real-time car price prediction models.

Users can input car make, model, mileage, horsepower, and other characteristics, and the system returns a price predicted based on the trained model. The system is constantly

International Journal For Recent Developments in Science & Technology
2581-4575
Crossref
A Peer Reviewed Research Journal

updated to reflect market changes by retraining the model with new data as needed.

RESULT

Fig 1: User Login

Fig 2:Accuracy Bar Chart

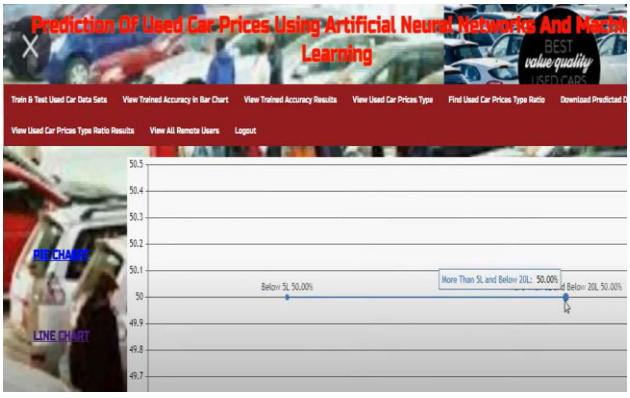

Fig 3:Line Chart

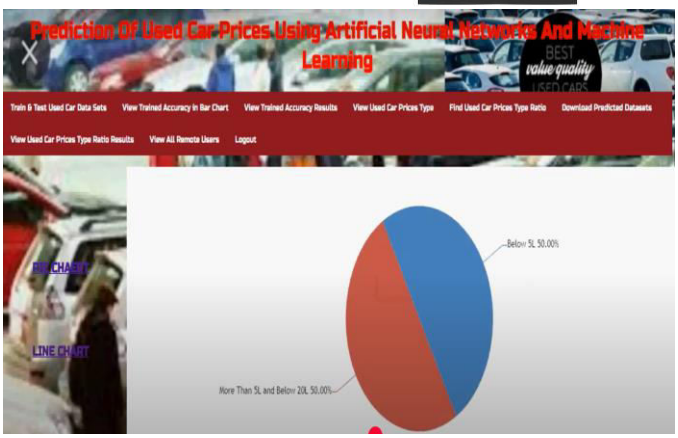

Fig 4:Pie Chart

CONCLUSION

In summary, the car price prediction systems show how the proper collection of data, pre-processing, and feature selection is crucial in improving the accuracy of prediction. Studies show that a single machine learning engine might not be enough to handle complex data sets because it will lead to lower accuracy. Using multiple machine learning engines, the prediction accuracy increased, demonstrating the power of mixed methods in solving complex problems like car price prediction. However, this increases the demand for computing resources.

In spite of these challenges, the system performed remarkably well in predicting car prices. Future work will focus on the testing and refinement of the system with multiple data sets to achieve greater scalability, efficiency, and usability. The key is to reach a balance between the precision of the prediction and the use of computing resources so that the system is more efficient and closer to practical application in the automotive market.

REFERENCES

[1] C Jin, "Machine Learning for Used Car Price Prediction", 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT), pp. 223-230, November 22.

[2] MU Sumeyra and K YILDIZ, "Linear Regression Is mainly Used To Predict Used Car Prices", International Journal of Computational and Experimental Science and Engineering., vol. 9, no. 1, pp. 11-6, Mar 2023.

[3] BC Arjun and AR Chitra, KNIME Analytics Platform Performance Analysis of Regression Algorithms for Used Car Price Prediction.

[4] R. V, A. C, S. B. M, A. Kumari P, V. S. Reddy R and S. Murthy R, "Custom Hardware and Software Integration: Bluetooth Based Wireless Thermal Printer for Restaurant and Hospital Management", 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), pp. 1-5, 2022.

[5] A, "In-Cabin Radar Monitoring System: Detection and Localization of People Inside Vehicle Using Vital Sign Sensing Algorithm", International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 8, pp. 104-9, Aug. 2022.

[6] Viswanatha, A. C. Ramachandra, P. T. Hegde, M. V. Raghunatha Reddy, V. Hegde and V. Sabhahit, "Implementation of Smart Security System in Agriculture fields Using Embedded Machine Learning", 2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC), pp. 1-6, 2023.

[7] A Joshi, C Vishnu and C K Mohan, "Early detection of earthquake magnitude based on stacked ensemble model", Journal of Asian Earth Sciences: X, vol. 1, no. 8, pp. 100122, Dec 2022.

[8] FE Ayo, JB Awotunde, S Misra, SA Ajagbe and N Mishra, "A Deep Learning Method Based on Rules for Predicting Used Car Prices", International Conference on Machine Intelligence and Signal Processing 2022, pp. 845-857, Mar 12.

[9] M Hankar, M Birjali and A Beni-Hssane, "A case study of used automobile pricing prediction using machine learning", In 2022 the 11th International Symposium on Signal Image Video and Communications (ISIVC) will be held on, pp. 1-4, May 18, 2022.

[10] .L Bukvi, J Paagi krinjar, T Fratrovi and B Abramovi, Supervised Machine Learning is used to predict and classify used vehicle prices., vol. 14, no. 24, pp. 17034, Dec 2022.

[11] F Al-Turjman, AA Hussain, S Alturjman and C Altrjman, "Vehicle pricing categorization and prediction using machine learning in the IoT smart manufacturing age", Sustainability., vol. 14, no. 15, pp. 9147, Jul 2022.

[12] K Samruddhi and RA Kumar, "Prediction of used automobile prices using a k-nearest neighbor model", IJIRASE (Int. J. Innov. Res. Appl. Sci. Eng.)., vol. 4, pp. 629-32, Sep 2020.

[13] .Y Li, Y Li and Y Liu, "The study of used vehicle price prediction using random forest and LightGBM", IEEE 2022 2nd International Conference on Data Science and Computer Application (ICDSCA), pp. 539-543, Oct 28.

[14] A Wang, Q Yu, X Li, Z Lu, X Yu and Z Wang, "Machine Learning-based research on the problem of used car valuation", In the 2022 International Conference on Computer

Network Electronics and Automation (ICCNEA), pp. 101-106, September 23.

[15] S Pudaruth, "Machine learning algorithms for predicting used automobile prices", International Journal of Information and Computer Technology., vol. 4, no. 7, pp. 753-64, Jan 2014.

[16] V. Viswanatha, A. C. Ramachandra, G. L. Reddy, A. V. S. T. Reddy, B. P. K. Reddy and G. B. Kiran, "An Intelligent Camera Based Eye Controlled Wheelchair System: Haar Cascade and Gaze Estimation Algorithms", 2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC), pp. 1-5, 2023.