# Stochastic Portfolio Modelling Under Uncertainty: Machine Learning Approaches to Systemic Risk Assessment in Insurance Portfolios

*Dr. Andrei Tonkoshkur, Associate Professor of Computer Science, Belarusian State University of Informatics and Radioelectronics (BSUIR)*

## 1. Introduction to AI in Insurance

The prosperity of the insurance industry is quite sensitive and may be questionable due to indirect factors that are hard to control. To improve their positions in the market and increase profitability, insurers pay attention to adjustments in the insurance portfolio and the proper classification of risk. The risk of an insurance portfolio is driven by various factors, and the risk profile of a class should be a weighted combination of the mean and variance of several variables. Artificial intelligence is making inroads into the insurance industry. AI can be transformative and bring acceleration in the process of various human development activities. Insurance can leverage AI technologies to improve traditional regulatory and risk prediction, policy generation, and pricing analysis in order to offer customized portfolio risks as well as the consumption habits of customers.

In our research, we address AI-based insurance portfolio risk management, i.e., we focus on the problem of how to apply AI models and techniques in the insurance domain to properly evaluate not only the risk of a single consumer but also the portfolio risk. AI can contribute to optimizing the consumer susceptibility assessment by reasoning from many sources of knowledge. Deep learning can be used—more specifically, RNN and LSTM neural structures—to support insurers in forecasting the expected portfolio risk. AI insurance portfolio risk management is the subject of this research, which is applicable in various insurance fields, including health insurance, life insurance, and non-life insurance. In our studies, explanations of the different predictive models used in many AI areas are also presented, i.e., earlier different AI techniques of applications such as neural networks and optimization algorithms were used as well.

### 1.1. Overview of Machine Learning in Risk Management

Risk management is a fundamental concept of modern insurance. An insurance portfolio manager, known as an underwriter, is responsible for effective portfolio risk identification, measurement, pricing, and mitigation. Machine learning technology is one of the approaches to help improve this process. Machine learning is a subset of artificial intelligence that leverages algorithms and systems to identify inherent data patterns and explore them in an effort to predict outliers, opportunities, risks, and mitigate them. The most common task of basic machine learning applications is to classify data according to specific rules or forecast a certain variable value. Some more exotic models rely on clustering of data pattern recognition and anomaly detection as a means of insurance fraud prediction.

The key advantage of machine learning models is their ability to learn from experience or previously calculated cases and adapt to new or changing situations. Data processing capabilities of machine learning models are superior to human ones and could help sort through significantly larger scales of data. Lastly, there is an increased data precision, as machine learning algorithms can find obscure and rare patterns, which traditional risk estimators might miss. Underwriters were the first to adopt machine learning, seeking better risk estimates used in developing an insurance contract or setting the premium value required to operate as an insurance company. The process of categorizing records in different classes by subgrouping them in homogeneous risk groups is known as segmentation. As well as customer behavior predictions, many companies are investing in predicting costs derived from insurance claims. Examples include predicting the severity of an accident and credit default insurance claims caused by international defaults. A more general example is the estimation of the number of claims in a particular time frame. The same approach can be applied to the credit risk management field by predicting the amount of non-performing loans. More common applications include forecasting the probability of missing a payment for a loan. In the latter machine learning applications, the task is mainly classification, grouping records in the customer default rate they belong to. Risk managers also value unsupervised machine learning technologies that are able to cluster risk behavior; for example, at-risk vs. non-risk customers. Crucially, data quality remains the biggest risk in the model development process. Data bias is a close second because models must be continuously re-trained, yet

data bias often diverges at the source. Model validation never ends; it must continue throughout the full model life cycle.

**2. Insurance Portfolio Risk Management**

The portfolio of every insurance firm is a collection of risk and return profiles as seen in a multitude of contracts, each with a distinct regulatory regime. Insurance, as an industry, profits on the power of pooled risks. An insurer, beneath the law of large numbers and through appropriate long-term underwriting and portfolio diversification, can estimate a sufficient variety of ultimate policy amounts to establish premium rates that generally cover policyholder benefits, administrative costs, and profits. Insurance portfolio returns are outcomes of those forces that set premiums and could be summarized: a) Underwriting, or pricing and selection, generally reflects profitability and insolvent risk. Healthy risk management practices can produce profits from insurance operations. b) Mortality, morbidity, expenses, and policyholder payments are random and fluctuate around expectations throughout the contract range. Since a portfolio is characterized by the properties of each policy-based contract within it, this will result in addition to requirements in the fulfillment of obligations and in shareholders who require a specific risk-return profile. Therefore, portfolios also have to be managed in aggregate according to the specifications. These safety considerations have very significant damage impacts, and they are affected by considerable amounts of claimed loss from catastrophic external swings in their costs connected to the short-term risk. Negligence in handling such catastrophic danger that is never cost-saving may have catastrophic consequences. To handle these dynamics, several reinsurance organizations maintain a special, mostly distinct branch and sales profit and loss report. Traditional portfolio risk assessment methods have been used by insurance firms to evaluate the basic chance of policy underwriting and portfolio contact. However, it takes less into consideration internal and external operations that prevent organizations from knowing these risks rather than ultimate amounts of policy received. In an accident and health insurance field containing risk, the true claim for disability or loss of life may not be identified for several years. Trade-offs have therefore been slow in accommodating the various facets of risk. Nevertheless, the conversion to more data-driven techniques has been pointed out to be promising, which tackle limitations in the ratings. Super-scale calculations allow insurers to use simulations and bootstrapping to assess the distribution of actual risk and to investigate possible severity results. This allows us to

incorporate potential additive correlations between threats, which may be ignored in numerical techniques when utilizing methods of risk assessment.

### 2.1. Challenges in Traditional Risk Analysis

In the traditional insurance sector, risk analysis and policy development were based on historical data, such as past claims and valuable expert experience. However, this method had several limitations. Most of the traditional methods depend on historical data, which could not be applied to unseen events. Furthermore, these approaches cannot adequately identify all of the sources of risk as the world around us is constantly changing. Recently, the world has been changing rapidly, and information and markets are not as resilient as they once were. For a variety of reasons, companies now have data in separate silos rather than integrated systems. The collection of big data would show us future trends, but these types of data are frequently too technical and need to be converted into a form that is understandable. Statistical analysis was only for univariate data and did not consider the relationships between variables, resulting in a lack of sufficient predictive potential.

Risk in both a commercial and a strategic context is undergoing a significant shift. Risks are not only increasing in severity and frequency, but the range and combination of uncertainties are also expanding. There are developments worldwide, like geo-economic tensions and chemical weapon use. This is an existential risk that may decimate entire countries and regions due to increased interconnectedness. As a result, the insurance industry is beginning to find challenges with more intelligent AI-based tools, which are not exclusively the intellectual property of startups leading the dawn of a new insurance generation. The insurance industry distinguishes between predictable risk, e.g., correlated events or volatilities, and many non-predictable risks such as earthquakes and technological development. It is already apparent that the future requires more intelligent risk management in order to keep up with the complexity of these developments, as seen in 5G and AI, which could lead to added value for the industry.

## 3. Utilizing Machine Learning for Risk Analysis

It is increasingly becoming feasible for insurers to apply modern machine learning and artificial intelligence techniques in their portfolio risk management strategies. Machine learning algorithms can assist insurers in identifying risks and evaluating, forecasting, and managing future events. Some common machine learning tools include clustering,

classification, regression, anomaly detection, and ensembles, among others, and they have been used in diverse industries with compelling results. Appropriate application of these tools can drive organizations toward increased efficiency and growth. Improved analytics, clearer identification of leading predictors of financial fraud, and advanced modeling of customer behavior could help insurers to underwrite, evaluate, and price risk more accurately.

At its core, real-time data should enable earlier detection and response throughout the organization. Predictive analytics are instrumental in pinpointing prospective risks and providing insight to guide changes within a business, averting potential crises. By highlighting areas that are breeding risk, predictive models enable insurers to have time to prepare and protect against the undesirable consequences. The developers create machine learning systems in a way that enables the workers to pick up on the presented digital or real data in close to real time as they make separate decisions simultaneously with the computer, and then watch if their decisions are correct. Another benefit of machine learning as a portfolio risk tool is its capacity for continuous learning. Once it is trained by a large data set of certain scenarios, the model can then be used to predict future similar occurrences. As the training data becomes outdated, synthetic data can be used to augment the learning systems to adapt to the many changes in risk and growth already noted. Several case studies show that insurers have successfully utilized machine learning tools to improve their management of risk.

### 3.1. Data Collection and Preprocessing Techniques

1 Introduction

1.1 Data Collection and Machine Learning Techniques for Risk Analysis

Machine learning (ML) is a vital technology that has been employed in insurance risk analysis. AI techniques require data as input to solve a problem. The fundamental underlying principle for successful ML system design is to acquire a dataset containing a large amount of accurate, relevant, diverse, and unbiased data. Data collected by humans, or secondary data, contain biases due to the absence of an objective technique to measure these phenomena. Consequently, there is a need to deal with the problems that arise from biases in the data. The "garbage in, garbage out" principle also applies: if the quality of input data is low, poor output will be produced. Techniques adopted as

preprocessing techniques focus on making data available for application in ML algorithms to apply risk analysis.

Data scientists often explore distinct sources of data such as sensors, databases, and text to solve AI problems. Data can be in the form of unstructured, semi-structured, or structured data. Data cleaning, transformation, and normalization techniques are often used to prepare datasets for ML algorithms. The essence and importance of crucial data acquisition techniques have been widely documented. Although data preprocessing techniques that have been developed over the years have been reviewed, they did not comprehensively address risk analysis in insurance portfolios. Machine learning (ML) research in risk management must be underpinned by sensible data preprocessing techniques to boost model performances. Moreover, acquiring data in the modern world is known as big data as it hails from several knowledge areas for analysis. Big data preprocessing, which is presented in this section, brings forth the evolution of technologies such as data warehousing, cloud computing, IoT, and NoSQL databases, which house structured, semi-structured, and unstructured big data. By exploring and combining primary and secondary data using these emerging technologies, researchers can acquire data for insurance risk analysis. Also, this section briefly presents tools that are complementary to data preprocessing in big data.

**4. Case Studies in AI-Based Insurance Risk Management**

Myriad are the use cases demonstrating the practical applications of AI-based technologies in order to improve insurance policies' risk management. There are several actual companies that have worked to improve their assessment capabilities either independently or through collaborations to develop the most adequate AI-based systems to suit their specific needs. Notable case studies in this area of investigation focus on the development of AI software capable of grouping different kinds of insurance policies, such as car insurance, health insurance, life insurance, and travel insurance, generally serving different groups of potential customers for various enterprises across many countries through overall risk assessments and cluster analysis.

A leading car insurance company has created the technology to keep a regular watch on their current portfolio and has been the first to integrate this tool into their pricing department. One of the most prominent aspects of such companies is that once the tool has been validated, carriers interested in the technology will have access to a database,

and indeed similarly to the tool, the company will not keep a database of all the uses of these services by operational companies. Like these companies, many are moving, collaborating with experts and companies in different sectors around the world to better understand global market dynamics and the optimal strategies and AI tools tailored to their specific needs. The main goal is to serve many types of clients spread across different countries, providing current assessments of the degree of portfolio risk portraying different states of clients and their relationships.

**4.1. Auto Insurance**

AI technologies have significantly transformed the auto insurance business regarding risk analysis methods in order to innovate new ways of underwriting and improving the claims process. The main innovations developed include telematics, or the usage of sensors and devices to collect data on the manner a car is used and driven, and usage-based insurance, or the pricing of a car insurance premium according to how the vehicle is used and specifically how the consumer drives, either by adopting a pay-as-you-drive approach, where miles driven are an input for the risk assessment, or pay-how-you-drive/manage, where the style of driving and harsh driving are the driver value and impact on insurance premium. These insurance business models assume the acquisition of data mainly through on-board devices and mobile apps, with further integration of AI capabilities.

Indeed, the literature shows two case studies that have been very successful. The first regards an American auto insurance comparison website that had more than 4 million registered users as of January 2021. By acquiring another InsurTech, it is further evolving by integrating a recommendation robot employing AI and processing data collected from driving licenses, registrations, or car insurance quotes to assist in the auto vehicle sales business. The latter is the cooperation between a personal insurance agency and a car manufacturer, which developed a digitized car-shopping and online insurance-search integration in the U.S. to suggest one or more insurance deals to potential car buyers. The partnership processes dipped into drivers' telematics histories and came up with customized vehicle purchase suggestions and quotes for car insurance policy savings based on driving behavior. Based on the classification chosen earlier for the table of AI strategies, business models of AI applications in car insurance were mainly about dataset analysis for risk identification of policyholders, drivers, and

damages. This delicate paradigm shift involves, among others, the risk of the "black-boxing" of drivers with the consequent distortion of competition in the distribution of car insurance products on the insurance liberalization market. However, a few doubts in this regard do not invalidate the positive effects of AI in the auto insurance business. While the sharing of driving data allows the adoption of a behavior-based pricing model, usable data produced in the event of an accident can optimize the claims adjusting process and prevent fraud attempts. For example, telematics data can be used at claim time to reconstruct the mechanics of the car crash, providing timely responses on the responsible or accountable party in direct insurance lines. More generally, they are helpful to predict the outcome of the claim: this is particularly impactful in the case of car insurance, where the cost of claims is much higher than the cost of the policy. Also, vehicle theft has been successfully mitigated with the integration of telematics and machine learning that would detect anomalies in the cars' use. As a result, these systems allowed for quicker detection and recovery of stolen vehicles and more accurate reconstructions.

**5. Future Trends and Implications**

5. Future Trends and Implications

Over the last two years, AI has seen rapid advancements, and it is anticipated that some of these will have an impact on insurance practices in the near future. As these trends continue to permeate the area of customer-facing insurance, the changes in risk management might become more pronounced.

Personalized AI-based insurance products are making personalized insurance products available, given the advanced predictive power. Based on this technology, some insurers are offering usage-based policies or pay-as-you-go insurance, which may benefit policyholders who pose a lower risk. On the contrary, AI technologies might be used to identify policyholders who claim more often and, as a result, will be charged higher premiums. This may raise questions about auto-alignment to a preferred risk charge. The customer journey in insurance or claim analytics, including enriched data through AI, is becoming easier to analyze, facilitating improved customer experience and operational efficiency. Retail insurers might use AI for fast and pleasant customer experiences, from onboarding to a data-driven pricing proposal and claim handling.

Regulatory challenges in a virtual future insurance landscape are evolving along with the trajectory of regional regulators and mainstream regulators to protect the customer. After privacy concerns, the insurance industry must manage customer opinions and new insurance regulations. In order to maintain customer trust, businesses have to take into account AI regulation. AI applications will transform businesses' insurance offerings in terms of product and service offerings and the customer journey. However, it also calls into question the underlying ethics, particularly with regard to data protection, privacy, and transparency. Furthermore, several AI applications may be discriminatory, which would be ethically and legally unacceptable. AI solutions are evolving in such a manner that the insurance product is becoming a companion in the customer's journey, thereby making the pricing affordable as well.

### 5.1. Ethical Considerations in AI Implementation

It is important to consider the ramifications of implementing AI in risk management, as there are various ethical considerations that one needs to take into account. One of the primary concerns with AI is customer data and the possible violation of privacy. Considering the sensitive nature of such goods, there is a very high necessity for legal instruments to ensure protection for customers. It is further important to ensure transparency in advanced AI systems to improve the understanding of their decision-making process. Manufacturers of AI-driven systems also have to deal with the concept of fairness, with the likelihood that allowing such systems to be developed and implemented would lead to customer dissatisfaction and a decreased level of trust. One can observe significant public debate concerning bias and unfairness in AI applications. The introduction of AI in cost estimations and claim settlements would, apart from other concerns, extend these discussions to liability.

Unfair discrimination is frequently linked to concerns over AI-based algorithms. Policymakers and organizations advocate that these guidelines should be considered and addressed to account for AI-specific issues in areas as important as life and health insurance. However, it is predicted that updates would be issued in an unexpected way, every 3 to 5 years. The Council recommends that the entire industry undergo initial and ongoing training as part of its ethical adoption and utilization of emerging technology. Based on the framework concepts outlined in this text, this should include a thorough understanding of the potential risks and benefits of AI systems, as well as integrating

ongoing ethical assessments of tools in use to grasp the changing contours of risk in their evolving applications. Robust, inclusive, and ongoing stakeholder engagement, including regular consultation with customers and policyholders, will be crucial to earning trust.

## 6. Conclusion

The paper addresses the application of AI technologies for solving one of the core tasks in the insurance industry—insurance portfolio risk management. It provides an overview of the existing state-of-the-art in insurance data analysis, discusses various applications for AI and machine learning in enhancing insurance risk management, and presents a case study for enhanced insurance portfolio valuation by using an unsupervised neural network model. The paper also outlines challenges in the deployment of machine learning in the insurance industry, such as regulatory requirements, ethical issues, transparency in decision-making, and challenges in the explainability of models. The paper builds on the existing research and offers a comprehensive analysis of AI solutions in insurance. It also provides clear directions for future studies and applications. In conclusion, this paper clarified the background and status of insurance portfolio risk management. We not only showed the partners a lot of knowledge in AI and data science fields, but also provided various solutions at the business model and algorithm level. The era of AI is coming in the insurance industry, and many traditional insurance practitioners are making various attempts to create new opportunities and explore the possibilities that AI technology can offer. The combination of emerging technology and insurance policy on risk management will bring transformative changes in insurance products and customer perception. It is an inevitable trend for financial institutions to upgrade their risk prevention tools with the help of machine learning technology. Even for competitive reasons, we will make significant explorations in this regard. The results of this research are rich in content at all levels of IT technical strategy. We hope AI and ethics practitioners in insurance focus on the standardized collection and analysis of insurance business data.