IBM Team ID: PNT2022TMID21714
Team Leader: Silambarasan S (142219104121)
Team Members: Siddharth P (142219104120)
Srinath R (142219104127)
Samuel jayaraj J (142219104108)

Domain Name: Applied Data Science

Use Case Name: Car Resale value Prediction

Paper 1

Authors: Nitis Monburinon Faculty of Information Technology
Year : *2018*

Title*: *Prediction of Prices for Used Car by Using Regression Models

Methodology: The used car data used in this research were collected from www.kaggle.com which uploaded by Orges Leka under the public domain license. This dataset consists of 371,528 carobservations and the attributes of used car are from eBayKleinanzeigen .

Advantage: Several related works have been done previously on the subject of used car price prediction. Pudaruth [1] predicted the price of used cars in Mauritius using multiple linear regression, k-nearest neighbors, naive Bayes and decision trees.

Disadvantage: Although their results was not good for prediction due to a less number of car observation. Pudaruth concluded in his paper that the decision tree and naive Bayes are unable to use for variable with a continuous value.

Paper 2
Authors: Shen Gongqi, Wang Yansong, Zhu Qiang

Year : *2011*

Title: A New Model for Residual Value Prediction of the Used Car Based on BP Neural Network and Nonlinear Curve Fit

Methodology: Some modeling assumptions were set as following: (1) the used car we mentioned here was private car only, not including the car used as commercial car such as taxi or as the chauffeur-driven car in government or company, which is different from the private car in appraisal conditions. (2) The data we acquired were from the used car trade market in Shanghai which may be a little different from the other place.

Advantage: According to the analysis of the recorded data, several factors which were considered effect on the used car residual value substantially were chosen as following: the car manufacturer, model, mileage, age, maintenance record, physical condition, market occupancy, after-sale service, and the driving habit of owner. The above-mentioned factors were the original factors which were interconnected with each other, affecting the residual value coherently. If the relationship among these factors was neglected and the function of factors to residual value was calculated respectively, it will lower the accuracy of the prediction model. To solve this issue, according to the literature [2], four strengthened factors were adopted to replace the original factors, the strengthened factors were shown .

Disadvantage: . Distribution curve of residual value In the common sense, a car starts to lose its value since the day it was bought, at the beginning of the first several years, the slope of distribution curve of residual value over time is sharp, which then slowly turn to smooth. It is widely accepted in the field of used car appraisal that the lifetime of the car ,usually 15 years, is divided into 5 parts sequentially and averagely, each part constitutes the 5/15, 4/15, 3/15, 2/15, 1/15 of the whole new car value respectively

Paper 3
Authors: Chejarla Venkata Narayana

Year : *2021*

Title: Second Sale Car Price Prediction using Machine Learning Algorithm

Methodology: 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 modeled for prediction analysis and Extra Trees Regression fits the number of decision trees. The results are so encouraging with our approach.

Advantage: Car Price Prediction Using Machine Learning Techniques. Considerable number of distinct attributes are examined for the reliable and accurate prediction. To build a model for predicting the price of used cars in Bosnia and Herzegovina, they have applied three machine learning techniques (Artificial Neural Network, Support Vector Machine and Random Forest).

Disadvantage: Predicting the price of Used Car Using Machine Learning Techniques.In this paper, they investigate the application of supervised machine learning techniques to predict the price of used cars in Mauritius. The predictions are based on historical data collected from daily newspapers. Different techniques like multiple linear regression analysis, k-nearest neighbours, naïve bayes and decision trees have been used to make the predictions.