



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 7, July 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Prediction of Prices of Used Cars Using Machine Learning

Manthan Mishra, Prof.Sandarsh Gowda M M

MCA Graduate(2nd year),Dept. of MCA, Bangalore Institute of Technology, Bangalore, Karnataka, India

Asst. Professor, Dept. of MCA Bangalore Institute of Technology, Bangalore, Karnataka, India

ABSTRACT: Given the huge increase in car practise, freshly built cars are unable to reach purchasers owing to a variety of issues such higher pricing, a restricted inventory, limited finances, as well as other considerations. As result, the market for second-hand vehicles is expanding quickly everywhere, except for India, which was decided where it nonetheless remains quite young and heavily dominated by the unorganised sector. This raises the possibility of fraud whenever buying an automobile that is used. In order to predict the cost of a used car without favouring either the customer or the merchandiser, an amazingly precise computation is required. In this case, two models that both can train a specified automotive information are built: a Random Forest- Based ML strategy and a Supervised learning-based Artificial Neural Network network.

Many different distinct factors are taken into account to produce trustworthy and precise projections. The results obtained are superior to approaches that rely just on simple linear models and are in accordance with the predictions of theory. The Keras software Crs procedure, also known as the Keras (pronounced Regressor, along with other algorithms for machine learning like Random Forest, Lasso, Ridge, and Linear regressions are used to build an ANN (Artificial Neural Network). The collection of data for automobiles is utilised to examine those approaches. Results from experiments show that the

Random Forest model has yielded the least amount of error out of all the other algorithms, with a Mean Absolute Error value of 1.0970472 and R2 error value of 0.772584.

Everyday, the world and everyone's expectations are expanding. Out of all the anticipation, one of them will buy an automobile. Everyone will eventually purchase a used automobile and they are unable to consistently purchase new ones. However, another individual is unfamiliar knowing the selling prices for the selected automobile on the marketplace for second hand cars. While the price for fresh cars increased due to inflation, the value case for used cars became strengthened. greater technology costs.

KEYWORDS: Artificial Neural Network, Keras Regression algorithm

I. INTRODUCTION

How will educational computing operate?

Unlike written programming, a system of computer algorithms known as "machine learning" possesses the capacity to gather information from knowledge and improve themselves. ML is a subset of AI that employs statistical approaches and data to predict a result and can be used to produce insightful findings.

Compared to traditional computing, data mining is significantly different. In traditional programming, every directive was implemented after an engineer consulted a specialist in the industry for which the product were being made. Every rule includes a logic basis, and the computer will apply the conclusion that follows the rational assumption. The greater the complexity of the system is, the more rules must be specified. It can be maintained. quickly become impossible.

All learning happens in the cerebral area of a machine studying system. A computer teaches in a manner similar to how a person learns. People learn by encounter. Forecasting becomes better as our knowledge increases. By metaphor, once we meet an unknown circumstance, our chances of success are lower than they would be in a known situation. The

identical instruction is given to machinery. For the purpose of to make an exact forecasting a machine looks at a specific instance. Whenever we give the system a scenario that is similar, it can anticipate the outcome. Nevertheless, if the equipment offers a fresh For instance, it struggles to forecast, much like a person.

II. LITERATURE SURVEY

Synthetic neural connections for Second-Hand Car Price Prediction

[1]Over the past ten years, the number of cars on Mauritius' roadways has steadily increased by 5%. 173 954 automobiles were registered with the National Transport Authority in 2014. As a result, one in six Mauritians has a car, the majority of which are old or reconditioned second-hand vehicles. The primary goal of the present investigation is to investigate whether abused automobile prices are possible to predicted using neural networks that are artificially created. In order to do this, data on 200 autos was gathered to multiple sources and fed into four distinct neural network models.

They found that support vector machine regression produced conclusions which was slightly better than those produced by neural network analysis or linear regression models. a few of the predicted values, particularly for more expensive cars, are, however, fairly far from the actual pricing. A larger data set and additional study are required to be able to produce improved projections. additional testing with various network types and topologies.Utilising anticipate the retail value of an automobile using the Random Forest Machine Learning Algorithm [2]The country that constitutes India is the largest automobile markets in the world. Every day, numerous customers often sell their vehicles to new owners after using them for a while. These purchasers have access to a variety of venues, including cars24.com, cardekho.com, and OLX.com, where they may sell their old cars. However, the hardest question of all is how much the automobile should be worth. Algorithms developed using machine learning might be ready to address this problem. Here, I forecasted the purchase cost of the preowned vehicle using powerful Language package Sci Kit Learn and machine learning methods like Random Forest and Extra Tree regression analysis. A reasonable price for the car might be predicted by combining data from already sold utilised automobiles and machine learning techniques like directed learning. No matter how large or nothing the information being analysed is, the results show that both of these algorithms are exceedingly precise when developing judgements. Calculating the Retail Value of Second hand Vehicles Using Supervised Learning Technologies.[3]Over the previous ten years The amount of of cars produced has gradually increased; in 2016, more than 70 million passenger cars were created. The used automobile market, which has grown into a thriving business on its own, has resulted from this.We will attempt to create A statistical technique that uses data on customers as well as regression models from neural networks, such as Lasso Regression, Multiple Regression, and Regression Vegetation, to forecast the pricing of a vehicle that is used. according to the order of ascending to choose the best model, we will also compare these models' predictive accuracy. Regression models deployed to forecast used automobile prices.[4]Second-hand vehicles are now the most According to the quick growth in the number of private vehicles and the expansion of the resale markets, used automobiles have become a popular option for people looking to buy cars. The web-based used vehicle market gives buyers and sellers the chance to conduct P2P transactions. This study proposes an expense appraisal mechanism based on big data analysis. It employs an enhanced neural network software for examining the price data every every sort of car utilising extensively distributed car information and a substantial amount of vehicle transaction data. As a way to ascertain frequency price that best fits the car, it attempts to build a model for evaluating used car prices. The fitting curve of the forecast price is compared with the actual transaction price produced from the optimised model through the sample simulation experiments. As a result, the accuracy and fitting of the optimised model are both improved. Machine Learning Methods for Used Car Price Prediction: Predictive Analytics in Retail. It is well recognised that making difficult but intelligent judgements [5] is an essential duty for every organisation. Making poor decisions can result in significant losses or possibly the closure of a corporation. The primary aim of this research programme is one of the retail industries, namely the used automobile sales company, in an effort to suggest a creative solution to this difficulty. According to the proposed research, predictive analytical models will be a tremendous asset to businesses, especially for aiding in decision-making. Predictive analytics is a process where firms analyse their previous data using statistical methods and technologies to produce fresh insights and plan the future appropriately.

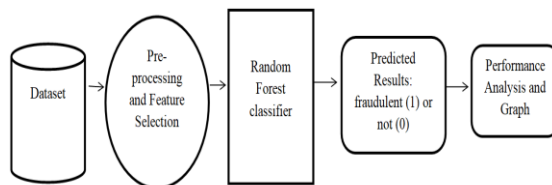


Fig 1. Proposed architecture

III. EXISTING WORK

The current system was developed utilising an ANN (Artificial Neural Network) and the Keras Regression algorithm, also known as Keras Regressor, along alongside additional methods that use machine learning, such as Straight Regressions, the Random Forest, Lasso, and Ridge. All of these approaches are tested using the car information. The experiments show that the Random Forest model has generated: least amount of error out of all the other algorithms, with a Mean Absolute Error value of 1.0970472 and R2 error value of 0.772584. Rectification Linear Unit (ReLU) activation functions are employed in a contemporary framework. Different level quantity events are collected given effectiveness measures. The secret layer's total quantity of neural networks is fixed at 10,000 epochs on every occasion. set at 20, the input dataset is provided as the first layer's input, and the selling price is used as the output layer's last layer's ANN. Artificial Neural Networks require a significant amount of processing power. Due to its numerous parameters, ANN also requires larger training datasets. These factors make ANN computationally intensive.

The current system model is considerably more intricate. They are made up of several processing nodes joined together. Understanding how the node weights produce the projected outcome is difficult. Neural networks might not be the ideal option if you need to provide model findings that are simple to explain to a non-technical audience.

IV. PROPOSED METHODOLOGY

Using the Decision Tree Regressor, we create used car price predictions in the suggested system. The data came from the website Kaggle.com, which uses Cardekho's automotive data sets for automobile sales and purchases. The following information was included in the dataset: automobile name, year, selling price, current price, miles travelled, fuel type (diesel, gasoline, or CNG), seller type (dealer or person), gearbox type (automatic or manual), and owner (number of prior owners).

The import and reading of the research's csv file comes initially. The dataset's null values, shape, columns, numerical and categorical features, dataset columns, Its feature's distinctive qualities, details of data, and other elements has been carefully examined. Following the connected and separate attributes are transferred, we split the dataset into training and testing data. We training using 80% of the data and test with 20% on the information that is available. test our model.

Then, a web application is created, with the Using the Python Flask Framework for the rear end and Java script, HTML, and CSS for the front part. This app for the web allows any individual to enter characteristics to estimate a previously owned car's cost of sale. The client must pick option for factors is like gasoline sort, engine category, and supplier sort & enter values for variables like year, starting price (in lakhs), kilometres journeyed, and the previous owner. display the results.

V. IMPLEMENTATION

Data collection

We create the Data Collection module in the first module. The real procedure of producing a machine learning model and accumulating input starts now. This stage is critical while the amount and quality data that we can gather will determine how effectively the predicted outcome behaves.

Website scraping as well as other procedures are examples of gathering knowledge techniques. The dataset in question is available in the well-known information archive Kaggle as well referred to. The dataset's link is as follows.

www.kaggle.com/datasets/jayaprakashpondy/car-price is the link to Kaggle.

Data preparation

Amass information in order to prepare it for analysis. Trim counterfeits, correct errors, handle missing numbers, normalise, convert data types, and anything else who may require cleaning up.

Randomising information removes the effects of the precise order in which we collected and/or otherwise prepared our data.

Conduct additional exploratory research, like data visualisation, to find important relationships between variables or class imbalances (bias alert!).

There are various sets for training and evaluating.

Model selection

Choice Tree Linear was the machine learning algorithm we used. Regression and classification difficulties can be resolved using a supervised learning method known as a tree of choices, however this approach is frequently preferred. It is a tree-structured classifier, where each node in the leaf represents the prediction outcome and inside components represent the distinguishing characteristics of the collected data.

The two distinct nodes in this foliage are the Decision Node and a leaf node. The nodes known as decision nodes are used to make judgements that consist of numerous roots, whereas nodes with leaves are the outcomes of decisions and do not have any more forks.

Saving the trained model

As soon as you've felt confident with you constructed and verified designs in a ready for use setting, the first thing to do is to put it into an.h5 or.pkl file using a library like pickles.

Check to make sure Vinegar is configured for the surroundings.

The following process involves dumping the model into an a.pkl file as well as installing the module from there.

Dataset

The database includes 301 unique pieces of data. There are nine data rows in the dataset, and every is described under.

- Name-Car Manufacturer
- Year: The year when it was bought
- Selling_Price - the cost to sell it after using the car
- Present_price - the cost at the time the car was bought
- Kms_Driven - the distance the vehicle has travelled in kilometres



Fuel_Type indicates the type of fuel (petrol, diesel, or CNG).

- Seller_Type: Indicates whether an individual or a dealer is selling it.
- Transmission - the vehicle's gear transmission, whether automatic or manual.
- Owners - The total number of owners who used the vehicle

VI. CONCLUSION

It can be determined that the rising cost of new cars and the constrained financial choices for consumers to buy such contribute to the problem. The consumer market selling used vehicles is expanding globally. The necessity for an approach that is capable of estimating the cost of an older car using a variety of features is therefore critical. Predicting the value of used cars needs great care and an understanding of the types and industry of autos. We accurately anticipated the used automobile cost and carried out an efficiency study using our proposed Decision Tree as the Regressor. Tree Scores on Education Set is 1.0, whereas it is 0.916727846049827 on the Check Set.

REFERENCES

- Peerun, Saamiyah, Chummun, Nushrah, and Pudaruth, Sameerchand are the first. (2015). Artificial Neural Networks for Second-Hand Car Price Prediction.
- [2] Car's Selling Price Prediction Using Random Forest Machine Learning Algorithm by Abhishek Pandey, Vanshika Rastogi, and Sanika Singh (March 1, 2020). 2019 Next Generation Computing Technologies Conference (NGCT-2019), Available at SSRN: <http://dx.doi.org/10.2139/ssrn.3702236> or <https://ssrn.com/abstract=3702236>.
- [3] Pattabiraman Venkatasubbu and Mukkesh Shankar. (2019). Using Instructor-Assisted Methods to Predict Refurbished Car Prices. 9. The International Journal of Engineering and Advanced Technology (ijeat.A1042.1291S319).
- [4] Notes regarding the distribution: The following piece appeared in the International Journal of Trend in Scientific Research and Development (ijtsrd), Volume 5 | Issue 4, June 2021.
URL of the paper: <https://www.ijtsrd.com/computer-science/dataprocessing/42421/prediction-of-car-price-using-linear-regression/ravishastri>
- [5] <https://datasets.kaggle.com>
- [6] N. Monburinon, P. Chertchom, T. Kaewkiriya, S. Rungpheung, S. Buya, and P. Boonpou, "Prediction of prices for used cars by using regression models," 2018 5th International Conference on Business and Industrial Research (ICBIR), pp. 115–119, doi: 10.1109/ICBIR.2018.8391177.
- [7] "Price evaluation model in secondhand car system based on BP neural network theory," 2017 18th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), pp. 431–436, doi: 10.1109/SNPD.2017.8022758.
- [8] C. V. Narayana, C. L. Likhitha, S. Bademiya, and K. Kusumanjali, "Machine Learning Techniques To Predict The Price Of Used Cars: Predictive Analytics in Retail Business," 2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 1680–1687.
- [9] Using SVM Techniques, Prediction of Vehicle Prices S.E. Viswapriya, Durbaka Sai Sandeep Sharma, and Gandavarapu Sathya Kiran published an article in the June 2020 issue of the International Journal of Innovative Technology and Exploring Engineering (IJITEE), volume 9 issue 8.



SJIF Scientific Journal Impact Factor

Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details