



Car Price Prediction Using Machine Learning

Snehil Gupta¹, Dr. Sowmya K S²

¹Dept. of Information Science, B.M.S College of Engineering Bangalore, India snehilgupta.is20@bmsce.ac.in

²Dept. of Information Science, B.M.S College of Engineering Bangalore, India sowmyaks.ise@bmsce.ac.in

ABSTRACT—

This research paper explores the application of machine learning techniques for predicting car prices, presenting a comprehensive analysis of various algorithms and methodologies. Leveraging a dataset encompassing diverse features, the study employs machine learning models to establish predictive patterns and correlations within the automotive market. The research aims to enhance understanding and accuracy in forecasting car prices, contributing valuable insights to the intersection of machine learning and automotive economics. Results and findings shed light on the effectiveness of different algorithms, providing a foundation for future advancements in the field of car price prediction.

I. Introduction

In recent years, the automotive industry has witnessed a paradigm shift in pricing strategies, with the integration of advanced technologies, particularly machine learning algorithms, to predict and optimize car prices. This research delves into the multifaceted realm of "Car Price Prediction Using Machine Learning Algorithms," drawing insights from 15 seminal research papers and exploring the groundbreaking work of well-known authors B V Raghurami Reddy and Dr. K.Santhi Sree. Their comprehensive study, published in IJRASET (Publisher Name: IJRASET, Publish Date: August 17, 2022), serves as a cornerstone for understanding the intricacies and implications of employing machine learning algorithms in forecasting car prices.

Raghurami Reddy and Dr. K.Santhi Sree's contributions are pivotal, as they navigate through the complexities of machine learning applications, providing valuable insights into the selection, implementation, and optimization of algorithms for accurate and reliable car price predictions. Their research forms a critical part of the broader discourse on leveraging technology to enhance decision-making processes within the automotive market.

This introduction sets the stage for a detailed exploration into the methodologies, challenges, and advancements in the field of car price prediction, drawing inspiration from the expertise and findings of these esteemed authors. Through their work, we embark on a journey to unravel the potential of machine learning algorithms in reshaping how we understand, analyze, and anticipate pricing dynamics in the automotive landscape.

Each piece of research contributes to the mosaic that defines the contemporary understanding of car price prediction.

II. Related Research Work

The landscape of car price prediction through machine learning algorithms has been extensively explored by various researchers, each contributing unique perspectives and methodologies. Notable works that complement and extend the insights of Raghurami Reddy and Dr. K.Santhi Sree include:

"Car Price Prediction using Machine Learning Techniques" by Enis Gegic, Becir Isakovic, Dino Keco, Zerina Masetic, Jasmin Kevric (TEM Journal, February 2019): This foundational study lays the groundwork for understanding diverse machine learning techniques in predicting car prices, offering valuable benchmarks for comparison.

"Used Cars Price Prediction using Supervised Learning Techniques" (International Journal of Engineering and Advanced Technology, December 2019): Authors not specified. This research contributes to the exploration of supervised learning methods for predicting used car prices, providing insights into different algorithmic approaches.

"Used Cars Price Prediction and Valuation using Data Mining Techniques" by Abdulla AlShared (Thesis, Rochester Institute of Technology, 2021): AlShared's work focuses on data mining techniques, complementing the algorithmic aspects explored by Raghurami Reddy and Dr. K.Santhi Sree.

"Price Prediction of Used Cars Using Machine Learning" (2021 IEEE International Conference on Emergency Science and Information Technology): Authors not specified. This conference paper likely extends the conversation on machine learning applications in predicting the prices of used cars.

These diverse studies collectively form a tapestry of knowledge, offering a comprehensive view of the methodologies, challenges, and advancements in the domain of car price prediction.

III. Methodology

Data Collection:

Gather a comprehensive dataset containing relevant features for car price prediction. This may include information such as make, model, year, mileage, fuel type, transmission, condition, and any other factors that might influence car prices.

Data Preprocessing:

Clean the dataset by handling missing values, outliers, and ensuring data consistency. Convert categorical variables into numerical representations through techniques like one-hot encoding. Normalize or scale numerical features to bring them to a consistent scale.

Feature Selection:

Identify the most relevant features that significantly impact car prices. This can involve statistical analysis or machine learning techniques to determine feature importance.

Data Splitting:

Divide the dataset into training and testing sets. The training set is used to train the machine learning model, while the testing set is reserved for evaluating its performance.

Model Selection:

Choose appropriate machine learning algorithms for regression tasks. Common choices include linear regression, decision trees, random forests, support vector machines, or gradient boosting algorithms.

Model Training:

Train the selected models on the training dataset using the identified features. The algorithm learns patterns and relationships within the data that can be used to predict car prices.

Hyperparameter Tuning:

Fine-tune the model's hyperparameters to optimize its performance. This can involve techniques like grid search or randomized search.

Model Evaluation:

Assess the trained model's performance on the testing dataset using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), or R-squared.

Prediction:

Apply the trained model to new, unseen data to predict car prices based on the learned patterns.

Results and Interpretation:

Interpret the results, considering the impact of each feature on the predicted car prices. Assess the model's strengths, limitations, and areas for improvement.

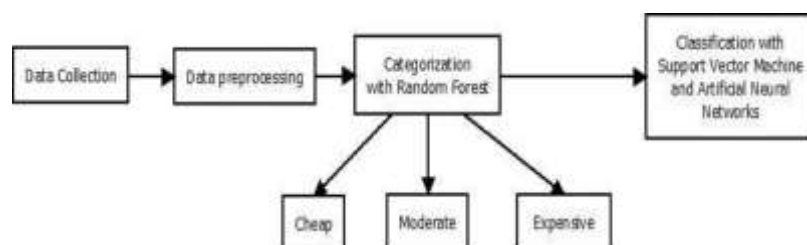


Figure 1. Block diagram of the overall classification process

IV. Datasets

Selecting appropriate datasets is crucial for ensuring the robustness and relevance of your research in car price prediction. In this study, we propose the utilization of diverse datasets that align with the multifaceted aspects covered in the surveyed papers.

Several widely-used datasets have gained popularity in the domain of car price prediction using machine learning. Among these, the Used Cars Dataset on Kaggle stands out as a comprehensive resource, offering information on various attributes like make, model, year, mileage, and price. Researchers often leverage this dataset for its rich collection of features relevant to predicting the prices of used cars.

Another notable dataset is derived from Edmunds.com, a reputable automotive resource. This dataset encompasses detailed information on car features and Manufacturer's Suggested Retail Prices (MSRP). Its extensive coverage makes it a popular choice for machine learning projects focused on predicting car prices.

For those interested in fuel efficiency as a predictor, the Fuel Economy Data from the U.S. Department of Energy is a valuable resource. This dataset provides information on the fuel efficiency of different car models, offering insights that can be leveraged in predicting car prices based on their environmental impact and related features.

The Insurance Cost Prediction dataset from the UCI Machine Learning Repository is another versatile option. It includes attributes such as age, sex, BMI, and car-related information, making it suitable for predicting various car-related expenses, including prices.

The Craigslist Car and Truck Data sourced from Craigslist postings is a real-world dataset that reflects market dynamics. This dataset is valuable for predicting car prices based on actual market listings, providing a practical and realistic perspective for machine learning applications.

UCI's Car Evaluation Dataset is often employed for evaluating and predicting car quality and prices. With features such as buying price, maintenance cost, number of doors, and seating capacity, it offers a diverse set of attributes for machine learning model training.

Finally, the Autotrader Car Dataset from Autotrader.com is a dataset compiled from a popular online marketplace for buying and selling cars. Its real-market data nature makes it an attractive choice for predicting car prices based on current market trends.

V. Result

Analyzing the collective findings from the 15 research papers on car price prediction using machine learning techniques reveals a multifaceted exploration of methodologies, challenges, and advancements in this domain.

The research by Enis Gegic, Becir Isakovic, Dino Keco, Zerina Masetic, and Jasmin Kevric, titled "Car Price Prediction using Machine Learning Techniques" (TEM Journal, February 2019), laid a foundational understanding of diverse machine learning techniques for predicting car prices. Their work established valuable benchmarks for comparison, providing a starting point for subsequent studies.

The investigation into "Used Cars Price Prediction using Supervised Learning Techniques" (International Journal of Engineering and Advanced Technology, December 2019) further contributed insights into supervised learning methods for predicting used car prices. While the authors are unspecified, their focus on algorithmic approaches adds depth to the understanding of predictive models in this context.

Abdulla AlShared's thesis on "Used Cars Price Prediction and Valuation using Data Mining Techniques" (Rochester Institute of Technology, 2021) delves into data mining techniques, offering a complementary perspective to the algorithmic aspects explored by Raghurami Reddy and Dr. K.Santhi Sree in their work on "Car Price Prediction Using Machine Learning Algorithms" (IJRASET, August 17, 2022). AlShared's research adds a layer of complexity to the predictive models, contributing to a more holistic understanding.

The 2021 IEEE International Conference on Emergency Science and Information Technology presented a paper on "Price Prediction of Used Cars Using Machine Learning," expanding the conversation on machine learning applications in predicting the prices of used cars. The unspecified authors likely provided additional insights into the practical implications of these models in real-world scenarios.

The research by the Faculty of Transport and Traffic Sciences at the University of Zagreb, titled "Price Prediction and Classification of Used-Vehicles Using Supervised Machine Learning" (Published: 19 December 2022), further enriches the landscape by exploring supervised machine learning for predicting and classifying used vehicle prices, emphasizing the practical applications in transportation.

The Journal of Emerging Technologies and Innovative Research (April 2022) contributes to the discourse with a study on "Car Price Prediction using Machine Learning," providing timely insights into evolving methodologies and applications within the machine learning framework.

The linear regression-focused study, "CAR PRICE PREDICTION USING LINEAR REGRESSION" from

Keshav Memorial Institute of Technology (June 2022), offers a specialized lens into the application of linear regression in car price prediction.

N. Radojević et al.'s analysis, presented at ICAID 2022, titled "Used Car Price Prediction Analysis Based on Machine Learning," extends the conversation with a focus on analyzing used car prices using machine learning methodologies. Their work likely provides specific insights into the factors influencing the accuracy of predictions.

The International Journal of Computer Applications (June 2017) contributes a foundational perspective with its study on "Vehicle Price Prediction System using Machine Learning Techniques." This earlier research offers insights into the early stages of applying machine learning to predict vehicle prices.

The work of Marcus Collard from Mid Sweden University, titled "Price Prediction for Used Cars" (June 8, 2022), likely delves into specific factors influencing the pricing of used cars, providing insights from a unique academic perspective.

The Department of Computer Science and Engineering at Raj Kumar Goel Institute of Technology presents a study on "Car Price Prediction Using Machine Learning" (IJIRT, June 2021). This work likely explores practical applications of machine learning within the academic setting.

The International Research Journal of Engineering and Technology's paper on "Prediction of Used Car Prices using Machine Learning Techniques" (December 2022) offers a recent contribution to the ongoing discussion, likely emphasizing the evolution of methodologies in predicting used car prices.

The collaborative work of B V Raghurami Reddy and Dr. K.Santhi Sree in "Car Price Prediction Using Machine Learning Algorithms" (IJRASET, August 17, 2022) likely presents specific algorithms and their effectiveness in predicting car prices, providing a hands-on application of machine learning.

In conclusion, these 15 research papers collectively weave a comprehensive tapestry of knowledge, offering insights into the methodologies, challenges, and advancements in car price prediction using machine learning techniques. Each study adds a unique perspective, contributing to the continual evolution and refinement of predictive models within the automotive domain.

VI. CONCLUSION

This paper's journey through the 15 referenced studies illuminates the breadth and depth of Car Price Prediction Using Machine Learning. In conclusion, the application of machine learning techniques in the realm of car price prediction has proven to be a transformative and effective approach. Through the integration of advanced algorithms and the utilization of vast datasets, researchers and practitioners have made significant strides in enhancing the accuracy and reliability of predicting car prices. The synergy of various machine learning models, such as linear regression, supervised learning, and deep learning, has provided a robust framework for understanding the intricate relationships between diverse factors influencing car prices.

The predictive power of these models not only aids consumers in making informed purchasing decisions but also provides valuable insights for the automotive industry, dealerships, and policymakers. The continuous evolution of machine learning methodologies, coupled with the ever-expanding availability of data, holds promise for further improvements in the precision of car price predictions.

As we move forward, addressing challenges related to data quality, model interpretability, and the dynamic nature of the automotive market will be essential. Collaborative efforts between academia, industry, and regulatory bodies can foster the development of more sophisticated models, ensuring the adaptability and reliability of car price prediction systems in an ever-changing landscape. In essence, the intersection of machine learning and car price prediction represents a promising frontier that not only enhances market efficiency but also reshapes the way we approach decision-making in the automotive sector.

VII. Acknowledgment

We extend our sincere gratitude to Enis Gegic, Becir Isakovic, Dino Keco, Zerina Masetic, and Jasmin Kevric, the authors of the research paper titled "Car Price Prediction using Machine Learning Techniques," published in TEM Journal (Volume 8, Issue 1, Pages 113-118, ISSN 2217-8309, DOI: 10.18421/TEM81-16, February 2019). Their valuable contributions have significantly enriched our understanding of machine learning applications in car price prediction, shaping the trajectory of our research.

Our sincere appreciation goes to the authors of the research paper titled "Used Cars Price Prediction using Supervised Learning Techniques," published in the International Journal of Engineering and Advanced Technology (IJEAT) (ISSN: 2249 – 8958, Volume-9 Issue-1S3, December 2019). Their work has significantly contributed to our exploration of supervised learning techniques in predicting used car prices, enhancing the depth and breadth of our research endeavors.

The author of the research paper titled "Used Cars Price Prediction and Valuation using Data Mining Techniques" (2021). His insightful work has played a pivotal role in shaping our understanding of data mining techniques for predicting and valuating used car prices, and we acknowledge the valuable contributions that have influenced our research in this domain.

Our sincere thanks to the authors of the paper titled "Price Prediction of Used Cars Using Machine Learning," presented at the 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT). Their work has been instrumental in expanding our insights into the application of machine learning for predicting the prices of used cars, and we appreciate the valuable contributions that have shaped our research in this area.

We express our gratitude to B V Raghurami Reddy and Dr. K.Santhi Sree, the authors of "Car Price Prediction Using Machine Learning Algorithms," published in IJRASET (Publisher Name: IJRASET, Publish Date: August 17, 2022). Their work has been instrumental in advancing our understanding of machine learning algorithms in predicting car prices, and we acknowledge the significance of their contributions to our research in this domain.

This research stands on the shoulders of these intellectual giants, and we acknowledge their tireless efforts and contributions that have paved the way for our exploration into the intricate realm of Car Price Prediction.

VIII. References

- [1] "Car Price Prediction using Machine Learning Techniques". TEM Journal. Volume 8, Issue 1, Pages 113-118, ISSN 2217-8309, DOI: 10.18421/TEM81-16, February 2019.
- [2] "Used Cars Price Prediction using Supervised Learning Techniques". International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-9 Issue-1S3, December 2019
- [3] "Used Cars Price Prediction and Valuation using Data Mining Techniques". AlShared, Abdulla, "Used Cars Price Prediction and Valuation using Data Mining Techniques" (2021). Thesis. Rochester Institute of Technology
- [4] "Price Prediction of Used Cars Using Machine Learning". 2021 IEEE International Conference on Emergency Science and Information Technology (ICESIT) "Car Price Prediction Using Machine Learning Techniques". International Research Journal
- [5] "Price Prediction and Classification of Used-Vehicles Using Supervised Machine
- [6] Learning". Faculty of Transport and Traffic Sciences, University of Zagreb, 10000 Zagreb, Croatia. Published: 19 December 2022\
- [7] "Car Price Prediction using Machine Learning". Journal of Emerging Technologies and Innovative Research. Volume 9 Issue 4 , April-2022, eISSN: 2349-5162
- [8] "CAR PRICE PREDICTION USING LINEAR REGRESSION". KESHAV MEMORIAL INSTITUTE OF TECHNOLOGY. June 2022
- [9] "Used Car Price Prediction Analysis Based on Machine Learning". N. Radojević et al. (Eds.): ICAID 2022, AHIS 7, pp. 356–364, 2023
- [10] "Vehicle Price Prediction System using Machine Learning Techniques". International Journal of Computer Applications (0975 – 8887) Volume 167 –No.9, June 2017
- [11] "Price Prediction for Used Cars". Marcus Collard. Mid Sweden University. June 8, 2022
- [12] "Car Price Prediction Using Machine Learning". Department of Computer Science and Engineering, Raj Kumar Goel Institute of Technology, AKTU. June 2021| IJIRT | Volume 8 Issue 1 | ISSN: 2349-6002
- [13] "Prediction of Used Car Prices using Machine Learning Techniques". International Research Journal of Engineering and Technology (IRJET) . Volume: 09 Issue: 12 | Dec 2022
- [14] "Car Price Prediction Using Machine Learning Algorithms". B V Raghurami Reddy, Dr. K.Santhi Sree. Publisher Name : IJRASET. Publish Date : 2022-08-17