Institute
and Faculty
of Actuaries

Insurance analytics: prediction, explainability, and fairness

Kjersti Aas[1], Arthur Charpentier[2] (iD), Fei Huang[3] (iD) and Ronald Richman[4] (iD)

[1]Norwegian Computing Center, Oslo, Norway; [2]Université du Québec à Montréal, Montreal, Canada; [3]University of New South Wales, Sydney, Australia; and [4]Old Mutual Insure and University of the Witwatersrand, South Africa
Corresponding author: Fei Huang; Email: feihuang@unsw.edu.au

Abstract

The expanding application of advanced analytics in insurance has generated numerous opportunities, such as more accurate predictive modeling powered by machine learning and artificial intelligence (AI) methods, the utilization of novel and unstructured datasets, and the automation of key operations. Significant advances in these areas are being made through novel applications and adaptations of predictive modeling techniques for insurance purposes, while, concurrently, rapid advances in machine learning methods are being made outside of the insurance sector. However, these innovations also bring substantial challenges, particularly around the transparency, explanation, and fairness of complex algorithmic models and the economic and societal impacts of their adoption in decision-making. As insurance is a highly regulated industry, models may be required by regulators to be explainable, in order to enable analysis of the basis for decision making. Due to the societal importance of insurance, significant attention is being paid to ensuring that insurance models do not discriminate unfairly. In this special issue, we feature papers that explore key issues in insurance analytics, focusing on prediction, explainability, and fairness.

1. Introduction

Actuaries have been using econometric and statistical models for decades. And just as statistical learning has fundamentally changed the way predictive models are built, actuaries have had to adapt to these new techniques. Neural networks, a concept rooted in the 1940s and inspired by the structure of the human brain, have exploded in recent decades with the advent of massive data, enabling increasingly sophisticated architectures and capturing more complex effects. This progress has made it possible to implement the universal approximation theorem, which had previously existed only in theory. The 2024 Nobel Prize in Physics, awarded to Hopfield and Hinton, underscores their pivotal role in this revolution.

But the arrival of these artificial intelligence (AI)/machine learning models has not been without its problems. Breiman (2001) spoke of a cultural difference between data modelers and algorithmic modelers. But the difference is more profound. Econometric and statistical models are deeply probabilistic, whereas learning algorithms are not. In a Support Vector Machine (SVM), we try to place a separating plane in a cloud of points, based on distance, and if this allows us to separate images of dogs and cats or individuals who are sick and others who are not, the question of the *probability* of belonging in a given group rarely arises. Yet it is this quantity that is essential for actuaries, in order to construct a tariff. The actuary is not trying to predict who will die in a life insurance portfolio, but to estimate, as accurately as possible, the probability of death for

each individual. Recent advances in insurance predictive analytics have brought new challenges, such as handling high-cardinality features, incorporating Poisson and Tweedie loss functions into machine learning models, and enforcing smoothness and monotonicity constraints.

As a highly regulated industry, insurance often requires models to be explainable, enabling regulators and stakeholders to understand the basis for decision-making. However, emerging machine learning and AI models is usually too complex and opaque to meet these explainability standards. This has created a need for new models and techniques that can harness the predictive power of these black-box models while maintaining the transparency and interpretability that insurance demands.

Issues of discrimination and fairness in insurance have long been debated. Yet, AI and Big Data have added layers of complexity, as opaque algorithms and proxy discrimination introduce new concerns. Addressing these challenges requires multiperspective and cross-disciplinary collaboration. Importantly, even to start addressing these challenges, interpretability and explainability is a fundamental prerequisite.

To encourage further research in this area and support recent innovations, the Annals of Actuarial Science (AAS) has launched a special issue titled "Insurance Analytics: Prediction, Explainability, and Fairness."

2. Predictive analytics in insurance

Statistical models have long been used in insurance, but their use raises profound epistemological questions. Von Mises (1939) explained that the "probability of death" applies to a group or class of individuals, not to any single person, as it has no meaning when referring to an individual, even with detailed knowledge of their life and health. In the frequentist approach, probabilities are constructed as asymptotic limits of frequencies, grounded in the law of large numbers. By reasoning in terms of "homogeneous risk classes," actuaries historically relied on robust statistical techniques, both mathematically and philosophically, to rate policyholders. However, modern machine learning techniques now allow for pricing individual risks and personalizing premiums with increasing granularity. This shift introduces new challenges when applying advanced analytics in insurance, including managing high-cardinality features, classifying policyholders into unique subgroups, and incorporating Poisson and Tweedie deviance loss functions in boosting and tree-based methods.

In this special issue, we present four papers that focus on predictive analytics in insurance.

Campo & Antonio (2024) proposed the data-driven Partitioning Hierarchical Risk-factors Adaptive Top-down (PHRAT) algorithm to reduce hierarchically structured risk factors to their essence by grouping similar categories at each level. They also utilize embeddings to encode textual descriptions of economic activities, aiding in the grouping of categories for inputs.

Lee & Jeong (2024) modified the alternating direction method of multipliers (ADMM) for subgroup analysis to classify policyholders into unique groups. They interpret the credibility problem using both random effects and fixed effects, which correspond to the ADMNM approach and the classic Bayesian approach, respectively.

Willame et al. (2024) reviewed the use of boosting under the Poisson deviance loss function and log-link (following Wüthrich & Buser, 2019) and apply boosting with cost-complexity pruned trees on Tweedie responses (following Huyghe et al., 2022). They introduced a new Boosting Trees package in R designed for insurance applications.

Wu et al. (2024) extended the traditional Lee-Carter model using Kernel Principal Component Analysis (KPCA) to enhance mortality rate predictions. They demonstrated the robustness of this model, particularly during the COVID-19 pandemic, showing its superior performance in volatile conditions.

3. Explainability and interpretability

Insurance, as a high-stakes business, faces stringent regulatory requirements, particularly regarding explainability and interoperability. Traditional statistical models, such as Generalized Linear Models (GLMs) and Generalized Additive Models (GAMs), are typically more interpretable than modern machine learning models like Gradient Boosting Machines, Random Forests, or Neural Networks. While a range of interpretability tools have been developed to increase transparency in these black-box models, they have not been without criticism (Hooker et al., 2021; Rudin, 2019; Xin et al., 2024). Balancing regulatory demands for explainability with the use of advanced machine learning models has become a pressing challenge for the insurance industry. Recent literature emphasizes the growing importance of model interpretation and transparency in this field, as highlighted by Aas et al. (2021), Delcaillau et al. (2022), and Richman & Wüthrich (2023).

In this special issue, we present four papers that focus on explainability and interpretability in insurance analytics:

Jose et al. (2024) developed a zero-inflated Poisson neural network (ZIPNN) by following the combined actuarial neural network (CANN) approach to model admission rates. They extend this with zero-inflated combined actuarial neural network (ZIPCANN) models and adopt the LocalGLMnet method (Richman & Wüthrich, 2023) to interpret the models.

Lindholm & Palmquist (2024) proposed a method for constructing categorical GLMs guided by information derived from a black-box predictor. They use partial dependance (PD) functions to create covariate partitions based on the black-box predictor, followed by an auto-calibration step and a lasso-penalized GLM fitting.

Maillart & Robert (2024) explored an approach to estimate a GAM with non-smooth feature functions. This method distills knowledge from an Additive Tree model, partitions the covariate space, and fits a GLM using binned covariates for each decision tree, followed by an ensemble approach and final GLM fitting for auto-calibration.

Richman & Wüthrich (2024) introduced ICEnet, a method that enforces smoothness and monotonicity constraints in deep neural networks. To train neural networks with these constraints, they augment datasets to produce pseudo-data that reflect the desired properties. A joint loss function is used to balance accurate predictions with constraint enforcement.

4. (Algorithmic) fairness

As in other industries, insurers are redefining their practices with the rise of Big Data and advanced AI algorithms, enabling them to detect previously unknown patterns, incorporate more rating factors, improve predictive accuracy, and move toward more granular risk classification. While these technologies expand the scope of what is possible, they do not fundamentally change the longstanding issues of insurance discrimination. In fact, in this rapidly evolving landscape, old challenges are becoming more pronounced. Concerns about indirect discrimination and the use of algorithmic proxies are growing, as insurers increasingly leverage vast datasets and sophisticated models.

Avraham (2017) argued that insurance faces unique moral and legal challenges. While policymakers seek to prevent discrimination based on factors like race, gender, and age, the insurance business inherently involves distinguishing between risky and non-risky insureds, which often correlates with those sensitive characteristics. Actuaries must remain vigilant to these issues and actively contribute to solutions that mitigate the risks of discrimination.

Recent research has begun addressing these issues for insurance applications from multiple perspectives (such as ethical, actuarial, statistical, economic, and legal perspectives), with contributions from Prince & Schwarcz (2019), Baumann & Loi (2023), Lindholm et al. (2022, 2024), Frees & Huang (2021), Xin & Huang (2023), Barry & Charpentier (2023), Charpentier (2024), Araiza Iturria et al. (2024), and Fahrenwaldt et al. (2024), as examples.

We encourage more contributions to this crucial area of research, addressing the ongoing challenges of discrimination and fairness in insurance via multidisciplinary research and collaborations.

5. Conclusion

The integration of advanced analytics, machine learning, and AI into the insurance industry has presented significant opportunities to enhance predictive accuracy, streamline operations, and deliver more personalized services. However, with these advances come complex challenges, particularly in maintaining the explainability and fairness of increasingly opaque models. As insurers adopt these powerful tools, the responsibility to ensure ethical and responsible use becomes even more critical.

The papers in this special issue of the Annals of Actuarial Science highlight cutting-edge approaches to prediction and explainability in insurance analytics. They collectively demonstrate the potential of advanced methods to address industry challenges, but they also emphasize the need for further research to reconcile the power of these models with business, regulatory, and ethical considerations. We encourage continued research contributions to this critical area.

References

Aas, **K.**, **Jullum**, **M.**, & **Løland**, **A.** (2021). Explaining individual predictions when features are dependent: More accurate approximations to Shapley values. *Artificial Intelligence*, **298**, 103502.

Araiza Iturria, **C. A.**, **Hardy**, **M.**, & **Marriott**, **P.** (2024). A discrimination-free premium under a Causal framework. *North American Actuarial Journal*, 1–21. doi: 10.1080/10920277.2023.2291524.

Avraham, **R.** (2017). Discrimination and insurance. In *The Routledge handbook of the ethics of discrimination* (pp. 335–347).

Barry, **L.**, & **Charpentier**, **A.** (2023). Melting contestation: Insurance fairness and machine learning. *Ethics and Information Technology*, **25**(4), 49. doi: 10.1007/s10676-023-09720-y.

Baumann, **J.**, & **Loi**, **M.** (2023). Fairness and risk: An ethical argument for a group fairness definition insurers can use. *Philosophy & Technology*, **36**(3), 45. doi: 10.1007/s13347-023-00624-9.

Breiman, **L.** (2001). Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, **16**(3), 199–231.

Campo, **B. D. C.**, & **Antonio**, **K.** (2024). On clustering levels of a hierarchical categorical risk factor. *Annals of Actuarial Science*, 1–39. doi: 10.1017/S1748499523000283. Published online.

Charpentier, **A.** (2024). *Insurance, biases, discrimination and fairness*. Springer.

Delcaillau, **D.**, **Ly**, **A.**, **Papp**, **A.**, & **Vermet**, **F.** (2022). Model transparency and interpretability: Survey and application to the insurance industry. *European Actuarial Journal*, **12**(2), 443–484.

Fahrenwaldt, **M.**, **Furrer**, **C.**, **Hiabu**, **M. E.**, **Huang**, **F.**, **Jørgensen**, **F. H.**, **Lindholm**, **M.**, & **Tsanakas**, **A.** (2024). Fairness: Plurality, causality, and insurability. *European Actuarial Journal*, **14**(2), 1–12.

Frees, **E. W.(Jed)**, & **Huang**, **F.** (2021). The discriminating (Pricing) actuary. *North American Actuarial Journal*, **27**(1), 2–24. doi: 10.1080/10920277.2021.1951296.

Glenn, **B. J.** (2000). The shifting rhetoric of insurance denial. *Law and Society Review*, **34**(3), 779–808.

Hooker, **G.**, **Mentch**, **L.**, & **Zhou**, **S.** (2021). Unrestricted permutation forces extrapolation: Variable importance requires at least one more model, or there is no free variable importance. *Statistics and Computing*, **31**(6), 1–16.

Huyghe, **J.**, **Trufin**, **J.**, & **Denuit**, **M.** (2022). Boosting cost-complexity pruned trees on Tweedie responses: The ABT machine for insurance ratemaking. *Scandinavian Actuarial Journal*, **2024**(5), 417–439. doi: 10.1080/03461238.2023.2258135.

Jose, **A.**, **Macdonald**, **A. S.**, **Tzougas**, **G.**, & **Streftaris**, **G.** (2024). Interpretable zero-inflated neural network models for predicting admission counts. *Annals of Actuarial Science*, 1–31. doi: 10.1017/S1748499524000058. Published online.

Lee, G. Y., & **Jeong**, H. (2024). Nonparametric intercept regularization for insurance claim frequency regression models. *Annals of Actuarial Science*, 1–26. doi: 10.1017/S1748499523000271. Published online.

Lindholm, M., & **Palmquist**, J. (2024). Black-box guided GLM building with non-life pricing applications. *Annals of Actuarial Science*. Published online.

Lindholm, M., **Richman**, R., **Tsanakas**, A., & **Wüthrich**, M. V. (2022). Discrimination-free insurance pricing. *ASTIN Bulletin*, **52**(1), 55–89. doi: 10.1017/asb.2021.23.

Lindholm, M., **Richman**, R., **Tsanakas**, A., & **Wüthrich**, M. V. (2024). What is fair? Proxy discrimination vs. demographic disparities in insurance pricing. *Scandinavian Actuarial Journal*, **2024**(9), 935–970. doi: 10.1080/03461238.2024.2364741.

Maillart, A., & **Robert**, C. Y. (2024). Distill knowledge of additive tree models into generalized linear models: A new learning approach for non-smooth generalized additive models. *Annals of Actuarial Science*. Published online.

Prince, A. E., & **Schwarcz**, D. (2019). Proxy discrimination in the age of artificial intelligence and big data. *Iowa Law Review*, **105**, 1257.

Richman, R., & **Wüthrich**, M. V. (2023). LocalGLMnet: Interpretable deep learning for tabular data. *Scandinavian Actuarial Journal*, **2023**(1), 71–95.

Richman, R., & **Wüthrich**, M. V. (2024). Smoothness and monotonicity constraints for neural networks using ICEnet. *Annals of Actuarial Science*, 1–28. doi: 10.1017/S174849952400006X. Published online.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, **1**(5), 206–215.

Von Mises, R. (1939). *Probability, statistics, and truth*. Macmillan.

Willame, G., **Trufin**, J., & **Denuit**, M. (2024). Boosted Poisson regression trees: A guide to the BT package in R. *Annals of Actuarial Science*, 1–21. doi: 10.1017/S174849952300026X. Published online.

Wu, Y., **Chen**, A., **Xu**, Y., **Pan**, G., & **Zhu**, W. (2024). Modeling mortality with kernel principal component analysis (KPCA) method. *Annals of Actuarial Science*. Published online.

Wüthrich, M. V., & **Buser**, C. (2019). Data analytics for non-life insurance pricing. Lecture notes available at SSRN. Lecture notes available at SSRN. http://dx.doi.org/10.2139/ssrn.2870308.

Xin, X., & **Huang**, F. (2023). Antidiscrimination insurance pricing: Regulations, fairness criteria, and models. *North American Actuarial Journal*, **28**(2), 285–319. doi: 10.1080/10920277.2023.2190528.

Xin, X., **Huang**, F. & **Hooker**, G. (2024), *Why You Should Not Trust Interpretations in Machine Learning: Adversarial Attacks on Partial Dependence Plots*. arXiv preprint arXiv: 2404.18702.