



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

Insurance Data Science Conference

Milan, 16th June 2022

Dr. Jürg Schelldorfer

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

Senior Actuarial Data Scientist, Swiss Re





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.



Perspective	Insurance Risk Modelling	Standard Machine Learning
Foundation	Distribution and uncertainty	Point estimate and algorithm
Mathematical foundation	Statistical model	Numerical optimization
Data foundation	Small and medium; structured data	Big; unstructured data
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»	<ul style="list-style-type: none"> • Predictability • Stability and robustness (long-term) • Smoothness • Parsimony • Interpretability / explainability • - • - 	<ul style="list-style-type: none"> • Predictability • Stability and robustness (short-term) • - • Anti-parsimony • Black-box • Computability • Calibration
Non-mathematical model selection «criteria»	<ul style="list-style-type: none"> • Causality / truth between predictors and predictant • Inclusion of expert knowledge • Human adjustability of models 	<ul style="list-style-type: none"> • Correlation • - • -
Non-technical considerations	<ul style="list-style-type: none"> • Regulatory framework • Political and social aspects 	<ul style="list-style-type: none"> • Ethics and fairness • Accountability and transparency
Professional associations	<ul style="list-style-type: none"> • Professional standards 	<ul style="list-style-type: none"> • Ethical codes of conduct

*still a hypothesis, additional research needed



Links

www.actuarialdatascience.org

<https://www.youtube.com/watch?v=uHGICi9jOWY>



Appendix



References

1. B. Efron (2020) Prediction, Estimation, and Attribution. *Journal of the American Statistical Association* **115:539**, 636-655
2. G. Shmueli (2010) To explain or to Predict? *Statistical Science* **25/3**, 289-310
3. L. Breimann (2001) Statistical Modeling: The Two Cultures. *Statistical Science* **16/3**, 199-215



www.actuarialdatascience.org



Actuarial Data Science
An initiative of the Swiss Association of Actuaries

Home	Home	Updates
ADS Tutorials	<p>The main purpose of this website is to make the work and results of the working group "Data Science" of the Swiss Association of Actuaries (SAA) / Schweizerische Aktuarvereinigung (SAAV) easily available to interested people. Actuarial Data Science (ADS) is defined to be the intersection of Actuarial Science (AS) and Data Science (DS).</p> <p>The core targets are:</p> <ul style="list-style-type: none"> • ADS Tutorials: Writing tutorials for actuaries which provide a thorough and yet easy introduction to various methods from Data Science. We provide methodological papers together with the code, such that everyone can easily learn the methods on his own data. • ADS Strategy: We have worked out a strategy for the Swiss Association of 	<p>Below, we provide the most recent changes to the website:</p> <ul style="list-style-type: none"> • 19th July 20: Publication of our ninth tutorial: Convolutional neural network studies. (1) anomalies in mortality rates. (2) image recognition (incl. code) • 7th May: Publication of our eighth tutorial: Peeking into the Black Box: An
ADS Strategy		
ADS Lectures / Courses		
ADS Regulatory / Ethics		
DS Lectures / Books		
External Courses		
Newsletter		
About Us		

Actuarial Data Science in the Swiss Association of Actuaries
Insurance · Zurich, Zurich · 1,671 followers

An initiative of the Swiss Association of Actuaries and its Data Science Working Group.

[Visit website](#)

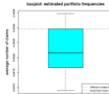


1



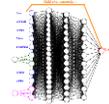
Recap of GLM's, introduction to **regression trees** and **neural networks**

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

- AdaBoost
 - AdaBoost.M1
 - LogitBoost
 - SAMME.R
- GradientBoost
- XGBoost

Overview of **boosting algorithms**

5



Unsupervised learning: Clustering data into homogeneous groups

6



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

7

1) Overview of the three approaches to preprocess text data with NLP

2) Word embeddings

3) Word embeddings in NLP

4) Word embeddings in NLP

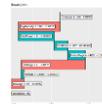
5) Word embeddings in NLP

6) Word embeddings in NLP

7) Word embeddings in NLP

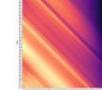
Overview of the three **approaches to preprocess text data with NLP**

8



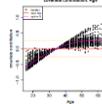
Overview of **tools for explaining and interpreting black box machine learning models** like boosted trees or deep neural networks for P&C pricing

9



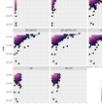
Convolution Neural Networks (CNN) to detect anomalies in **mortality rates**

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.

12



Two actuarial case studies which demonstrate the use **Natural Language Processing Using Transformers**.