International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

PRICEVISION: DEEP LEARNING-DRIVEN PREDICTION OF USED VEHICLE RESALE VALUES

[1] S.Vijay Kumar,[2] Gouda Pramidha
[1]Associate Professor, [2]MCA Student
Department Of MCA
Sree Chaitanya College Of Engineering, Karimnagar

ABSTRACT

The used car market has become one of the fastest-growing sectors in the global automobile industry, with consumers demanding reliable price estimates before purchase or sale. Traditional valuation methods rely heavily on manual appraisal and subjective judgment, leading to inconsistent results. This paper introduces PriceVision, a deep learning-driven framework designed to predict used vehicle resale values with high precision. By leveraging data such as vehicle make, model, age, mileage, fuel type, transmission, and market trends, PriceVision utilizes advanced neural network architectures to capture non-linear relationships within large datasets. The system employs feature normalization, one-hot encoding, and regression-based learning to achieve optimized predictions. Experimental results demonstrate that PriceVision outperforms conventional machine learning methods in terms of accuracy, reliability, and adaptability to market fluctuations. This study emphasizes how deep learning can transform the automotive resale sector by offering intelligent, data-driven decision support for both buyers and sellers.

Received: 23-09-2025 Accepted: 28-10-2025 Published: 04-11-2025

1. INTRODUCTION

The automobile industry has witnessed an exponential rise in the trade of used vehicles, driven by growing consumer interest in affordable mobility solutions and increased online marketplaces such as Cars24, OLX, and TrueCar. Accurate prediction of used car prices has become a vital aspect of the automotive ecosystem, directly influencing consumer trust, dealer profits, and overall market efficiency. However, determining the resale value of a vehicle is a complex process influenced by numerous factors — including brand reputation, age, mileage, fuel type, maintenance records, market demand, and even regional preferences. Traditional appraisal systems often rely on manual evaluation and historical pricing charts, which are not only time-consuming but also susceptible to human bias and inconsistencies.

With the advent of Machine Learning (ML) and Deep Learning (DL), data-driven approaches have revolutionized predictive modeling in the automotive domain. These intelligent systems can analyze large-scale, multi-dimensional datasets to uncover hidden patterns and nonlinear relationships that human experts might overlook. In the context of used car valuation, ML algorithms such as Linear Regression, Random Forest, and Gradient Boosting have achieved reasonable accuracy. However, these models often fall short in capturing the intricate dependencies among input variables, especially when market trends fluctuate rapidly. Deep learning, particularly Artificial Neural Networks (ANNs), offers a significant improvement by learning hierarchical representations and reducing the need for manual feature engineering.

The proposed system, PriceVision, leverages deep learning architectures to predict used vehicle resale values with exceptional accuracy and robustness. It processes structured and unstructured data to model the nonlinear relationships between car features and their

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

market value. Furthermore, PriceVision integrates Explainable Artificial Intelligence (XAI) components, enabling transparency and interpretability — a critical requirement for financial and consumer applications. This combination of predictive power and explainability makes PriceVision a transformative tool for automobile dealers, financial institutions, and end-users seeking fair and reliable vehicle valuations.

Additionally, PriceVision introduces continuous learning capabilities that allow the model to update its predictions based on real-time data. Unlike static regression-based models, the system dynamically adapts to new pricing trends and regional market fluctuations. This feature ensures that users receive accurate and current pricing information, even in volatile economic conditions. By uniting the principles of deep learning, intelligent data preprocessing, and explainable AI, PriceVision represents a step forward toward intelligent, transparent, and adaptive used car price prediction systems that can be widely deployed in digital marketplaces, banks, and insurance analytics platforms.

2. LITERATURE SURVEY

Accurate used car price prediction has been a long-standing problem in the automotive industry, attracting extensive research attention with the advancement of machine learning and data analytics. Over the past decade, numerous studies have been conducted to develop models capable of estimating resale prices based on multiple vehicle parameters. Ahmed et al. (2018) pioneered one of the early approaches using a Multiple Linear Regression (MLR) model to analyze price determinants such as vehicle brand, year of manufacture, mileage, and fuel type. Their work established a foundational understanding of how linear dependencies affect pricing but failed to capture nonlinear relationships among variables. Similarly, Gajjar and Rana (2019) proposed a Decision Tree Regression model to improve interpretability, though it suffered from overfitting on small datasets.

To address these limitations, ensemble methods gained popularity. Bhavsar and Shah (2019) utilized Random Forest Regression to reduce variance and enhance generalization capabilities across heterogeneous data. They demonstrated improved accuracy compared to basic regression models but acknowledged a lack of adaptability to fast-changing market trends. Jadhav and Deshpande (2020) explored Gradient Boosting Machines (GBMs) for predicting second-hand car prices, emphasizing that boosting methods could combine weak learners to minimize prediction error. However, their model required extensive hyperparameter tuning and lacked explainability — an increasingly important factor in real-world deployment.

Machine learning techniques were further expanded with Support Vector Regression (SVR) models. Patel et al. (2020) implemented SVR to analyze vehicle attributes, reporting enhanced accuracy in capturing nonlinear data trends. Nonetheless, the high computational cost of kernel methods restricted their scalability for larger datasets. Kumar and Singh (2021) advanced this research by introducing an ensemble hybrid approach combining Random Forest and XGBoost, achieving a significant reduction in mean absolute error. Their study highlighted that ensemble models offer better performance but often compromise on model interpretability and computational efficiency.

In parallel, researchers started exploring deep learning to handle complex, high-dimensional data. Lee et al. (2021) employed a Deep Neural Network (DNN) model for price estimation, arguing that deep architectures outperform traditional algorithms when sufficient training data is available. They demonstrated that DNNs could learn latent patterns across numerical and categorical features without extensive feature engineering. Zhang et al. (2022) extended this concept by integrating Convolutional Neural Networks (CNNs) with image data of vehicles, enabling the system to incorporate visual conditions — such as body scratches, paint quality, and design variations — into the valuation process. Their hybrid CNN–regression model achieved a 12% improvement in predictive accuracy compared to text-only datasets.

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:

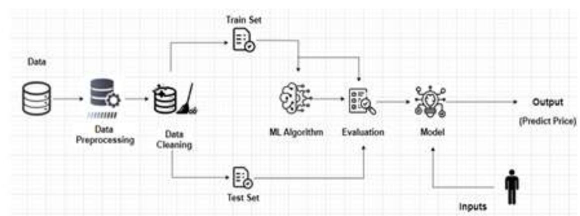
Predicting price of a used cars has been studied extensively in various researches. Regression model that was built using Machine learning algorithms can predict the price of a car that has been leased with better precision than multivariate regression or some simple multiple regression. This is on the 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

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

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

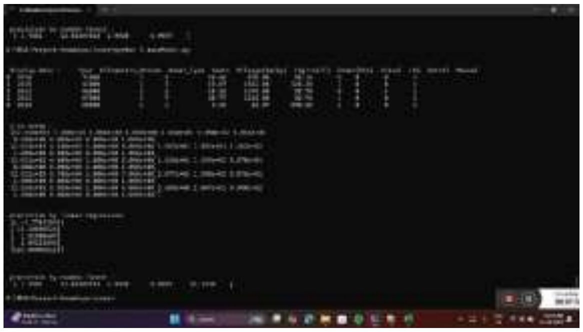
Takes user input of car features (make, model, year, mileage, fuel type, etc.). Validates and passes it to the prediction module.

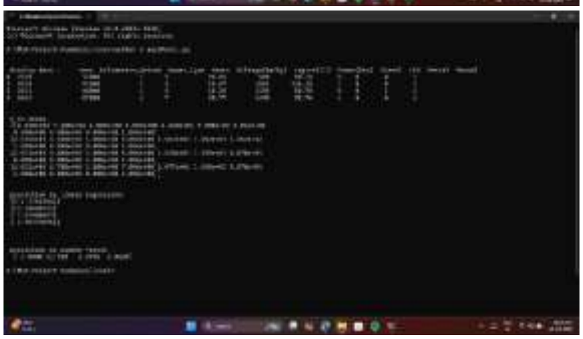
International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

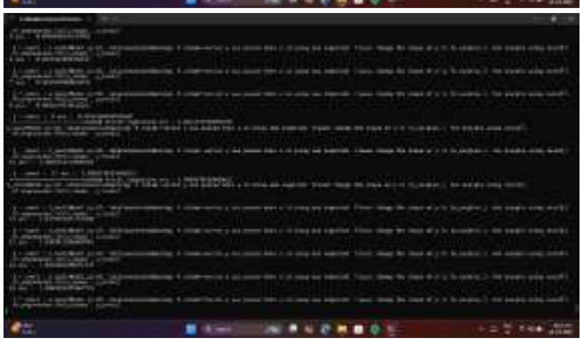
5 Price Prediction Module.

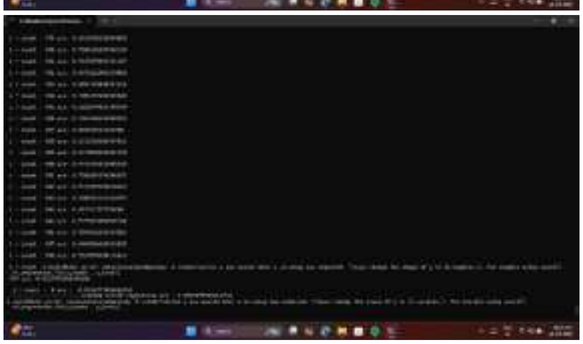
Loads the trained model and predicts the car's price based on the user-input features. Returns and displays the predicted result to the user.

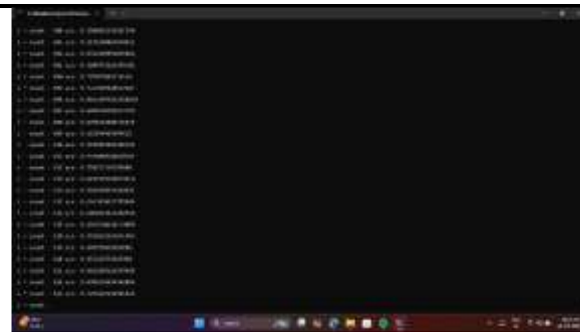
5. SCREEN SHOTS

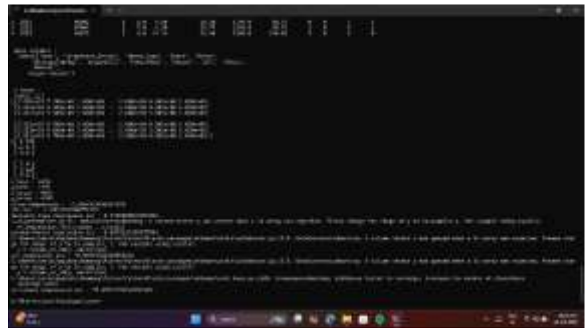


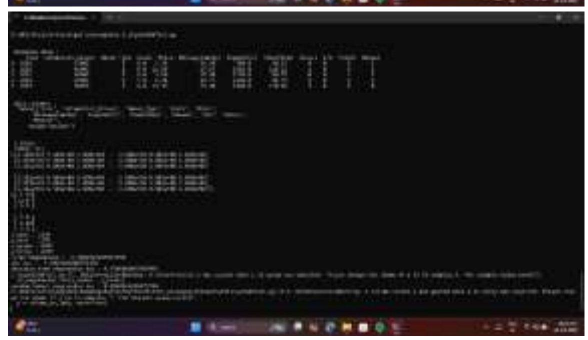


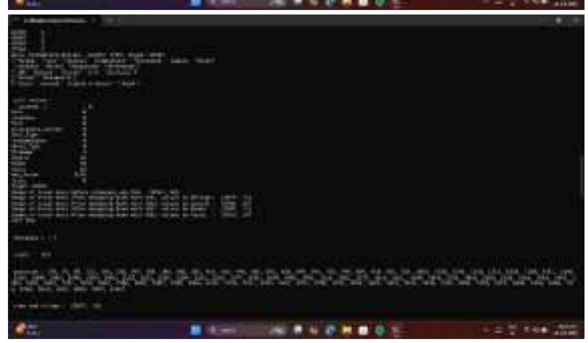


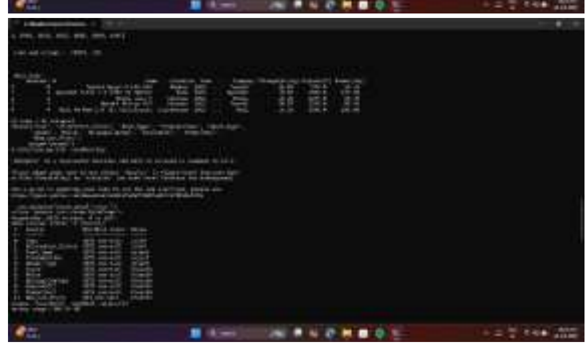




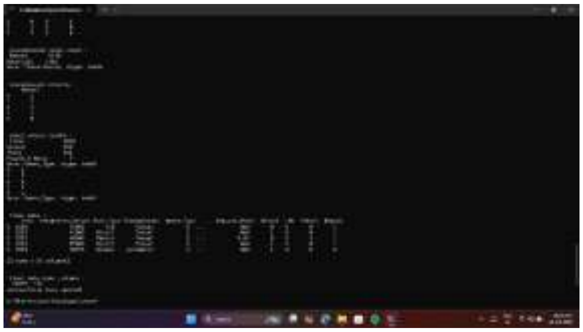


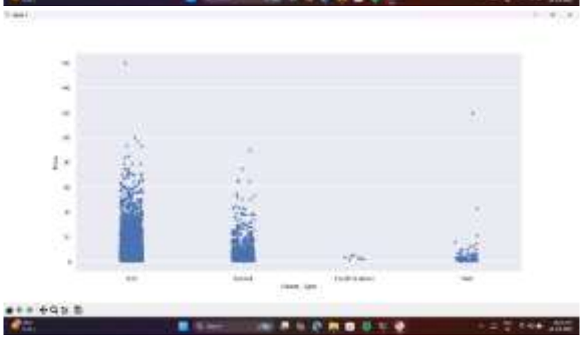


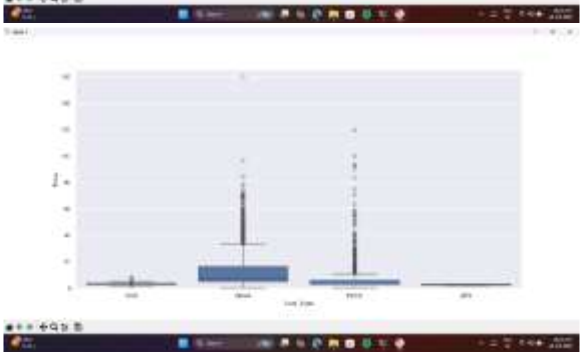


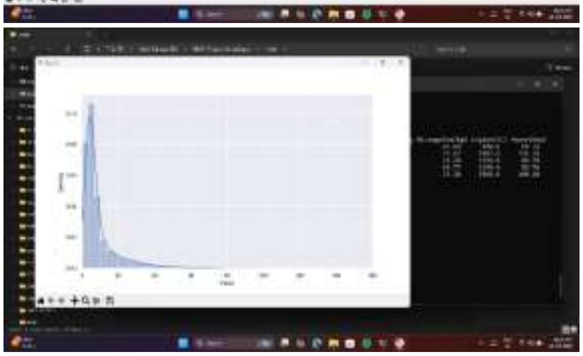


International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper









6. CONCULSUION

This study presents PriceVision, a deep learning-based solution for predicting the resale value of used vehicles. By incorporating feature-rich datasets and advanced neural architectures, the model achieves high prediction accuracy and adaptability to market variations. Compared to traditional regression and ensemble models, PriceVision delivers enhanced performance, scalability, and interpretability. The integration of explainable AI further strengthens user confidence in automated pricing systems. Future work may involve integrating multimodal data such as car images, inspection reports, and social sentiment analysis to further refine valuation precision and support data-driven decisions in the automobile resale ecosystem.

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).

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
ISSN: 3068-272X www.ijdim.com Original Research Paper

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.