Deep Science Publishing
https://doi.org/10.70593/978-93-49910-74-4_2

Chapter 2: Intelligent underwriting: applying machine learning to risk assessment

2.1. Introduction to Intelligent Underwriting

Intelligent underwriting (IU) is a new, specialized area of work in the smart insurance space that seeks to harness the latest advancements in artificial intelligence (AI) and machine learning (ML) to deliver near-term rewards for the global insurance community and consumers. Policies can be better priced and discounted, thereby relinquishing the untimely and undue transactions between creators of risk and insurers. The insurance business will apply these advances to deliver higher corporate net profits and shareholder valuations. Service empowers insurers to positively interact more quickly, adequately, and accurately with consumers and policy owners throughout the insurance process life cycle, from the selection of the product to coverage settlement. No technology disruption will rescue laggard insurers, but for those insurers that proactively embrace intelligent underwriting, the advantages it brings will raise their business fortunes considerably (Ribeiro et al., 2016; Wüthrich, 2020; Henckaerts et al., 2022).

For consumer-facing insurers, improved consumer experiences, driven by information technology for acquiring and servicing insurance needs, are enhancing loyalty towards the insurance brands. Back-office operations driven by core technology systems, however, have been slow in transforming. Processes in many operational and support areas are particularly in need of radical overhaul. These areas work best with speedy touchless processing of clean business that is the low frequency by definition. Intelligent underwriting answers directly to this history of disruption in back office operational and support areas. For reasons that we will cover, these core processes - operations, risk assessment, pricing, actuarial, grievance settlement, and so on - have missed the AI-ML train in recent years when a majority of the world's focus has been on digital enablers of front office customer-facing business (Yoon et al., 2018; Wüthrich & Merz, 2021).

2.1.1. Overview of Intelligent Underwriting Concepts

Over the past few decades, the impact of technology has changed how insurance acts and how people, as appointed risk-takers, are affected by risk assessment. In the modern era, actuaries and underwriters both carry the task of solving risk selection problems. Actuaries do so at the macro level and are responsible for setting cost-efficient premiums that reflect the underlying risk distribution. In contrast, underwriters make final decisions on new and existing business at the micro-level, within the context of one specific policy. It is generally accepted that evaluating an applicant's risk at an insurance policy level leads to better decisions than evaluating risk at an aggregate level to set a general price. However, compared to the large influence of modern predictive analytics on actuarial science, no significant innovation has been made on the core problems of underwriting over the last five decades. This paper aims to enhance this void by showing how the advancements made in predictive analytics and machine learning can transform the underwriting procedure towards "intelligent underwriting".

Around the world, many successful implementations of the predictive modeling pipeline at an actuarial level can be found. The actuary states the decision problem at hand, obtains data, explores the data, prepares the data, develops a predictive model, validates the model, documents the findings, deploys the solution, and finally, maintains the implementation. However, what most remain unaware of is that this pipeline and its components may be enhanced by an automation-based approach. The principal challenge of developing predictive models remains data quality — an automatic approach through data mining and cleaning could spend the most time and money but is not deemed exciting. In situations where large and continuous amounts of structured datasets are available, data preparation could be done via automatic cleaning tools. In this work, we propose a more intelligent take on these guidelines. Our version of intelligent underwriting relies on three fundamental principles: machine learning, usable automation, and real-time updating of predictive models. We believe that the combination of these principles will allow actuaries and underwriters to embrace automation without losing their control over the underwriting procedure.

2.2. The Role of Machine Learning in Risk Assessment

Machine learning is a branch of artificial intelligence that can be used to automatically extract knowledge from data and convert it into useful information. In finance, the successful application of machine learning spans customer loan underwriting and screening, risk modeling, credit risk rating, and corporate debt rating, among others.

Risk evaluation can be thought of as an exploration procedure to analyze the risk character of clients. At the preliminary stage, it is an intuitive review that mainly relies

on domain expertise and experience, which is inherently subjective, talking to clients, and checking for any clear signals, like credit records and company financial statements. In the middle stage, we typically use statistical techniques to analyze available data quantitatively, which serves as a prerequisite for tailored modeling and formal scoring procedures that combine the result with professional expertise and experience. This is usually done under supervision. In the latest stage, we adopt decision models based on qualitative and quantitative data, with AI/machine learning techniques as major weapons. The models are typically applied to high-volume and relatively lower-risk underwriting decisions. At last, we conduct remedial measures and code deciphering.

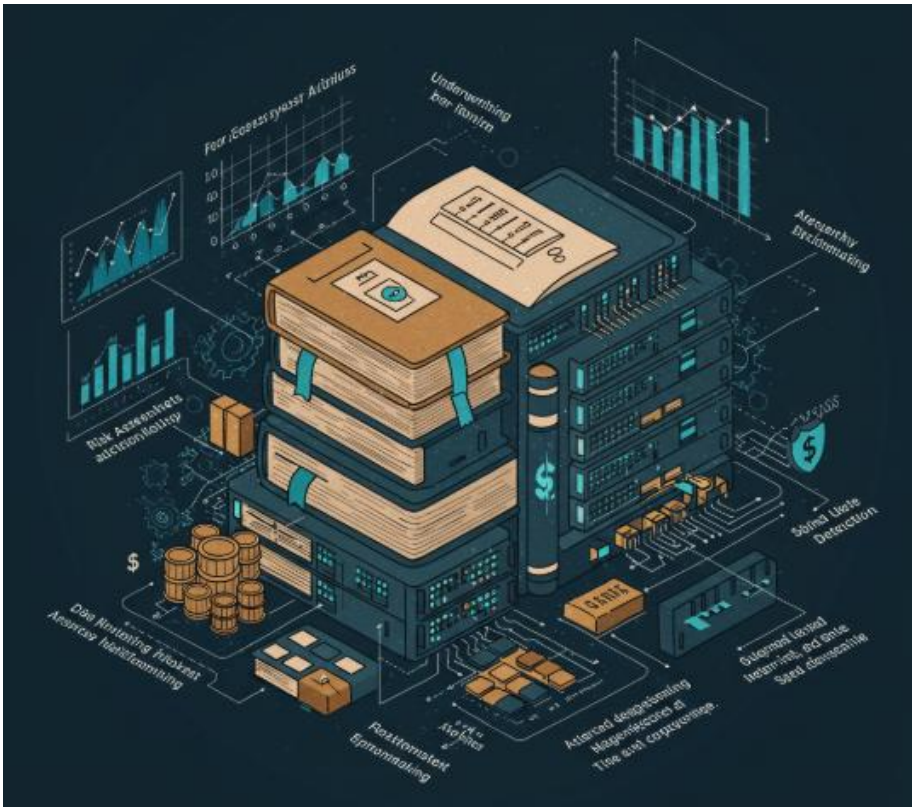

Fig 2 . 1 : Underwriting & Machine Learning

Data-driven intelligence is redefining every significant aspect of risk evaluation, including risk identification, risk prediction, risk quantification, customer interaction/nurturing, and risk change monitoring. New risk signals and prediction variables/quantities are belatedly identified and used. Hybrid decision mechanisms and monitoring frameworks have been implemented. Data-empowered clients are influencing original risk management and shaping new relationships. Large amounts of unstructured data produced by external data providers, and internal nonstandardized and semi-structured sources like merchant transaction records are being used. Traditional

risk evaluation models have been enhanced or enlarged by more informative variables and altered by different models/learning instruments incorporating machine learning techniques. These new models/techniques have been employed to protect against higher risks and for more coverage. A marketplace for unused risk modeling expertise is being created, enhancing traditional intelligence-powered risk assessment with data-driven machine learning tools.

2.2.1. Exploring the Impact of Machine Learning on Risk Evaluation

The emergence of machine learning (ML) has ushered in a new era for risk assessment in a variety of fields, including medical care and marketing because it improves predictive power. For actuaries, the predictive aspects of ML's power are particularly appealing. However, ML originated in computer science, where the algorithms and methods differ considerably from those used in traditional modeling techniques such as generalized linear models. As a result, ML has posed questions about whether the existing approaches used by actuaries are adequate for risk evaluation going forward. More urgently, actuaries must investigate this new instrument to comprehend its significance while carrying out the discipline's fundamental obligations. We provide an overview of the ML revolution and discuss its potential additionally, we present a research agenda that focuses on the fundamental issues posed by ML in the context of risk assessment, and risk prediction.

According to various accounts from several fields of study, machine learning (ML) has sparked a sea change in predictive modeling methods available to researchers and practitioners. Because the earliest implementations were even made available in a commercially available modeling package, it is frequently claimed that no new algorithmic breakthroughs propelled this transition. Certain items were incorporated into more widely known machine learning accounts as "algorithms." However, rather than technical characteristics, the innovation was the extensive practical application of certain procedures, frequently with substantial amounts of carefully curated data and expert tuning of the methods. The peer review procedures that support traditional academic publications have only gradually started to accommodate this new innovator, which has incited debate and concern over potential risks associated with these evolving techniques.

2.3. Data Sources for Underwriting

Machine learning models can be built and applied only if there is data available to describe how the world works. An attractive feature of machine learning models is that they perform well even on data that is not representative of the population that the model

will be applied, a drawback of many of the traditional models and approaches used in our industry. While this allows companies using machine learning to apply their models in countries or regions that are not their own, differences in available data will still typically limit such transfers to broad strokes, leaving the situation in the destination country to be handled in a manner mainly driven by local preferences, market conditions, and cultures, or requiring the model to be retrained to also be effective in the new country or region. In the context of applying machine learning to underwriting, we can follow the same approach and perform the risk assessment of either a proposal or a portfolio based on whatever data sources we have available to us while imposing limitations or applying adjustments when necessary for lack of complete information.

Machine learning models can absorb large quantities of data from a variety of sources in a mostly automated fashion. In this section, we will describe the main types of data, offering observations and considerations to help practitioners and companies build the data environment that will lead to successful implementations of machine learning methods to assess insurance risk. Data sources are typically classified into three categories: structured data, unstructured data, and external sources of data. With structured data, the data is presented consistently or organized into a tabular form, similar to the data that traditional statistical models operate on. Unstructured data, as the name suggests, is information that has not been structured, stored, or presented in a way that facilitates its analysis. External sources of data refer to data from companies or organizations other than those connected to the insured or the proposal.

2.3.1. Structured Data

Risk assessment traditionally relied heavily on structured data: well-defined, numeric information such as an applicant's credit score or prior premium payment history. For a typical insurance underwriting decision, the following structured data may be used:

• Risk feature data. This includes information specific to the risk(s) being insured, such as loss history, claimant characteristics, value, coverage limits, and deductible levels.

• Applicant data. This includes information about the person(s), entity, or group applying for coverage, such as age, sex, marital status, credit scoring, occupation, education, business strategy, financials, or insurance history.

• Event feature data. This includes information about the event or peril being insured against, such as location, timeframe, and limits or exclusions regarding the peril.

• Market data. This includes information about the insurance market in which the transaction is taking place, such as market capacity, competition, pricing, and regulatory issues.

• Other information. This includes information about aspects peripheral to the insurance transaction that may be relevant, such as environmental risks, political risks, industry or national economic factors, and ideas of public policy.

In many cases, structured data required for underwriting risk assessment is available from the data management capabilities of the insurer, agent, or third-party supplier. Such information may be stored in standard commercial software databases for insurers or even in data warehouses; however, such data may be difficult to access and impossible to share with other stakeholders without major cost or time investments. It is commonly far easier to use the data supplied in real-time during the underwriting process, even when not comprehensive.

2.3.2. Unstructured Data

Unstructured data can be defined as information that lacks a pre-defined format. In this definition, we consider unstructured not only free text or videos but also images and categorical variables. Categorical variables generate strings of characters that can be squeezed into hyper-cube representations of fixed dimensionality and contain a low amount of information: they are the most primitive representations of semantic classifiers. These variables exist for all the underwriting fields for which choices are made, some choices are binary, and others are categorical with a finite number of choices. An example of a variable with a finite number of choices is the type of commercial property: convenience store, office complex, warehouse, etc. It can be expected that any data-driven model will benefit from the step of changing the type of categorical variable into the associated dummy variables. Free text is more complicated, because it deals with large vocabulary files, resulting for instance from the use of an insurance underwriting system. Such a system takes in all of the passive and active steps associated with assessing insurance underwriting information: it generates the strings of characters that define underwriting queries and defines a flow of the queries over time, guiding the underwriter to conclusions. Images include photographs, which typically illustrate the physical location of an insured building and business-related pictures that explain key aspects of the insured commercial activity, such as the key ingredients used, the product type, and its internal and external packaging and presentation. Finally, videos record in real-time people performing acts that relate to the risk associated with the insured building. Videos of businesses in operation allow the most detailed and accessible information to be collected for real-time evaluations.

2.3.3. External Data Sources

Data Sources for Underwriting

Underwriting has traditionally relied on internal company knowledge and trust in the submissions for information. However much external information is gradually becoming available and augmenting this internal data; it is important to consider how this external information can help underwriting. Data for underwriting comes from two main types: structured and unstructured data.

Structured data is, for example, data warehoused in a database, where the relation between different data points is known. An example is information on a risk use—such as business details, prior loss experience, reinsurance program, and coverage limits, with details encoded in different lines of business codes. This data lends itself to traditional statistical analysis and predictive modeling techniques. While the dimensional details vary across different lines of businesses, the general relations between insurance company data pools for the different lines of businesses are generally understood. Modeling is done separately for each line of business.

External Data Sources

External data is complementary to internal data, especially when the company has little experience in underwriting risks of a particular type. Apart from the data required to grow the company's intellectual property pool, it is important to look at when and how to use external information in risk assessment based on data availability. Data can vary in many dimensions: accessibility, timeliness, reliability, frequency, update cycle, granularity, and cost.

External weather data is available from companies on which tracking featured data is central to their business models. Companies specialize in more physical aspects of data. Timely, accurate weather data access for the underwriting decision is of the highest importance, given that this information can rapidly change on short time scales.

2.4. Machine Learning Techniques

Machine learning is a subfield of artificial intelligence that uses statistical techniques to automatically learn with data without being explicitly programmed. Machine learning can be considered a generalization of rule-based systems, where instead of encoding the rules, the system automatically learns the rules by looking at the data. Machine Learning can be broken down into three main types: supervised learning, unsupervised learning and reinforcement learning.

Supervised Learning

Supervised learning is the kind of machine learning where labels are available for the training data. The goal is to learn a model that accurately transforms the inputs into the desired outputs. For example, if the decision is whether to insure a certain client or not, the training data would have historical information on various clients and whether or not any claims were made by them while they were insured. The decision would be based on a set of features taken from this information and the model would learn to map from the features of the insured client to the decision of whether to insure or not.

The model learned by the machine learning algorithm would then be used on data for new client applications to make a decision. This is known as a classification problem, since the outcome is a discrete variable (the decision is either yes or no). Many different machine-learning algorithms can be used to achieve this classification. The trained model can also predict the likelihood associated with the decision. This would be calculating the probability distribution over the output classes instead of predicting the most likely outcome.

Unsupervised Learning

In unsupervised learning, labels for the outputs are not available for the training data. The goal is to learn on the input training data only. For example, if we have lots of data where we do not know whether a client made a claim or not and we do not know what features differentiate such clients, we can use unsupervised techniques to cluster the clients into groups.

2.4.1. Supervised Learning

Supervised learning uses a labeled dataset, where the algorithm learns to predict an outcome by mapping the relationship between the predictors (inputs) and the outcome (target) variable, which means that exact correct outputs are available for some training samples to supervise the learning algorithm. More specifically, let X be the d-dimensional input space from where the predictors are drawn, $X \in Rd$, Y be the output space based on which predictions are made, $Y \in R$, and N be the number of input/output pairs available for training. In the case of supervised learning, a training set consisting of N pairs of input/output values is made available to the algorithm: {(X1, Y1), (X2, Y2), ..., (XN, YN)}, $Y = f(X) + \varepsilon$ is called the underlying relationship that is either deterministic, where $\varepsilon = 0$, or probabilistic, where ε is drawn from a stochastic process. The learning algorithm's goal is to output a predictive function $f\hat{}$ that can approximate f based on the training sample. This predictive function $f\hat{}$ can be either a regression function for continuous outcomes or a classification function for categorical outcomes.

Supervised learning is arguably the most well-known paradigm of machine learning due to its broad applicability to many practical problems in various fields of human endeavor, including risk assessment in underwriting. In this section, we discuss how supervised learning can fit a specific application of risk assessment, namely the construction of prediction models for estimating the incidence of losses based on data from previously underwritten policies. The emphasis is placed on why supervised learning is such a powerful way to apply machine learning without going into the theoretical details of model fitting, validation, and selection.

2.4.2. Unsupervised Learning

Despite the importance of researchers, in assessing the risk of insurance claims, and looking for structured segmented patterns in the portfolio database, unsupervised machine learning techniques have received little attention in the insurance literature. Unsupervised learning has the potential to group policyholders or claims into well-defined risk segments. The information regarding policyholder behavior or risk attracted many researchers in the last two years, given the availability of a deep pool of data. In general, unsupervised learning techniques can be classified into two categories: relational clustering and pattern analysis. Although clustering creates data segments of similar observations to optimize decision-making, segmentation by clustering does not optimize decision-making. Pattern analysis uses non-clustering methods to discover simple, easily understandable models. The second option facilitates decision-making by segmentation and can be more efficient. Shapley value decomposition and unsupervised decision trees are examples of pattern analysis methods.

Relational clustering concentrates either on similarity measures or combination measures. The first method is based on the distance between observations, and the second method creates the clusters to minimize a combination cost at once. These methods can be tree-based or graphical clustering methods. Hierarchical and k-means clustering are tree-model structures. While hierarchical clustering is computationally inefficient for hundreds of observations, k-means clustering requires the prior definition of the number of segments. Graphical methods use decision tree models or graphical distribution models. Graphical segmentation methods have several drawbacks, for instance, they cannot provide information about clusters or segments beyond the predictions.

2.4.3. Reinforcement Learning

Reinforcement Learning enables machines to learn through experience and feedback, much like humans do. While the concept may be familiar from fields such as behavioral

psychology and behavioral economics where benefits motivate people to conduct beneficial and rational behavior, reinforcement learning implements such ideas for machine learning on a non-verbal level. A crucial difference to supervised learning is the way we provide feedback to the algorithm. In supervised learning, we explicitly tell our algorithm how to correctly map input to result by providing both input and classified result. In reinforcement learning, we only provide feedback concerning the result, not the way of mapping. We want our machine to systematically explore different opportunities to find the most efficient mapping.

While reinforcement learning is an easy way to realize such feedback loops – and indeed equally intuitive and human-like in its basic principle as deep neural networks for feature recognition in images – implementation is non-trivial. The machine needs to learn to compare and quantify risks because choices might pay out only after a significant amount of time. When training the machine, in a lot of cases, it is not known what impact a particular action has on long-term outcomes. We learn, as humans, that smoking influences health negatively but we do not know the statistics of that influence. Therefore the reinforcement learning algorithm needs to learn both the probabilities and the consequences, a double task which makes it computationally harder. For certain cases, the influence on long-term estimates is so small that humans can incur the risk of waiting until certain proof exists.

2.5. Feature Engineering in Risk Models

Feature engineering is the process of creating meaningful attributes from transactional, behavioral, or demographic information. The attributes created in this way are known as features, and together they help the model to become better at distinguishing between good and bad outcomes. Poorly engineered features can limit the potential of the model to give good predictions. Examples of features to predict whether a non-listed company will default include the last-year assets of the company relative to the number of months the company has been in operation, and the growth of the company, measured as the ratio between last-year sales and the last-three-year average sales. Expansion-contraction cycles are critical to understanding company behavior. For a limited liability company, allowed losses are part of society's potential losses and need to be monitored very carefully.

In the context of banking, relevant features include any demographic information available on the borrower; any transactional, historical information; policies regarding the collection of overdue debts; recovery information; complaints; and any information that can help identify milestones in the credit cycle of the borrower. Thus, features are critical for building models that make better predictions concerning the default event and where policy changes perceived under a behavioral perspective constitute possible

sources of changes in the predicted probability of occurrence. In the context of insurance and risk protection, financial characteristics, products purchased and the history of claims and complaints are particularly relevant.

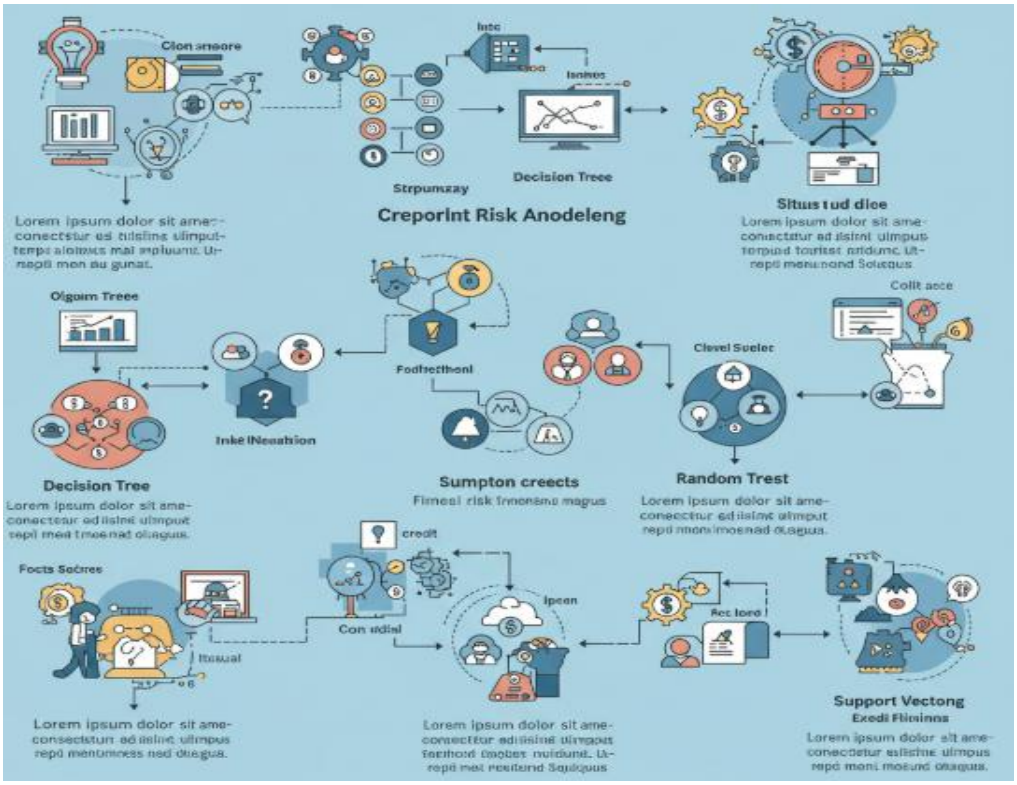

Fig 2 . 2 : Credit Risk Models with Machine Learning

2.5.1. Essential Elements of Feature Engineering in Risk Assessment

Feature engineering is a vital process in any supervised predictive model but it becomes particularly important in critical areas like risk assessment where quality is paramount. Without clever feature engineering using all relevant domain knowledge, no amount of advanced mathematical algorithm wizardry will guarantee good performance and may deliver poor results in risk assessment. In fact, for risk assessment problems there are thus far no default easy off-the-shelf solutions with applied ML methods that produce reliable predictive performance results or express shortcuts algorithmically taking care of the feature engineering for us. Rather feature engineering provides us with some shortcuts to “be smarter” by asking the right questions about the data, like searching through dubious transactions and asking which engine parts or specifications or combinations of features allow us to best forecast risk. This reduces the complexity of the learning problem at hand without guaranteeing optimal performance. It is unlikely

that deep models without feature engineering would produce satisfactory results, the interpretability plus an abundance of features requiring some additional filtering are crucial for the admissibility of ML methods. Additionally, feature engineering allows us to use existing models from the statistical toolbox. Furthermore, besides the performance of prediction with ML methods, it is vital that feature engineering is not done to just improve numerical results of predictive accuracy of separate methods using similar features. In the end, it is important that the whole risk analysis is plausible and the advisement of authorities has an interpretive and actionable effect when it comes to making important practical loss prevention decisions.

Employing a well-conceived feature engineering process with an emphasis on domain knowledge, suitable ML algorithms, and smart additional treatment of important features will help reduce the path to a working risk model. While the final implementation of any specific model may be and often is done with just a few features, providing a substantial amount of features reflecting the uncertainties of an underlying economic problem improves the explanations provided by the model and also the risk analysis.

2.6. Model Training and Validation

Training machine learning (ML) models is a nontrivial process that requires the proper selection of training data to achieve the desired results. Various validation techniques determine if the results of the models can be trusted to make real-world predictions. In this chapter, we will explore both areas to aid practitioners in effectively and accurately training ML models.

Training Data Selection

The training data is perhaps the most essential facet of any machine learning solution. Data drives all aspects of machine learning, yet practitioners spend the least amount of time selecting the data. The models themselves are trained for a few hours, yet data selection can take months. Often the model must be retrained in secondary development phases, as the selected data is not the best or most appropriate for the business problem being solved. Even defining the business problem is difficult until worse models are built. The early phases of development are filled with uncertainty, and naive decisions can lead to wasted time and resulting solutions with no practical use. Mirror the real world as closely as possible.

When selecting training data, one major consideration is the predictive feature set. Which features are correlated to the outcome of interest? This first means defining the potentially predictive features and using domain knowledge is incredibly important. If the feature set is already known, then data selection is relatively straightforward. Selecting predictive features is one of the few areas of model building in which domain

knowledge is best. Most model training is best done by those familiar with ML and will result in better and faster results.

Model evaluation must also be considered when selecting the training data. If the only possible evaluation is error rates or predictive metrics, then care must be taken to ensure that the training data represents a possibly complete subset of the data across all outcome variables. By understanding the evaluation techniques, decision can be made about how to partition data into training and test sets to influence model selection and hyperparameter tuning.

2.6.1. Training Data Selection

Machine Learning (ML) models are built from data that exemplify the relationship of interest. By exposing the model to samples of input features along with their correct outcomes, the ML implements the desired function $y = F(X)$, ready to apply to future samples X test. Therefore, the training data has a predominant role in model building, whether the training is implemented with a supervised or unsupervised machine learning approach, reformulating the quality of data, and how much data is available for training. ML algorithms usually require significantly larger datasets than traditional statistical modeling methods.

The challenge, in this case, is to fill the dataset with more instances, that don't exhaust the available examples' dataset. This is understood because the amount of available data for crucial business decisions in the insurance industry is usually tight. Balancing between the need for larger data and the available data will need creativity on the business analyst or risk experts' side. The industry is moving to more socializing moments, which makes it easier for companies to extract more data from their clients. Better data leads to better models. From traditional unstructured or biased data, to what has been called the Big Data, and more recently to complete transparency offered through Blockchain technology.

The data available through these many methodologies is the stepping stone to better and better-performing ML underwriting models. However, given that these models perform non-transparently and non-explainable functions, they may end up being better-performed tools for the incumbents than for the startups, as the better data will always be with the clients with most business with the traditional player. It is important to highlight that these non-linear underwriters shouldn't be seen as substitutes for traditional underwriting. It is extremely unlikely they will manage to be the entire underwriter process.

2.6.2. Validation Techniques

While the classification model we are building predicts business outcomes such as 'Will a customer default on loan payment?' and 'How likely is the customer to churn?', the best model performance may be hidden in the associated metrics. We build various models, and tune the hyperparameters to achieve the best metric score, but how certain are we that our trained model will generalize? Generalization is essential for any supervised model that we build: the trained model must perform well on the real-world data points it has not been exposed to during the training. In simple terms, generalization is ensuring that we solve the problem for the stakeholders in the intended manner. It is crucial to have a mechanism to assess whether these built models indeed generalize. As real-world use cases, we perform a very simplistic exploratory analysis to show how many factors estimator styles capture model predicting scores – such as when the model is deployed to detect fraudulent transactions or in customer journey scorecards when assessing the likelihood of customer retention for marketing promotions. In this chapter, we usually partition the data into training and validation sets at an early stage in a typical application. The validation data is used to gain insight into how the model may perform on unseen data points, and then the results are used to select among competing models, their parameters, or both. Even though its name seems to suggest its primary use for validation after the model has been built, the validation set is pivotal for the entire model development effort, and the major variables are selected or tuned over iterations. However, the model needs to generalize well and the objective of these collections of data is how well these results mimic what a model would do on future observations. Thus, if we don't properly set up the partitions, we cannot trust the results and the main advantage of data resampling procedures is to bring some trust into these validation results – when it generates an estimate of the generalization effort. Data splitting is usual because it is simple, quick, and easy to follow.

2.7. Ethical Considerations in Machine Learning

Machine learning has ushered in a new epoch of innovation, with the potential to radically transform any field that generates and uses data in significant ways. However, this ground-breaking technology can also be misused, leading to societal harm as well as impact related to data privacy. Many of our daily activities, from when we put out the garbage to our credit ratings, and our social media footprint to our resume for a job application, are being monitored and evaluated by proprietary algorithms that affect our activities in ways that we may not understand or be aware of. Bias and discrimination on sensitive attributes such as gender, race, and ethnicity have long been issues of concern, particularly when it comes to fairness for marginalized and under-resourced communities. In underwriting and related areas, the additional ethical questions of life-

and-death issues and accountability, especially when algorithms are used to deny individuals healthcare or credit or charge them more based on algorithmically determined risk scores, loom large.

Just what constitutes bias and discrimination in the context of machine learning? Traditionally, policy decisions on matters such as hiring of employees or car loans have been based on human judgments, and the courts have ruled on whether or not such decisions are discriminatory. With the increased use of algorithmic assessments, the equal protection clause has been challenged. The inconsistency in finding fairness criteria stems from the fact that machine learning models generate risk estimates for subpopulations based on a particular set of measurable attributes, usually referred to as "features" in machine learning parlance. Machine learning models do not explicitly predict sensitive attributes, yet these attributes may well be the basis for determining risk-related decisions.

2.7.1. Bias and Fairness

Machine learning is a fascinating but complicated new direction for statistical methodology, as it introduces flexible and novel supervised risk estimators, typically based on least squares minimization or penalty minimization and based on nonparametric empirical process concepts such as multidimensional average covering numbers. Many such estimators are now in vast use in countless disciplines of business and policy-making that affect the outcome of how hundreds of millions of people live their lives each day. In this essay, we discuss two important ethical considerations when utilizing machine learning in such institutional practices. The first ethical concern is whether the predicted event rates form a nonlinear function of the predictor variables such as age, income, or gender, so that the model treatment would be more favorable to a certain set of groups. This is the bias and fairness issue. The second ethical concern is whether, for such large data sizes, these event rates are valid estimators of the true predicted probabilities of the predicted event and so as used they are not a source of model-based bad predictions and false alarms. This is the transparency and accountability issue.

We then assume that the goal is to predict the probability of the event of interest in a specific population based on a large data set, which includes the population in which to predict the event occurs for some of the subjects. In practice, one does not have to exclusively utilize the locational likelihood estimator, as other point estimators could be used, but one computes the predicted probabilities using special plug-in functions whose output are estimates of the true conditional probabilities. No matter what unique procedure from among thousands is used to compute the event probabilities, once derived, these predicted probabilities are utilized algorithms to create policy

recommendations concerning an individual or group of individuals. That is, a treatment is based on current predicted probabilities.

2.7.2. Transparency and Accountability

Many machine learning systems are "black boxes," providing little indication of how their internal mechanics translate data into predictions. This opacity presents a problem in high-stakes fields since a decision that would radically alter a person's life could easily be made based on inscrutable logic. As such, recent moves to increase accountability have prompted both renewed interest in model transparency and also in audit frameworks that can help inform users about how a given model functions, and the kinds of mistakes it is likely to make. Model transparency encompasses two different types of explanations; one, which we might call "internal transparency," refers to models that are simple enough and used in a manner that allows a human observer to gauge their behavior. The second kind, "external transparency," refers to procedures that can be applied to arbitrary models to generate insights about their modes of function and their particular failures.

A well-grounded criticism of black box models is that they enable abuses of decision-making procedures. By allowing operators of machine learning algorithms to maintain the ability to depend on inscrutable logic, accountability is undermined and bad actors can escape without consequences. In the interest of ensuring that putatively moral actors are engaging in moral decision-making processes, it is necessary to determine the value of a model's predictions. Output evaluations can be made using tools like data audits, which describe the domains of a model's predictions, and error analyses, which assess the nature or extent of its shortcomings. These preliminary evaluations set the stage for deeper inquiry possible with input-output probes that examine how model predictions vary with different data inputs, as well as layerwise relevance propagation, which determines the contribution of individual features to layerwise activations.

2.8. Conclusion

In this chapter, we summarize the major points of the previous discussion, recapitulating the meta-research, problem definition, tasks, data, and predictive pipeline building upon which we founded the first complete study to apply machine learning to medical insurance underwriting risk classification and screening. In this light, we provide practical suggestions for similar development efforts in the future and speculate on some possible avenues for further research. We finally subject the technology underlying our predictive system to a little deeper scrutiny and give answers to some discussions on its implications and impact on humanity. The last standalone chapter is a more general

conclusion in which we stress and discuss the openness, exploratory nature, and practical focus of the research project supporting our work. We argue that the present modeling framework is interpretable enough to be useful in practice, even if some of its assumptions are by the way quite strong and somewhat simplistic. We conclude with considerations about the critical importance of performing transparency enforcing validation and robustness testing of similar models before one can think of deploying them for risk management or risk transfer purposes in life and health. Despite the questions asked, intelligent or not, underwriting is here to stay and eyewitnessing important developments shortly. It will be greatly facilitated and synergistically enhanced by recent machine learning advances in algorithms and increasing data volumes and granularities, not only in the insurance industry. However, while machine learning may help improve pricing accuracy, access better loss history, and alleviate – possibly reducing – loss adjustment costs, it is hardly likely to be used seriously for building long-term relationships with clients.

Fig 2 . 3 : Machine learning for risk assessment

2.8.1. Final Thoughts on the Future of Intelligent Underwriting

In summary, machine learning and other complex methodological approaches will significantly impact underwriting and risk assessment shortly. We are confident that these methods, with their unique capabilities, will be valuable additions to - if not substitutes for - traditional statistical approaches. Likely, an intelligent underwriting process making extensive use of machine learning methods will create additional bonding factors between insurers and clients. We believe that in many loss-assessment dimensions, underlying risks will be more precisely measured than is currently the case. The limitations of catastrophic event modeling will be majorly overcome by an intelligent approach to risk pricing. Better assessments will increase client acceptance and modeling trustworthiness and will reduce moral hazard - as less risk populations will be subsidizing riskier partners within pools and planner's margins will shrink due to fewer deductibles. As a by-product of better assessments of individual risk, the proportion of process-induced errors and executive loss decisions will be reduced. Ideally, the final decisions will be left to the planner's human judgment as supervised but not restricted by quantitative methodologies.

We believe that prudent company executives should institute the intelligent underwriting ways recommended within their companies in a trial-and-error fashion. Senior management must recognize that there is no one-size-fits-all solution and that a wide range of options is available. With the assistance of capable consultants, planners should get started as soon as possible with constructive work programs. Companies that fail to do so will be left behind as innovators and early adopters reap the benefits of these methods. Insurers that cannot afford the massive investment involved should consider outsourcing underwriting and diary management and at least remain in touch with intelligent underwriting methods for the more expensive above-retention limits share of loss assessment to appropriately evaluate the performance of the external vendor.

References

Yoon, J., Zame, W. R., & van der Schaar, M. (2018). Estimating Counterfactual Treatment Outcomes Over Time Through Adversarially Balanced Representations. NeurIPS, 2018, 1–11. https://doi.org/10.48550/arXiv.1805.09501

Wüthrich, M. V. (2020). Machine Learning in Individual Claims Reserving. Scandinavian Actuarial Journal, 2020(1), 1–23. https://doi.org/10.1080/03461238.2019.1681390

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

Wüthrich, M. V., & Merz, M. (2021). Statistical Learning for Actuaries: Theory and Computational Methods. Springer. https://doi.org/10.1007/978-3-030-78309-6

Henckaerts, R. J., Verbelen, R., Antonio, K., & Claeskens, G. (2022). Machine Learning for Mortality Modeling. Insurance: Mathematics and Economics, 106, 154–170. https://doi.org/10.1016/j.insmatheco.2021.10.005