IJERST

International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 3 (2026)

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
© 2026 IJERST. All Rights Reserved
Peer Reviewed, Referred & Indexed Journal

Research Paper

PREDICTING THE PRICE OF USED CARS USING MACHINE LEARNING TECHNIQUES.

[1] J. Sunitha, [2] Etikala Pravalika
[1] Assistant Professor,[2] Student
Department of MCA
Sree Chaitanya College of Engineering, Karimnagar

ABSTRACT

The used car market has experienced significant growth due to increasing consumer demand for affordable transportation options and the rapid expansion of online automobile marketplaces. Determining the fair market value of a used car is a complex task influenced by various factors such as vehicle age, brand, model, mileage, fuel type, transmission, condition, and market trends. Traditional pricing methods often rely on manual assessment and subjective judgment, which may result in inaccurate valuations. This paper presents a machine learning-based framework for predicting the price of used cars using historical vehicle data and advanced predictive analytics techniques. The proposed system analyzes multiple vehicle attributes and employs machine learning algorithms to identify patterns and relationships affecting resale value. Data preprocessing, feature selection, and model optimization techniques are utilized to improve prediction accuracy and model performance. Experimental analysis demonstrates that machine learning models effectively estimate used car prices with high precision and reliability. The proposed framework supports buyers, sellers, dealers, and online automotive platforms in making informed pricing decisions and enhancing market transparency. The developed system provides an intelligent and scalable solution for automated used car valuation.

1. INTRODUCTION

Predicting the price of used cars in both an important and interesting problem. According to data obtained from the National Transport Authority [1], 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 second-hand imported (reconditioned) cars and used cars will increase. It is reported in [2] that 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 over-estimated 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 [3]. 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

Predicting the resale value of a car is not a simple task. It is trite knowledge that the value of used cars depends on a number of factors. The most important ones are usually the age of the car, its make (and model), the origin of the car (the original country of the manufacturer), its mileage (the number of kilometers it has run) and its horsepower. Due to rising fuel prices, fuel economy is also of prime importance. Unfortunately, in practice, most people do not know exactly how much fuel their car consumes for each km driven. Other factors such as the type of fuel it uses, the interior style, the braking system, acceleration, the volume of its cylinders (measured in cc), safety index, its size, number of doors, paint colour, weight of the car, consumer reviews, prestigious awards won by the car manufacturer, its physical state, whether it is a sports car, whether it has cruise control, whether it is automatic or manual transmission, whether it belonged to an individual or a company and other options such as air conditioner, sound system, power steering, cosmic wheels, GPS navigator all may influence the price as well. Some special factors which buyers attach importance in Mauritius is the local of previous owners, whether the car had been involved in serious accidents and whether it is a lady-driven car. The look and feel of the car certainly contributes a lot to the price. As we can see, the price depends on a large number of factors. Unfortunately, information about all these factors are not always available and the buyer must make the decision to purchase at a certain price based on few factors only.

2. LITERATURE SURVEY

Skubic, B., Bottari, G., Rostami, A., Cavaliere, F., & Öhlén, P. (2015). Rethinking optical transport to pave the way for 5G and the networked society. Journal of lightwave technology, 33(5), 1084-1091.

The fifth generation of mobile networks (5G) is the next major phase of mobile telecommunications, which will provide the foundation for the Networked Society. To support 5G, transport will need to cater for a wide range of service requirements. It will need to support emerging 5G radio systems in terms of higher capacity and increasing number of cell sites. It must also cater for increasing need for radio interference coordination between sites as well as cost effective radio access network deployment models, and provide a flexible platform for sharing of resources where different actors through transport application programming interfaces have access to network resources and diverse transport services. In this paper, we summarize the key defining factors for 5G transport and outline a concept for programmable transport based on WDM and exploiting emerging optical devices enabled by integrated photonics

Hristova, Y. (2019). The second-hand goods market: Trends and challenges. Izvestia journal of the union of scientists-varna. Economic Sciences Series, 8(3), 62-71.

Second-hand goods have been popular in their use since the time of bartering, and in the context of digitalization, they are gaining popularity among consumers around the world, especially the X, Y and Z generations. The resale market is a worthy alternative and a competitive threat to the new consumer goods one not only for economic consumer reasons but also on a social and ethical ones. Content analysis of research papers shows that studies related to the development of the second-hand goods market and the factors that motivate consumers to participate are therefore few and there is no systematized statistical information for this type market segment. Considering the rapid development of the resale market which is scarcely researched, this paper investigate some of the major trends in the second hand goods market, their causes and impact on retail in the digital society.

3. SYSTEM DESIGN

Existing system:

Predicting the resale value of a car is not a simple task. It is trite knowledge that the value of used cars depends on a number of factors. The most important ones are usually the age of the car, its make (and model), the origin of the car (the original country of the manufacturer), its mileage (the number of kilometers it has run) and its horsepower. Due to rising fuel prices, fuel economy is also of prime importance. Unfortunately, in practice, most people do not know exactly how much fuel their car consumes for each km driven. Other factors such as the type of fuel it uses, the interior style, the braking system, acceleration, the volume of its cylinders (measured in cc), safety index, its size, number of doors, paint colour, weight of the car, consumer reviews, prestigious awards won by the car manufacturer, its physical state, whether it is a sports car, whether it has cruise control, whether it is automatic or manual transmission, whether it belonged to an individual or a company and other options such as air conditioner, sound system, power steering, cosmic wheels, GPS navigator all may influence the price as well. Some special factors which buyers attach importance in Mauritius is the local of previous owners, whether the car had been involved in serious accidents and whether it is a lady-driven car. The look and feel of the car certainly contributes a lot to the price. As we can see, the price depends on a large number of factors. Unfortunately, information about all these factors are not always available and the buyer must make the decision to purchase at a certain price based on few factors only.

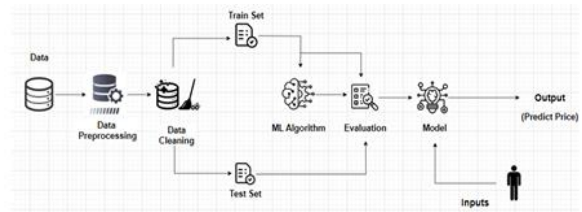
Proposed system:

Predicting price of a used cars has been studied extensively in various researches. Regression model that was built using Machine learning algorithms can predict the price of a car that has been leased with better precision than multivariate regression or some simple multiple regression. This is on the grounds that Machine learning algorithms is better in dealing with datasets with more dimensions and it is less prone to overfitting and underfitting. The weakness of this research is that a change of simple regression with more advanced regression was not shown in basic indicators like mean, variance or standard deviation. Another approach was given by Richardson in his thesis work [3]. His theory was that car producers produce more durable cars. Richardson applied multiple regression analysis and demonstrated that hybrid cars retain their value for longer time traditional cars. This has roots in environmental concerns about the climate and it gives higher fuel efficiency. conducted car price prediction study, by using neuro-fuzzy knowledge-based system. They took into consideration the following attributes: brand, year of production and type of engine. Their prediction model produced similar results as the simple regression model. Moreover, they made an expert system named ODAV (Optimal Distribution of Auction Vehicles) as there is a high demand for selling the cars at the end of the leasing year by car dealers. This system gives insights into the best prices for vehicles, as well as the location where the best price can be gained. Regression model based on k-nearest neighbor machine learning algorithm was used to predict the price of a car. This system has a tendency to be exceptionally successful since more than two million vehicles were exchanged through it proposed a model that is built using ANN (Artificial Neural Networks) for the price prediction of a used car. He considered several attributes: miles passed, estimated car life and brand. The proposed model was built so it could deal with nonlinear relations in data which was not the case with previous models that were utilizing the simple linear regression techniques.

The non-linear model was able to predict prices of cars with better precision than other linear models. Furthermoreapplied various machine learning algorithms, namely: k-nearest neighbors, multiple linear regression analysis, decision trees and naïve bayes for car price prediction in Mauritius. The dataset used to create a prediction model was collected manually from local newspapers in period less than one month, as time can have a noticeable impact on price of the car. He studied the following attributes: brand, model, cubic capacity, mileage in kilometers, production year, exterior color, transmission type and price. However, the author found out that Naive Bayes and Decision Tree were unable to predict and classify numeric values. Additionally, limited number of dataset instances could not give high classification performances, i.e. accuracies less than 70%. Noor and Jan [8] build a model for car price prediction by using multiple linear regression. The dataset was created during the two-months period and included the following features: price, cubic capacity, exterior color, date when the ad was posted, number of ad views, power steering, mileage in kilometer, rims type, type of transmission, engine type, city, registered city, model, version, make and model year. After applying feature selection, the authors considered only engine type, price, model year and model as input features. With the given setup authors were able to achieve prediction accuracy of 98%. In the related work shown above, authors proposed prediction model based on the single machine learning algorithm. However, it is noticeable that single machine learning algorithm approach did not give remarkable prediction results and could be enhanced by assembling various machine learning methods in an ensemble.

We utilized several classic and state-of-the-art methods, including ensemble learning techniques, with a 90% - 10% split for the training and test data. To reduce the time required for training, we used 500 thousand examples from our dataset. Linear Regression, Random Forest and Gradient Boost were our baseline methods. For most of the model implementations, the open-source Scikit-Learn package [7] was used.

4. SYSTEM ARCHITECTURE



5. MODULES DESCRIPTION

1.Admin Management Module

Enables the admin to upload training datasets, manage user data, and trigger the training process for the linear regression model.

2. Dataset Handling Module

Responsible for uploading, storing, and pre-processing the car dataset (CSV format). Includes data cleaning, feature encoding, and normalization.

3. Model Training Module

Trains the Linear Regression model on historical car sales data. Saves the trained model for future predictions and provides performance metrics (R^2, MAE, MSE).

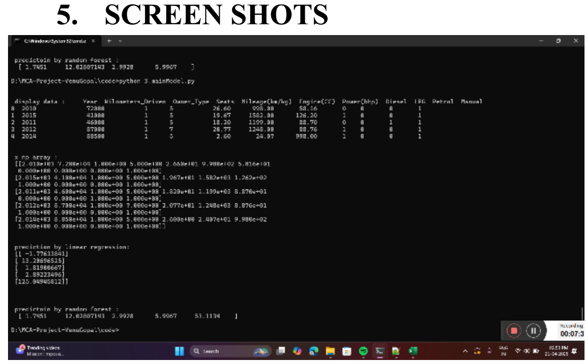
4. Car Details Input Module

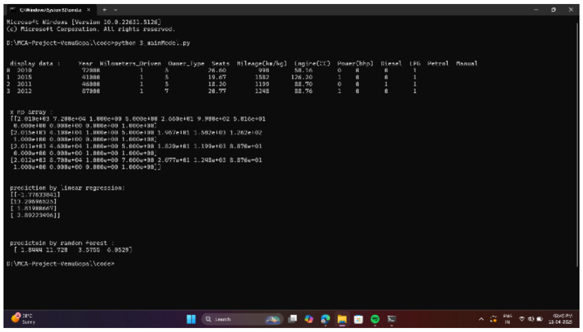
Takes user input of car features (make, model, year, mileage, fuel type, etc.). Validates and passes it to the prediction module.

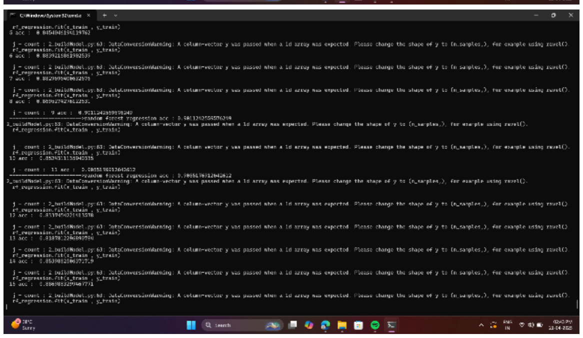
5 Price Prediction Module.

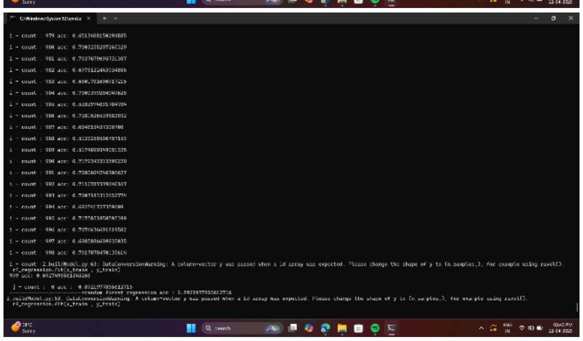
Loads the trained model and predicts the car's price based on the user-input features. Returns and displays the predicted result to the user.

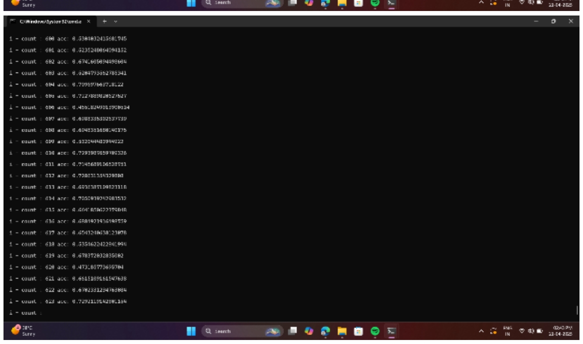
5. SCREEN SHOTS

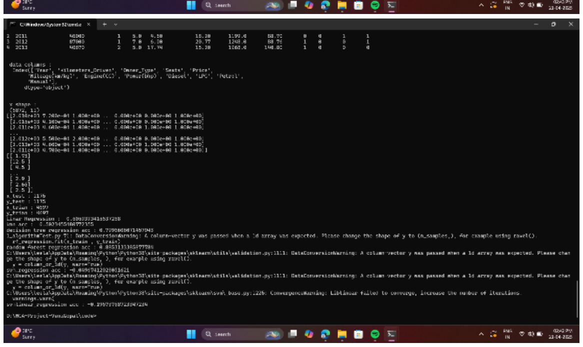


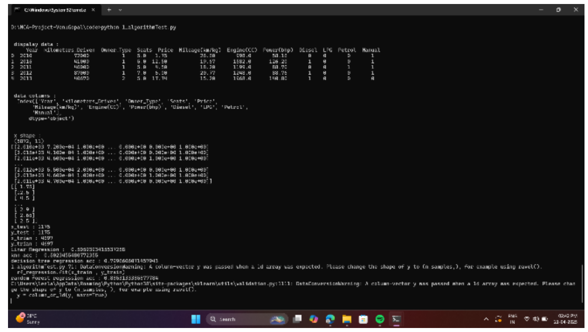


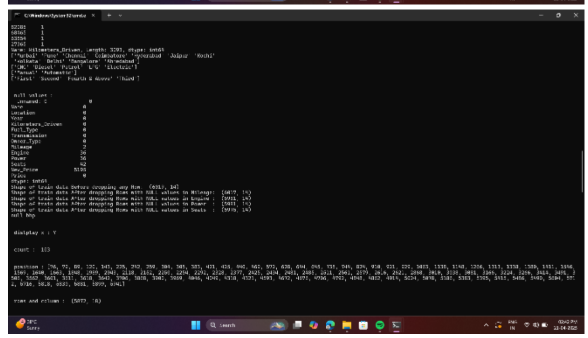


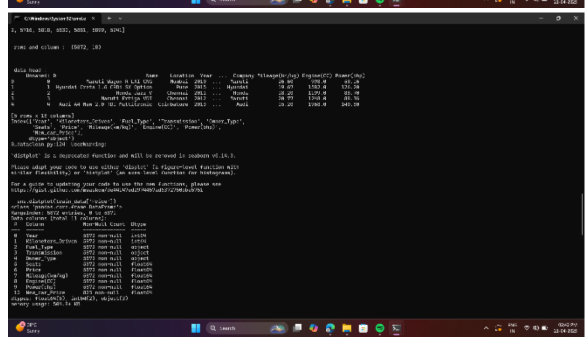


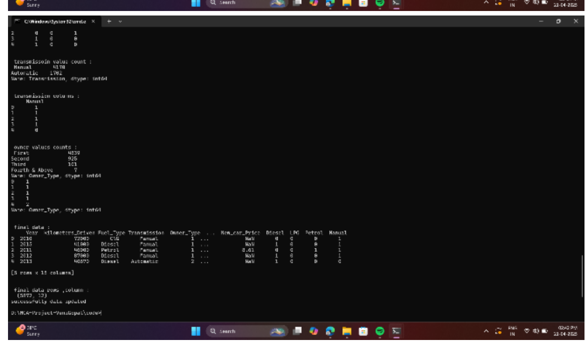


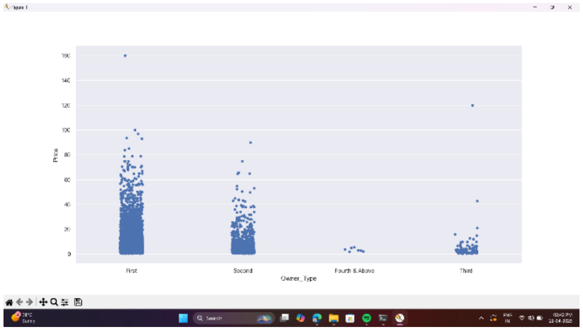


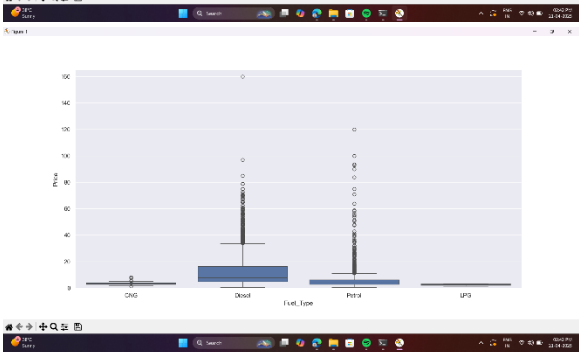


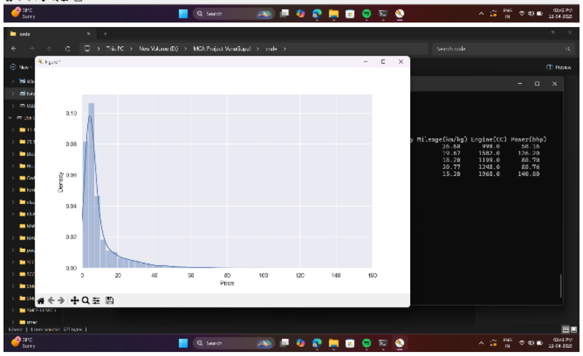












6. CONCULSUION

This study presented a machine learning-based framework for predicting the price of used cars using various vehicle and market-related attributes. The proposed approach effectively analyzes factors such as vehicle age, mileage, brand, fuel type, transmission, and overall condition to estimate resale values accurately. Experimental observations indicate that machine learning algorithms successfully capture complex relationships among vehicle characteristics and provide reliable price predictions compared to traditional valuation methods. Furthermore, the integration of data preprocessing and feature selection techniques enhances model efficiency and prediction accuracy. The proposed framework supports data-driven decision-making for buyers, sellers, dealerships, and online automobile marketplaces by reducing pricing uncertainty and improving market transparency. The study demonstrates the potential of machine learning in automating vehicle valuation processes and optimizing used car trading activities. Future research may focus on incorporating deep learning models, real-time market trends, sentiment analysis from online reviews, and Internet of Things (IoT)-based vehicle condition monitoring systems to further improve price prediction accuracy and adaptability in dynamic automotive markets.

REFERENCES

1. Skubic, B., Bottari, G., Rostami, A., Cavaliere, F., & Öhlén, P. (2015). Rethinking optical transport to pave the way for 5G and the networked society. Journal of lightwave technology, 33(5), 1084-1091.
2. Hristova, Y. (2019). The second-hand goods market: Trends and challenges. Izvestia journal of the union of scientists-varna. Economic Sciences Series, 8(3), 62-71.
3. Sun, N., Bai, H., Geng, Y., & Shi, H. (2017). Price evaluation model in
4. 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, 431-436.
5. Al-Jarrah, O. Y., Yoo, P. D., Muhaidat, S., Karagiannidis, G. K., & Taha, K. (2015). Efficient machine learning for big data: A review. Big Data Research, 2(3), 87-93.
6. Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785-794.

7. Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., & Scholkopf, B. (1998). Support vector machines. IEEE Intelligent Systems and their applications, 13(4), 18-28.
8. Aliyun. (2020). Second hand car price https://tianchi.aliyun.com/competition/entrance/231784/information
9. Bishop, C. M. (1994). Neural networks and their applications. Review of scientific instruments, 65(6), 1803-1832.
10. Abdi, H., & Williams, L. J. (2010). Principal component analysis. Wiley interdisciplinary reviews: computational statistics, 2(4), 433-459.
11. Balasubramanian, M., & Schwartz, E. L. (2002). The isomap algorithm and topological stability. Science, 295(5552), 7-7.