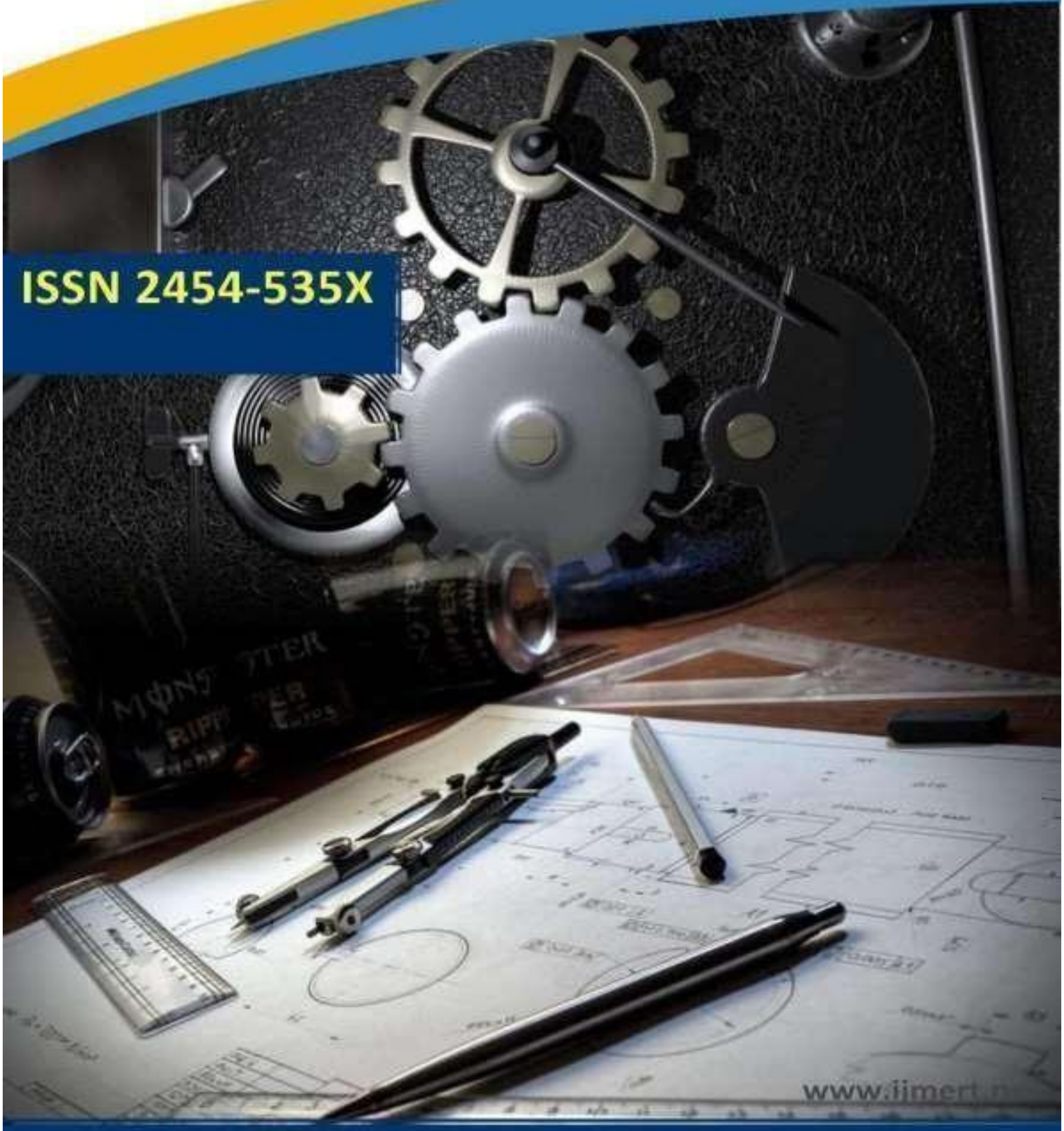
International Journal of
Mechanical Engineering Research and Technology



ISSN 2454-535X

Email ID: info.ijmert@gmail.com or editor@ijmert.net

ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 2, 2024

PREDICTION OF USED CAR PRICES

AVINASH JADHAV[1], ANDUGALA PRASHANTH[2], RAIPALLY MADHURI[3], GUTHA AJAYKUMAR[4], KUNCHALA SRIHARI[5]

[1]Assistant professor, Dept. of CSE, Malla Reddy College of Engineering HYDERABAD.

[2,3,4,5]UG Students, Department of CSE, Malla Reddy College of Engineering HYDERABAD.

ABSTRACT

Over the last decade, the number of cars driving on Mauritius's roads has actually been steadily increasing, rising by 5% annually. The National Transport Authority received 173,954 vehicle registrations in 2014. As a result, 1 in 6 Mauritian adults own a vehicle, with many of those vehicles being pre-owned or replaced vehicles. In this study, we want to find out whether it's possible to utilise artificial semantic networks to forecast the value of used vehicles. As a result, four separate maker learning models were given data pertaining to 200 automobiles culled from diverse sources. Results were somewhat better using support vector equipment regression compared to semantic networks and straight regression, according to our findings. However, for more expensive vehicles in particular, several of the predicted values are far off from the actual pricing. Consequently, more investigations using a bigger dataset are necessary, as is a great deal of trial and error with other types of networks and frameworks, to get much improved predictions.

Keywords: ML, DL, SVM, car price, Linear regression.

I INTRODUCTION

Obtained from the National Transportation Authority (2014), the number of cars increased by 254% from 2003 (68,524) to 2014 (173, 954), as shown in Number 1. Given that new cars and trucks make up such a small percentage of the overall number of vehicles sold each year, it is reasonable to assume that the sale of used imported (refurbished) automobiles and pre-owned previously owned vehicles has eventually increased. When purchasing a new automobile in Mauritius, most

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 2, 2024

buyers are also interested in finding out how much their vehicle will be worth when they're ready to sell it on the used car market. A number of factors influence the prediction of used car prices. Year of production, manufacturer, model, gas economy, horsepower, and nation of origin are the most crucial ones. Additional considerations include the following: fuel type and quantity per use, braking system type, acceleration, interior style, vehicle condition, number of cylinders (in cubic centimetres), vehicle size, number of doors, vehicle weight, customer reviews, paint colour and type, gearbox type, sports car status, audio system, planetary wheels, power steering, air conditioning, GPS navigation, safety and security index, and so on. Whether the vehicle has been in any serious accidents and the identities of its former owners are two of the distinctive factors often considered in a Mauritius setting. Predicting the rate of used cars is, therefore, a quite commendable company. This research will investigate the feasibility of using semantic networks for used car rate prediction. Additional approaches, such as support vector regression and straight regression, will also be used to compare the results. In accordance with, this paper continues. Several service semantic networks and cost forecasts have been consolidated in this system. In this system, the procedure and the data gathering are detailed. The technology calculates what used car prices are likely to be. A conclusion and some suggestions for further research round out the report.

II SURVEY OF RESEARCH

[1] Predicting the value of pre-owned vehicles has been the subject of several studies. It is common practice for researchers to use historical data to estimate product prices. In order to forecast the pricing of used automobiles in Mauritius, Pudaruth used a number of methods, including decision trees, k-nearest neighbours, Ignorant Bayes, and direct regression. The cars in question were not brand new. When several methods' predictions were compared, it became clear that their prices are rather close to one another. Decision trees and the Naive Bayes method were both shown to fail miserably when it came to predicting numerical values. The small sample size does not provide excellent

International Journal of Mechanical Engineering Research and Technology

ISSN 2454 – 535X

http://www.ijmert.com

Vol.16 Issue 2, 2024

prediction accuracy, according to Pudaruth's study.

[2] in A multivariate regression model that helps with numerical value classification and forecasting was presented by Kuiper, S. (2008). In it, the author shows how to predict the GM truck rate in 2005 using this multivariate regression model. Automobile rate forecasting does not need specialised knowledge. So, it is possible to predict prices using the data that is readily accessible online. In the same vein as the article's car rate forecast, the author offered variable selection techniques to help identify which variables would work best in the model.

As a method for predicting the prices of secondhand vehicles using Random Forest, Pal et al. [3] found in 2019. With an accuracy of 83.62% for test data and 95% for train-data, the Kaggle information established was utilised to forecast used-car prices in this research. After eliminating outliers and uninteresting characteristics from the dataset, the most relevant functions used for this prediction were price, km, brand name, and truck type. As a novel implementation, Random Woodland outperformed earlier work with similar datasets in terms of accuracy.

[4] The need to create a model to predict the cost of used cars in Bosnia and Herzegovina is shown by Gegic, E. et al. (2019). They used AI methods including made-up semantic networks, assistance vector devices, and made-up woods. Still, all of those methods were used in tandem. To get the data for the projection, we used a web scraper that was made in PHP to collect information from the autopijaca.ba website. In order to determine which method was the best match for the given data, we compared the individual algorithms' results. The final prediction model was a Java programme. The design was further evaluated using examination data, which confirmed its accuracy of 87.38%. In 2019, Dholiya et al. showcased a method for reselling cars that relies on equipment learning.

Giving the person a realistic idea of how much the vehicle may cost is the main objective of the method developed by Dholiya, M., et al. The system, which is an online application, may also provide the user with a list of options for different types of cars based on the details of the vehicle they are attempting

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 2, 2024

to discover. By doing so, it helps provide the buyer or seller with useful information upon which to make a decision. The system trains utilising past data acquired over a long period of time, and it uses the multiple direct regression technique to create projections. At first, the KDD (Knowledge Exploration in Databases) method was used to gather the raw data. Afterwards, it cleaned and preprocessed the data in an effort to find meaningful patterns, which it subsequently used to draw conclusions.

III PROPOSED SYSTEM

Because, like the prices of other items, the prices of autos also change with time, data for this research has been sourced from a variety of automotive websites as well as the small ads sections seen in regular newspapers. One hundred and twenty records were gathered. Old vehicles' production year (YEAR), make (MAKE), engine capacity (ENGINE) measured in cubic centimetres, paint type (typical or metallic), gearbox type (guidebook or automatic), mileage (GAS MILEAGE) and price (RATE) in Mauritian rupees are all part of the detailed information. A large number of tests were conducted to determine the optimal semantic network criteria and network topology. Out of all the semantic network frameworks that were tested, we found that a network with just one hidden layer and two nodes made the most insignificant error. However, out of these four approaches, k-Nearest Neighbour had the lowest accuracy, while Support Vector Regression and a multi layer understanding with back-propagation made somewhat better predictions than linear regression. A cross validation value of ten folds was used to conduct all trials.

IV WORKING METHODOLOGY

This part contains the research study methodology. In order to compile the vehicle dataset for this analysis, olx.com was used. Every vehicle was filmed along with its make, version, vendor, mileage, year of production, fuel type, and price. Table 1 displays a sample of the collected data. There will probably be a lot of used car data in these databases, so they will need some engineering and tweaking. Excluding duplicate data is a good first step since

International Journal of Mechanical Engineering Research and Technology

ISSN 2454 – 535X

http://www.ijmert.com

Vol.16 Issue 2, 2024

they could affect the model's output. Rate predictions are greatly aided by power and engine. Additionally, new_price is a very good cost forecaster. In predictive statistics and AI, top traits with a high connection coefficient have a stronger impact on the prediction variable. However, this is not always the case. The correlation coefficient, as its name implies, is a statistical measure that describes the relationship between two or more variables. Always between 1 and -1 (positive to negative), with 0 indicating complete disorganisation, is the range of values for the variation of the correlation coefficient between 2 criteria.

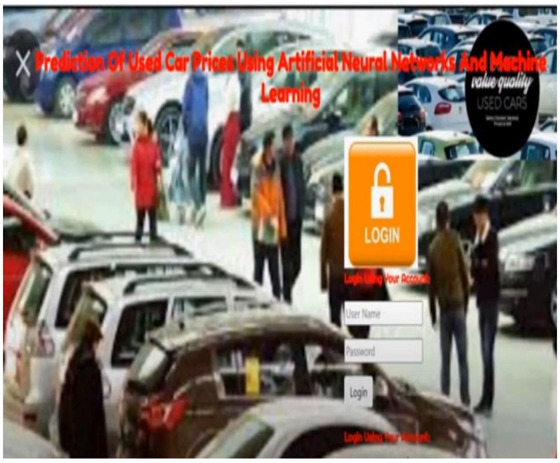


Fig.1. Home page.

This task makes use of the Scikit-learn machine learning package to build a number of AI formulae. We use the same training data to train each design, and we utilise the same screening data to assess it. The findings are defined and compared in the section that follows. For predictable variables, the regression-based method in monitored machine learning is reliable. When X is the independent variable and Y is the dependent variable, it is clear that a single straight regression model can predict Y. In order to predict Y's future value, the design will use the Y-intercept, slope, and noise of the regression line.

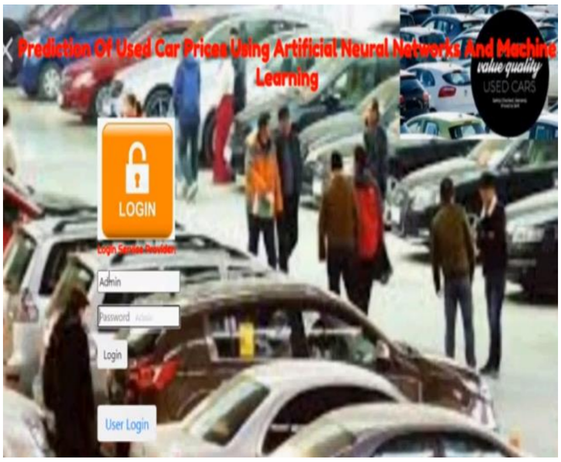


Fig.2. Admin login page.

Fig.3. user registration page.

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 2, 2024

Fig.4. Login page sign in page.

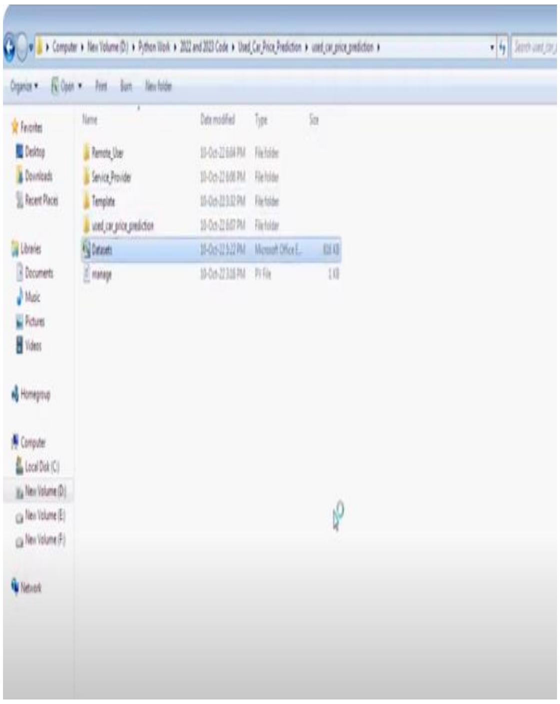


Fig.5. Upload dataset page.

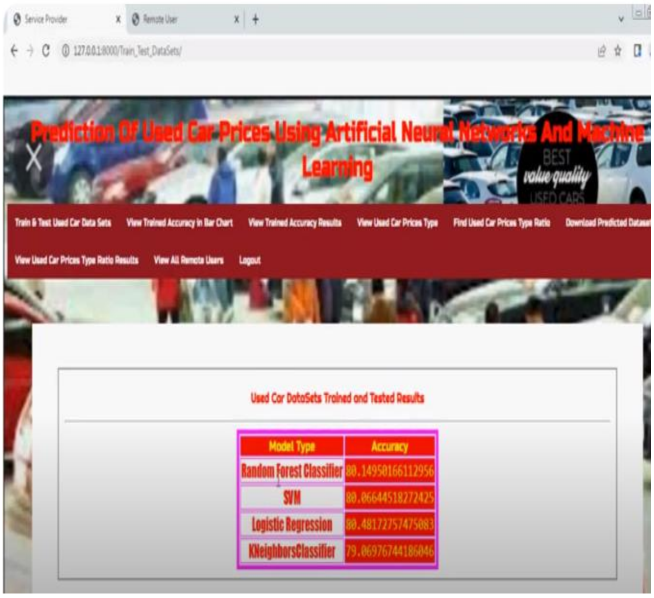


Fig.6. Accuracy details.

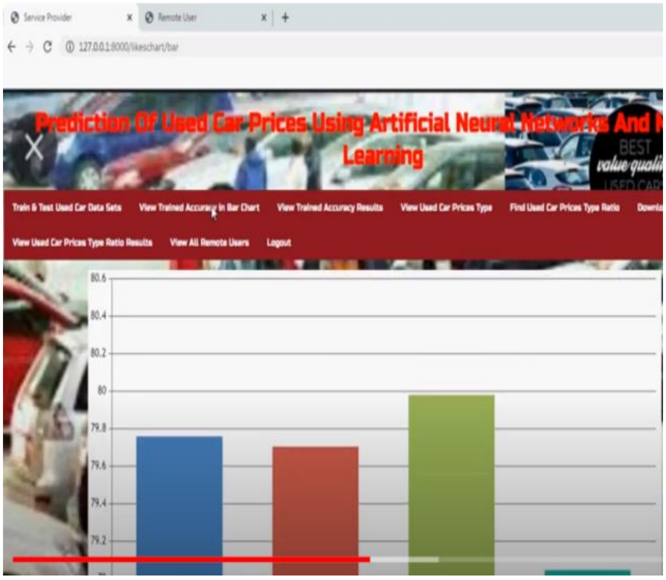


Fig.7 . Graph results.

International Journal of Mechanical Engineering Research and Technology

ISSN 2454 – 535X

http://www.ijmert.com

Vol.16 Issue 2, 2024

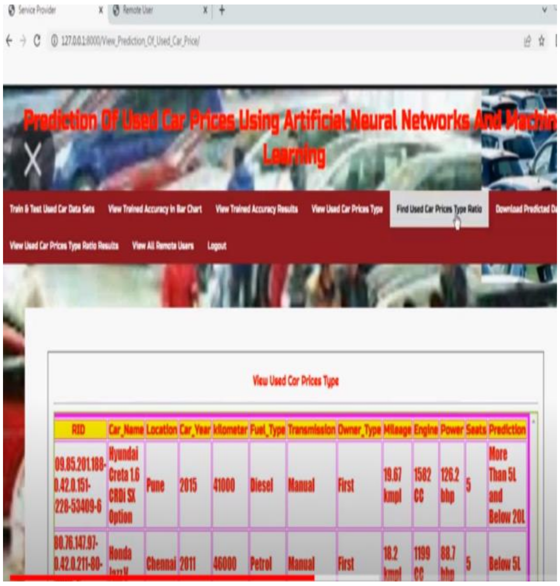


Fig.8. Output results.

CONCLUSION

Predicting the rate of used, refurbished, and second-hand automobile sales in Mauritius was the primary objective of this work. The fact that the car market has been steadily increasing by around 5% over the last decade is evidence of the significant demand for cars among Mauritian citizens. There are a plethora of automotive websites on Mauritius, but none of them provide a tool to estimate the value of pre-owned vehicles according to their specifications. We employed the cross-validation method with a ten-fold increase on our dataset of 200 documents. Using four separate machine learning methods, we can predict the price of used cars based on their make, year, colour, gearbox type, engine capacity, and mileage. With each of the four approaches, we saw a considerable drop in the ordinary residual value. Consequently, we draw the conclusion that predicting the rate of used vehicles is a highly risky but potentially lucrative endeavour. Dealers and owners of automobiles looking to get an idea of their worth will find this approach incredibly useful. We want to use a wider variety of machine learning formulae for future predictions, as well as to gather additional data and characteristics.

ACKNOWLEDGMENT

We thank CMR Technical Campus for supporting this paper titled "**PREDICTION OF USED CAR PRICES USING ARTIFICIAL NEURAL NETWORKS AND MACHINE LEARNING**", which provided good facilities and support to accomplish our work. I sincerely thank our Chairman, Director, Deans, Head of the Department, Department Of Computer Science and Engineering, Guide and Teaching and Non- Teaching faculty members for giving valuable

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 2, 2024

suggestions and guidance in every aspect of our work.

REFERANCES

[1] NATIONAL TRANSPORT AUTHORITY. 2015. Available at: http://nta.govmu.org/English/Statistics/Pages/Arch ives.aspx. [Accessed 24 April 2015].

[2] Bharambe, M. M. P., and Dharmadhikari, S. C. (2015) "Stock Market Analysis Based on Artificial Neural Network with Big data". Fourth Post Graduate Conference, 24-25th March 2015, Pune, India.

[3] Pudaruth, S. (2014) "Predicting the Price of Used Cars using Machine Learning Techniques". International Journal of Information & Computation Technology, Vol. 4, No. 7, pp.753- 764.

[4] Jassibi, J., Alborzi, M. and Ghoreshi, F. (2011) "Car Paint Thickness Control using Artificial Neural Network and Regression Method". Journal of Industrial Engineering International, Vol. 7, No. 14, pp. 1-6, November 2010

[5] Ahangar, R. G., Mahmood and Y., Hassen P.M. (2010) "The Comparison of Methods, Artificial Neural Network with Linear Regression using Specific Variables for Prediction Stock Prices in Tehran Stock Exchange". International Journal of Computer Science and Information Security, Vol.7, No. 2, pp. 38-46.

[6] Listiani, M. (2009) "Support Vector Regression Analysis for Price Prediction in a Car Leasing Application". Thesis (MSc). Hamburg University of Technology.

[7] Iseri, A. and Karlik, B. (2009) "An Artificial Neural Network Approach on Automobile Pricing". Expert Systems with Application: ScienceDirect Journal of Informatics, Vol. 36, pp. 155-2160, March 2009.

[8] Yeo, C. A. (2009) "Neural Networks for Automobile Insurance Pricing". Encyclopedia of Information Science and Technology, 2nd Edition, pp. 2794-2800, Australia.

[9] Doganis, P., Alexandridis, A., Patrinos, P. and Sarimveis, H. (2006) "Time Series Sales Forecasting for Short Shelf-life Food Products Based on Artificial Neural Networks and Evolutionary Computing". Journal of Food Engineering, Vol. 75, pp. 196–204.

[10] Rose, D. (2003) "Predicting Car Production using a Neural Network

International Journal of Mechanical Engineering Research and Technology
ISSN 2454 – 535X
http://www.ijmert.com
Vol.16 Issue 2, 2024

Technical Paper- Vetronics (Inhouse)". Thesis, U.S. Army Tank Automotive Research, Development and Engineering Center (TARDEC).

[11] LEXPRESS.MU ONLINE. 2014. [Online] Available at: http://www.lexpress.mu/ [Accessed 23 September 2014].

[12] LE DEFI MEDIA GROUP. 2014. [Online] Available at: http://www.defimedia.info/ [Accessed 23 September 2014].

[13] He, Q. (1999) "Neural Network and its Application in IR". Thesis (BSc). University of Illinois.

[14] Cheng, B. and Titterington, D. M. (1994). "Neural Networks: A Review from a Statistical Perspective". Statistical Science, Vol. 9, pp. 2-54.

[15] Anyaeche, C. O. (2013). "Predicting Performance Measures using Linear Regression and Neural Network: A Comparison". African Journal of Engineering Research, Vol. 1, No. 3, pp. 84-89.