

IBM – NALAYATHIRAN

CAR RESALE VALUE PREDICTION

LITRETURE SURVEY

CAR RESALE VALUE PREDICTION

The first paper is 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.

The Second paper is 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).

The Third paper is Price Evaluation model in second hand car system based on BP neural networks. In this paper, the price evaluation model based on big data analysis is proposed, which takes advantage of widely circulated vehicle data and a large number of vehicle transaction data to analyze the price data for each type of vehicles by using the optimized BP neural network algorithm. It aims to establish a second-hand car price evaluation model to get the price that best matches the car.

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 [2] has introduced and explained how bias-variance tradeoff is achieved in both regression and classification. The selection of variables/attribute plays a vital role in influencing both the bias and variance of the statistical model. Robert Tibshirani [3] 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. Trevor Hastie and Daryl Pregibon [4] have explained the concept of pruning in their research paper. Moreover, hypothesis testing using ANOVA is needed to verify whether the different groups of errors really differ from each other. This is explained by TK Kim and Tae Kyun in their paper [5].

Conclusion

The increased prices of new cars and the financial incapability of the customers to buy them, Used Car sales are on a global increase. Therefore, there is an urgent need for a Used Car Price Prediction system which effectively determines the worthiness of the car using a variety of features. The proposed system will help to determine the accurate price of used car price prediction. This paper compares 3 different algorithms for machine learning : Linear Regression, Lasso Regression and Ridge Regression.

List of References

- [1] Kanwal Noor, 2017, Vehicle Price Prediction System using Machine Learning Techniques International Journal of Computer Applications. Volume 167 - Number 9
- [2] Mariana Lusitania et al, (2009). Support vector regression analysis for price prediction in a vehicle leasing application
- [3] Richardson, M. S. (2009). Determinants of used vehicle resale value.
- [4] Listiani, M. (2009). Support vector regression analysis for price prediction in a car leasing application (Doctoral dissertation, Master thesis, TU Hamburg-Harburg).
- [5] Doan Van Thai, Luong Ngoc Son, Pham Vu Tien, Nguyen Nhat Anh, Nguyen Thi Ngoc Anh, "Prediction car prices using qualify qualitative data and knowledge-based system" (Hanoi National University)
- [6] Wu, J. D., Hsu, C. C., & Chen, H. C. (2009). An expert system of price forecasting for used cars using adaptive neuro-fuzzy inference. Expert Systems with Applications, 36(4), 7809-7817
- [7] Gongqi, S., Yansong, W., & Qiang, Z. (2011, January). New Model for Residual Value Prediction of the Used Car Based on BP Neural Network an Nonlinear Curve Fit. In Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on (Vol. 2, pp. 682.
- [8]. Pudaruth, S. (2014). Predicting the price of used cars using machine learning techniques. Int. J.Inf.Comput. Technol, 4(7), 753-764.
- [9]. Noor, K., & Jan, S. (2017). Vehicle Price Prediction System using Machine Learning