JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

MACHINE LEARNING IN CREDIT RISK ASSESSMENT: ANALYZING HOW MACHINE LEARNING MODELS ARE TRANSFORMING THE ASSESSMENT OF CREDIT RISK FOR LOANS AND CREDIT CARDS

Teja Reddy Gatla

Associate Director at DTCC, AI Expertise, Department of Information Technology

***ABSTRACT*—The main purpose of this paper is to review the use of machine learning in credit risk assessment, focusing on the data pre-processing aspects which are often the most time-consuming activities in the development of machine learning models. The report begins with an overview of machine learning methods, assessing their advantages and disadvantages compared to traditional credit scoring, such as the Statistical Discrimination Method. An investigation of how the choice of performance measure used in credit scoring can influence the model that is selected for classification. Model settings and the issues regarding classification of the response variable are discussed, followed by an example of a test/train data split [1]. The main body of this document presents a demonstration of data pre-processing techniques and an evaluation of their effects on classification accuracy. Effects on model assessments and comparisons are discussed throughout. An example of variable selection method is investigated, demonstrating the accessibility to high-level statistical tools due to inter-platform data compatibility. Finally, we discuss the future use of predictive models that are automatically implemented using model classification strategies[1]. Variable coding is an important step of pre-processing which this report does not cover due to keeping to the constraints of the main objective of developing classification models. Machine learning is a technique of data analysis that makes automated model building by employing different algorithms. It is a subfield of artificial intelligence that involves some systems learning from data, since no humans are necessary for providing help on the most important decisions [1,2]. Over the last few years, banks and other financial organizations have been pushing harder to use machine learning to improve their credit risk models. One of the most attractive things about using machine learning is that it can help make better credit decisions in a shorter amount of time, which can be critical in an industry where being the first to make a decision on a potential customer can be the difference between millions of pounds [2]. Additionally, predictive models can help automatically implement the best strategies to consumers, again saving time and resources for the organization. This has led to great interest in developing and testing new machine learning models to be used as benchmarks for more traditional credit risk models.**

***Keywords*— Machine learning, Credit, Finance, Loans, Credit risk, Automation, Model building, Classification, Statistics, banks, Deby, Workflow SOAR platforms, Predictive models**

INTRODUCTION

Credit scoring and decision is an imperfect and a priori classification problem. Accurate classification of good and bad credit risks is of great practical importance as this is the main object of credit scoring algorithms. The costs of misclassifications are not symmetric here which means in resolving this issue one has to focus on both sides of the problem. The cost paid for mislabeling a wrong customer as a good one is higher compared to the price incurred for misclassifying the good customer as a bad one. Poor customer classification can easily lead to lending to a bad borrower who can't pay the loan [3,4]. This cost includes the loss of the loan amount and interest on the loan, as well as the cost of the debt recovery process. The necessity of incurring this cost is thus to be contrasted with the damage caused by failing to give a credit to a good client, which is merely interest lost by the bank on the amount of loan. Since the implications of the two types of misclassifications are unequal, asymmetric cost theory has been proposed in statistics [5]. The p-value method does this by comparing the chance of a bad customer's classification to the rate of acceptable levels of bad customer identification, for example, P0. A test is then set to have a type I error rate (classifying as a bad customer a real good customer) but being accepted if the type II error rate (giving as a good customer a bad one) is at a level lower than a specific [6] value. This approach is acceptable and beneficial, although not flexible as it does not reflect credit situations with continuous decisions from very bad ones to very good ones.

Credit risk assessment lies at the core of a bank's loan-granting decision. It is therefore not surprising that creating economic systems aimed at improving credit risk assessment is an important topic in financial research [5]. The fair treatment of credit applicants and the increased marketing efficiency of the lender are the primary economic motivations behind improved credit risk assessment technology. However, new risk assessment technologies must be thoroughly analyzed for their impact on the distribution of profits and risks between the lender and the applicants. This is a widely recognized difficult problem and there is no one answer on what the "best" credit risk assessment technology is.

What credit scoring has done is provide a relative measure of the risk of lending to specific demographic at a point in time. We expect any new technology to provide a relative risk measure superior to that of credit scoring so that profitability from lending to different demographic groups is not adversely affected. With advancements in machine learning, there are numerous possibilities when it comes to enhancing credit risk models and the data analyzed.

Machine learning, with its inherent complexity, proves to be of significant benefit in the realm of credit risk assessment [5,6]. It provides a flexible environment to build robust credit risk models, covering a wide array of machine learning approaches. From the simplicity of decision trees to the intricacy of support vector machines, these methods offer immense potential for accurately classifying and reclassifying potential or existing borrowers. By harnessing the power of machine learning, financial institutions can revolutionize the credit risk assessment process.

II. RESEARCH PROBLEM

The main research problem addressed in this study was to assess the intrigues of machine learning in credit assessment to understand how it analyzes credit risk and loan approval. Similarly, this research will focus on credit scoring within the consumer lending industry. The consumer lending industry has been rapidly increasing in the use of machine learning algorithms to automate the loan approval process. The increased efficiency of the process has led to increased profits. Here, credit scoring is the assigning of a class label (good, bad) to a consumer based on his credit worthiness as determined by his credit report. The objective is to approve the loans of the good creditworthy consumers that will be repaid and reject the loans of the bad creditworthy consumers that will default. The benefit is maximized when a lending institution approves all of the good creditworthy consumers and minimizes the approval of bad creditworthy consumers [6]. Approving a loan for a bad creditworthy consumer results in a financial loss to the lending institution in the form of a defaulted loan. The two types of losses from the lender's perspective can be defined in terms of the confusion matrix as: False Negative (Approve, when we should not) – approving a loan of a bad credit consumer False Positive (Reject, when we should have approved) – rejecting a loan of a good credit consumer.

III. LITERATURE REVIEW

A. Overview of Machine Learning in Credit Risk Assessment

Machine learning provides a new tool for improving the predictive accuracy of credit scoring systems. The prediction of credit risk is a very attractive area for the application of machine learning technology for a number of reasons. Traditional credit scoring systems have been based on linear discriminant analysis[7], though logistic regression is more commonly used today. These methods have limitations in their predictive ability. The primary disadvantage of these techniques is the linearity of the model assumed. It is now widely accepted that the relationship between the attributes of an instance and its class are often non-linear and that non-linear models are more realistic and accurate. In addition, complex interactions between the influences of different attributes on the probability of the credit event occurring are almost impossible to specify and measure using human judgment [8]. Machine learning techniques can automatically detect these interactions and are capable of fitting highly non-linear models. High predictive accuracy is the ultimate goal of credit risk assessment, and it is here that machine learning can make a real contribution to the credit industry. Another aim of credit risk assessment is to separate good from bad credit applicants. Traditional credit scoring systems are relatively poor at this task, mainly because of a class imbalance in the instances used to train the model. The class to be predicted is the credit event occurring (default, bankruptcy) with the instance being an applicant for credit [9]. Default is usually a rare event and in the case of a specific company or portfolio of credit applicants, there will be very few bad instances relative to good ones. In this context, the data used to train the credit scoring model will be unbalanced and the model produced will be biased towards the majority class, leading to an underprediction of the minority class. While a lot of machine learning happens as an exhibition of the ability of these automated systems to learn from imbalanced data, this is not a particular flaw of machine learning algorithms. Many have built-in methods for handling skewed class distributions and various techniques can be used to resample the data to produce a balanced set and train a model that can give good predictions for both classes. The ability to accurately predict the minority class is especially important for credit risk assessment, as the aim is to avoid bad instances. High false negative rates are very costly in this domain.

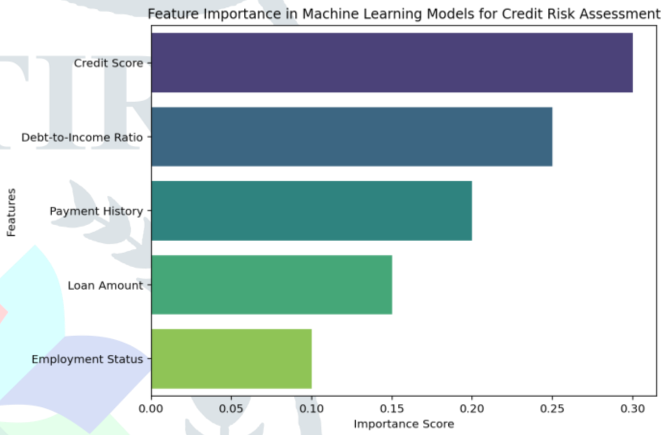

Fig. 1 Feature Importance in Machine Learning Models for Credit Risk Assessment

B. Benefits of Machine Learning Models in Credit Risk Assessment

High predictive accuracy is essential to a credit decision, as both Type I and Type II errors are costly. This is also a paramount goal in statistical discrimination, to accurately identify a group or individual to offer or restrict opportunities, resources or life chances [10] This will aid in the continual progress to alleviate adverse selection processes in credit lending. A study comparing the two methods by using the same data inputs into two models. The neural network, which acts as a universal function approximator of an equation or decisions, showed empirical evidence of better predictive accuracy in both identifying good credit and bad credit applicants [10]. This was shown by an increase in the net profit for the credit grantor. The first benefit, increased predictive accuracy, directly addresses the current limitations exhibited by statistical modeling techniques [10]. Empirical studies on the usage of neural networks (one type of machine learning algorithm) have shown significant evidence of increases in predictive accuracy over traditional linear regression models[11].

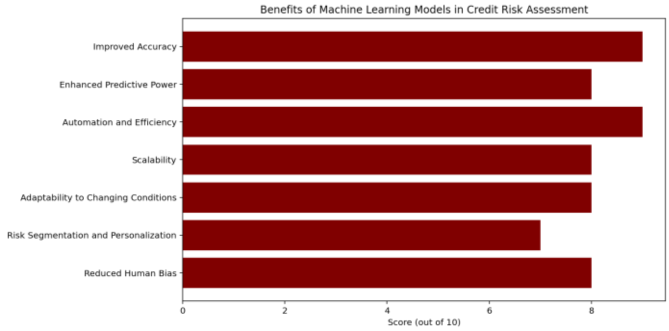

Fig. 2 Benefits of Machine Learning Models in Credit Risk Assessment

As shown in the figure above, the benefits of machine learning in credit risk assessment can be divided into seven categories. These are: increased predictive accuracy, increased model transparency, feature automation, improved model robustness, time and cost efficiency, enhanced decision making, and profitable innovation.

C. Challenges and Limitations of Machine Learning in Credit Risk Assessment

Credit scoring is a well-developed application area in finance with a good number of theoretical and empirical results. Many researchers have demonstrated that the predictive accuracy of the statistical credit scoring can be superior to that of the traditional methods of credit scoring such as expert judgment or the simple yes/no decision rules. In fact, statistical credit scoring uses a model built on an applicant's credit file and additional information that may be gathered at the time of application to rank-order the applicants by their likelihood to repay the credit obligation. This ranking helps decision makers to process the credit applications more efficiently. If the rank ordering can be ordered in increasing or decreasing order of risk, credit grantors can cut off the sample at any point, accepting the applicants above that point and rejecting the applicants below that point. This is very useful in situations where the credit grantor has constraints on the number of applicants to approve. Finally, the credit scoring models can be used to monitor the existing customers' credit behavior and can be used to predict changes in their credit risk.

D. Current Trends and Developments in Machine Learning for Credit Risk Assessment

AI has revolutionized the security systems with SOAR platforms and automatic machines being the part of the system. SOAR adopts machine learning and natural language processing and represents a fully automated incident response workflow, so the security operations flow seamlessly, and threat detection and response times are enhanced considerably. For instance, the SOAR research of Cisco Systems are evident that AI can prioritize concerns from various sources by automating type of regular people’s work. Even more, intelligent self-operated security systems which AI enables have the function of analyzing, spotting, and reacting for the purpose of cybersecurity threats in real-time reduce the requirement of human intervention and eliminate the shortcoming brought by human error to have better efficacious cybersecurity. For instance, the researches conducted by [12] aim at the AI robots' surveillance capacity in mitigating cyber risk as well as creating resilience and efficiency.

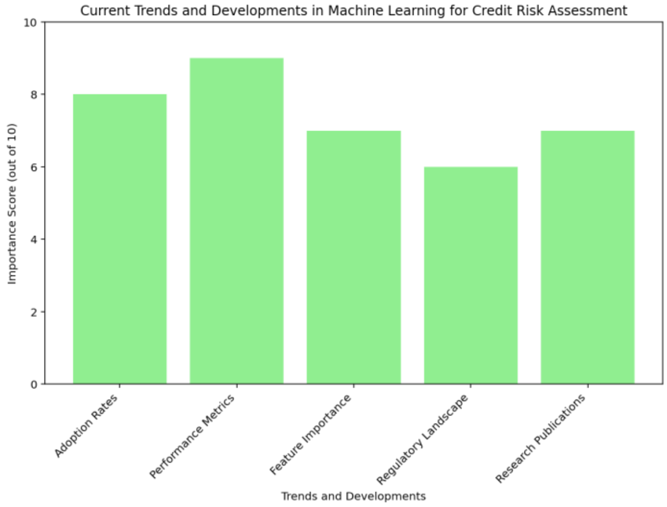

Fig. 3 Current Trends and Developments in Machine Learning for Credit Risk Assessment

E. Economic Impact of Improved Credit Risk Assessment

A failure to assess credit risk accurately translates to pricing that does not correspond to the risk, adverse selection, and moral hazard. Each of these can reduce the volume of credit extended and can lead to an adverse selection death spiral. Reduced credit volume in turn hampers economic growth and stability. Considering the recent subprime mortgage crisis, the combination of adverse selection along with the securitization of poorly understood mortgage-backed securities amplified the severity of the recession and led to a freezing of credit availability [12,13]. On the micro level, adverse selection and moral hazard cause creditors to lose confidence in the quality of the debt or the counterparty, often leading to preventative measures in the form of hard to overcome terms and conditions or increased collateral requirements. These too can be detrimental to economic growth by reducing capital expenditures and increasing the cost of consumption smoothing. Finally, the increased accuracy of risk assessment can facilitate unsecured lending to low-risk borrowers. Since unsecured loans are often the most efficient means of consumption smoothing, facilitating these transactions can have a large impact on the welfare of the borrowers.

F. Reduction in Default Rates and Loan Losses

There have been three separate studies to analyze the effects of traditional credit scoring and the added predictive power of data mining. The first study can be found in a white paper by FICO. In the paper, two-thirds of the sample use the FICO score in the origination process, and the other third had a series of random application processes (Image already assessed by the time you read this paper). The sample findings showed that for all the new models used on the random sample, there was a 15% increase in credit to the high-risk population. This increased revenue was seemingly good; however, the new models were also decreasing the performance of the low-risk population [13]. The default rate increased for the low-risk groups, and the population with the best scores had an increase in the default rate from 3.9% to 6.5 percent. The second and third studies have been conducted by German research teams. [14] studied 12,000 German individuals over a period of 5 years. The data was presented to Risk Managers, and they were told to accept the population with the best credit scores within the new sample. The results of both studies found that in comparison to the old methods, there was an increase in credit to the high-risk population and higher default rates. While the objective of increasing revenue with a higher credit issuance to high-risk individuals has been achieved, the group most affected has been the individuals who initially had bad credit ratings. High-income, high-risk individuals ended up paying more for credit, and the group of poor individuals with a good credit history who wanted to re-establish their creditworthiness were being turned down for credit.

G. Enhanced Financial Inclusion and Access to Credit

Unparalleled to traditional credit risk assessment methods, innovative usage of ML doesn't only disrupt the traditional approach but also extend its benefits to previously underserved or unserved population segments. These benefits often accrue through automatic, data-driven decisioning to provide more consistent, less biased access to credit in comparison with decisions made by humans. It is fruitful to highlight a few specific scenarios in regard. Many individuals and small business owners that lack a credit history or have credit histories that are stale (i.e., outdated or limited recent information) are often at a huge disadvantage. In developed countries, these individuals are frequently offered credit with higher interest rates, because risk-based pricing strategies in the credit industry often result in those without recent credit information being penalized with higher risk assessments [15]. High risk assessments are translated to higher interest rates or rejections on the credit applications. In the worst-case scenarios, the poorest credit risks may be offered credit at such high interest rates that it is usurious in nature, or they may be offered inappropriately small sized credit products because risk assessment of these types of credit products is often insufficient. In the countries with emerging credit systems, large segments of the population may be unable to obtain credit of any kind. Credit availability is a key driver to economic growth and poverty reduction, but traditional manual underwriting processes are often prohibitively expensive relative to the size of the credit transactions being contemplated. In all of these scenarios, the problem is essentially that the individuals seeking credit are not able to qualify for credit products at rates commensurate with their actual underlying credit risk[16]. ML offers to solve this problem by building more predictive models using alternative data, or in some cases using no data at all on the actual credit outcome but basing the credit decision on a comparison of the credit applicant to past and present credit applicants with known credit outcomes. Alternative data can be used to assess credit risk on individuals without traditional credit data and provide a few consumers with an initial credit history a chance to qualify for a better priced credit product. Learning from past credit decisions can more optimally match credit products and pricing to the underlying credit risk, and automatic decisioning greatly reduces the cost of credit for small sized transactions.

IV. SIGNIFICANCE AND BENEFITS

By employing machine learning in credit risk assessment, the list of potential benefits is extensive. Reduction of human biases, minimal data requirements, increased speed and efficiency, ability to generalize and the possibility of detecting complex interactions and non-linear impacts of variables are all items in the wish list of a modeler. However, it can be noted that many of these benefits are potential and are not yet realized. For example, although algorithms such as neural networks and genetic algorithms are capable of realizing complex interactions between variables, in reality this feature may not have a practical advantage over a simpler piecewise linear model since the increased complexity may be difficult to interpret and rationalize [17]. It can also be argued that complex interactions, though accurately modeled, do not necessarily need to be detected if they do not impact the variable being predicted. Although the importance of building a model that most accurately resembles reality should not be understated, the pragmatic aim in credit risk assessment is usually to maximize predictive accuracy of the variable being modeled, for example default probability of a loan, subject to practical and statistical constraints. Despite this, the ability to tackle complex interactions remains an attractive feature of machine learning [18]. A related benefit is that models can be updated more easily to reflect a changing environment (e.g. economy) through simply using more recent data, which is an important consideration in a dynamic and evolving industry such as credit. This stands in contrast to the drawn out and sometimes unsuccessful process of redeveloping traditional models to make them more accurate and/or parsimonious.

V. FUTURE ENHANCEMENTS

The future implications and significance of implementing machine learning within credit risk assessment in the United States can be depicted as having an overwhelmingly enormous effect on various relevant parties involved here. Not only will it revolutionize the way credit risk is evaluated, but it will also bring about numerous benefits that cannot be ignored. Primarily, the deployment of machine learning would undeniably cause a diminished number of credit defaulters, leading to a more stable financial system. By employing this new technology, financial institutions would be able to sift out the "good" applicants from the "bad" with much higher accuracy. This would result in a significant reduction in the number of defaults and an overall improvement in the quality of credit portfolios [19]. Credit scoring would become a much easier and cheaper process due to the automation of machine learning. Machine learning algorithms have the capacity to process heavy data at high numbers and efficiency. They can, therefore, generate predictive models just by feeding them with relevant data. This would eliminate the need for manual analysis and subjective decision-making, ultimately saving financial institutions substantial time and resources. Although it may require a hefty amount of cost to begin with in order to configure all systems to machine learning, the return on investment would undoubtedly be positive in the long run. The initial expenses would be outweighed by the enhanced accuracy and efficiency of credit risk assessment, resulting in better-quality loan portfolios and reduced losses for financial institutions. This, in turn, would create a more stable and sustainable lending environment.

As for the debtors, these first two points may not be seen in a positive light initially. However, the implementation of machine learning opens many doors in regards to credit risk assessment for those that may have had a financial mishap in the past. With a traditional statistical model, it is extremely difficult to cater a special predictive model for those that do not have consistent data. People who have migrated from overseas, or students that have just started working will have no credit history to speak of [20]. This would result in an automatic rejection as lack of credit history is associated with high risk. Machine learning would make it easier for these people to acquire lower interest rates on loans, as their creditworthiness could be assessed based on alternative data sources such as educational background, income potential, and payment behavior in other areas of their life.

VI. CONCLUSION

This paper aims to explore various machine learning methods and shed light on their potential adverse impacts upon borrowers. While the use of advanced technologies undoubtedly brings numerous benefits to credit risk assessment, it is crucial to consider the possible consequences for the individuals seeking credit. Striking a balance between accurate risk assessment and fair treatment should be a focal point in the implementation of these methods. Analyzing the potential drawbacks will enable stakeholders to make informed decisions and address any unintended consequences that may arise from the use of machine learning in credit risk assessment..It can be conclusively shown that machine learning can be used to solve parts of the credit risk assessment problem much more accurately than the current methods which are in place by banks. This increase in accuracy can provide banks with better tools for making decisions on whether to lend to potential clients. With the classification and clustering methods, machine learning provides a continuous scale which can be used in client profiling. This enables banks to reduce their exposure to those clients who are in the higher risk brackets. Simulation of the outcomes from lending activities provides for assessing the risk of new clients, without affecting the current level of risk that the bank's customers pose. All of these methods provide much better decision-making tools than currently used discriminant analysis methods and clearly define the potential clients and the risk associated with each of them.

REFERENCES

[1] N. Kagiyama, S. Shrestha, P. D. Farjo, and P. P. Sengupta, "Artificial Intelligence: Practical Primer for Clinical Research in Cardiovascular Disease," *Journal of the American Heart Association*, vol. 8, no. 17, Sep. 2019, doi: https://doi.org/10.1161/jaha.119.012788. Available: https://www.ahajournals.org/doi/10.1161/JAHA.119.012788. [Accessed: Dec. 14, 2019]

[2] D. Gefen, "E-commerce: the role of familiarity and trust," *Omega*, vol. 28, no. 6, pp. 725–737, Dec. 2000, doi: https://doi.org/10.1016/s0305-0483(00)00021-9

[3] J. B. Caouette, E. I. Altman, and P. Narayanan, *Managing credit risk : the next great financial challenge*. New York: Wiley, 1998.

[4] R. A. Ippolito, "Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry," *The Journal of Law and Economics*, vol. 35, no. 1, pp. 45–70, Apr. 1992, doi: https://doi.org/10.1086/467244

[5] R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning for Network Intrusion Detection," *2010 IEEE Symposium on Security and Privacy*, 2010, doi: https://doi.org/10.1109/sp.2010.25. Available: https://ieeexplore.ieee.org/abstract/document/5504793/. [Accessed: Dec. 06, 2019]

[6] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine Learning Interpretability: A Survey on Methods and Metrics," *Electronics*, vol. 8, no. 8, p. 832, Jul. 2019, doi: https://doi.org/10.3390/electronics8080832

[7] S. Lessmann, B. Baesens, C. Mues, and S. Pietsch, "Benchmarking Classification Models for Software Defect Prediction: A Proposed Framework and Novel Findings," *IEEE Transactions on Software Engineering*, vol. 34, no. 4, pp. 485–496, Jul. 2008, doi: https://doi.org/10.1109/tse.2008.35

[8] C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 31, pp. 685–695, Apr. 2021, doi: https://doi.org/10.1007/s12525-021-00475-2. Available: https://link.springer.com/article/10.1007%2Fs12525-021-00475-2

[9] M. Bazarbash, *FinTech in Financial Inclusion: Machine Learning Applications in Assessing Credit Risk*. International Monetary Fund, 2019.

[10] L. C. Thomas, *Consumer credit models : pricing, profit, and portfolios*. Oxford ; New York: Oxford University Press, 2009.

[11] F. Bergadano and Luc De Raedt, *Machine learning : ECML-94 : European Conference on Machine Learning, Catania, Italy, April 6-8, 1994 : proceedings*. Berlin ; New York: Springer-Verlag, 1994.

[12] T. Caelli and W. F. Bischof, *Machine learning and image interpretation*. New York: Springer Science+Business Media, Llc, 2013.

[13] I. H. Witten and E. Frank, *Data Mining : Practical Machine Learning Tools And Techniques With Java Implementations*. San Francisco, Calif.: Morgan Kaufmann, 2000.

[14] IvezićZ., A. Connolly, J. T. Vanderplas, and A. Gray, *Statistics, Data Mining, And Machine Learning In Astronomy : A Practical Python Guide For The Analysis Of Survey Data*. Princeton, N.J.: Princeton University Press, 2014.

[15] M. Doumpos, C. Lemonakis, D. Niklis, and C. Zopounidis, *Analytical Techniques in the Assessment of Credit Risk : an Overview of Methodologies and Applications*. Cham: Springer International Publishing : Imprint : Springer, 2019.

[16] B. Baesens, D. Roesch, and H. Scheule, *Credit risk analytics : measurement techniques, applications, and examples in SAS*. Hoboken, New Jersey: Wiley, 2016.

[17] T. V. Gestel and B. Baesens, *Credit risk management : basic concepts : financial risk components, rating analysis, models, economic and regulatory capital*. Oxford ; New York: Oxford University Press, 2009.

[18] C. Joseph, *Advanced credit risk analysis and management : an indispensable book on credit and finance for better risk management*. Chichester, West Sussex: John Wiley & Sons Inc., 2013.

[19] J. Callaghan, A. Murphy, and H. Qian, *First International Conference on Credit Analysis and Risk Management*. Newcastle upon Tyne: Cambridge Scholars, 2012.

[20] S. Chishti, I. Bartoletti, A. Leslie, and MillieS. M., *The AI Book : The Artificial Intelligence Handbook for Investors, Entrepreneurs and FinTech Visionaries*. Newark: John Wiley & Sons, Incorporated, 2020.