Insurance Data Science Conference 16 - 18 June 2021

Programme and Book of Abstracts

2021-06-16

Scientific Committee:

- Katrien Antonio (KU Leuven and University of Amsterdam)
- Christophe Dutang (Université Paris Dauphine)
- Arthur Charpentier (Université du Québec à Montréal)
- Markus Gesmann (Insurance Capital Markets Research, London)
- Ioannis Kyriakou (City, University of London)
- Davide de March (Markel, London)
- Silvana Pesenti (University of Toronto)
- Rosalba Radice (City, University of London)
- Ronald Richman (SA Taxi, South Africa)
- Jürg Schelldorfer (Swiss Re, Zurich)
- Giorgio Alfredo Spedicato (Unipol Group, Milan)
- Andreas Tsanakas (City, University of London)
- Mario Wüthrich (ETH Zurich)
- Rui Zhu (City, University of London)

Conference sponsors:

Swiss Re Institute

MiraiSolutions

smarter analytics - better decisions

Programme

UK Times (GMT + 1)

16/06/21

14:00-15:00

Opening remarks: Andreas Tsanakas

Keynote (Chair: Andreas Tsanakas) — **Thomas Wiecki** — **Bayesian decision making lifts off with PyMC3**

Short talks (15min each)

Session 1A: Pricing (Chair: Arthur Charpentier)		Session 1B: Life & Mortality (Chair: Pietro Millossovich)	
Jonathan Sedar	A Novel Bayesian Pricing Model with Risk-Level Freq-Sev Decomposition, running in Production	Lina Palmborg	Efficient use of data for LSTM mortality forecasting
Tsz Chai (Samson) Fung	Mixture composite regression models with multi-type feature selection	Salvatory Kessy	Mortality Forecasting Using Stacked Regression Ensemble
Emiliano A. Valdez	Synthetic Dataset Generation of Driver Telematics	Mike Ludkovski	Multi-output Gaussian Processes for longevity analysis

15:15 – 16:15

Lightning talks (7min each)

Session 2A: Claims modelling (Chair: Markus Gesmann)		Session 2B: Simulation/QF/meta models (Chair: Silvana Pesenti)	
Ranjini Vaidyanathan	Enhancing Auto Claim Review using Machine Learning	Rodrigo S. Targino	Risk Budgeting Portfolios from Simulations
Ziyi Li	Neural Network Embedding of the Negative Binomial Regression Model for Claim Frequencies	Fernando Mierzejewski	Modelling Credit Structures and Securitisations with Data Science
Mohammad Zoynul Abedin	A Novel Hybrid Method to Predict Insurance Claim by Mining Imbalanced Datasets	Alexandre Carbonneau	Deep Hedging of Long-Term Financial Derivatives
Ivan Sergienko	Deep Learning for Stochastic Policy-Level Modelling	Fabio Viviano	Monte Carlo Valuation of Future Annuity Contracts
Symeon Koumoutsaris	Modelling the subsidence risk in France taking into account the effects of climate change	Giovanni Rabitti	From local to structural input importance in variable annuities data
Seema Sangari	Correcting Reporting Delays in Cyber Events at Industry level	Xiaochen Jing	Metamodeling for Variable Annuity Valuation: What works and what does not

16:30 – 17:30

17/06/21

Opening remarks: Markus Gesmann

Panel discussion: AI – Coming of age? (Chair: Trevor Maynard) — **Luca Baldassarre** (Lead Data Scientist, Swiss Re), **Trevor Maynard** (Head of Innovation, Lloyd's), **Liz McFall** (Director of Data Civics & Chancellor's Fellow, University of Edinburgh), **Yves-Alexandre de Montjoye** (Associate Professor at Imperial College London), **Maria Óskarsdóttir**, (Assistant Professor in Computer Science, Reykjavik University)

Session 3A: Reserving (Chair: Mario Wüthrich)		Session 3B: ML (Chair: Davide de March)	
Steve Guo	Patterns and Anomalies of Loss Development in P&C Insurance Market	Veronica Coronel Vera	Person at the centre: how AutoML boosts behavioural modelling in P&C insurance
Henning Zakrisson	Gradient Boosting Machines in Collective Reserving Model for Reserves Prediction	Mouloud Belbahri & Olivier Gandouet	A Twin Neural Model for Uplift
Markus Gesmann	Hierarchical Compartmental Reserving Models	Anna Kozak	The use of denoising autoencoders for categorical and continuous variables

Session 4A: Programming/industry deployment (Chair: Jürg Schelldorfer)		Session 4B: Social media/covid/fraud (Chair: Arthur Charpentier)	
Patrick Laub	Approximate Bayesian Computation and Insurance	John Ng	Gompertz network and Lasso regularisation in modelling age-specific impact of COVID-19 vaccination
Chibisi Chima-Okereke	Introducing the D programming language to Insurance	Rei England	Spreading the word: The effect of word-of-mouth networks on insurance customer choices
Daniel Murphy	Month arithmetic in R with the mondate package	Rohan Yashraj Gupta	A comparative study of using various Machine Learning and Deep Learning based fraud detection models for Universal Health Coverage schemes and assessing the impact of COVID-19 in healthcare fraud
Francesca Vitalini	Agile; the right answer for the 'next normal' in the insurance sector	Shrinivas Shikhare	Next Generation LTC - Life insurance Underwriting using Facial Score Model
Kenneth Lim & Maxime Allard	Processing Insurance Claims with Automated, Scalable and Fair AI	Tim King	Process Mining Applied to Complex Medical Claims Management
[redacted]avier Marechal, Anja Friedrich	Investigating Applications of Data Science in UK and non-UK Actuarial Teams		

18/06/21

Opening remarks: Ioannis Kyriakou

Keynote (Chair: Ioannis Kyriakou) — **Bettina Grün** — **Advances in Model-Based Clustering**

Session 5A: Pricing (Chair: Ioannis Kyriakou)		Session 5B: Life & Mortality (Chair: Silvana Pesenti)	
Robert Matthijs Verschuren	Customer Price Sensitivities in Competitive Automobile Insurance Markets	Qiqi Wang	Multi-State Health Transition Modeling Using Neural Networks
Montserrat Guillen	Number of claims and number of near-misses for telematics pricing in automobile insurance	Hang Nguyen	Scenario selection with Lasso regression for the valuation of variable annuity portfolio
Łukasz Delong	Gamma Mixture Density Networks and their application to modelling insurance claim amounts	Mario Marino	Deepening Lee-Carter for longevity projections with uncertainty estimation

Session 6A: Statistical modelling (Chair: Andreas Tsanakas)		Session 6B: Life (Chair: Pietro Millossovich)	
S.R.Pranav Sai	Capturing the power of ensemble learning using GLM and Artificial Neural Network for insurance pricing	Katja Hanewald	Multi-population modeling with economic trends: A hybrid neural network approach
Pierre-Oliver Goffard	Sequential Monte Carlo Samplers to fit and compare insurance loss models	Francesco Ungolo	A Dirichlet Process Mixture model for the analysis of competing risks
Rui Zhu	Copula model selection using image processing	Simon Schnürch	Point and Interval Forecasts of Death Rates Using Neural Networks
Giles Stupfler & Abdelaati Daouia	Extremile Regression	Claudio Giorgio Giancaterino	Unsupervised Learning applied to the Customer Lifetime Value (CLV)
Himchan Jeong	A non-convex regularization approach for stable estimation of loss development factors	Zhiyu Quan	Tree-based Models for Variable Annuity Valuation: Parameter Tuning and Empirical Analysis
Queensley Chukwudum	Relativities in the Over-Dispersed Poisson Bootstrap Claims Reserves		

Virtual conference centre

Throughout the conference you can meet other attendees, speakers and sponsors in our virtual conference centre on Gather.

Join Gather any time, but particularly before and at the end of each day's proceedings and during the breaks to meet and talk to other attendees, just as you would do in person.

Joining instructions for Gather will be shared with registered attendees via email.

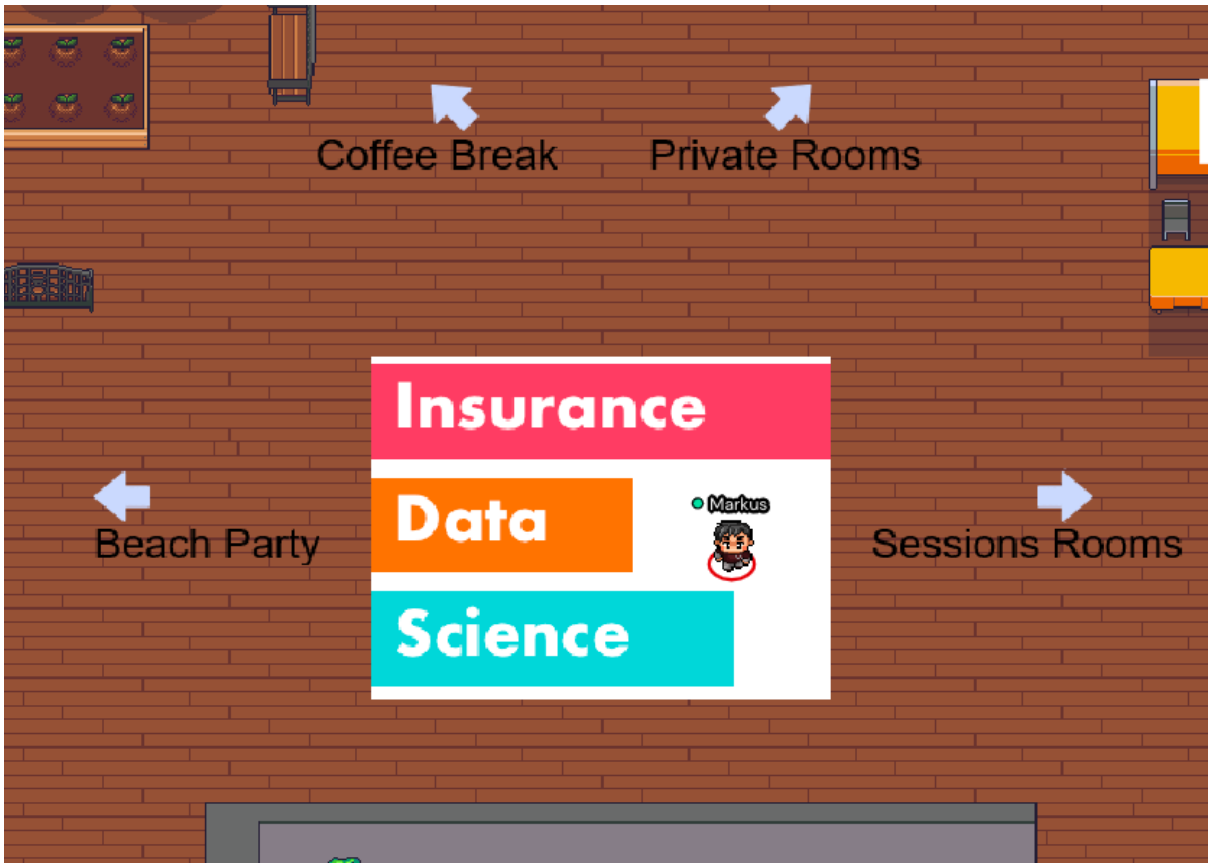


Figure 1: Join Gather during the conference and meet other attendees.

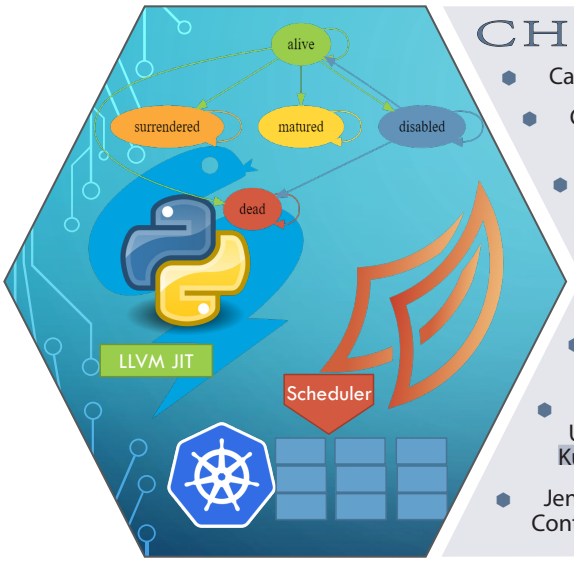
Life Probabilistic Modeling

Mirai Solutions

Highly performant Python code with JIT compilation and parallel/distributed computing.

We have successfully implemented a framework to allow the Life team of an Insurance Company to perform the implementation and testing of life actuarial models, as required by regulators, in a highly streamlined and cost-effective manner.

Current performance is comparable to highly optimized commercial solutions written in C++



CHALLENGES

- Calculations in pure Python are slow
- Complex state models based on time inhomogeneous Markov chains
- Scaling up to 1000s of CPUs
- Reduce risk of failure and improve the quality of the delivery

SOLUTION

- Python framework makes use of Numba to accelerate calculations
- Dask provides advanced parallelism for analytics. Using Dask for distributed calculations in Kubernetes cluster enables to scale up to 1000s of CPUs
- Jenkins and Bitbucket play a key role in the Continuous Integration pipeline

BENEFITS

- 70% cost saving:
  Replacing expensive proprietary solutions with open source software and the ability to scale computational resources on demand will allow up to 70% cost saving after an initial transition period.
- CI set-up contributes to minimizing the number of bugs and regressions and gives confidence when modifying the codebase.

"Continuous Integration is very rare in actuarial modeling. Designing and implementing the CI pipeline was one of my key contributions to the project success. I also enjoyed the extensive usage of open source, cutting-edge technologies."

contact@mirai-solutions.com

Keynotes

Bayesian decision making lifts off with PyMC3

Thomas Wiecki, PyMC Labs

Thomas Wiecki is the Chief Executive Officer and founder of PyMC Labs (www.pymc-labs.io) - a Bayesian consultancy. He is an author of the popular PyMC3 package — a probabilistic programming framework written in Python. He holds a PhD from Brown University.

Abstract:

Bayesian modeling is a powerful paradigm with key properties relevant to the insurance industry: uncertainty quantification, pooling of information from different data sources, the ability to include expert knowledge in a principled way, and the flexibility to model a wide variety of data sets.

However, while a great tool to assess the risk of certain events to occur as well as how uncertain we are about this risk estimate, we still have to make a final decision about which contracts to insure at what price. In this talk I will demonstrate how Bayesian Decision Making is the missing piece to solving this problem. With an applied problem from insuring rocket launches I will show how PyMC3 can be used to balance various trade-offs such as risk, uncertainty, correlations, and upside to find the optimal allocation of our insurance budget.

Keywords: Bayesian Modelling, PyMC3,

Advances in Model-Based Clustering

Bettina Grün, WU Vienna

Bettina Grün is an applied statistician at WU Vienna, Austria. She is an ordinary member of the R Foundation and maintains and co-authored several R packages on the Comprehensive R Archive Network (CRAN) covering finite mixture models, topic models and mining association rules. Her research interests are in statistical computing and in the application of statistical methods in different areas such as tourism, marketing and economics.

Abstract:

Cluster analysis aims a grouping objects and is a main task in exploratory data analysis, statistical data analysis and machine learning. Model-based approaches have the advantage that model specification and selection are performed within a principled statistical framework, facilitating interpretation, improving validation and including uncertainty quantification.

We will give an overview on recent advances in model-based clustering including challenges due to big data settings, new model extensions as well as developments regarding estimation and validation methods. We will also discuss implementations in software available in particular for the R environment for statistical computing and graphics.

Keywords: finite mixture models, model-based clustering,

References

1. Grün, B. (2019). Model-based clustering. In *Handbook of Mixture Analysis*, 157-192. Chapman and Hall/CRC. https://arxiv.org/abs/1807.01987
2. Leisch, F., Grün, B. (2020). *CRAN Task View: Cluster Analysis & Finite Mixture Models.* https://CRAN.R-project.org/view=Cluster.

Contact details

- Email: Bettina.Gruen@wu.ac.at
- Homepage: https://statmath.wu.ac.at/~gruen

Panel discussion

AI - Coming of age

This panel will explore how we can harness the best of new AI technology and control the risks of it.

Panel members:

- Trevor Maynard (Head of Innovation, Lloyd's) – Moderator/Chair
- Luca Baldassarre (Lead Data Scientist, Swiss Re)
- Liz McFall (Director of Data Civics & Chancellor's Fellow, University of Edinburgh)
- Yves-Alexandre de Montjoye (Associate Professor at Imperial College London – Head of the Computational Privacy Group)
- María Óskarsdóttir (Assistant Professor, Department of Computer Science, Reykjavik University)

Abstracts/Session_1A

A Novel Bayesian Pricing Model with Risk-Level Freq-Sev Decomposition, running in Production

Jonathan Sedar, Oreum Industries

Abstract: Oreum Industries has developed a general approach to pricing commercial lines that replaces the disjointed process of conventional model development with an integrated & holistic framework based on probabilistic programming. At the core this is a hierarchical Bayesian model of expected loss-cost with risk-level zero-inflated freq-sev decompositions and coupled covariance; built for offline inference and online probabilistic price prediction in Production; implemented in Python using the industry leading pymc3 framework and arviz diagnostic tooling.

The solution is tailored to individual clients, where modelled effects due to risk characteristics are interpretable by underwriters and the probabilistic price output allows for constrained underwriter discretion. The model is open-box, and the technical work is designed to be entirely reproducible and extensible by the client's own data science teams, including data processing, EDA & modelling modules, and reproducible reporting and documentation via Jupyter notebooks. We also guide the client's in-house dev teams to integrate the model into their Production platforms.

This presentation will be a description of the problem space, the general model solution, discussion of choices made and how this fits into an end-to-end Production process.

Keywords: commercial-auto, Bayesian-inference, pricing, frequency-severity, probabilistic-programming, pymc3, python

References

1. McElreath & Koster. (2014). Using Multilevel Models to Estimate Variation in Foraging Returns: Effects of Failure Rate, Harvest Size, Age, and Individual Heterogeneity. Human Nature, 25, 100-120. https://github.com/rmcelreath/mcelreath-koster-human-nature-2014
2. Jones JH, Bird RB, Bird DW. (2013) To kill a kangaroo: understanding the decision to pursue high-risk/high-gain resources. Proc R Soc B 280: 20131210. http://dx.doi.org/10.1098/rspb.2013.1210
3. Salvatier J., Wiecki T.V., Fonnesbeck C. (2016) Probabilistic programming in Python using PyMC3. PeerJ Computer Science 2:e55 DOI: 10.7717/peerj-cs.55. https://docs.pymc.io

Contact details

- Email: jon.sedar@oreum.io
- Homepage: https://oreum.io

Mixture composite regression models with multi-type feature selection

Tsz Chai (Samson) Fung, ETH Zurich (presenter)

George Tzougas, London School of Economics and Political Science

Mario Wüthrich, ETH Zurich

Abstract: The aim of this paper is to present a mixture composite regression model for claim severity modelling. Claim severity modelling poses several challenges such as multimodality, heavy-tailedness and systematic effects in data. We tackle this modelling problem by studying a mixture composite regression model for simultaneous modeling of attritional and large claims, and for considering systematic effects in both the mixture components as well as the mixing probabilities.

For model fitting, we present a group-fused regularization approach that allows us for selecting the explanatory variables which significantly impact the mixing probabilities and the different mixture components, respectively. We develop an asymptotic theory for this regularized estimation approach, and fitting is performed using a novel Generalized Expectation-Maximization algorithm. We exemplify our approach on real motor insurance data set.

Keywords: Splicing, Generalized Expectation-Maximization algorithm, Variable selection, Asymptotic normal theory, Multimodal and heavy-tailed claim losses

References

1. Fung, T.C., Tzougas, T., Wuthrich, W. (2021). Mixture composite regression models with multi-type feature selection. Submitted.

Contact details

- Email: tszchai.fung@math.ethz.ch

Synthetic Dataset Generation of Driver Telematics

Emiliano A. Valdez, University of Connecticut (presenter)

Banghee So, University of Connecticut

Jean-Philippe Boucher, Université du Québec à Montréal

Abstract: This article describes techniques employed in the production of a synthetic dataset of driver telematics emulated from a similar real insurance dataset. The synthetic dataset generated has 100,000 policies that included observations about driver's claims experience together with associated classical risk variables and telematics-related variables. This work is aimed to produce a resource that can be used to advance models to assess risks for usage-based insurance. It follows a three-stage process using machine learning algorithms. In the first stage, a synthetic portfolio of the space of feature variables is generated applying an extended SMOTE algorithm. The second stage is simulating values for the number of claims as multiple binary classifications applying feedforward neural networks. The third stage is simulating values for aggregated amount of claims as regression using feedforward neural networks, with number of claims included in the set of feature variables. The resulting dataset is evaluated by comparing the synthetic and real datasets when Poisson and gamma regression models are fitted to the respective data. Other visualization and data summarization produce remarkable similar statistics between the two datasets. We hope that researchers interested in obtaining telematics datasets to calibrate models or learning algorithms will find our work valuable.

Keywords: Bayesian optimization, Gaussian process, Neural network, SMOTE, Usage-based insurance (UBI), Vehicle telematics

References

1. Baecke, P. and Bocca, L. (2017) The value of vehicle telematics data in insurance risk selection processes. *Decision Support Systems* **98**, 69–79.
2. Chawla, N.V., Bowyer, K.W., Hall, L.O. Hall, and Kegelmeyer, W.P. (2002) SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* **16**, 321–357.
3. Dalkilic, T.E., Tank, F., and Kula, K.S. (2009) Neural networks approach for determining total claim amounts in insurance. *Insurance: Mathematics and Economics* **45(2)**, 236–241.
4. Goodfellow, I., Bengio, Y., and Courville, A. (2016) *Deep Learning*. MIT Press.
5. McCulloch, W.S. and Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* **5**, 115–133.

Contact details

- Email: emiliano.valdez@uconn.edu
- Homepage: http://www2.math.uconn.edu/~valdez/
- Social media: Twitter @evaldez2021

Abstracts/Session_1B

Efficient use of data for LSTM mortality forecasting

Lina Palmborg, Stockholm University (presenter)

Abstract: We consider a simple long-short term memory (LSTM) neural network extension of the Poisson Lee-Carter model, with a particular focus on different procedures for how to use training data efficiently, combined with ensembling to stabilise the predictive performance. We compare the standard approach of withholding the last fraction of observations for validation, with two other approaches: sampling a fraction of observations randomly in time; and splitting the population into two parts by sampling individual life histories. We provide empirical and theoretical support for using these alternative approaches.

Furthermore, to improve the stability of long-term predictions, we consider boosted versions of the Poisson Lee-Carter LSTM. In the numerical illustrations it is seen that even in situations where mortality rates are essentially log-linear as a function of calendar time, the boosted model does not perform significantly worse than a simple random walk with drift, and when non-linearities are present the predictive performance is improved. Moreover, boosting allows us to obtain reasonable model calibrations based on as few data points as 20 years.

This is based on joint work with M. Lindholm.

Keywords: sequential neural networks, mortality forecasting, ensemble models, boosting, Lee-Carter model

References

1. Lindholm, M., Palmborg, L. (2021). Efficient use of data for LSTM mortality forecasting. Preprint available at https://www.su.se/profiles/lipa5748

Contact details

- Email: lina.palmborg@math.su.se

Mortality Forecasting Using Stacked Regression Ensembles

Salvatory R. Kessy, University of New South Wales (presenter)

Michael Sherris, Andrés M. Villegas, and Jonathan Ziveyi, University of New South Wales

Abstract: We present a stacked regression ensemble method that optimally combines different mortality models to reduce the mean squared errors of mortality rate forecasts and mitigate model selection risk. Stacked regression uses a supervised machine learning algorithm to approximate the horizon-specific weights by minimizing the cross-validation criterion for each forecasting horizon. The horizon-specific weights facilitate the development of a mortality model combination customized to each horizon. Unlike other model combination methods, stacked regression simultaneously solves model selection and estimates model combinations to improve model forecasts. Our numerical illustrations based on 44 populations from the Human Mortality Database demonstrate that stacking mortality models increases predictive accuracy. Using one-year-ahead to 15-year-ahead out-of-sample mean squared errors, we find that stacked regression improves mortality forecast accuracy by 13% - 49% and 19% - 90% over the individual mortality models for males and females, respectively. Therefore, combining the mortality rate forecasts provides lower out-of-sample point forecast errors than selecting the single best individual mortality method. Stacked regression ensemble also achieves better predictive accuracy than other model combination methods, namely Simple Model Averaging, Bayesian Model Averaging, and Model Confidence Set. Our results support the stacked regression ensemble approach over individual mortality models and other model combination methods in forecasting mortality rates. We also provide a user-friendly open-source R package, `CoMoMo`, that combines multiple mortality rate forecasts using different model combination techniques.

Keywords: Stacked regression, ensemble learning, cross-validation, model uncertainty, model combination, age-period-cohort model, mortality forecasting.

References

1. Kessy, Salvatory, Michael Sherris, Andrés Villegas, and Jonathan Ziveyi. 2021. "Mortality Forecasting Using Stacked Regression Ensembles." SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3823511.
2. John, M. Bates and Clive W. J. Granger. 1969. "The Combination of Forecasts." Journal of the Operational Research Society 20 (4): 451–68. https://doi.org/10.1057/jors.1969.103.
3. Clarke, Bertrand. 2004. "Comparing Bayes Model Averaging and Stacking When Model Approximation Error Cannot Be Ignored." Journal of Machine Learning Research 4 (4): 683–712. https://doi.org/10.1162/1532443047739360 90.
4. SriDaram, Dilan, Michael Sherris, Andrés Villegas, and Jonathan Ziveyi. 2021. "A Group Regularisation Approach for Constructing Generalised Age-Period-Cohort Mortality Projection Models." SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3790991.
5. Wolpert, David H. 1992. "Stacked Generalization." Neural Networks 5: 241–59.

Contact details

- Email: s.kessy@unsw.edu.au
- LinkedIn: https://www.linkedin.com/in/salvatory-kessy-a9b7a615b/
- Repository: https://github.com/kessysalvatory/CoMoMo

Multi-output Gaussian Processes for longevity analysis

Mike Ludkovski, University of California Santa Barbara (presenter)

Nhan Huynh, University of California Santa Barbara

Abstract: I will discuss several interrelated projects on the use of multi-output Gaussian Process (GP) models for analyzing longevity jointly across several populations. The spatial covariance framework underlying GPs offers a flexible way to capture cross-population dependence within an Age-Period setup. In turn GPs generate rigorous uncertainty quantification, non-parametric forecasts, and a transparent method for information fusion and extrapolation. In order to handle 5+ populations, we moreover investigate coregionalization ideas for dimension reduction. Our applications include analysis of Human Mortality Database data across EU nations and genders, analysis of cause-of-death datasets across a handful of countries, and analysis of national Covid-19 excess deaths. All the models are implemented in R and illustrations include 2 RShiny apps posted on the web.

Keywords: Gaussian Processes, Mortality Modeling, RShiny

References

1. N. Huynh, M. Ludkovski, Multi-Output Gaussian Processes for Multi-Population Longevity Modeling (2021). *Annals of Actuarial Science*, to Appear, arxiv.org/2003.02443
2. N. Huynh, M. Ludkovski, *Tutorial on Multi-Output GPs*, (2020). https://nhanhuynh46.github.io/MOGPTutorials/.

Contact details

- Email: ludkovski@pstat.ucsb.edu
- Homepage: http://www.pstat.ucsb.edu/faculty/ludkovski
- Social media: https://www.linkedin.com/in/mike-ludkovski-9623a927, twitter.com/MLudkovski

Abstracts/Session_2A

Enhancing Auto Claim Review using Machine Learning

Ranjini Vaidyanathan, CCC Information Services

Abstract: Auto insurance carriers receive thousands of estimates from repair facilities every day. And though some estimates are reviewed by the insurer in person at the vehicle's location, many estimates are sent to the insurer to be reviewed remotely, based on images of the damaged vehicle and the detailed estimate. Sometimes, carriers do not review a submitted estimate. Carriers may instead apply rule-based priority schemes to pick the subset of estimates to review.

This talk will discuss a modeling approach that uses large historical datasets to train a predictive model to automatically recommend a subset, based on the insurer's preferences, of estimates to review. By training machine learning and deep learning algorithms on complex estimate line data and on claim images that show vehicle damage, we see increased performance compared to traditional rule-based methods. The result is a more efficient claim review process for insurance carriers.

Keywords: Machine Learning, Auto Claim Estimates

Neural Network Embedding of the Negative Binomial Regression Model for Claim Frequencies

George Tzougas, London School of Economics and Political Science

Ziyi Li, London School of Economics and Political Science (presenter)

Abstract: The aim of this paper is to embed the overdispersed Negative Binomial regression model for claim counts into a neural network architecture following the Combined Actuarial Neural Network approach (CANN) approach of Wüthrich and Mertz (2019). The implementation of the blended model is illustrated by a real data application which involves fitting French motor third-party liability (MTPL) insurance data. We demonstrate that the neural net boosting of the Negative Binomial regression model allows us to explore missing interactions of nonmultiplicative type that cannot be captured by the Negative Binomial Regression model.

Keywords: Neural networks, Architecture, MTPL insurance, Negative Binomial regression model, Nesting, Embedding layers, Claim frequency, Machine learning, Deep learning

References

1. Wüthrich, M.V., Merz, M. (2019). The art of modeling article. *ASTIN Bulletin: The Journal of the IAA* **49(1)**, 120-122.

Contact details

- Email: g.tzougas@lse.ac.uk

A Novel Hybrid Method to Predict Insurance Claim by Mining Imbalanced Datasets

Mohammad Zoynul Abedin, Ph.D. Department of Finance and Banking Hajee Mohammad Danesh Science and Technology University

Abstract:This proposal aims to predict the insurance claim appropriately to eliminate the devastating effect in an insurer's insurance portfolio[1]. The number of claims exposes the intrinsic risk of a policyholder. When an insurer fails to predict insurance claim accurately, it consequences a negative effect on their insurance portfolio. Therefore, to accurately predict the insurance claim in insurance data science, this proposed study will mine the insurance data by selecting significant feature, balances the data and then classifies the claims. Firstly, this proposal will use Gradient Boosting (GB), Random Forest (RF), XG Boost (XGB), Least Angle Regression, Lasso, and Linear SVM for insurance data mining. Secondly, to rectify the data imbalance problem, this study proposes Synthetic Minority Oversampling Technique (SMOTE) and Random Over-Sampling Examples (ROSE) algorithms. This proposal will also apply Gradient Boosting (GB), Random Forest (RF) and XG Boost (XGB) for data classification purpose. The applied performance measures such as accuracy, Area Under the Curve (AUC), F-Measure, Maximum Profit (MP) and Expected Maximum Profit (EMP) will evaluate the classifiers' performance. The maximum profit of a classification technique refers to the profit deriving from the classification outcome when the optimal cutoff is used. The expected maximum profit refers to the anticipation of the maximum profit of a classifier regarding the distribution of classification costs (Verbraken et al., 2013). Applied insurance data will come from kaggle[2],[3] and leidenuniv[4]. The possible findings may apply to lessen the negative impact on insurance portfolio. It may also provide guidelines to the insurers and other stakeholders to evaluate the probable insurance claims to cope up with insurance portfolio.

Keywords: Insurance data science, imbalanced data, feature engineering, claim prediction, performance measure

References

Verbraken, T., Verbeke, W., and Baesens B (2013). A Novel Profit Maximizing Metric for Measuring Classification Performance of Customer Churn Prediction Models. *IEEE Transactions on Knowledge and Data Engineering*, 25 (5): 961-973

Contact details

- Profile: https://team-bigdata.com/
- LinkedIn Profile: https://www.linkedin.com/in/abedin-mohammad-zoynul-b70170132/

[1]Insurance portfolio represents company obligations and so insurers keep an equivalent amount of assets to meet these obligations.

[2]https://www.kaggle.com/c/porto-seguro-safe-driver-prediction, https://www.kaggle.com/c/bnp-paribas-cardif-claims-management, https://www.kaggle.com/c/bnp-paribas-cardif-claims-management

[3]https://www.kaggle.com/c/ClaimPredictionChallenge

[4]http://liacs.leidenuniv.nl/~puttenpwhvander/library/cc2000/

Deep Learning for Stochastic Policy-Level Modelling

Ivan Sergienko, Riskfuel Analytics (presenter)

Maxime Bergeron, Riskfuel Analytics

Abstract: Valuation, hedging, and financial reporting of insurance products with embedded derivatives, such as Variable Annuities and Universal Life, often relies on Monte Carlo simulation of markets and events. IFRS 17 requires market-consistent measurement of options and guarantees, increasing the industry's computational burden. Deep Learning promises to deliver significant computational speed-up compared to traditional stochastic models [1]. In this talk we demonstrate that the technology has evolved enough to handle multi-dimensional problems encountered
in realistic industrial applications.

Keywords: Deep Learning, Monte Carlo, Stochastic Modelling, IFRS 17

References

1. Doyle, D.; Groendyke, C. (2019). Using Neural Networks to Price and Hedge Variable Annuity Guarantees. Risks 2019, 7, 1. https://doi.org/10.3390/risks7010001

Contact details

- Email: ivan@riskfuel.com
- Homepage: https://riskfuel.com
- Social media: https://www.linkedin.com/in/isergienko/

Modelling the subsidence risk in France taking into account the effects of climate change

Symeon Koumoutsaris, Guy Carpenter (presenter)

Abstract: Droughts have important impacts on many sectors of the economy, such as in agriculture and the energy sector, but are also closely related to hazards with growing insured loss relevance, such as wildfires and drought-induced soil movements or subsidence. In France in particular, damage to properties due to subsidence account for one third of the insured losses from natural catastrophes (CCR, 2018). A statistical model to assess the risk of subsidence is presented here. Relative soil moisture deficit from reanalysis is modelled using a 1-inflated beta generalized additive model, with the help of the gamlss R package, including time as predictor in order to account for the non-stationarity in the data. We assess the risk under future climate using the EURO-CORDEX regional climate model simulations.

Keywords: Subsidence, model, climate change

References

1. Caisse Centrale de Réassurance (CCR), Les Catastrophes Naturelles en France, 2018 (https://www.ccr.fr/documents/35794/35836/plaquette_Bilan_4P_A4_05-2018_V6.pdf/bf643039-faf5-4aa4-bb69-ffaf33b87674?t=1531151583000)

Contact details

- Email: symeon.koumoutsaris@guycarp.com

Correcting Reporting Delays in Cyber Events at Industry level

Seema Sangari, School of Data Science and Analytics, Kennesaw State University (Presenter)

Dr. Eric Dallal, AIR Worldwide, Verisk Cyber Solutions

Scott Stransky, AIR Worldwide, Verisk Cyber Solutions

Abstract: Having reporting delays is a well-known problem in the cyber insurance industry. With rapidly evolving cyber incidents, cyber insurers find it challenging to estimate potential future cyber claims appropriately in the absence of complete information. Healthcare data breach statistics from HIPAA accepted that cyber events remain undetected for months/years after they occur- Marriott's cyber event occurred in 2014 but was reported in 2018. Furthermore, unlike the public sector, private organizations frequently have less restrictive reporting requirements. Often private sector breaches come into picture only once they become news headlines, as they don't want to disclose incidents for fear of reputation risk. As a result, cyber data is biased and needs to be corrected to capture the real level of incidents. Current data might show a decrease in the number of events but, in reality, it's been increasing and we are just not aware of it yet. Researchers have been investigating this problem for the last few decades from different perspectives. Reporting delays are not restricted to the cyber world but the healthcare industry too – COVID-19 is classic example. One can't find out whether s/he is infected until symptoms show, which might take some time. Direct estimation of reporting delays from raw data suffers from two problems: it overestimates short delays, since recent events with long reporting delays will not yet be in data sets; and it precludes reporting delays longer than the longest reporting delay in the data set. Additionally, cyber incident data shows non-stationarity in the reporting delay distribution. We present an algorithm that deals with all three problems via estimation of a series of parametric reporting delay distributions. We validate our method by using our algorithm to estimate a one year ahead correction of cyber incident counts reported as of 2018 and comparing these to counts reported as of 2019.

Keywords: Cyber Events, Cyber Insurance, Delay Distribution, Reporting Delays

References

1. Healthcare Data Breach Statistics. https://www.hipaajournal.com/healthcare-data-breach-statistics/
2. Ayers, E. (2021). Healthcare data breaches jumped 55% in 2020; cyber threats continue to worsen. *Advisen Cyber FPN*.
3. Brookmeyer, R and Liao J (1990). The analysis of delays in disease reporting: Methods and results for the acquired immunodeficiency syndrome. *American Journal of Epidemiology*, 132(2):355-365.

Contact details

- E-mail: SSangar1@students.Kennesaw.edu, EDallal@air-worldwide.com, SStransky@air-worldwide.com

Abstracts/Session_2B

Risk Budgeting Portfolios from Simulations

Bernardo F. P. da Costa, Instituto de Matemática, Universidade Federal do Rio de Janeiro (UFRJ)

Silvana Pesenti, Department of Statistical Sciences, University of Toronto

Rodrigo S. Targino (presenter), Escola de Matemática Aplicada (EMAp), Fundação Getulio Vargas (FGV)

Abstract: Large pension plans face the difficulty of investing premiums in a financially prudent but economically efficient way. An investment concept used in industry, in particular by pension funds, is to construct *risk budgeting portfolios* (RBPs). RBPs are diversified portfolios where the diversification pertains to the value of the risk contribution of each asset to the total portfolio loss. The risk contribution of an asset to the portfolio loss is the added risk of marginally increasing the portfolio's position in that asset. Hence, RBPs are diversified in terms of the risk stemming from each asset and therefore ideal for pension plans which face stringent investment requirements. Advances in the academic literature and practical implementations of RBP strategies, however, make the strong and unrealistic assumption that assets' returns follow a multivariate Gaussian distribution. We propose an efficient stochastic optimization framework that calculates RBPs for assets with arbitrary joint distributions. For coherent risk measures, such as the expected shortfall, the resulting model is convex, allowing for very efficient algorithms. Furthermore, the RBP is constructed using only sampled scenarios, making it applicable in a variety of settings such as data-driven statistical models, arbitrary dependency structures, or company internal loss models. and the resulting algorithms are applicable to non-Gaussian losses.

Keywords: Risk Budgeting Portfolios (RBPs), Stochastic Optimization, Non-gaussian losses

References

1. Roncalli, T. (2013). *Introduction to Risk Parity and Budgeting.* CRC Press
2. Shapiro, A., Dentcheva, D. and Ruszczyński, A. (2009). *Lectures on Stochastic Programming, Modeling and theory.* MPS-SIAM.
3. Tasche, D. (1999). *Risk contributions and performance measurement*, **Report of the Lehrstuhl für mathematische Statistik, TU München.**

Contact details

- Email: rodrigo.targino@fgv.br
- Homepage: https://rtargino.netlify.app/

Modelling Credit Structures and Securitisations with Data Science

Fernando Mierzejewski, AG Insurance (presenter)

Abstract: Although the use of credit structured products and securitisations may have been reduced after the credit crisis of 2007-2008, they remain as a major source of funding to the financial system. Yet, as highly customised contracts, there is no consensus about the method to derive the price adopted in actual transactions. Main players in the industry, such as credit rating agencies, investment banks, and institutional investors, have tended to rely on valuation methods based on complex and opaque mathematical modeling, strongly dependent on qualitative assessment, and thus prone to sudden adjustments. This has been raised as a main factor causing the crisis episode (IMF, 2008).

Major challenges faced when pricing structured products pertain the big amount of data to be processed in short time periods, and the way to properly combine the output of mathematical models with qualitative aspects raised by expert judgement and management requirements. We claim that the Data Science theoretical framework provides with a natural environment to deal with these two challenges. A Data Science frameword is proposed to describe the problem of the management of these products, aiming to output reliable estimations of main variables involved in the process, namely, the probability of default and thus the credit rating attached to the portfolio of collateral assets, the rating of the component credit tranches, and the cost at issuance.

Keywords: structured finance, securitisations, credit ratings, data science

References

1. International Monetary Fund (2008). *Global Financial Stability Report. Containing Systemic Risks and Restoring Financial Soundness.* https://www.imf.org/en/Publications/GFSR/Issues/2016/12/31/Containing-Systemic-Risks-and-Restoring-Financial-Soundness
2. Lancaster, B.P., Schultz, G.M., Fabozzi, F.J. (2008). *Strucured Products and Related Credit Derivatives.* Wiley & Sons Inc.
3. Murphy, K. (2012). *Machine Learning: A Probabilistic Perspective.* MIT Press.

Contact details

- Email: femierze@gmail.com
- https://be.linkedin.com/in/fernando-mierzejewski-phd-b098604

Deep Hedging of Long-Term Financial Derivatives

Alexandre Carbonneau, Concordia University - PhD student (presenter)

Abstract: This study presents a deep reinforcement learning approach for global hedging of long-term financial derivatives. A similar setup as in [1] is considered with the risk management of lookback options embedded in guarantees of variable annuities with ratchet features. The deep hedging algorithm of [2] is applied to optimize neural networks representing global hedging policies with both quadratic and non-quadratic penalties. To the best of the author's knowledge, this is the first paper that presents an extensive benchmarking of global policies for long-term contingent claims with the use of various hedging instruments (e.g. underlying and standard options) and with the presence of jump risk for equity.

Monte Carlo experiments demonstrate the vast superiority of non-quadratic global hedging as it results simultaneously in downside risk metrics two to three times smaller than best benchmarks and in significant hedging gains. Analyses show that the neural networks are able to effectively adapt their hedging decisions to different penalties and stylized facts of risky asset dynamics only by experiencing simulations of the financial market exhibiting these features. Numerical results also indicate that non-quadratic global policies are significantly more geared towards being long equity risk which entails earning the equity risk premium. Results presented are from the paper [3].

Keywords: Reinforcement learning; Global hedging; Variable annuity; Lookback option; Jump risk.

References

[1] Coleman, T., Kim, Y., Li, Y., and Patron, M. (2007). Robustly hedging variable annuities with guarantees under jump and volatility risks. *Journal of Risk and Insurance*, **74(2)**, 347-376.

[2] Buehler, H., Gonon, L., Teichmann, J. and Wood, B. (2019). Deep hedging. *Quantitative Finance*, **19(8)**, 1271-1291.

[3] Carbonneau, A. (2020). Deep Hedging of Long-Term Financial Derivatives. arXiv preprint arXiv:2007.15128.

Monte Carlo Valuation of Future Annuity Contracts

Fabio Viviano, University of Udine and University of Trieste (presenter)

Anna Rita Bacinello, University of Trieste

Pietro Millossovich, University of Trieste and The Business School (formerly Cass)

Abstract: The valuation of future annuity contracts is getting ever-more prominent since it is implicitly present in many contexts (e.g. pricing guaranteed annuity options or pension de-risking strategies such as pension buy-ins and buy-outs). To overcome the straightforward and time consuming approach based on nested simulations, which is frequently used by practitioners (see Boyer & Stentoft, 2012), several alternatives were proposed in literature.

For instance, Denuit (2008) suggests comonotonic approximations of the life annuity conditional expected present value, while Dowd et. al. (2011) propose a Taylor series approach. In this regard, we present an application of the well-established Least-Squares Monte Carlo method which was firstly proposed in literature by Longstaff & Schwartz (2001) in the context of American-type option pricing.

The idea is to approximate the involved conditional expectations through a linear combination of some basis functions depending on relevant risk-factors influencing the value of the desired contract. Then, the coefficients are estimated by means of an Ordinary Least Squares procedure. This allows to drastically reduce the computational effort needed (since a lower number of simulations are required) and preserves the accuracy of the desired estimates. This methodology is implemented in the `R` function `calculate.Annuity` which provides tools for pricing future annuity contracts for individuals aged x at a future time T. The function is connected with the well-known and widely used `StMoMo` `R` package. Moreover, it gives also the possibility to exploit self-customed mortality data object and to include stochastic interest rate models. The returning class of such a function is of the type `sim.Annuity` on which it is possible to call other basic functions to produce quantitative and qualitative analysis of the output; for instance, `plot`, `hist`, `summary`, `quantile`, etc.

Keywords: LSMC, Life annuities, Longevity risk, Stochastic mortality

References

1. Boyer, M., Stentoft, L. (2012). If we can simulate it, we can insure it: an application to longevity risk management. *Insurance Mathematics and Economics* **52**, 35-45.
2. Denuit, M. (2008). Comonotonic approximations to quantiles of life annuity conditional expected present value. *Insurance: Mathematics and Economics* **42**, 831–838.
3. Dowd, K., Blake, D., Cairns, A. J. G. (2011). A computationally efficient algorithm for estimating the distribution of future annuity values under interest-rate and longevity risks. *North American Actuarial Journal* **15**, 237–247.
4. Longstaff, F. A., Schwartz, E. S. (2001). Valuing american options by simulation: a simple least-squares approach. *The Review of Financial Studies* **14**, 113–147.

Contact details

- Email: bacinel@units.it (Anna Rita Bacinello)
- Email: pietro.millossovich.1@city.ac.uk (Pietro Millossovich)
- Email: viviano.fabio@spes.uniud.it (Fabio Viviano)

From local to structural input importance in variable annuities data

Giovanni Rabitti, Heriot-Watt University (presenter)

Emanuele Borgonovo, Bocconi University

Elmar Plischke, Clausthal University of Technology

Abstract: Variable annuities are special life insurance policies which include financial guarantees. Since these embedded guarantees can be very complex, the pricing model for the fair value of these variable annuity contracts cannot be written in closed form. Consequently, the fair values are typically calculated using black-box simulations.

In this work we want to identify the most important inputs which determine the fair values of these contracts. Considering the variable annuity benchmark dataset of Gan and Valdez (2017), we first introduce the local importance indices based on the finite-difference decomposition. Then we move to the global approach considering the uncertainty over the inputs, modeled by the joint input distribution. Since the inputs are dependent, we consider the structural importance indices to eliminate the impact of spurious effects driven by this dependence (Li et al., 2010). With these structural indices the analysis can better interpret the simulator output (Borgonovo et al., 2020), i.e. the fair value of the variable annuity. Moreover, we present the connection between the local and the global approaches. Finally, we compare our results on the input importance rankings with those of Gan et al. (2018).

Keywords: Sensitivity analysis, Dependent inputs, Structural effects

References

1. Borgonovo, E., Plischke, E., Rabitti, G. (2020). Interactions and Computer Experiments. *submitted*.
2. Gan, G., Quan, Z. Valdez, E. A. (2018). Machine Learning and Its Application in Variable Annuity Valuation. In *Proceedings of the 2018 SIAM International Conference on Data Mining*.
3. Gan, G., Valdez, E. A. (2017). Valuation of large variable annuity portfolios: Monte Carlo simulation and synthetic datasets. *Dependence Modeling* **5**, 354-374.
4. Li, G. et alii (2010). Global Sensitivity Analysis for Systems with Independent and/or Correlated Inputs. *The Journal of Physical Chemistry A* **114**, 6022-6032.

Contact details

- Email: g.rabitti@hw.ac.uk

Metamodeling for Variable Annuity Valuation: What works and what does not

Xiaochen Jing, University of Wisconsin-Madison

Abstract: Variable Annuities have become popular retirement products with various options of guarantees, but their complex design also makes liability management a difficult task for insurers. There have been several dozen papers published in the past years on exploring the use of statistical learning and metamodeling approaches for Variable Annuity valuation and risk management in the actuarial science and quantitative finance literatures. However, they all focus on specific techniques in the context of synthetic data. In this paper, I investigate the effectiveness of metamodeling approaches with different experimental designs and metamodels with real-world Variable Annuity contracts. In particular, I use textual analysis to extract and formulate value-related features and develop a flexible and comprehensive simulation-based scheme for Variable Annuity valuation. I find that (1) real-world variable annuity contracts are very complex and the intricate relations between their valuation and features are difficult to obtain. And (2) the overall performance of a metamodeling method depends on the employed machine learning methods as well as the sample size—though not substantially on the sampling methods. Both improve performance at the cost of longer runtime.

Keywords: Variable Annuities, Textual Analysis, Metamodeling, Machine Learning

References

1. Gan, G. (2013). Application of Data Clustering and Machine Learning in Variable Annuity Valuation. *Insurance: Mathematics and Economics*, **53**, 795-801.
2. Gan, G. and Valdez, E. (2016). An Empirical Comparison of Some Experimental Designs for the Valuation of Large Variable Annuity Portfolios. *Dependence Modeling*, **4(1)**, 382-400.
3. Gan, G. and Valdez, E. (2019). *Metamodeling for Variable Annuities.* Chapman and Hall/CRC Press.
4. Lin, S. and Yang, S. (2020). Fast and efficient nested simulation for large variable annuity portfolios: A surrogate modeling approach. *Insurance: Mathematics and Economics*, **91**, 85-103.

Contact details

- Email: xjing6@wisc.edu

Abstracts/Session_3A

Patterns and Anomalies of Loss Development in P&C Insurance Market

Arthur Charpentier, Université du Québec à Montréal

Steve Guo, Ball State University (presenter)

Michael Ludkovski, University of California, Santa Barbara

Abstract: We analyze loss development in NAIC Schedule P loss triangles using functional data analysis methods. Relying on robust principal component analysis (RPCA), we study the incremental loss ratio curves of workers' compensation lines across hundreds of companies and 24 years. RPCA helps us to find out patterns of loss development, including (i) identifying outlier loss triangles; (ii) providing a dimension reduction tool to interpret the functional loss development data via a few factors. As one example of a relevant insight, we document distinctive loss development patterns between the late 1980s, 1990s and late 2000s periods. Moreover, our approach provides novel visualization tools. In the latter part of the article, we propose a functional model for generating probabilistic forecasts of incomplete cumulative loss ratio curves based on historical and similar development patterns.

Keywords: Loss triangle; loss development; loss reserving; loss ratio; unsupervised learning; functional data; robust principal component analysis; outlier detection; IBNR

References

1. Croux, Christophe, Peter Filzmoser, and Maria Rosario Oliveira (2007), "Algorithms for projection–pursuit robust principal component analysis." Chemometrics and Intelligent Laboratory Systems, 87, 218–225.
2. Hyndman, Rob J and Han Lin Shang (2010), "Rainbow plots, bagplots, and boxplots for functional data." Journal of Computational and Graphical Statistics, 19, 29–45.
3. Rousseeuw, Peter J, Ida Ruts, and John W Tukey (1999), "The bagplot: a bivariate boxplot." The American Statistician, 53, 382–387.
4. Shang, Han Lin (2013), "Functional time series approach for forecasting very short-term electricity demand." Journal of Applied Statistics, 40.
5. Shi, Peng and Brian M Hartman (2016), "Credibility in loss reserving." North American Ac- tuarial Journal, 20, 114–132.

Contact details

- Email: qguo@bsu.edu
- Homepage: https://sites.google.com/site/qguosteve

Gradient Boosting Machines in Collective Reserving Model for Reserves Prediction

Henning Zakrisson, Stockholm University

Abstract: In this presentation we consider regression based reserving models that allow for separate RBNS and IBNR reserves based on aggregated discrete time data containing information for different combinations of accident years, reporting delay, and payment delay. The models that are analysed either describe both claim count and payment amount dynamics, or only payment amount dynamics.

All introduced models will be closely related to the cross-classified over-dispersed Poisson (ODP) chain-ladder model.

Further, these general ODP models will be estimated using regression functions defined by tree-based gradient boosting machines (GBM). This will provide us with machine learning based reserving models that have interpretable output, and that are easy to bootstrap from. We will give a brief introduction to GBMs, including basic calibration and model selection, and illustrate the reserve performance based on complex simulated data.

The presentation is based on joint work with M. Lindholm, R. Verrall, and F. Wahl.

Keywords: Claims reserving, Reported But Not Settled Claims, Incurred But Not Reported Claims, Gradient Boosting Machines

References

1. Wahl. F, Lindholm M, Verall R. (2019) The collective reserving model. *Insurance: Mathematics and Economics* **87**, 34-50
2. Lindholm M, Verall, R., Wahl, F., Zakrisson, H. (2020) Machine Learning, Regression Models and Prediction of Claims Reserves. *CAS E-forum Summer*

Contact details

- Email: zakrisson@math.su.se
- LinkedIn: https://www.linkedin.com/in/henningzakrisson/

Hierarchical Compartmental Reserving Models

Markus Gesmann, Insurance Capital Markets Research

Abstract: Hierarchical compartmental reserving models provide a parametric framework for describing the high-level business processes driving claims development in insurance using differential equations.

We will discuss how those models can be presented in a fully Bayesian modelling framework for the aggregated claims settlement process to capture trends observed in paid and outstanding claims development data reflecting the random nature of claims and latent underlying process parameters.

We show how the experienced modeller can utilise her expertise to describe the volatility of the underlying risk exposure profile and uncertainty on prior parameter assumptions and highlight in particular the subtle, but important difference between modelling incremental and cumulative claims payments.

Keywords: Multi-level models, Bayes, Dynamical Systems, Reserving

References

[1] Gesmann, M., and Morris, J. "Hierarchical Compartmental Reserving Models." Casualty Actuarial Society, CAS Research Papers, 19 Aug. 2020, https://www.casact.org/research/research-papers/Compartmental-Reserving-Models-GesmannMorris0820.pdf

Contact details

- Email: markus.gesmann@insurancecapitalmarkets.com
- Homepage: https://insurancecapitalmarkets.com

Abstracts/Session_3B

Person at the center: how AutoML boosts behavioral modelling in P&C insurance

Veronica Coronel Vera, Munich Re (presenter)

Davide Burlon, Munich Re

Massimo Cavadini, Munich Re

Abstract: Risk and behavioural pricing are based on a number of features which are either collected or engineered in order to be able to -firstly- project the cost of claims an insurer has to expect in the future, and -secondly- to weigth this by the likelihood of contributing to an insurance top/bottom line. This sets the basis for solid pricing of P&C portfolios. The result of this technical exercise are price changes which reflect a customer's risk quite precisely. But how does an existing customer react to such price changes? And how do competitors prices influence a potential customer? How do the respective models interact with models projecting cost of claims? The scope of the seminar is to show how applied actuarial science is reacting to the changing times, for example by incorporating implementable AutoML approaches to risk and behavioral pricing.

Keywords: demand, elasticity, AutoML

References

1. Burlon, D., Cavadini, M. (2020) Pricing Survey H12020. https://www.munichre.com/content/dam/munichre/contentlounge/website-pieces/documents/Global_consulting_Pricing_survey_2020-09-09.pdf/_jcr_content/renditions/original./Global_consulting_Pricing_survey_2020-09-09.pdf

Contact details

- Email: VCoronelVera@munichre.com
- LinkedIn: https://www.linkedin.com/in/veronica-coronel-vera
- Email: dburlon@munichre.com
- Linkedin: https://www.linkedin.com/in/davideburlon/

A Twin Neural Model for Uplift

Mouloud Belbahri, University of Montreal & TD Insurance

Abstract: Uplift modeling is a particular case of causal inference. Such models deal with cause-and-effect inference for a specific factor, such as a marketing intervention. In practice, these models are built on individual data from randomized experiments where the goal is to partition the participants into heterogeneous groups depending on the predicted uplift. Most existing approaches are adaptations of random forests for the uplift case. Several split criteria have been proposed in the literature, all relying on maximizing heterogeneity. However, in practice, these approaches are prone to overfitting. In this work, we bring a new vision to uplift modeling. We propose a new loss function which can be interpreted as an augmented Binomial likelihood. Our solution is developed for a specific twin neural network architecture allowing to jointly optimize the uplift as well as marginal probabilities of success for treated and control individuals. We show that this model is a generalization of the uplift logistic interaction model and modify the stochastic gradient descent algorithm to allow for structured sparse solutions. This helps training our uplift models to a great extent. We show our proposed method is competitive with the state-of-the-art in simulation setting and on a real data from large scale insurer's marketing randomized experiment.

Keywords: causal inference, heterogeneous treatment effects, loss functions, gradient descent, regularization.

References

1. Belbahri, M., Gandouet, O., Kazma, G. (2020). Adapting neural networks for uplift models. *arXiv preprint arXiv:2011.00041*.
2. Belbahri, M., Murua, A., Gandouet, O. and Partovi Nia, V. (2019). Qini-based uplift regression. *arXiv preprint arXiv:1911.12474*.
3. Künzel, S.R., Sekhon, J.S., Bickel, P.J. and Yu, B. (2019). Meta-learners for estimating heterogeneous treatment effects using machine learning. *In Proceedings of the National Academy of Sciences, 116(10):4156–4165*.
4. Nie, X., Wager, S. (2017). Quasi-oracle estimation of heterogeneous treatment effects. *arXiv preprint arXiv:1712.04912*.

Contact details

- Email: mouloud.belbahri@tdassurance.com

The use of denoising autoencoders for categorical and continuous variables

Anna Kozak, Quantee, R&D Department (presenter)

Łukasz Delong, SGH Warsaw School of Economics, Institute of Econometrics, Quantee, R&D Department

Maciej Tomaszewski, Quantee, R&D Department

Abstract: In this presentation we discuss the use of denoising autoencoders as a tool for reducing the dimension of data in unsupervised problems and for representation of explanatory variables in supervised problems. We consider categorical and continuous variables. For categorical variables, we consider autoencoders based on a neural network fitted to all categorical variables with two types of softmax function acting on all variables and on individual variables. We test different types of noise for categorical data which could be applied to data matrix. For a supervised objective, we compare the results with the entity embedding which has gained the most popularity in insurance applications. For continuous variables, we also consider autoencoders based on neural networks but we replace the mean-square loss, used to measure the replication error of the decoder, with the Gamma deviance loss for skewed variables. We also test different types of noise which could be applied to data matrix with continuous covariates. The experiments and their results are illustrated on an actuarial data set.

Keywords: denoising autoencoders, softmax, Gamma deviance, representation learning, dimension reduction

References

1. Wuthrich, Mario. (2019). From Generalized Linear Models to Neural Networks, and Back. SSRN Electronic Journal.
2. Pascal Vincent, Hugo Larochelle, Isabelle Lajoie, Yoshua Bengio, and Pierre-Antoine Manzagol. (2010). Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion. J. Mach. Learn. Res. 11 (3/1/2010), 3371–3408.

Contact details

- Email: anna.kozak@quantee.ai
- Social media: https://www.linkedin.com/in/kozakanna

Abstracts/Session_4A

Approximate Bayesian Computation and Insurance

Patrick Laub, University of Melbourne (presenter)

Pierre-Olivier Goffard, Université Lyon 1

Abstract: In this talk, I will give an overview of Approximate Bayesian Computation (ABC). ABC is a statistical learning technique to calibrate and select models by comparing observed data to simulated data. This technique bypasses the use of the likelihood and requires only the ability to generate synthetic data from the models of interest. In work with Pierre-Olivier Goffard (ISFA, Lyon), we apply ABC to fit and compare insurance loss models using aggregated data.

ABC is extremely computationally demanding, and the process of making our optimised Python implementation was non-trivial. If time permits, I would briefly mention some of the strategies I used to create an efficient ABC implementation, which can run on any OS, and can utilise the full power of a large high-performance computer or a simple laptop.

Keywords: Approximate Bayesian Computation, censored data, likelihood-free statistics, model selection, Python

References

1. Pierre-Olivier Goffard, Patrick J. Laub (2021), *Approximate Bayesian Computations to fit and compare insurance loss models*, Insurance: Mathematics and Economics (under revision), available at https://arxiv.org/abs/2007.03833

Contact details

- Email: patrick.laub@unimelb.edu.au
- Homepage: https://pat-laub.github.io/
- Social media: https://twitter.com/PatrickLaub

Introducing the D programming language to Insurance

Chibisi Chima-Okereke, MEng, MSc, PhD, Active Analytics (presenter)

Abstract: D is a static compiled programming language created in 2001. At the point of inception, the language was proposed as a modernised and more productive version of C++, retaining the performance profile of C++ but having much improved capabilities for templates and metaprogramming. The current motto of the language is "write fast, read fast, run fast", referring to its simple but yet familiar style of syntax and inference, built-in arrays and range semantics, a form on screen that is easy to read, the fact that the reference compiler is particularly fast, and compiles to efficient native code, giving the language great performance characteristics.

D has built-in unit testing features, full C and some C++ compatibility, Fortran subroutines can also be called. It ships with Dub, D's official and very easy to use package manager. D has a lively and helpful community, as well as a fantastic introductory text "Programming in D" by Ali Çehreli.

D is being used in companies large and small, from companies such as Mercedes Benz where it is used for software development tools, through companies such as Facebook, ebay, and Netflix where it is used to build a deep modelling framework VectorFlow. Many more companies use D in varied, interesting, and highly productive, rewarding, and profitable endeavours.

D has an embarrassment of riches and this presentation is a flying visit of some of its features, which might provide useful for data science functions in the insurance industry.

Keywords: D, Programming Language

References

1. Ali Çehreli, edited by Luís Marques (2021). Programming in D.
2. The D programming language website, https://dlang.org/
3. Walter Bright, Andrei Alexandrescu, Michael Parker, Origins of the D programming language, Proceedings of the ACM on Programming Languages, June, 2020 Article No. 73

Contact details

- Email: chibisi@active-analytics.com
- Homepage: active-analytics.com
- Repository: https://github.com/ActiveAnalytics
- Social media: https://twitter.com/aanalyticsltd

Month arithmetic in R with the mondate package

Daniel Murphy, FCAS, MAAA, Trinostics LLC

Abstract: `mondate` is an R package that measures dates in units of months to fill a gap in the R environment. Inspired by the paper "Time Calculations for Annualizing Returns: the Need for Standardization" (Damien Laker, in *The Journal of Performance Measurement*, 2008), `mondate` interprets a date as the fractional number of elapsed calendar months since December 31, 1999 (the "origin"). Month-arithmetic is essential when aging data from accounting records evaluated as of month- and year-ends.

This talk will quickly look at month arithmetic for R's `POSIX` and `Date` objects, why R's `difftime` has no `units = "months"`, and how a few `mondate` functions can simplify an actuary's life.

Keywords: Date, POSIXt, evaluation date

References

1. Laker, Damien (2008). Time Calculations for Annualizing Returns: the Need for Standardization. *The Journal of Performance Measurement*, 2008.

Contact details

- Email: chiefmurphy@gmail.com
- Homepage: trinostics.com
- CRAN: https://CRAN.R-project.org/package=mondate
- Repository: https://github.com/chiefmurph/mondate
- Linkedin: https://www.linkedin.com/in/danmurphy123/

Agile, the right answer for the 'next normal' in the insurance sector

Francesca Vitalini, Mirai Solutions GmbH (presenter)

Stephanie Gehring, Mirai Solutions GmbH

Abstract: To succeed in the new, fast-changing, digital age, even the traditionally hierarchical and heavily regulated business of insurance has to rethink itself in more flexible, technologically advanced and agile terms. With the disruption led by the Covid-19 pandemic highlighting the need to accelerate the change, and increasing the demand for a structural shift, it has become even more evident that the 'next normal' for the insurance sector will have to be agile.

Often starting at team-level, mainly in IT and data science projects, agile practices have successfully found their way into the insurance world, to ensure product innovation, meet customer needs, and shorten time to market. However, many insurance companies are still struggling to undergo a full agile transformation, finding it difficult to adopt the agile mindset and practices in non-IT projects, as well as embracing the new culture at organizational level.

In this presentation we would like to discuss the challenges the insurance industry faces today and how agile is the right response to them. We will analyse how an agile cross-functional team in insurance data science looks like and how such a team approaches product development. Finally, we will identify the key factors in the long journey of enterprise-wide transformation towards an agile organization.

Keywords: Agile, Digital Transformation

References

Contact details

- Email: francesca.vitalini@mirai-solutions.com
- Homepage: https://mirai-solutions.ch/
- Social media:
  - Linkedin:
    - personal: https://www.linkedin.com/in/francesca-vitalini-phd-csm-61896034/
    - company: https://www.linkedin.com/company/mirai-solutions-gmbh/
  - Twitter:
    - personal: @Pisita11
    - company: @MiraiSolutions

Processing Insurance Claims with Automated, Scalable and Fair AI

Kenneth Lim, International Business Machines (IBM) (presenter)

Maxime Allard, Imperial College London & International Business Machines (IBM)(presenter)

Abstract: Processing claims are a major cost driver for insurers, accounting for 30% of operating costs on average [1]. Inefficient claims processing leads to claims leakage, which is money lost from spending more on resolving a claim than necessary, through decision making errors and additional touchpoints. Insurers can use AI to improve efficiencies and reduce costs. However, longstanding ethical and scalability concerns have hindered AI adoption in the industry.

We present an end-to-end AI pipeline architecture motivated by real non-life insurance projects to process claims more efficiently and minimise claims leakage. It optimises the allocation of thousands of claims to handlers with state-of-the-art decision optimisation models and produces high-quality automated predictive models with AutoAI[2], which address the scalability concerns. The predicted outcomes are also explainable and debiased, addressing the ethical concerns[3].

We contribute to the literature through the novel application and modification of academic data science research (e.g. [2],[3],[4]) to real-world non-life insurance problems, addressing longstanding ethical and scalability concerns in the insurance industry.

Keywords: Claims Leakage, AutoAI, Explainable AI, Decision Optimisation

References

1. McKinsey & Co. 2016. From transparency to insights: McKinsey's insurance cost benchmarking 2016. Report, https://www.mckinsey.com/industries/financial-services/our-insights/from-transparency-to-insights-mckinseys-insurance-cost-benchmarking-2016
2. Dakuo Wang, Parikshit Ram, Daniel Karl I. Weidele, Sijia Liu, Michael Muller, Justin D. Weisz, Abel Valente, Arunima Chaudhary, Dustin Torres, Horst Samulowitz, and Lisa Amini. 2020. AutoAI: Automating the End-to-End AI Lifecycle with Humans-in-the-Loop. In Proceedings of the 25th International Conference on Intelligent User Interfaces Companion (IUI '20). Association for Computing Machinery, New York, NY, USA, 77–78. DOI:https://doi.org/10.1145/3379336.3381474
3. P. K. Lohia, K. Natesan Ramamurthy, M. Bhide, D. Saha, K. R. Varshney and R. Puri, "Bias Mitigation Post-processing for Individual and Group Fairness," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 2019, pp. 2847-2851, doi: 10.1109/ICASSP.2019.8682620.
4. Sood, A., Elder, B., Herta, B., Xue, C., Bekas, C., Malossi, A.C.I., Saha, D., Scheidegger, F., Venkataraman, G., Thomas, G. and Mariani, G., 2019. Neunets: An automated synthesis engine for neural network design. arXiv preprint arXiv:1901.06261.

Contact details

- Email: kenneth.lim@ibm.com, maxime.allard3@ibm.com
- Social media: https://uk.linkedin.com/in/kennethtklim, https://uk.linkedin.com/in/maxime-allard, https://twitter.com/allardmaxime079

Investigating Applications of Data Science in UK and non-UK Actuarial Teams

Xavier Marechal, CEO at Reacfin (Presenter)

Anja Friedrich, Manager at Synpulse Management Consulting (Presenter)

Abstract: Significant changes in technology, regulation, markets, customer behaviour, the environment and other global trends are influencing the actuarial department. The increasing availability of big data, The availability of technical data science skills, and the application thereof; are changing how insights are being derived and continuing to shape the operating model of the actuarial department.

We performed benchmarking exercises which involved structured interviews with senior first line actuarial department representatives from different UK, South Africa, Belgium, Luxembourg & Switzerland life and non-life insurance organisations to investigate how the insurance industry is utilising data science, with a focus on application and use cases within an actuarial context. We investigated the strategy and the operating model within which data science is used including the types of tools and techniques being used.

Within our benchmarking exercise we also included themes around the types of data; the technical nature of machine learning techniques and software being used; and wider considerations including risks, risk management, governance, and ethics related to data science.

We investigated trends impacting the skill set required by those working within Data Science and the barriers to adopting data science.

We interviewed representatives from first line actuarial departments mainly with Heads of Actuarial Reporting and Pricing Departments, Heads of Actuarial Systems and Heads of Actuarial Transformation and Strategy including direct insurance organisations and group entities.

Our talk will summarise the findings from this Actuarial Data Science benchmarking exercise.

Our talk will go further and explore, based on the above investigations of use of data science in actuarial findings the actionable steps and recommendations to optimise the use of data science within insurance and non-insurance industries.

We will explore the changes that we are expecting to see in order to extract value and how those changes could be managed and implemented.

Keywords: Value from data science, barriers to adopting data science, machine learning techniques, actuarial use cases

Contact details

Xavier Marechal

- Email: xavier.marechal@reacfin.com. Homepage: https://www.reacfin.com

Anja Friedrich

- Email: anja.friedrich@synpulse.com. Homepage: https://www.synpulse.com

Abstracts/Session_4B

Gompertz network and Lasso regularisation in modelling age-specific impact of COVID-19 vaccination

John Ng, Reinsurance Group of America (RGA)

Abstract: The U.K. launched the world's first mass COVID-19 vaccination on 8th December 2020. This research aims to estimate the mortality impact of vaccination in England by applying machine learning, epidemiological and actuarial techniques to publicly available data.

COVID-19 mortality rates by age in England followed a log-linear pattern consistently throughout 2020, and historical data shows the mortalities of the younger age groups are great predictors of the mortalities of the 80+ age group. Under the counterfactual "unvaccinated" scenarios, the expected mortality of the 80+ age group could be predicted from the actual mortality experience of the largely unvaccinated younger age groups, whose Gompertz projections are used to construct the weighted average prediction in our proposed age-structured mortality model, known as 'Gompertz Network'. Lasso regularisation was applied to handle multi-collinearity. Then, the mortality impact of 80+ age group is estimated from actual versus expected analysis.

The vaccination is estimated to have prevented 8,000 deaths by end of February 2021. Mortality impact studies by Warwick Universtiy and Public Health England also adopted counterfactual approaches, and reported broadly consistent results. However our Gompertz Network model is based on real world data and does not rely on assumptions for vaccine uptake and effectiveness, hence yielding empirical evidence that the vaccines are already preventing deaths at scale in England.

Keywords: Gompertz Network, Lasso, COVID-19 Vaccination, Mortality Impact

References

1. Ng, J., Reid, S. (1 Mar 2021). Mortality Impact of COVID-19 Vaccination in England: Counterfactual Insights from Gompertz to Machine Learning. The Institute and Faculty of Actuaries.
2. Ng, J., Reid, S. (29 Mar 2021). Mortality Impact of COVID-19 Vaccination in England: Publication of Models and an update. The Institute and Faculty of Actuaries.
3. Public Health England (26 Mar 2021). Impact of COVID-19 vaccines on mortality in England: December 2020 to February 2021. GOV.UK.
4. Moore, S., Hill E.M., Tildesley M.J., Dyson L., Keeling M.J. (18 Mar 2021). Vaccination and non-pharmaceutical interventions for COVID-19: a mathematically modelling study. The Lancet Infectious Diseases (Online).

Contact details

- Email: wui_hua@cantab.net
- Social media: Linkedin https://www.linkedin.com/in/wui-hua-ng/

Spreading the word: The effect of word-of-mouth networks on insurance customer choices

Rei England, The Business School (formerly Cass), City, University of London (presenter)

Iqbal Owadally, The Business School (formerly Cass), City, University of London

Douglas Wright, The Business School (formerly Cass), City, University of London

Abstract: Attracting and retaining loyal customers is a key driver of insurance profit. Though price remains the main factor in customer choice, customer service experience is a significant part of customer decision-making. Customers are also highly influenced by recommendations from friends (Ghodrati and Taghizad, 2014). These word-of-mouth networks allow information to spread between customers in a market. Berger (1988) uses a model of an insurance market where customers who have a bad experience switch according to recommendations from their friends, but is not able to model the network explicitly and does not include the effect of price changes. Agent based models (ABMs) use simulations to model systems of interacting heterogeneous agents. Owadally et al (2019) use this method to investigate the impact of regulatory regimes on insurance market cycles that arise endogenously from market competition. We combine their work with models of opinion transmission (Bianchi and Squazzoni, 2015) to build an ABM with a small-world word-of-mouth network. Insurers choose how much to spend on their service quality, and customers evaluate insurers based on premium, brand preference, and their perceived service quality. We find that the presence of the network produces a systemic bias in customer assessment of insurer quality. Early experiences tend to persist as they are passed continually around clusters, resulting in a distribution of insurer market share similar in shape to real-world data. This suggests that newer insurers may benefit more from a higher service quality as they build their reputation. The FCA report on renewals behaviour (FCA, 2020) suggests a regulatory change which would prevent insurers from charging renewing customers differently than new customers. The model is run with this scenario to explore the effect on renewal rates, customer satisfaction, and the implications for customer service quality. Preliminary results suggest an increase in market concentration. Additionally renewal rates increase significantly, though renewal decisions become more sensitive to changes in customer opinion and word-of-mouth information.

Keywords: word-of-mouth, ABM, networks, customer service, renewal

References

1. Berger, L. A. (1988). Word-of-mouth reputations in auto insurance markets. *Journal of Economic Behavior & Organization*, **10(2)**, 225–234.
2. Bianchi, F., and Squazzoni, F. (2015). Agent-based models in sociology. *Wiley Interdisciplinary Reviews: Computational Statistics*, **7(4)**, 284–306.
3. FCA. (2020). General insurance pricing practices: Final report. *Financial Conduct Authority*, **Vol. MS18**; Tech. Rep. No. 1.3
4. Ghodrati, H., and Taghizad, G. (2014). Service quality effect on satisfaction and word of mouth in insurance industry. *Management Science Letters*, **4(8)**, 1765–1772.
5. Owadally, I., Zhou, F., Otunba, R., Lin, J. and Wright, D. (2019). An agent-based system with temporal data mining for monitoring financial stability on insurance markets. &Expert Systems with Application*s, **123**, 270-282.

Contact details

- Email: rachael.hampden-turner@city.ac.uk

A comparative study of using various Machine Learning and Deep Learning based fraud detection models for Universal Health Coverage schemes and assessing the impact of COVID-19 in healthcare fraud

Rohan Yashraj Gupta, Sri Sathya Sai Institute of Higher Learning (SSSIHL)

Abstract: Fraud poses major risk in any healthcare systems due to its financial implications. To mitigate this, most companies have fraud detection models in place. In recent times, use of Machine Learning and/or Deep Learning based model are being explored in a pursuit of building a better fraud detection model. In this study, we are performing a comparative study of using various Machine Learning and Deep Learning based fraud detection models for Universal Health Coverage schemes. However, building such fraud detection models requires two major challenges to be overcome – handling data-imbalance and identifying an appropriate classification model. We have tackled these challenges using three different data-imbalance techniques and six classification models, we have also used six variants of neural network models. For this, we have used data from part of the world's largest universal health coverage scheme called Ayushman Bharat (PM-JAY India). A total of 26 models were tested as part of this study. The performance of these models was measured using standard performance metrics and was identified that a neural network model trained on undersampled data performed better than other models in this study.

We have also performed an analysis of the impact of COVID-19 on healthcare fraud. For this purpose, we identified various COVID-19 specific triggers which enabled us to identify new fraud cases. We found that the month on month rate of fraudulent cases is highly correlated with month on month rate of COVID-19 cases. To determine this we calculated the Pearson correlation coefficient and fitted a logarithmic regression model between fraud in healthcare and COVID-19 cases. Our experimental analysis shows a Pearson correlation coefficient of 0.86. The logarithmic regression performed on the data gave the r-squared value of 0.91.

Keywords: Machine learning, Deep learning, Data imbalance, Actuarial techniques, Data embedding, Classification models, Ayushman Bharat, PM-JAY India, Largest universal health coverage scheme, Fraud detection framework, Pearson correlation, Logarithmic regression, COVID-19, Fraud detection

References

1. R. Y. Gupta, S. S. Mudigonda, P. K. Baruah, and P. K. Kandala (2020). "Implementation of Correlation and Regression Models for Health Insurance Fraud in Covid-19 Environment using Actuarial and Data Science Techniques," Int. J. Recent Technol. Eng., vol. 9, no. 3, pp. 699–706.
2. R. Y. Gupta, S. Sai Mudigonda, P. K. Kandala, and P. K. Baruah (2019). "Implementation of a Predictive Model for Fraud Detection in Motor Insurance using Gradient Boosting Method and Validation with Actuarial Models," IEEE International Conference on Clean Energy and Energy Efficient Electronics Circuit for Sustainable Development (INCCES) pp. 1–6
3. R. Y. Gupta, S. S. Mudigonda, and P. K. Baruah (2021). "TGANs with Machine Learning Models in Automobile Insurance Fraud Detection and Comparative Study with Other Data Imbalance Techniques," Int. J. Recent Technol. Eng., vol. 9, no. 5, pp. 236–244, Jan. 2021
4. R. Y. Gupta, S. S. Mudigonda, P. K. Kandala, and P. K. Baruah (2019). "A Framework for Comprehensive Fraud Management using Actuarial Techniques," Int. J. Sci. Eng. Res., vol. 10, no. 3, pp. 780–791.

Contact details

- Email: rohanyashrajgupta@sssihl.edu.in
- Repository: Github
- Social media: LinkedIn

Next Generation LTC/Life underwriting using 'Facial Score' model

Shrinivas Shikhare, Director Transformation Consulting Insurance Business Unit. Capgemini Financial Services UK

Abstract: Life and Long-Term Care (LTC) insurers require prospective policyholders to undergo a medical examination, as part of the underwriting process. It is imperative to understand (a) current medical conditions and (b) life expectancy of the prospects for classification, rating and pricing. It is time consuming and costly process. This happens only once at the time of inception of insurance policy which puts significant risk to insurer as health condition and score may change in lifetime.

Technical advancement in wearable technologies and AI based facial recognition techniques are pushing the boundaries of LTC and Life Insurance Product requirements such as periodic upload of photo or video file provides key inputs to facial score which can be translated to risk score for post issue underwriting or premium rebalance purpose. Base features like face detection, age, gender, BMI (empirical) and heart rate (based on Photopletysmograph) are calculated based on machine learning models to arrive 'facial score'. This is modified further with advance features like facial morphology, facial asymmetry etc. Hence moving from classical approach to Active Risk Management approach.

(Demo available for base feature including heart rate measurement using webcam)

References

[1] Moraes JL, Rocha MX, Vasconcelos GG, Vasconcelos Filho JE, de Albuquerque VHC, Alexandria AR. Sensors (Basel). *Advances in Photopletysmography Signal Analysis for Biomedical Applications*. 2018 Jun 9;18(6):1894. doi: 10.3390/s18061894. PMID: 29890749; PMCID: PMC6022166.

[2] Q. Zhao, N. Werghi, K. Okada, K. Rosenbaum, M. Summar, and L. M. Zhao. *Ensemble learning for the detection of facial dysmorphology*, International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), vol. 36, pp. 754-757, 2014.

[3] Q. Zhao, K. Okada, K. Rosenbaum, L. Kehoe, et al. *Digital facial dysmorphology for genetic screening: Hierarchical constrained local model using ICA*. Med. Image Anal., vol. 18, no. 5, pp. 699-710, 2014.

[4] R. P. Kosilek, R. Frohner, R. P. Würtz, C. M. Berr, J. Schopohl, M. Reincke, H. J. Schneider. *Diagnostic use of facial image analysis software in endocrine and genetic disorders: review*, current results and future perspectives, Eur J Endocrinol. vol. 173, no. 4, pp. 39-44, 2015

[5] T. Mutsvangwa and T. S. Douglas. *Morphometric analysis of facial landmark data to characterize the facial phenotype associated with fetal alcohol syndrome.* J Anat., vol. 210, no. 2, pp. 209-220, 2007

Contact details

- Email: Shrinivas.Shikhare@Capgemini.com

Process Mining Applied to Complex Medical Claims Management

Tim King, Insurance Industry Consultant at Teradata

Abstract: Process mining is an analytical technique with emerging prevalence within the healthcare space [1]. Beginning with defining the multi-industry business problem facing companies and governments across the globe, this presentation will illustrate with granular, publicly available medical claims data how R's bupaR and TraMineR packages can be used to accelerate fraud detection and large loss prediction. These packages are utilized to create a spectrum of process mining oriented descriptive statistics and dynamic end-user dashboards to visualize individual claim pathways and portfolio level patterns. The evolution of ICD-10 diagnoses over the life of a medical claim, for example, provides insight into anomaly detection and medical provider behavior. Visualizations are created both through "stock" functions from the aforementioned R packages as well as customized views. We will also demonstrate how trace clustering can be utilized to categorize large volumes of individual claims traces into meaningful buckets for business decision-making [2].

Keywords: Fraud detection, Large loss prediction, R

References

[1] Ian Litchfield, Ciaron Hoye, David Shukla, Ruth Backman, Alice Turner, Mark Lee, Phil Weber (2018). Can process mining automatically describe care pathways of patients with long-term conditions in UK primary care? A study protocol. *British Medical Journal*

[2] Mathilde Boltenhagen, Thomas Chatain, Josep Carmona (2019). Generalized Alignment-Based Trace Clustering of Process Behavior. *Petri Nets 2019 / ACSD 2019 - 40th International Conference on Application and Theory of Petri Nets and Concurrency, Jun 2019, Aachen, Germany. ffhal-02176771*

Contact details

- Email: tim.king@teradata.com

Correcting Reporting Delays in Cyber Events at Industry level

Seema Sangari, School of Data Science and Analytics, Kennesaw State University (Presenter)

Dr. Eric Dallal, AIR Worldwide, Verisk Cyber Solutions

Scott Stransky, AIR Worldwide, Verisk Cyber Solutions

Abstract: Having reporting delays is a well-known problem in the cyber insurance industry. With rapidly evolving cyber incidents, cyber insurers find it challenging to estimate potential future cyber claims appropriately in the absence of complete information. Healthcare data breach statistics from HIPAA accepted that cyber events remain undetected for months/years after they occur- Marriott's cyber event occurred in 2014 but was reported in 2018. Furthermore, unlike the public sector, private organizations frequently have less restrictive reporting requirements. Often private sector breaches come into picture only once they become news headlines, as they don't want to disclose incidents for fear of reputation risk. As a result, cyber data is biased and needs to be corrected to capture the real level of incidents. Current data might show a decrease in the number of events but, in reality, it's been increasing and we are just not aware of it yet. Researchers have been investigating this problem for the last few decades from different perspectives. Reporting delays are not restricted to the cyber world but the healthcare industry too – COVID-19 is classic example. One can't find out whether s/he is infected until symptoms show, which might take some time. Direct estimation of reporting delays from raw data suffers from two problems: it overestimates short delays, since recent events with long reporting delays will not yet be in data sets; and it precludes reporting delays longer than the longest reporting delay in the data set. Additionally, cyber incident data shows non-stationarity in the reporting delay distribution. We present an algorithm that deals with all three problems via estimation of a series of parametric reporting delay distributions. We validate our method by using our algorithm to estimate a one year ahead correction of cyber incident counts reported as of 2018 and comparing these to counts reported as of 2019.

Keywords: Cyber Events, Cyber Insurance, Delay Distribution, Reporting Delays

References

1. Healthcare Data Breach Statistics. https://www.hipaajournal.com/healthcare-data-breach-statistics/
2. Ayers, E. (2021). Healthcare data breaches jumped 55% in 2020; cyber threats continue to worsen. *Advisen Cyber FPN*.
3. Brookmeyer, R and Liao J (1990). The analysis of delays in disease reporting: Methods and results for the acquired immunodeficiency syndrome. *American Journal of Epidemiology*, 132(2):355-365.

Contact details

- E-mail: SSangar1@students.Kennesaw.edu, EDallal@air-worldwide.com, SStransky@air-worldwide.com

Abstracts/Session_5A

Customer Price Sensitivities in Competitive Automobile Insurance Markets

Robert Matthijs Verschuren, University of Amsterdam

Abstract: Insurers are increasingly adopting more demand-based strategies to incorporate the indirect effect of premium changes on their policyholders' willingness to stay.

However, since in practice both insurers' renewal premia and customers' responses to these premia typically depend on the customer's level of risk, it remains challenging in these strategies to determine how to properly control for this confounding. We therefore consider a causal inference approach in this paper to account for customer price sensitivities and to deduce optimal, multi-period profit maximizing premium renewal offers. More specifically, we extend the discrete treatment framework of Guelman and Guillén (2014) by Extreme Gradient Boosting, or XGBoost, and by multiple imputation to better account for the uncertainty in the counterfactual responses.

We additionally introduce the continuous treatment framework with XGBoost to the insurance literature to allow identification of the exact optimal renewal offers and account for any competition in the market by including competitor offers. The application of the two treatment frameworks to a Dutch automobile insurance portfolio suggests that a policy's competitiveness in the market is crucial for a customer's price sensitivity and that XGBoost is more appropriate to describe this than the traditional logistic regression. Moreover, an efficient frontier of both frameworks indicates that substantially more profit can be gained on the portfolio than realized, also already with less churn and in particular if we allow for continuous rate changes.

A multi-period renewal optimization confirms these findings and demonstrates that the competitiveness enables temporal feedback of previous rate changes on future demand.

Keywords: Causal inference, renewal optimization, price sensitivity, customer churn prediction, Extreme Gradient Boosting, automobile insurance.

Full working paper available online at https://arxiv.org/pdf/2101.08551.pdf.

References

Guelman, L. and Guillén, M. (2014). A causal inference approach to measure price elasticity in automobile insurance. *Expert Systems with Applications* **41/2**, 387–396.

Contact details

- Email: r.m.verschuren@uva.nl
- Homepage: https://www.uva.nl/profile/r.m.verschuren
- Social media: https://www.linkedin.com/in/robertverschuren

Number of claims and number of near-misses for telematics pricing in automobile insurance

Montserrat Guillen, Universitat de Barcelona (presenter)

Ana M. Perez-Marin, Universitat de Barcelona

Jens P. Nielsen, City, University of London

Abstract: We present a method to integrate telematics data in a pay-how-you-drive insurance pricing scheme that penalizes near-miss events. We illustrate our method with a sample of drivers for whom information on near-miss events and claims frequency records are available. We discuss the implications for motor insurance ratemaking.

New telematics devices provide vast amounts of information that create a challenge for traditional statistical techniques, or machine learning techniques (Barry and Charpentier, 2020). In this paper, we present a new methodology that breaks down telematics data into strategic components that can be used both for pricing and for attempting to alter driver behavior. By introducing the concept of "near-miss telematics", we change the focus of motor insurance from a low frequency model based on claims to a high frequency model based on near-misses, and reduce the complexity of further data analytics for data scientists, statisticians, and actuaries. Near-miss telematics offer a solution to the big data problem of modern telematics information; it also provides a new way of presenting the car insurance bill, which now resembles a phone bill. This paper considers billing on a weekly basis, but there is no reason why bill information should not be processed in real time, thus providing drivers with a direct and clear incentive to drive more safely in order to save insurance costs .

Our pricing principle is to combine a baseline insurance premium with added extra charges for near-miss events indicating risky driving (or discounts) that can be updated on a weekly basis. This procedure provides an incentive for safe driving.

In our real-case study, hard-braking and acceleration events as well as smartphone use while driving increase the cost of insurance. One area where we will explore more insights is the risk tracking of drivers via near-miss information over time. This is in line with the recent analysis proposed by Guillen et al. (2021), which offers a new perspective of the use of percentile charts for accident prevention.

Keywords: Claims frequency, Pricing, Dynamic ratemaking, Poisson model, Speed.

References

1. Barry, L., Charpentier, A. (2020). Personalization as a promise: Can Big Data change the practice of insurance? *Big Data & Society* **7(1)**, 2053951720935143.
2. Guillen, M., Nielsen, J. P., Ayuso, M., Perez-Marin, A. M. (2019). The use of telematics devices to improve automobile insurance rates. *Risk analysis* **39(3)**, 662-672.
3. Guillen, M., Nielsen, J. P., Perez-Marin, A. M., Elpidorou, V. (2020). Can automobile insurance telematics predict the risk of near-miss events? *North American Actuarial Journal* **24(1)**, 141-152.
4. Guillen, M., Perez-Marin, A. M., Alca?iz, M. (2021). Percentile reference charts for speeding based on telematics information. *Accident Analysis and Prevention* **150**, num. 105865.

Contact details

- Email: mguillen@ub.edu. Homepage: www.ub.edu/riskcenter/guillen
- Repository: www.ub.edu/irea. Social media: https://www.linkedin.com/in/montserrat-guillen-09631020/, @mguillen_estany

Gamma Mixture Density Networks and their application to modelling insurance claim amounts

Łukasz Delong, Institute of Econometrics, SGH Warsaw School of Economics (presenter)

Mathias Lindholm, Department of Mathematics, Stockholm University

Mario V. Wüthrich, RiskLab, Department of Mathematics, ETH Zürich

Abstract: We discuss how mixtures of Gamma distributions with mixing probabilities, shape and rate parameters depending on covariates can be fitted with neural networks. We develop two versions of the EM algorithm for fitting so-called Gamma Mixture Density Networks, which we call the EM network boosting algorithm and the EM forward network algorithm, and test their implementation together with the choices of hyperparameters.

A simulation study shows that our algorithms perform very well on artificially constructed data sets. We further illustrate the application of the Gamma Mixture Density Network on a real data set with MTPL claim amounts and conclude that Gamma Mixture Density Networks can improve the fit of the regression model and the predictions of the claim severities used for rate-making compared to classical actuarial techniques.

Keywords: Expectation-Maximization algorithm, neural networks, mixtures of distributions, regression models, rate-making.

References

1. Delong, Ł., Lindholm, M., Wüthrich, M.V. (2021). Gamma Mixture Density Networks and their application to modelling insurance claim amount. Available at SSRN.

Contact details

- Email: lukasz.delong@sgh.waw.pl
- Homepage: www.lukaszdelong.pl

Abstracts/Session_5B

Multi-State Health Transition Modeling Using Neural Networks

Qiqi Wang, Renmin University of China (presenter)

Katja Hanewald, UNSW Sydney

Xiaojun Wang, Renmin University of China

Abstract: This article proposes a new model that combines a neural network with a generalized linear model (GLM) to estimate and predict health transition intensities. The model allows for socioeconomic and lifestyle factors to impact the health transition processes, and captures linear and nonlinear relationships. A key innovation is that the model features transfer learning between different transition rates. It autonomously finds the relationships between factors and the links between the transition processes. We apply the model to individual-level data from the Chinese Longitudinal Healthy Longevity Survey from 1998–2018. The results show that our model performs better in estimation and prediction than standalone GLM and neural network models. We thus provide new estimates of the life expectancies for a range of population subgroups. The model can be easily applied to other datasets, and our results confirm that machine learning techniques are promising tools to model insurance risks.

Keywords: Neural networks, Transfer learning, Multi-state health transitions

Contact details

- Email: k.hanewald@unsw.edu.au

Scenario selection with Lasso regression for the valuation of variable annuity portfolio

Hang Nguyen, University of New South Wales, Sydney (presenter)

Michael Sherris, University of New South Wales, Sydney

Andres Villegas, University of New South Wales, Sydney

Jonathan Ziveyi, University of New South Wales, Sydney

Abstract: A variable annuity is an insurance contract that allows policyholder to choose an investment fund and provide guarantees that are often tied to the fund performance. Performing valuation on large portfolios of these variable annuities is very time consuming because the fair market values of the contracts are usually calculated using Monte Carlo simulations. There have been several studies on how to reduce the computational time, which involves the selection of representative contracts and constructing a metamodel to estimate the values of the remaining contracts in the portfolio. Increasing the number of representative contracts usually increases the accuracy of the method because the metamodel is trained with more data, however, more contracts mean higher computational cost as more Monte Carlo simulations are required to compute the fair market value of the additional contracts. In this paper, after the first step of selecting the representative contracts, we propose using lasso regression to select a set of representative scenarios which we use to build a linear model to predict the fair market values. This step allows us to expand the set of representative contracts without significant increase in the computational load. The proposed approach leads to remarkable improvement in the computational efficiency and accuracy of the metamodel.

Keywords: variable annuities, portfolio valuation, computational time

References

1. Gan, G., Lin, X.S. (2015). Valuation of large variable annuity portfolios under nested simulation : A functional data approach. *Insurance: Mathematics and Economics* **62**, 138-150.
2. Gan, G., Valdez, E.A. (2018). Regression Modeling for the Valuation of Large Variable Annuity Portfolios. *North American Actuarial Journal* **22**, 40-54.
3. Gan, G. (2018). Valuation of Large Variable Annuity Portfolios Using Linear Models with Interactions. *Risks*. **6(3)**, 1-19.
4. Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society*. **Series B (Methodological) 58(1)**, 267-288.
5. Minasny, B., McBratney, A.B.. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences* **32**, 1378-1388.

Contact details

- Email: hang.nguyen@unsw.edu.au
- Social media: https://www.linkedin.com/in/hang-nguyen-849b7a72/

Deepening Lee-Carter for longevity projections with uncertainty estimation

Mario Marino, Sapienza University of Rome

Abstract: Undoubtedly, several countries worldwide endure to experience a continuous increase in life expectancy, extending the challenges of life actuaries and demographers in predicting future mortality evolution. In past literature, several stochastic mortality models have been proposed, starting from the pivotal Lee-Carter model (LC). Recently, various research works encourage the adequacy of deep learning models to achieve more accurate mortality projections.

A proper forecasting model provides robust point predictions, outlining the future mortality trend, as well as confidence ranges for estimates variability. Uncertainty measures associated with the expected values are necessary to sufficiently inspect the phenomenon and, at the same time, to judge both the model adequacy and the results reliability. As in actuarial assessments, uncertainty measures, such as prediction intervals, are imperative. This is a crucial issue, since learning models such as NNs furnish point predictions. However, to the best of our knowledge, machine and deep learning literature in mortality forecasting lack for studies about uncertainty estimation.

As new advance in mortality forecasting, we formalize the integration of deep learning techniques in the LC model framework, in terms of both point estimates and prediction intervals for future mortality rates. In doing so, we refer to a Recurrent Neural Network with Long Short Term Memory (LSTM) architecture to forecast the LC time-index. The resulting integrated model, namely LC-LSTM, fill the gap between the deep learning integrated mortality models and the uncertainty estimation, getting suitable ranges of variability. We test the proposed model in a numerical application considering both genders and three representative countries worldwide, scrutinizing two different learning periods. Our findings confirm the suitability of deep learning models to improve the predictive capacity of the canonical Lee-Carter model, providing reliable mortality boundaries on the long-run forecasts useful for actuarial evaluations.

Keywords: Mortality Forecasting, Recurrent Neural Networks, Prediction Interval

References

[1] Brouhns, N., Denuit, M., Vermunt, J. (2002). A Poisson log-bilinear approach to the construction of projected life tables. *Insurance: Mathematics and Economics*, **31**.

[2] Khosravi A., Creighton D. (2011). Comprehensive review of neural network-based prediction intervals and new advances. *IEEE Transactions on Neural Networks*, **22(9)**.

[3] Nigri, A., Levantesi, S., Marino, M., Scognamiglio, S., Perla, F. (2019). A deep learning integrated Lee-Carter model. *Risks*, **7(1)**.

[4] Perla, F., Richman, R., Scognamiglio, S., Wuthrich, M. (2021). Time-series forecasting of mortality rates using deep learning. *Scandinavian Actuarial Journal*.

[5] Richman, R., Wuthrich, M. (2019a). A neural network extension of the Lee-Carter model to multiple populations. *Annals of Actuarial Science*.

Contact details

- Email: mmarino@uniroma1.it
- Social media: Twitter, @MarioShawMarino

Abstracts/Session_6A

Capturing the power of ensemble learning using GLM and Artificial Neural Network for insurance pricing

S.R.Pranav Sai, Doctoral Research Scholar Actuarial Science, Sri Sathya Sai Institute of Higher Learning

Abstract: Neural network modelling is yet to see significant growth in the insurance sector. The reason for this is attributed to the lack of explainability of neural network models. But neural networks on the other hand can improve the performance of actuarial models if they are appropriately integrated. Much work has appeared in the areas of integrating actuarial models with neural networks. Mario in their works in [1] and [2], propose the Combined Actuarial Neural Networks (CANN). In the CANN architecture, the GLM model is integrated with the neural network with a help of a skip connection thus giving a neural net boosting of the actuarial model. We take this work forward by proposing new architectures for getting a seamless integration of the neural networks and the actuarial models for getting better results. we use the principles of ensemble building in deeplearning for building these models. Ensemble learning is proving to be an effective strategy in predictive analytics and is increasingly being used in solving many real-world problems. The fundamental principle in ensemble building is that the best model knows less about the data than all the models combined. Combining several base models to create a more powerful ensemble model then rises as a natural by-product of this workflow. Ensembles range from simple to weighted averages up to complicated 2nd and 3rd level meta-models. In our work, we propose ensemble building of the GLM with neural networks and inherently learning how to best combine predictions from the existing models called the base learners. We use three types of ensembles using these models. 1) Stacking Ensemble 2) Weighted Average Ensemble 3) 3rd level stacking Ensemble In the stacking ensemble, the output of the base-learners is taken as inputs for training a meta-learner, that learns how to best combine the base learners' predictions. In the weighted average ensemble, the model weights the contribution of each ensemble member based on their performance on a hold-out validation dataset. Models with better contribution receive a higher weight. In the 3rd level stacking ensemble, the multiple meta learners are obtained from the base learners and a 3rd level meta learner is obtained. It was found that on the freMTPL2freq dataset, these ensemble models gave improved results and reduced Poisson deviance compared to the traditional GLM models. These ensembles on average gave about 80% improvement over the 'GLM2' model of [2] which was used as a baseline model for comparison.

Keywords: GLM, Neural Network, Ensemble Deeplearning, CANN

References

1. Wüthrich, M., & Merz, M. (2019). EDITORIAL: YES, WE CANN! ASTIN Bulletin, 49(1), 1-3. doi: 10.1017/asb.2018.42
2. Schelldorfer, Jürg and Wuthrich, Mario V(2019). Schelldorfer, Jürg and Wuthrich, Mario V., Nesting Classical Actuarial Models into Neural Networks (January 22, 2019). Available at SSRN: https://ssrn.com/abstract=3320525 or http://dx.doi.org/10.2139/ssrn.3320525

Contact details

- Email: srpranavsai@sssihl.edu.in

Sequential Monte Carlo Samplers to fit and compare insurance loss models

Pierre-O. Goffard, ISFA, Université Claude Bernard Lyon 1 (presenter)

Abstract: Insurance losses are generally divided into low intensity claims and extreme claims. A composite model takes this into account by combining two models, one for the left tail and one for the right tail of the loss severity distribution. The cutoff point between the two parts of the loss data is a particularly interesting parameter. The calibration of such a model results either from maximum likelihood estimations or techniques borrowed from the theory of extreme values. Here, a Bayesian approach is presented. A sequential Monte Carlo sampler is used to both fit the models to the data and find the most suitable model. The methodology is illustrated on a real world insurance dataset.

Keywords: Loss models, Bayesian statistics, sequential Monte Carlo sampler

References

1. Pierre Del Moral, Arnaud Doucet, and Ajay Jasra (2016). Sequential monte carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **68(3)**, 411-436.
2. Yinzhi Wang, Ingrid Hobaek Haff, and Arne Huseby (2020). Modelling extreme claims via composite models and threshold selection methods. *Insurance: Mathematics and Economics*, **91**, 257-268.

Contact details

- Email: pierre-olivier.goffard@univ-lyon1.fr
- Homepage: http://pierre-olivier.goffard.me/

Copula model selection using image processing

Rui Zhu, The Business School (formerly Cass), City, University of London (presenter)

Andreas Tsanakas, The Business School (formerly Cass), City, University of London

Abstract: Copulas are widely used for the modelling of dependent risks, with extensive applications in actuarial and financial risk management (Denuit et al., 2006; McNeil et al., 2015). Choosing a suitable copula family, when multivariate datasets are of small to moderate size, is a non-trivial task, as shown in the seminal study of Genest et al. (2009), with a wide array of goodness-of-fit statistics and information criteria deployed for that task.

Here we present our work on utilising image processing techniques to aid bivariate copula selection, without resorting to potentially complex likelihood-based criteria. We treat the heatmaps of smoothed bivariate densities (with standard normal margins) as RGB images and explore the spatial patterns presented in the images to help copula selection. More specifically, we extract image features from a pretrained convolutional neural network, AlexNet (Krizhevsky et al., 2012), and use these extracted features to train a classifier to classify copula models. The higher-level features from AlexNet are discriminative and can help distinguish between different heatmap patterns.

Experiments on simulated data show that our approach can beat the cross-validated log-likelihood criterion of Gronneberg and Hjort (2014) and the Cramer-von Mises statistic. When the sample size of the copula model is reduced to a small one, e.g. 100, the classification accuracy of our approach is 4% higher than the cross-validated log-likelihood criterion and 11% higher than the Cramer-von Mises statistic. There is also evidence showing that the classification accuracy can be further improved by (a) including in the classifier the estimated rank correlations and some summary statistics and (b) by pre-processing samples, via a Rosenblatt transform with respect to a fitted copula model.

Keywords: Copula model selection, classification, image processing, AlexNet

References

1. Denuit, M., Dhaene, J., Goovaerts, M., and Kaas, R. (2006). *Actuarial theory for dependent risks: measures, orders and models*. John Wiley & Sons.
2. McNeil, A. J., Frey, R., and Embrechts, P. (2015). *Quantitative risk management: concepts,techniques and tools-revised edition*. Princeton university press.
3. Genest, C., Rémillard, B., and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. *Insurance: Mathematics and Economics*, 44(2):199–213.
4. Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25 (2012): 1097–1105.
5. Grønneberg, S. and Hjort, N. L. (2014). The copula information criteria. *Scandinavian Journal of Statistics*, 41(2):436–459.

Contact details

- Email: rui.zhu@city.ac.uk (Rui Zhu); a.tsanakas.1@city.ac.uk (Andreas Tsanakas)

Extremile Regression

Abdelaati Daouia, Toulouse School of Economics (co-presenter)

Irene Gijbels, KU Leuven

Gilles Stupfler, ENSAI & CREST (presenter)

Abstract: Regression extremiles define a least squares analogue of regression quantiles. They are determined by weighted expectations rather than tail probabilities. Of special interest is their intuitive meaning in terms of expected minima and maxima. Their use appears naturally in risk management where, in contrast to quantiles, they fulfill the coherency axiom and take the severity of tail losses into account. In addition, they are comonotonically additive and belong to both the families of spectral risk measures and concave distortion risk measures. We provide the first detailed study exploring implications of the extremile terminology in a general setting of presence of covariates. We rely on local linear (least squares) check function minimization for estimating conditional extremiles and deriving the asymptotic normality of their estimators. We also extend extremile regression far into the tails of heavy-tailed distributions. Some applications to estimation and inference for tail risk associated with insurance payouts are provided.

Keywords: Asymmetric least squares, Extremes, Heavy tails, Regression extremiles, Regression quantiles, Tail index

References

Daouia, A., Gijbels, I., Stupfler, G. (2021). Extremile Regression. *Journal of the American Statistical Association*, to appear.

Contact details

- Email: gilles.stupfler@ensai.fr
- Homepage: http://ensai.fr/en/equipe/stupfler-gilles/
- Twitter: @GillesStupfler
- LinkedIn: http://www.linkedin.com/in/gilles-stupfler-5713651ab

A non-convex regularization approach for stable estimation of loss development factors

Himchan Jeong, Simon Fraser University (presenter)

Hyunwoong Chang, Texas A&M University

Emiliano A. Valdez, University of Connecticut

Abstract: In this article, we apply non-convex regularization methods in order to obtain stable estimation of loss development factors in insurance claims reserving. Among the non-convex regularization methods, we focus on the use of the log-adjusted absolute deviation (LAAD) penalty and provide discussion on optimization of LAAD penalized regression model, which we prove to converge with a coordinate descent algorithm under mild conditions. This has the advantage of obtaining a consistent estimator for the regression coefficients while allowing for the variable selection, which is linked to the stable estimation of loss development factors. We calibrate our proposed model using a multi-line insurance dataset from a property and casualty insurer where we observed reported aggregate loss along accident years and development periods. When compared to other regression models, our LAAD penalized regression model provides very promising results.

Keywords: Insurance reserving, log-adjusted absolute deviation (LAAD) penalty, loss development, non- convex penalization, robust estimation, variable selection

Contact details

- Email: himchan_jeong@sfu.ca
- Homepage: https://ssauljin.github.io/hjeong/

Relativities in the Over-Dispersed Poisson Bootstrap Claims Reserves

Queensley C. Chukwudum, Department of Insurance and Risk Management, University of Uyo, Akwa Ibom State, Nigeria

Abstract: The over-dispersed Poisson bootstrapped claims reserves structure is studied in this paper. The main objective is to examine the structural relationships between the total estimated outstanding loss or case reserves (CR), the total current year's (CTY) payments and the total immediate next year's (NTY) projected loss payments. These variables are extracted from the bootstrapped reserves and the simulated incremental payments.

The analysis makes use of three paid claims run-off triangle datasets from Nigeria, the United States and the United Kingdom. In all cases, the Spearman's rho and Kendall tau indicate moderate to high positive correlation of CR with respect to NTY payments as expected. Additionally however, it confirms a weak correlation with CTY payments.

The mean CR is then predicted with the aid of a generalized linear model (GLM) which considers a Poisson distribution with log link function. The CR totals are taken as the target variable and the CTY and NTY incremental payment sums as the predictors. Two GLM models are trained and tested - model 1 includes all variables and model 2 accounts for the potential interaction between CTY and NTY. Both models provide fairly rough estimates of the mean CR.

Although a very weak correlation is observed between CTY and NTY in each dataset, the predictive power of the GLM is slightly improved when their interaction is accounted for as is implied by the second model's lower Akaike information criterion and its root mean square errors in comparison to model 1. This outcome suggests that the interaction between CTY and NTY is an important feature that affects CR hence, it should be factored in when reserving claims.

Keywords: Case reserves, GLM, loss reserving, run-off triangles

References

1. Zhang, Y. (2010). A General Multivariate Chain Ladder Model. Insurance: Mathematics and Economics. 46. 588-599. 10.1016/j.insmatheco.2010.03.002.
2. Nieto-Barajas, L.E. and Targino, R. S. (2020). A Gamma Moving Average Process for Modelling Dependence across Development Years in Run-off Triangles. ASTIN Bulletin, 10.1017/asb.2020.36, (1-22).
3. Avanzi, B., Taylor, G., Vu, P.A. and Wong, B. (2016). Stochastic Loss Reserving with Dependence: A Flexible Multivariate Tweedie Approach. Insurance: Mathematics and Economics, 10.1016/j.insmatheco.2016.08.006, 71, (63-78).
4. Merz, M. and Wuthrich, M. (2008). Prediction Error of the Multivariate Chain Ladder Reserving Method. North American Actuarial Journal. 12. 10.1080/10920277.2008.10597509.

Contact details

- Email: queensleyv@yahoo.com

Abstracts/Session_6B

Multi-population modeling with economic trends: A hybrid neural network approach

Qiqi Wang, Renmin University of China

Katja Hanewald, UNSW Sydney (presenter)

Xiaojun Wang, Renmin University of China

Abstract: This paper proposes a new model for estimating and predicting the mortality of multiple populations using a hybrid neural network model that combines convolutional neural network (CNN) layers with a long short-term memory (LSTM) structure. This model incorporates economic conditions represented by estimates of gross domestic product (GDP) per capita to link with mortality trends and assist mortality forecasts. The new hybrid model predicts gross domestic product (GDP) per capita and mortality rates simultaneously for multiple populations. The model detects and incorporates the complex linear and nonlinear relationships between mortality rates and economic conditions autonomously. We apply the model to data from the Human Mortality Database for 34 countries over the period 1971-2018. The results show that the new model gives good estimates and forecasts. We provide mortality rate predictions for countries and the future trend of mortality. Our results confirm the hybrid neural network approach is promising in mortality modeling.

Keywords: CNN Long Short-Term Memory Networks, Multi-population mortality, GDP

Contact details

- Email: wangqiqi_ruc@outlook.com

A Dirichlet Process Mixture model for the analysis of competing risks

Francesco Ungolo, Technische Universiteit Eindhoven (presenter)

Abstract: A regression model for the analysis of competing risks is developed. The joint distribution of the time to events is characterized by a non-parametric frailty component which explains their variability whereas unexplained by available covariates. The model is analysed in a fully Bayesian setting, yielding a flexible Dirichlet Process Mixture model for the joint distribution of the time to competing events. The modelling apprach allows to easily derive the crude and the net hazard from first principles. An efficient MCMC sampler is provided for inference in order to deal with large datasets and large parameter spaces.

The approach is applied to the analysis of the lapse risk for the life-insurance dataset used in Milhaud and Dutang (2018).

Keywords: Competing Risks, Bayesian nonparametrics, MCMC, Lapse risk

References

1. Milhaud, X., Dutang, C. (2018). Lapse tables for lapse risk management in insurance: a competing risk approach. *European Actuarial Journal* **8**, 97-126.

Contact details

- Email: francescoungolo@gmail.com
- Twitter https://twitter.com/FraUng89Stats

Point and Interval Forecasts of Death Rates Using Neural Networks

Simon Schnürch, Fraunhofer Institute for Industrial Mathematics ITWM and University of Kaiserslautern (presenter)

Ralf Korn, Fraunhofer Institute for Industrial Mathematics ITWM and University of Kaiserslautern

Abstract: The Lee-Carter model (Lee and Carter, 1992) has become a benchmark in stochastic mortality modeling. However, its forecasting performance can be significantly improved upon by modern machine learning techniques (Richman and Wüthrich, 2019; Perla et al., 2020). We propose a convolutional neural network architecture for mortality rate forecasting, empirically compare this model as well as other neural network models to the Lee-Carter model and find that lower forecast errors are achievable for many countries in the Human Mortality Database. We provide details on the errors, forecasts and global behavior of our model to make it more understandable and, thus, more trustworthy. As neural networks by default only yield point estimates, previous works applying them to mortality modeling have not investigated prediction uncertainty. We address this gap in the literature by implementing a bootstrapping-based technique and demonstrate that it yields highly reliable prediction intervals for our neural network model. Finally, we find that annuity values can strongly differ depending on the model, highlighting the important issue of model risk.

Keywords: Mortality forecasting, neural networks, convolutional neural networks, uncertainty quantification, prediction intervals, Lee-Carter model, mortality of multiple populations.

References

1. Lee, R. D., Carter, L. R. (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association* **87/419**, 659-671.
2. Perla, F., Richman, R., Scognamiglio, S., Wüthrich, M. V. (2020). Time-Series Forecasting of Mortality Rates using Deep Learning. *Scandinavian Actuarial Journal* **(to appear)**.
3. Richman, R., Wüthrich, M. V. (2019). A Neural Network Extension of the Lee-Carter Model to Multiple Populations. *Annals of Actuarial Science* **(to appear)**.
4. Schnürch, S., Korn, R. (2021). Point and Interval Forecasts of Death Rates Using Neural Networks. Available at SSRN: https://ssrn.com/abstract=3796051.

Contact details

- Email: simon.schnuerch@itwm.fraunhofer.de; korn@mathematik.uni-kl.de

Unsupervised Learning applied to the Customer Lifetime Value (CLV)

Claudio Giorgio Giancaterino, Catholic University of Milan (presenter)

Abstract: The core business of Insurance Companies is to enable individuals and firms to protect themselves against rarely events paying a small premium compared to the eventually damage incurred.

Customer Lifetime Value (CLV) evaluates the value of the customer for the Company, in other words, it's the Net Present Value of the cash flows ascribed to the relationship with a customer.

In this work, from the collection of portfolio contracts by one insurance year, will be predicted the Customer Lifetime Value of the last three months of the year, also looking at the effects coming from the use of Unsupervised Learning.

Unsupervised Learning describes tasks that involves using a model to discover a good internal representation of input data useful for subsequent Supervised Learning. In this job Unsupervised Leaning are used to provide a low-dimensional representation of inputs and clustering numerical variables to provide a better portfolio analysis of customers. In both situations, Unsupervised Learning can be used as feature engineering to improve Machine Learning performances.

Keywords: Unsupervised Learning, Machine Learning, Gradient Boosting Machine, Principal Component Analysis, Autoencoders, ISOMAP, t-Distributed Stochastic Neighbor Embedding, K-Means, Hierarchical Clustering, DBSCAN, Gaussian Mixture Models.

References

1. T. Hastie, R. Tbishirani, J. Friedman (2017).*The Elements of Statistical Learning*,Springer.
2. M. V. Wutrich, S. Rentzmann (2019).*Unsupervised Learning: What is a sport car?*,SSRN.
3. I. Goodfellow, Y. Bengio, A. Courville (2016). *Deep Learning*,MIT Press.
4. Yoon Hyup Hwang (2019).*Hands-On Data Science for Marketing*, Packt.
5. G. Bonaccorso (2019).*Hands-On Unsupervised Learning with Python*, Packt.

Contact details

- Email: c.giancaterino@gmail.com
- Repository: https://github.com/claudio1975/Customer_Lifetime_Value
- Social media: https://www.linkedin.com/in/claudiods/

Tree-based Models for Variable Annuity Valuation: Parameter Tuning and Empirical Analysis

Zhiyu Quan, University of Illinois at Urbana-Champaign (presenter)

Guojun Gan, University of Connecticut

Emiliano Valdez, University of Connecticut

Abstract: Variable annuities have become popular retirement and investment vehicles due to their attractive guarantee features. Nonetheless, managing the financial risks associated with the guarantees poses great challenges for insurers. One challenge is risk quantification, which involves frequent valuation of the guarantees. Insurers rely on the use of Monte Carlo simulation for valuation as the guarantees are too complicated to be valued by closed-form formulas. However, Monte Carlo simulation is computationally intensive. In this paper, we empirically explore the use of tree-based models for constructing metamodels for the valuation of the guarantees. In particular, we consider traditional regression trees, tree ensembles, and trees based on unbiased recursive partitioning. We compare the performance of tree-based models to that of existing models such as ordinary kriging and GB2 regression. Our results show that tree-based models are efficient in producing accurate predictions and the gradient boosting method is considered the most superior in terms of prediction accuracy.

Keywords: tree-based model; variable annuity; portfolio valuation; metamodeling

References

1. Gan, G. and Valdez, E. A. (2017b). Valuation of large variable annuity portfolios: Montecarlo simulation and synthetic datasets.Dependence Modeling, 5:354–374.
2. Breiman, L. (2001). Random forests.Machine Learning, 45(1):5–32.
3. Hothorn, T., Hornik, K., and Zeileis, A. (2006). Unbiased recursive partitioning: A condi-tional inference framework.Journal of Computational and Graphical Statistics, 15(3):651–674.
4. Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine.Annals of Statistics, 29(5):1189–1232.
5. Strobl, C., Malley, J., and Tutz, G. (2009). An introduction to recursive partitioning:rationale, application, and characteristics of classification and regression trees, bagging,and random forests.Psychological methods, 14(4):323.

Contact details

- Email: zquan@illinois.edu
- Homepage: http://zquan.pages.math.illinois.edu/
- Linkedin: https://www.linkedin.com/in/zhiyufrankquan/