

DAV/DGVFM  
**Jahrestagung**  
2026

*Friedrich Loser, Techniker Krankenkasse*

---

# Maschinelles Lernen und Schadenreservierung

---

ASTIN Fachgruppe, 29. April 2026

## Agenda

- Schadenreservierung mit Methoden des Maschinellen Lernens: Weltreise
- Erzeugung und Vergleich synthetischer Schadendaten
- Sonderthemen: Saisonale Strukturen; Abwicklungsgeschwindigkeit

## “Machine Learning in Reserving”: Weltreise

- Berlin
  - USA
  - Südafrika
  - Schweiz
  - Schweden
  - Australien
  - UK (IFoA MLR Working Group)
- 

**“Erlebnis”:  
ML-Verfahren  
führen zu  
deutlich  
genaueren  
Prognosen**

## ICA 2018 (Berlin): Best Paper Award

### M.V. Wüthrich „Neural networks applied to chain-ladder reserving“

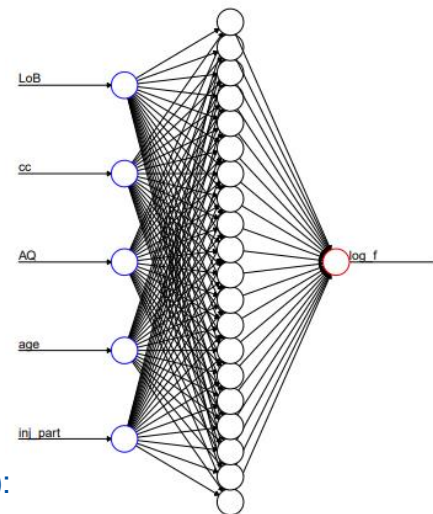
Artikel (Preprint 2017, EAJ 2018)

- Startpunkt: Macks Chain–Ladder Modell
- Künstliche neuronale Netze „in aktuarieller Sprache“
- Enthält R-Code für **simulation machine** (Datengenerierung) und kleines künstliches neuronales Netz

Vortrag bei ICA 2018 (Berlin), Preisträger

- Vorstellung der simulierten Daten (auf Basis von 10 Mio. Schweizer Unfallschäden):
  - Cumulative claim  $C_{i,j}$  aggregates many different claims (of different types).
  - Available information in our example (accident insurance):
    - line of business  $LoB \in \{1, \dots, 4\}$
    - claims code  $cc \in \{1, \dots, 53\}$  Berufs-/Gefahrengruppe
    - accident quarter  $AQ \in \{1, \dots, 4\}$
    - age of injured  $age \in \{15, \dots, 70\}$
    - injured body part  $inj\_part \in \{10, \dots, 99\}$
  - Each claim has covariate information  $\mathbf{x} = (LoB, cc, AQ, age, inj\_part) \in \mathcal{X}$ .

simulation  
machine



Fortsetzung  
ab Folie 7

## K. Kuo (USA): “DeepTriangle: A Deep Learning Approach to Loss Reserving”

Quelle und Code:  
arxiv-preprint 2018;  
Risks 2019, 7(3), 97;

### Modell:

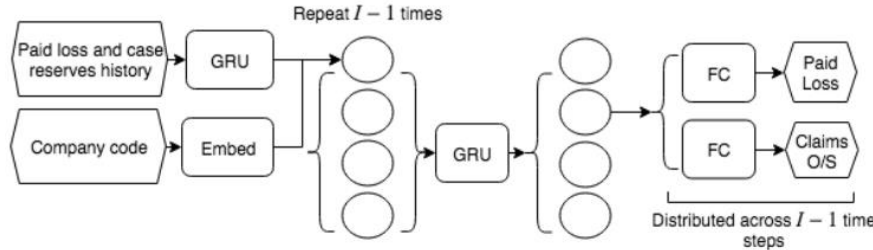


Figure 2. DeepTriangle architecture. *Embed* denotes embedding layer, *GRU* denotes gated recurrent unit, *FC* denotes fully connected layer.

### „DeepTriangle“:

- Rekurrentes neuronales Netz mit Embeddings (zeitgemäße Deep-Learning-Architektur)

### Güte:

Table 1. Performance comparison of various models. DeepTriangle and AutoML are abbreviated to DT and ML, respectively. The best metric for each line of business is in bold.

Line of Business	Mack	ODP	CIT	LIT	ML	DT
<b>MAPE</b>						
Commercial Auto	0.060	0.217	0.052	0.052	0.068	<b>0.043</b>
Other Liability	0.134	0.223	0.165	0.152	0.142	<b>0.109</b>
Private Passenger Auto	0.038	0.039	0.038	0.040	0.036	<b>0.025</b>
Workers' Compensation	0.053	0.105	0.054	0.054	0.067	<b>0.046</b>
<b>RMSPE</b>						
Commercial Auto	0.080	0.822	0.076	0.074	0.096	<b>0.057</b>
Other Liability	0.202	0.477	0.220	0.209	0.181	<b>0.150</b>
Private Passenger Auto	0.061	0.063	0.057	0.060	0.059	<b>0.039</b>
Workers' Compensation	0.079	0.368	0.080	0.080	0.099	<b>0.067</b>

### Ergebnis:

- Mit Abstand genaueste Prognosen

### Nachteile:

- Undurchsichtig („black box“)
- Unverständnis (AI-Engineer vs Actuary)

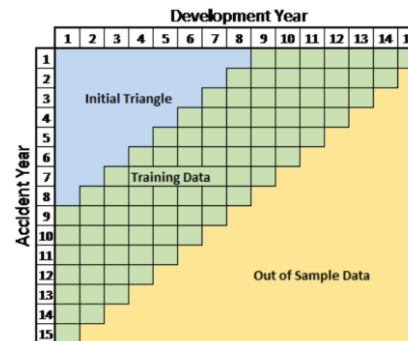
## Balona & Richman (Südafrika, 2020): The Actuary and IBNR Techniques. ML approach

### ML approach to selecting method and parameters

- Partition data into “training” and “testing” subsets
- Determine ability of each model and parameter set to predict unseen data using the “test” subset
- Choose the model and parameter set that results in the best predictive performance on the unseen “test” subset

### Apply this to traditional IBNR reserving

1. Define a set of IBNR calculation methodologies (such as CL, BF, CC with varying parameters)
2. Fit these methodologies on sub-triangles starting from a small initial triangle, and then increasing the triangle by one calendar year until the end of the available data
3. At each sub-triangle, calculate the performance of the methodology using some performance metric (AvE/CDR)
4. Select the methodology that results in the best score across all sub triangles



### Bringing it all together

1. Select a reasonably sized **initial triangle**
2. Select several of the most recent calendar periods as the **training set**
3. For each reserving methodology, **M**, from a collection of possible methodologies:
  - a. Apply the reserving methodology to the **initial triangle**
  - b. Calculate the performance metric on the first diagonal of the **training set**
  - c. Include the **first** diagonal of the **training set** in the **initial triangle** and apply **M**
  - d. Calculate the performance metric against the second diagonal of the **training set**
  - e. Repeat until all diagonals of the **training set** are exhausted
  - f. Calculate the average performance metric for **M** across all diagonals in the **training set**
4. Select **M** that achieves the best performance metric



## A. Gabrielli (Zürich), “Claims Reserving and Neural Networks”

[Dissertation](#) ETHZ 2020. Co-Autoren der Paper A-C: MV Wüthrich; C: R Richman

simulation  
machine

„In **Paper A** konstruieren wir auf der Basis von neuronalen Netzwerken eine stochastische Simulationsmaschine, die Einzelschadenfälle generiert. Diese synthetischen Einzelschadenfälle ermöglichen, klassische Schadenreservierungsmodelle zu evaluieren, sowie auch neue Schadenreservierungsmethoden zu entwickeln.“

back-testing  
chain-ladder

„In **Paper B** verwenden wir die stochastische Simulationsmaschine von Paper A, um das Chain-Ladder Modell zu evaluieren. Diese Studie gibt uns ein Gefühl für das Verhalten des Chain-Ladder Modells, insbesondere bei wachsender Portfoliogröße.“

NN-embedding  
ODP-model

**Paper C:** Over-Dispersed Poisson Modell wird in neuronales Netz eingebettet (siehe S. 8)

NN-boosted  
double ODP

**Paper D:** Zusätzlich Schadenzahlen in deutlich komplexerer Architektur berücksichtigt (siehe S. 9)

individual claims  
reserving model

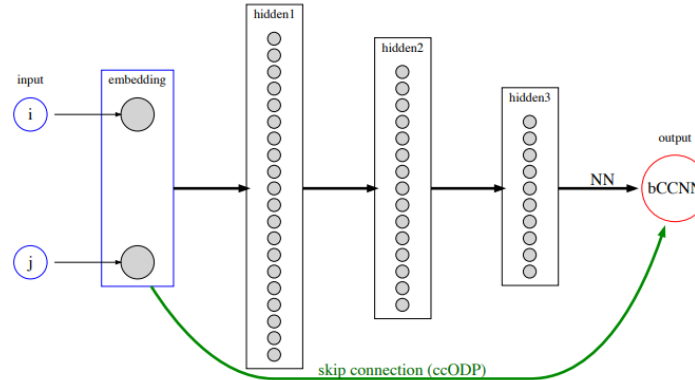
„In **Paper E** entwickeln wir ein Schadenreservierungsmodell für gemeldete Einzelschäden. Mit Hilfe eines einzigen neuronalen Netzwerks schätzen wir die erwarteten zukünftigen Zahlungen für die gemeldeten Einzelschäden auf der Basis von schadenspezifischen Informationen.“

## A. Gabrielli, Paper C (NN-embedding ODP-model)

### NN-embedding ODP-model

„In **Paper C** betten wir das Over-Dispersed Poisson Modell in ein neuronales Netzwerk ein. Dabei verwenden wir das Over-Dispersed Poisson Modell als Startpunkt für die Kalibrierung des neuronalen Netzwerks. Mit diesem verschachtelten Modell können wir zusätzliche Struktur in den Daten jenseits des klassischen Schadenreservierungsmodells erkunden.“

#### Architektur:



**Modell (Code) und Daten werden verwendet und mit “Gradient Boosting Machines” (GBM) u.a. verglichen in:**

M. Lindholm et al., "Machine Learning, Regression Models, and Prediction of Claims Reserves", Casualty Actuarial Society E-Forum, Summer 2020.

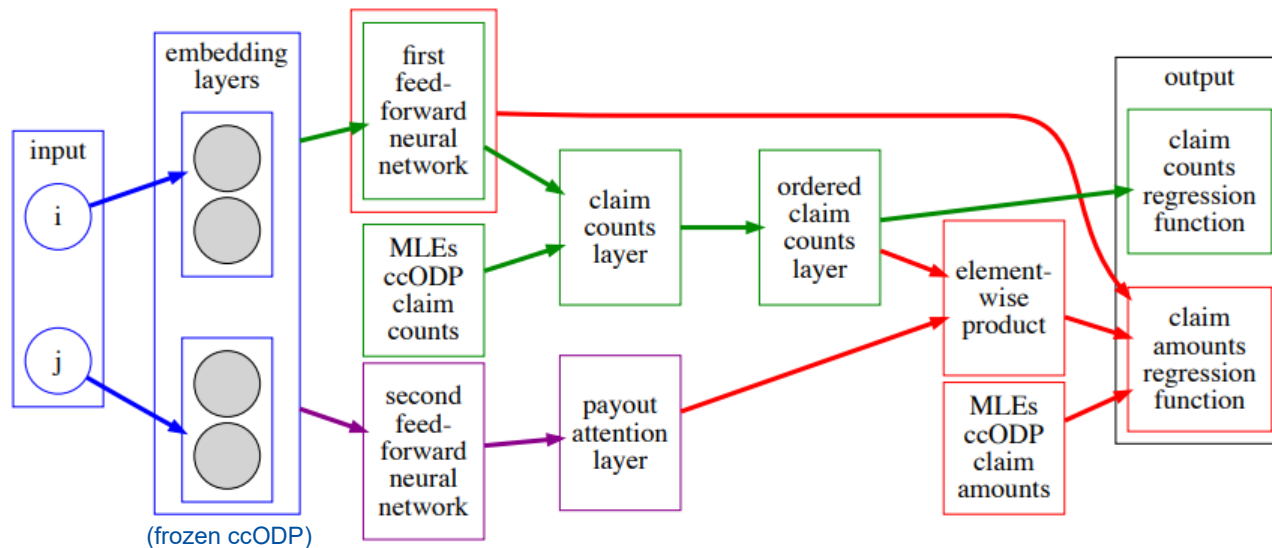
Vilma Härkönen, "On Claims Reserving with Machine Learning Techniques", 2021

## A. Gabrielli, Paper D (NN-boosted double ODP)

### NN-boosted double ODP

„In **Paper D** erweitern wir die Einbettung des Over-Dispersed Poisson Modells für Schadenhöhen von Paper C zu einer gemeinsamen Einbettung von je einem Over-Dispersed Poisson Modell für Schadenhöhen sowie für Schadenanzahlen. Ein solches Modell ermöglicht, die zusätzliche Information der Schadenanzahlen zu benutzen.“

#### Architektur:



Originalcode auf GitHub verfügbar.

Notebook zum ausprobieren: <https://www.kaggle.com/code/floser/claims-reserving-neural-nets-and-payout-attention>

## V. Härkönen (Stockholm): “On Claims Reserving with Machine Learning Techniques”

- [Masteruppsats 2021:4](#), Försäkringsmatematik, Stockholms universitet, Juni 2021 (Supervisor: M. Lindholm)
- SCOR Actuarial Prize 2024, Presentation: <https://actview.com/videos/on-claims-reserving-with-machine-learning-techniques-4554> (Mai 2025) bzw. [https://www.institutdesactuaire.com/global/gene/link.php?doc\\_id=19581&fg=1](https://www.institutdesactuaire.com/global/gene/link.php?doc_id=19581&fg=1)

### „Predicted Reserves Simulation“

Model	Type	LoB					
		1	2	3	4	5	6
True	Reserve	39 689	37 037	16 878	71 630	72 548	31 117
CL	Reserve	38 569	35 460	15 692	67 574	70 166	29 409
CL	Bias %	-2.82	-4.26	-7.02	-5.66	-3.28	-5.49
ODP	Reserve	38 308	35 151	15 452	67 055	69 470	29 115
ODP	Bias %	-3.48	-5.10	-8.45	-6.39	-4.24	-6.44
GBM	Reserve	39 697	37 229	16 367	72 667	71 433	32 114
GBM	Bias %	<b>0.02</b>	<b>0.52</b>	<b>-3.03</b>	1.44	-1.54	<b>3.20</b>
C → Simple NN	Reserve	41 268	34 779	15 356	71 682	70 649	29 336
Simple NN	Bias %	3.98	-6.10	-9.02	<b>0.07</b>	-2.62	-5.73
Simple NN TW	Reserve	41 005	39 135	15 949	75 646	75 438	29 640
Simple NN TW	Bias %	3.31	5.66	-5.50	5.54	3.98	-4.75
D → Double NN	Reserve	40 029	35 959	15 686	69 509	72 512	30 047
Double NN	Bias %	0.85	-2.91	-7.06	-2.96	<b>-0.05</b>	-3.44

### „Predicted Reserves Folksam“ (Schadenversicherer)

Model	Type	LoB		
		1	2	3
True	Reserve	734 200	135 241	486 714
CL	Reserve	401 572	131 799	375 972
CL	Bias %	-45.30	-2.55	-22.75
ODP	Reserve	459 873	141 439	374 627
ODP	Bias %	-37.36	4.58	-23.03
GBM	Reserve	783 878	145 219	394 719
GBM	Bias %	<b>6.77</b>	7.38	<b>-18.90</b>
Simple NN	Reserve	593 404	135 031	388 580
Simple NN	Bias %	-19.18	<b>-0.15</b>	-20.16
Double NN	Reserve	474 484	138 474	375 805
Double NN	Bias %	-35.37	2.39	-22.78

Einfache GBMs in 6 von 9 Fällen am genauesten

Siehe auch (Anhang) Master-Arbeit von C. Hunsiker an Kuo-Datensatz: XGBoost deutlich überlegen

Bestätigt durch eigene Erfahrungen mit GBMs (bei größeren Datenmengen)

**Ergebnis: Gradient Boosting (GBM) meist am genauesten. CL und ODP teils weit abgeschlagen.**

Siehe auch Simulationsergebnisse und R-Codes in „Machine Learning, Regression Models, and Prediction of Claims Reserves“ by Mathias Lindholm, Richard Verrall, Felix Wahl and Henning Zakrisson, Casualty Actuarial Society E-Forum, Summer 2020 <https://www.casact.org/sites/default/files/2021-08/Complete-Summer-2020.pdf> (R-Code p.37-40)

## Avanzi et al. (Australien): SynthETIC, SPLICE, ADLP & RL

### „SynthETIC: an individual insurance claim simulator with feature control” (2020)

- Transparenter, quelloffener Simulator mit “Feature-Strukturierung”
- Default: Generiert indiv. KFZ-Personenschadendaten

### „SPLICE: A Synthetic Paid Loss and Incurred Cost Experience Simulator” (2021)

- Erweiterung von SynthETIC, kann z.B. Rückstellungsentwicklung simulieren

### „Ensemble distributional forecasting for insurance loss reserving” (2024)

- Optimierung von linearen Modell-Ensembles (GLMs und GAMLSS)
- R-Package ADLP (Accident and Development period adjusted Linear Pools)

### „Reinforcement Learning for Micro-Level Claims Reserving” (2026, u.a. mit Richman & Wüthrich)

- Sehr erfolgreiche, komplexe Deep-Learning-Methode: Agent lernt durch Belohnungen/Strafen
- “the proposed Soft Actor–Critic implementation delivers competitive claim-level accuracy and strong aggregate OCL performance” (Outstanding Claim Liability)

Synthetische  
Einzel-  
schaden-  
daten

R-Package  
ADLP

RL-Paper!

## MLR WP: Website (IFoA)

### Machine Learning in Reserving Working Party

The General Insurance Machine Learning in Reserving Working Party (MLR WP) is a group of over 70 volunteers, bringing together a range of data scientists, actuaries and academics from around the globe.

When we started out in 2019, our premise was to find out why, whilst machine learning techniques are widespread in pricing, they are not being adopted 'on the ground' in reserving (certainly in the UK). Since then we have been working to help develop data science skills, and are looking at ways that machine learning can be incorporated into reserving practice.

“The Institute and Faculty of Actuaries (IFoA) is the UK’s only chartered professional body dedicated to educating, developing, and regulating actuaries based both in the UK and internationally. The IFoA regulates and represents over 34,000 members worldwide.”

**Machine Learning Journey**  
In our latest blog, Adam Stanley takes you on his journey to learning about using machine learning in reserving.

**MLRWP book**  
Check out our new [book](#)! It's a collection of most of our blog posts over the last 4 years, organised by topic.

**Looking for a particular topic? Check out:**

- Our [about](#)
- Our [publications](#)
- Our [blog](#)

**Learn more about the working party**  
Check out our [about](#) page for:

- information on the working party
- how to join us



We find we often get asked the same questions, so we have collated our answers to them here

[FAQs >](#)



### Foundations – Learning Materials

Our Foundations page is for those new to Machine Learning. Here you can find resources including:

- > A suggested roadmap of how to learn
- > Learning materials that we recommend
- > Sample notebooks in R and Python where methods are applied to reserving data sets.

[Foundations – Learning Materials >](#)



### Our 'Book'

With well over 30 blogs and articles we know it can be hard to find your way through all of the material on our site. We have compiled a book to make it easier to navigate.

It is still a work in progress, so bear with us, but hopefully it provides a clearer oversight of our output, which includes help and advice on learning ML, highlights some of the latest research, as well as offering thoughts on some of the practical considerations.

[Book >](#)

[Back To Top](#)

<https://ve.actuaries.org.uk/mod/page/view.php?id=52123>

<https://actuaries.org.uk/about-us/>

[MLR GitHub site >](#)

## MLR WP: Überblick und Auszug aus Ergebnissen

### Organisation:

Große internationale Arbeitsgruppe  
(> 70 Personen) des IFoA mit  
Fachexperten und Gästen

### Ergebnisse:

- Moderne Website
- FAQs
- Einige Vorträge
- **Blog-Post: Über 30 Artikel und Notebooks**  
(1-2 Autoren. Viel in R,  
zunehmend in Python)
- Online-Buch und weitere  
Lernmaterialien



### Beispiele

#### Reserving using machine learning - an advanced example in R (Part 3)

This article introduces the final of a series of three R notebooks. The notebooks provide R code to replicate the central scenario in the paper of Maximilien Baudry "NON-PARAMETRIC INDIVIDUAL CLAIM RESERVING IN INSURANCE".

Nigel Carpenter  
Nov 30, 2020 · 2 min read · 0 Comments



#### Getting to grips with GLM, GAM and XGBoost

Introduction Congratulations! You've just completed your introductory course in data science and analytics in R, and you're ready to start enriching your day-to-day actuarial work with new and exciting models.

Thomas Saliba  
Jul 5, 2021 · 3 min read · 0 Comments

#### A brief history of papers looking at using neural networks in reserving

This content was first presented by Kevin Kuo at the CAS Casualty Loss Reserving Seminar, the CIA Appointed Actuary Virtual Seminar and at GIRO in the Autumn of 2020.

Kevin Kuo  
Jul 26, 2021 · 7 min read · 0 Comments



#### Probabilistic Forecasting with Neural Networks Applied to Loss Reserving

Here we highlight the 2020 thesis by Muhammed Taheer Al-Mudafer, which applies an innovative approach to reserving using probabilistic forecasting and neural networks.

Sarah MacDonnell  
Sep 30, 2021 · 2 min read · 0 Comments

#### ML modelling on triangles - a Python example

Here we share a Python version of our previous R worked example on using ML modelling on triangles, which we originally presented at the 2021 IFoA Spring Conference

Jacky Poon, Grainne McGuire  
Nov 5, 2021 · 52 min read · 0 Comments



#### Practical Considerations Part 1: Time & Resource Limitations

Why is Machine Learning not commonly being adopted in the UK Reserving Market? There are many reasons why Machine Learning ("ML") techniques are currently not being adopted in the UK reserving market.

Isabelle Williams  
Apr 23, 2022 · 13 min read · 0 Comments

# A. Stanley (IFoA): „My Machine Learning in Reserving Journey“

„An ML beginner recounts his first steps into using Machine Learning for reserving“

Auszüge aus  
[mlr-blog](#)

(1)  Machine Learning in Reserving Working Party 🔍

There are plenty of ways to simulate data, with many examples throughout the various blogs covered. For this exercise I have used [SPLICE](#). SPLICE is a synthetic paid loss and incurred cost experience simulator built in R. It is built upon its predecessor SynthETIC, with added capabilities on simulating case estimates. SynthETIC would be sufficient for this exercise; however, it makes more sense to use and become familiar with the most up to date method. Hence using SPLICE to generate a dataset is the most appropriate method. SPLICE

On this page

- Introduction
- Deciding Data**
- Reserving with GLMs
- Reserving with LASSO
- Machine Learning Triangles

Here is the code to generate the data. Since it takes a long time to run, I have saved the data set. It is available [here](#) as an RDS file.

```
# Generate the dataset
datasets <- generate_data(
  n_claims_per_period = 2000, # expected claims per period parameter
  n_periods = 30, # number of accident and development periods
  complexity = 2, # scenario 2
  random_seed = 42)

saveRDS(datasets, file="_sim_data.RDS") # Save the dataset
```

(2)  Machine Learning in Reserving Working Party 🔍


## Machine Learning Triangles

The next [blog](#) brings together everything we've done so far, and compares a number of machine learning and non-machine learning methods:

- GLM
- LASSO
- Decision trees
- Random Forest
- XGBoost

On this page

- Introduction
- Deciding Data
- Reserving with GLMs
- Reserving with LASSO
- Machine Learning Triangles**
- Conclusion
- About the author
- Appendix

 Machine Learning in Reser... 🔍

regr.xgboost_25	6,848,350,784	110%
regr.rpart	6,504,165,134	105%
actual	6,221,167,904	100%
<b>regr.xgboost_500</b>	<b>6,212,036,569</b>	<b>100%</b>
regr.glm	6,110,647,241	98%
regr.ranger	5,971,603,530	96%
regr.cv_glmnet	5,454,727,465	88%

Ergebnis: XGBoost ist am genauesten (wenn man es richtig trainiert)

## Agenda

- Schadenreservierung mit Methoden des Maschinellen Lernens: Weltreise
- Erzeugung und Vergleich synthetischer Schadendaten
- Sonderthemen: Saisonale Strukturen; Abwicklungsgeschwindigkeit

## Erzeugung synthetischer Daten aus Echtdaten: Software

Ergebnisse der Arbeitsgruppe  
“Synthetisierung von Daten”  
des Ausschusses ADS/AI:

1. Vortrag auf Herbsttagung 2025  
[actuview.com](https://actuview.com) (Video + Foliensatz)

2. Notebooks (u. a. Software-Vergleich) →  
[github.com/DeutscheAktuarvereinigung/](https://github.com/DeutscheAktuarvereinigung/)

Anwendung auf Datensatz “freMTPL2freq“: Wird das sehr auffällige  
Segment\* in den synthetischen Daten korrekt abgebildet?

Dataset	Overall Claim Frequency	Segment Claim Frequency	Segment Size
Original	0.1	0.66	0.03
DataSynthesizer	0.8	1.05	0.01
SDV	0.0	0.00	0.06
synthpop R	0.1	0.60	0.03
Synthomat Python	0.1	0.59	0.03

### Beispiel: Anwendung von Synthpop auf Einzelschadendaten

Code für Daten-  
synthetisierung  
(vollständig):

```
library(synthpop)
df_synth <- syn(df_echtdaten)
write.csv2(df_synth$syn, "synth.csv")
```

Ergebnisdatei  
(Auszug):

	A	B	C	D	E	AA
1	AnfallJahr	AnfallMonat	NLM_0	NLM_1	NLM_2	NLM_24
2	2022	5	0	0	2045	0
3	2024	6	0	6215	0	0
4	2020	4	0	0	2716	0
5081	2023	7	0	0	0	1280

Aggregation zum  
Abwicklungsdreieck:  
Siehe Folgeseite

# Saisonalstruktur: Echte und synthetischen Daten führen zum gleichen Ergebnis

## Zahlungen (wachstumsbereinigt)

Für-Monat	N00	N01	N02	N03	N04	N05	N06	N07	N08	N09	N10	N11	N12	N13	N14	N15	N16	N17	N18
2024M01	1	148	113	29	26	1	3	15	13	1	0	0	0	0	1	0	0	0	0
2024M02	6	182	109	19	16	6	16	7	1	0	0	2	0	0	0	0	0	0	0
2024M03	5	168	205	32	7	26	19	2	3	7	4	0	0	0	0	0	4	0	0
2024M04	6	265	152	26	28	12	10	8	9	0	0	0	1	0	0	0	0	0	9
2024M05	7	237	200	52	17	14	3	5	11	1	0	1	6	5	0	0	0	9	0
2024M06	5	249	275	45	23	3	1	0	4	2	0	0	0	0	2	0	10	0	0
2024M07	9	307	194	26	4	4	2	2	0	1	25	12	0	0	0	4	3	0	0
2024M08	6	389	237	19	13	2	1	2	0	17	10	1	2	0	6	1	3	0	0
2024M09	5	357	143	18	9	2	7	2	32	1	2	0	10	0	2	2	0	0	0
2024M10	5	270	119	11	12	3	3	30	13	0	0	1	9	0	2	2	0	0	0
2024M11	4	132	93	26	4	0	8	0	1	0	0	6	0	0	0	0	0	0	0
2024M12	2	171	91	17	1	12	4	0	0	0	0	7	0	0	0	0	0	0	0
2025M01	5	128	125	35	56	34	1	3	0	0	1	5	0	0	0	0	0	0	0
2025M02	4	167	120	84	38	1	0	0	2	4	7	3	0	0	0	0	0	0	0
2025M03	2	266	187	43	5	7	2	0	3	11	0	0	0	0	0	0	0	0	0
2025M04	4	328	239	36	17	1	1	4	15	0	0	0	0	0	0	0	0	0	0
2025M05	13	364	209	43	18	2	3	14	7	0	0	0	0	0	0	0	0	0	0
2025M06	1	472	244	43	7	1	6	14	0	0	0	0	0	0	0	0	0	0	0
2025M07	14	432	212	18	5	8	5	0	0	0	0	0	0	0	0	0	0	0	0
2025M08	10	454	231	72	24	20	0	0	0	0	0	0	0	0	0	0	0	0	0
2025M09	7	322	223	39	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2025M10	2	294	186	35	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2025M11	1	190	155	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2025M12	2	172	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2026M01	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

## Zahlungen auf Basis synthetischer Einzeldaten

Für-Monat	N00	N01	N02	N03	N04	N05	N06	N07	N08	N09	N10	N11	N12	N13	N14	N15	N16	N17	N18
2024M01	2	164	124	36	40	0	2	16	14	1	0	0	0	0	1	0	0	0	0
2024M02	5	176	101	23	22	4	12	5	1	0	0	0	2	0	0	0	0	0	0
2024M03	3	160	202	32	11	27	16	1	4	5	0	0	0	0	0	0	6	1	0
2024M04	8	285	144	26	24	12	12	8	8	0	0	0	0	0	0	0	0	0	7
2024M05	5	250	203	37	22	14	5	2	17	2	0	1	8	0	0	0	0	0	11
2024M06	2	233	278	27	22	0	1	0	6	3	0	0	1	0	1	0	11	0	0
2024M07	5	332	185	18	1	12	4	2	0	1	28	10	0	0	0	8	5	0	0
2024M08	7	361	219	22	12	2	1	0	19	9	1	2	0	5	2	3	0	0	0
2024M09	4	386	125	23	12	5	8	3	37	1	2	0	9	0	4	1	0	0	0
2024M10	7	254	132	4	13	0	2	17	12	0	0	2	8	0	4	3	0	0	0
2024M11	4	124	112	35	1	0	7	0	4	0	0	5	0	0	0	1	0	0	0
2024M12	1	159	93	11	1	10	5	0	0	0	0	10	0	0	0	1	0	0	0
2025M01	4	167	135	33	55	43	2	21	0	0	2	7	0	0	0	0	0	0	0
2025M02	5	155	103	101	31	2	0	1	0	2	6	10	0	0	0	0	0	0	0
2025M03	1	250	193	45	2	7	1	0	5	7	0	0	0	0	0	0	1	0	0
2025M04	1	340	228	32	45	1	0	2	15	0	0	0	0	0	0	0	0	0	0
2025M05	10	378	197	41	12	0	8	11	8	0	0	0	0	0	0	0	0	0	0
2025M06	1	471	218	36	3	1	3	11	0	0	0	0	0	0	0	0	0	0	0
2025M07	16	425	193	4	7	8	2	0	0	0	0	0	0	0	0	0	0	0	0
2025M08	4	450	265	58	22	18	0	0	0	0	0	1	0	0	1	0	0	0	0
2025M09	5	326	291	42	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2025M10	0	269	188	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2025M11	1	185	134	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2025M12	3	177	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2026M01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

## Endstands-Prognose

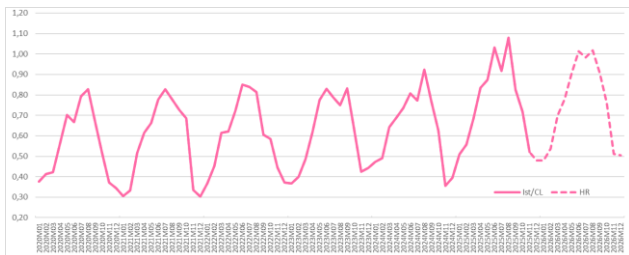
1. Wenn Abrechnungsstand ausreichend hoch (ca 80%):

### Chain-Ladder-Verfahren

- Kumulative Zahlungen
- CL-Fenster wählbar
- Restnachlauf einschätzen

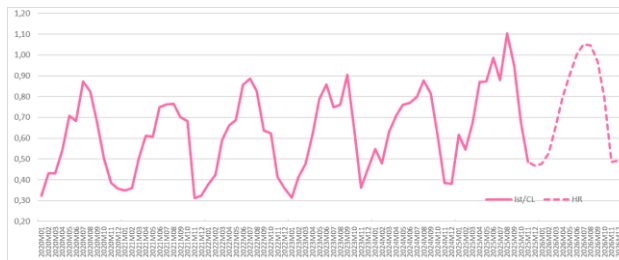
Ist/CL: Chain-Ladder-Verfahren  
HR: Zeitreihenprognose

## Endstand (Chain-Ladder und Zeitreihenprognose):



Statistisch  
gleiches  
Ergebnis

## Endstand (Chain-Ladder und Zeitreihenprognose):



## 2. Sonst Zeitreihenprognose

### Parameter:

- Welche Vorjahre?
- ... welche nicht?
- Zu-/Abschläge?  
(für externe Einflüsse)

## Herausforderung: Änderung der Abwicklungsgeschwindigkeit

### Kumulierte Zahlungen (wachstumsbereinigt)

Für-ZR																				Progn. Endstand (ohne Restlachauf 19+)		
	N00	N01	N02	N03	N04	N05	N06	N07	N08	N09	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	CL	CL321
A01	165	395	587	731	823	874	903	923	937	945	957	963	973	979	984	987	990	991	992		992	992
A02	174	459	616	742	836	879	909	928	937	951	958	969	975	979	983	986	988	989			990	990
A03	237	475	616	751	835	886	914	928	946	955	967	974	979	983	986	988	989				991	991
A04	245	461	613	743	841	890	911	933	947	961	971	976	982	986	988	990					994	994
A05	217	464	613	753	847	886	915	935	950	962	968	974	979	982	984						990	990
A06	223	461	614	755	839	888	915	936	951	966	973	979	983	985							995	993
A07	221	469	623	746	841	887	917	937	950	958	965	970	972								985	983
A08	222	467	610	748	836	888	918	938	948	958	965	968									987	982
A09	216	461	626	757	845	895	925	941	955	962	966										991	984
A10	202	483	636	761	843	896	922	941	952	959											992	982
A11	238	502	672	780	863	903	927	941	950												994	981
A12	246	543	694	811	871	909	929	941													998	982
A13	271	545	710	797	866	899	927														1.002	982
A14	311	620	721	808	857	894															996	973
A15	408	635	743	804	854																1.002	968
A16	403	662	742	803																	1.042	970
A17	436	668	748																		1.143	984
A18	443	669																			1.303	993
A19	434																				1.636	992

### Zahlungen

Entwicklungsfaktor für N01								
CL321	CL	Faktor	Für-ZR	N00	N01	N02	N03	N04
	2,39	2,39	A01	165	230	192	144	92
		2,52	A02	174	285	157	126	94
2,28	2,31	2,00	A03	237	238	141	135	84
2,05	2,18	1,88	A04	245	216	152	130	98
2,03	2,17	2,14	A05	217	247	149	140	94
2,06	2,15	2,07	A06	223	238	153	141	84
2,11	2,15	2,12	A07	221	248	154	123	95
2,10	2,14	2,10	A08	222	245	143	138	88
2,12	2,14	2,13	A09	216	245	165	131	88
2,26	2,17	2,39	A10	202	281	153	125	82
2,21	2,16	2,11	A11	238	264	170	108	83
2,21	2,16	2,21	A12	246	297	151	117	60
2,09	2,15	2,01	A13	271	274	165	87	69
2,04	2,13	1,99	A14	311	309	101	87	49
1,78	2,07	1,56	A15	408	227	108	61	50
1,67	2,03	1,64	A16	403	259	80	61	
1,57	1,98	1,53	A17	436	232	80		
1,54	1,94	1,51	A18	443	226			
			A19	434				

Beobachtung: Sehr starker Anstieg des CL-prognostizierten Endstands!

Ursache: ab Zeitraum A13 elektronisches Meldeverfahren

Empfehlung: Zusätzlich Verfahren mit Aktualitätsgewichtung (z.B. CL -3-2-1) einsetzen und Abweichung verfolgen

Geänderte Abwicklungsgeschwindigkeit kann zu erheblichen Fehleinschätzungen führen. Auch für ML-Verfahren kurzfristig meist schwer erkennbar

## Zusammenfassung

### Maschinelles Lernen in der Schadenreservierung

- Ermöglicht genauere Prognosen
- Erfolgsmodell: Gradient Boosting
- Interessante Einzelschadenmodelle auf Basis neuronaler Netze

Sehr gute Einarbeitungsmaterialien verfügbar

### Datenbasis für Modellentwicklung

- Synthetische Einzelschadendaten via *simulation machine* und *SPLICE* verfügbar/erzeugbar
- Abwicklungsdaten sind gut aus Echtdateen synthetisierbar (und ggf. geeignet für Veröffentlichung)

Sehr gute Einarbeitungsmaterialien verfügbar

### Sonderthemen

- Chain-Ladder mit Zeitreihenprognose verknüpfbar
- Änderung Abwicklungsgeschwindigkeit auch bei ML-Verfahren kritisch

## Anhang

- C. Hunsicker (2023), “On machine learning approaches to forecast non-life insurers’ loss reserves”
- Embeddings: Dimensionsreduktion bei nominalen Merkmalen
- Richman & Wüthrich, “One-Shot Individual Claims Reserving” (Preprint 13.3.2026)

## C. Hunsicker (Delfte, NL): “On machine learning approaches to forecast non-life insurers’ loss reserves”

[Master Thesis \(2023\)](#)

- Gleicher Datensatz wie Kuo (2018)
- Genauestes XGBoost-Modell verwendet Embeddings
- Fazit: “Loblieb” auf XGBoost (eXtreme Gradient Boosting)

### Architektur:

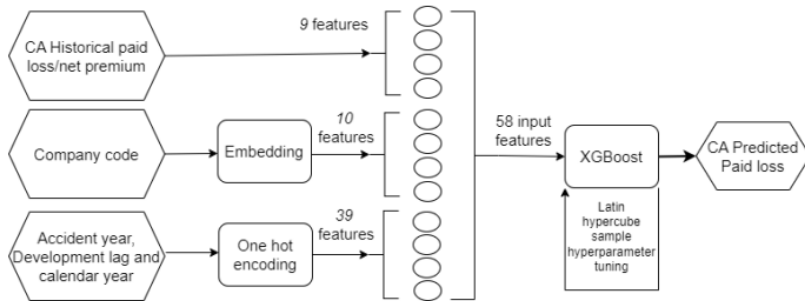


FIGURE 4.13: XGBoost model architecture for commercial auto data set.

### Modellgüte:

Models	Results							
	Line of business							
	Commercial auto		Private passenger auto		workers compensation		other liability	
MAE	RSME	MAE	RSME	MAE	RSME	MAE	RSME	
Mack chain ladder	0,33	0,49	0,22	0,37	0,29	0,39	0,43	0,65
Bootstrap chain ladder	0,42	0,67	0,26	0,57	0,30	0,40	0,46	0,67
Clark’s LDF	0,52	0,69	0,39	0,51	0,32	0,43	0,58	0,81
XGBoost architecture	0,31	0,42	0,20	0,33	0,26	0,34	0,55	0,71
XGBoost full data	0,26	0,40	0,21	0,33	0,26	0,35	0,40	0,56
XGBoost split data	0,24	0,38	0,19	0,28	0,21	0,30	0,42	0,57
AutoML full data	0,26	0,37	0,36	0,43	0,24	0,32	0,44	0,64
AutoML split data	0,35	0,47	0,65	0,73	0,85	0,98	0,50	0,70

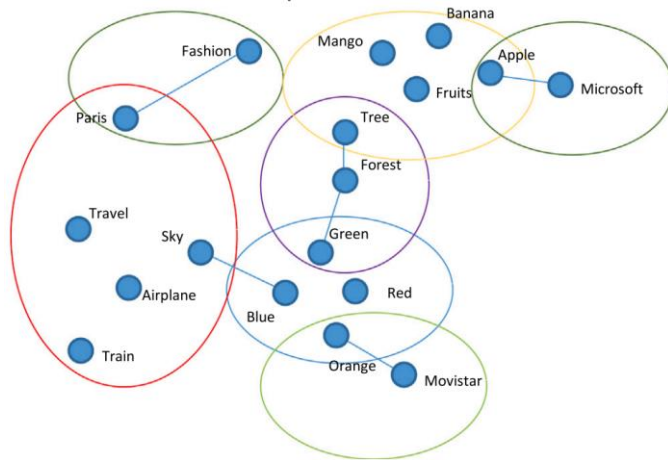
TABLE 5.1: The RSME and MAE results of all described models and architectures.

“The research has shown that machine learning models, particularly the XGBoost algorithm, have demonstrated significant improvements in predictive accuracy over traditional methods.”

## Embeddings: Dimensionsreduktion bei nominalen Merkmalen

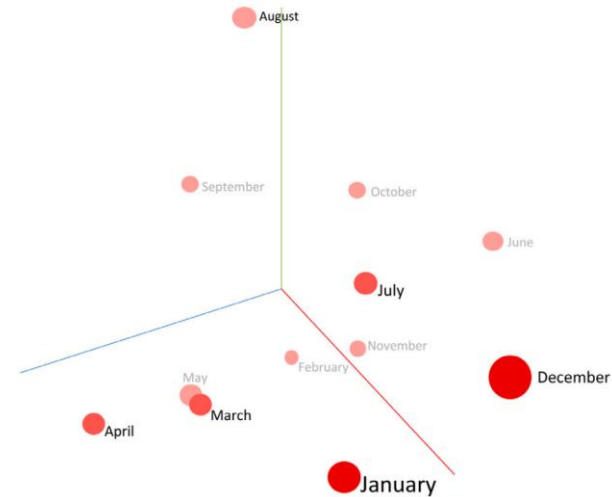
### Word-Embedding:

Wortvektoren, projiziert auf zwei Dimensionen



### Embeddings für Energiepreise:

Kalenderdaten-Embedding für Monat (3D)



Quelle: Wagner et al. (2022), "Short- and long-term forecasting of electricity prices using embedding of calendar information in neural networks"

Weitere Beispiele mit Global Vectors for Word Representation (GloVe, 2014) siehe <https://www.kaggle.com/code/floser/examples-of-similar-word-embeddings-in-glove>

## Richman & Wüthrich, “One-Shot Individual Claims Reserving” ([Preprint 13.3.2026](#))

### Abstract

Individual claims reserving has not yet become established in actuarial practice. We attribute this to the absence of a satisfactory methodology: existing approaches tend to be either overly complex or insufficiently flexible and robust for practical use. Building on the classical chain-ladder (CL) method, we introduced a new perspective on individual claims reserving in Richman and Wüthrich [[arXiv:2602.15385](#)]. This manuscript has sparked considerable discussion within the actuarial community. The aim of the present paper is to continue and deepen that discussion, with the ultimate goal of advancing toward a new standard for micro-level reserving.

**Keywords.** Claims reserving, chain-ladder method, individual claims reserving, micro-level reserving, granular reserving, neural networks, Mack’s method.