Pattern 6 Research Proposal

Topic Proposal

Exploring the used cars market of Belarus and building up a regression model to predict the price of a car based on the specific features that we choose.

SMART

Specific; Build a model to find the relationship between car prices by looking at different factors that include numerical, categorical values and further use the model to predict car prices. Features we will focus on are the following; color, transmission, odometer value, year of production, body type, number of photos duration of days, etc. **Question:** Is it possible to build a model to find a relationship between car prices by looking at different factors that include numerical, categorical values and further use the model to predict car prices?

Measurable: We will use specific measuring metrics such as r-square, MAE, MSE and RMSE for modeling in our price prediction. **Question:** Is it possible to measure metrics such as r-square, MAE, MSE and RMSE with the data set categories in towards coming up with a model?

Achievable: **Question:** Based on the preliminary analysis that the team concluded is it possible to find a pattern between target variable(car price) and the independent variable?

Relevant: Can the research help the sellers and buyers in the used car market to make an informed decision about the price of the vehicle? By this analysis, buyers/sellers would be able make good real time decisions.

Time Oriented: We completed EDA in the first phase of our project work. Now, we are going to proceed in further analysis by modeling of certain features versus price. **Question:** Will the final modeling analysis be completed by December, 7th with the presentation?

Source

The source of our data set: Used cars catalog-Belarus on Kaggle: (https://www.kaggle.com/lepchenkov/usedcarscatalog?select=cars.csv).

Number of Observations; 38,500.

GitHub

The link to our GitHub repository: https://github.com/laihanel/PatternSix.