

LITERATURE SURVEY FOR CAR RESALE VALUE PREDICTION

[1] Doan Van Thai et al. This paper covers procedures for determining importance, information derivation, and subjective information rules. The primary reason for the flow research is to investigate various types of vehicle information, with the goal of developing a mechanized procedure to forecast used vehicle costs.

[2] Nitish Monburinon et al. report on execution of relapse dependent on directed ML models. Each model is prepared using trade-in vehicle showcase data gathered from a German business website. As a result, proclivity is aided. Relapse trees provide the best execution, with a mean total error (MSE) = 3D 0.28. . Following that, there was a random woodland relapse with MSE = 3D 0.35 and a random direct relapse with MSE = 3D 0.55.

[3] Ning Sun et al. The value assessment model based on massive information investigation is proposed in this paper, which exploits generally circled vehicle information and numerous vehicle exchange data to dissect the value information for each type of vehicle by utilizing the improved BP neural system calculation. It plans to set up a reused vehicle value assessment model to determine the best value for the vehicle.

[4] Listiani demonstrated that a regression model built with support vector machines (SVM) can estimate the residual price of leased cars with greater accuracy than simple multiple regression or multivariate regression. SVM Predicts Used Car Prices Using Machine Learning Techniques 755 better at dealing with very high dimensional data (number of features used to predict the price) and avoiding both overfitting and under-fitting She used a genetic algorithm in particular to find the optimal SVM parameters in less time. The only disadvantage of this study is that the superiority of SVM regression over simple regression was not expressed in simple measures such as mean deviation or variance.

[5] Gonggie proposed a model for used car price prediction based on ANN (Artificial Neural Networks). He considered several options. attributes include the number of miles driven, the

estimated life of the vehicle, and the brand. The proposed model was designed to deal with nonlinear data relationships, which previous models that used the simple model could not. linear regression methods The non-linear model predicted car prices with greater accuracy than other linear models.

[6] Pudaruth used machine learning algorithms such as k-nearest neighbors, multiple linear regression analysis, decision trees, and naive bayes to predict car prices in Mauritius. The dataset used to build a prediction model was gathered manually from local newspapers over a period of less than one month, as time can have a significant impact on the price of a car. He researched the following characteristics: brand, model, and cubic capacity, kilometerage, year of manufacture, exterior color, transmission type, and price The author discovered, however, that Naive Bayes and Decision Tree were unable to predict and classify numeric values. Furthermore, due to the limited number of dataset instances, high classification performances, i.e. accuracies less than 70%, were not possible.

[7] Noor and Jan used multiple linear regression to create a model for predicting car prices. The dataset was created over a two-month period, price, cubic capacity, exterior color, date when the ad was posted, number of ad views, power steering, mileage in kilometers, rims type, transmission type, engine type, city, registered city, model, version, make, and model year Following feature selection, the authors only considered engine type, price, and model. As input features, year and model are used. The authors were able to achieve a prediction accuracy of 98% using the given setup.

[8] Wu et al. conducted a used car price prediction study, by using a neuro-fuzzy knowledge-based system. They considered the following factors: attributes: brand, year of manufacture, and engine type Their prediction model produced results that were similar to the simple regression model. Furthermore, they developed an expert system known as ODAV (Optimal Distribution of Auction Vehicles) because car dealers have a high demand for selling cars at the end of the leasing year. This system provides information on the best vehicle prices as well as the location. where the best possible price can be obtained To predict the price of a car, a regression model based on the k-nearest neighbor machine learning algorithm was used. This

system has a history of being extremely successful. since more than two million vehicles have been manufactured.

[9] Richardson tested the hypothesis that automakers are more willing to produce vehicles that do not depreciate rapidly. He demonstrated, in particular, that hybrid cars (cars that use two different power sources to propel the car, i.e. they have both an internal combustion engine and an electric motor) retain their value better than traditional vehicles. This is most likely due to increased environmental concerns about climate change and higher fuel efficiency. The significance of other factors such as. This study also took into account age, mileage, make, and MPG (miles per gallon). He gathered all of his information from various websites.

[10] Author Nabarun Pal et al. Random forest, a supervised learning method, was used to forecast used car prices. The model can accurately predict car prices by selecting the most correlated feature. The researchers experimented with both linear and random forest regression methods. According to their findings, random forest outperforms linear regression. The optimal number of trees was determined using a grid search algorithm. They concluded that forest accuracy is good with 500 decision trees. In the case of a regression problem, they extracted as many features as possible from the input data set. When it came to classification, they used the Square root of some features. As random forest is used for cost prediction, they converted the problem into a regression problem.

[11] B.Lavanya et al. by Utilizing Machine Learning Algorithms like Linear Regression, Multiple Regression attempted to foster a factual model which will actually want to anticipate the cost of a pre-owned vehicle, in light of past shopper information and a given arrangement of highlights. Likewise they contrasted the forecast precision of these models to decide the ideal one. The fundamental target is to utilize three distinct expectation models to anticipate the retail cost of an utilized vehicle and to think about their degrees of precision.

[12] Fahad Rahman Amik et al. developed a forecasting system (using machine learning techniques) that helps a potential buyer to estimate the price of a pre-owned car he is interested in. A dataset is collected and pre-processed. Exploratory data analysis has been performed.

Following that, various machine learning regression algorithms, including linear regression, LASSO (Least Absolute Shrinkage and Selection Operator) regression, decision tree, random forest, and extreme gradient boosting have been applied. After evaluating the performance of each method, the best-performing model (XGBoost) was chosen. This model is capable of properly predicting prices more than 91% of the time. Finally, the model has been deployed as a web application in a local machine so that this can be later made available to end users.

[13] Yadav et al. developed machine learning models which make it possible to accurately predict the price of a second - hand car according to its parameters or characteristics. The implementation techniques and evaluation methods are used on a Car dataset consisting of the selling prices of various models of car across different cities of India. The outcome of this experiment shows that clustering with linear regression and Random Forest model yield the best accuracy outcome. The machine learning model produces a satisfactory result within a short duration of time compared to the aforementioned self.

[14] Pattabiraman Venkatasubbu and Mukkesh Ganesh tried to develop a statistical model which will be able to predict the price of a used car which is based on previous consumer data and a given set of features by using Machine Learning Algorithms such as Lasso Regression, Multiple Regression and Regression trees. Using these techniques, they have created, trained and tested the effectiveness of our statistical models. They have also compared the prediction accuracy of these models to determine the optimal one.

[15] Enis Gegic et al. applied three machine learning techniques (Artificial Neural Network, Support Vector Machine and Random Forest) to build a model for predicting the price of used cars in Bosnia and Herzegovina. However, the mentioned techniques were applied to work as an ensemble. The data used for the prediction was collected from the web portal autopijaca.ba using web scraper that was written in PHP programming language. Respective performances of different algorithms were then compared to find one that best suits the available data set. The final prediction model was integrated into the Java application. Furthermore, the model was evaluated using test data and the accuracy of 87.38% was obtained.