



Mathematical Statistics  
Stockholm University  
Master Thesis **2021:18**  
<http://www.math.su.se>

# Tree based methods for non-life insurance pricing

Viktor Johansson\*

September 2021

## Abstract

This study covers the process of modeling claim frequency in non-life insurance. A performance comparison was made between a selection of models for claims frequency in non-life insurance. A portfolio of 163212 motor third party liability policies was analysed for the purpose of making a pricing model. The framework of generalised linear models, well established and still extensively used in the insurance industry, was taken as the standard against which more recent techniques were compared. The models evaluated were regression trees, bagging with trees, random forest and gradient boosting machines. As a gauge of predictive performance, Poisson out-of-sample deviance was used. Other practical considerations from the point of view of an insurance company, such as legal issues and difficulty of implementation, were discussed. Results suggest that in cases where the model assumptions of a generalised linear model appear to met, there is little utility in using a tree based method.

**Keywords:** Non-life insurance pricing, GLM, CART, Random Forest, Gradient Boosting Machine, MTPL insurance.

---

\*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden.  
E-mail: [viktor.johansson@gmail.com](mailto:viktor.johansson@gmail.com). Supervisor: Filip Lindskog.