Recent Achievements and Perspectives in Actuarial Data Science

Risk Day, ETH Zurich

13th September 2019

Dr. Jürg Schelldorfer, Actuary SAA

Senior Analytics Professional, 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.

Machine Learning in the insurance industry

Dr. Tobias Büttner, Head of Claims, Munich Re, mentioned the following[1]:

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.**

Implications of the use of Machine Learning (ML) in insurance:

- **ML can affects operations**, which impact the data actuaries use (i.e. claims, underwritten risks,…)
- **ML can affect** the underyling risks
- ML can be used to strenghten the core skills
- **Automation** (not necessarily ML) can help to improve efficiency

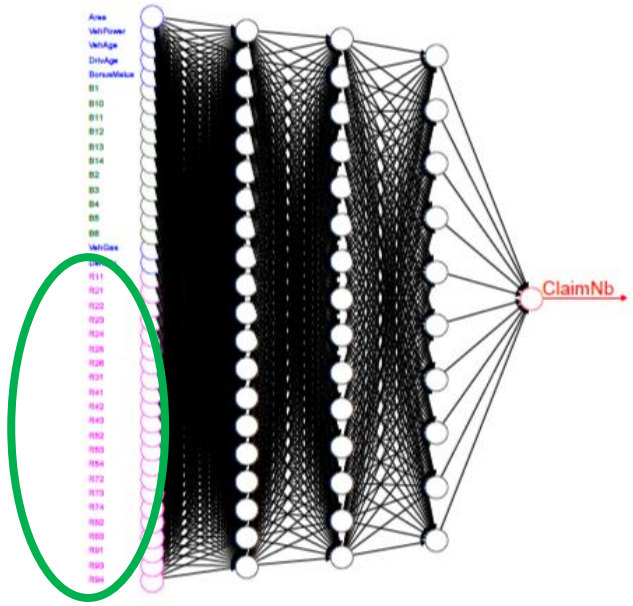
[1] SZ-Fachkonferenz: KI und Data Analytics in der Versicherungsbranche; Data Analytics im Management von Großschäden, Büttner T. (2019), Munich Re

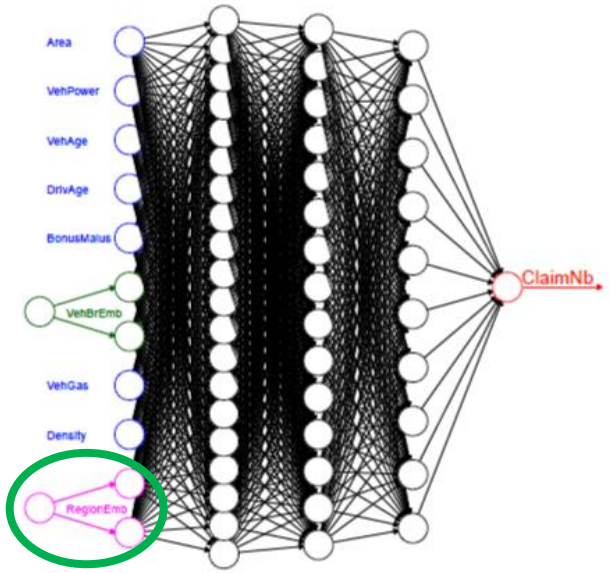
Table of Content

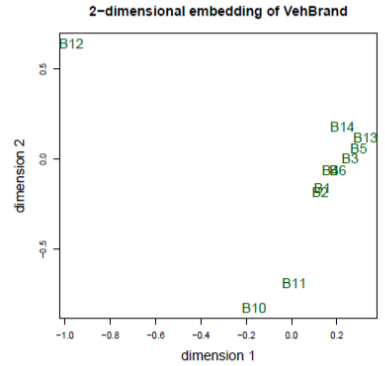
Recent achievements

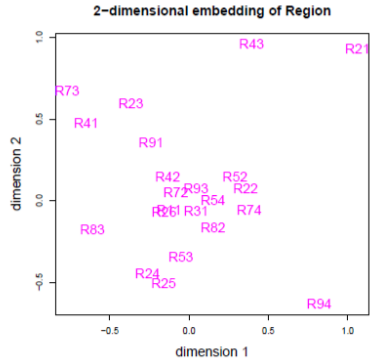
1 – Factor embeddings in neural networks[1]

- In insurance pricing, factor variables (i.e. vehicle brand, region,…) consist of many levels and are often encoded as dummy variables (or one-hot encoding), i.e. the levels are orthogonal in the feature space.
- With neural networks, we use (factor) embeddings which make the fitting of neural networks with many factors and levels feasable and natural.



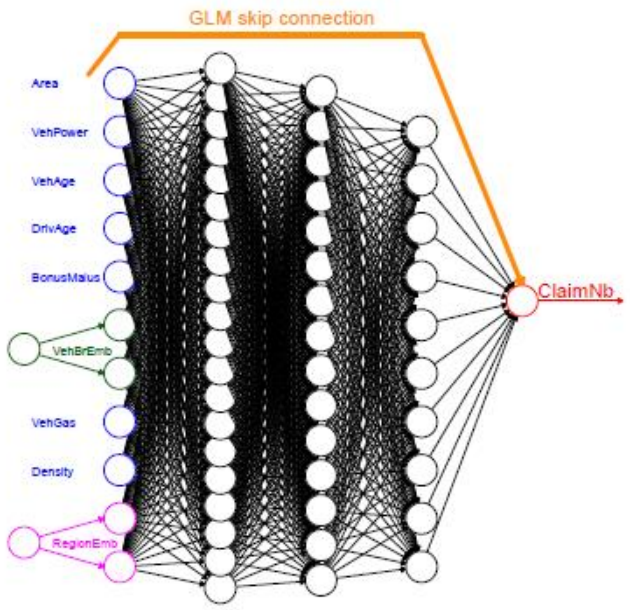






[1] Paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3320525

2 – Combined Actuarial Neural Networks (CANN)[1]



- Linear predictor with regression parameter $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_q)^\top \in \mathbb{R}^{q+1}$

$$\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^q \quad \mapsto \quad \theta^{\mathrm{GLM}}(\boldsymbol{x}; \boldsymbol{\beta}) = \langle \boldsymbol{\beta}, \widetilde{\boldsymbol{x}} \rangle \in \boldsymbol{\Theta}.$$

▷ Feature pre-processing is done by the actuary/statistician.

- Choose network of depth $d \in \mathbb{N}$ with network parameter $\boldsymbol{w} = (W_{1:d}, \boldsymbol{w}_{d+1})$

$$\boldsymbol{z} \in \mathbb{R}^{q_d} \quad \mapsto \quad \theta^{\mathrm{NN}}(\boldsymbol{z}; \boldsymbol{w}_{d+1}) = \langle \boldsymbol{w}_{d+1}, \widetilde{\boldsymbol{z}} \rangle \in \boldsymbol{\Theta},$$

with neural network function (feature pre-processing $\boldsymbol{x} \mapsto \boldsymbol{z}$)

- Choose linear predictor with parameter $(\boldsymbol{\beta}, \boldsymbol{w}) = (\boldsymbol{\beta}, W_{1:d}, \boldsymbol{w}_{d+1})$

$$\boldsymbol{x} \;\mapsto\; \theta^{\mathrm{CANN}}(\boldsymbol{x}; \boldsymbol{\beta}, \boldsymbol{w}) \;=\; \langle \boldsymbol{\beta}, \widetilde{\boldsymbol{x}} \rangle + \left\langle \boldsymbol{w}_{d+1}, \widetilde{\boldsymbol{z}}^{(d:1)}(\boldsymbol{x}) \right\rangle.$$

Advantages:

- Extension of GLM
- GLM as starting point for optimization
- Enables uncertainty quantification

[1] Paper: https://doi.org/10.1017/asb.2018.42

3 – Portfolio bias in neural networks[1]

GLM provide unbiased estimates on a portfolio level, and the GLM provides exactly the same unbiased estimated portfolio average as the homogeneous model.

Corollary 2.3. *Assume that true model is given by (2.3). The MLE of the GLM provides an unbiased portfolio average, that is,*

$$\mathbb{E}\left[\bar{\mu}^{\mathrm{GLM}}\right] = \bar{\mu}^{\star}, \qquad \textit{with uncertainty} \quad \mathrm{Var}\left(\bar{\mu}^{\mathrm{GLM}}\right) = \frac{\phi^{\star}}{\left(\sum_{i=1}^{n} w_i\right)^2} \sum_{i=1}^{n} w_i (b^{\star})''(\theta_i^{\star}).$$

Due to early stopping in neural networks model calibration, the model has a bias on the portfolio level!

Extract from an example for claims frequencies:

Figure 1.2. Remark that the evaluation of (3.5) requires knowledge of the true means $\mu_i^{\star}$ which are available in our special set-up.

	# param.	in-sample loss $\overline{\mathcal{L}}_{\mathcal{D}}(\boldsymbol{\mu})$	estimation error $\mathcal{E}_{\mu^{\star}}(\boldsymbol{\mu})$	portfolio average $\bar{\mu}$
(a) true model $\mu_i^{\star}$		27.7278	0.0000	10.1991%
(b) homogeneous model $\widehat{\mu}_i^{\mathrm{hom}} = \bar{\mu}^{\mathrm{hom}}$	1	29.1065	1.3439	10.2691%
(c) GLM $\widehat{\mu}_i^{\mathrm{GLM}}$	57	28.1282	0.4137	10.2691%
(d1) neural network $\widehat{\mu}_i^{\mathrm{NN}}$ run no. 1	780	27.7204	0.1566	10.2973%
(d2) neural network $\widehat{\mu}_i^{\mathrm{NN}}$ run no. 2	780	27.7484	0.1795	10.0661%
(d3) neural network $\widehat{\mu}_i^{\mathrm{NN}}$ run no. 3	780	27.7621	0.1669	10.0605%

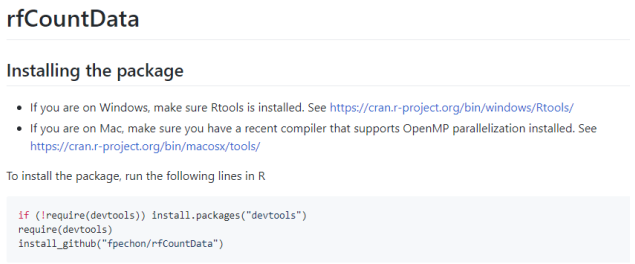
Remedies are proposed in the corresponding papers.

[1] Paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3347177

4 – Random Forest and Boosting

Achievements:

- *rfCountData*: random forest for Poisson distribution (GitHub)
- Review of most relevant boosting algorithm (AdaBoost, LogitBoost, XGBoost)[1]



Perspectives:

- Usage of random forest for claims severities (L2 is not a good loss function) and total loss amounts?
- *rfSeverityData* package?
- How to make random forest (better) interpretable?
- Are random forest / boosting appropriate for uncertainty quantification?

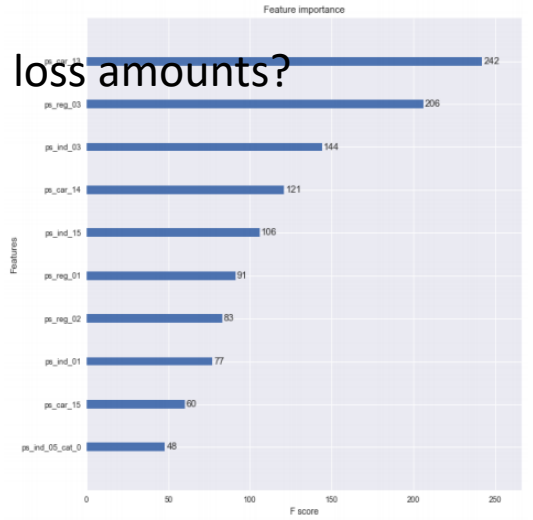


Figure 2: `XGBoost` best configuration in run III.: feature importance

[1] Paper: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3402687

Perspectives

(Our) Topics in Actuarial Data Science

We have written the following six tutorials:

1. French Motor Third-Party Liability Claims: Introduction, boosting and neural networks for P&C Pricing
2. Insights from Inside Neural Networks: Guidance how to fit neural networks for insurance data
3. Nesting Classical Actuarial Models into Neural Networks: Embedding of GLM's into neural networks
4. On Boosting: Theory and Applications: Boosting and its variant illustrated with a P&C Kaggle dataset
5. Unsupervised Learning: What is a Sports Car?: Unsupervised learning techniques applied in P&C
6. Lee and Carter go Machine Learning: Recurrent Neural Networks: LSTM NN applied to mortality forecasting

We are working on the following:

- Natural Language Processing and RNN's
- Segmentation using decision trees
- Mortality forecasting, Part II

Further topics and ideas:

- Missing data and data imputation
- Dissimilaritiy measures for categorical variables
- Convolutional Neural Networks and images
- Explainability / Interpretability of machine learning models
- Graphical Models / Causality?
- GAN?
- Performance measures and visualizations?
- Spatial modeling and random (Gaussian) fields?

Selected L&H business application

(including my personal biases)

Network-based approach to medical health used for underwriting[1,2].

Individual-based Mortality Forecasting[3]

1. Swiss Re, Understanding medical risk: a network-based approach (Link)
2. SZ-Fachkonferenz: KI und Data Analytics in der Versicherungsbranche; Expore your health, schnell und smart durch die Gesundheitsfragen, Dannenberg T. (2019), RISK-CONSULTING Prof. Dr. Weyer GmbH
3. Euroforum: Rethinking Insurance; Big Data – Mehrwerte durch Data Analytics generieren, Caro G. (2019), Swiss Re

Selected P&C business application

(including my personal biases)

Behavioural and situational data for the vessels in marine insurance[1].

Satelitte imagineries in agroculture insurance[1].

There is a move from..

- …pure claims modeling to..
- …claims + behavioural + lapse modelling[2]

1. Sigma 4/2019: Advanced Analytics: unlocking new frontiers in P&C insurance, Swiss Re, 2019
2. Driving data for automobile insurance: will telematics change ratemaking? , Monserrat Guillén, SAV Mitgliederversammlung 2019, Lucerne

Non-quantitative aspects

Model Risk Management

‘Machine Decisions’: Governance of AI and Big Data Analytics; CRO Forum (2019)

- «…that model governance techniques and frameworks that exist today **do not need to be fundamentally altered, but can be enhanced and adjusted** to meet the evolving needs of complex tools and machine learning developments»
- Model Management Framework
- Ethical Framework

Believing the Bot - Model Risk in the Era of Deep Learning

Ronald Richman* Nicolai von Rummell† Mario V. Wüthrich‡

Version of August 29, 2019

Abstract

Deep Learning models are currently being introduced into business processes to support decision-making in insurance companies. At the same time model risk is recognized as an increasingly relevant field within the management of operational risk that tries to mitigate the risk of poor business decisions because of flawed models or inappropriate model use. In this paper we try to determine how Deep Learning models are different from established actuarial models currently in use in insurance companies and how these differences might necessitate changes in the model risk management framework. We analyse operational risk in the development and implementation of Deep Learning models using examples from pricing and mortality forecasting to illustrate specific model risks and controls to mitigate those risks. We discuss changes in model governance and the role that model risk managers could play in providing assurance on the appropriate use of Deep Learning models.

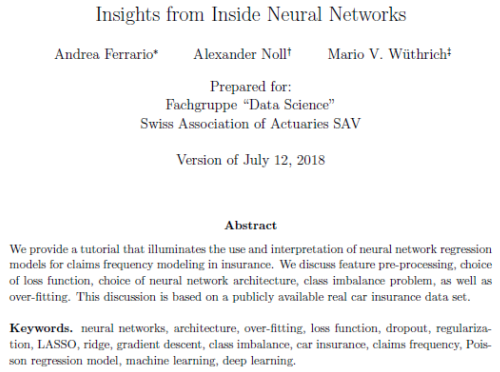
Keywords. Deep learning, Model Risk, Pricing, Mortality Forecasting, Insurance Modelling

1 Introduction

Deep learning refers to a modern approach to designing and fitting neural networks that recently has achieved state of the art results on machine learning problems in computer vision, natural language processing, machine translation and speech recognition, and has become the main avenue for solving unstructured data problems [1, 2]. In addition to unstructured data, deep learning approaches have produced promising results on structured data problems [3], as well as time series forecasting [4]. Modern neural networks are generally characterized by specialized architectures that are adapted to domain-specific problems, as well as by the depth of the networks, meaning to say, that these networks are composed of multiple layers of non-linear functions. Recently, deep learning techniques have been applied to problems within actuarial science such as pricing, reserving, analysis of telematics data and mortality forecasting. For a recent review of these applications, see [5]. The benefits of applying deep learning to actuarial

Model Risk for NN

Guideline for fitting a NN for Pricing



Insights from Inside Neural Networks

Andrea Ferrario* Alexander Noll† Mario V. Wüthrich‡

Prepared for:
Fachgruppe “Data Science”
Swiss Association of Actuaries SAV

Version of July 12, 2018

Abstract

We provide a tutorial that illuminates the use and interpretation of neural network regression models for claims frequency modeling in insurance. We discuss feature pre-processing, choice of loss function, choice of neural network architecture, class imbalance problem, as well as over-fitting. This discussion is based on a publicly available real car insurance data set.

Keywords. neural networks, architecture, over-fitting, loss function, dropout, regularization, LASSO, ridge, gradient descent, class imbalance, car insurance, claims frequency, Poisson regression model, machine learning, deep learning.

Ethics and Company-internal Training Guide

Ethics:

Publications on Ethics in ML/AI applications:

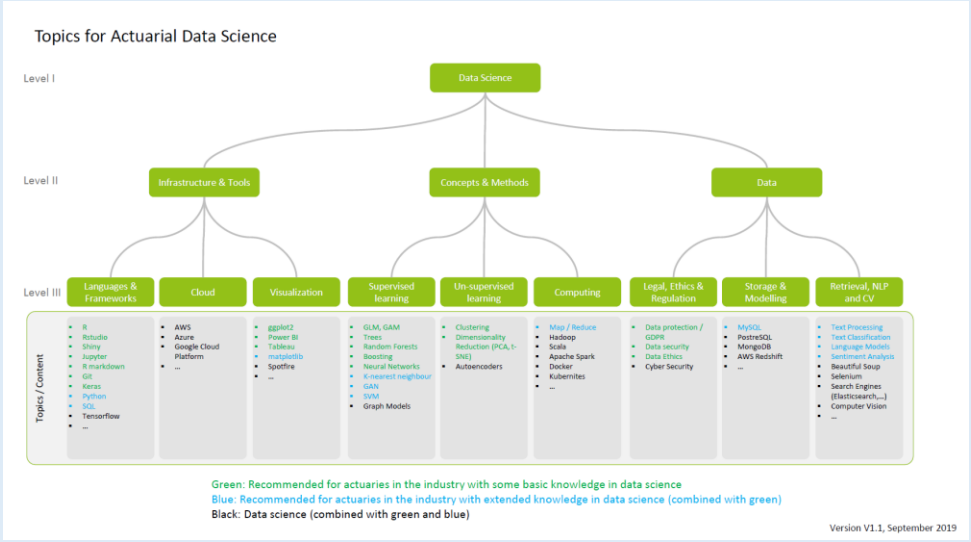
- Ethical Codex for Data-Based Value Creation, Swiss Alliance for Data-Intensive Services, 2019
- Ethics Guidelines for Trustworthy AI, European Commission, 2019
- Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's financial industry, MAS, 2019
- Ethically Aligned Design, IEEE, 2019

These papers raise some questions w.r.t. to the role and responsiblity of the actuaries:

- Should an actuary be fullfilling the relevant ethical codex for a Data Scientist? Or is he already doing it?
- What should be expected from an actuary w.r.t. to ethics?

Company-internal Training Guide

For already fully qualified actuaries in industry (demand from smaller actuarial associations and companies) we have summarized the topics to start with ADS…



Summary

Conclusions

- Statistical learning methods and neural networks allow to fit dependency structures naturally beyond the (currently used) GLM.
- 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.
- Model risk management needs to be addressed carefully for machine learning models
- There are many business challenges ahead which require machine learning skills.

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

Visit

www.actuarialdatascience.org

Article, data and code of the tutorials

References to literature

Acknowledgements

People:

- All members of the SAA working group
- Dr. Alexander Noll
- Dr. Simon Renzmann
- Ron Richman

Insititutions:

- Swiss Association of Actuaries (SAA)
- RiskLab at ETH Zurich
- MobiLab for Analytics at ETH Zurich

Companies:

- Swiss Re

References

- www.actuarialdatascience.org
- Nesting Classical Actuarial Models into Neural Networks, Schelldorfer J. and Wüthrich M.V. (2019), SAA
- Editorial: Yes, we CANN!, Wüthrich, M.V., Merz, M. (2019). ASTIN Bulletin 49/1
- Bias Regularization in Neural Network Models for General Insurance Pricing, Wüthrich M.V. (2019), SSRN
- rfCountData, Pechon F. (2018), GitHub
- On Boosting: Theory and Applications, Ferrario A. and Hämmerli R. (2019), SAA
- Understanding medical risk: a network-based approach, Caro G. (2019), Swiss Re
- SZ-Fachkonferenz: KI und Data Analytics in der Versicherungsbranche; Data Analytics im Management von Großschäden, Büttner T. (2019), Munich Re
- SZ-Fachkonferenz: KI und Data Analytics in der Versicherungsbranche; Expore your health, schnell und smart durch die Gesundheitsfragen, Dannenberg T. (2019), RISK-CONSULTING Prof. Dr. Weyer GmbH
- Euroforum: Rethinking Insurance; Big Data – Mehrwerte durch Data Analytics generieren, Caro G. (2019), Swiss Re
- Sigma 4/2019: Advanced Analytics: unlocking new frontiers in P&C insurance, Swiss Re (2019)
- Driving data for automobile insurance: will telematics change ratemaking?, Monserrat Guillen (2019)
- 'Machine Decisions': Governance of AI and Big Data Analytics, CRO Forum (2019)
- Believing the Bot – Model Risk in the Era of Deep Learning, Richman R., von Rummell N, Wüthrich M.V. (2019), SSRN
- Insights from Inside Neural Networks, Ferrario A., Noll A., Wüthrich M.V. (2018), SSRN

Appendix

ADS basics: Articles and repositories

The following articles/repositories are fundamental for entering the topic of Actuarial Data Science (ADS):

- Data Analytics for Non-Life Insurance Pricing, ETH Zurich, M.V. Wüthrich and C. Buser
- AI in Actuarial Science, R. Richman, SSRN, 2018
- ADS Tutorials, SAA, 2018-present
- Insurance Analytics – A Primer, International Summer School of the Swiss Association of Actuaries, 2018
- Insurance Data Science: Use and Value of Unusual Data, International Summer School of the Swiss Association of Actuaries, 2019

And do not forget the fundamentals of Statistics vs. Machine Learning:

- Statistical Modeling: The Two Cultures. L. Breimann, Statistical Science 16/3, 199-215, 2001
- To explain or to Predict?, G. Shmueli, Statistical Science 25/3, 289-310, 2010

ADS basics: R packages[1,2]

ML meta packages:
- *caret*
- *mlr*

Neural Networks:
- *keras*

Machine/Statistical Learning excl. NN
- *rpart*
- *ranger*, *randomForest*, *rfCountData*
- *xgboost*, *gbm*
- *cluster*, *clusterR*, *tsne*, *umap*, *kohonen*
- *glmnet*

data:
- *tidyverse*
- *data.table*

Visualisations:
- *ggplot2*
- *DataExplorer*
- *esquisse*

Interpretability:
- *iml*

Others:
- *Rmarkdown*
- *Rshiny*

Insurance data:
- *CASdatasets*

[1] R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
[2] CRAN Task View: Machine Learning & Statistical Learning, T. Hothorn, 2019

Outlook and Call for Action

Outlook Working Group:

- Additional tutorials
- Offering an ADS block course
- Dedicated SAA working group on ethics and providing a structure for Data Scientists to become member of the SAA.

Call for Action:

- Insurance analytics and actuarial data science should be strengthened at actuarial education and research institutions.
- Foster research and developments in actuarial data science between companies and universities.
- Synthetic data generation (Simulation Machine, GAN?,…) techniques to allow collaborations with research institutions and actuarial associations.
- How to generate publicly available and yet well calibrated actuarial data sets for machine learning?