Deep Science Publishing
https://doi.org/10.70593/978-93-49910-41-6_6

Chapter 6: Artificial intelligence in credit risk assessment: Algorithms that learn and adapt

6.1. Introduction to Credit Risk Assessment

Financial institutions need to know how much credit risk they take when lending money. This is key because high credit risk can cause large financial losses. Making informed decisions about whom to lend money to and under what terms also benefits the long-term financial position of corporations and governments. Financial institutions that set fair terms for the credit risk they take are more likely to make a profit. A lender who makes large sums of money is better able to guarantee its customers' deposits and investments. Economic systems, economies, and financial markets are also more stable over the long term if banks make careful credit decisions (Crook et al., 2007; Baesens et al., 2015; Addo et al., 2018).

This paper is based on the assumption of this necessary connection between the result of a credit risk assessment and decisions on the allocation of resources. Credit risk assessment is based on statistical methods. The type of evaluation to be undertaken influences the particular statistical model, approaching to investigate factors and their relationships. There is no single best method for assessing credit risk. The successful operation of various statistical models relies on the possession of accurate data, though this is a high standard that cannot always be met. The accuracy of the data affects how accurately the models can interpret the relationship between different factors that influence credit risk. Two factors have improved the relevance of the use of statistics in credit risk assessment for banks: computing capacity and an increasing interest in how technology can be used to optimize and organize business operations (Lessmann et al., 2015; Yurdakul, 2019).

Research has been published on technology that scans companies' financial statements and how it can be used to analyze credit risk. The accuracy of these instruments in predicting bankruptcy and credit risk is very high. Not only do these instruments show that there is an interest in the use of technology in credit risk assessment, they can also be seen as an example of the wisdom of the operation of financial and loan operations. The recent study demonstrates that technological solutions are available that show how business activities are planned. These aforementioned companies are specialized in providing technology that can assess the entire credit risk of a bank's customer base. Technology has been developed around the globe to familiarize banks with the credit risk posed by an individual customer. Artificial intelligence technology is able to combine and resell credit evaluation rules from different banks around the globe. As a result, the technology can evaluate the credit risk that the company in question represents within different loan project banks under different loan concession criteria.

6.1.1. Significance of Credit Risk Assessment in Financial Decision-Making

Credit risk assessment is a fundamental business practice that has a direct impact on practices and strategies for borrowers and lenders alike. When deciding whether to issue a loan, lenders factor in the possibility that a borrower might not pay it back in full or on time, also known as credit risk. Lenders often approve a borrower based on what percentage they believe would repay the loan. It only follows then that they would charge higher interest rates to borrowers more likely to default than for those less likely to default. Credit risk assessment differs from creditor to creditor and is a vital component of consumer or borrower finance. In an increasingly interconnected world, it is important to focus on them all. The purpose of financial capital is to prevent the credit risk levels of these forms from rising too high.

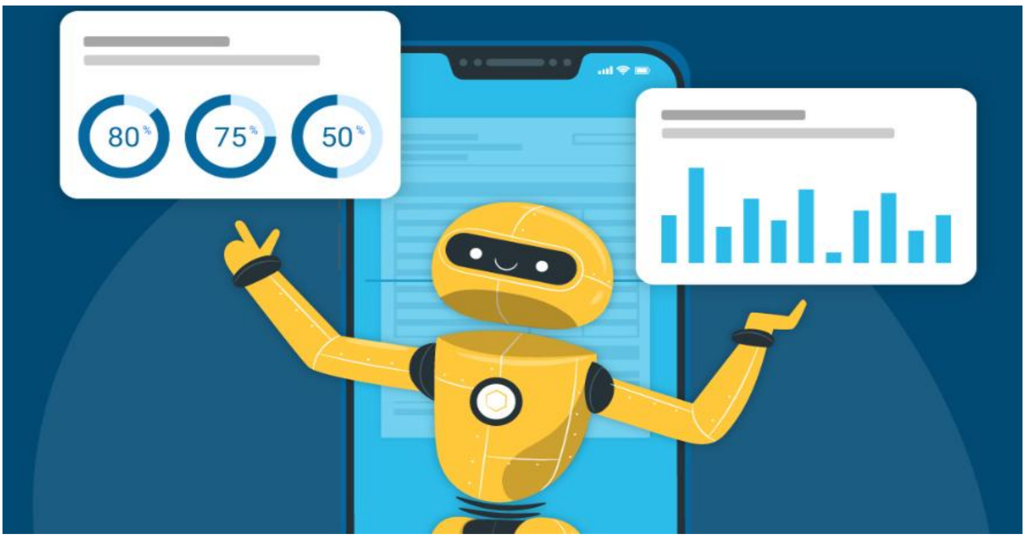

Fig 6 . 1 : Understanding AI for Credit Risk Assessment

Credit risk assessment is also an important concept from the banks' point of view. Credit risk involves the loss of profits from insufficient loan approvals, the collection of denied loans, and continuing to make loan payments. Those with the most loans, such as financial institutions or banks, must adopt methods and tools for resolving these debts due to loan default. Market interest rates fall and the investor's assets decrease if the trading counterparty finds this debt insurmountable and loss-making. A recession or rapid economic expansion is considered a financial system result. Businesses and households stipulate services provided by financial organizations. As a component of the initial credit estimate, banks are expected to have expertise on the latest and most precise credit risk model. Banks are now expected to perform loan surveys as a common requirement. Banks will prepare financial reports using the exact method. Regulations will dictate whether government companies will issue financial information. Protecting financial institutions from financial instability is one goal of precautionary policy. The credit risk evaluation framework will need to be updated as time moves on and when new evidence develops. The recent revolution of information devoted to evidence-based guidance is based on piecemeal investigation and subjective reasoning. Efforts are being made to increase the authenticity, relevance, and level of opinions, and implement necessary tools. This has made it easier to monitor the calculating activities and risk levels in a more current location. It is important to develop and execute the organizations devoting themselves to ensuring that they can conduct ongoing evaluations in uncertain and difficult operating environments.

6.2. The Role of AI in Financial Services

The financial sector is being broadly transformed by artificial intelligence (AI). Innovative AI applications are already being utilized for responding to customer inquiries, automating processes, and providing insight for customer-related offers and credit risk assessment. One important application of AI in the financial services industry is in making the credit assessment processes more effective and efficient. Automated AI-based credit risk assessment is able to process a substantial volume of data and support the efficiency of the credit approval processes to assist in processing large volumes of credit applications. This technology allows many relevant pieces of information to be integrated and processed to create a creditworthiness profile of the borrower. This can improve predictive empirical models used for risk and create a unique profile of the borrower for a better pricing strategy on portfolio risk, which is important for the management of financial institutions.

AI involves the use of automatic adaptation machines, together with the application of learning algorithms to upgrade the models after training, to make a prediction, conclusion, or AI-driven decision. These capabilities have broad applications in the business sector, including in credit operations. The development of high-cost digital data systems that operate continuously allows many complex operations to be applied to this data to make credit operations more effective and efficient, supporting lower expected defaults and more controlled profitability. Automated credit scoring reduces human bias, contributes to process efficiency, and can be integrated into customer engagement strategies, such as websites and mobile applications. AI systems can be an important aid in the decision-making process for potential borrowers and provide rapid feedback on potential lending opportunities. Yet, there are potential issues in managing, developing, and utilizing AI systems. Regulators around the world are attracting attention to the risks of not ensuring that these systems are developed, utilized, and managed to ensure their responsible and beneficial integration. Management of these problems is a task for AI managers and corporate boards in the financial services sector.

6.2.1. The Impact of Artificial Intelligence on Credit Risk Evaluation

The application of artificial intelligence in financial risk assessment has brought about unprecedented progress in the effectiveness of credit risk evaluation compared to traditional techniques. One of the most significant impacts is the improvement in data analysis. AI greatly improves the understanding and use of the diverse characteristics of a broad variety of information held by a financial institution. Modern financial institutions develop record-keeping at an unprecedented pace, generating massive data sets for assessing potential borrowers. Apart from previously collected data, it is now possible to consider numerous economic and financial factors identified by AI as relevant, for which rules do not yet exist. The use of emerging, unconventional data can broaden the content of the dataset, making it possible to construct projections tailored to the client and the purpose of credit. Ultimately, the objective of credit risk management is to find the potential debtor and then to offer them credit. Fundamental AI capabilities that make this feasible include predictive analysis and risk segmentation. The process of predicting potential loan default and risk stratification of a broad range of targets was successfully automated and expressed by nonlinear, complex algorithms.

Modern assessments improve over time when using machine learning. The endowment will thrive relative to the obtained information through the marking of new information from the computer learning platform, ultimately looking more like actual results. This dynamic makes it helpful and inventive. Indeed, a view of the dominant long-term risk factors significantly decreases the possibility of incorrect decisions. Financial institutions and credit providers evaluating the borrower's risk of default are more likely to discover and quantitatively attribute new, emerging factors or indications. The capacity of computer models to adjust over time to new evidence in a continually unfolding economic scenario turns into a risk managerial approach that can take immediate account of the detected changes. Rapid detection can also stop the use of data that are useless and irrelevant for pricing risk and, as a result, the expense of investing in information that no longer holds any value. Computer-based models can have lasting, clear benefits for risk evaluation if the length of the use of outdated information or data resources and their modification becomes predictable. Data that have become obsolete after the onset of external regulations or severe risk are factored into model creation and performance evaluation in order to protect the risk assessment process from prying and ethical management. Furthermore, a select amount of data is also retained in the study, which is available to the public. Data from the presidential executive long-term credit payment are retained, as the purpose and necessity of processing this data encourage the legitimacy of processing this data.

6.3. Overview of Traditional Credit Risk Models

To measure the default risk of a borrower, a set of methodologies has emerged and been developed over time. Logistic regression is often used to define risk using classical models. A logistic model assumes a linear relationship between a set of independent variables and the inverse of the probability of a particular characteristic or event: in this case, defaulting on a loan. Another popular choice is to implement a credit-scoring system. By creating such a system, a lender can more effectively rank different borrowers in terms of risk. They often come in two types: application and behavioral. Application scoring raises a decision to grant a loan in the first place, while behavioral scoring measures the likelihood of default after the loan has been granted. These tools are often limited in their outcome. They are best used when the borrower has a documented credit history and when large amounts of data are available for credit risk prediction.

Despite their use, classical models come with limitations. Behaviors and contingencies not documented in the data are not always taken into consideration. Most of these models rely on one particular dataset and are limited in cross-sectional and longitudinal prediction. This results in the historical data used being volatile and built upon assumptions that presumably hold true only for that time. Also, some subjective judgment is sometimes placed to create these models and can be biased. The world is much different than it was 40 years ago within banking and credit services. Lenders now have access to an abundance of data, such as analytics, social media, purchase history, etc.—all things that can give lenders better insights into borrowers and their borrowing, but also power their credit assessment techniques differently. Therefore, learning some of the traditional techniques and philosophies of these techniques is necessary to compare with machine learning and artificial intelligence models.

6.3.1. Insights into Conventional Credit Risk Assessment Models

The main purpose of credit risk models is to evaluate credit risk for loans to individuals. The two major types of credit risk models are the credit scoring system and the behavior scoring system. Credit scoring systems primarily use the prospective borrower's previous credit history as the information used in the loan decision process. Behavior scoring systems are much more intrusive and attempt to measure psychological and mental imbalances of individuals by evaluating the stability of job, address, and behavior at the bank.

Mainly, two types of credit risk models or scores – credit scoring and behavior scoring methods – are used for the evaluation of credit risk. It is argued that most of the credit risk models discussed in the branch of econometrics and financial economic literature use the credit scoring techniques in the developed credit markets. Credit scoring models are used by bankers as an eminent and effective tool for the quantification or measurement of credit risk. The speed of pioneer credit scoring models gained impetus after two landmark papers. In these works, piecewise interpolation formulas for the inverse cumulative of certain percentiles of the standardized logistic normal distribution were provided. These interpolation estimates give closer results of the inverse cumulative logistic when compared with previous methods of simplex lookup tables.

6.4. Machine Learning Techniques in Credit Risk

Systematic risk evaluation is key to the creditworthiness of borrowers. Currently, banks use traditional scoring models, which have a high degree of operational risk involved with many subjective interpretations. The machine learning technique is understood as an innovative methodology to support the empirical studies of credit risk assessment. The advantages of these algorithms include high predictive accuracy and model adaptation. The accuracy level on the testing dataset is usually determined by the parameters of a given search. During the training stage, one searches for the model with the highest level of predictive accuracy. However, this strategy may lead to an overfitted solution. It is important to select the algorithm that is the best fit for the purpose and quality of the dataset. The default prediction is a clear example of when non-linear models are used in a study. In this case, the machine learning technique enables one to predict potential credit loss. The behavior of the dependent and independent variables is often volatile, which is difficult to capture by traditional logit or probit models. This has been referred to as the issue of capturing framework complexity, i.e., capturing space-time interdependence between borrower characteristics and macroeconomic influences on credit risk. Various machine learning techniques take into account the full set of interdependencies and thus make accurate predictions of the probability of default.

ML Advantages and Techniques

The economic literature on credit risk management postulates the use of machine learning algorithms in evaluation research. The machine learning algorithms offer increased efficiency, accuracy, and model interpretability. They are capable of capturing the relationship between dependent and independent variables. Decision trees and random forests algorithms have been used and evaluated as potential benchmarks for credit analysis. The model development lifecycle from data preprocessing to model deployment can be automated for the analysts. Post-consumption of the results becomes relatively quicker in predicting good deals from bad ones, with significantly increased accuracy. Moreover, as an outcome of such techniques, analysts can recommend marginal conditions for classification. To summarize, further research is recommended in the machine learning area and model interpretability to enable use, further development, and popularization of this methodology in credit risk assessment. Model results should move from 'black box outcomes' to 'model results messages' strategies. Data gathering and quality control issues are seen as main research directions to enable the dissemination of the results in academia and the financial sector.

6.4.1. Supervised Learning Approaches

Supervised learning, a critical component of machine learning, provides a framework to develop predictive models of empirical regularities that can be used to assess the risk of a future, unknown outcome. The models are trained on recorded historical outcomes, or labels, to then predict the risk of future, unknown outcomes given relevant features, or attributes associated with the current and past environment. Credit risk assessment is one of the first areas where advanced statistical learning methods have been applied. The resulting Z-score model defines the boundary that separates defaulted firms from non-defaulted firms. Since then, the technique was extended and a similar approach, namely the logistic regression analysis, a classification technique, was used in a credit risk modeling context to compute probabilities of default. Support vector machines and boosting algorithms have been used to further advance logistic regression for risk modeling purposes.

A variety of other credit scoring models can be found in the literature, making use of diverse machine learning algorithms. Note that the performance of these credit risk models is critically influenced by accurate feature selection, data quality, and sample size. Several finance institutions also use these models for credit risk assessment. The purpose of a credit scoring model is to classify customers into groups based on characteristics of their accounts. The dependent variable in a credit scoring model typically equals 1 if the loan defaults and 0 if not. The model outputs a risk score which estimates the probability of default for each new customer. This risk score is used in the iterative credit decision process, helping identify defaulting customers and thus avoid untimely credit losses.

One of the most important aspects in developing a credit scoring model is the balance between predictive power and model interpretability. The more advanced a statistical method becomes, the less transparent the effect of the estimated coefficients is. The trade-off between complexity and interpretability is very important when designing the model to be used in practice. The process of model building is extremely challenging. Models are trained on previous data and used for future predictions, so they must be robust to unexpected changes in input data and are supposed to be white swans: a white swan model is supposed to predict not only the probability of default in normal times, but also the probability of default for given compound assignment in the fat tail event. There is no doubt that complex models could capture relationships in the input data that are invisible to more parsimonious models.

Overall, supervised learning is an essential part of advancing the credit assessment models. However, more sophisticated models must be well validated before they are implemented. Using them only in seed approaches is recommended to avoid shot hits.

6.4.2. Unsupervised Learning Techniques

As noted earlier, supervised learning constitutes the main approach in assessing credit risk. In contrast to these methods, unsupervised algorithms deal with correlating aspects of a phenomenon with no pre-labeled knowledge about the outcomes to be estimated. These methods have much appeal in credit risk assessment, given they can help us discover hidden structures and relationships that might predict default but are unaccounted for using traditional credit scoring and other models. They are effective at grouping borrowers who share some similarity in their characteristics. The most common unsupervised learning techniques in credit risk assessment are clustering and anomaly detection.

Clustering helps in grouping random observations into meaningful subclasses using centroid-based, hierarchical, or density-based methods, thereby assisting in identifying loan applicants with similar characteristics. They can also be applied to discover clusters of defaulters, or those whose characteristics change to indicate higher default probabilities when they start to become less compliant. Anomaly detection expects no clear patterns or rules, limits the number of mathematical assumptions, and detects, in a population, those members who are different. They are applicable to identify hidden and unverified risks within a relationship. Unsupervised methods are useful for segmenting customers and targeting marketing variables, providing insights into behavior to be further investigated using supervised methods. They are not influenced by human judgment, which can reflect racial, gender, or ethnic factors but focus on the relationship between good and bad credit risk decision outcomes.

Unlike supervised learning, unsupervised techniques are easily scalable to accommodate big data, thus providing support for decision-making to manage risk by segmenting and rating indicators of time-related developments. However, their interpretative ability depends on a human's experience in that field of application, meaning that human involvement is necessary when one dictates what to derive from the model and more encompassing knowledge is required to assess the outputs. Unsupervised techniques cannot be used in credit risk assessment as standalone tools and so are applied as complements to aid supervised methods.

6.5. Deep Learning in Credit Scoring

1. The development of deep learning models has recently established itself as a next, more advanced step in the application of machine learning in various areas, where large amounts of data with complex patterns create a challenge. Such a setup is present in credit scoring, motivating an attempt to apply deep methods in this area.

2. Principles of Deep Learning What distinguishes deep learning from shallow and classical machine learning is the use of multilayer networks. Each subsequent layer uses the output of the previous one as an input for creating hierarchical abstractions of data, presenting more complex inputs. By doing so, it allows modeling intricate, non-linear patterns and features present in, for example, the behavior of a borrower. With the correct architecture and sufficient number of neurons, the network can capture some of the underlying, highly non-linear probability distribution of the response being $Y = 1$. This is an advantageous approach to the previously discussed machine learning applications. Furthermore, it can be efficiently and effectively parallelized, which is even more appealing given the enormous size of credit scoring datasets. Deep learning algorithms such as convolutional neural networks or recurrent neural networks can extract and learn complex features automatically based on the input data, a reason for the ever-growing popularity of the methods in fields such as computer vision or natural language processing.

3. Application in Credit Scoring Despite the relative novelty of deep learning, the existing body of academic research has already shown great promise for the application of deep learning techniques to improve the precision of risk predictions in various credit scoring and risk-related problems. Some studies focused on assessing the robustness of deep architectures on general acceptance, while others attempted to incorporate advanced methods in specific scoring contexts, comparing their results with the industry’s standard predictors. The beneficial applications concern micro-sized loans,

ensuring banks can make a profit while reducing the need for manual assessments. However, the added value in dual deep learning is diminished when analyzing defaulters, preventing this model from replacing traditional models. As the area rapidly evolves, more financial institutions are trying to develop and apply their deep learning models, not just out of necessity to keep up with competition, but out of enthusiasm for a novel approach.

These claims make it clear that deep learning methods can improve the accuracy of credit scoring systems in the right areas of application, where traditional methods fail to provide adequate performance. Two elements are crucial to a successful application of deep learning in credit scoring: data quality and data preprocessing. For deep learning methods, the broadness, richness, and quality of the data are far more significant than in any other model. Data preprocessing in the context of deep learning can include digitizing all the usable data of the bank and funneling it through a cloud platform into a scaled-down data lake, where AI investigates, finds patterns, and learns automatically. The results suggest that the use of broad and deep information and the intricate architecture of deep learning models indeed result in higher efficiency until the cost of setting up and maintaining models exceeds the additional gain and technological benefit. To this end, the solution appears elusive given how the evaluation of a deep learning model's performance and setting up a commercial credit-scoring model is entirely different in practice. This refers to the necessity of risk management decisions, the use of model governance, and the validation of internal models, among others.

Fig 6 . 2 : AI in Credit Scoring

6.5.1. Neural Networks Overview

Introduction Neural networks are one of the key algorithms used within deep learning, a branch of machine learning that facilitates models to learn progressively complex patterns or features from the raw input data. Artificial neural networks are constructed based on the structure of human brains: basic structural units are neurons, and their main components are weights, capturing connections to other neurons. The basic form of neural networks consists of three types of layers: input, hidden, and output layers. Neurons in weaker layers transfer information to subsequent layers, and the latent ones have the processing power. Such an organization ensures that information is conveyed through a sequence of transformations. The relevant information about the output layer can be obtained after feeding a sufficient amount of data. Thanks to their architecture, neural networks are widely acclaimed for their ability to model even complex nonlinear relationships that traditional algorithms are unfit to capture. Thus, deep learning can be valuable in tackling modern credit risk challenges, as raw input data nowadays come in various formats, such as text, images, and sounds. Neural network architectures, however, are multifaceted. The feedforward artificially recurrent neural network is the most basic one, but credit risk-related applications have consecutively employed more advanced architectures. The convolutional neural network is excellent at capturing high-dimensional, image-like data, and it has gained increasing popularity in machine learning and big data areas. Respective layers in a convolutional neural network have an interconnected pattern, whereas neural layers in an artificial neural network are fully connected. Additionally, some especially structured credit risk analytics integrate both spatial and temporal features. The recurrent neural network with long short-term memory units is widely theorized as being suitable for prediction models, with particularly favorable drift control and adjustable input data embeddings. At the same time, researchers often opt to use activation functions to introduce non-linearity into the decision equation. Sigmoid, rectified linear unit, and tangent are the most common ones. There exist many optimization algorithms to train these highly complex deep architectures, including stochastic gradient descent, Nesterov accelerated gradient, or adaptive gradient algorithms. Each algorithm presents unique trade-offs between speed and accuracy, or between exploitation and exploration, thus offering variable rates of improvement in key metrics. They lead to better data fitting to hone the model’s accuracy in anticipating patterns and changes in the data. However, in usage for credit risk analytics, limitations come from the trade-off between complexity and interpretability. The high complexity of these models and the standard desire for simplicity can lead to obstacles in model integrity and trust. They also require a large number of training examples to curb overfitting when handling high-dimensional data and reiterative functions. Furthermore, even though single or advanced models present promising results in accuracy and empirical gain, they entail the use of valuable computational resources. For instance, training a convolutional neural network with an average of five million parameters may take approximately twenty-four hours on a standard computer. With regard to data, using a single model for an image classification system may require thousands of data instances for adequate performance; it is important to ensure that the elements of the training data set are representative of the entire population. As such, the deployment of neural networks in business operations calls for consideration of cost-effectiveness, model transparency, and confirmation of value attributed to its application. Some popular applications of neural networks include default prediction, fraud detection, and image and text retrieval. In credit risk analytics, models using structured data, such as financial and demographic data, are dominant, although the recent big data movement has instigated the use of unstructured data models.

6.5.2. Applications of Deep Learning

The applications of deep learning in credit scoring are numerous and diverse. It has been used for assessing customer needs, identifying their main characteristics and consumer behavior, as well as for overall creditworthiness scores. Its ability to extract and work with alternative data allows for targeted risk management, credit scoring of young borrowers, and identifying those using dissimulation. In recent years, competition between fintechs and financial services has grown, revealing the need for better credit scoring models – focusing not just on the black box but also on the methodology for improved cooperation with current procedures. Throughout the review, we refer to applications of deep learning to predict borrower default and to use the modified borrower score for assessing the likelihood of defaulting.

Deep learning has some benefits over traditional machine learning in terms of higher accuracy in in-sample and out-of-sample predictions, better ability to handle unstructured data, and to automatically process feature extraction. In finance, it can capture the relationship between different economic, political, and financial factors associated with vast data. However, this also implies the difficulty – or in most cases impossible – to be audited as it is unknown, in great detail, how a deep neural network arrives at its output. The predictive component of deep learning allows adaptation to the fast-changing environment, where new predictors always emerge while robust risk management employs statistical audits and judgments to avoid quick judgments leading to a decision bottleneck. Compliance and regulation issues will also be discussed as these are very important in finance. The literature that employs deep learning for credit scoring is typically agnostic in terms of producing policies to be implemented. In this literature review, we examine the potential and limitations of real-time credit scoring using deep learning models because it is more generally possible now as more banks and NBFCs move from rules to bank examiners using machine learning algorithms. In finance, usually, humans make risk decisions based on business intuitions, and their statistical significance is independent of system design, although banks also stress test out-of-the-box events as cyclic components besides measuring current risk. The review concludes by providing a showcase in real-time credit scoring to demonstrate deep learning implementation in a business setting. This example in various banks has shown more than 10 percent new impactful information in real-time credit scoring and finance to improve cost, profit, as well as GDP growth.

6.6. Data Sources for Credit Risk Assessment

The scope of this book is the application of machine learning algorithms in credit risk assessment. To achieve this goal, the user should have enough data to build a model evaluation. There are many data sources we can use for this purpose; these sources are divided into traditional and alternative data sources. Traditional data sources include credit reports, application forms, and financial statements. They are mainly used for assessing the risk of borrowers when an institution applies for a standard loan. The use of financial statements data is possible in the case of corporate borrowers; in turn, as far as lending to individuals is concerned, income certificates and tax returns are required. An additional type of data used by financial institutions to assess the credit risk of corporate clients is qualitative data, i.e., regarding social and environmental risk.

In turn, predictive models adequate for this type of assessment require the use of data regarding the behavior of borrowers from sources other than credit reports. These data are purchased or collected by banks mostly for their own purposes. It is possible to distinguish five main types of alternative data sources: 1) Retail payment history of the potential borrower; 2) Public registers and databases: data that provide information on the background of the borrower. This includes checking for criminal records; 3) Records of judgments, foreclosures, and executions: the potential borrower's prior history of loan repayment may be obtained from public records that contain filings or court judgments regarding bankruptcy, levy, or foreclosure; 4) Discretionary credit data: other information about the borrower(s) that they opted to divulge. An example might be a past payment performance history in sectors such as utilities or insurance. Moreover, credit scoring models, in addition to the use of credit reports, also use advanced credit data, e.g., such data from social welfare offices, which provide data on the payment discipline of individuals, or such life insurances are also used as a source of information on the time of obtaining insurance, the sum insured, the amount of the premium, and the personal data of the policyholder and the insured. These data are obtained with the policyholder's consent after specifying the policy using electronic data transmission.

6.6.1. Traditional Data Sources

The process of assessing the likelihood of borrowers meeting their credit obligations is an essential function of modern financial markets. Systematic approaches to measuring and modeling the risk of default in borrower assessments have deep historical roots. Over time, and in the absence of modern computing power, lenders have used readily available and easy-to-digest data to support the creditworthiness assessments of potential borrowers. At the core of credit risk assessments is the credit report provided by credit bureaus. This central document displays the repayment history of an individual’s previous debts. It is important to keep in mind that a credit report inherently measures an individual’s ability to pay bills, not his or her willingness to.

In a small business setting, lenders might have focused more on the financial statements or bank account transaction data of a company or owner of the business, and may have discussed operations, management, strategy, and industry specifics with the business owner. Ultimately, as the process functions today, the credit assessment process is influenced by the rules and regulations of the relevant state, the size and purpose of the loan under consideration, as well as the policy and strategy of the lender or investor involved. The location of the borrower also impacts the scope of credit underwriting by the lender, as local culture can affect the norms of borrowing and repayment. Overall, the three credit bureau reports used today, in combination with annual and financial data submitted by the borrower, past account aging data, collections data, and rare supplemental qualitative data acquired through further discussions, internal databases, and a variety of public data record databases, illicit proper portfolio for a lender. Furthermore, this repository of financial, demographic, and behavioral data will continue to evolve as time goes on and the spread of character and capacity data continues.

6.6.2. Alternative Data Sources

Traditional inputs are slowly being, but not completely, replaced by the more novel data sources that seek to contribute a specific and potentially complementary aspect to a borrower’s risk profile. These include utility transactions from mobile network carriers for prepaid phone services, as well as direct deposit histories that shed light on income stability, inflows, and the presence of other sources of wealth. Other alternative data inputs include applicant-reported rental history information, social media profiles, utility consolidating auto-pay services, and underbanked individuals’ mobile payment platforms.

Leveraging information from someone's online presence at scale can provide deeper insights into the propensity of the borrower's creditworthiness. By doing so, aspiring customers who do not have access to traditional credit data on one hand or lack collateral to secure loans on the other hand are suddenly on the banks’ radar. This is in harmony with the sustainable development goals of financial inclusion. While demographic data shared through non-traditional sources makes understanding a customer in a better light possible, it is important to obtain consent from customers and ensure compliance. Since alternative data isn't mandated by regulators, ensuring that this is sufficiently explained and having clearly stated terms and conditions before collecting it are mandatory to ensure that this is done in a fair manner.

There have been concerns about the ethical use of alternative data in various studies and news stories that have depicted potential biases in the profiling of prospective borrowers based on these new available customer data sets that are often grounded in demographic attributes. Ensuring that the data we use is inclusive is the Achilles' heel of this model ingestion since measuring bias and fairness, considering that there are no clear feelings about what might be considered unfair versus what is fair, blurs the lines. However, a pronounced agreement in the literature indicates that the process does allow, to an extent, for the prediction of credit risk of individuals who haven't borrowed in the past. Consequently, a more balanced and optimized credit decision and customer experience can be achieved through a combination of conventional and contemporary data sources. Some challenges include internet coverage in various parts of the world that may lead to an unclear or incomplete online trail and, across some demographics, a more aggressive or laissez-faire online presence. Hence, the models and data pipelines being built include cleansing and preprocessing data as well. Firms need to be aware of the regulatory and contractual risks of using alternative data and/or vendors. Not complying with data privacy laws or credit reporting legislation may lead to lawsuits or regulatory penalties.

6.7. Feature Engineering for Credit Risk Models

Credit risk models require a range of input data from applicants, including personal information, collateral information, and payment history. The number of features a credit risk model should consider is an essential aspect of the data preprocessing phase. This step aims to select and transform variables to extract maximum model performance. The main objective of feature engineering for credit risk assessments is to build variables (features) that are meaningful to predict creditworthiness. This importance is expressed through correlation measures with the target, which can be used as an input for interpretability and reliability.

The choice of a particular feature generation and transformation technique is based on a business context and is a case-by-case solution. Nominal features can be transformed through different property encoding because the coefficients depend on the chosen base level. The choice of integrating the business domain into the development of the predictive model is emphasized in the banking sector because domain relevance can provide an understanding of what to investigate. The integration of domain expertise into the feature engineering process means that the collaboration of financial economists and data scientists is of great importance. Statistical techniques need to be used to verify the validity of domain relevance. The feature selection process using domain relevance alone can lead to biases. The complexity of the feature should be re-examined to address the issue of overfitting. The encoding of new features as a combination of feature outputs can improve the correlation and relevance. In practice, choosing between interpretability and stronger performance is necessary.

One of the advantages of feature engineering for model interpretability is that it can help explain the relationships and reasoning behind model predictions to end users. Finally, feature engineering requires further examination with a large volume of non-linear relationship data, and this can be carried out using machine learning-generated features to see the relationships between variables. Controlling the process of feature engineering iteratively to maintain feature accuracy is a mandatory process in developing reliable credit risk models. Model performance variations occur easily due to the fluctuation of data and should be regularly monitored. Ensure that new data features are created every time the structure changes, taking into account the volume and number of features in the model. The re-derivation of features varies depending on the business case; the final random effect result is obtained from the recent data. The meta model using a weighted scorecard helps improve the integration of decay effects and the initial scorecard's time dependencies until 12 months from the date of origin. This study also explains the importance of different types of features such as numerical, categorical, ordinal, and target for good model performance.

6.8. Model Validation and Testing

Evaluating models is a key part of the model development process. This critical ability is required to detect problems that can lead to model failures. Model evaluations are for verifying that models are designed correctly. Model testing primarily consists of analysts developing models and applying the models and observations from the training data set to predict against a holdout set that was not involved in the model development. The fitting behavior of the predictive model should be maintained when the model is used on the holdout set.

It is performed using some validation techniques such as cross-validation, random sampling, and the holdout method. These are a very important step while segmenting the data. The evaluation process consists of selecting the appropriate performance metrics that will provide a comprehensive understanding of the effectiveness of the model-specific organizational practices, rather than focusing on only one outcome such as accuracy. Adaptations of the model development process, such as optimization of model parameters, were required to improve the effectiveness of models. It is also important to stress test models based on datasets for which expected performance is known, and then see how models would perform in an adverse environment. Validating a predictive model for origination over multiple unknown time periods and business volumes will help to identify the potential usefulness over a range of potential scenarios.

The development and implementation of predictive models that equity lenders will be using to analyze client portfolios in the dynamic market will be supported by this research. Credit risk examines how lender portfolios change over time and what the future holds for lender profits. The dynamic markets require predictive models that are effective and flexible; they will therefore be based on the application of artificial intelligence algorithms. However, predictive models have a lot of rules and restrictions. Over time, it is expected that model drifts will occur. It is still important to validate models in a preliminary stage and offer options to handle it. It is often vital to tackle model drift as it is a vital part of the risk management process. It is often vital to validate models in a primary stage in order to check the potential for probable future results.

6.8.1. Performance Metrics

When validating models, defining the performance metrics to be used is crucial to evaluate the real effectiveness of the model in assessing credit risk. The simplest metric to use is the accuracy, which is the fraction of correctly predicted instances. More extended analysis can be conducted by using precision, which is the fraction of relevant instances among the retrieved instances, and recall, which is the fraction of retrieved

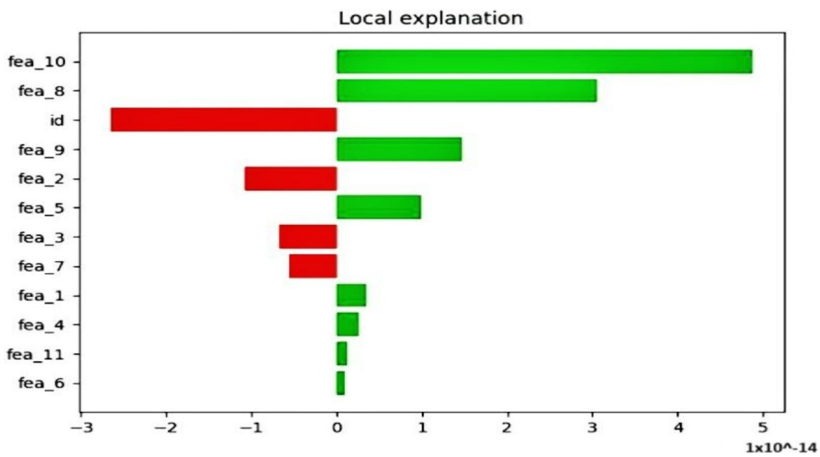

Fig 6 . 3 : Credit Risk Assessment and Financial Decision

relevant instances among all relevant instances. The F1 score is the harmonic mean of precision and recall. So, the confusion matrix includes false negatives, false positives, true negatives, and true positives. In practice, the choice of the metric to be used to measure the performance of a model strongly depends on the purpose of the model. To make a proper trade-off between false positive and false negative, it is crucial to associate consequences to them: in a credit risk model, for example, the consequences of a false positive and of a false negative are not symmetrical. Another possible issue is the imbalance of the data: it may lead to misleading results in the performance evaluation of classification models. Some strategies such as under-sampling or over-sampling can be performed to mitigate the problem of imbalanced datasets. In building predictive models for credit risk, one should adopt a strategy that adheres to compliance with regulatory requirements. In each universe in which financial activities operate, banks must comply with laws and regulations established by the supervisory authorities. Furthermore, every day when we estimate these coefficients, or we obtain the outputs, we deal with another challenging issue: considering that the number of default payments in a timeframe changes depending on the reference universe, we may encounter low stability in the model built on one subset when trying to apply it to a recent subset. That is why we always need to constantly monitor and re-estimate the previously mentioned metrics to detect any potential deviation and make the required adjustments to stabilize the scores.

6.8.2. Stress Testing Models

Model users and regulators increasingly require that financial institutions have tested their models under severe macroeconomic scenarios to diagnose the impact of faced challenges and adverse conditions. Stress testing is used to evaluate how the model can perform under a series of adverse conditions and how these conditions could potentially enhance losses and impair the accuracy of expected results. Stress testing is done by using stress testing scenarios and techniques. Scenario analysis uses one or more predefined probable negative external events and registers every required change in the model's outputs after it has been noted. On the other hand, an attempt to record the changes in model output caused by the variation of one or a group of inputs is rooted in a sensitivity method. Due to the examination of a large portion of data, stress testing covers assumptions that involve making one to five-year zero loss assumptions. Indeed, in model decision-making, extreme observations should be considered to create a model that is immune to the potentially very large loss values caused by these. Measurement of recession and recovery of the position should also be done because revealing the impact of the clubs has only a drawback at zero point and credit.

Scenarios and stress tests should not be the average result carried to the future of any economic and financial evaluation but should extend the determinant along all unobservable and observable financial markets that describe the robustness of the published premises. As a result, a financial institution could run meaningful and valid simulations for the individual stakeholders to combine various outcomes. Therefore, a base run is also used in stress testing, i.e., how the institution can survive a specific recession or series of material currency changes. It is important since banks require estimating what the outcome could be if they fail in order to save a more disastrous outcome from bankruptcy and collection of debt if the institution is likely to follow insolvency. One of the main objectives of the stress test is to decide on the suitability of these establishments. Therefore, a sound stress test should supply the financial institution with data on whether the procedures and functions it needs are suitable and correct. As a result, government agents or management bodies are recommended to initiate a stress test, which should make the results transparent. However, the interest of consumers and suppliers of the credit institution should be provided. In addition, any responsibilities, risks, or possible results that could arise from the stress test should be shared with the parties participating. Furthermore, enhance the communication of the results and the consequences of the exercise. These significant actions could have a substantial impact on the readiness of the different consumer classes in an attempt to change threat portfolios if one of the principal consumers or dealers announced that the relevant lending bank could be influenced.

6.9. Conclusion

AI brings a huge potential to improve the established procedures for assessing credit risk by means of either transforming the traditional statistical and econometric methodologies or improving the machine learning and scoring systems. AI exacts continuous adaptation of the methods used in AI procedures and invests in very expensive IT infrastructure and human resources. Innovating in AI means not only coping with an evolving data landscape but also with the continuously changing regulatory environment. Financial institutions that wish to adopt this technology must do so by adopting a comprehensive approach that embraces both the traditional techniques of the banks, where longstanding know-how and knowledge of their clients have a dominant weight in the credit approval attitude, and the most recent AI techniques. The integration of new methodologies and the use of less exploited alternative data sources seem to be the promising directions for future research on credit risk assessment. There is a need for future research on the evolution of banking practices and risk management techniques, particularly those based on AI from an ethical and social perspective because it is unclear to what extent the effectiveness of a shift away from traditional commerce and banking practices, towards an interest in interdisciplinary approaches to credit scoring and AI-driven assessment of credit risk. The evolving standards in data protection, privacy, and fairness for the domain of credit and finance are areas for potential fruitful research. In conclusion, we encourage greater respect and an even more fruitful and multidisciplinary dialogue on common interests.

6.9.1. Summary and Future Directions in Credit Risk Assessment

Since the inception of the first credit rating agency in 1916, credit risk assessment has seen countless technological advancements in decision-making frameworks, databases, and statistical methods. In the last decade, more than 150 new AI research papers and two dozen patent applications for AI techniques have surfaced in the field of finance. AI, in particular, has improved the richness and variety of datasets available for credit evaluation, eased the burden of stringent model assumptions, and supported advanced non-linear and non-parametric model structures. However, AI is not a panacea. It thrives where traditional evaluation systems fail, and it is only successful if developed in collaboration with product designers, regulators, financial analysts, IT managers, consumers, policymakers, and equality scholars. With the ability to adapt, iterate, and consider competitions, smart AI designs can put manual methods to the test. Stakeholder collaboration is especially important in risk assessment. Credit risk assessments will likely continue to evolve in the years to come, emphasizing computer learning's unique income types that take different prediction and judgment outcomes into account. Indeed, more complete data scans can uncover new themes. For instance, the exclusion of alternative finance information about credit scoreholders could provide a unique qualitative and quantitative look at behaviors not discerned from the analysis of preceding customer credit relationships. Machine learning technological advances are expected to continue to evolve in the next few years. Machine learning and AI algorithms will continue to expand the use of several hundred detailed data metrics but are now able to make better estimates for those who cannot fully explain their technical core. Model approval will continue to be an opportunity for research. Additionally, the top priority in the coming years is to establish agreed-upon regulatory and ethical frameworks to increase AI usage in finance. AI's limitations must not result in the routine overrunning of legal rules for economic purposes. Research on AI discrimination will also increase. Looking further forward, more detailed audit trails will be launched to give customers greater visibility into the credit rating process. The credit rating body is now the sole determinant of all decision-making procedures based on anonymous data which it supplies. AI that offers greater transparency in lending judgment, which ensures strong feedback from customer acceptance, could democratize entry into finance and thereby reveal great opportunities. In the long run, the combination of strategic data and AI's thorough design would probably enable us to switch to capital markets' revolutionary mechanisms.

References

Crook, J., Edelman, D., & Thomas, L. (2007). Recent Developments in Consumer Credit Risk Assessment. European Journal of Operational Research, 183(3), 1447–1465. https://doi.org/10.1016/j.ejor.2006.09.100

Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). Analytics in a Big Data World: The Essential Guide to Data Science and Its Applications. Wiley, 1–208. https://doi.org/10.1002/9781119203618

Yurdakul, M. (2019). Credit Risk Assessment with Artificial Intelligence: A Comparative Study Using Neural Networks and Support Vector Machines. Applied Artificial Intelligence, 33(7), 621–642. https://doi.org/10.1080/08839514.2019.1637136

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking Classification Algorithms for Credit Scoring: A Review. European Journal of Operational Research, 247(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit Risk Analysis Using Machine and Deep Learning Models. Risks, 6(2), 38. https://doi.org/10.3390/risks6020038