CAR RESALE VALUE PREDICTION

Literature Survey:

In order to meet the needs of second-hand car value assessment, the used car value assessment system has been designed based on the improved replacement cost method. The system includes system management module, used car parameter management module, used car evaluation management module and evaluation information inquiry module. We enter the relevant basic information of second-hand car information, and figure out the used car's new rate, the purchase price and the selling price through the calculation to the system.

The global used car market size was valued at USD 1.57 trillion in 2021 and is expected to expand at a CAGR of 6.1% from 2022 to 2030. The used car shipment was recorded at 120.3 million units in 2021. The market has witnessed significant growth in the last few years as the price competitiveness among the new players has been one glowing spot in the used car industry. The inability of customers to buy new cars became one of the reasons for the growing used cars sales volume, which is complemented by the investments made by the industry participants to establish their dealership network in the

market. These dealership networks helped market participants to brand and make used car options viable.

The paper investigates statistical models for forecasting the resale prices of used cars. First, a comparative analysis of alternative prediction methods provides evidence that random forest regression is particularly effective for resale price forecasting. It is also shown that the use of linear regression, the prevailing method in previous work, should be avoided. Second, the empirical results demonstrate the presence of heterogeneity in resale price forecasting and identify methods that can automatically overcome its detrimental effect on the forecast accuracy. Finally, the study confirms that the sellers of used cars possess informational advantages over market research agencies, which enable them to forecast resale prices more accurately.

The paper discussed about the multivariate used for the prediction of value need in the developed model. One of the most important uses of multivariate modeling is precisely that 'of controlling for confounding' to let emerge the effect of the risk factor of interest on the study outcome. In this paper, they discuss linear regression analysis for the examination of continuous outcome data and logistic regression analysis for the study of categorical outcome data. Furthermore, it focuses on the most important application of multiple linear and logistic regression analyses.

The study examined and compared the performances of the Random Forest, k-Nearest Neighbour and SVM classifiers for land use/cover classification using Sentinel-2 image data. An area of 30 × 30 km² within the Red River Delta of Vietnam with six land use/cover types was classified using 14 different training sample sizes, including balanced and imbalanced, from 50 to over 1250 pixels/class. All classification results showed a high overall accuracy (OA) ranging from 90% to 95%. Among the three classifiers and 14 sub-datasets, SVM produced the highest OA with the least sensitivity to the training sample sizes, followed consecutively by RF and kN. The high accuracy was achieved with both imbalanced and balanced datasets.