

Predicting Evaluation Prices of the Danish Real Estate Property

Exam-numbers 115, 144, 146, 204

25. august 2019

The intention of this project is to predict which features are most relevant for evaluation of real estate properties in Denmark. Two prediction models will be used, one with a classification output by price-intervals. The second model with a continuous output. Data will be scraped from boliga.dk, hvorlangterder.dk, statistik.politi.dk, Danmarks Statistik. The data from boliga.dk and hvorlangterder.dk are undoubtedly essential to the model. Whereas the data scraped from statistik.politi.dk and Statistikbanken contributes with potential to include socio-economic factors.

The model will take variation in pricing caused by geographical matters, as well as differences in types of housing into account. Machine learning will be applied, for the purpose of prediction of prices. Because of the large amount of terms and corresponding interaction-terms in the prediction-model, regularization by the Lasso-feature will be used.