CAR RESALE VALUE PREDICTION

TEAM ID:

PNT2022TMID15858

TEAM MEMBERS:

1) MUGUNTHAN M   (927619BEC4121)
2) MUHIN V   (927619BEC4122)
3) KESAVA SAI RAAM C N   (927619BEC4102)
4) MOHANRAJ T R   (927619BEC4120)

INDUSTRY MENTORS NAME:

1) SOWJANYA
2) SANDEEP DOODIGANI

FACULTY MENTOR NAME:

1) Dr.K KARTHIKEYAN

ABSTRACT

Car price prediction using AI / Machine Learning techniques has picked researchers interest since it takes a significant amount of work and expertise on the part of the field expert. For a dependable and accurate forecast, a large number of unique attributes are analyzed. We employed 6 different machine learning approaches to develop a model for forecasting the price of used automobiles.

PROPOSED METHODOLOGY

Car resale value prediction system is made with the purpose of predicting the correct valuation of used cars that helps users to sell the car remotely with perfect valuation and without human intervention in the process to eliminate biased valuation. Due to limited data, system only takes into account limited features for predicting the resale value of the car. Since this is an online system, current system does not take into account any physical damage to the car body or engine while predicting the resale value. To achive accurate prediction and better model training, huge dataset of resale cars of Swift Dezire of 5 cities is gathered via web scraping cars24 website. This dataset contains data of 5 main features i.e., fuel type, kms driven, city, car purchase year and resale value. Here resale value becomes our target column whereas other columns served as features for our model. Data scraped consists of many unwanted characters like comma, whitespaces etc. which has to be removed as model can only understand numbers. Moreover, fuel type was converted into numerical codes via one-hot encoding.

LITERATURE REVIEW

TOPIC : PREDICTION OF RESALE VALUE OF THE CAR USING LINEAR REGRESSION ALGORITHM

AUTHOR : Kiran S (2020)

DESCRIPTION: The Linear Regression model for prediction of resale value of the car is provided. Linear Regression model is better suited for prediction of target attribute that is msrp (car price) and it is performing very good . Further this work can be implemented using different machine learning algorithms and approaches in order to get higher accuracy rate and lower error percentage.

TOPIC: CAR PRICE PREDICTION USING MACHINE LEARNING TECHNIQUES

AUTHOR : Enis Gegic (2019)

DESCRIPTION: Applying single machine algorithm on the data set accuracy was very less. Therefore, the ensemble of multiple machine learning algorithms has been proposed and this combination of ML methods gains the highest accuracy,This is significant improvement compared to single machine learning method approach. However, the drawback of the proposed system is that it consumes much more computational resources than single machine learning algorithm.

TOPIC : PRICE EVALUATION MODEL IN SECOND-HAND CAR SYSTEM BASED ON BP NEURAL NETWORK THEORY

AUTHOR : Ning Sun(2017)

DESCRIPTION: In this paper, they introduce an application of the online second-hand car price evaluation model, and analyze the drawbacks of the traditional formula method and the accuracy of the proposed model. on the basis of the BP neural network algorithm, we optimize the algorithm locally, which greatly improves the speed and the accuracy of this model and makes the system have the practical application value.

TOPIC : VEHICLE PRICE PREDICTION SYSTEM USING MACHINE LEARNING TECHNIQUES

AUTHOR : Kanwal Noor(2017)

DESCRIPTION: Using Minitab, we get the price being predicted in the additional column "FIT". Apart from it, the residual value being the difference between actual and predicted response variable is also calculated. Results' samples of actual and predicted price indicating the number of observations, FIT, Resid and standard Resid.

TOPIC : A METHODOLOGY FOR PREDICTING USED CARS PRICES USING RANDOM FOREST

AUTHOR : Nabarun Pal(2018)

DESCRIPTION: In this paper they used many techniques like supervised learning method known as Random Forest. Kaggle's dataset was used as a basis for predicting used car prices. In order to determine the price impact of each feature, careful exploratory data analysis was performed. 500 Decision Trees were trained with Random Forests. It is most commonly used for classification, but they turned it into a regression model by transforming the problem into an equivalent regression problem.

TOPIC : PREDICTING THE PRICE OF USED CARS USING MACHINE LEARNING TECHNIQUES

AUTHOR : Sameerchand Pudaruth (2018)

DESCRIPTION: In this paper, four different machine learning techniques have been used to forecast the price of used cars in Mauritius. The mean error with linear regression was about Rs51, 000 while for kNN it was about Rs27, 000 for Nissan cars and about Rs45, 000 for Toyota cars. J48 and NaiveBayes accuracy dangled between 60-70% for different combinations of parameters. The main weakness of decision trees and naïve bayes is their inability to handle output classes with numeric values. Hence, the price attribute had to be classified into classes which contained a range of prices but this evidently introduced further grounds for inaccuracies. The main limitation of this study is the low number of records that have been used. As future work, we intend to collect more data and to use more advanced techniques like artificial neural networks, fuzzy logic and genetic algorithms to predict car prices.

REFERENCE

[1]   Annina S, Mahima SD, Ramesh B, "An Overview of Machine Learning and its Applications". International Journal of Electrical Sciences & Engineering (IJESE), 2015. pp. 22-24.

[2]   Karl Storchmann, "On the depreciation of automobiles: An international comparison". Department of Economics, Yale University, 2004. pp. 372-373.

[3]   Sakshi Gupta, "Regression vs. Classification in Machine Learning", October 6, 2021 [Online] Available:https://www.springboard.com/blog/data-science/regression-vsclassification/ [Accessed March 6, 2022].

[4]   Jim Frost, "Overfitting Regression Models," Statistics by Jim, 2022. [Online],Available:https://statisticsbyjim.com/regression/overfitting-regressionmodels/ [Accessed April 18, 2022].

[5]   Skikit-learn,supervised-learning.[Online]Available: https://scikitlearn.org/stable/supervised_learning.html#supervised-learning [Accessed March 12, 2022]

[6]   Aarshay Jain, "A Complete Tutorial on Ridge and Lasso Regression in Python",January,28,2016.[Online],Available:https://www.analyticsvidhya.com/blog/2016/01/ridge-lassoregression-python-complete-tutorial/ [Accessed March 12, 2022]

[7]   Leo Breiman, "Random Forests". University of California Berkeley, 2001.

[8]   Ajitesh Kumar, "Mean Squared Error or R-Squared – Which one to use?". [Online], Available: https://vitalflux.com/mean-squareerror-r-squared-which-one-to-use/[accessed March 12, 2022].

[9]   Ahmad Abdulal, Nawar Aghi, "House Price Prediction". C.S. Bachelors Thesis, Kristianstad's University, 2020.