Used Car Price Prediction

By performing different models, it was aimed to get different perspectives and eventually compared their performance. This study, its purpose was to predict Dealer Listing prices of used cars. With the help of the data visualizations and exploratory data analysis, the dataset was uncovered and features were explored deeply. The relation between features was examined. At the last stage, predictive models were applied to predict the price of cars in order: random forest, linear regression, ridge regression, lasso, Lasso, XGBoost.
Which variables are significant in predicting the price of a car? newest car, vehicle with 1 owner, buyback and protection eligible, extended warranty, Vehicle exterior color with Bright White Clearcoat and vehicle with interior black.

- The Jeep manufacture has a higher price than Cadillac.
- The price of the newest car is higher than the oldest car.
- The vehicle with 1 owner, buyback, and protection eligible has more demand and the price of the vehicle is more than others.
- Some sellers in states like IL, MD, IN, NC, PA, GA, MN, and WI sell their vehicles at the highest price.
- The price of all vehicles with an extended warranty is lower than other vehicles.
- The SUV is more in demand than other styles.
- The Vehicle exterior color with Bright White Clearcoat is more in demand and to be sold followed by other colors. The most vehicle with interior black is in demand and sold more than other colors

Future Work

- Keeping the current model as a baseline, we intend to use some advanced techniques algorithms to predict used car prices as our future work.
- To improve model performance, I will use hyperparameter tuning like GridSearchCV, and to improve prediction performance I will use ensemble methods like voting classifiers to get the better result.
- Cross-Validation to improve the R-Squared and use adjusted R Squared to compare fits.
- Adding more independent variables to a regression model tends to increase our metrics.

References

Jupyter Notebook, ML

1. https://scikit-learn.org/stable/supervised_learning.html#supervised-learning
2. https://www.kaggle.com/yogidsba/predict-used-car-prices-linearregression
3. https://www.kaggle.com/iabhishekmaurya/used-car-price-prediction/notebook
4. https://towardsdatascience.com/predicting-car-price-using-machine-learning-8d2df3898f16