Dean&Francis

ISSN 2959-6157

Research on the Influencing Factors of Used Car Price in the Market

Tianrun Ma [1, *]

[1] College of Letters and Science, University of California, Davis, 95616, America

*Corresponding author: trma@ucdavis.edu

Abstract:

As people are traveling longer and longer distance by now a transportation become a part of people's life. The demand of car has been increased and many people do need cars to help them go to school, work, and purchase daily supplements. However, not everyone can afford a new car and a used car can be an option. Like people who just graduate from school or people need their own car but they cannot pay as much as the price of a new car. A used car can be cheaper and higher Cost-effectiveness to most of the customers. With many different used cars in the market, what is the best price for a used car and how can people predict the price of a used car can be a problem. This paper will mainly discus about the prediction on the used cars and the method that is used for predict used car price.

Keywords: Used car; price; prediction; influencing factors.

1. Introduction

The used car market has become an integral part of the automotive industry, offering a more affordable alternative for those who cannot afford new vehicles. With transportation becoming a crucial part of daily life, the demand for cars continues to grow. However, purchasing a new vehicle can be financially out of reach for many, particularly for recent graduates, young professionals, or individuals on tighter budgets. For these consumers, used cars provide a cost-effective solution, balancing affordability with functionality.

A key challenge in the used car market is determining a fair price. Many factors contribute to the price of a used car, with its condition being one of the most important. Customers generally believe that the price of a car is linked to its history, including how it was driven, maintained, and whether it was involved in accidents. Alongside these factors, external elements such as market conditions, supply and demand, and broader economic forces play a significant role in shaping used car prices. The ability to predict these prices accurately is essential for both buyers and sellers in navigating the market.

Research has highlighted a range of factors that influence used car prices. For example, Sallee et al. found that engine specifications, such as the number of cylinders, directly impact fuel consumption, with more cylinders often leading to higher fuel costs, thereby lowering a car's resale value [1]. Additionally, Gutknecht pointed out that dealer pricing strategies can lead to significant variations in used car prices, particularly due to discrepancies in dealer knowledge

and experience [2]. Another important factor is vehicle performance; Nafi'iyah and Mauladi noted that horsepower and weight influence not only a car's fuel efficiency but also its overall market value [3]. Furthermore, Boranova et al. discussed how external economic events, such as the COVID-19 pandemic, disrupted production and supply chains, which had a ripple effect on the price of both new and used cars [4]. The age of a car is another significant determinant, as older cars typically experience higher rates of depreciation.

In response to these pricing complexities, predictive models are increasingly being used to analyze and forecast the prices of used cars. Linear regression has been a popular tool for examining the relationship between variables like car condition and price. Kavita and Mathur argued that linear regression is particularly useful in identifying correlations between variables, making it an effective method for predicting used car prices [5]. However, as data becomes more complex, more advanced techniques are being adopted.

Machine learning, as described by Naqa and Murphy, is gaining popularity in predicting used car prices due to its ability to handle large amounts of data and identify patterns that traditional methods may miss [6]. Machine learning models can analyze numerous variables simultaneously, providing deeper insights into how factors such as market conditions, fuel economy, and vehicle specifications interact to influence price. Haque et al. highlighted the efficiency of machine learning in processing large datasets, enabling quicker and more accurate predictions [7]. Similarly, Hapsari et al. demonstrated how machine learning models can adjust to real-time changes in the market, offering dynamic and adaptable pricing models [8].

An even more advanced approach to price prediction is the use of deep learning algorithms. Imram et al. talked about how machine learning can help collect data and filter the useless information [9]. Nikou et al. explored the potential of deep learning in identifying non-linear relationships between factors, making it a valuable tool for predicting complex pricing outcomes [10]. Deep learning models can consider a broader range of variables-such as the car's age, brand, and condition-while accounting for interactions between these factors. This leads to more accurate price forecasts, helping both consumers and dealers make better-informed decisions.

This paper will explore the factors that affect used car prices and evaluate the effectiveness of different methods for predicting these prices. This paper will focus on traditional models like linear regression, as well as more advanced techniques such as machine learning and deep learning. By applying these methods to historical data, this paper aims to forecast future trends in the used car market, providing valuable insights for both consumers and industry professionals.

2. Methods

2.1 Data Source

The dataset used in this study was obtained from Kaggle's "Car Price" dataset, which provides detailed information on 205 different car models. The dataset includes various attributes related to each car, such as technical specifications and pricing information. For this research, the entire dataset, comprising all 205 car models, was utilized to ensure comprehensive analysis. The dataset was provided in CSV format, which was loaded into analytical tools for further processing and analysis.

2.2 Variable Selection

The price of a used car is influenced by a wide range of factors, many of which reflect consumers' preferences and priorities when making a purchase decision. The dataset includes a diverse set of variables, such as fuel type, wheelbase, car length, car width, car weight, and other physical and technical attributes. Additionally, there are categorical variables with limited values, such as aspiration (standard or turbocharged), number of doors, and body type.

Given the large number of variables, not all of them are likely to play an equally significant role in determining car prices. Consumers may focus on certain key attributes when purchasing a vehicle, such as performance, safety, or brand reputation, while other factors may have less influence. In this study, each variable will be analyzed in relation to price, aiming to identify the elements that have the most significant impact on used car prices. Understanding these relationships can help uncover the major factors driving price differences between various car models.

2.3 Method Introduction

To analyze the relationship between car attributes and their prices, this study uses linear regression as the primary analytical method. Linear regression is a widely used statistical technique for modeling the relationship between a dependent variable (in this case, the car price) and one or more independent variables (the car's features). This method allows us to assess whether each variable in the dataset has a statistically significant impact on price and, if so, how much influence it exerts.

Furthermore, the model's goodness-of-fit metrics, such

ISSN 2959-6157

as the R-squared value, will be used to evaluate how well the selected variables explain the variance in car prices. Variables with little or no effect on price may be excluded in further stages of the analysis to focus on the most influential factors. The results of the regression analysis will provide insights into which car features are most closely associated with price trends in the used car market, helping to identify consumer preferences and priorities.

3. Results and Discussion

3.1 Linear Regression Analyze

The linear regression analysis conducted on the dataset revealed notable patterns in how different car elements affect the price of used vehicles. Some elements of used cars have only little correlation with price, while others demonstrated a stronger influence. Those can happen because of the cost of the vehicle or what factors consumers care about. This section discusses both cases and the potential reasons behind the varying levels of significance observed.

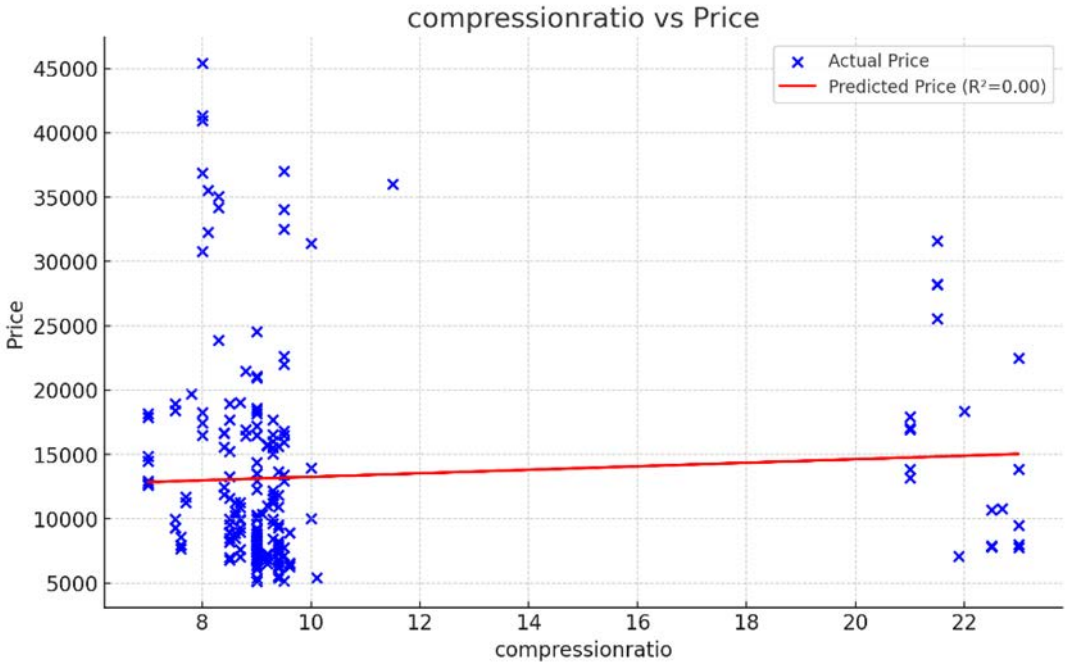

Fig. 1 Price vs Compression Ratio

When each variable in the dataset was analyzed for its impact on price, some elements do not have a large impact on the price of a used car. As Figure 1 shows, the compression ratio only has a small effect on the price of used vehicles. A higher compression ratio did not result in significant price changes for used cars. The R-squared value was quite low in this case, and showing that this compression ratio does not play an important role in used car prices.

This may be because when consumers purchase used cars, consumers are not focusing on many technical details, especially since most consumers are not expecting a high performance of used cars and consumers most likely looking for cars that are cheap for both the price of the vehicle and daily cost. In the context of used cars, buyers are often more interested in factors such as the car's overall condition, mileage, and brand reputation. Instead of looking into advanced technical advantage of a car like compression ratio, meeting the needs of normal consumers is the main thing that affect the price.

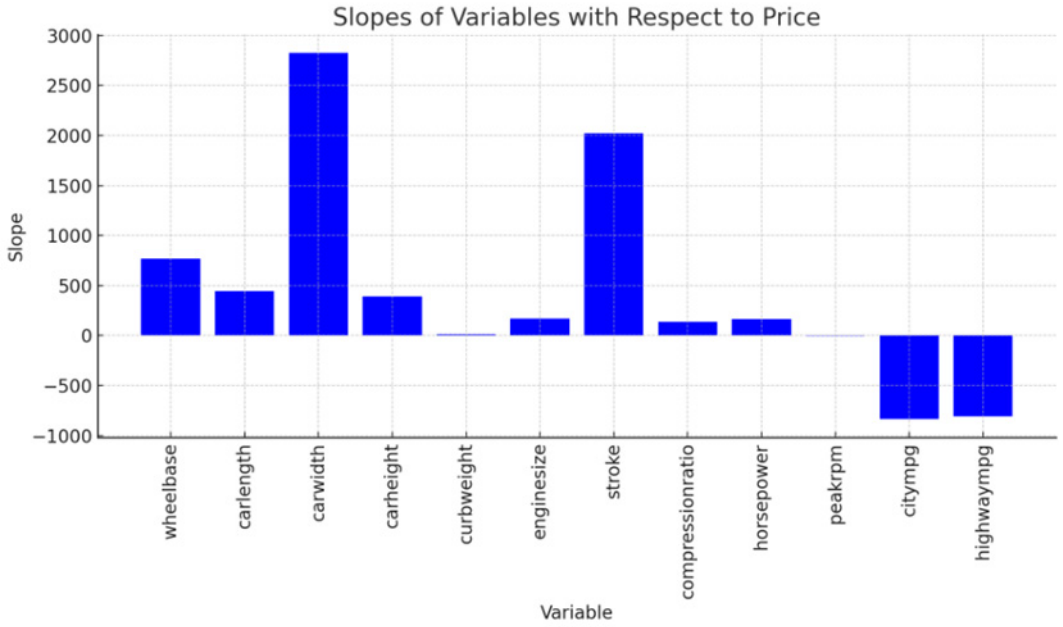

Fig. 2 Slopes of Variables with Respect to Price

The trend observed for compression ratio was also seen with other variables such as peak RPM, curb weight, engine size, and horsepower. As Figure 2 shows, those elements do not have a great impact on the price of used cars, and this means either the consumers do not care or the cost of those elements is not significant. These attributes are typically associated with the performance and power of a vehicle, but in the case of used cars, they do not seem to have a substantial impact on pricing. The linear regression results for these variables also showed low R-squared values, indicating that they do not explain much of the variation in price.

3.2 Price Analyze according to Elements

It is possible that the used car consumers are not looking for the same thing as a new car consumer. A used car consumers are looking for an affordable and cheap car. what will lead a lower demand with higher price vehicles. For new cars, consumers may expect the performance of the car such as horsepower and engine size, which directly affect a car's driving experience. However, for used cars, buyers more likely to focus on more cost of the use like fuel efficiency, reliability, or maintenance history rather than performance.

Additionally, factors like curb weight and engine size may be less relevant because the buyers in this market are not looking for high-performance or heavy-duty vehicles. However, cars that offer value at a reasonable price. As a result, these variables have little impact on the price of used cars.

On the other hand, some variables have a stronger correlation with price. The stroke of the car's engine is one of the elements that have a big impact on price. Stroke had a significant influence on the price of a used car. This is not usual because it suggests that stroke is a technical specification that refers to the distance the piston travels inside the engine cylinder, and has a measurable impact on car pricing. However, this is likely not due to consumer awareness of this particular specification. Instead, the cost of a vehicle with a higher stroke is higher, and this might lead to a higher price on used cars.

Most consumers purchasing used cars are unlikely to notice advanced technical information about a car. Instead, the reason why a stroke can influence the price of used cars can be the cost of production. A higher or more complex stroke configuration can lead to increased production costs, and these higher costs are reflected in the car's resale price even in the used car market.

Figure 2 also shows that the city mile per gallon (mpg) and highway mpg have a negative effect on the price of the used cars. This is because, with a higher MPG, the cost of daily use will be increased. That will lead to the demand for used cars decreasing because consumers are looking for used vehicles that are affordable and cheap.

4. Conclusion

This study has shown various factors influencing the price of used cars and explored the effectiveness of different predictive models in analyzing these factors. The results

ISSN 2959-6157

show that while certain technical specifications, such as compression ratio, peak RPM, curb weight, engine size, and horsepower, have limited impact on used car prices, other elements like engine stroke significantly influence pricing. This difference suggests that used car buyers consider about affordability first and practicality over high-performance features, focusing on cost-effective vehicles that meet their daily transportation needs.

The linear regression analysis provided valuable ideas into which elements are most relevant to used car pricing. However, the study also highlighted the potential benefits of using more advanced predictive techniques, such as machine learning and deep learning, to capture complex, non-linear relationships between variables. These methods help for future research, as they can adapt to market changes and provide dynamic pricing forecasts.

In conclusion, understanding the factors that drive used car prices is important for both consumers and dealers navigating the market. By focusing on the attributes that matter most to buyers and utilizing predictive models, stakeholders can make more informed decisions, resulting in fairer pricing and a more efficient market overall. As the used car market continues to develop, research and advanced analytical methods will be essential for staying attuned to changing consumer preferences and economic conditions.

References

[1] Sallee J M, West S E, Fan W. Do consumers recognize the value of fuel economy: Evidence from used car prices and gasoline price fluctuations. National Bureau of Economic Research, 2015.

[2] Gutknecht S. Determinants of residual values: Empirical analyses of price effects in the used-car market. Cuvillier Verlag, 2020.

[3] Nafi'iyah N, Mauladi K F. Linear regression analysis and SVR in predicting motor vehicle theft. 2021 International Seminar on Application for Technology of Information and Communication (iSemantic), 2021, 98: 54-58.

[4] Boranova V, Huidrom R, Ozturk E O, Stepanyan A, Topalova P, Zhang S. IMF Working Papers Volume 2022 issue 006: Cars in Europe: Supply chains and spillovers during COVID-19 times (2022). IMF eLibrary, 2022.

[5] Kavita M, Mathur P. Crop yield estimation in India using machine learning. 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), 2020.

[6] El Naqa I, Murphy M J. What Is Machine Learning. Machine Learning in Radiation Oncology. Springer, 2015.

[7] Haque F F, Abdelgawad A, Yanambaka V P, Yelamarthi K. Crop yield analysis using machine learning algorithms. 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), 2020, 1-2.

[8] Hapsari I, Surjandari I K. Visiting time prediction using machine learning regression algorithm. 2018 6th International Conference on Information and Communication Technology (ICoICT), 2018, 1: 495-500.

[9] Imran I. Using machine learning algorithms for housing price prediction: The case of islamabad housing data. Soft Computing and Machine Intelligence, 2021, 1: 11-23.

[10] Nikou M, Mansourfar G, Bagherzadeh J. Stock price prediction using deep learning algorithm and its comparison with machine learning algorithms. Intelligent Systems in Accounting, Finance and Management, 2019, 26(4): 164-174.