



Mathematical Statistics
Stockholm University
Master Thesis **2025:4**
<http://www.math.su.se>

Using autoencoders to initialize neural networks for claims prediction

Jonathan Bengtsson*

February 2025

Abstract

In this thesis we discuss predicting number of insurance claims under a Poisson model assumption, using fully connected artificial neural networks. We initialize the weights of these networks using autoencoders, with special attention to the handling of categorical features in the data. More specifically we primarily use a joint embedding of categorical features to learn numerical representation of categories, instead of using the somewhat older and more established way of handling this - separate entity embeddings. We then use these representations together with numerical features to learn representations of all features - representations which we then use to initialize hidden layers in fully connected feed-forward networks. We use denoising autoencoders and undercomplete autoencoders. We evaluate prediction power on a real car insurance data set and find evidence of improvement in comparison to standard methods.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden.
E-mail: jonathan.e.bengtsson@gmail.com. Supervisor: Mathias Millberg Lindholm.