PRICE PREDICTION OF USED CARS

ORIE 4741 PROJECT PROPOSAL

Jiahe Chen *
jc3472@cornell.edu

Wei Cheng †
wc655@cornell.edu

October 3, 2021

1 Project Description

The question we want to explore from this data set is how to predict the price of used cars based on features like make, year, wheel drive, etc. Although such kind of price prediction can be found everywhere on used car websites, these websites never release their prediction methods and we do not know which features impact the price most.

Answering this question can provide a more transparent guideline to customers and also help the car owners know which parts of the car should be maintained properly or when the car should be sold so that the owner can argue for a better price.

Answering this question can also help companies like local car dealer make better decisions on buying/selling used cars since car price can be very location specific. For example, cars with all-wheel drive (AWD) can be very popular in places with snowy winter due to the better performance in snow or slick surfaces. Cars with high mile-per-gallon (MPG) can be very popular in places with high gas price due to the potential saving of gas consumption. However, most car priced predictions provided by public websites do not incorporate these details and the predicted results could be useless or even misleading for companies.

We use this dataset from Kaggle, which contains the most all relevant information that Craigslist provides on car sales including columns like price, condition, manufacturer, latitude/longitude, and 18 other categories. We may also add other features including but not limited to: gas price of the trading location, weather (snowy? warm?) of the trading location, how many integrated circuits (ICs) are required to manufacture the car, etc. These additional features can help the model make more accurate predictions. The dataset also has more than 400,000 data points. Therefore, it provides enough data for us to train, validate and test different models.

*School of Electrical and Computer Engineering, Cornell University
†Ann S. Bowers College of Computing and Information Science, Cornell University