# Enhancing Insurance Risk Scoring with AI

By Dr. Andreas Petrou

Professor of Electrical and Computer Engineering, Aristotle University of Thessaloniki (AUTH), Greece

## 1. Introduction to Insurance Risk Scoring and AI

Insurance risk scoring involves the assimilation, calibration, and interpretation of diverse data pertinent to insurance coverages into a single, simple score for a single class of loss under consideration, and a policyholder within that scope. Prediction of those individual losses is a step toward risk pricing. When underwriting and ratemaking require accurate risk assessment, traditionally, insurers relied on risk class and experience-rating approaches to estimate the total losses and trends in order to set fair and affordable premiums. This requires an understanding of which features are fundamentally related to the setting of different prices. In many cases, these are the features that pricing regulations allow insurers to use in setting premiums or liabilities for ratemaking, as well as to enable insurance firms to match rates to risk characteristics accurately. Other features may instead be relied upon as a proxy for other information that has a regularization impact on the cost. The expense linked to the generation of this needed data often leaves those with less comprehensive and consistent details priced higher. This is especially the case in a modern setting that uses both credit and behavioral scoring tools. As the use of technology increases, it is particularly important for the people and businesses concerned to limit utilization and data use in research and initial work to be vigilant against the accuracy reinforcement of these procedures that are generally more favorable to those with the most insight. To simplify the problem concerning fairness, we consider a model with all class members having the same covariate distribution in the section. Time is one possible important confounder. In order to weigh whether the model should be modified to account for this, it would be important to see if there is evidence on the pooling assumption by checking for changing effects of loans. When the models relax this assumption, the result may also alter the predictions of the AI significantly.

### 1.1. Overview of Insurance Risk Assessment

Insurance risk assessment is a way that helps insurance companies determine the cost for potential risks. It is an important step for underwriting; it evaluates the features and habits of policyholders. The process of risk assessment consists of deciding the premium and policy content according to the risk level of the applicant or policyholder. Traditionally, underwriters in insurance companies evaluate the risk of a policyholder, such as the number of claims, according to many factors, such as region, age, health, and so on. Then, from actuarial science, they calculate the approximate tendency of policyholders to relate to specific risks from historical data and determine the price of insurance. However, this traditional way of predicting the future only depends on historical information. Therefore, emerging technologies are important to gather information to be beneficial for risk.

When the insurance business is connected to emerging technologies such as AI and the insurance market filters into new models, the insurance industry also experiences radical changes. InsurTech involves the addition of new age technology to the insurance business, which operates with the traditional insurance model to bring a better solution for customers and policyholders. Present insurance risk assessment involves deploying sophisticated models constructed by domain knowledge expertise applied in the data science field. With the growing domain knowledge expertise, risk scoring techniques are also evolving. However, they still hold alleged feedback mechanisms like any heuristic mechanism's downside. They generate the risk profile based on historical parameters evaluated and hence miss the current possible effects adding to the risk profiling; thus, they affect the insurance sustainability principle. Therefore, there is always a need for monitoring, rectifications, and validation in the mechanism.

## 2. Fundamentals of Machine Learning in Insurance

Machine learning is a branch of artificial intelligence (AI) that deals with the development of algorithms whose performance on a task improves with experience. In the insurance sector, this means the ability to analyze gigabytes of data and extract potentially useful insights about the future by learning from evidence. In this document, we use 'insurance AI', 'insurance analytics', 'insurance machine learning', and similar phrases to refer to the use of machine learning in the insurance sector.

Different AI algorithms learn in different ways. The two main learning paradigms are supervised learning (SL) and unsupervised learning (UL). In SL, the algorithm receives a labeled dataset together with a measure or object to be learned and is trained to make predictions. In UL, only features are received, and the algorithm tries to discover some hidden structure in the data itself. There is also a wide range of intermediate learning paradigms, such as semi-supervised learning, reinforcement learning, and more. Practical experience shows that in the insurance industry, as in other sectors, SL algorithms often deliver superior results. Supervised learning methods, particularly the combination of statistical models with heuristics and expert knowledge that characterizes the field of data mining, are therefore the main focus of our attention in this document. The main reason the insurance industry is only now taking steps towards the development of AI learning systems is that until recently, the data they have collected about policyholders, claimants, and the world in general have been too sparse. Machine learning algorithms require very large quantities of data, and the methods used to obtain them must be at least reasonably stable over time. Systems such as those underlying uplift modeling, used to model the aggregate behavior of theft-ridden areas, are therefore unattractive for the development of long-term revenue intelligence techniques. In practice, this constraint also prevents us from constructing AI systems to interact with important parts of the insurance supply chain, such as the damage-control underwriters. Even though these subcontracts negotiate very few claims per year, each tolerance or refusal decision plays an enormous role in shaping the future risk portfolio. Then in 2010, a transformation began. European anti-money laundering laws obliged all insurers to collect and enter into a common data-sharing database hundreds of millions of dollars' worth of personal claims data.

### 2.1. Supervised vs Unsupervised Learning

Machine learning algorithms can be divided into two main categories: unsupervised and supervised learning. In supervised learning, models learn from a labeled dataset, using labeled data to predict or approximate the outcome of unseen data. In other words, the model learns to perform complex computations before the desired output is known. In the insurance domain, labeled data consists of historical policy data, including personal, policy, and portfolio data, from which the output is known. By learning from past voided policy data, the model can predict whether similar future policies will be voided. This approach aims to support risk assessment practices. Different from supervised learning, unsupervised learning models deal with unlabeled data. The models extract hidden patterns and relationships in the data based on input. The technique is also used in the insurance industry, for example, for customer segmentation or fraud detection based on past claim transactions.

The strength of a supervised learning method, when applied to the accumulation of unsystematic risk, seems to be the company's capability to provide sufficient labeled data and the model's generalization towards very similar data facing similar or equal problems. The unsupervised learning method, on the other hand, can only be applied if the model can identify more data points or relationships between input and output requirements than a labeled set could guarantee. For example, an unsupervised model can be applied where we want to identify a common group of new upcoming customers facing a very tailored problem, where labeled data is hardly available yet. This needs to be ensured synthetically, through hypotheses. The unsupervised learning method can also be applied to create reduced forms or synthetic training datasets. The following allows for common pre-processing methods, which can be easier due to a less complex behavior in unsupervised rather than supervised cases.

## 3. Applications of AI in Insurance Risk Scoring

Artificial intelligence (AI) contributes significantly to the optimization and enhancement of many processes in various domains. However, the most transformative features of AI, such as artificial neural networks (ANNs), have been increasingly explored in many industrial sectors relatively recently. Today, they are primarily employed for the identification of patterns and regularities in large data samples, which eventually aids in the prediction of potential dangers or trends. These capabilities have been leveraged to make significant improvements in insurance risk scoring processes. In many cases, traditional scoring is based on expert judgment, questionnaires, or focused manual assessments of financial or asset-related data. Data that has been assessed is converted into numerical scores by summing up points allocated according to predefined criteria and then adjusted for additional risk factors or discounts.

AI potentially provides a number of exploitation paths. From the process point of view, AI-segmented risks reduce the volume of files that underwriters see and rate. The risks suitable for automation are processed without human intervention. Commercial underwriters will not be involved if their risks and attachments are standard. This means a free flow for new business and a reduction in time-to-quote from days to minutes. Underwriters are primarily contacted by risks identified by predictive analytics as complex or showing unusual behaviors. Scores visualized by technology in the hands of producers and brokers will further improve their knowledge and help build a better closing strategy. Finally, AI-generated information will help to improve loss control services and better assess the loss adjustment reserve. Moreover, data analytics and AI will help actuarial departments set pricing strategies and more dynamic and competitive rates that are consistent with profitability.

### 3.1. Automated Underwriting Systems

Insurance is a trust-based economy where an insurer promises to provide financial protection in case of potential future events or losses. The risk-taking is limited, and insurers earn a very low profit margin, which is why optimizing operational costs is so crucial. Traditional underwriting systems, which rely mostly on manual processes, are time-consuming, leading to longer processing times and manual errors that could result in legal and financial consequences. Automation in the insurance industry may also reduce human bias in decisions, creating an equal and fair evaluation of each application. Today, some insurance carriers are using automation for pricing and underwriting risk products in real time. An automated underwriting system utilizes real-time data feeds from vehicle sensors, flight patterns, and shipping tracking data in these processes. Such data-driven insurance can allow pricing flexibility for new and safer risks that can be priced dynamically and more efficiently.

While analyzing these implementations, it must be stressed that the algorithms are only used in a given 'fair' application space. All outcomes are examined by underwriters who have the authority to override the system based on further tests. While using AI, the growth rate is mainly due to the shift in the methods used to solve these problems. The ambiguity present in the decision-making process or changes in incoming variables and processes makes traditional programming methods difficult to implement. The underwriter of the future is likely to leverage AI algorithms in pricing the more commoditized insurance products.

Automating the process not only gives the ability to reduce the time spent evaluating these risks and allows for more time devoted to analysis, but it can also remove human errors from some of the decision-making processes. Additionally, consumers are likely to see a reduced long timeframe, leading to a more pleasant interaction.

## 4. Challenges and Considerations in Implementing AI for Risk Scoring

When implementing AI, there are numerous factors and challenges that a company faces. One of the main challenges that needs to be addressed is the data on which the algorithms act, as in practice there are weak points concerning its quality, the interpretation of the results, and its long-term use, but also the proof of complementarity between techniques and the definition of the criteria on which to rely to judge the interests of adopting a method for innovating Risk Score. Another major challenge concerns the ethical use of AI. Indeed, the use of AI raises questions in terms of respect for privacy and the protection of personal data, as well as on the transparency and explainability of the decisions made by the algorithms, which are essential to establish the trust of the various stakeholders.

In the face of new and morally hazardous types of behavior, new forms of prevention should emerge, and we must therefore prepare for them in light of the current challenges. There is, however, always a certain resistance to change in a company, on the part of both business leaders, managers, and the employees themselves. Moreover, such an initiative may not be appropriate if the company is not "culturally ready" enough and does not have the necessary means and skills for this new mode of operation. To implement this project, an appropriate plan and process should be set up. It is necessary that employees well-prepared and sustainably integrated into change be trained in the use of new technologies: employees must be aware of the risk associated with this data model management and the importance of respecting choices, habits, and rights.

### 4.1. Data Privacy and Ethical Considerations

Data Privacy and Insurance. Ensuring data privacy and regulatory compliance is a top concern. The insurance industry has enacted legal and ethical guidance so that any data collected from the consumer (such as internal, external, sale of data, or biometric information) is held to industry standards for governance, confidentiality, availability, privacy, and security. This includes the latest regulations that are implemented internationally in various regions as well as other privacy laws in multiple countries. These regulations provide accountability to effectively reduce the process time and limit any damages to its citizens in terms of data breaches.

Algorithmic Ethical Considerations in Artificial Intelligence. When using AI to automate insurance risk scoring, the following ethical and privacy considerations relevant to this model need to be addressed. Fairness and accountability are among the top concerns for insurers and consumers. Fairness ensures that people who are similar are treated the same in insurance risk scoring. This includes using both structured and unstructured data. Accountability for the criteria used to assess risk and charging, as well as the ability to audit these, would result in more open and ethical practices. Bias needs to be addressed within the AI scoring model. Descriptive explanations in the AI models are essential as opposed to the model being defined merely as a "black box." Furthermore, delivering critical information directly to those facing the issue would be crucial. Given the rapid growth of new entrants into insurance, it is important to have an ongoing dialogue around ethical and privacy oversight regarding the use of AI in personal insurance. Indeed, there is an ongoing conversation at a transnational level to develop regulatory frameworks. At this juncture, the industry sources consulted in this study were mindful of existing and emerging regulations. For example, the use of AI in insurance in certain regions is governed by regulations that specify the minimum requirements that affect how AI techniques are used with regard to fairness, transparency, and accountability.

**5. Case Studies and Best Practices**

Our case studies illustrate different successful AI implementations in insurance risk scoring based on both quantitative results and reflections from the companies. All implemented AI-based technologies were able to improve risk assessment in terms of predictive performance and efficiency. Initially faced with skepticism, they ultimately became accepted as part of the companies' standard sales and underwriting process and initiated similar projects in other areas of the company. If you are an insurer, it is important to understand that an off-the-shelf AI product promising improvement in predictions may not be the best solution for your organization because of different starting points in the use of risk scores, standards for neural networks, and requirements for surveillance. However, our case studies provide valuable insights and key takeaways from practitioners for insurers considering an AI project. These best practices can be summarized as follows. Firstly, choose an AI solution tailored to your business model, specific customer needs, and risk assessment process. Secondly, choose an AI solution to address priority challenges or to improve the most important parts of the insurance application process. Finally, plan ahead for the use of AI and set up your monitoring and maintenance system to optimize your solution over time and keep it up to date.

### 5.1. Success Stories in AI-Enhanced Risk Assessment

AI risk assessment models can achieve tremendous improvements over standard rating tables. Thus far, numerous practitioners have reported successful adoption of artificially enhanced risk scoring. Below, we present three success stories, put these results into context, and discuss each case's specific challenges and success factors. The case studies are based on expert interviews, and we accompanied data from first-mile metrics to illustrate the success in concrete in each case. Before AI, risk scoring was measured in terms of explained variance, and the resulting improvement reflects the gain in accuracy. Following AI adoption, bottom-line improvements like reduced claims costs are more relevant and direct KPIs.

A common thread across the stories was the importance of involving many stakeholders in the projects and sharing the insights AI delivers. For Petplan, a direct-to-consumer insurer, AI analysis of different locations' risks prompted a radical strategy change: Petplan sold five branches to reduce the risk cluster that the data had unveiled. Petplan was keenly aware of the need to avoid possible regulatory consequences of these practices and emphasized ethical considerations. Another key determinant of success was the capacity to pilot and scale like the device insurer did. They initiated the project but thereby only impacted the fringe of their risk underwriting. When they witnessed success, they ramped up. It is crucial to present the WorkPlan as a high-potential journey, not a high-potential threat.

## 6. Future Direction

The journey of insurers in integrating AI into their individual risk strategies is just at the beginning of a very exciting phase. The trends discussed in this section are anticipated to further shape how insurance risk assessment of the future will look. The connected and mobile devices are expected to integrate with technologies to provide robust and accurate data in real time. These devices are expected to employ new machine learning algorithms, data analytics, and AI systems. Such systems are expected to use facial expressions, eye tracking, EEGs, and GSRs to identify the specific emotional characteristics of an individual, which are often not reported or misinterpreted.

The increased market demand for ethical and responsible AI underlines the necessity of maintaining a diligent and robust governance, risk, and compliance framework. The framework is expected to gain prominence over the next few years. The new control and risk framework is anticipated to shift its focus to manage AI ethics risk and identify data management weaknesses. To maintain competitive advantages in the era of AI, continuous learning must be core cultural values. Insurers are expected to embrace technologies that can use large datasets for making predictions, hedging bets, and pricing depending on real time data. Banks, governments, and security agencies are expected to use innovation to identify and prioritize different fraud risks. By leveraging big data, it is expected that the insurance industry will adopt a new approach to underwriting and reinsurance. In the underwriting stage, the insurance industry is anticipated to expand its techniques in profiling, segmentation, and pricing, with systems that are forecasting-based and market-driven. This is in addition to including the ability to transact through digital channels while managing channels that are affected by COVID-19.

## 7. Conclusion

The present report has explored numerous aspects of the topic of the enhancement of insurance risk scoring with AI. The current state of AI and risk scoring is underpinned by the recognition of the importance of integrating authentic AI into systems, instead of being stuck with merely rules-based algorithms. Successful AI systems exhibit exceptional accuracy in the face of uncertainty, which by its very nature is afforded to the world of insurance. AI-based systems are capable of preventing new types of fraud, though there are certain structural issues involved with the pairing of AI and rules-based methods in insurance operations. Along with the primary components of AI and risk scoring in insurance, this report thematically considered a set of challenges involved with the topic. From the perspective of ethics, inequality and insurance discrimination are focal points. Practices such as undesirable risk pricing are also considered ethically troubling. Transparency in rule-making, and more importantly, in data, is identified as being mandatory in order to foster consumer and stakeholder trust. Displacement of low-earning employees could be a near-future consequence with regard to the deployment of AI-based systems in insurance. A lack of consumer awareness underpins customer resistance towards interconnected insurance e-commerce AI systems. It is suggested that a regulatory code can be employed to manage and foster a general understanding of the legal consequences of using interconnected e-commerce and AI.

In conclusion, the section argues that technological flux means that insurers must constantly update ethical considerations in conjunction with continuous learning to keep in step with regulatory and industry changes. AI is necessary for successful business for both customer services and for internal operations.

AI is a necessary component of the future for insurance, wherein the goal is not only to optimize new insurance products, but to also enhance customer service and improve operational effectiveness.

**Reference:**


1. S. Kumari, “Cybersecurity in Digital Transformation: Using AI to Automate Threat Detection and Response in Multi-Cloud Infrastructures ”, J. Computational Intel. & Robotics, vol. 2, no. 2, pp. 9–27, Aug. 2022
2. Tamanampudi, Venkata Mohit. "Automating CI/CD Pipelines with Machine Learning Algorithms: Optimizing Build and Deployment Processes in DevOps Ecosystems." *Distributed Learning and Broad Applications in Scientific Research* 5 (2019): 810-849.

3. Machireddy, Jeshwanth Reddy. "Data-Driven Insights: Analyzing the Effects of Underutilized HRAs and HSAs on Healthcare Spending and Insurance Efficiency." Journal of Bioinformatics and Artificial Intelligence 1.1 (2021): 450-470.

4. Singh, Jaswinder. "Social Data Engineering: Leveraging User-Generated Content for Advanced Decision-Making and Predictive Analytics in Business and Public Policy." Distributed Learning and Broad Applications in Scientific Research 6 (2020): 392-418.

5. Tamanampudi, Venkata Mohit. "AI and DevOps: Enhancing Pipeline Automation with Deep Learning Models for Predictive Resource Scaling and Fault Tolerance." *Distributed Learning and Broad Applications in Scientific Research* 7 (2021): 38-77.

6. J. Singh, "Combining Machine Learning and RAG Models for Enhanced Data Retrieval: Applications in Search Engines, Enterprise Data Systems, and Recommendations", J. Computational Intel. & Robotics, vol. 3, no. 1, pp. 163–204, Mar. 2023.

7. Tamanampudi, Venkata Mohit. "AI Agents in DevOps: Implementing Autonomous Agents for Self-Healing Systems and Automated Deployment in Cloud Environments." Australian Journal of Machine Learning Research & Applications 3.1 (2023): 507-556.