Analytics exercise: Estimating the winning market price

Estimating the price at which an item will be sold in an auction is one of the main everyday challenges in business. In this example you are asked to estimate a market price for a car and understand how it changes depending on car characteristics, market trends etc.

Data

The data provided consists of approx. 5,000 cars that were sold via a b2b auction in 2018. The price shown in the table is the highest bid that was reached during the auction.

Some data clean-up has already happened (e.g. cars with engine damage were filtered out). However, there may still be minor damages like scratches, but we do not have more information about that.

We have also extracted 8 criteria based on the equipment of car that we think might have a good impact on the value of a used car (e.g. leather seats, air condition, automatic gears, alloy wheels etc.). These features have been labelled *feature1* to *feature8* and are shown in the data below.

Task at hand

We would like you to find the main determinants of the value of a used car depending on the basic description and the 8 provided features. The following questions are of special interest to us:

1. What are the most important characteristics and features that determine the selling price of a used car?
2. How does the estimated value of a car change over time?
    a. Does a relative change in selling price over time differ significantly with respect to any of the car characteristics, e.g., color, price range or features?
    b. Are there any statistically significant seasonality patterns in pricing, e.g. certain car types being more expensive in summer than winter?
3. Assume you need a car for a year (buy it now and sell in 1 year) and will drive approx. 10,000 miles during this time frame. You want to spend at least $20k.

    What particular car from this data set would you buy (selling price has to be at least $20k at the moment of purchase) if you want to minimize the loss in absolute value in $ when you sell it. (Loss in value is $ defined as: Price at purchase – price after a year)

4. Please share the out-of-sample accuracy metric for the model you used to answer the above questions.

5. Feel free to share any other interesting insights worth mentioning.

Deliverables

We would like you to present the answers to the above questions to us (could be a notebook or in powerpoint) including the codes you used to answer these questions. Feel free to include any additional insights you gained through this analysis.

Thank you, have fun and good luck!