

Actuarial Data Science: An Overview

Actuarial Data Science Après-Midi, Zurich
16th September 2020

Dr. Jürg Schelldorfer

Senior Data Scientist, Swiss Re

Chair of the «Data Science» working group, 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.

Table of Content

1. Motivation
2. Strategic considerations
3. Tutorials
4. Insurance Risk Modeling vs Machine Learning*
5. Neural Networks and Random Forest: Learnings*
6. Closing

* if time allows

Motivation

Motivation (1/2)

Patrizia Kern, Head Marine, Swiss Re Corporate Solutions⁽¹⁾

Neue Probleme ahoi

Die Vernetzung und Beschleunigung im Bereich Marine nimmt rasant zu. Fortschritte in der Technologie steigern nicht nur die Effizienz, sondern bergen auch neue Gefahren.

VON PATRIZIA KERN

Die Transportindustrie steht vor einem tiefgreifenden Wandel. In 15 Jahren werden wir sie kaum wiedererkennen. Zum einen werden die Handelsvolumina signifikant steigen; bis 2030 ist beinahe eine Verdreifachung des Handelsverkehrs zu erwarten. Haupttreiber hierfür sind die demografische Entwicklung in Süd-Asien, der Aufstieg der Mittelschichten in Indien, Indonesien und China sowie die Entstehung von modernen Grossstädten (Smart Cities), welche die globale Nachfrage nach Rohstoffen und dem Güterhandel exponentiell ansteigen lassen werden.

Zum anderen wird der technologische Fortschritt eine nahtlose Vernetzung der globalen Marine-Lieferkette ermöglichen, wenn nicht sogar erzwingen. Dafür verantwortlich ist die rasante Entwicklung im Bereich der Sensorik, welche gekoppelt mit der Miniaturisierung und Integration neuer Übertragungstechniken eine permanente digitale Kommunikation zwischen der Herstellung von Gütern, deren Handel, Bereitstellung und Nutzung ermöglichen wird.

Der Weg zu intelligenten Schiffen (Smartships) sowie zu teil-/autonomen Strassen- und Wasserfahrzeugen ist nicht mehr weit. Diese neue Art von Fahrzeugen, gekoppelt mit automatisierten Beladungssystemen, wird nicht nur die Wertschöpfungskette der Schifffahrt erheblich verändern, sondern auch die entsprechende gesamte Risikolandkarte. Es ist davon auszugehen, dass die Seefahrt durch Effizienzsteigerungen sicherer wird, jedoch aufgrund gesteigerter Kapazitäten, Verdichtung des Schiffsverkehrs und Mengen an transportierten Waren die Transportrisiken bedeutend zunehmen werden. Hängen werden wiederkehrende kleinere

Risiken aufgrund der Möglichkeit von vorausschauender Instandhaltung (Predictive Maintenance) sowie rechtzeitiger präventiver und Ad-hoc-Vorbeugungsmassnahmen reduziert werden können.

Gerade die Ermittlung von genaueren Risiko-Szenarien mittels Big-Data- & Smart-Analytics-Techniken stellt eine besondere Herausforderung dar. Die Anzahl an zu untersuchenden Kausalszusammenhängen, der Zugang, die Sammlung, Bereinigung und Analyse von geeigneten Daten sowie die Gewinnung von signifikanten Merkmalen für die Entwicklung von neuen, vorausschauenden Risiko-Modellen (Predictive Modellig) erfordern die Mitwirkung und das Zusammenspiel einer Vielzahl unterschiedlicher Engineering-, Analytics- und Technologie-Disziplinen.

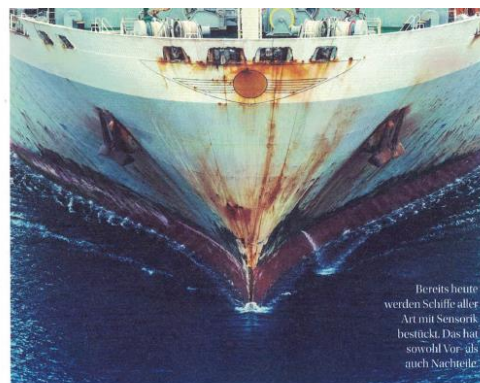
Digitale Transformation

Die Vernetzung und Beschleunigung der Transportindustrie nimmt also rasant zu. In diesem neuen digitalen Kontext werden Ver-

sicherer vor einem gewaltigen transformativen Veränderungsprozess stehen, da ihre Wertschöpfungskette stark herausgefordert, wenn nicht sogar zertrennt wird. Wieso ist dem so? Zunächst müssen wir verstehen, dass die Leistungserstellung der Versicherung vereinfacht ausgedrückt auf der Erfassung der Risiko-Informationen, der Modellierung der Risiko-Szenarien, der Ermittlung der erwarteten Schadenhöhe und der Schadenfrequenz zur Berechnung der jeweiligen Risikoprämie basiert. Traditionell stützen sich Versicherer primär auf verhältnismässig wenig Informationen über den zu versicherten Gegenstand. Anhand historischer Schadensereignisse wird ein «statistisches» Risiko abgeschätzt.

Im digitalen Zeitalter werden Sensor-Daten eine dynamische Betrachtung des Risikos – zum Beispiel durch eine kontinuierliche Überwachung der Zustände der bewegten Güter gekreuzt mit Navigationsdaten, Verkehrsdaten und Wetterdaten – in Echtzeit ermöglichen. Diese neue Ausgangslage wird nicht nur die Leistungserstellung der Versicherung verändern, sondern auch die Leistungsbringung.

Die sich entwickelnde digitale Lieferkette, die primär auf dem Sammeln und Auswerten von Echtzeitdaten basiert, wird die Teilnehmerlandkarte und die Machtverhältnisse vollkommen verändern. Ein Beispiel: Wenn Versicherungsmakler und Versicherer heute ein gutes Verständnis für den Transportmarkt und die entsprechenden Risiken haben, werden morgen Reederei, Schiffbetreiber, Logistiker sowie Industriehersteller ihre Betriebsrisiken besser einschätzen können. Dadurch werden sich die Versicherungsanfrage und der Deckungsbedarf signifikant verändern.



In diesem Kontext müssen Versicherer ihre Kernkompetenzen im Bereich der Analyse von gewaltigen Datenströmen zur Gewinnung neuer Erkenntnisse und Einblicke in der sich laufend verändernden globalen Risikolandkarte fundamental ausweiten. Zudem müssen sie zur Sicherstellung ihrer Beteiligung und Mitwirkung Kooperationen mit Dateneigentümern, Datenanbietern sowie Technologieanbietern eingehen und ausweiten, um Wettbewerbsvorteile und – unter Umständen – Existenzen sicherzustellen.

Konkurrenz von ausserhalb

Digitale Transformationen, ungeachtet der Branche, öffnen eine Vielzahl neuer Opportunitäten und Geschäftsmodelle auch für branchenfremde Akteure. Das gilt vor allem für jene, die eine breite Kundenreichweite, Kundennähe und Kundenbindung bzw. einen weitgreifenden technologischen Vorsprung besitzen. Ausschlaggebend wird künftig jedoch nicht nur das Beherrschen von

neuen technischen Fähigkeiten sein, sondern auch die Bereitschaft, in einen kapitalintensiven Geschäftsbereich einsteigen zu wollen und zu können.

Das Vorhandensein von Daten sowie die Möglichkeit und Fähigkeit, diese in Erkenntnisse (Insights) umzuwandeln, werden sich rasch ausbreiten und diese Nutzung wird sich entlang der Kernkompetenzen eines Unternehmers entfalten. So wird die Fähigkeit, strategische Partnerschaften mit Kunden, Daten-Providern und Technologie-Partnern eingehen zu können, zunehmend auch für Versicherer ausschlaggebend sein, um das digitale Eco-System zu erschliessen. Durch kontinuierliche Innovationen kann so ein Vorsprung geschaffen bzw. profitabel gewachsen werden.

Patrizia Kern ist Head Marine bei Swiss Re Corporate Solutions.

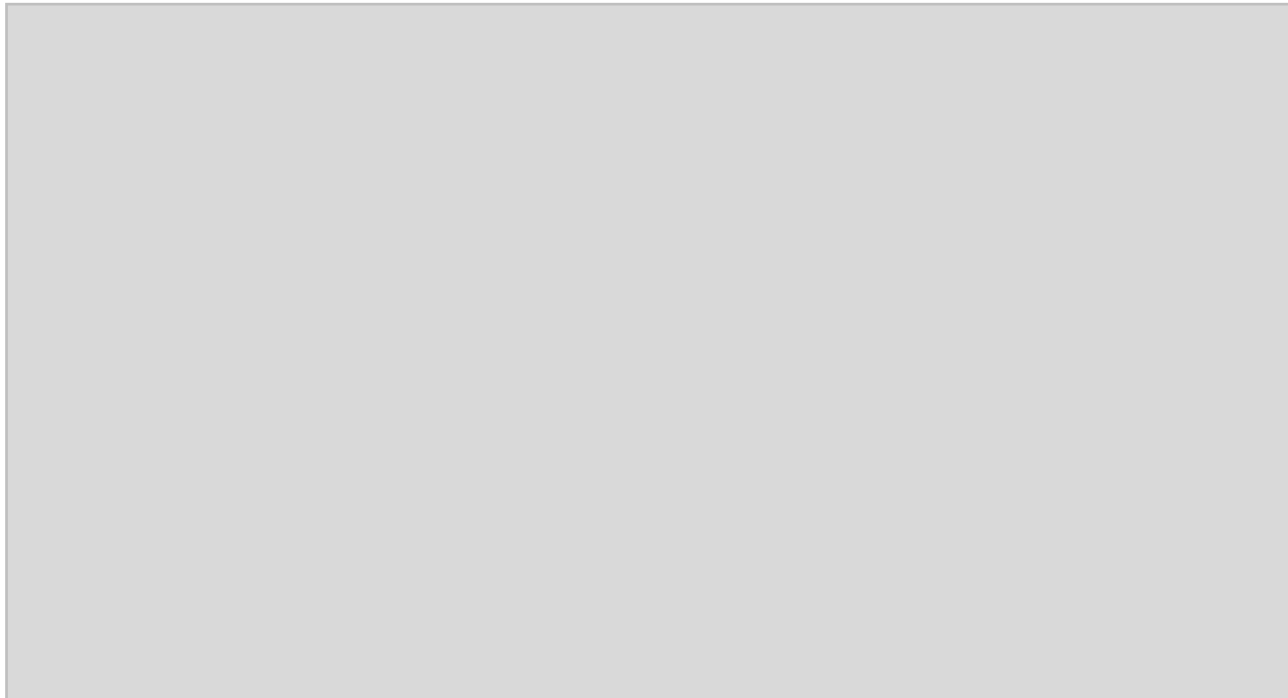
Core messages:

- «Traditionell stützen sich Versicherer primär auf verhältnismässig wenig Informationen über den zu versicherten Gegenstand.»
- «[...] Zusammenspiel einer Vielzahl unterschiedlicher Engineering-, Analytics und Technologie-Disziplinen.»
- « [...] Sammeln und Auswerten von Echtzeitdaten [...]»
- « [...] ihre Kernkompetenzen im Bereich der Analyse von gewaltigen Datenströmen zur Gewinnung neuer Erkenntnisse und Einblicke in [...] Risikolandkarte [...]»
- «[...] Beherrschen von neuen technischen Fähigkeiten sein [...]»

⁽¹⁾ Risk Management Guide 2018, 02.05.2018

Motivation (2/2)

Dr. Tobias Büttner, Head of Claims, Munich Re, mentioned the following¹:



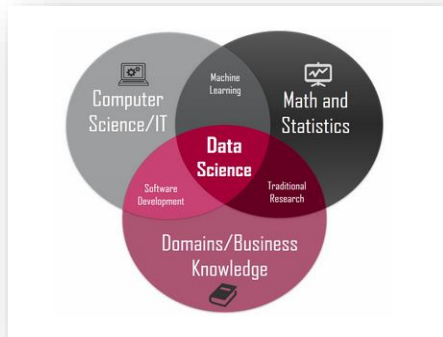
Property claims were assessed using images.

But later the reserves had to be increased significantly. **Damages below/hidden in the roofs have not been appropriately estimated.**

¹ SZ-Fachkonferenz: KI und Data Analytics in der Versicherungsbranche; Data Analytics im Management von Großschäden, Büttner T. (2019), Munich Re

Strategic considerations

What is (Actuarial) Data Science?



Data Science is the intersection of ...

- Mathematics and statistics
- Computer science
- Domains/business knowledge

Data Science is the intersection of ...

- Statistical modelling and/or machine learning
- High computing power
- Big data management
- Domains/business knowledge

Actuarial Data Science = [...] of Actuarial Science and Data Science

	Actuarial Science	Data Science
Basics	Mathematical Basics	
Data	Small Data	Small and Big Data
	Structured & static Data	Unstructured & dynamic Data
	Internal Data	External Data
Mathematics & Statistics	Probability Theory	Computational Statistics
	Life and Non-Life Insurance Mathematics	Algorithm
	Quantitative Risk Management	Information Theory
Computer science		Machine Learning & Visualisations
		Numeric Optimization
		Data Management
Programming languages	SAS, S Plus, R	Python, R
	SQL	SQL
	Excel / VBA	Julia, Spark, Scala
Domains / business knowledge	Reserving, Pricing	
	ERM, ALM, Solvency	
	Accounting, Economics, Law	

Strategy of the SAA (1/2)

DATA SCIENCE STRATEGY
Data Science working group
of the Swiss Association of Actuaries (SAA)

Version 2.0, August 2018
Published version

5th kind Big Data, Analytics & Unstructured Data (2012)

4th kind Enterprise Risk Management (2005)

3rd kind Asset Liability Management (1989)

2nd kind Non-Life Actuary (20th c.)

1st kind Life Actuary (17th c.)

Logo: SAV ASA ASA Schweizerische Aktuarvereinigung Associazione Svizzera degli Attuari

Data Science Strategy

Fokusthema 9 | IWW+IS Trendmonitor 4-2019 | 35

Wie Data Science zukünftig den Beruf des Aktuars prägen wird

Anja Friedrich

Olivier Steiger

Die Digitalisierung sowie die neuen technologischen und datenwissenschaftlichen Fortschritte haben einen starken Einfluss auf die Gesellschaft und die Wirtschaft. Auch der Versicherungsbranche stehen grosse Veränderungen bevor. Damit verändern sich auch die Tätigkeiten und das Berufsbild des Aktuars. Wir zeigen in diesem Artikel, wie der Berufsverband der Aktuarinnen der Schweiz, die Schweizerische Aktuarvereinigung (SAV), diese Veränderungen einschätzt und mit welchen Massnahmen er ihnen begegnet. Damit stellt die SAV sicher, dass die Aktuarinnen in der Schweiz alle notwendigen Kenntnisse und Fähigkeiten besitzen, um die Anforderungen der Gesellschaft, der Wirtschaft und der Politik an den Berufsstand erfüllen zu können.

Die Digitalisierung sowie die neuen technologischen und datenwissenschaftlichen Fortschritte haben einen starken Einfluss auf die Gesellschaft und die Wirtschaft. Auch der Versicherungsbranche stehen grosse Veränderungen bevor. Damit verändern sich auch die Tätigkeiten und das Berufsbild des Aktuars. Wir zeigen in diesem Artikel, wie der Berufsverband der Aktuarinnen der Schweiz, die Schweizerische Aktuarvereinigung (SAV), diese Veränderungen einschätzt und mit welchen Massnahmen er ihnen begegnet. Damit stellt die SAV sicher, dass die Aktuarinnen in der Schweiz alle notwendigen Kenntnisse und Fähigkeiten besitzen, um die Anforderungen der Gesellschaft, der Wirtschaft und der Politik an den Berufsstand erfüllen zu können.

unterdessen verabschiedet und verschiedene Massnahmen befinden sich in der Umsetzung.

Das Verstehen, das Modellieren und das Managen von Versicherungs- und Finanzrisiken bilden die Kernkompetenz von Aktuarinnen. Deshalb sind vertiefte Kenntnisse in den Bereichen Mathematik, Statistik und Computerwissenschaften die Grundlage für die aktuellen Tätigkeiten, ergänzt mit Kenntnissen der Wirtschafts- und Rechtswissenschaften (siehe Abb. 1). Durch die Digitalisierung und die neuen technologischen Möglichkeiten kommt den Bereichen Statistik und Computerwissenschaften aber eine immer grösser werdende Bedeutung zu.

Wirft man einen Blick auf die historische Entwicklung des Aktuarberufs, so kann man diesen in fünf Generationen unterteilen (siehe Abb. 2). Der Tätigkeitsbereich von Aktuarinnen erweiterter sich von Generation zu Generation, wobei die zeitlichen Abstände zwischen den Generationen im Laufe der Zeit immer kürzer wurden. Angesichts dieser historischen Entwicklung erkennen Aktuarinnen der fünften Generation an, dass sich der Beruf des Aktuars immer wieder verändert und das Lernen von neuen Fähigkeiten sowie der Umgang mit neuen Anforderungen für den Aktuar selbstverständlich sind.

Nachfolgend möchten wir die wichtigsten von der SAV definierten Handlungsfelder erläutern und auf die umzusetzenden Massnahmen genauer eingehen.

Die Autoren
Anja Friedrich ist Managerin bei Syngula und Mitglied der Fachgruppe Data Science der Schweizerischen Aktuarvereinigung (SAV).
Dr. Olivier Steiger ist Abteilungsleiter Versicherungstechnik der Swiss und Präsident des Bereichs Ausbildung der Schweizerischen Aktuarvereinigung (SAV).

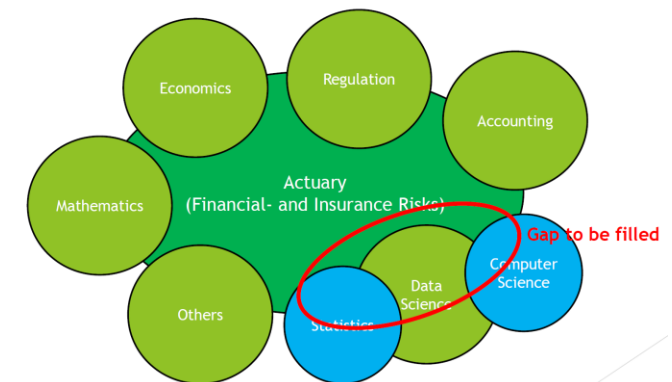
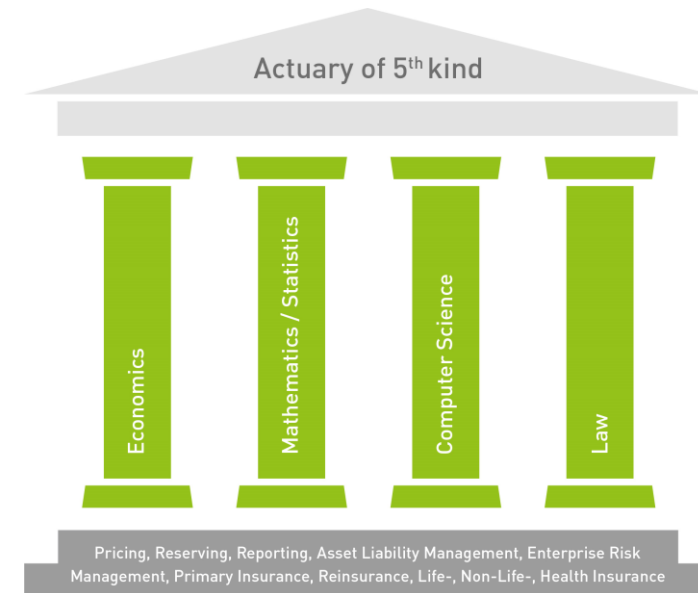
Ausbildung zum Aktuar
Die Inhalte der Aktuaranbildung folgen den Standards der Actuarial Association of Europe (AAE) und der International Actuarial Association (IAA).

Wie Data Science zukünftig den Beruf des Aktuars prägen wird

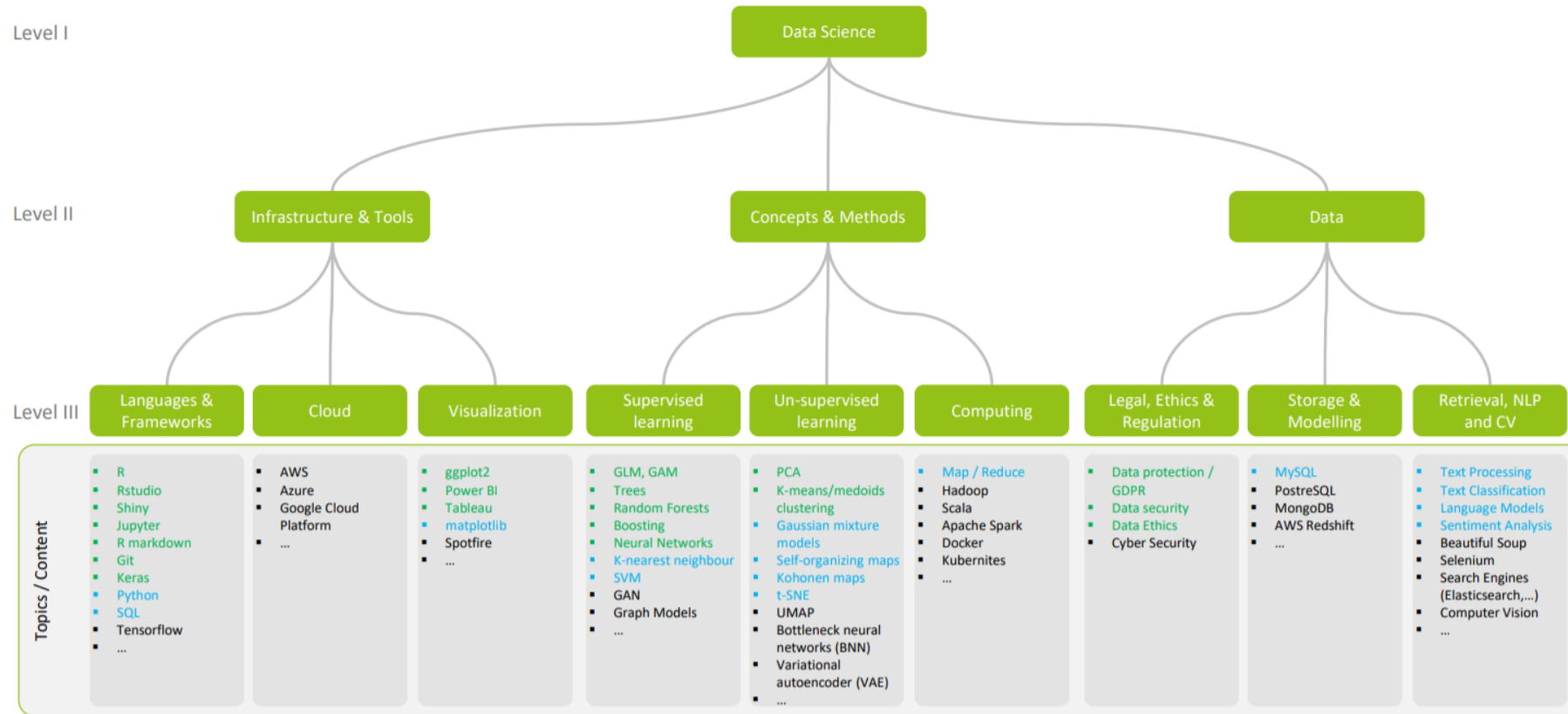
Strategy of the SAA (2/2)

Core considerations:

- (1) Insurance and financial risks are the core business of actuaries.
- (2) In-depth expertise in mathematics, statistics and computer science as a foundation, augmented by literacy in economics and law.
- (3) These skills are needed in non-life, life and health insurance and also in reinsurance.
- (4) Co-existence and collaboration with Data Scientists in the insurance industry.



Topics in Actuarial Data Science



Green: Recommended for actuaries in the industry with some basic knowledge in data science

Blue: Recommended for actuaries in the industry with extended knowledge in data science (combined with green)

Black: Data science (combined with green and blue)

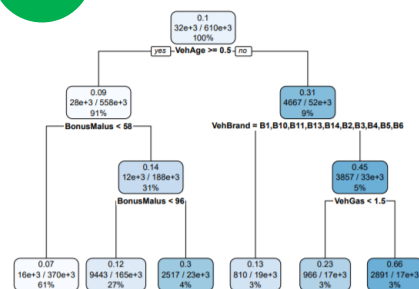
Version V1.1, Oktober 2019

[Upskilling Guide](#)

Tutorials

(Our) Topics in Actuarial Data Science (1/3)

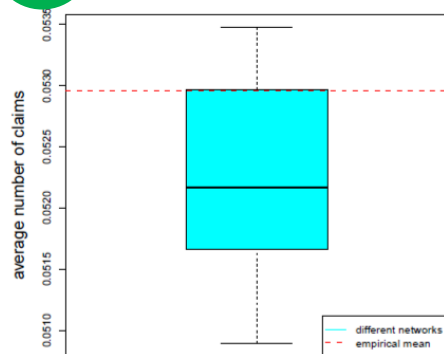
1



Introduction: From GLM to regression trees and neural networks, applied to P&C frequency modeling.

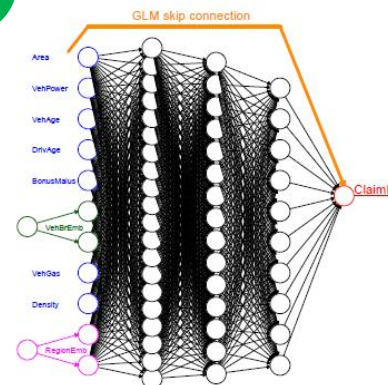
2

boxplot: estimated portfolio frequencies



The tutorial **illuminates different aspects that need to be considered when fitting neural network regression models**. We discuss feature pre-processing, choice of loss function, choice of neural network architecture, class imbalance problem, as well as over-fitting and bias regularization.

3



CANN provide the framework for extending the GLM's, allowing to improve the accuracy of the model as well as providing a framework to **assess the uncertainties**.

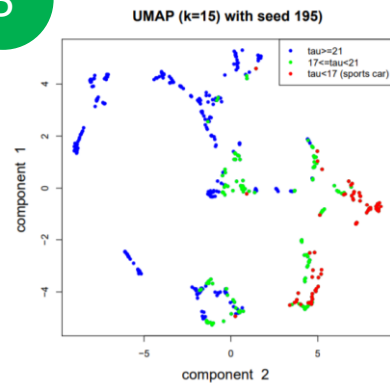
4

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

Boosting: Introduction and a comparison of the many variants of boosting algorithm (implemented in R/Python libraries).

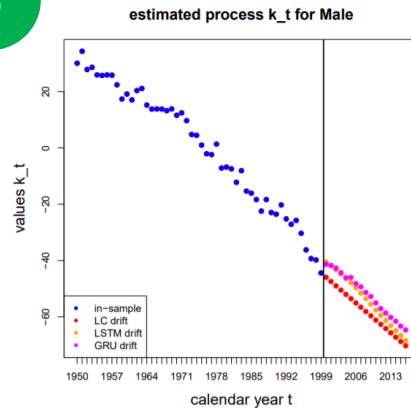
(Our) Topics in Actuarial Data Science (2/3)

5



Unsupervised learning methods are techniques that aim at reducing the dimension of data (covariables, features), cluster cases with similar features, and graphically illustrate high dimensional data.

6



In Recurrent neural networks (RNNs) are introduced, by describing the two most popular RNN architectures, the long short-term memory (LSTM) network and gated recurrent unit (GRU) network. Their common field of application is time series modeling. It is demonstrated their use on a **mortality rate prediction problem**.

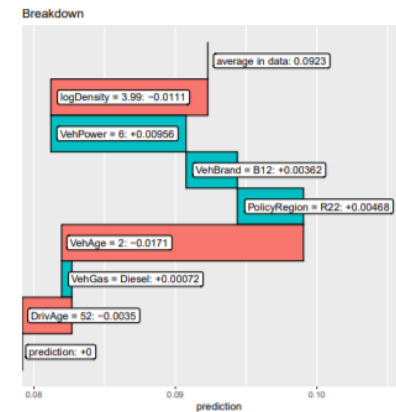
7

- 1) **Classical Approach:** we generate bag-of-words and bag-of-POS2 numerical representations of text documents to be fed into the machine learning classifiers;
- 2) **Modern Approach:** we use word embedding algorithms to compute real-valued numerical representations for each document to be fed into the machine learning classifiers;
- 3) **Contemporary Approach:** we leave text preprocessing to the minimum by training recurrent neural networks directly on text documents.

Three **approaches to preprocess text data** with NLP and perform text document classification using machine learning.

- (1) 'bag-of-' models,
- (2) word embeddings
- (3) the two most popular Recurrent Neural Networks (RNNs).

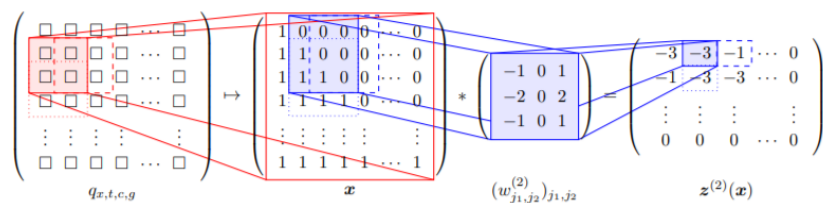
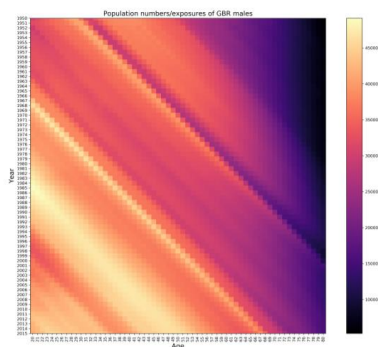
8



This tutorial gives an overview of **tools for explaining and interpreting black box machine learning models** like boosted trees or deep neural networks. All our methods are illustrated on a publicly available real car insurance data set on claims frequencies.

(Our) Topics in Actuarial Data Science (3/3)

9



Convolutional neural networks (CNNs) are particularly well suited to find common spatial structure in images or time series. It is illustrated how to use a **CNN to detect anomalies in mortality rates** taken from the Human Mortality Database (HMD); the anomalies are caused by migration between countries and other errors.

Insurance Risk Modeling vs Machine Learning

Insurance Risk Modeling vs Machine Learning (1/2)

- What are the commonalities and the differences between
 - Insurance risk modeling
 - and
 - Machine learning?
- A simple, but fundamental question:
 - Limited interest of actuaries
 - Machine learning experts are struggling to understand insurance risk modeling
 - Statisticians discuss about statistics vs. machine learning
- References:
 - [Prediction, Estimation, and Attribution](#), B. Efron, Journal of the American Statistical Association 115:539 , 636-655, 2020
 - [To explain or to Predict?](#), G. Shmueli, Statistical Science 25/3, 289-310, 2010
 - [Statistical Modeling: The Two Cultures](#). L. Breimann, Statistical Science 16/3, 199-215, 2001

And thanks to Christian Lorentzen, Swiss Re colleagues and others for contributing!

Disclaimer: The comparison is of course not as strict as indicated using the subsequent table. In many applications it is much more in the middle. And there are many more commonalities (e.g. GLM belong to both).

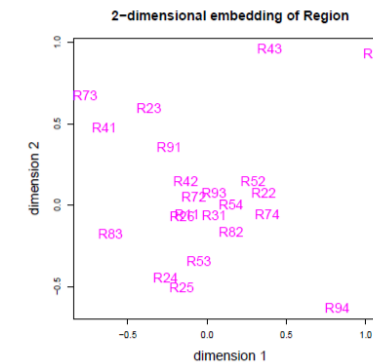
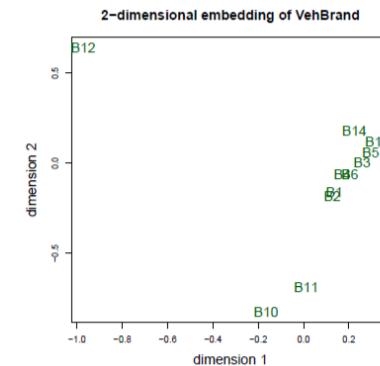
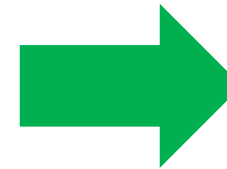
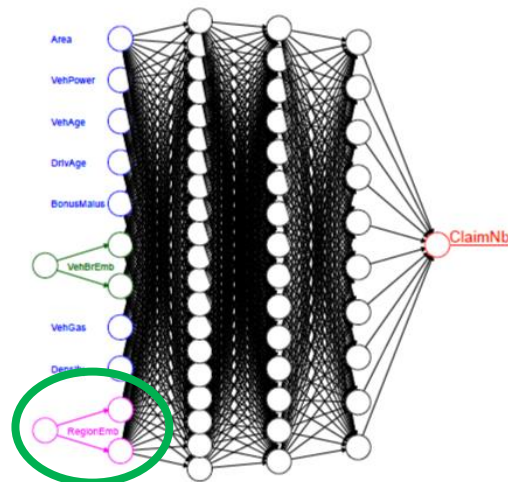
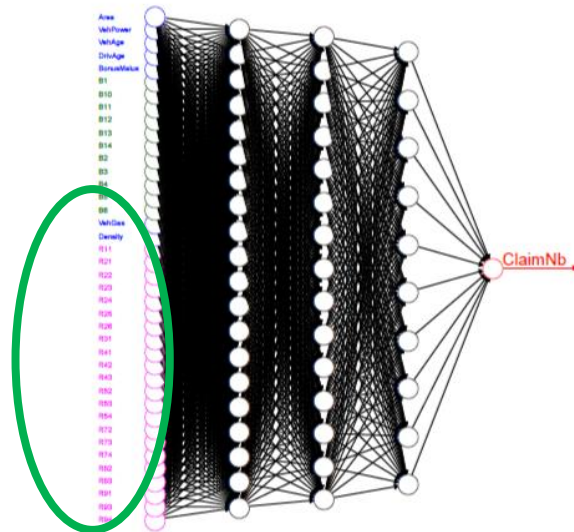
Insurance Risk Modeling vs Machine Learning (2/2)

	Insurance Risk Modeling	Pure Prediction Algorithm
Foundation	Distribution and uncertainty	Point estimate and algorithm
Mathematical foundation	Statistical model	Numerical optimization
Modelling target	Probabilistic forecast Uncertainty quantification	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 (in-sample) • Stability and robustness (long-term) • Smoothness • Parsimony • Interpretability / explainability • - 	<ul style="list-style-type: none"> • Predicatability (out-of-sample) • Stability and robustness (short-term) • - • Anti-parsimony • Black-box • Computability
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, train/test paradigm • - • -
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

Neural Networks and Random Forest: Learnings

1 – Embeddings for categorical features¹

- In insurance pricing, there are often many categorical features (i.e. vehicle brand, region, age group,...) which consist of many levels.
- Usually, the categorical features are encoded as dummy variables (or one-hot encoding), i.e. the levels are orthogonal in the feature space.
- With neural networks, one should use (feature) embeddings:
 - Considerable smaller number of model parameters
 - Weakening the orthogonality assumption
 - Graphical representation in low dimension
 - Prediction performance is not necessarily better with embeddings

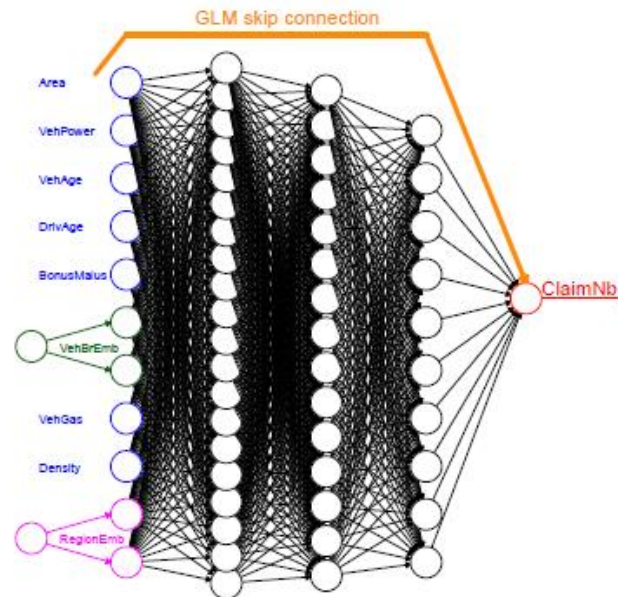


¹ Paper(s): https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3320525

2 – Combined Actuarial Neural Networks (CANN)²

- In most cases, there already exists a productive GLM.
- As the GLM has been developed and fine-tuned over years, there is a very good understanding of the pros and cons of the GLM.
- Instead of building a neural network from scratch, the actuaries are looking for an approach of using neural networks to improve the current GLM.

- The idea is to nest the GLM into a network architecture using a **skip connection** that directly links the input layer to the output layer.
- This approach is called CANN.



- We start the gradient descent algorithm for fitting the CANN model with the GLM solution.
- By that, the **algorithm explores the network architecture for additional model structure that is not present in the GLM**.
- Analyzing the results can hence be used to identify where the GLM needs to be improved, e.g. which interactions are missing.
- In this way we obtain an improvement of the GLM by network features. This provides a more systematic way of using network architectures to improve the GLM.
- CANN allow for **uncertainty quantification** due to its low computational effort.
- This approach is not restricted to GLM's, any regression algorithm can be chosen for the skip connection. However, this is not possible for all regression models.

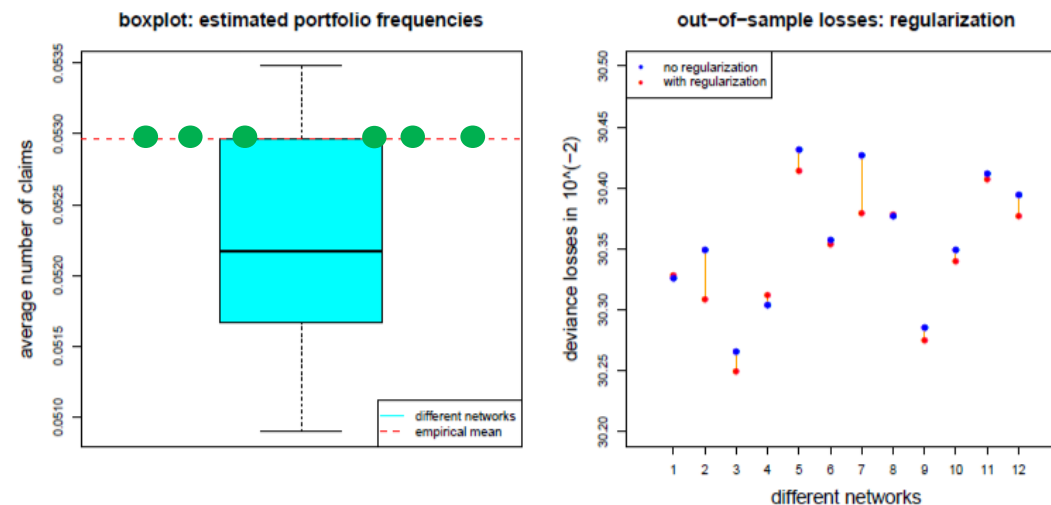
² Paper(s): <https://doi.org/10.1017/asb.2018.42>; https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3320525

3 – Portfolio bias in neural networks³

Let us examine various pricing models:

- The same price μ_0 (empirical portfolio average) for all policyholders provides the empirical **portfolio premium** ϕ_{GLM} .
- Using a GLM, the price gets differentiated between the policyholders according to their risk characteristics (=features). Overall, the sum of all predicted prices μ_i gives again the empirical portfolio premium ϕ_{GLM} .
- Mathematically, GLM provide unbiased estimates on a portfolio level under the canonical link.
- Using a neural network for determining the individual prices μ_i provides price differentiation, but the neural network provides a portfolio average ϕ_{NN} which is different than ϕ_{GLM} .
- Is this an issue? YES, **network calibrations have a bias** and one needs to correct for these biases, the insurance company does not earn the price it needs to cover its liabilities!

- The neural networks show a bias (mainly negative) whereas the green dots illustrate the level of various GLMs (●).

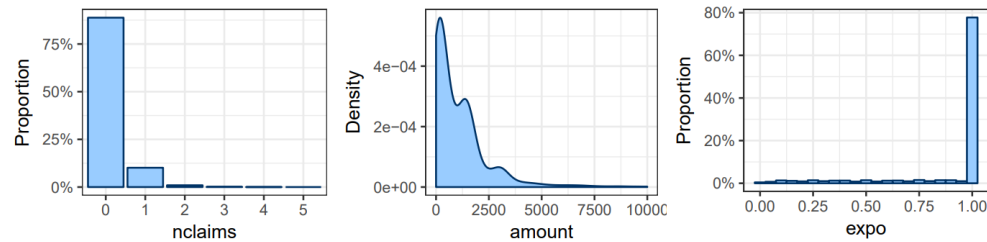


- Reason for the bias of neural networks: **early stopping criteria**.
- Solution: If we work in a GLM with canonical link function, this can simply be achieved by an additional MLE step using the **neuron activations in the last hidden layer as new covariates in the GLM**.

4 – Random forest for insurance data⁴



- Typical insurance data (MTPL from the CASdatasets R package):



- The plots are characteristic for (non-life) insurance data: highly unbalanced count data with excess zeros (left) and varying exposure on the frequency side (right) combined with scarce, but potentially **long- or even heavy-tailed continuous data** on the severity side (middle).
- The default random forest implementation in R (e.g. randomForest, ranger) or Python (e.g. sklearn) are based on the standard squared error loss function.
- The squared error loss function is not necessarily a good choice when modeling integer-valued frequency data or right-skewed severity data.

- The loss function used in the algorithm needs to be adjusted **such that the specific characteristics of insurance data are carefully considered.**
- Claim frequency modeling involves count data, typically assumed to be **Poisson** distributed. Therefore, an appropriate loss function is the Poisson deviance.
- **The exposure needs to be taken into account** in the expected number of claims. The Poisson deviance loss function can account for different policy durations.
- Right-skewed and long-tailed severity data is typically assumed to be **gamma** or **log-normally** distributed.
- What is the issue using gaussian-based random forest? → The tails of the distribution are not modelled accurately.
- Use of the **distRforest** R package.

⁴ Paper(s): <https://arxiv.org/abs/1904.10890> ; <https://github.com/henckr/distRforest>

Conclusions

Neural networks and random forests may substantially improve classical insurance risk models, if appropriately applied.



Embeddings of categorical features reduce the neural network size and allow for visualisations of the categorical feature levels in a low-dimensional space.



Neural networks need to be corrected for its bias to determine the correct technical price. The bias stems from the using early stopping criteria.



CANN provide the framework for extending the GLM's, allowing to improve the accuracy of the model as well as providing a framework to **assess the uncertainties**.



Random forest (and also neural network) loss function needs to be aligned with the characteristics of insurance data.

And yet, a complex and very well calibrated GLM may still be as good as an advanced machine learning model in terms of accuracy.

Closing

Visit

www.actuarialdatascience.org

strategy documents;
Tutorials with article, data and code;
and much more!

Acknowledgements

Current members working group:

- Andrea Ferrario
- Anja Friedrich
- Frank Genheimer
- Thomas Hull
- Christian Lorentzen
- David Lüthi
- Michael Mayer
- Daniel Meier
- Jürg Schelldorfer
- Alessandro Torre
- Frank Weber
- Mario Wüthrich

Contributors:

- Roger Hämmerli
- Mara Nägelin
- Alexander Noll
- Simon Renzmann
- Ron Richman
- Robert Salzmann

Institutions:

- [Swiss Association of Actuaries \(SAA\)](#)
- [RiskLab at ETH Zurich](#)
- [MobiLab for Analytics at ETH Zurich](#)

Companies:

- [Swiss Re](#)