# What are the basic similarities and differences between insurance risk modelling and standard machine learning techniques? – An Overview

EAA e-Conference on Data Science & Data Ethics

12th May 2022

Dr. Jürg Schelldorfer, Actuary SAA

Senior Actuarial Data Scientist, Swiss Re

Chair of the «Data Science» working group of the Swiss Association of Actuaries (SAA)

# Disclaimer

The opinions expressed in this presentation are those of the author only. They are inspired by the work that the author is doing for both Swiss Re and the SAA, but they do not necessarily reflect any official view of either Swiss Re or the SAA.

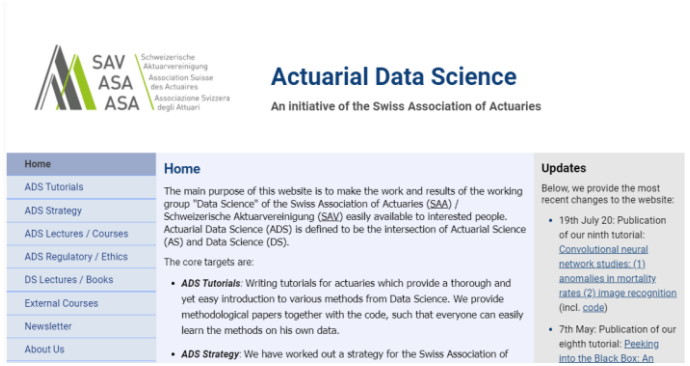
# Insurance risk modelling and machine learning

| | Insurance Risk Modelling | Standard Machine Learning |
|---|---|---|
| Foundation | Distribution and uncertainty | Point estimate and algorithm |
| Mathematical foundation | Statistical model | Numerical optimization |
| Modelling target | Probabilistic forecast | Point forecast |
| Statistical distributions | Non-Gaussian (asymmetric, skewed) | Gaussian (symmetric) |
| Signal-to-noise (SNR) ratio | Small | High |
| Mathematical model selection «criteria» | • Predictability (in-sample)<br>• Stability and robustness (long-term)<br>• Smoothness<br>• Parsimony<br>• Interpretability / explainability<br>• -<br>• - | • Predicatability (out-of-sample)<br>• Stability and robustness (short-term)<br>• -<br>• Anti-parsimony<br>• Black-box<br>• Computability<br>• Calibration |
| Non-mathematical model selection «criteria» | • Causality / truth between predictors and predictant<br>• Inclusion of expert knowledge<br>• Human adjustability of models | • Correlation, train/test paradigm<br>• -<br>• - |
| Non-technical considerations | • Regulatory framework<br>• Political and social aspects | • Ethics and fairnesss<br>• Accountability and transparency |
| Professional associations | • Professional standards | • Ethical codes of conduct |

# Appendix

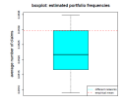
# SAA Data Science Working Party (1/2)

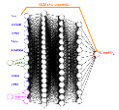
www.actuarialdatascience.org

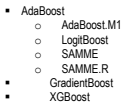



# SAA Data Science Working Party (2/2)

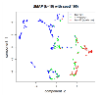
1.  Recap of GLM's, introduction to **regression trees** and **neural networks**

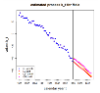
2.  **Best practice** step-by-step guide to fit **neural networks for actuarial applications**

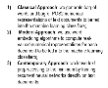
3.  How to embed a GLM into a neural networks explore additional structure

4. 

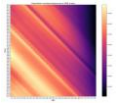
Overview of **boosting algorithms**

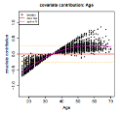
5.  **Unsupervised learning:** Clustering data into homogeneous groups

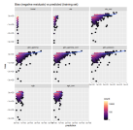
6.  Recurrent neural networks (RNNs) are used use on a **mortality rate prediction problem**

7.  Overview of the three **approaches to preprocess text** data with **NLP**

8.  Overview of **tools for explaining and interpreting black box machine learning models** like boosted trees or deep neural networks for P&C pricing

9.  **Convolution Neural Networks** (CNN) to detect anomalies **in mortality rates**

10.  **LocalGLMnet:** a deep learning architecture for actuaries, allowing variable selection and nice interpretations

11.  Revisiting and clarifying (i) statistical techniques to assess the calibration of a model, and (ii) how to compare and rank different models.