Explainable Feature Attribution and Temporal Credit Scoring: AI-Enhanced Models for Borrower Credit Risk Assessment

Dr. Katarzyna Szymkowiak, Professor of Computer Science, Poznań University of Technology, Poland

1. Introduction

Credit risk evaluation within banking systems has always been and is still a moot issue. Traditional credit risk assessment models quickly become irrelevant because they do not meet modern trends in banking and new financial processes, instruments, and products. Many recent discussions expose the problems of the existing credit risk assessment methodologies and discuss the need for innovation and the use of the possibilities offered by artificial intelligence and machine learning. The originality of this article consists in the endeavor to develop an innovative solution by improving the existing credit risk assessment models with an artificial intelligence component, more precisely with the capabilities provided by artificial neural networks. Thus, the aim with which the elaboration of this paper started was to enhance traditional credit risk assessment models with artificial intelligence. Several research questions developed on this aim: What are the latest trends in credit risk assessment? What is the contribution of traditional credit risk assessment to banking? How can synthetic indicators created by using artificial neural networks improve existing credit risk assessment models? To provide relevant answers to these questions, this paper is organized as follows: the introduction is followed by one section which presents the role and purpose of credit status assessment within the support of banking operations, as acknowledged by specialized literature. In section three, we will provide an overview of the regulatory requirements applicable to credit institutions regarding the provision of capital to cover credit risk in the credit portfolio, which encompasses the more important normative acts. Section four describes the use of discriminant scoring models in the evaluation of credit status from the perspective of recent changes in the risk environment, illustrated with a relevant practical study. This is then followed by some reflections on creating individual probability of default indicators and synthetic credit risk assessment indicators through artificial neural networks integrated into scoring models, and the model can also be assimilated into the rating models of the listed companies as well. Furthermore, potential improvements and limitations of the approach presented will be indicated. The final section provides a summary of the study.

2. Understanding Credit Risk Assessment in Banking Systems

Credit risk assessment in the banking system plays a pivotal role in maintaining the financial stability of banks and other financial institutions. Such an assessment is critical to conceiving strategies to mitigate the default risks that emanate primarily from banks' lending practices. Currently, traditional ways of measuring credit risk present various limitations. Some of the traditional measures generally used for credit risk assessment are merger models, ratings, and other assessments. Various techniques that are widely used across the globe for the purpose of any new domestic and foreign entry may include CAMEL rating, risk-adjusted rates, lower return on equity, discriminant analysis for off-balance sheet items, and registers analysis not based on accounting data.

The primary objective of the merger model for assessing the credit risk of banks is to evaluate the probability of bankruptcy or distress of a firm in the short run. The expected bankruptcy cost in the form of legal and administrative expenses and the dilution effect on assets resulting from a hasty sale under forced liquidation must be estimated. The consequences of a firm's failure often lead to the firm losing assets worth reorganization and reinvestment. The main objective of credit statistics is to ensure a safe, reliable, and efficient clearing and settlement system. Measures for reducing the risk of credit extension should be observed and recorded. The main purpose of the credit facility is to protect capital resources and earnings against the adverse effects of credit risk. There are a few commonly used ratios in banks that can serve as yardsticks when analyzing a bank's performance. The credit risk of a bank can be minimized, and it will cause the loss of important revenue if banks become risk-averse. In credit risk analysis, we use teaching value at risk and expected shortfall as risk measures. In the increasingly competitive financial market, the quality and value of banks increase in importance. The purpose of this text is to present AI approaches for credit risk assessment and analyze them for assurance support in decision-making.

2.1. Traditional Approaches to Credit Risk Assessment

Quantitative and qualitative methodologies have been used to ascertain the creditworthiness of borrowers. Banks have been employing traditional techniques to evaluate potential borrowers seeking loans. Among the quantitative methodologies used in the process of credit risk assessment are credit scoring and the estimation of the probability of default. Even rating techniques determine the obligor's probability of default; that is, the well-known credit metrics have focused on a borrower's historical ability to repay specifically. On the other hand, for the qualitative methodologies, risk rating has been employed to assess the relative risk of different sectors, industries, and various loans with different amounts and maturities. Therefore, various credit scores and risk rating levels suggest differing historical performance and future default risk.

The banking innovations may create some dimensions of default risk that, in turn, would map into risk rating levels that are not currently captured in the model. The financial ratios-based possibility of default estimate can be biased from discriminatory variable usage, favoring one class of borrowers over another, deriving from the subjective identification of the key drivers or from other quantitative measures used. In addition, by elaborating on discriminant analysis results, other issues clearly arise where seasoned data alone are employed in stability velocity; that is, no account is taken of a dynamic reaction to changes in circumstances in the application data. We state that the analysis of historical data may not be highly informative about the direction of future changes, for whereas history may never repeat, it certainly has a tendency to rhyme. Both of these issues lead to the conclusion that the present risk assessment literature is not well equipped to presage any movements in financial credit risk perception or to underpin an early borrower default warning system. Given these reasons, a call to action is needed for the deployment of modern statistical techniques to more adequately model these credit drivers.

2.2. Challenges and Limitations of Traditional Models

Traditional financial credit risk assessment models are more rigid and often detect credit risk intermittently. This is due to the slow response from such traditional models owing to weeks to months of updates. Such models are also often developed with very limited data. In the valuation arena, the model is highly affected in the fast market. Also, this data is taken from history, and there is no single model that can predict the future perfectly based on historical data. Even in the best of times, when economic conditions are good and borrowers behave as in the past, this is still a limitation. The characteristics of the loan can change; borrowers change as the economy's conditions and respond to taxes, employment, and inflation change. Consequently, the main concern is the sensitivity of the economic conditions associated with public loans or company financial statements obtained using historical data. That is why rating one-year-old data is a common complaint from loan officers. Many analytical rating systems have a difficult time capturing this type of effect.

The traditional models are often not developed to detect the new type of risk in credit and financial business. The current risk measurement models are limited when detecting and fully evaluating the output of technological innovation. Besides these standard risk conventional models, the intensive competition in the world is now crying for other creative ideas to challenge the traditional credit risk assessment models. To fulfill the criteria of this credit assessment, a more analytic-based approach driven by data is highly needed.

3. The Role of AI and Machine Learning in Credit Risk Assessment

Artificial intelligence (AI) and machine learning are marked for their high disruptive potential in credit scoring assessments. When a massive amount of data is analyzed and relies on already created patterns that can be used by the computer to predict outcomes, machine learning can automatically adjust or modify its algorithms. Successful big data analysis can lead to powerful predictions that the human mind cannot see using a traditional way of doing business. Risks that can be simply calculated manually are pointed out by machine learning and presented in a visually more compelling and accessible manner. In machine learning, data can be evaluated by several thousand units, where correlations start to follow patterns. It can help in identifying hidden or unhanded trends and can truly play a crucial role in finding a better solution in turning raw data into profit.

Machine learning-based algorithms can process and analyze millions of data points and find correlations that are often missed by hand. These machine learning algorithms can identify significant correlations, which can be excluded or reported in human analysis based on their great volumes. The more data sets are provided, the better these algorithms are, and they are capable of increasingly accurate predictions. Statistical credibility is distinct from the numerous data series processed by machine algorithms and is a determining factor that allows them to predict with astonishing accuracy. In doing so, they calculate complex models. Furthermore, these advanced technologies can predict future behavior and, if necessary, correct incorrect automated models in real time. Bias can also be minimized, which is usually present during traditional credit assessment operations and is often exaggerated or reduced by human prejudices.

3.1. Overview of AI and Machine Learning

This article discusses a variety of cutting-edge AI and machine learning technologies for different AI and ML interpretations, such as AI, machine learning, and deep learning. Here we show that AI is generally defined as the capacity of an artifact to carry out a task that would otherwise require human intelligence. Machine learning is a method that provides systems the capacity to learn and master data patterns over time. By offering devices exposure to data and helping them deduce for themselves what tasks they should accomplish and how to approach them, ML fosters more intelligent systems. Supervised learning, unsupervised learning, and reinforcement learning are the three primary types of ML. DL is the three-tiered neural network known as deep belief networks. Every higher layer extracts features from the lower ones to master fundamental concepts. This encapsulated database of features is then used to train a classifier and make predictions in the top layer.

AI is growing procedural capacity using analysis, which previously had been the domain of science. In the context of using artificial intelligence for credit risk assessment, two kinds of AI techniques are frequently mentioned: supervised learning and unsupervised learning. In building an ML model, data is essential. In specifying the result, it is necessary to implement ML strategies for credit risk assessments through a strong data set. The quality and quantity of training data help to produce a strong and highly precise model. AML requires an extensive training data set to establish a strong model, which can cover the majority of data patterns. Unlike traditional deterministic procedures that enable strategies to be developed based on hardcoded values, AML needs a different strategy. AML relies on large volumes of training data to deduce what a positive or negative score might look like. Unstructured space creates knowledge beyond human limitations. AI can actually handle vast extents of such data.

3.2. Benefits and Advantages in Credit Risk Assessment

Credit risk assessment conducted with the help of artificial intelligence offers a range of benefits. These benefits are discussed in the following paragraphs. Banks often make up for a lack of credit risk assessment accuracy by setting interest rates at a level that compensates for additional risks involved. Therefore, most banks have set this price, i.e., interest rate, higher than it would be based on the simple relationship between credit rating and default frequency. The AI-driven credit risk model can improve the prediction of interest rates and, as a consequence, improve banks' business through improved accuracy. Agility. Conventional credit risk models based on multiple logit or hazard models can only be updated once a year due to their complex nature. Credit risk models based on AI, or more specifically machine learning models, are less complex and can be recalibrated on a rolling basis. This means that a bank can internalize market signals on how credit risk changes in each loan segment. Given the diverse range of data that AI methods can process, it is easier to customize credit risk assessment. At its best, such customization means that banks, through AI, can offer each borrower a unique risk profile and price. Cost can be reduced through the automation of reporting and processing at the operations level – e.g., handling credit applications – and the process can be made more efficient. Some large banks are known to be working to develop loan pricing platforms based on AI. The advanced analytical aspects of this area will be deepened in theme 4. Technology will improve the credit risk component in the management and reporting on financial risk, thus fostering risk-based rules for reporting and portfolio strategy. In particular, AI technologies will facilitate risk management rules.

4. Key Machine Learning Algorithms for Credit Risk Assessment

In a logistic regression model, we define a linear combination of the input features and use a logistic function to transform the result between 0 and 1. This allows us to make good use of the tools of statistical inference to estimate the coefficients of our model. In risk analysis, this is often used as a risk score model, where loans with a high risk score are rejected. The main advantage of logistic regression is that the production of a prediction is quick compared to other methods, and it is easily interpretable. Decision trees and their more sophisticated extensions, random forests, help improve model accuracy, but at the cost of complex algorithmic models that are not as easy to interpret. The primary advantage of decision trees and random forests in credit risk is in describing simple patterns that are easy for practitioners and policymakers to understand, coupled with their ability to interpret the importance of different features in influencing a decision. Decision trees are able to divide the input space in a series of splits, leading to great flexibility in handling highly non-linear data. This creates a unique business opportunity when managing risk that depends on a specific combination of factors. Random forests essentially involve training many decision trees and using the average of multiple trees to make a prediction, which reduces the tendency to overfit with a single model and is called ensemble learning. Support vector machines are useful in high or very high dimensional spaces. With an increase in the number of dimensions, the performance of the model deteriorates and for this reason, it requires data preprocessing, for example by using PCA. Although EN and SVM are very popular, they are data-hungry algorithms that require significant preprocessing and regularization. The value of these methods, however, can be limited by this focus, as the true value of data science is not in designing an advanced model to meet a particular data set, but in the validation process and, when needed, to innovate the models themselves to anticipate rather than imitate risk. Thus, current methods reflect this data-centric approach. However, as the technique of using machine learning for risk modeling for credit scoring improves, the number of methods and the special cases described should increase as they are able to recognize patterns that previously were hidden.

4.1. Logistic Regression

Logistic regression is an extension of linear regression, where we transform the continuous response variable into a categorical one (e.g., 1 if the outcome has some characteristic such as 'default' or 'fraud', and 0 if it does not). In credit risk or other kinds of risk, the outcome is often binary, such as 1 for default and 0 for non-default. In this case, logistic regression can be used to model the probability for each individual using, for instance, the person's income and credit history. The mathematical form is logit(P(default)) = B0 + B1 * feature1 + B2 * feature2 + ... + Bm * featurem. Just like linear regression, these coefficients (denoted as 'B') are adjusted to generate the optimal outcome possible. Consequently, we could assess what factor is most important in predicting the outcome by comparing the estimated coefficients/weights.

Logistic regression has a lot of advantages. First, the model equation is very simple and easy to interpret, thus it can be readily implemented in any software. Second, it is particularly suitable for the binary situation. However, there are several assumptions of logistic regression. For instance, logistic regression assumes that there is a linear relationship between the dependent and one or more independent variables. This could be violated when the dependent variable does not have a common difference. In addition, logistic regression also assumes 'independent observations' to prevent influencing each other. Non-independence results in biased estimates. Logistic regression is widely used in the banking sector, not only for its traditional model interpretation quality and implementational ease, but also AI technology such as machine learning and deep learning, considering current industry trends.

4.2. Decision Trees and Random Forests

Decision trees have been one of the most frequently used algorithms for credit risk assessment in the banking sector. Decision trees are a tremendously popular approach in credit risk. The trees are extremely transparent and easily understandable to borrowers, thus satisfying regulatory requirements for meaningful disclosure of the risk factors influencing a credit decision or interest rate offer. By looking at all possible decision pathways and the decision variables sorted according to importance within the decision tree workflow, a decision tree can indicate the factors influencing an individual borrower's default risk and, thus, connect to the decision factors that are used in pricing and risk-based pricing functions within predictive models.

Random forests are the ensemble technique of decision trees. An ensemble technique simply infers multiple decision trees before combining their predictions that classify or associate an unseen data instance into a target class. Random forests, by building multiple decision trees and combining their predictions, have the effect of decreasing overfitting and smoothing the predictions. By comparing predictions that consist of several large ensembles of decision trees, which have been built on different subsets of the same data, the technique has the ability to create a standardized metric, or variable importance measure, which enables the data scientist or model developers to discern which of the dozens to hundreds of predictor variables that go into a random forest model are making the model deviate to a greater extent from its prediction. Random forests can be utilized in virtually any machine learning task that involves the use of decision trees. The advantage of decision trees and random forests is that they are much simpler to implement and quicker to run than neural networks or some traditional models, and are, therefore, easier to understand and to explain to business users and regulators who need everything within a bank's risk or compliance ecosystem to be transparent. Additionally, decision trees represent a bridge between static scoring models and advanced techniques, such as the use of binaries.

4.3. Support Vector Machines

Support Vector Machines (SVM) are another powerful technique to develop a credit risk assessment model. In SVM, the algorithm is generally used for separating hyperplanes, where the maximum margin between classes is the basic principle. The hyperplane optimally separates the data and captures the differences between classes. To handle real-world data, where every dataset is thought to have some kind of noise and error, it would be very difficult to use many potential hyperplanes. Hence, SVM is used to find a hyperplane with a maximum margin that robustly performs. In SVM, a hyperplane is used that is different from the classical linear regression models. It is a frontier that best segregates the different groups or classes, whereas the remaining data points are defined as support vectors, and the perpendicular distance calculated from the support vector to the frontier is referred to as the margin. This technique handles the curse of high dimensionality and is used when the financial dataset is expected to have complex behavior by finding the best hyperplane in the higher-dimensional space. Non-linear functions appear in most of the financial data, and the practical uses of SVM are very diverse. SVM is used to analyze and predict credit defaults and to effectively measure debtor type, compare different behaviors of borrowing, connect consumer integration and financial behavior, identify high- and low-risk financial investors, and correctly identify the borrowers who have the same creditworthiness. It is more stable and resistant to overfitting than traditional classification algorithms and easily suits the requirements of a modern portfolio management information evaluation system. However, training SVM on a large dataset becomes computationally very expensive. It also features algorithmic complexity accompanying operations related to matrices, coefficients, and training time.

5. Case Studies and Applications in Banking Systems

Several case studies and applications have showcased the enhancement of credit risk models with AI, focusing on machine learning models. This section demonstrates the practical application of machine learning in the banking system, considering the vertical that is closest to the guidelines for evaluating intelligent credit risk. It describes how six financial institutions have integrated machine learning into their credit decision process, from the earliest phases to the most mature deployment in the sector.

At the heart of AS24's real-time credit scoring model lies a powerful machine learning algorithm. Specifically, it is a random forest, which has demonstrated its predictive accuracy according to tests conducted. Delta Bank, together with data science and credit experts, developed an underwriting decision engine. The pilot has scored 1,000 applications with the new system, and the results demonstrated a clear improvement in performance for the AI-enhanced decision-making engine in terms of accuracy, operational efficiency, and adherence to regulatory requirements. To use this machine learning model, modifications were necessary in the data integration process, privacy rules, and team responsibilities for monitoring it. The learning curve of the organization as a whole can be considered one of the main difficulties faced in the implementation. After the implementation, some of the lessons learned were the need for continuous model maintenance, adjustments and improvements, and regular recalibration of risk strategies based on the data updated in the model.

Visa, one of the largest global payment networks, refers to different use cases where machine learning has been used for credit risk evaluation. For instance, the company describes an application that uses machine learning predictive engines to evaluate fraud risks related to payment authorizations using a worldwide database of transactions. The machine learning model has demonstrated its superiority in comparison with traditional models in terms of reducing the false positive rate. The company lists superior predictive capabilities of machine learning models, enhanced business rules, better insights and explainability, and speed of deployment as some of the advantages of using machine learning in the context of risk evaluation.

5.1. Real-world Examples of AI Implementation in Credit Risk Assessment

There are many real-world examples of AI implementation in credit risk assessment within banking systems. Projects span from global banks investing in modern infrastructure to start-up fintech services. Each of the examples provides a high-level overview of the AI design and describes the improvements or challenges from adopting the AI-improved credit risk model. Several banks across the globe are currently early in developing and deploying AI-enhanced designs and products to automate credit risk assessment. Each bank is developing proprietary products employing different AI algorithms in the credit risk assessment engines. In these examples of the best implementation of AI, we can see that AI outperforms the traditional model. Despite difficulties with the bank's adoption of AI in the credit risk concept, such as risk infrastructure, model governance, and investment, AI has brought a concrete outcome for credit risk prediction. Results of the AI implementation in the credit risk context show a significant improvement in the predictive power of the credit risk models. The basis of the model is a fully trained deep learning model. The AI-enhanced product reduces the default rates and increases the active users. Additionally, the bank is accelerating digitization through the approval of a large number of personalized offers. Despite difficulties with maintaining the volume of approval offers, the AI model concept brought a range of benefits for credit risk assessment.

6. Challenges and Ethical Considerations in AI-Enhanced Credit Risk Assessment

There are several challenges associated with AI-based credit risk assessments that extend well beyond technical considerations. One of the key challenges involves interpreting what constitutes fair and unbiased assessments. In particular, insipient biases in historical data can become reproduced and possibly aggravated by machine learning models. Another concern is that AI processes could black-box credit risk algorithms, potentially shifting the focus from transparency to accountability. Furthermore, the integration of AI in credit risk assessment raises important issues in relation to consumer privacy and data security, as well as the difficulty of assessing these privacy and data security issues. The widespread adoption of AI-enhanced decision-making processes is also restrained by several regulatory challenges and related ethical considerations that differ between national jurisdictions.

A related but broader sociotechnical challenge is ensuring public confidence in the fairness of automated decision-making processes. This issue is thrust into the spotlight when the outcome of automated decision algorithms severely impacts human life, as in healthcare or criminal justice decision-making. Decisions about lending are often as consequential as those made by judges, especially for people living below the poverty line. Accounts of people whose loans were declined by AI algorithms emphasize the need to consider how judgments are made and who is held accountable when assumptions underpinning such practices fail. These concerns motivate recommendations stipulating that AI decision-making processes should be interpretable for human decision-makers and the affected individuals, who ideally should be involved in validating these systems. It is recommended that if financial firms intend to use AI-enhanced methods, they should apprise the regulator of this decision and demonstrate the technical abilities of their model to assess the distributional impacts of possible gaps. In summary, it is important to balance the opportunities offered by the AI-based tools against the ethical, regulatory, and technical challenges in this field.

7. Future Direction

6. FUTURE DIRECTION

This paper examined and critically discussed the broad and complex range of issues associated with integrating AI technologies into credit risk assessment. These two technologies are harbingers of the next generation of AI and machine learning risk assessment models. As the technological opportunities unfold, there are some important emerging trends. Over time, access to new data sources will continue to develop further commercially valuable insights about consumers. We also anticipate considerable growth in the capability to support these technologies with enhanced data analytics, especially in the combination of paired and unpaired causal inference models. We may also see increased emphasis on system architectures that address vulnerabilities. Building ethical AI for credit scoring exposed by adversarial attacks, for example, by keeping discerning features hidden.

Banks and other providers also have to provide explanations for decision-making logic, controls on bias, and importantly, monitoring for potential harm. Alongside these developments, we expect to see the rise of AI and the development of associated regulations and frameworks to guide its use. A significant majority of stakeholders are of the firm view that AI-specific regulations and guidelines are currently fragmented and piecemeal. In turn, banks, the major generators and deployers of AI, require guide rails for normative guidance. Ethical AI in credit scoring can reduce the risk of model misuse. Banks are, of course, also key to furthering the development of solutions by AI technologists. We anticipate the potential for banks to collaborate with technology providers to roll out these solutions in production.

As a range of issues are worked out, including additional financial, privacy, and regulatory concerns, we anticipate that many of the most powerful techniques may remain classified, proprietary AI and machine learning secrets. An additional frontier in this debate concerns ethical considerations. Ethical considerations in data use and analysis have long been a key concern of information technologists, especially where systems have the ability to make judgments and infer privacy-sensitive information about users. Approaches and frameworks for evaluating existing models for fairness, transparency, and accountability exist and are increasingly being integrated into commercial systems. These requirements are particularly important in the development and deployment of AI in customer-facing settings such as credit risk assessment. Many of these methods have good cause to shift to make way for new, more intelligent, and socially equitable credit assessment systems. In general terms, it appears available technology is providing us with many of the tools that may well be necessary to transition to such a paradigm in the—albeit distant—future. Firms looking to integrate new models and populations of training data might like to consider as a new technique, relying on Generative Adversarial Networks to investigate whether the properties of the training population could be improved to more clearly meet wider cross-sectional definitions of fairness. Finding areas for general agreement between the performance of AI and an international consensus on human fairness is, as they say, still a work in progress. The world awaits.

8. Conclusion

Credit risk models are leveraged to optimize bank operations and to make potentially critical decisions about consumers. Consequently, it is crucial to evolve credit risk models, adapting them to modern banking challenges. Using a traditional, unsegmented approach and historical data, financial institutions can no longer address the growing number of applicants. As long as they keep this method up, they will grant too good a rate to certain borrowers and too high a rate to certain others. The introduction of AI can bring a lot to a credit risk assessment model. It enables, on one hand, a major gain in terms of predicting accuracy and, on the other hand, an agile and efficient evaluation of a borrower's risk in real-time. Though AI is a revolution, integrating it into a financial institution requires a perfect treatment of the issues. Indeed, one should only consider that models are good and fair as long as they are perfectly transparent and their predictions are perfectly interpretable by humans to avoid any act of discrimination. This is a first analytical exploration of the use of AI in the sector of credit risk assessment. This introductory paper aims to open up new areas of investigation. However, a plethora of potential academic investigations can stem from it. It would be relevant to investigate actual bank practices in integrating AI in the risk analysis of companies, small businesses, and consumers. Given the intimate connection between banking and digital technology, it would be interesting to question the social acceptability of these methods in the population. In particular, it would also be interesting to study the relationship between ethics and credit risk assessment practices in the future. This research can indeed be located at the crossing between technology and ethics. The profile of risk assessment officers is therefore bound to change. To conclude, it is clear that AI is revolutionizing the standards in credit risk assessment.