Deep Science Publishing
https://doi.org/10.70593/978-93-49910-74-4_1

DeepScience
Open Access Books

Chapter 1: Foundations of artificial intelligence and deep learning in the insurance ecosystem

1.1. Introduction to AI in Insurance

What is left to say about AI? The hype is over. In recent years we have seen AI as an emerging technology become a disruptive technology. What started as an exotic tool for techies and data scientists has become a must-tool for every industry requiring support at every operational step: data gathering, data cleaning and preparation, data enrichment, data-driven predictive analysis, optimization towards desired objectives, quasi-real-time insights delivery, etc. How could the Insurance ecosystem be an exception to this? In terms of data, the Insurance Ecosystem handles the largest data trove. In their core business model, Insurance Companies, Reinsurers, Brokers, and Agencies in Life and Nonlife run on risk quantification and modeling, as well as on anticipating customers' behavior. In the adjacent spaces around their core business model, Insurance Companies and Agencies exploit another massive data trove for upselling and cross-selling ancillary services around their core business offering: customer data captured during years of policy underwriting and claim processing (Ng, 2016; Chen et al., 2017; Panigrahi & Borah, 2021).

And yet… coming back to our previous statement, what is left to say about AI? The answer to this somewhat provocative question is simple: although the Insurance Ecosystem is no exception to the paradigm shift represented by AI and despite most of the businesses have been implementing AI-driven optimization and insights, we are still at the beginning of the adoption curve. Indeed, despite all the buzz surrounding AI, what most companies have been doing during the last 10 years is only the first phase of a journey that started with the Global Financial Crisis of 2008. The first phase was about adopting AI techniques to replace legacy approaches based on statistical methods largely due to the lack of a larger amount of labeled data. Most initiatives still rely on small-size

pilot initiatives, testing AI capabilities generically proclaimed as superior vs traditional methods. What the industry is slowly realizing is that AI is not a panacea but it is a new toolbox that has to be put in place (Ribeiro et al., 2016; Weng et al., 2020).

1.2. Overview of Deep Learning

Modern AI is primarily driven by Deep Learning (DL), with astonishing success across multiple fields. DL is deeply rooted in Artificial Neural Networks (ANN), a specific type of mathematical function approximator inspired by biological neural networks and related cognitive discovery. ANNs demonstrate near-universal approximation power on the function spaces they support with vastly fewer parameters than classical function approximators. However, ANNs were generally too impractical for near-term applications, suffering from high training time and low generalization accuracy. This changed when key engineering design principles were introduced over the last decade, leading to attentive DL architectures with hundreds of millions of parameters fed by massive training datasets on advanced Distributed Optimization architectures.

To better understand these architectural innovations, we first survey the ANNs' mainstream notions. Variations of ANNs thus far implemented represent functions mapping Euclidean vector spaces to other vector spaces. They achieve their function approximator status through the nonlinear composition of learnable transforms mapping vector spaces to and from lower-dimensional subspaces. The transforms are learned from given labeled function observations using stochastic gradient descent. DL, which leverages ANNs with additional internal higher abstraction layers to support greater data complexity, is a mainstream technological platform that enables large-scale large-data AI systems. Label observations are either supervised with domain labels or self-supervised.

Function Learning – The function approximating the natural landmarks for supervised sampling are posterior classes with the conditional distribution's related expectations if color and depth distribution are given as conditional parameters of total distribution. Such latent variable suboptimal supervision can be developed to be locally optimal up to arbitrary precision at any observation space location. For supervised classification, the landmark's associated score function, defined as the class probability log-derivative, allows efficient learning through backpropagation smoothing. A majority of applied researchers rely on the Mean-Field approximation which leads to popular cross-entropy training using betrayals.

1.2.1. Foundations of Deep Learning in Insurance

Deep Learning (DL), a specialized aspect of Machine Learning (ML), is receiving spirited attention in the wider AI, ML, and DL community in general. The insurance ecosystem is no exception. It is stated that the industry is transforming through data analytics as a new generation of insurers builds the capacity to collect, manage, mine, and act on massive stores of internal and external data, advancing to a new level of efficiency. There is engagement with peer organizations in developing a new platform to provide predictive modeling via trusted company data of the membership.

Behind the flurry of activity in developing a DL infrastructure, discussion is needed about the accepted foundations of DL so that developers can purposefully pick and choose platforms, tools, practitioner guidance documents, and potential classes and development teams experienced with existing tools, technologies, and good practices. The next two sections introduce elements of DL where deep is interpreted broadly — represented in its very title, Deep Learning. The examples use the insurance ecosystem to help direct decisions and efforts to fully develop capable systems early on without intervention-heavy wrangling later in a project. The nuggets found here are relevant for all domains.

Fig 1 . 1 : AI and ML in Insurance

1.3. Historical Context of AI in Insurance

Over the last several decades, Artificial Intelligence (AI) in general, and its subfield of Machine Learning (ML) in more recent years, has become a pillar of support for several industries. AI enables the automation of processes through the utilization of large amounts of data internecine with advanced statistical algorithms. The insurance industry is no stranger to this advancement in technology; insurance has long been considered a trailblazer in the application of data to the field of underwriting and risk selection. These earliest attempts could be considered the precursors to the introduction of AI in the insurance industry. The statistical techniques that were employed were considered revolutionary for their time. Unfortunately, they received little in the way of further advancements to build upon. Lo, and behold, several decades later the rest of the world "discovered" the power of these methods and rapidly embraced them in a variety of disciplines.

So, what happened during the intervening years between these two "discoveries"? For one, the advancement of technology in the form of better computing power, increased access to data, and enhanced statistical algorithms probably were the distant cousins to the initial adoption of "AI" like technologies in the insurance space. The term "artificial intelligence" was originally coined in 1956 and was initially focused on symbolic processing. Theoretical advances in the relevant fields culminated in the rise of numerous AI-based technologies such as expert systems, computer vision, and natural language processing. Research efforts supported these technologies, which were widely adopted across multiple industries. However, by the late 1990s, a "disappointment" occurred in the field, where the lofty expectations of AI technologies diverged from reality. Research funding dwindled. Edge technologies were still commonly used in the real world due to, ironically, being more reliable due to their simpler and proven success record.

1.3.1. Evolution of AI Applications in the Insurance Industry

AI is a relatively young field, with the first implementations only being applied less than 75 years ago. During its evolutionary steps, it developed and morphed into two factions: "traditional" AI, which is connected to what is called "Good Old-Fashioned AI" and is based on symbolic logic and knowledge representation; Deep Learning. Deep Learning, and its ancestor Neural Networks, are based on computational models that try to simulate the brain processing mechanism to solve complex problems (especially in the recognition domain, such as vision and voice).

Using the above taxonomy, we present the different applications of AI - and the respective evolution of both Good Old-Fashioned AI and Deep Learning - in the

insurance domain. Historically, AI has been foundational in the field of expert systems. These powerful products, generally structured in a knowledge base and an inference mechanism, were built during the early decades of this technology. For three decades from the late 1980s through the 1990s, the leading applications in insurance were expert systems devoted to decision making and media selection. They functioned as automated tools to allow claims processors to efficiently select the optimal sequence of decisions (each may have several alternatives) and corresponding media, including phone calls and faxes, needed to effectively move claims through their life cycles. Another early area of influence was data mining using statistical techniques.

However, during the first decade of the new millennium, there was a hiatus from the large-scale use of symbolic AI in insurance. This was a period when deep learning applications were also not deployed in the industry. We believe that this gap allowed for relative inertia in insurance operations. As a result, the insurance ecosystem was behind what happened in other industries at that time.

1.4. Key Components of AI Technologies

The major technology and algorithms that comprise AI and ML can be categorized into three broad classes: machine learning, natural language processing, and computer vision. It is also important to note that several other algorithms and technologies enable innovative and cool AI use cases, such as digital twins, simulation, and optimization, recommender engines, generative AI, and clustering and pattern recognition. In insurance, while all of the enablers are useful, it is often the case that the deep learning use cases mostly fall into three classes.

A Machine Learning algorithm is a method or a class of methods that finds insights and statistical learning from data. In the age of Big Data, as we are accumulating terabytes of data, developing the ability for our computer systems to learn from the data in an unsupervised way is becoming important. While machine learning is an important topic and is often used interchangeably with AI, machine learning is a branch of AI technology that enables computational learning from data through biological-inspired programming. All AI depends on learning algorithms to do the learning from data. In practice, there are a wide variety of algorithms that power various use cases and applications of AI, including supervised, unsupervised, semi-supervised, and reinforcement learning, natural language processing, computer vision, and recommender systems that are used for applications like credit scoring and insurance claim processing.

Natural Language Processing is the field of machine learning that supports the interactions between humans and computer systems through natural languages. Natural language processing allows computers to read, dissect, and understand all of the

languages that people speak. NLP encapsulates a combination of text analysis, syntactic and semantic analysis through annotated data sets, and text generation, which enables chatbots to imitate human conversation. At the same time, a particular subfield is speech recognition, which converts human speech into a format that machines can understand.

1.4.1. Machine Learning Algorithms

The key components of an AI model are the data, features or variables, and the ML algorithm. Feature engineering, or creating good quality features from data engineers' intuition, domain knowledge, and experience, is often the most important component that differentiates high-performance models in various real-world scenarios. Domain expertise, data selection, and data preprocessing are also very important parts of the ML pipeline. Selecting the right ML algorithm is necessary, of course, but often secondary. In some cases, a relatively simpler and less computationally expensive model can provide better performance than a more complex and computationally intensive one. Creating a good model is often an iterative process of preprocessing, feature engineering, and retraining of the ML model. Dropout, or reducing overfitting without diversification of the data, is a strategy used in the training of most deep neural networks.

An ML algorithm is a set of tools and techniques that tries to establish a relationship between the target variable and predictor variables. ML algorithms can be broadly classified into two types: supervised and unsupervised. Supervised ML methods create embeddings or quasi-representations of features that map the input variables to a single value or a class label. In contrast, unsupervised techniques try to cluster the features into mutually exclusive segments to use their relationships with each other to make certain predictions. The most popular unsupervised ML methods currently are clustering, variational autoencoders, and generative adversarial networks.

1.4.2. Natural Language Processing

Natural Language Processing, or NLP, is a collection of goals that aim to enable computers to take into consideration the human languages that we communicate in our day-to-day lives. Since these languages are not mathematical- or code-based, they require models of cognitive functioning. Some of these goals include general domain sentence parsing, question answering, mathematical angle analysis of gloomy matters, humor analysis and generation, generating praise or questioning someone in a sarcastic tone, cleaning, endorsing, or rejecting social media information, text question answering comprehensions, inferencing the intent behind a specific multi-line dialogue, etc. Like the full AI goal, these goals are a consequence of the functioning of both nature and nurture.

NLP is an interesting area of AI research because it directly deals with utilizing human language, which may be the most accurate and expressive medium that we have so far to record and code the concepts in our minds, as well as near and far reality. The desire to communicate in a common vernacular language predates the benefits of having done our AI engineering very many decades ago. The earlier efforts in AI went under the classification of "symbolic AI", perhaps best represented in the public forum by the two debut AI programs. Given that both of these natural language expert systems had been showing such long-lasting commercial exploitation in their narrow domains, an ancillary debatable question would be: what has prevented the more comprehensive AI true spoken/typed vernacular language conversational expert systems from being deployed in every day enterprise operations?

1.4.3. Computer Vision

Computer Vision (CV) is a field closely related to computer graphics and is focused on finding patterns in visual input from the world around us. To this degree, it is related to all of the other component technologies presented in this section of the essay: a camera is a sensor in a physical sense (like a microphone), the input could equally be received from a synthetic generation process, each camera has a different underlining operating envelope and processes must be developed for dealing with noise and uncertainty present in input data across that envelope. What makes CV unique, however, is the fact that we are processing visual imagery – something that is not structurally or temporally organized in the traditional sense of discrete, vector data. This means that while we can apply the same machine learning methodologies to a CV that we can to traditional data, the operating assumptions of those methodologies are working against us, and achieving good performance requires careful thought and perhaps additional assistance.

In addition, recent experimental advances in deep convolutional neural networks have led to amazing success in overcoming such assumptions across a very wide range of computer vision problems – so much so that the general tools developed can be “fine-tuned” for a specific application in only a few cases and reportedly available pre-trained CV solutions can often be used for generic – if not application-specific – feature extraction with little performance loss. Recent advances potentially mitigate, if not remove, some of the challenges of using traditional techniques in domains different from those defined in our training sets while keeping the advantages of working in those domains. Using such pre-trained models to extract features or performance optima from other domains that differ only at the level of input structure would seem a reasonable expectation.

1.5. Data Sources in the Insurance Sector

Insurance companies have historically sat atop a mountain of proprietary data but have leveraged only a fraction of this data in machine learning model development and implementation. As a result, most machine learning applications in insurance center around claims data – the richest source of insight into either the probabilities of interest or the economic consequences of those probabilities. For example, the details in claims records for automobile accidents could be used to develop underwriting models for personal automobile insurance or for parametric insurance contracts that cover the same risk for commercial property insurance. Similarly, claims data can be leveraged to rate construction defects or earthquake insurance contracts, among others.

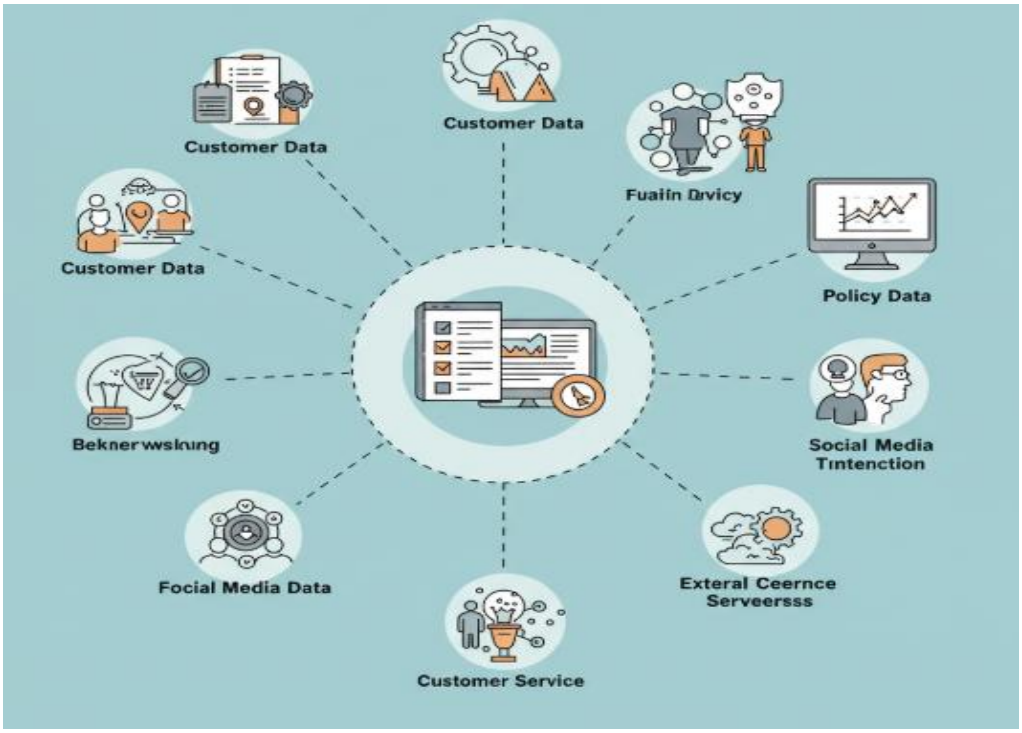

Fig 1 . 2 : Data Sources in the Insurance Sector

However, the historical claims exposures do not span exposures to all types of risks so they can never be the only source of information. Other historical data – in particular, information about the insureds or applicants – are also critical. This is an actuarial lens on data inputs in underwriting. A fate where historical claims data are the only predictors of loss severity happens when a set of actual and parametric insurance contracts are priced to have the same margins and are therefore useless to the insurer. Like any competitive pricing strategy, such data homogenization cannot persist indefinitely.

Claims data may not even be the most important input variables in every application. In particular, customer data are likely to dominate in Marketing applications. Marketing loss ratio estimates will be key for management decisions on the marketing budget. Because of the strategic role of marketing, additional datasets – such as markets, competitors, and the economy as a whole – offer additional value when assessing conversion rates and loss cost ratios, and so these could also play a role together with customer data and macro-economic variables in every other machine-learning application.

1.5.1. Claims Data

Insurance claims databases, which are an important subset of corporate data, have major advantages over alternative sources of data used in empirical work in the insurance sector. Most importantly, they contain detailed information about actual choices made by insurance consumers, which may differ from reported preferences. For example, customers may purchase a vehicle insured by a different company than the one covering the rest of their household. If they seek coverage for the first time, they may switch to another company for unitary or annual costs starting five, ten, or twenty percent lower than those requested the previous year by their current insurers.

In addition, claims data contain unit prices of risk defined for each claim made during the period considered. In the case of property insurance, the losses paid are adjusted by time, for rehospitalization and colpine factors, to represent their expected present value when the claim is closed. In the case of health insurance, which generally does not have the concept or value of a closed claim but the costs paid for the health events covered by the policies, most policymakers and agencies do not use the time of cost realization, only the probability that the insured event occurs and are not time-dependent when calculating the total estimated risks. These factors imply that costs are only time-dependent, a presumption of a one-dimensional aggregate loss-driven distribution method.

1.5.2. Customer Data

Normally, customer data refers to personal information, such as the name, age, gender, education level, occupation, and income of insurance policyholders. These data can be obtained at the policy inception or renewal. Customer data also include some "dynamic" properties, which change over time, such as the insured amount and premium level, coverage options, type of policy, and years of policyholder's age. These data are recorded regularly during the life of policies and provide crucial information for distinguishing profitable from unprofitable policies. Customer data in life insurance are, for example,

the lapse period, the length of time the policy has been in force, and the historical claims of the insured person.

Other customer data that can be used in the life insurance problem include the number of dependents covered by the policyholder, any severe history of diseases in the policyholder's family, and the health checkup results. These attributes are believed to have crucial relevance to the possibility that the insured dies or the possibility of any future severe diseases occurring, and can also be utilized to evaluate the policyholder's health. In general, there are three types of policyholders' behaviors: first, a policyholder renews the policy and dies after some time; second, the policy is in force until the insured policyholder's death; and third, the policyholder lapses the policy before his or her death. If there are dependents for the policyholder to take care of after the policyholder's passing, one would expect that the policyholder would not have any lapse behavior, which leads to a long-lasting insurance contract.

There are major differences between the data for general insurance and life insurance. The data for general insurance mainly include the characteristic attributes of the insured. By contrast, the data for life insurance comprise information on both the policyholder and the insured. Here are five notes about customer data in insurance. First, customer data are heterogeneous data and difficult to obtain. Second, customer data evolve quickly and frequently change over time. Third, customer data are costly to obtain or create. Fourth, customer data have varying accuracy, quality, and currency. Fifth, customer data have different levels of usefulness and have wide coverage.

1.5.3. Market Data

In practice, market data is mostly used when a bias correction model is considered in prediction taste. Data of market outcome is usually the transaction price of houses or commercial estate at the time of transaction. For the residential property market, the most publicly accessible data is sales data based on the registry of real estate transactions. The second source is the Register of Property Prices and Transaction Valuations owned by a government agency and has information about every transaction of property priced by independent property valuers. Access is limited. More detailed data including the market and transaction price of each property is very often collected by real estate agencies.

When considering the commercial estate market, the certified registries of real estate funds owned by the state or property register of treasury include commercial estate market price data. However, the data considers only big commercial estate transactions primarily with public financing. The second source is real estate transaction financial databases primarily used for investment purposes.

Usually, the market price value is not evaluated conditionally but is defined based on previous research or data or by analogy of previous transactions of similar (or not so similar) property. In practice, the temporal aspect of market data is crucial. For development period predictions, property prices are predicted based on market transactions conducted no longer than 6 to 12 months before. However, the problem is that for remediation predicting, property predictive market price is not based on prospective property options only but also on retrospective property transactions conducted with similar properties a long time ago. The explanation for such a notion is that it is difficult to sell property of similar importance. Thus, the property can be sold for a similar price after the remediation is finished or otherwise.

1.6. Regulatory Considerations

In this section, we'll discuss how AI applications in insurance need to balance innovation and monetization alongside ensuring compliance with current regulations. Until relatively recently, AI use was unregulated. However, as insurance companies have increasingly relied on AI applications in higher stakes areas, there has been increased scrutiny from regulatory agencies and growing sensitivity from consumers. On the one hand, AI use in insurance often involves using sensitive data from consumers, ranging from health disclosures to driving habits. On the other hand, AI models are notoriously hard to interpret. These two facts together create a scenario where the AI model can be discriminative or otherwise not serve the best interest of the consumer while in some cases not allowing for easy identification of the failure mode. As such, relying on AI in insurance requires careful consideration of regulatory requirements for compliance. In this section, we validate and summarize the key considerations any stakeholder should keep in mind when designing, deploying, and using an AI model, both from a regulatory and industry perspective. We'll first focus on the key regulations that today impact AI use in insurance, and expand into ethical considerations around using AI and deep-learning methods.

Regulatory oversight over privacy and regulatory rules dictating equitable treatment of consumers concerning adverse actions taken based on AI outputs are currently the two most critical areas specific to AI use in insurance. Other areas of regulatory and industry-required compliance will require additional considerations for insurance stakeholders but are more generally applicable across AI use today. The impact of being subjected to these broader compliance mandates will often dictate the solution stack richness and operational readiness available to any insurance stakeholder.

1.6.1. Data Privacy Regulations

machine learning (ML) revolution in Insurance will provide economic, structural, and operational efficiency, but this efficiency and technology promise comes with a set of legal and regulatory hurdles and considerations. Any AI and ML modeling effort in Insurance will need to go through legal and regulatory scrutiny before its models become production-ready. Multiple questions need to be answered before doing so: What are the existing laws and regulations that govern data use for AI/ML in Insurance? Are applicable laws specific to model outputs or model inputs? What consideration must be given to models that rely on scraped data from social media or similar sources? What are the prospective laws that will have repercussions on AI/ML in Insurance? This section identifies and discusses data privacy regulations that will determine what data can be used for model input, the model output, and the effect on the target variable. The section also identifies some of the ethical concerns and considerations to be kept in mind when developing an ML product to mitigate unintended harm to society at large. Other regulatory issues are discussed, such as the regulation of technology that determines the decision rules to accept or reject an insurance policy, and the disparate impact of technology, defined as the adverse effect on historically disadvantaged racial or ethnic minority groups.

Algorithms, like other human inventions and scientific research, are to be used for the good of humanity and in the best interest of society. Any argument justifying model complexity and an ethically ambiguous post-hoc explanation is counterproductive. We should, therefore, start with simple models that create explainable rules, akin to regression or rule trees, that determine the underwriting, claims, or pricing decisions of insurance. Ethics requires that the factors determining the decision help or at the very least do not harm the provider or policyholder.

1.6.2. Ethical AI Use

Not only specific regulations governing the use of AI are still evolving, but it is also worth noting that no regulation can cover every possible aspect of ethical AI use. Therefore, the principles of ethical AI use are broader than compliance, and compliance is only one component of an overall ethical approach. An ethical approach is in place, especially when companies actively take care of guidelines and leading practices: Human-AI interaction should be transparent and understandable; intelligence systems should be robust, secure, and reliable to minimize environmental and human harm; AI systems should be trained and tested in a way that avoids inequitable effects on individuals or groups; Artificial Intelligence should operate within the legal framework; technical robustness and safety play a crucial role in an ethical approach to AI; and a business should operate AI for common good.

Depending on the company, the quality of its data and systems, and related ethical considerations, AI may even require further measures that go beyond the recommendations. For instance, in the insurance industry, financial motives and data access get even more intertwined than in many other applications of AI. At the same time, the financial implications of extracting information from customer data may become substantial, especially in life insurance due to the low margins in this market. Seems to be, that there are no fundamental conflicting forces at work, considering purely ethical AI use either. This may be different for non-life insurance since isolated applications may indicate different implications of ethical AI.

1.7. Risk Assessment and Pricing Models

Foundational in the insurance business, underwriting procedures govern risk acceptance, risk retention, and risk mitigation by effectively predicting the probability of a loss. Pricing refers to the process of finding the best insurance premium charge or value. The guiding principles for pricing are that the prices must be fair and adequate for the loss exposures, that is, the estimated costs of covering the respective claims must be adequate, and the price should be affordable to potential customers. The art and science in risk assessment and pricing suggest that regulatory guidelines must be observed, as an important cost in Canada is the administrative expense of the current compliance of each province with government rate making.

Predictive analytics involves all those models and artificial intelligence tools used to build predictive, informative, or diagnostic answers in any domain or field based on current and historical available information. Predictive analytics in insurance pricing models is the process that creates predictive answers to the technical question “What is going to happen in the future?” Predictive modeling can be defined as the process that yields information or answers to the question “What is the probability of an event happening?” The event is related to the transaction or interaction of persons with the insurance company, such as filing claims from the purchased policy, customers canceling the policy, finding more appropriate coverages, and clients considered likely to consume certain services.

1.7.1. Predictive Analytics

As the insurance sector finds itself under increasing pressure to optimize its risk selection and pricing for underwriting, important questions arise, such as how exactly, and to what extent, can companies optimize their operations thanks to innovative technological solutions. It is clear that doing so is not a simple task; many active insurance companies appear to be committed to the task, which is not without an associated pressure on

earnings and solvency. As alluded to in earlier chapters, traditional prudential assessment models are only relevant in the operational context if complemented with models that can go beyond traditional mixed approaches, which disregard much of the available information and its dynamic structure inside the formula. Within this context, it is important to differentiate between Pricing and Risk Assessment, as the models associated with each of the tasks are different considering their operational objective, structure, and the business context to which they are intended to be applied. Historically, the traditional approach to Computing Insurance Pricing consisted of latent models based on compounded claims development loss models again conditioned only to a set of exogenous risk pricing factors. Those traditional approaches have been challenged by the rapidly evolving data analytics capabilities and more particularly the steadily increasing convenience of big data and predictive analytics.

Predictive analytics, defined here as data-driven predictive modeling, based on Machine Learning and Data Mining, applied to data from a wide number of dynamic tempering, static, and exogenous factors, is playing an increasing role in the risk assessment function. The shift towards predictive models has particularly affected the traditional task of birthdate underwriting. A key characteristic of the development of predictive analytics in the insurance sector both overall and especially about the traditional task of prospecting potential clients, is its facility of comparatively easy implementation in a given category than its specialized actuarial counterparts. Simple predictive models are already offering attractive input for optimum pricing from the perspective of the insurance company. More complex predictive modeling, supervised in supervised risk pricing tasks, is already being implemented to optimize the effort of the loss assessment consultant or damage adjuster.

1.7.2. Underwriting Automation

The underwriting process deals with risk selection and classification. It's an essential part of an insurer's business strategy. However, the traditional pathway is often cumbersome for both parties. Flexibility constraints may be troublesome for customers while still exposing the insurance company to losses associated with price inconsistencies. Automated underwriting uses predictive analytics to highlight the most significant risk factors to improve decision-making. Machines can take over the process to enhance customer experience as long as they can solve the problem of adverse selection and customers are straightforward. It leads to more accurate pricing of risks, a better customer experience, and a significant reduction of costs.

Artificial intelligence systems powered by machine learning capabilities are already performing predictive analytics across many industries. However, the task of automating complex human decision-making in areas like underwriting, claims, and fraud detection,

or detecting outliers in customer transactions across all core insurance functions is much more difficult. Predictive behavior models create new categorical segmentation instead of predicting past behavior but these models typically fail to deliver the required level of precision. New AI models are much better at predictive segmentation but despite their ability to make better predictions than traditional models, even these machine learning methods fail to deliver perfect predictions required for operating a highly profitable automated underwriting program. Additionally, underwriting considers many different kinds of risk factors.

1.8. Fraud Detection Techniques

Fraud detection in networks is a critical and challenging task that serves both economic and regulatory purposes. The large financial impact of fraud incidents makes the prevention and detection of these events a high priority for insurance organizations. Fraud detection is an evolved application of anomaly detection and behavioral analysis, both of which serve to promote security and safety in a variety of contexts. While different in objective and approach, both endeavors explore the underlying structure of observed data or behavior to detect outliers, the latter of which usually corresponds to a malicious agent such as a cyber-criminal, hacker, or scam artist. Anomaly detection seeks to uncover patterns that deviate from the norm to identify potential fraud, which then leads to other supports like predictive modeling. Predictive modeling creates a model based on training data that are already hijacked or drawn from compromised information and then tested on newly generated data. It requires that the historical data are sufficiently large to capture a sufficient number of hijacked events.

Behavioral analysis studies the users' activity or patterns over time to create profiles of normal user behavior. User profiles can then be used as signatures to identify fraud. This approach relies on the performance of pattern recognition. It assumes that compromised users generate different behavioral signatures from normal ones so that statistical methods or other artificial techniques can detect the difference in behavior. Fraud detection using state-of-the-art anomaly detection and behavioral analysis techniques usually requires a separate pre-clustering step to reduce the search space and increase the accuracy of detected fraud. This is because, for most malicious advanced persistent threats, the alerts returned by the detection sensors comprise a small fraction of the total number of alerts or represent only a subset of different users. While clusters are formed based on behavioral similarity, not all users in the identified clusters are involved in fraud.

1.8.1. Anomaly Detection

Anomaly detection is the identification of rare items, events, or observations that raise suspicions by differing significantly from the majority of the data. Anomaly detection is both a fascinating and a challenging research problem. A number of applications involving anomaly detection have emerged due to the rapid development of the Internet, and it is arguably one of the most heavily researched problems in statistical learning today.

Fig 1 . 3 : Artificial Intelligence (AI) In Insurance Market Size

Detecting fraudulent insurance claims is a difficult problem because fraud is a rare event when compared with the total number of claims. Estimating the probability of fraud on a claim requires a model that can completely characterize the normal health of the patient, considering the clinical aspect of the claim, and identifying those patients who are unique or different from the general population. Semi-supervised learning can utilize a set of previously classified claims to learn representational features in a higher dimensional input feature space that captures the characteristics of the normal dynamic activity in a patient. Plus, autoencoders have been successfully employed in this way for many domains, such as surveillance video, by utilizing reconstruction errors in the features learned by the encoder.

Additionally, adversarial training and constrained neural networks can parameterize the features learned by the autoencoders to increase their specificity to the normal activity, while adversarial discriminative domain adaptation is advantageous in the insurance domain as it utilizes identification information from the labeled fraud claims to improve the accuracy of the classification. Anomaly detection is an essential task in contemporary computer vision research, as supervised training sets for diverse general tasks in the real world can be prohibitively expensive and labor-intensive.

1.8.2. Behavioral Analysis

Behavioral analysis refers to the observation and interpretation of how a person or group behaves within a specific setting. These behavioral characteristics can reveal information such as whether the individual is of high or low risk or whether their behavior appears fraudulent. In simple terms, the behavioral analysis of a person is no different than the diagnostic phase of a physician assessing a patient. Fortunately, unlike a physician, with machines, we do not need to take regular notes. We can track the behavior of an individual for thousands of different processes. This is especially true for our digital lives, in which we create our timelines without really trying. Yet privacy concerns about digital footprints are often raised. The important point, though, is that behavioral analytics can inform us whether a person is of high or low risk, appears to be cheating the system, or is engaging in activities that warrant further investigation.

As with other techniques discussed in earlier chapters, behavioral analytics and anomaly detection do not compete for one another's attention. Much like not all fraud techniques can be assessed only using behavioral means, not all fraud situations go undetected from other detection techniques. Instead, behavioral analytics sits next to these techniques, helping to improve the situation. It enhances the ability of other detection systems to detect side cases for which they often issue false positives. With services that analyze historical datasets and tie the activities together, it is sometimes possible to retroactively identify the accounts that operated more erratically than others and, by doing so, have more fraud.

1.9. Conclusion

Pioneers of Deep Learning noted that neural networks 'have been applied to all the hardest pattern recognition problems and have solved them all'. This is no different in the implemented solutions of AI in Insurance - from automated data entry to improving value predictive models hidden within complex datasets. However, with the production version systems or supply chain embedding these solutions in Insurance, it is a unique and quite different story. As with our previous comparison of other technologies, AI is

much harder to implement in a production system, particularly with the volume of real-world cloud-based data pipelines feeding Insurance's business ecosystem, and the complexity involved in core processes and embedding these into systems. This is not a new story as observed on innovation in embedded production applications years ago - and IT service firms providing advanced analytics or AI embedded solutions, are but a new flavor of consultancy dollar being burnt.

Systemically, Insurance cannot afford to embed work processes based on software innovation that centuries of human resources have cost (though, it was worth hundreds of years of families and life's work). We need to realize, that AI and Deep Learning are but a contemporary elixir, as our Economics of Being Small Society discusses - selection or transformation of existing core primitive functions. Realizing so allows AI to move back to its origins - AI as a "bare technology" of primitive "core functions or procedures" that can be "wired together" to create practical problem sensible full applications or products like others. AI should be seen as organic modular building blocks that can evolve into different shapes and forms - like a Meccano system of kits. Implemented solutions within Insurance, should reflect that.

1.9.1. Final Thoughts on the Integration of AI in the Insurance Landscape

Artificial Intelligence has emerged as a pivotal force in the insurance sector, wielding the capacity to drive efficiency, enhance productivity, and augment overall performance across the entire insurance value chain - from underwriting, risk prediction, and pricing sophistication, fraud detection and mitigation, claims processing, customer service delivery, agent engagement, and distribution network management. By enabling near-perfect underwriting, pricing accuracy, and customer service, along with faster and better claims processing, the integration of AI in insurance operations has far-reaching implications on performance metrics such as reduced loss ratios for insurers, thus improving bottom lines; enhanced customer satisfaction, which is measurable by metrics such as customer satisfaction scores; and enhanced customer retention ratios. Additionally, the efficiency dividends have positive implications for shareholder value, thus making the case for compelling strategic reasons for investing time, effort, and partnership resources in designing and executing AI strategies and initiatives.

Despite the considerable promise of AI, there are significant implementation challenges, especially for larger, more complex multi-line insurers. Success requires seamless buy-in from each business function or Line of Business, in addition to close implementation collaboration with technology partners, to achieve the desired implementation outcomes. This, combined with the need for capital allocation, makes AI a tougher business bet, especially for regional players with constrained financial and organizational resources. That said, the rewards of successful implementation make the case for embarking on a

journey for adopting AI-powered solutions, because the advantages of being exposed to the latest in AI technologies and solutions are bound to have major omnichannel borrowings to reap the benefit of the leapfrogging potential, to experience a virtuous funding cycle and hence the motivation to undertake the journey.

References

Ng, A. Y. (2016). Artificial Intelligence and Deep Learning: The Future of Insurance. Communications of the ACM, 59(7), 56–65. https://doi.org/10.1145/2904272

Weng, W., Lin, X., & Zhang, H. (2020). Deep Learning for Insurance Analytics: Risk Assessment and Customer Profiling. Expert Systems with Applications, 158, 113438. https://doi.org/10.1016/j.eswa.2020.113438

Panigrahi, S., & Borah, S. (2021). AI-Powered Claims Processing in Insurance: A Deep Learning Framework. Journal of Big Data, 8(1), 44. https://doi.org/10.1186/s40537-021-00433-2

Chen, M., Ma, Y., Li, Y., Wu, D., Zhang, Y., & Youn, C. H. (2017). Wearable 2.0: Enabling Human-Cloud Integration in Next Generation Healthcare Systems. IEEE Communications Magazine, 55(1), 54–61. https://doi.org/10.1109/MCOM.2017.1600363CM

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778