

Predicting Valuation Prices of Danish Real Estate Property

Exam-numbers 115, 144, 146, 204

25. august 2019

The intention of this project is to examine which features are most relevant for predicting the valuation of real estate properties in Denmark. Supervised learning by classification- and regression prediction-model are applied to data scraped from boliga.dk, hvorlangterder.dk, statistik.politi.dk, Danmarks Statistik. The data from boliga.dk is undoubtedly essential to the model. Whereas the data collected from statistik.politi.dk and Statistikbanken contributes with potential to include socio-economic factors. The scraped data from hvorlangterder.dk attempts to add further geographical attributes to the prediction-model.

The model will take variation in pricing caused by geographical matters, as well as differences in types of housing into account. Because of the large amount of terms and corresponding interaction-terms in the prediction-model, regularization by the Lasso will be used in our ML-process.