Mathematical Statistics
Stockholm University
Bachelor Thesis **2019:3**
http://www.math.su.se

# Living with Trees
## Predicting Swedish Apartment Prices with eXtreme Gradient Boosting

Andreas Hörnqvist*

January 2019

**Abstract**

The recent advent of statistical learning methods promises accurate predictions and insights. The ideas of what would later be the underpinnings of decision trees was first introduced in the nineteen-fifties. Today the use of serial ensembles of decision trees deliver unparalleled performance on a wide range of learning problems. One method that leverages such a procedure is eXtreme Gradient Boosting. The fair market value of real estate property is usually determined by a licensed appraiser, such as a real estate agent. While knowledgable professionals appraisers are due to make subjective estimates resulting in uncertain assessments. The purpose of this thesis is to produce a model that rivals the predictive performance of appraisers and provides both buyers and sellers objective price estimates. Applying eXtreme Gradient Boosting to a dataset containing records of sold apartments in Sweden during 2013 to 2018 results in such a model. Leveraging bespoke predictors and tuning meta-parameters a predictive model that in fact outperforms appraisers estimates, in terms of list prices, is achieved. The most important features for assessing the market value is determined to be local price point, living area and rent.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: andreas.hornqvist@gmail.com. Supervisor: Chun-Biu Li and Disa Hansson.