

CAR RESALE VALUE PREDICTION LITERATURE SURVEY

Paper 1: Used Cars Value Prediction and Valuation using Machine Learning Techniques

- **Publication year: DEC 2021**
- **Author: Abdulla AlShared**
- **Summary:** Several studies and related works have been done previously to predict used car prices around the world using different methodologies and approaches, with varying results of accuracy from 30% to 89%. In (Pudaruth, 2014) the researcher proposed to predict used car prices in Mauritius, where he applied different machine learning techniques to achieve his results like decision tree, K-nearest neighbors, Multiple Regression and Naïve Bayes algorithms to predict the used cars prices, based on historical data gathered from the newspaper. Achieved results ranged from accuracy of 60-70 percent,

Paper 2: Used Cars Price Prediction using Supervised Learning Techniques

- **Publication Year: DEC 2019**
- **Author: Mukesh Ganesh**
- **Summary:** Overfitting and underfitting come into picture when we create our statistical models. The models might be too biased to the training data and might not perform well on the test data set. This is called overfitting. Likewise, the models might not take into consideration all the variance present in the population and perform poorly on a test data set. This is called underfitting. A perfect balance needs to be achieved between these two, which leads to the concept of Bias-Variance tradeoff. Pierre Geurts has introduced and explained how bias-variance tradeoff is achieved in both regression and classification. The selection of variables/attributes plays a vital role in influencing both the bias and variance of the statistical model. Robert Tibshirani proposed a new method called Lasso, which minimizes the

residual sum of squares. This returns a subset of attributes which need to be included in multiple regression to get the minimal error rate. Similarly, decision trees suffer from overfitting if they are not pruned/shrunk.

Paper 3: Car Price Prediction Using Machine Learning

- **Publication Year: JUNE 2021**
- **Author: Ketan Agrahari, Ayush Chaubey, Mamoor Khan, Manas Srivastava**
- **Summary:** With the recent arrival of internet portals, buyers and sellers may obtain an appropriate status of the factors that ascertain the market price of a used automobile. Lasso Regression, Multiple Regression, and Regression Trees are examples of machine learning algorithms. We will try to develop a statistical model that can forecast the value of a pre-owned automobile based on prior customer details and different parameters of the vehicle. This paper aims to compare the efficiency of different model predictions to find the appropriate one. On the subject of used automobile price prediction, several previous studies have been conducted. To anticipate the value of pre-owned automobiles in Mauritius, Pudaruth employed naive Bayes, k-nearest neighbors, multiple linear regression, and decision trees. However, because there were fewer cars observed, their results were not good for prediction. In his article, Pudaruth concluded that decision trees and naive Bayes are ineffective for continuous-valued variables.

Paper 4: Vehicle Price Prediction using SVM Techniques

- **Publication Year: June 2020**
- **Author :S.E.Viswapriya, Durbaka Sandeep Sharma,Gandavarapu Sathya Kiran**
- **Summary:** According to authors et al, in this paper they mainly concentrate on collecting various data from web portals by using web scrap techniques. And those have been compared with the help of

different machine learning algorithms to predict the vehicle price in an easy manner. They classified the price according to different ranges of price that is already given. Artificial neural network, support vector machine, random forest algorithms were used on different datasets to build classifiers model. Another approach was given by Richardson in his thesis work. In his theory it states more durable vehicles will be produced by vehicle producers. He compared the hybrid vehicles and traditional vehicles in how it actually retains their value for longer time using multiple regression techniques. This improves the environmental conditions, and also it helps to provide huge efficiency of using fuels. Wu et al, in this paper they have used neuro fuzzy knowledge-based systems to demonstrate vehicle price prediction. By considering the following attributes such as brand, year of production and type of engine they predicted a model which has 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 by vehicles at the end of the leasing year by vehicle dealers. This system gives insights into the best prices for vehicles, as well as the location where the best price method.

Paper 5: USED CAR PRICE PREDICTION AND LIFE SPAN

- **Publication year: Dec 2021**
- **Author: Aditya Nikhade, Rohan Borde.**
- **Summary:** The main objective of this project is to predict the Prices of Used Cars, compare the prices and also estimate the life span of a particular car, keeping in mind various statistics of that car. It is said that a new car loses its value by 10% the moment the car is taken out from the showroom. We can easily say that the main predictor of prices in this scenario is the number is kilometers the car has been driven. secondly, we need also need to keep in mind the brands of a car, each car company have their own way of pricing their car and so the prices differ from one car to another. So, the main motive of this project is to assure that the money they would invest in the car will be worthy. For the prediction of the price of used cars we applied the supervised

machine learning techniques. The predictions are based on dataset collected from various website and Kaggle Website mostly. Different techniques like multiple linear regression analysis, decision trees and k-nearest neighbors have been used to make the predictions. The predictions are then rate and compared Data Which we are collected in order to find those which provide the best performances. From this we can see that this easy problem turned out to be indeed very difficult to resolve with high accuracy. All these four methods provided performance and comparable. In the Upcoming life span, we intend to use more sophisticated algorithms to make the predictions.

Paper 6: Prediction of The Prices of Second-Hand Cars

- **Publication Year: May 2019**
- **Author: Ozer Celik, U. Omer Osmanoglu.**
- **Summary:** In today's economic conditions, interest in second hand products has increased. Especially second-hand car or vehicles have a wide customer base. In the sector which has a workshop market, it is very important to make fast sales, to make the right pricing and to calculate the ideal prices of the cars in order to exchange at the right price. With linear regression analysis second-hand in such cases first determination of variables with effect on price, then it is possible to calculate the price by establishing estimating model. In this study, the model was established by determining 23 of 78 variables affecting the price such as price, brands and model years of 5041 second-hand cars. The Determination Rate (R2) of these 23 variables was found to be 89.1%. Then, by using this regression model, second hand prices of the cars were estimated via machine learning algorithm. The data set is divided into two as training and test data (70-30% and 80-20%). As a result of the study, it was determined the affinities between the real values and the estimated values. The proximity rate ($\pm\%$) calculated in result of study shows affinity intensity of the estimation results to the true results. Via the prediction model established as a result of machine learning, the predictive accuracy rate was found to be 81.15%

according to the 10% proximity of the correct results (upper limit; 110%, lower limit; 90%). According to the results, it is thought that machine learning technique could be second-hand to estimate second hand car prices. However, it is possible to reach a better estimation rate with a data set with more units and different variables.

Paper 7: Used Car Price Predicting Analysis System (UCPAS)

- **Publication Year: 27 Oct 2021**
- **Author: Yadav, A., Kumar, E., & Yadav, P. K.**
- **Summary:** The highly interesting research area that noticed in the last few years is object detection and find out the prediction based on the features that can be benefited to consumers and the industry. In this paper, we understand the concept of object detection like the car detection, to look into the price of a second-hand car using automatic machine learning methods. We also understand the concept of object detection categories. Nowadays, the most challenging task is to determine what is the listed price of a used car on the market, Possibility of various factors that can drive a used car price. The main objective of this paper is to develop machine learning models which make it possible to accurately predict the price of a second-hand car according to its parameter or characteristics. In this paper, 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.

Paper 8: Fair Price Prediction System for Used Cars in Sri Lanka Using Machine Learning and Robotic Process Automation.

- **Publication Year:** 22 Nov 2019
- **Author:** T P Jayadeera, D J Jayamanne.
- **Summary:** The prices of most of the brand-new cars are rapidly increased by car manufacturers due to the increment of prices in raw materials and inflation rates. Also, the taxes for importing brand new cars have been rising consistently during the last decade in Sri Lanka. Due to these factors, most of the time Sri Lankan middle and lower-class people tend to buy used cars (Toyota Corolla) rather importing brand-new cars. With this increment of demand in used cars, some of the used car sellers take advantage of this scenario by listing unrealistic prices for the used cars and most of the used car buyers getting caught on this. This work focuses on creating a used car price prediction system for Sri Lanka using supervised learning techniques. For the study, Analysis with different suitable machine learning models is performed using an online data set to discover the best suitable regression models. The selected models are trained using an actual Sri Lankan used car data set which is expected to be extracted from online car advertisement websites with the help of Robotic Process Automation (RPA) technology. In the end, a web application will be presented to the users (Used car buyers) which accepts features of the car as inputs and provide a predicted fair price for the expected particular used car.

Paper 9: Vehicle Price Prediction System using Machine Learning Techniques.

- **Publication Year: June 2017**
- **Author: Kanwal Noor, Sadaqat Jan.**
- **Summary:** This paper presents a vehicle price prediction system by using the supervised machine learning technique. The research uses multiple linear regression as the machine learning prediction method which offered 98% prediction precision. Using multiple linear regression, there are multiple independent variables but one and only one dependent variable whose actual and predicted values are compared to find precision of results. This paper proposes a system where price is dependent variable which is predicted, and this price is derived from factors like vehicle's model, make, city, version, color, mileage, alloy rims and power steering.

Paper 10: Predicting Used Car Prices.

- **Publication Year: 2011**
- **Author: Kshitij Kumbar, Pranav Gadre, Varun Nayak.**
- **Summary:** Determining whether the listed price of a used car is a challenging task, due to the many factors that drive a used vehicle's price on the market. The focus of this project is developing machine learning models that can accurately predict the price of a used car based on its features, in order to make informed purchases. They implement and evaluate various learning methods on a dataset consisting of the sale prices of different makes and models across cities in the United States. Their results show that Random Forest model and K-Means clustering with linear regression yield the best results, but are compute heavy. Conventional linear regression also yielded satisfactory results, with the advantage of a significantly lower training time in comparison to the aforementioned methods.