sciendo

Modern Approaches in Credit Risk Modeling: A Literature Review

Luana Cristina ROGOJAN
Bucharest University of Economic Studies, Bucharest, Romania
Institute for Economic Forecasting, Bucharest, Romania
rogojanluana17@stud.ase.ro

Andreea Elena CROICU
Bucharest University of Economic Studies, Bucharest, Romania
Institute for Economic Forecasting, Bucharest, Romania
croicuandreea17@stud.ase.ro

Laura Andreea IANCU
Bucharest University of Economic Studies, Bucharest, Romania
Institute for Economic Forecasting, Bucharest, Romania
lauraandreeaiancu@yahoo.com

Abstract. *In the financial industry, models are pervasive, and their quantity and complexity continue to increase. Constant advancements are made in econometric and statistical theory, but a fast-developing body of rules and regulations governing their use needs modeling specialists to remain vigilant and adaptable. The tendency of these regulations to be ambiguous necessitates that industry professionals and institutions interpret them independently and jointly. This leads in what is referred to as a "industry standard," or a set of procedures that are recognized among modeling professionals but not necessarily to those outside of the industry. Non-practitioners in the industry may view the modeling department as a "black box" for these reasons. The accurate evaluation of financial credit risk and the forecasting of bankruptcy are crucial to both the economy and society. In recent years, more and more approaches and algorithms have been advanced for this purpose. At this point, it is of the highest concern to investigate the current credit risk assessment methods. In this paper, we review the traditional statistical models and cutting-edge intelligent methods for forecasting financial distress, with a focus on the greatest advances in the academic literature, as the promising trend in this field. Lastly, the paper will conclude with an overview of the evolution of methodologies and conceptual frameworks in credit risk management research, as well as possible future research directions.*

Keywords: credit risk modeling, credit risk, machine learning models, traditional models, bibliometric analysis.

Introduction

The financial credit risk denotes the risk linked with financing, i.e., a borrower is unable to repay the lenders or defaults. Anderson (2013) defines credit risk as “the probability that a legally enforceable contract may become worthless (or at least substantially reduced in value) because the counterparty defaults and goes out of business.” According to Saunders and Cornett (2011), it is the “risk that the promised cash flows from loans and securities held by financial institutions may not be paid in full.” Consequently, the objective of financial credit risk assessment is to resolve the following issue: certain a number of companies identified as having bad/good credit or bankrupt/healthy, as well as a set of financial factors that define the situation of a company over a given period, forecast the probability that the company will belong to a high-risk category or go

DOI: 10.2478/picbe-2023-0145
© 2023 L. C. Rogojan; A. E. Croicu; L. A. Iancu, published by Sciendo.
This work is licensed under the Creative Commons Attribution 4.0 License.

bankrupt in the subsequent years. The first issue is known as credit rating or scoring, while the second is known as bankruptcy (failure) prediction or corporate financial difficulty prognosis. Both are solved in a manner comparable to a binary classification task.

In general, credit risk assessment models can be categorized as qualitative or quantitative (Saunders and Cornett, 2011). When evaluating credit risk, qualitative considerations primarily focus on borrower characteristics, such as financial leverage, track record, earnings volatility, collateral, and market conditions like level of interest rates, business cycle, etc. Depending on these variables, a subjective judgment on whether or not to offer credit to an applicant is made. Quantitative models attempt to get a credit score, which is then used to categorize borrowers into different default risk groups or to estimate the likelihood of default (Saunders and Cornett, 2011). Due to the rarity of default risk, it is hard to model credit risk statistically (Anson, Fabozzi, and Choudhry, 2000). Nevertheless, numerous quantitative credit risk models have emerged over time since the concept of credit risk first surfaced. Like other models, they frequently draw on a variety of conceptual frameworks, such as ratios, option theory, econometrics, or expert systems (Caouette, Altman, and Narayanan, 2008).

Historically, traditional statistical methods were mainly used to forecast credit risk. However, these models continue to be inefficient because they depend on a number of restrictive assumptions, such as a large sample size, normally distributed independent variables, and a linear relationship between all variables. To far, a wide number of approaches have been used to analyze the creditworthiness of applicants using either conventional statistical methods or advanced machine learning methods.

The rest of the paper is organized as follows: the literature review of the traditional and modern approaches in credit risk modeling will be discussed in section 2. The third section highlights the data and methodology used for conducting the bibliometric analysis. In section 4 we summarized the results of the bibliometric analysis and finally, in section 5 is dedicated to the conclusions.

Literature review

Advancement of traditional credit risk modeling

Credit risk is a fundamental type of risk in finance and is commonly recognized as the earliest type of risk in financial markets, heading back to 1800 B.C. (Caouette, Altman, and Narayanan, 2008). FitzPatrick (1932) and the well-known Altman models (1968) were the pioneers in the field of financial credit risk assessment. According to Zamore et al. (2018) the Altman's Z score, based on multivariate discriminant analysis of five financial indicators, is the first quantitative credit risk model in use nowadays. In accordance with Benzschawel (2012), despite being 50 years old, the Z-score remains a valuable tool for many market participants. The methodology is criticized, however, among other things, for being inconsistent and backward-looking because accounting ratios are dependent on historical data. Other credit risk models, like the structural and reduced-form models, have emerged as a result of this. The writings of Black and Scholes (1972, 1973) and Merton (1974) are the earliest sources of structural credit risk models. Conforming to structural models, which are based on the capital structure theory of Modigliani and Miller (1958), a default happens when a firm's assets are worth less than its debt. In instance, Black and Scholes (1973) show that call options on equity have some effects on debt value by using the options pricing model to price debt and equity. The issue with the Black-Scholes model is that firm asset values cannot be observed immediately. Building on this research, Merton (1974) demonstrates that, under certain conditions, it is possible to calculate the asset value and use the results to calculate the

distance to default, which he refers to as the likelihood of default. The Merton's model is currently the most significant model in credit risk modeling, at least in terms of structural models.

Reduced-form models, in contrast to structural models, can calculate the probability of default without assuming anything about the origin of the credit risk premium (Benzschawel 2012). The theory of risk-neutral pricing, which is the foundation for reduced-form models, states that the market value of a hazardous investment is equal to the present value of future cash flows discounted at the risk-free rate. Among other authors, Jarrow and Turnbull (1995) and Duffie and Singleton (1999) made significant contributions to the application of the risk-neutral pricing theory to the modeling of credit risk. In credit risk modeling, the reduced-form method has also been a widely accepted approach (Weigel and Gemmill, 2006).

Machine learning – New approaches in credit risk modeling

Machine learning is a data analysis technique and component of artificial intelligence based on the idea that a system can learn from previous data, recognize patterns or distributions in datasets, and make decisions. In general, it is a system that simplifies the creation of analytical models with minimal assistance.

The overwhelming evidence from multiple recent studies has demonstrated that intelligent approaches without the use of constrictive assumptions can significantly increase the accuracy of statistical methods. In comparison with other statistical techniques mentioned in the paper, algorithms that use machine learning as a main stream of classifiers are also widely used in consumer and corporate credit modeling. These methods include, among others, artificial neural networks (ANNs), support vector machines (SVMs), K-Nearest Neighbor classifier (KNN), case-based reasoning (CBR), decision trees (DTs), Random Forest (RF), genetic programming (GP) and hybrid learning. Crook, Edelman, and Thomas (2007) and Ravi (2007) provide a full analysis of Neural Networks and Genetic Algorithm and their extensions, respectively.

Artificial Neural Networks

Artificial neural networks denote the family of neural networks with various structures, such as multi-layer perceptron (MLP) or backpropagation neural network (BPNN), self-organizing map (SOM), learning vector quantization (LVQ), radial basis function network (RBFN), probabilistic neural network (PNN), and others. A basic neural network is composed of numerous simple, interconnected processors known as neurons, each of which generates a sequence of real-valued activations. Sensors that perceive the environment activate input neurons, whereas weighted connections from previously active neurons stimulate other neurons. Early supervised neural networks were essentially versions of linear regression techniques that date back to at least the early 1800s. Consequently, learning networks with several nonlinear layers date back to at least 1965. In the 1960s and 1970s, backpropagation, an effective gradient descent method for teacher-based supervised learning in discrete, differentiable networks of arbitrary depth, was developed and used to neural networks. With the help of unsupervised learning, deep learning models become partially applicable in the real world. During the 1990s and 2000s, there were numerous advances in supervised deep learning. Deep neural networks have finally received widespread interest in the new millennium, mostly by surpassing alternative machine learning techniques, such as kernel machines Vapnik (1995) and Scholkopf et al. (1998), in a variety of significant applications. Deep neural networks are now applicable to the broader field of reinforcement learning, in which there is no supervising instructor. Deep learning techniques involve the addition of multiple layers to a neural network, though these layers may be repeated. Regarding this topic, the majority of deep

learning strategies rely on the four architectures listed below: (i) convolutional neural networks, which are regular neural networks that have been extended throughout space using shared weights; (ii) recurrent neural networks, which are regular neural networks that have been extended across time by having edges that feed into the next period step rather than into the next layer in the same time step; (iii) recursive neural networks, which are stratified networks where there is no time aspect to the input sequence, and (iv) standard deep neural networks, which consist of layers of varying types that are not repeated or arranged in any particular order.

When compared to other alternative approaches, they have demonstrated amazing result in the application of bankruptcy prediction. According to Charalambous et al. (2000) and Fu-yuan (2008) finding the underlying patterns in complex financial data benefits from their outstanding ability to handle nonlinear data. In principle, ANNs are capable of approximating any function, but they have a number of limitations, such as a lack of explanation capabilities as black-box algorithms, a lengthy training period, the potential inability to produce optimal solutions, and an overfitting to the training data. MLP produced great overall performance on the Taiwan and American markets, according to Huang et al. (2004). As shown by Blanco et al. (2013), MLP outperformed the conventional statistical models like Latent Dirichlet Allocation (LDA), Quadratic Discriminant Analysis (QDA), and Linear Regression (LR) in terms of misclassification costs for credit scoring in the microfinance sector. Rafiei et al. (2011) showed that, similar to Multiple discriminant Analysis (MDA), MLP accurately forecasted the bankruptcy of Iranian businesses. Because of its straightforward topology and quick computation, RBFN performed very well in tests of financial risk (Chakraborty and Sharma 2007). A stable credit rating model built on LVQ was used in Chen et al. (2011) to predict corporate failure and analyze credit risk. SOM is a non-parametric neural network that combines data abstraction and visualization in a way that is desirable. Through a two-step clustering technique, it was used to assess and illustrate the dynamic trajectory patterns of businesses across a number of years defined on the 2D space (Chen et al. 2013c).

Support vector machines

Since the first application of support vector machines (SVMs) in bankruptcy prediction (Min and Lee, 2005), there has been considerable interest in the application of SVMs to financial problems. SVMs transform the input vectors nonlinearly into a high-dimensional feature space by means of a kernel function, allowing linear models to separate the data. SVM is based on the search for an optimal hyperplane capable of classifying two classes and maximizing the margin of separation. Hua, Wang, Xu, Zhang, & Liang (2007), Boyacioglu, Kara, & Baykan (2009), Bellotti & Crook (2009), Yeh, Chi, & Hsu (2010), Yang, You, & Ji (2011) have been applied SVM to assess credit risk, because of its mathematical simplicity and least constraints. Due to their good generalization performance on high-dimensional and relatively small amounts of data, according to Yang et al. (2011) SMVs have gained widespread popularity in recent years. SVMs have been shown to be especially effective when the underlying data is typically nonlinear and non-stationary, yielding reasonably accurate models that can aid in bank lending decisions and profitability (Jayanthi et al., 2011). Huang, Thang, Lee, and Change (2012) utilized semi-supervised SVM to estimate the likelihood of financial distress. Bae (2012) demonstrated that an SVM model with radial basis kernel function (RSVM) performed exceptionally well for predicting Korean manufacturing distress.

K-Nearest Neighbor classifier

Regression and classification are investigated using KNN, a non-parametric and lazy learning method. It is non-parametric because the modeling approach is dependent on the dataset and there are no assumptions regarding data distributions. Because real-world datasets frequently do not conform to theoretical assumptions, this property is useful. In a lazy algorithm, all training datasets are implemented during the testing phase. This classifier is simple and easy to comprehend. As a result, it is broadly applied for classification due to its low calculation time and straightforward interpretation. KNN is utilized in various industries, including finance, healthcare, and handwriting detection. In credit scoring, financial and banking institutions evaluate the credit information of their customers to determine if they meet the standards for credit financing. KNN's operational mechanism consists of computing distance, locating the nearest neighbors, and voting for labels. This method's controlling variable is the number of nearest neighbors, denoted by the variable K. Each dataset has a unique optimal K value. If K is small, noise would have a significant impact on the outcome, whereas if K is large, it would be costly to compute. According to research, a small K is more flexible with a high variance and a low bias, whereas a large K is not flexible with a low variance and a high bias.

The training phase of KNN is significantly faster than that of other classifiers, and it is intuitive and straightforward because it makes no assumptions. KNN is useful for solving multi-class problems and regression problems. According to Genesis (2018), this method's weakness is that the testing phase of K-nearest neighbor classification is more expensive and slower due to the large memory required to store the entire training dataset. In addition, KNN is sensitive to magnitudes and outliers because it chooses neighbors directly based on distance. As shown by Avinash (2018), features with a high magnitude will weigh more than those with a low magnitude. In addition, it is unsuitable for large-dimensional data, and dimension reduction is required to improve performance. Furthermore, this method cannot handle the problem of missing values.

Case-based reasoning

By recalling similar experiences, case-based reasoning (CBR) can solve new problems and offer solutions. On the basis of the K-nearest neighbors (KNN) principle, similar inputs should result in identical outputs. A traditional CBR system consists of four steps: (1) retrieve KNN from the case base using a similarity kernel function (Retrieval); (2) predict the class of a new sample based on its neighbors (Reuse); (3) validate the prediction of the new sample (Revise); and (4) add the new sample and its predicted class to the case base (Retain). CBR is applicable to small data sets and enables adding or deleting data from samples. Although it was criticized for its low accuracy in comparison to other methods such as neural networks (Jo et al., 1997), the discriminant capability of CBR as a viable alternative method for bankruptcy forecasting is confirmed. Park and Han (2002) enhanced a CBR model with feature weights derived from analytic hierarchy processing (AHP) for case indexing and retrieval, while Li et al. (2011) developed a new CBR forecasting method based on similarities to both positive and negative ideal cases. Since the pioneering study of Schank and Abelson (1977), case-based reasoning (CBR) has been successfully implemented in numerous fields, such as credit scoring (Bryant, 1997; Buta, 1994; Shin & Han, 2001; Wheeler & Aitken, 2000).

Decision Trees

Decision Trees are a simple way to visually represent a decision. It is a non-parametric supervised learning technique and decision support instrument used for regression and classification. It can

define and derive a sine curve from the dataset using a series of if-then-else decision rules. The model is more accurate if the tree is deeper because there are more complex if-then-else rules. This method is useful for variable screening and feature selection, and it can accurately handle high-dimensional data with minimal data preparation from the user. This method requires less data cleaning than others because it is not affected by outliers. In addition, this method can handle both quantitative and qualitative data, which is an advantage over other approaches because other classifiers require a single attribute of the data. Nevertheless, there are weaknesses, such as the problem of overfitting, that can be eliminated by employing random forest. The decision tree is likely to have the problem of over-fitting if authors do not set a maximum depth limit because, as shown by Avinash (2018), the tree can keep growing until it has exactly one leaf node for every single observation, perfectly classifying them all, so the tree is so complex that it does not classify the datasets well. This method lacks stability because even minor variations in the datasets may result in the generation of an entirely different decision tree.

Random Forest modeling

Random forest is the most widely used algorithm; it gathers a large number of decision trees from a training dataset and uses bagging to perform classification and regression tasks (Krishni, 2018). This method collects the votes from these decision trees and determines the final class based on the class receiving the most votes (Savan, 2017). During training, each base model is created independently by acquiring knowledge from distinct random subsamples of the data. The samples are drawn using a method known as bagging or boot strapping, which means that some samples may be utilized multiple times in a single decision tree. Using different subsamples to train each decision tree, the entire forest will have a low variance but a high bias, despite the fact that each tree has a high variance in terms of its own training dataset. The process of averaging the predictions of each decision tree to derive the overall predictions at the time of testing is known as bagging. The underlying reason this approach outperforms the decision tree is that many uncorrelated decision trees can protect each other from individual errors to derive the ensemble predictions, thereby reducing the over-fitting problem and producing the most accurate prediction results among the currently available algorithms (Tony, 2016). Other strengths include the model's ability to run efficiently on a large number of datasets and its ability to determine which variables are significant in the classification (Saimadhu, 2017). Random forest outperforms linear models because it can identify nonlinear object-to-feature relationships.

Hybrid learning models

Several heterogeneous techniques are utilized in the analysis of a hybrid system, but typically only one is used for the final prediction. In the first method, heuristic techniques are combined with classification models to optimize prediction performance from multiple perspectives: (1) adjust the learning algorithm's parameters; (2) select the relevant features; and (3) refine the learning samples. Recent studies have demonstrated that the performance of learning models can be substantially improved by combining heuristic techniques such as genetic algorithm (GA), particle swarm optimization (PSO), annealing simulation (AS), and ant colony optimization (ACO). Min et al. used genetic algorithm (GA) to optimize the input futures and parameters of an SVM model in order to enhance its predictive ability (Min et al. 2011). Similarly, Tsai et al. (2013) used GA to perform both feature selection and instance selection before applying an SVM and KNN classifier. They proved that feature selection followed by instance selection during data preprocessing aids classification. According to Lin et al. (2008) and Marinakis et al. (2009) swarm intelligence

exemplified by ant colony optimization and particle swarm optimization was used to automatically select the optimal features utilized by subsequent machine learning models. By combining artificial bee colony (ABC) and support vector machine (SVM), the hybrid classification model improved America's corporate credit rating (Chen et al. 2013a). Yeh et al. (2012) proposed a hybrid model that combines random forests (RF) and RST to obtain valuable information, and applied KMV to assess the credit rating of corporate entities using market-based data. The second method for constructing a hybrid system is to cascade different classifiers, where the output of the first-level classifier serves as input for the second-level classifier. Ravisankar et al. (2010) constructed various two-phase hybrids utilizing MLP, PNN, Real-time substructure testing (RST), and Genetic Programming (GP), with each hybrid employing one technique as a feature selection method and another as a classifier.

Data and Methodology

According to Fetscherin and Heinrich (2015) and Apriliyanti and Alon (2017), bibliometric analysis identifies key articles and objectively highlights the connections between and among articles on a particular subject or field by analyzing the frequency with which they have been cited by other research papers. The findings from the study are useful not only for assessing the popularity of prominent authors and their works, but also for gauging their influence. Thus, a bibliometric citation analysis enables, through meta-analytic evaluation, the development of a particular research area or discipline, as well as assisting in the identification of essential research streams and their underlying conceptual approaches.

Bibliometric analysis is largely based on the premise that researchers publish their most significant findings in peer reviewed journals and rely heavily on previously research papers in similar journals (Raan 2003). The conceptual framework in any bibliometric analysis is the citation (Kim and McMillan, 2008), which goes beyond the simple counting of articles published to include centers of excellence (Fetscherin and Usunier, 2012) and to map out links between and among papers in a particular research field (Kim and McMillan 2008, Fetscherin and Heinrich 2015, Apriliyanti and Alon 2017).

The papers used in this study are published research data from Scopus-indexed journals. Journal articles are the type of articles that we endorse. Every article's data indexed by Scopus, formatted as journal articles, and conforming to the themes required for this study are backed up in a file that is utilized by VOSviewer.

In this study, only articles relating to "credit risk" were filtered. We conduct a data search on Scopus using the keyword "Credit risk" in accordance with the title, keyword, and abstract criteria. Thus, 4923 articles were obtained and evaluated in accordance with the selected theme. The articles used in this study were published between 2000 and 2023. The collected articles then were saved in the *.CSV file format. After that, we used the VOSviewer application to visualize and analyze bibliometric maps of trends. VOSviewer is utilized for mapping analysis of papers, countries, and journals based on co-citation and keywords visualized on a network of existing terms' relationships. In addition, we filter the keywords that are going to be included in the network mapping visualization produced by VOSviewer.

Results and discussions

Recent developments in credit risk research

Table 1 shows the development of research in the field of credit risk published in Scopus indexed journals. Based on the data shown in Table 1, it can be seen that the number of credit risk researches filtered by certain keywords is 4923 articles from 2000-2023. The keywords with a significant number of publications include risk assessment, credit risk, risk management, credit risks, finance, machine learning, credit scoring, decision trees, credit risk assessment, support vector machines, credit risk evaluation, machine-learning, random forests, genetic algorithms, credit scoring model, credit risk analysis, deep learning and credit risk modeling.

Table 1. Number of papers published in Scopus database between 2000 and 2023 filtered by keywords

Keywords	Number of articles filtered by keywords
Risk Assessment	2083
Credit Risk	1892
Risk Management	848
Credit Risks	790
Finance	713
Machine Learning	430
Credit Scoring	408
Decision Trees	278
Credit Risk Assessment	259
Support Vector Machines	232
Credit Risk Evaluation	120
Machine-Learning	108
Random Forests	89
Genetic Algorithms	88
Credit Scoring Model	86
Credit Risk Analysis	82
Deep Learning	81
Credit Risk Modeling	66

Source: Scopus database.

Figure 1 illustrates the yearly growth or development of research in the field of credit risk. Based on Fig. 1, it is known that the development of research related to credit risk has increased in the last 23 years. This is shown by the number of articles in 2000 as many as 24 articles, which continued to grow until it increased by 563 articles at the end of 2022 to 587 articles. A very significant growth occurred from 2017 to 2022 were from the total number of research articles 226 in 2017 to 587 in 2022. This shows an increase by 361 articles between 2017 and 2022.

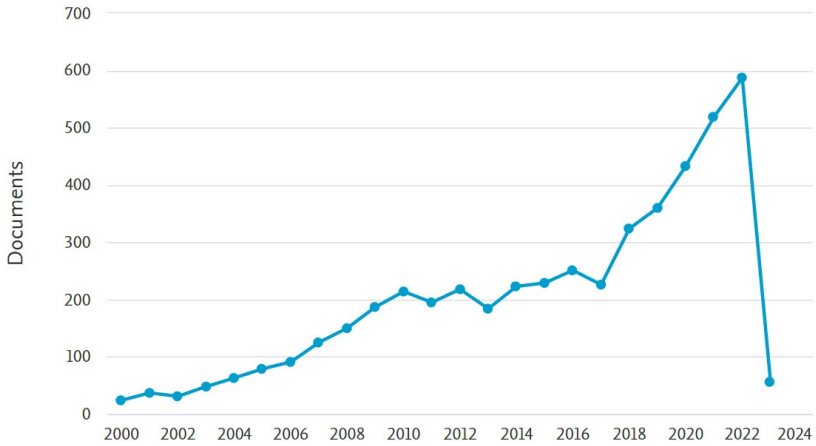


Figure 1. The evolution of credit risk research

Source: Scopus database.

According to the article data that has been collected, there are a total of 66 items associated with credit risk. The 66 items are organized into 14 groups. Each cluster illustrates the connection between two terms. Each term is labeled with a circle of a different shade. There is a positive correlation between the size of the circle and the frequency of terms in the title and abstract. The size of characters and circles is defined by their occurrence rate. The larger the size of the letters and circles, the more frequently the term appears.

Figure 2 depicts the relationship between network-described terms. In network visualization, the relationship is represented by a line from one term to another. Fig. 2 illustrates the relationship between each studied term and credit risk topics. The term credit risk appears most frequently in this research on network visualization.

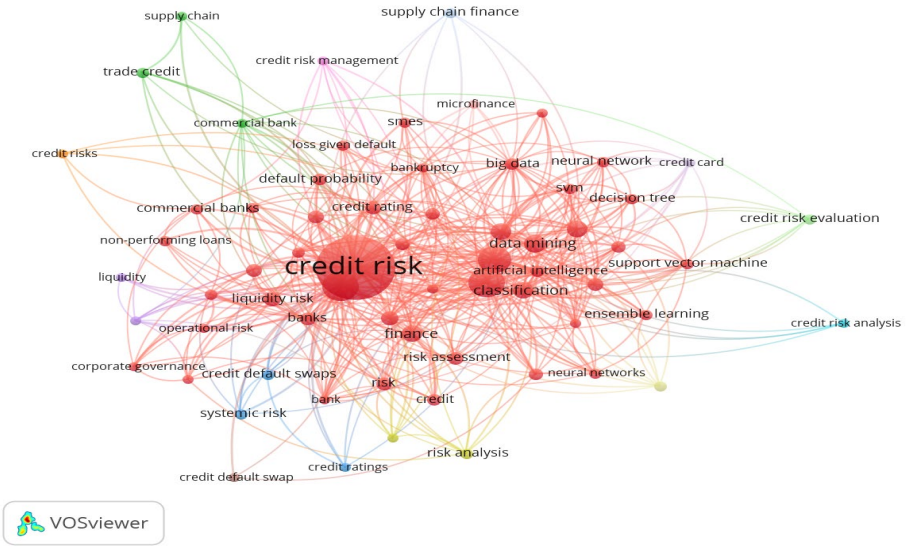


Figure 2. Network visualization of credit risk keywords

Source: Authors' own research.

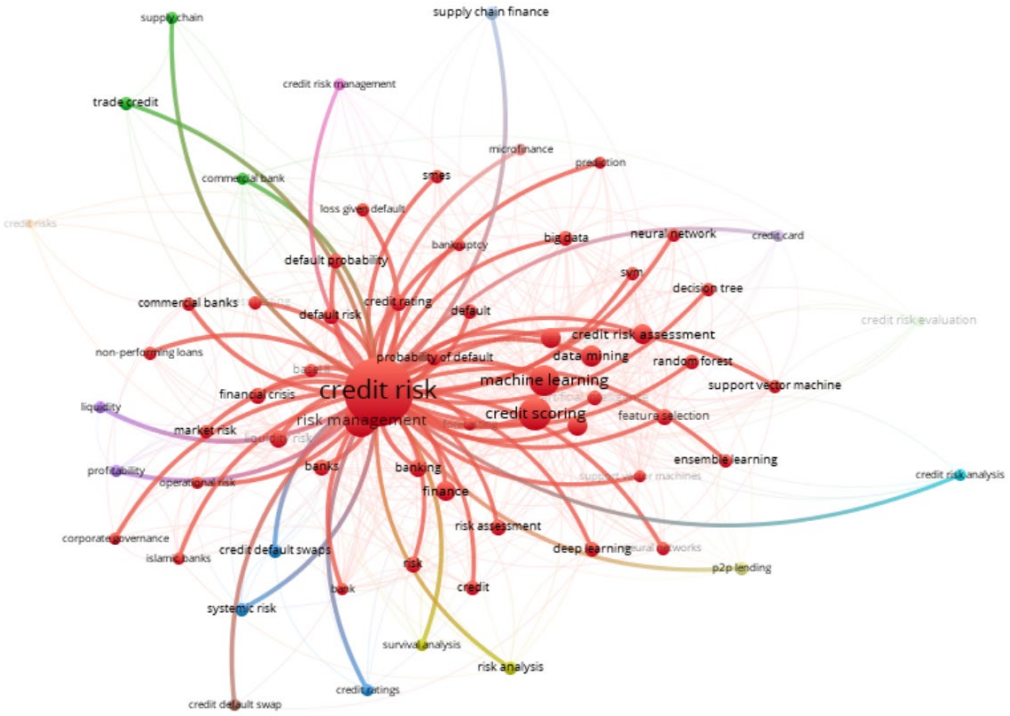


Figure 3. Credit risk keyword network visualisation

Source: Authors' own research.

According to Figure 3, the credit risk term is directly associated with 64 items with a total link strength of 1,413 links. As shown in Fig. 3, the term credit risk is associated with the following terms: artificial intelligence, credit risk assessment, credit scoring, machine learning, random forests, support vector machines, etc. It can be concluded from Figure 3 that the term credit risk has a very expansive relationship with other terms. Consequently, the possibility of research credit risk remains highly probable. The results of the network visualization also indicate that credit card research is still in its early stage and that credit card connectivity is the lowest, with approximately 20 links. Thus, the opportunity to conduct credit card research is quite high.

Conclusion

Throughout this review, we introduce the current state of the literature on financial credit risk assessment, discuss conventional and cutting-edge methods, and highlight some of the most important topics in latest studies.

The purpose of this investigation is to conduct a bibliometric analysis of the bibliographic data of credit risk research articles indexed by Scopus by trying to combine analysis in terms of the distribution of bibliometric maps with VOSviewer software. Using this methodology, we examined credit risk studies over the past 23 years in this paper. We utilized this method by analyzing 4923 credit risk papers published in 163 journals indexed in Scopus database. While this approach is widespread across numerous other fields of study, according to Kysucky and Norden (2015) it is relatively rare in finance.

In comparison to other scientific database, such as Web of Science and Google Scholar, Scopus does not cover all relevant publications. This may be a limitation of the paper. Thereby, it is probable that essential articles were omitted from our analysis. Another limitation of this research

is that bibliometric analysis is conducted only on research containing the keyword "credit risk." For future research, we will analyze data on credit risk in a broader manner, eschewing the Scopus database in favor of Web of Science and Google Scholar, for instance.

Acknowledgements

This paper was co-financed by the Bucharest University of Economic Studies during the PhD program.

References

Capotorti, A., & Barbanera, E. (2012). Credit scoring analysis using a fuzzy probabilistic rough set model. *Computational Statistics & Data Analysis*, *56*(4), 981-994.

Chen, N., Ribeiro, B., & Chen, A. (2016). Financial credit risk assessment: a recent review. *Artificial Intelligence Review*, *45*(1), 1-23.

Li, Z., Tian, Y., Li, K., Zhou, F., & Yang, W. (2017). Reject inference in credit scoring using semi-supervised support vector machines. Expert Systems with Applications, 74, 105-114.

Shi, S., Tse, R., Luo, W., D'Addona, S., & Pau, G. (2022). Machine learning-driven credit risk: a systemic review. *Neural Computing and Applications*, *34*(17), 14327-14339.

Uddin, M. S., Chi, G., Al Janabi, M. A., & Habib, T. (2022). Leveraging random forest in micro-enterprises credit risk modelling for accuracy and interpretability. *International Journal of Finance & Economics*, *27*(3), 3713-3729.

Wang, Y., Zhang, Y., Lu, Y., & Yu, X. (2020). A Comparative Assessment of Credit Risk Model Based on Machine Learning – a case study of bank loan data. *Procedia Computer Science*, *174*, 141-149.

Zamore, S., Ohene Djan, K., Alon, I., & Hobdari, B. (2018). Credit risk research: Review and agenda. *Emerging Markets Finance and Trade*, *54*(4), 811-835.

Zekic-Susac, M., Sarlija, N., & Bensic, M. (2004, June). Small business credit scoring: a comparison of logistic regression, neural network, and decision tree models. In *26th International Conference on Information Technology Interfaces, 2004.* (pp. 265-270). IEEE.