

LITERATURE SURVEY

1. Chuyang Jin, “Price Prediction of Used Cars Using Machine Learning”, 2021

This paper aims to build a model to predict used cars' reasonable prices based on multiple aspects, including vehicle mileage, year of manufacturing, fuel consumption, transmission, road tax, fuel type, and engine size. This model can benefit sellers, buyers, and car manufacturers in the used cars market. Upon completion, it can output a relatively accurate price prediction based on the information that user's input. The model building process involves machine learning and data science. The dataset used was scraped from listings of used cars. Various regression methods, including linear regression, polynomial regression, support vector regression, decision tree regression, and random forest regression, were applied in the research to achieve the highest accuracy. Before the actual start of model-building, this project visualized the data to understand the dataset better. The dataset was divided and modified to fit the regression, thus ensure the performance of the regression. To evaluate the performance of each regression, R-square was calculated. Among all regressions in this project, random forest achieved the highest R-square of 0.90416. Compared to previous research, the resulting model includes more aspects of used cars while also having a higher prediction accuracy.

2. Feng Wang; Xusong Zhang; Qiang Wang, “Prediction of Used Car Price Based on Supervised Learning Algorithm”, 2021

In this paper, we use machine learning algorithms to predict the price of used cars with less human intervention to make the results more objective. The method used is to pre-process the dataset through Python's Pycaret package and compare the performance of each algorithm through the algorithm comparison function, in this study Extra Trees Regressor, Random Forest Regressor performs relatively well. Finally, the algorithm was optimized by using the hyperparameter function. The results show that $R^2 = 0.9807$ obtained from extreme random numbers is the best performance. The algorithm was obtained and validated with new data to derive the final algorithm model. When new used car data flows into the used car system, used car prices will be automatically generated by this algorithm, which will make the workflow of the used car market faster and more competitive for that used car market.

3. Janke Varshitha; K Jahnavi; C. Lakshmi, “Prediction Of Used Car Prices Using Artificial Neural Networks And Machine Learning” , 2022

With the extensive growth in usage of cars, the newly produced cars are unable to reach the customers for various reasons like high prices, less availability, financial incapability, and so on. Hence the used car market is escalated across the globe but in India, the used car market is in a very nascent stage and mostly dominated by the unorganized sector. This gives chance for fraud while buying a used car. Hence a high precision model is required which will estimate the price of a used car with none bias towards customer or merchandiser. In this model, A Supervised learning-based Artificial Neural Network model and Random Forest Machine Learning model are developed which can learn from the car dataset provided to it. This project presents a working model for used car price prediction with a low error value. A considerable number of distinct attributes are examined for reliable and accurate predictions. The results obtained agree with theoretical predictions and have shown improvement over models which use simple linear models. An ANN (Artificial Neural Network) is built by using Keras Regression algorithm namely Keras Regressor and other Machine Learning Algorithms Namely Random Forest, Lasso, Ridge, Linear regressions are built. These algorithms are tested with the car dataset. Experimental results have shown that the Random Forest model with a Mean Absolute Error value of 1.0970472 and R2 error value of 0.772584 has given the less error among all the other algorithms. The work presented here has shown profound implications for future studies of Used Cars price Prediction using Random Forest and might one day help to solve the problem of frauds with one hundred percent accuracy.

4. Mustapha Hankar; Marouane Birjali; Abderrahim Beni-Hssane , “Used Car Price Prediction using Machine Learning: A Case Study”, 2022

In many business fields that are related to statistics and machine learning (ML), multiple linear regression (MLR) models are often used to estimate and fit a linear relationship between a continuous response variable and other explanatory variables. In our case study, we applied several regression techniques based on supervised machine learning to predict the resale price of used cars given many factors such as mileage, fuel type, fiscal power, mark, model, and the production year of the car. In all tested models, gradient boosting regressor showed a high R-squared score and low root mean square error.

**5. Shengqiang Han; Jianhua Qu; Jinyi Song; Zijing Liu Second-hand ,
“Car Price Prediction Based on a Mixed-Weighted Regression
Model”, 2022**

With the development of motor vehicles, the circulation demand of motor vehicles in the form of "second-hand cars" in circulation links is increasing. As a special "e-commerce commodity", second-hand cars are more complicated than ordinary e-commerce commodities. As a result, it is difficult to estimate the price of second-hand cars, which is not only influenced by the basic configuration of the car, but also by the car conditions. At present, the state has not issued a standard to judge the value of second-hand car. To solve this problem, in this paper, first making feature engineering, which includes data pre-processing and feature screening. Data pre-processing includes data cleaning and data transformation, data cleaning includes removing outliers and filling missing values, and data transformation is used to unify data format to improve data quality. The feature screening includes correlation analysis and feature extraction based on LightMBG, and the screened features provide the basis for model building, training and prediction. Then, five regression models are constructed by using the feature attributes obtained by the feature engineering for training, and evaluated. Then, Random Forest and XGBoost are weighted and mixed to get a novel regression model, and the effect of the model is better than that of the five regression models. Finally, the novel regression model is used to predict the price of second-hand cars.

6. Chejarla Venkat Narayana; Chinta Lakshmi Likhitha; Syed Bademiya; Karre Kusumanjali, “Machine Learning Techniques To Predict The Price Of Used Cars: Predictive Analytics in Retail Business”, 2021

It is generally known that, taking wise and challenging decisions is really a crucial task in every business. Taking improper decisions can cause huge loss and even lead to shutdown of business. To propose a novel solution for this challenge, this research work majorly focuses on one of the retail businesses i.e., used car sales business. The proposed research work shows that, the predictive analytical models will be a great add-on to business mainly for assisting the decision-making process. Predictive Analytics is a process, where the businesses use statistical methods and technologies to analyze their historical data for delivering new insights and plan the future accordingly. The major objective of our paper is to build a prediction model i.e., a fair price mechanism to predict the cars selling price based on their features like the car model, the number of years that a car is old, the type of fuel it uses, the type of seller, the type of transmission and the number of kilometres that the car has driven so far. This paper will help to get an approximation about selling price of a used car based on its features and reduces the seller and consumer risk in business. The proposed model utilizes the machine learning algorithms and regression techniques of statistics like linear, decision tree and random forest regressions to achieve this task.

7. Chejarla Venkata Narayana; Nukathoti Ooha Gnana Madhuri; Atmakuri NagaSindhu; Mulupuri Aksha; Chalavadi Naveen, “Second Sale Car Price Prediction using Machine Learning Algorithm”, 2022

Every business firm recognizes the need of making sound and challenging decisions. Poor decisions can lead to substantial losses and even the demise of a firm. This paper is focused on one of the retail enterprises, which deals with the used car sales. The major goal is to develop a prediction model that can estimate the selling price of used cars based on key factors. Machine learning techniques such as Random Forest Regression, Feature engineering technique such as Extra Trees Regression are employed to accomplish the goal as Random Forest Regression is modelled for prediction analysis and Extra Trees Regression fits the number of decision trees. The results are so encouraging with our approach.