Dynamic Risk Classification and Premium Calibration Through Streaming Analytics: A Real-Time Machine Learning Framework for Insurance Underwriting Optimisation

Dr. Bengi Gönen, Associate Professor of Electrical and Electronics Engineering, Middle East Technical University (METU)

1. Introduction

Underwriting processes have moved from the company's headquarters to the cloud. At the same time, financial and insurance markets are ardently adopting artificial intelligence-based solutions, with those technologies and processes enabling superhuman performance in many applications. The data-driven approaches, and thus the AI, are gradually yet significantly overthrowing the old deductive manner, which may have seemed to be quite economically justified in the recent past. Seeking to accelerate and optimize your organization's decision-making process by capitalizing on groundbreaking technologies is thus pursuing an irreversible, ever-changing trend. Efficient risk assessment becomes of utmost importance in the dynamically changing, globally present, and technologically driven insurance-financial markets. Consequently, our focus will be on the companies that, whether for the aforementioned reasons or others, have or will be willingly aiming to implement our proposed AI-based tool. Our contribution hereafter outlined is intrinsically related to the underwriting and finance expertise. The paper is very much focused on underwriting as a decision-making process, and the particular task it comes down to consists of fine-tuning an underwriting model by uncovering inefficiencies of current underwriting practices from the AI perspective. Future and ongoing trends bring to the spotlight the need to shed light on how AI will affect insurance practices in the underwriting domain, especially in its quest for optimization. The underwriting decision-making process may be iconic of sophisticated problem formulation, wherein remarkable components of approximating a good decision model are identified, such as understanding and extracting data.

1.1. Background and Significance

Underwriting is a method of assessing the risk associated with insuring individuals and groups. Traditional methods used to assess this risk often include a questionnaire or examination that targets lifestyle, family history, occupation, the financial position of individuals, and a report from a medical professional following a physical examination. As a result of the high level of questionnaire and examination completions, it may take months for an individual to receive underwriting feedback. To address the needs of the fast-paced underwriting market, real-time underwriting is in demand. Underwriting data is integrated into a rule-based engine developed by actuaries, resulting in real-time decisions. The digitalization of underwriting facilitates efficient exchange of information. Given the ease with which the target market, customers, and applicants can provide information, the difficulty lies in the underwriting team being inundated with a large amount of information that must be processed in order to make decisions. The volume of information becomes overwhelming, and the underwriting team may miss incomplete or excessive information, jeopardizing the robustness of the predictive actuarial material in support of the application assessment. The desire to provide a fast underwriting service, followed by the accumulation of a large amount of underwriting information, increases the opportunity for manual errors to be made. The modern underwriter must consider advanced tools from data mining, machine learning, and artificial intelligence to innovate their underwriting guidance or pricing.

The underwriter analyzes individual cases and assesses whether a proposal may be in jeopardy of falling into a more risky category. This prediction is based on previous customer interactions and trends. Information related to a customer's claim events is held by a single insurer and is not widely available. AI technologies can enhance these prediction models to improve the accuracy of predictions statistically. Some applications directly automate the quoting process, reviewing the underwriting materials, and providing price quotations based on specific customer applicants. To do this effectively, AI must handle complex relationships derived from customer data and their historical behavior, which are stochastically similar over time. Machine learning models are a key area of AI used for underwriting applications. In machine learning, the system learns from historical data and fundamental algorithms, and then applies these techniques to new, unprotected data, improving the accuracy of applications. Some of the standard machine learning subfields used can be in neural networks, deep learning, support vector machines, and random forests. The target market is based in Australia, and the research investigated a group of twenty insurance policies. Carriers have a predefined constant premium, and the research is focused on the acceptance criteria where an underwriter identifies a constant variable. Moreover, both fully underwritten and accelerated cases are studied.

2. Foundations of Underwriting

Underwriting is a critically important part of the spectrum of risk management. The function of underwriting is to ascertain the level of risk in an insurance proposal so that the insurer may either accept or decline the proposal. An additional objective is to predict the proposed loss frequency and loss severity, where the loss severity actually defines the payable claims to be distributed among loss events. The premium that an insurer charges should meet the following conditions: it should be reflective of the loss frequency and severity, and its probability distribution in order to meet future losses; and the revenue from premium ought to be consistent to cover the present value of the amount of claims or the expected loss.

In traditional actuarial loss modeling, it is assumed that the premium rate per exposure unit should be actuarially modeled according to the following principle: it is a linear function of the frequency and severity of the loss process (per unit), and it is loaded with a defined percentage in order to depend on future investment income. Where the principle of traditional insurance views underwriting as a linear model to calculate the premium as the sum of future losses and the insurer's margin of profitability, real-time underwriting assesses the nature of risk for pricing of competitive premiums. Therefore, it is based on deep learning as the prediction system in practice. One critical success factor for the improvement of underwriting policies in data-driven underwriting models is the validity of the risk data used in the estimation of the risk and consequently the eligible premium. From a regulatory point of view, insurers have to base their risk data and claims models on actuarial statistics from entities in the same risk portfolio, the same B class of monthly paid employees, or a temporarily established risk group.

2.1. Traditional Underwriting Practices

Historically, underwriting decisions have been influenced by both the data available and the methodologies that have been in use. When data were scarce, underwriters had to rely on manual evaluation of specific individual items and on blending in their underwriting capacities. Although this straightforward process was successful for a disposable workforce composed of highly skilled underwriters, the large number of people to be insured and the quickening of human life have required a cost-effective solution, scalable to millions in a few minutes. This drove the adoption of electronic rule-based underwriting. This simplified the practice, removing subjectivity from decision-making and refining, standardizing, and automating processes. Research showed that intended subjects were mostly healthy and that underwriting discrimination drove costs down by around 50%. Some early underwriting engines verified applicant responses to health and data questions, summing up the danger probabilities. This ensured that underwriting applied only to those avoided by rules, such as maximum age and amount limits, ensuring that the entire underwriting process was within a particular risk tolerance.

Underwriters were at ease with rule-based underwriting, but the principal challenges were in the efficiency of the process, the inability to expand products, the difficulty in enhancing the speed of operations, and a reduced number of risks classified as substandard. Additional challenges were found in the difficulty of maintaining the rule base for horizon products, insufficient initial applicants, or case submission anti-selection bias in the trial plan. According to recent research, 33 to 50 percent of underwriters were comfortable with the time, although 31 to 37 percent stated that it exceeded the planned time for an initial or improved case. According to analysis, 92% of clients highly value the ability to speed up the process.

3. Machine Learning in Underwriting

3.1. Introduction Pattern recognition and decision-making from data are the ultimate goals in data analysis, which has evolved into machine learning. In recent years, machine learning has made great strides in various industries and has been widely used in a number of fields. In underwriting, machine learning can be used as it is deemed principally beneficial in risk assessment and risk classification. Many models and techniques can be employed based on the amount of data and time in which decisions have to be made. Among the many machine learning algorithms, the success of many applications can be attributed to ensemble models that average the prediction scores from underlying predictive models. However, neural models and deep learning applications are usually used when dealing with highly diverse data, similar to individual life underwriting, needing more dimensions, such as text/image processing.

There are many case studies of machine learning approaches applied in underwriting, ranging from predictive modeling for risk stratification, developing a model to forecast daily mortgage prepayments, to claim fraud detection, and detecting dementia or Alzheimer's disease using non-invasive ocular imaging. Most of the trial challenges have highlighted this need for scale and adaptability. New models that can be developed within a tolerance for increased financial costs but not necessarily in a shorter time span are able to process a varying array of data inputs successfully. One primary benefit to companies embedding this AI-based underwriting technique is the opportunity to consider and evaluate data sets that would not have been previously available. Where data is provided, the in-built consideration into the model needs to be given to this data due to the potential for duplicitous information being presented within it.

3.1. Applications in Risk Evaluation

Rapid developments in machine learning and AI techniques have allowed data scientists to build models that are better at risk evaluation than traditional statistical algorithms. The availability of large and diverse data makes statistical analysis very complex, resulting in reduced accuracy of the risk score. With a machine learning algorithm, a large dataset reflecting different information relating to numerous geographical areas and demographics can be analyzed relatively easily. Techniques such as predictive analytics and neural networks enable the handling of more diverse and complex data, which identify connections at the required level of accuracy. Thanks to their ability to process large datasets quickly, machine learning models are well placed to enable real-time data processing to improve risk assessment accuracy, as they are able to assess new data immediately when it arises. In addition to enhanced risk evaluation accuracy, automated risk assessment can streamline underwriting as a whole, reducing time and operational costs. Further, the process can be more secure and able to uncover more fraudulent activities in applications. Continuous learning features in machine learning algorithms can improve model performance over time, instead of being fixed like traditional statistical models. In general, the use of artificial intelligence and machine learning in risk evaluation has enabled major advancements in insurance underwriting.

There are concerns surrounding the use of machine learning to evaluate applicants when these rejections may be based on data considered sexist, racist, ageist, or can be in other ways disputed. There are multiple debates and concerns regarding the use of machine learning models for risk evaluation, based on the potential biases that can be included in the housing data used to train these models. Given the controversy surrounding their use, it is essential that the benefits of these models in terms of enhancing risk assessment and reducing operational costs and time be weighed against the ethical difficulties they may cause.

4. Real-Time Underwriting Systems

Real-time underwriting refers to an underwriting system capable of producing decisions on risks or quotes within a timeframe not appreciably bounded by absent underwriter or consumer availability. Real-time has long been a capability of technology, where latency times between transmitting and receiving information in computer systems have approached continually decreasing values. But over the last decade, cloud computing and software infrastructure tools have enabled underwriting systems that can run many generated models in parallel, combine the results, enforce the underwriting guidelines, customize the conditions as needed, re-grade the quotes or risks, and issue the report of the transaction nearly instantaneously. This power has made it possible to build applications that support underwriting as well as case-specific advice and guidance to high-volume cases, providing the key speed and accuracy we have historically associated with personal services such as an individual relationship with an underwriter.

For insurers, real-time processing inside a rate generation system can lead to higher intake rates and more consistency along with the ability to detect and change adjusted factors, leading to quicker quotes. Over time, this real-time rate adjustment will also correspond more closely with the true cost because it does not need to account for the short-term variability in the market. It was once the case, and is still true in some cases, that the healthiest and least ill were once chosen first. The corrupted and severe cases received the underwriting attention later. Increasingly, markers and monitored health measures are being used to segregate risks as they are underwritten. A real-time underwriting system is better suited to this kind of incremental underwriting, making better use of any subsequently completed task until the final underwriting decision. An underwriting system using more instantaneous data can quickly localize underwriting complexity for the few cases that may not initially seem complex. The consumer should earlier in the underwriting process be confronted with somewhat more accurate knowledge about their complexity than previous decisions could bring.

4.1. Key Components and Technologies

A real-time AI-based underwriting system consists of multiple components and integrates up-to-date technological solutions to ensure efficiency, reliability, and high scalability of the system. First, the capability to handle and store terabytes of data is necessary. Typically, these are cloud-based or hybrid data storage solutions. The data processing capabilities in such systems are scalable, which is achieved through the use of multiple computing nodes. Very often, a combination of AI/ML algorithms is used for data processing. The AI module can determine whether an application needs to be considered as life insurance, critical illness, income protection, or hospital cash. Typically, AI inference is responsible for the underwriting recommendation based on the collected evidence and rules.

A client engagement tools component incorporates a user interface that interacts with the customer application platform for the application submission process. Additionally, this interface also assists underwriters in engaging directly with financial advisors or clients during the underwriting decision journey. Client engagement tools use existing out-of-the-box tools to advise the channels and sharing tools, combining them with domain-specific underwriting and risk knowledge. Data security is a key element, and strong encryption and effective isolation between services are employed to prevent data breaches. The introduction of a new real-time application requires close cooperation between deep technical specialists and software architects and engineers who have a vision of how to integrate a new technical stack with a complex and multi-system, possibly legacy environment. There is an ability to handle data analytics on the order of terabytes, helped by many independent compute nodes allowing parallel or distributed computations. The introduction of big data tools like this allows insurers to offer improved customer service and more efficient underwriting than before.

5. Challenges and Future Directions

The insurance industry faces some challenges in implementing AI-based solutions, primarily due to ethical, legal, and social issues of data and privacy, data quality, and cost for insurance companies. Implementation of any new system requires a certain amount of data. Reusability and connectivity are important for any data, but because of the nature of data or data quality, data are only allowed to be used for specific purposes. Thus, underwriting and claims data generated by insurance companies might not comply with some principles of this data sharing code. Insurance companies not only need experts in statistics and machine learning, and actuaries, but also need IT professionals to handle, develop, update, and maintain those systems.

Future Directions New research directions are required to address AI-based underwriting business issues, i.e., underwriting systems for specific insurance products, usage of telematics and wearable devices to assess policyholders, and systematic organization to fully utilize and exploit big data, customer-contributed data, and new emerging third-party data available from social media, smartphone apps, in combination. Participation and involvement of insurance companies, data protection agencies, actuaries, ML experts, and IT professionals in the research and contribution of their opinions can be a strength and can avoid some issues arising in practical implementations and reach some optimal viable solutions for everyone. The advantages of constructing a model including technology and domain expert knowledge in a short time, getting many ideas or the best possible "problem of truth" by using the experience and creativity of participants, and the model including actuaries will potentially quantify the uncertainties in the effect and effectiveness of the existing model/use of AI in the industry. There is more work needed to explore those papers in an organizational setting. AI systems in underwriting can bring benefits from at least two primary perspectives: customer services, by providing full coverage and improving service and speed of response, and operational efficiency, by providing cost-saving advantages.

6. Conclusion

To sum up, in the era of digital revolution, AI-based underwriting optimization has become a significant part of real-time technologies. The technology and scientific progress have the capability to provide powerful tools to automate revenue management operations in real time and thus are transforming the traditional underwriting process. AI and ML technologies continue to reinvent underwriting, with speed being at the heart of their development. AI enables us to impose real-time dynamic environmental changes and constraints for automated decision-making in interactive business processes such as underwriting. As firms continue to engage in product and customer targeting, optimizing human judgment and engaging more with customers from an underwriting perspective seems to be the logical course. AI technologies allow us to make changes to products and optimize risk for customer cohorts and not just the two or five big books of the organization. All these insights continue to propose that firms need to increasingly adopt new technologies to further improve the precision and complexity of risk selection for targeted customer segments, which are increasing on a monthly basis globally. There is a need to exploit the cost and opportunity that advanced technology provides. Their ability to provide data-oriented and analytics-based insights enhances customer engagement, customer lifecycle management, as well as redefining underwriting practices. It also addresses the financial services consumer's increased need for quicker, more efficient services with accurate risk assessment. The challenge is not the technology's capability, given its constant change and continuous innovation. Research and firm-level adaptation to practice, as well as the regulatory framework, will allow for best practices. There are growing concerns over the use of technologies that increasingly bear potential ethical issues. It is important, therefore, to investigate more collaborative approaches shared with industry players.