IEC
Institute of Electronics
and Computer

Journal of Artificial Intelligence and Systems, 2020, 2, 118–132
https://iecscience.org/journals/AIS
ISSN Online: 2642–2859

Artificial neural network and Bayesian network models for credit risk prediction

Germanno Teles[1], Joel J. P. C. Rodrigues[1,2,3,*], Ricardo A. L. Rabêlo[2], Sergei A. Kozlov[3]

[1]Instituto de Telecomunicações, Universidade da Beira Interior, Covilhã 6201-001, Portugal
Email: germanno.teles@ubi.pt
[2]Centro de Tecnologia, Federal University of Piauí (UFPI), Teresina – PI 64049550, Brazil
Email: joeljr@ieee.org, ricardoalr@ufpi.edu.br
[3]Photonics and Optoinformatics Department, ITMO University, St. Petersburg 197101, Russia
Email: kozlov@mail.ifmo.ru
*Corresponding Author: Joel J. P. C. Rodrigues, Email: joeljr@ieee.org

How to cite this paper: Germanno Teles, Joel J. P. C. Rodrigues, Ricardo A. L. Rabêlo, Sergei A. Kozlov (2020) Artificial neural network and Bayesian network models for credit risk prediction. Journal of Artificial Intelligence and Systems, 2, 118–132.
https://doi.org/10.33969/AIS.2020.21008

Received: January 5, 2020
Accepted: March 3, 2020
Published: March 9, 2020

Abstract

Credit risk threatens financial institutions and may result in irrecoverable consequences. Tools for risk prediction can be used to reduce bank insolvency. This study compares Bayesian networks with artificial neural networks (ANNs) for predicting recovered value in a credit operation. The credit scoring problem is typically been approached as a supervised classification problem in machine learning. The present study explores this problem and finds that ANNs are a more efficient tool for predicting credit risk than the aïve bayesian (NB) approach. The most crucial point is related to lending decisions, and a significant credit operation is associated with a set of factors to the degree that probabilities are used to classify new applicants based on their characteristics. The optimum achievement was obtained when the linear regression was equivalent to 0.2, with a mean accuracy of 85%. For the naïve Bayes approach, the algorithm was applied to four datasets in a single process before the entire dataset was used to create a confusion matrix.

Keywords

Artificial intelligence, Neural network, Bayesian network, Algorithms, Credit risk, Prediction

1. Introduction

Banks are exposed to a wide range of potential risks, ranging from those identified in the budgetary and technological structure to those related to brand reputation and those derived from the social and institutional environment. Because of the level of technology associated with Big Data, computing power, and data availability, most lending institutions have been compelled to renew their business models. Forecasting credit risk, active loan processing, and monitoring model reliability are vital for transparency and decision-making. From the viewpoint of machine learning, the problem has typically been approached as a problem of supervised classification. In this study, binary classifiers are built based on a machine model to predict the probability of loan default, and the results are compared with those from artificial neural networks (ANNs) to examine their reliability and efficiency. A practical case is used to exhibit the applicability, efficiency, flexibility, and accuracy of data mining approaches to model ambiguous events related to measuring credit risk for financial institutions. However, aside from the technical questions necessary to understand it, confidentiality issues raised from the use of personal data are also important. The application of the ANN algorithm has long raised numerous ethical questions [1]. According to Williams, Sweeney, and Anderson (2007), such problems continue to be addressed through discussions on artificial intelligence. The underlying concern is the fear that an algorithm may take the decision power away from a human. Given that these debates and questions are legitimate, the present study focuses on the algorithms relevant to decision-making in the financial sector, especially those used to simplify/increase fluidity and speed processes [2][41].

Algorithms are sets of codes designed to attain predefined objectives. For example, in a recruitment process, they can discriminate among people based on their profiles. According to Yaseen (2019) the all those classifiers introduce on credit score approach are well-known base classifiers in this domain are used o his work they show results, analysis, and statistical tests demonstrate the ability of the proposed combination method to improve prediction performance against all base classifier. Baghban (2019) considered classification an example of supervised learning as training data associated with class labels a focus on study of various classification techniques, showing its advantages and disadvantages. The work of Wu et.al. (2019) present a study on BI (Business Intelligence) on a bank institution showing the potentially benefit business, increasing the visibility and recognition of research achievements. The behavior of a BN can be reinforced by the work of Abid et.al.(2017) that use this classifier on analyses of costumers loans default payment allowing providing an effective decision support system for banks[66] [5][38][64] . The same approach is used in providing loans to an enterprise, where a bank's lending decisions are based on the algorithm used [3]. Therefore, it is critical to comprehend the underlying problems and establish ways to regulate the use of algorithms [4].

1.1. Artificial neural networks

The study of ANNs can be traced to Frank Rosenblatt (1958) who focused on perceptron algorithms for the development of smart automated systems and software [6]. ANNs are a reliable and efficient approach for predicting outcomes. The method received tremendous support following the development of machine language. Odom and Sharda [43] applied neural networks to the evaluation of credit risk. Initially, the network was based on the Hebb system, which aimed to improve the input vector through perception and focused on increasing the accuracy of the model. The perception neural network was followed by backpropagation, developed by Rumelhart and McClelland [48]. Backpropagation refreshes its weight through the maintenance of history, commonly known as neural processing. However, more interest was focused on deep learning in 2008 when Angelini and colleagues performed the first credit risk analysis used by banking management to compute capital requirements. ANNs have also been used to calculate the variables necessary to evaluate credit risk [22].

ANNs were used to study the nervous system and the way the brain processes information. An ANN entails processing algorithm to model the brains of humans and comprises a large number of interconnected nodes (neurons) working as a system to solve pattern recognition or data classification problems, particularly in the field of biology [25],[28]. The purpose of ANN research is to develop a computational system with relatively low computational cost and time commitment. A range of tasks can be performed by ANNs, including classification, pattern-matching, approximation, function optimization, data clustering, and vector quantization. According to Rumelhart and McClelland, ANN properties initially include the following [22]:

- The cycle or speed of the implementation of ANNs is in nanoseconds.

- The processing time is rapid, and numerous operations can be performed simultaneously.
- he complexity and size of an ANN is subject to the network design and application in use.
- The data in ANNs are stored in contiguous memory sites that can become overloaded when the limit is exceeded.

Learning is the primary property of ANNs, and this occurs in two forms: structure and parameter learning. Parameter learning improves the weight linked to the network, whereas structure learning concentrates on network topology and confirms whether there are any changes within the framework [7]. Moreover, learning can either be supervised or unsupervised depending on expert knowledge [56]. Within supervised learning, reinforcement learning contains an activation function useful for calculating the exact output [8]. This feature is applied to the overall input to determine the total network production [9]c. Some types of activation functions include the binary step, hyperbolic tangent, bipolar sigmoid curve, and identity functions [12].

1.2. Naïve Bayesian (NB) approach

The NB was initially studied by applying Bayes' theorem. The approach comprises a supervised statistical classifier based on the assumption of conditional independence, where probability models are estimated using labeled data (i.e., each instance is assigned to a class) [10]. Mutual conditional probability distributions are used in the Bayesian classifier, allowing class-based conditional independencies to function between variables, with a graphical model used to depict the underlying relationships [13][63]. The random variables form either a continuous or discrete relationship while the attribute within the data may have a real Boolean variable to formulate the relationship [14]. Every arc in the acyclic graph represents the dependence probability, and all variables are independent of the non-descendant.

The formula of Bayes theorem is given as follows[15]:

$$P(A|B) = P(B|A)P(A)/P(B) \tag{1}$$

where in a sample referred to as A, the chances of all events occurring is h. Further, P (h—A) is consistent with Bayes' theorem, which can be stated mathematically as in Eq. 1. This classification is considered the optimal one [16] [59].

When the network topology and data are given in the multiple variables of a sample, data training is impartial. The variables are then used to determine the entries in the continuous probability table. This approach lowers computational costs and is suitable for problems where a strong relationship exists between the variables [61]. The approach is also highly advanced in comparison with support vector machines, and is also applicable to medical diagnosis [62]. Compared with other algorithms such as particle swarm optimization, neural networks, and machine language algorithms, studies based on support vector machines have been promising for assessing credit risk [11].

The present study demonstrates that various algorithms can be used in parallel to address the issue in question, which in the case presented here is loan provision [65]. Multiple strategies for identifying the choice of features (or variables), algorithm, and criteria can provide a solution. For instance, in the new Big Data and digital era, transparency is critical [22]. Strategies based on deep learning are also necessary to train data on the application, and machine learning algorithms and their use must be regulated to ensure accuracy.

The particular focus of this study is credit risk scoring and how distinct machine learning models can help lenders identify default [17]. Further, the stability of these models is examined based on the choice of variables or subsets. Although the methods used by banks in their decisions to award loans remain unclear, the application of classical linear models in the banking sector is well known [29]. Finally, a transparent elastic approach is used as the benchmark, and its fit and decision rules are compared. To the best of the authors' knowledge, there are no solutions currently available in the literature for credit scoring based on ANNs and Bayesian networks. Further, this study aims to determine the best combination of parameters to use with ANNs and the BN approach to handle and precisely evaluate credit risk. The main contribution of this study is its proposed machine learning model or the combination of them, which is a rare approach to the problem of credit risk measurement. The study highlights the existing gap that prevents intelligence systems from addressing bank modeling concerns.

The paper is organized as follows. Section 2 presents the materials and methods, and describes the datasets and attributes used to forecast credit risk, the research data, and the prediction model. Section 3 reports the results and discusses the analysis when the training and test datasets were created using the Bayes and ANN approaches and presents the analysis defining a systematic method to handle and precisely evaluate credit risk. Finally, Section 4 summarizes the processes used to identify the best combinations and concludes the paper.

2. Data and methods

2.1. Research data

The dataset was collected from a financial institution, and a summary of 1,890 records was accessed and retrieved. As some attributes were missing, the dataset required preprocessing. The global mean approach was employed to replace the missing attributes (El-Shazly, 2002). Each record/instance was assigned to one of two classes: 1 (risky) or 2 (non-risky). The output of the processing represents the label as opposed to the value, where 1 indicates credit risk and 2 represents security (Demerjian, 2007). The predictor attributes include contract_value, balance_value, collateral_value, number_of_collateral, recovered_value, value_tx_rate, value_tx_interest_rate, value_rate_overdue, client_size, main_value_delay, seniority_level, duration_in_years, duration_in_months, duration_in_days, and delay_in_days. For accuracy, the attributes in the dataset were converted into classes, which is critical for preprocessing as it improves the accuracy of the algorithms [19]. Numerous banks have adopted these attributes to predict credit risk [67]. Moreover, to find the corresponding class, the data were normalized for ANNs, as the output ranges between 0 and 1. After normalization and conversion, the data were stored as a comma-separate values file and retrieved for further processing. The dataset was then classed into sets based on the age of the credit operation and trained. The weights were computed based on training and later tested.

2.2. Prediction model

The proposed model constitutes two complementary phases, the NB and ANN phases. The ANN phase is used to estimate the overall credit risk trend and establish the most significant factors, whereas the NB approach determines the probability of credit default when all variables have been measured [12][10]. Therefore, the two phases complement each other. From a technical perspective, although the same raw data were used to implement the two networks, no data flow exists between the systems [57]. Data implementation is independent because of the distinct rationales behind the two policies. Specifically, the output from one network cannot be used as an input of the other. The output was between 0 and 1 and the outcome was marked 2 when the result ranged between 0.75 and 0; otherwise, it was 1. The ANN was based on continuous raw data, most often coded in MATLAB, while the BN input data were converted into Boolean before being coded in MATLAB. The outcome of the two phases allows for validation and verification through the parallel and independent implementation of the dataset.

The proposed neural network is a multilayer perceptron (MLP) based on a feed forward architecture. The popularity of this architecture can be attributed to its link to the robust and powerful learning algorithm referred to as backpropagation learning. In the design of ANNs, the most critical element is accurate identification of the learning algorithm used in the training process [51][36]. In the present case, an active or supervised learning mechanism is used and an appropriate learning rule is applied. Gradient descent is used to adjust the relationship values and the Levenberg–Marquardt algorithm (LMA), one of the most common algorithms in computing and mathematics, was used to compute optimization problems that arise from generic curve fitting [24].

By selecting the chromosome with the lowest cost, the search process continued until the weights were turned into the target solution [37][40]. The learned credit risk function was configured using the autoregressive pattern to predict default risk via an MLP network [42]. The default condition of a given day is subject to the credit level of the previous days. Banks usually consider a time span of 30 days for credit strategies, and when confronting shortages, they invoke proxy funding resources [14][23].

NBs were used to identify the most critical risk indicators among those chosen as the model variables and their effect on each other and on the default risk measure was assessed. NBs are useful for graphically representing probabilistic relationships between variables [39]. They are crucial in data modeling, particularly when data are missing, because they characterize variables based on a combination of graphical models and statistical approaches. NBs can thus detect the possible relationships between variables and, thanks to causal inferences, help predict their trend using probability distribution functions regardless of the nature of the data [44]. Moreover, prior knowledge can be merged with existing data to provide accurate results, thereby leading to correct inferences. This method, together with Bayesian approaches, therefore prevents overfitting the data [49]. Hence, a Bayesian knowledge base allows researchers to draw conclusions and inferences about the relationships among the components of a system, making it the most suitable approach for achieving the second objective of this study.

2.2.1. Key parameters

The results generated by both the NB and ANN phases are the product of supervised learning. A network is produced starting from a random weight in the ANN classifier and the distributions in the NB classifier. The resulting system is trained using the dataset until the outcome is comparable to the distribution or pattern of the primary data[45]. The algorithms applied to the training data in the first and second phases were gradient descent and maximum likelihood estimation, respectively [13] [14] [20] [63]. Training based on algorithms is a standard procedure, and the parameters must be selected correctly. The primary concern is estimating the appropriate parameters for the trained function and probability distribution provided by ANNs and NBs to fit the data (Mileris, 2010). In the model, there are two sets of parameters, namely the sets of weights and binomial distribution parameters in ANNs and NBs, respectively.

During the first phase, the ANN characterizes a function of the datasets (input variables) and attempts to locate the most appropriate coefficients (weights) for the variables. Once the algorithm has learned the data, the target values can be estimated and hence default risk can be predicted[61] [18]. The learning process in the second phase occurs by applying the naïve Bayes rule. Nodes are identified using input variables and considered to contain the prior distribution. Similar parameters then define the previous characteristics of the network nodes.

2.2.2. Prediction and measurement of credit risk

Using the receiver operating characteristic (ROC) curve to evaluate a diagnostic test, an ANN interpolates between the Gauss–Newton algorithm and the gradient descent method [21]. Although the LMA is more robust than the Gauss–Newton algorithm, as with many fitting algorithms, it only finds the local minimum, which is not necessarily the global minimum [26]. To overcome this shortcoming, the genetic algorithm (GA) can be used to search the space of possible solutions [19]. GA first generates a random vector as the weight vector (chromosome), to which crossover and mutation are then applied. The output vector is calculated using inputs and weights, and the differences between the output and target values are introduced as the cost [36]. By selecting the lowest cost chromosome, the searching process continues until the masses evolve into a suitable final solution. Finally, to predict liquidity risk using an MLP network, the learned liquidity risk function is configured with an autoregressive pattern based on the type of risk under analysis[28]. The liquidity condition of a particular day strictly depends on the liquidity levels of the preceding days. Considering the nature of the problem, a powerful computational tool is needed to estimate and predict the credit risk function through the data provided. The ANN architecture with computationally intensive learning and massive parallelism through examples renders it suitable for the task at hand [34].

The ANN based on backpropagation and the naïve Bayes algorithm were implemented in MATLAB software. Initially, the data were divided into categories and the best combinations of the parameters were determined based on the learning rate, epoch, model checking rate, and number of neurons [27], as shown in Tables 1 and 2. The focus was not only on blending the parameters but also ensuring the accuracy of these combinations.

Table 1. Training dataset

Supervised Learning (Naïve Bayes Continuous)						
Parameters						
Lambda				0		
Homoscedasticity assumption				1		
Classifier Performances						
Error Rate			0.4037			
Value Prediction			Confusion Matrix			
	Recall	1-Precision		RISKY	NON- RISKY	Sum
RISKY	0.4721	59.63%	RISKY	642	446	1088
NON-RISKY	0.7227	63.85%	NON- RISKY	247	555	802
				889	1001	1890

Table 2. Test dataset

Supervised Learning (Naïve Bayes Continuous)						
Parameters						
Lambda				0		
Homoscedasticity assumption				1		
Classifier Performances						
Error Rate			0.3615			
Value Prediction			Confusion Matrix			
	Recall	1-Precision		RISKY	NON- RISKY	Sum
RISKY	0.6073	36.60%	RISKY	201	82	283
NON-RISKY	0.3735	42.63%	NON- RISKY	288	59	347
				489	141	630

3. Results and discussion

Twelve columns stood out in the data analysis: "Contract Value," "Balance Value," "Collateral Value," "Recovered Value," "Value Tax Rate," "Value Tax Interest Rate," "Main Value Delay," "Duration in Years," "Duration in Months," "Duration in Days," and "Delay in Days." These columns were eliminated immediately because the exercise focuses on categories or classes of data that operate as predictors of creditworthiness. Rather than hard figures and numbers, the true indicators of creditworthiness are scores. The columns that qualified as scores were "Value Rate Overdue," "Client Size," "Seniority Level," "Percent Used," and "Number of Collateral." The most important point is that these variables are related to lending decisions, and a significant credit operation is related to a set of factors to the degree that probabilities are used to classify new applicants based on their characteristics. Numeric data are removed, while specific classifications are preserved by creating an object that eliminates the identified columns. The training and test datasets were then created. The training dataset was used to train the model, whereas the test data were used to assess the model's accuracy. One-third and two-thirds of the data were allocated to the two sets, respectively. Next, the learning rate was iterated and plotted with different standards of accuracy to obtain a non-linear graph [31]. The optimum performance was obtained when the linear regression was equivalent to 0.2, with mean accuracy of 85%. For the naïve Bayes approach, the algorithm was applied to four datasets in a single process before the entire dataset was used to create a confusion matrix. The best outcomes of the iteration were sourced using the comparative approach [33].

3.1. Bayes approach

Table 1 describes the results of the two Bayesian models, showing that the classification rate improves when the indicators relating to "Value Rate Overdue," "Client Size," and "Seniority Level" are introduced. The best classification rates are 59.63% and 63.85% for the two classes. The criterion of the two types of errors (Type I and Type II) has been examined in numerous studies. The assignment of variables is as follows:

X1: Number_of_collateral
X2: Value_tx_rate
X3: Value_tx_interest_rate
X4: Value_rate_overdue
X5: Client_size
X6: Main_value_delay
X7: Percent Used
X8: Duration_in_years
X9: Seniority_Level

Type I error is also known as credit risk, which is the rate at which bad clients are classified as profitable. Therefore, when a bank has a significantly high rate, which implies that the rate of loan approval is too high, the potential for exposure to credit risk is considerable. Type II error is commercial risk, which is the rate at which applications of paying clients are rejected, with the bank experiencing an opportunity cost attributed to good customers. In the present study, the Type I error is exceedingly high at 52.79%. The introduction of seniority improves the outcomes and the classification rate rises. Further, the model based on the entire dataset indicator reduces the Type I and Type II errors to 32.97% and 39.27%, respectively. These findings show the correlation between value rate overdue and seniority and credit risk, concurring with previous results [35].

3.2. Artificial Neural Network approach

The objective of this approach was to estimate the credit risk function, and therefore continuous data were necessary. The only preprocessing of the dataset conducted was data normalization. As before, data were divided into two categories for training and testing at ratios of one-third and two-thirds [47][60]. The selected network comprised three layers: the MLP layer, the hidden layer, and an output layer [46]. The optimal structure was chosen through trial and error with the network assessed using the micro & small enterprises(MSE). The correlation between the output and target values, variance, mean residuals, learning process error, and root MSE were all used in the assessment of the network [46]. Because most of the hidden cases were sufficient for the network to perform optimally, several models were implemented using a single network. Figure. 1 exhibits the outcome of the assessment obtained from network training using the LMA.

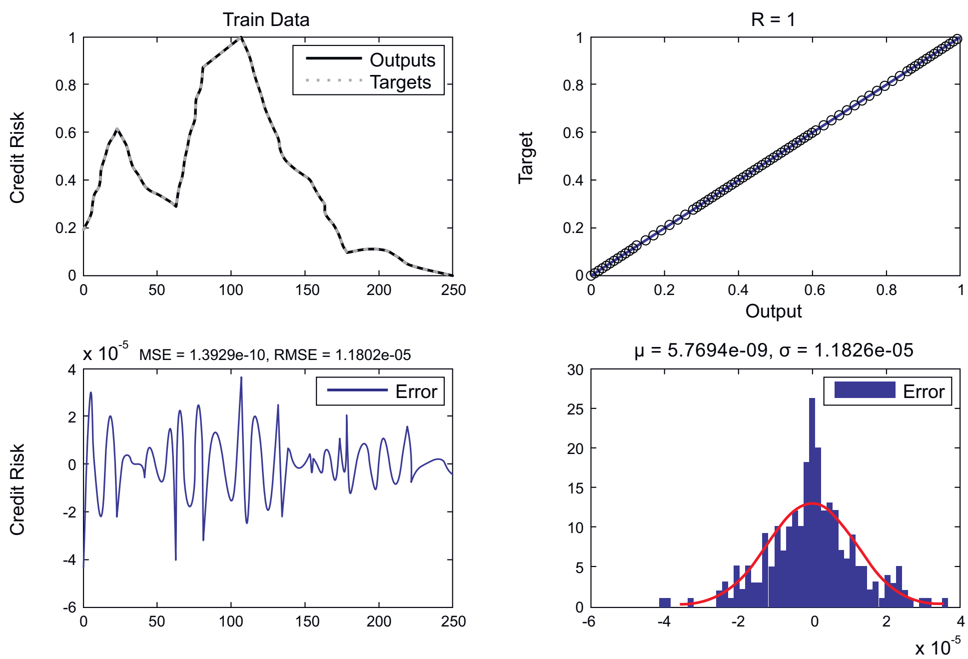

Figure 1. Analysis of the LMA-trained data.

The LMA was used to train the datasets; however, its inability to locate the global minimum influences the results from this approach[52]. Furthermore, the accuracy of the results was sufficient as long as the initial weights were a good approximation and the signal-to-noise ratio exceeded five [28]. For this reason, a meta-heuristic search algorithm, the GA, was used. Given that the GA has random behavior and is independent of its starting point, its application guaranteed that the LMA was functioning correctly. In addition, apart from overcoming the drawbacks of the LMA, the GA in figure 2 demonstrates that the dataset was sufficient to be modeled by any preferable algorithm. As shown in Table 3, the performance of the LMA was much better than that of the GA and it recognized data patterns accurately. As a result, credit risk was modeled using the LMA. Moreover, both the ANN and the Bayesian models provide reliable outcomes, but the former is more effective in the prediction of credit risk with an average score of 82% (Table 4).

The ROC curve in figure 3 aids the visualization and shows the trade-off between recall and precision. This allows the researcher to manipulate the false positive and true positive metrics [53][55]. The relationships among false positives, true positives, false negatives, and true negatives were further summarized using a simple confusion matrix. The probability of the above case occurring was 1, representing a 100% correlation, with 0 depicting no relationship.

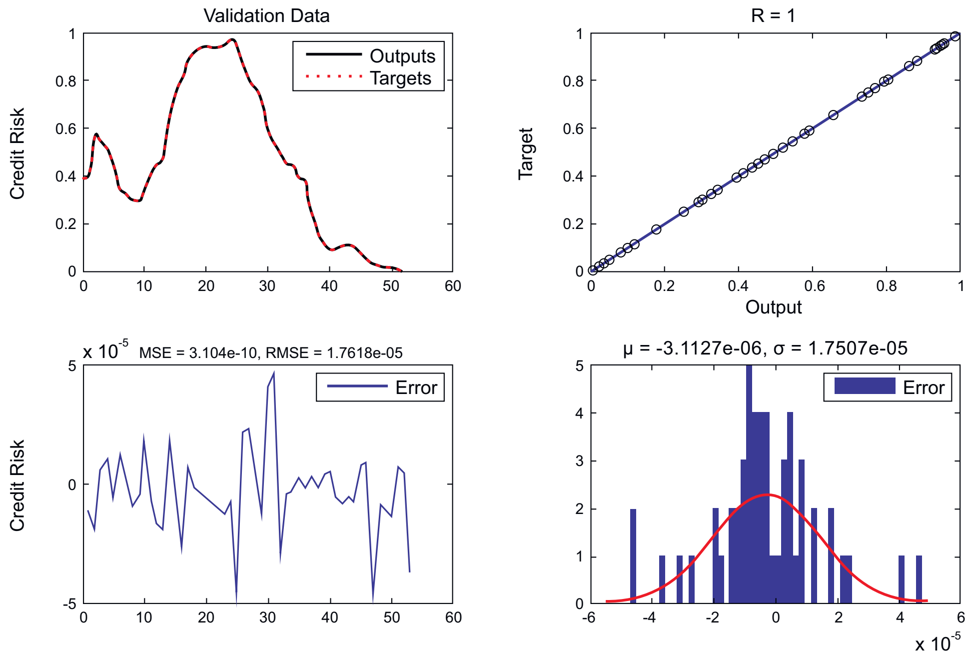

Figure 2. Assessment of the learning process: Validation based on the GA.

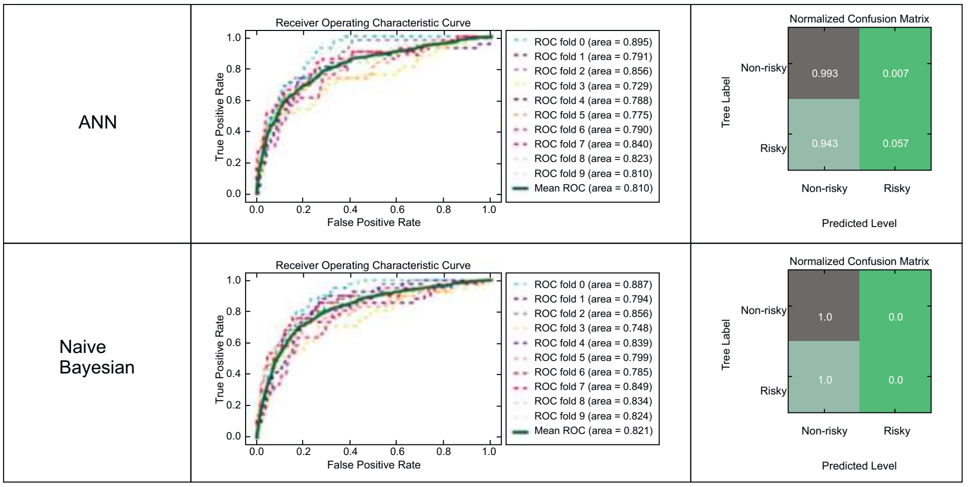

Figure 3. Classifier's ROC curve and confusion matrix of the ANN and NB.

Hence, compared with the NB, the ANN was more accurate and efficient as well as more promising for determining credit risk [30][58]. The Bayesian classification allocated the data classes into the tables where the values and attributes of an entity were predicted to be independent of others [31][32]. Because of the random nature of the ANN, the estimation accuracy was hugely dependent on the cases selected for training [34]. Therefore, the reported numbers could have changed slightly due to regular running. Similarly, the network structure becomes complex and quality is reduced because more time is needed to train the data[54].

This study addressed the issue of defining a systematic method to handle and precisely evaluate credit risk. The vagueness and ambiguity that characterize the credit risk concept complicate the formulation of an undisputable definition. Identifying the factors that

Table 3. Comparative analysis of BN and ANN performance by GA

Comparison Metric	Bayesian Networks	Artificial Neural Networks
Run Time	175s	6s
Training Data MSE	$9.1e^{-3}$	$13e^{-10}$
Validation Data MSE	$1.3e^{-2}$	$3.3e^{-10}$
Test Data MSE	$8.0e^{-3}$	$1.7e^{-10}$

Table 4. Comparison of classification accuracy

Algorithm used	Classification accuracy	Correctly classified cases	Incorrectly classified cases
Naïve Bayes	81.32%	1537/1890	353/1890
Neural Network	81.85%	1547/1890	343/1890

establish and influence credit risk to formulate a suitable functional form to estimate and predict its value is a difficult task [27]. Similarly, the spread and complexity of the credit risk phenomenon render traditional mathematical modeling techniques obsolete [50][37].

The present study proposed an approach that employs two of the most recent machine learning methods—Bayesian networks and ANNs—to address this concern. In the model, the variables were selected based on the data available from bank databases[52][55]. Despite the numerous capabilities of NBs and ANNs, machine learning methods, or a combination of them have been seldom used to approach the problem of credit risk measurement [54]. Therefore, this study bridged the existing gap that prevents intelligence systems from challenging bank modeling concerns. In particular, the focus was on the concept of insolvency as a characterization of credit risk [58]. As a result, inner factors were used to construct a model whose attributes permit prediction of credit risk issues. The case used bank data to demonstrate the accuracy, efficiency, flexibility, and rapidity of the data mining approaches when modeling events related to the measurement of credit risk. Implementation of NB and ANN can differentiate between the riskiest factors and estimate subject risk through the training and learning process. The results were highly consistent. Further, the binary outcomes gathered from the study depicted the ability of the NB–ANN approach to validate the findings through a parallel and independent implementation of the dataset.

4. Conclusion

The ANN algorithm based on backpropagation and the NB algorithm were implemented. Data were divided into categories to determine the best combination of the parameters. The best combinations were then observed, and the result was generated, gathered, and presented. The focus was not only on blending parameters but also on ensuring the accuracy of these combinations. As demonstrated, the optimum performance for ANNs was obtained when the linear regression was equivalent to 0.2, with mean accuracy of 85%. For the NB approach, the algorithm was first applied to four datasets in a single process and then the entire dataset was used to create a confusion matrix. The best outcomes of the iteration were sourced from the comparative approach. It was concluded that both the ANN and the NB models provide reliable outcomes, but the former is more effective for predicting credit risk with an average score of 82%. Future work may consider a comprehensive validation of the suggested method with other credit scoring databases identified by the high noise level sets method, with other methods such as forecasting routines used in the retail and consumer investment areas.

Acknowledgements

This work was supported by FCT/MCTES through national funds and when applicable co-funded EU funds under the Project UIDB/EEA/50008/2020; by the Government of the Russian Federation, Grant No. 08-08; by National Council for Scientific and Technological Development (CNPq), Brazil, Grant No. 309335/2017-5]; by Ciência sem Fronteiras of CNPq, Brazil [Grant number 200450/2015-8].

Conflicts of Interest

There is no conflict of interest.

References

[1] Edward I Altman. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4):589–609, 1968.

[2] A. C. Antonakis and M. E. Sfakianakis. Assessing naïve Bayes as a method for screening credit applicants. *Journal of Applied Statistics*, 36(5):537–545, 2009.

[3] Amir F. Atiya. Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on Neural Networks*, 12(4):929–935, 2001.

[4] B Baesens, T Van Gestel, S Viaene, M Stepanova, J Suykens, and J Vanthienen. Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, 54(6):627–635, jun 2015.

[5] Alireza Baghban, Ali Jalali, Mojtaba Shafiee, Mohammad Hossein Ahmadi, and Kwok-wing Chau. Developing an ANFIS-based swarm concept model for estimating the relative viscosity of nanofluids. *Engineering Applications of Computational Fluid Mechanics*, 13(1):26–39, 2019.

[6] William H. Beaver. Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4:71–111, 1966.

[7] Berk Bekiroglu, Hidayet Takci, and Utku Can Ekinci. Bank Credit Risk Analysis With Bayesian Network Decision. *IJAEST - INTERNATIONAL JOURNAL OF ADVANCED ENGINEERING SCIENCES AND TECHNOLOGIES*, 9(2):273–279, 2011.

[8] Leopold A. Bernstein. *Financial statement analysis : theory, application, and interpretation*. Irwin, 1993.

[9] Klaus Böcker. *Rethinking Risk Measurement and Reporting*, volume II. Risk Books, London, 2010.

[10] Andrew P. Bradley. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159, jul 1997.

[11] Xinying Zhang Chong Wu, Yingjian Guo and Han Xia. Study of Personal Credit Risk Assessment Based on Support Vector Machine Ensemble. *International Journal of Innovative Computing, Information and Control*, 6(5):2353–2360, 2010.

[12] R. H. DAVIS, D. B. EDELMAN, and A. J. GAMMERMAN. Machine-learning algorithms for credit-card applications. *IMA Journal of Management Mathematics*, 4(1):43–51, 1992.

[13] N Davutyan and S Özar. A credit scoring model for Turkey's micro & small enterprises (MSE's). In *13th Annual ERF Conference*, pages 16–18, Kuwait, 2006.

[14] Peter R. Demerjian. Financial Ratios and Credit Risk: The Selection of Financial Ratio Covenants in Debt Contracts. *AAA 2007 Financial Accounting & Reporting Section (FARS)*, 2007.

[15] Vijay S. Desai, Jonathan N. Crook, and George A. Overstreet. A comparison of neural networks and linear scoring models in the credit union environment. *European Journal of Operational Research*, 95(1):24–37, nov 1996.

[16] Douglas Diamond. Financial intermediation and delegated monitoring. *Review of Economic Studies*, 51(3):393–414, 1984.

[17] Ilia D. Dichev and Douglas J. Skinner. Large-sample evidence on the debt covenant hypothesis. *Journal of Accounting Research*, 40(4):1091–1123, 2002.

[18] Alaa El-Shazly. Financial Distress and Early Warning Signals: A Non-Parametric Approach with Application to Egypt. In *9th Annual Conference of the Economic Research Forum*, number October, pages 1–25, Emirates, 2002.

[19] D J Hand and W E Henley. Statistical Classification Methods in Consumer Credit Scoring: a Review. *Royal Statistical Society*, pages 523–541, 1997.

[20] James A Hanley and Barbara J McNeil. The Meaning and Use of the Area under a Receiver Operating (ROC) Curvel Characteristic. *Radiology*, 143(1):29–36, 1982.

[21] Martin F Hellwig. Risk aversion and incentive compatibility with ex post information asymmetry. *Journal of Economic Literature*, 438:1–25, 1998.

[22] Jih Jeng Huang, Gwo Hshiung Tzeng, and Chorng Shyong Ong. Two-stage genetic programming (2SGP) for the credit scoring model. *Applied Mathematics and Computation*, 174(2):1039–1053, mar 2006.

[23] Huseyin Ince and Bora Aktan. A comparison of data mining techniques for credit scoring in banking: A managerial perspective. *Journal of Business Economics and Management*, 10(February 2015):233–240, 2009.

[24] Michael Jacobs and Nicholas M Kiefer. The Bayesian approach to default risk: A guide. *Center for Analytical Economics*, (March 2010), 2013.

[25] Vita Jagric, Davorin Kracun, and Timotej Jagric. Does non-linearity matter in retail credit risk modeling? *Finance a Uver - Czech Journal of Economics and Finance*, 61(4):384–402, 2011.

[26] Liu Jie and Song Bo. Naive Bayesian Classifier Based on Genetic Simulated Annealing Algorithm. *Procedia Engineering*, 23:504–509, jan 2011.

[27] J. W. Kay and D. M. Titterington. *Statistics and Neural Networks Advance at the Interface*. Oxford University Press, 2000.

[28] Amir E. Khandani, Adlar J. Kim, and Andrew W. Lo. Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34(11):2767–2787, nov 2010.

[29] Adnan Khashman. Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. *Expert Systems with Applications*, 37(9):6233–6239, 2010.

[30] HC Koh, WC Tan, and CP Goh. A two-step method to construct credit scoring models with data mining techniques. *International Journal of Business . . .* , 1(1):96–118, 2006.

[31] K Komorád. *On credit scoring estimation*. PhD thesis, Humboldt University, 2002.

[32] Adel Lahsasna, Raja Noor Ainon, and Teh Ying Wah. Credit scoring models using soft computing methods: A survey. *The International Arab Journal of Information Technology*, 7(2):115–123, 2010.

[33] Tian-Shyug Lee and I-Fei Chen. A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with Applications*, 28(4):743–752, may 2005.

[34] Pawel Lewicki and Thomas Hill. *Statistics : Methods and Applications*, volume 1st. 2006.

[35] Russell James Lundholm and Richard G. Sloan. *Equity valuation and analysis*.

[36] Dominik Maltritz and Alexander Molchanov. Economic Determinants of Country Credit Risk: A Bayesian Approach. In *12th New Zealand Finance Colloquium*, 2008.

[37] D. Martens, B.B. Baesens, and T. Van Gestel. Decompositional Rule Extraction from Support Vector Machines by Active Learning. *IEEE Transactions on Knowledge and Data Engineering*, 21(2):178–191, 2009.

[38] Khalil Masmoudi, Lobna Abid, and Afif Masmoudi. Credit risk modeling using Bayesian network with a latent variable. *Expert Systems with Applications*, 127:157–166, aug 2019.

[39] Hamadi Matoussi and Aida Krichène Abdelmoula. Credit-risk evaluation of a Tunisian commercial bank: logistic regression vs neural network modelling. *International Journal of Accounting & Information Management*, 19(2):ijaim.2011.36619baa.005, jun 2011.

[40] Ricardas Mileris. Estimation of loan applicants default probability applying discriminant analysis and simple Bayesian classifier. *Economics and management*, 15(1):1078–1084, 2010.

[41] Antonietta Mira and Paolo Tenconi. Bayesian estimate of credit risk via MCMC with delayed rejection. *Stochastic Analysis, Random Fields and Applications IV*, (January 2003):277–291, 2004.

[42] T M Mitchell. Generative and Discriminative Classifiers : Naive Bayes and Logistic Regression Learning Classifiers Based On Bayes Rule. In *Machine Learning*, page 432. McGraw-Hill Education, 2005.

[43] M.D. Odom and R. Sharda. A neural network model for bankruptcy prediction. *1990 IJCNN International Joint Conference on Neural Networks*, pages 163–168 vol.2, 1990.

[44] James A. Ohlson. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1):109, 1980.

[45] Krisna G. Palepu, Paul M. Healy, Victor L. Bernard, Sue Wright, Michael Bradbury, and Philip Lee. *Business Analysis and Valuation Using Financial Statements*. 2000.

[46] Belief Revision, Financial Distress, Sumit Sarkar, and Ram S. Sriram. Bayesian Models for Early Warning of Bank Failures. *Management Science*, 47(11):1457–1475, 2001.

[47] Bernard Rosner. *Fundamentals of biostatistics*. 2012.

[48] David E. Rumelhart and James L. McClelland. An interactive activation model of context effects in letter perception: II. The contextual enhancement effect and some tests and extensions of the model. *Psychological Review*, 89(1):60–94, 1982.

[49] Okan Veli Safakli. Credit Risk Assessment for the Banking Sector of Northern Cyprus. *Banks and Bank System*, 2:21, 2007.

[50] Clifford W. Smith and Jerold B. Warner. On financial contracting. An analysis of bond covenants. *Journal of Financial Economics*, 7(2):117–161, 1979.

[51] Neda Soltani Halvaiee and Mohammad Kazem Akbari. A novel model for credit card fraud detection using Artificial Immune Systems. *Applied Soft Computing*, 24:40–49, 2014.

[52] Kent A. Spackman. Signal detection theory: valuable tools for evaluating inductive learning. *Proceeding Proceedings of the sixth international workshop on Machine learning*, pages 160–163, 1989.

[53] A. Steenackers and M. J. Goovaerts. A credit scoring model for personal loans. *Insurance Mathematics and Economics*, 8(1):31–34, mar 1989.

[54] Thomas Stibor. A study of detecting computer viruses in real-infected files in the n-gram representation with machine learning methods. *23rd International Conference on Industrial Engineering and Other Applications of Applied Intelligent Systems*, 6096 LNAI(PART 1):509–519, 2010.

[55] Lili Sun and Prakash P. Shenoy. Using Bayesian networks for bankruptcy prediction: Some methodological issues. *European Journal of Operational Research*, 180(2):738–753, jul 2007.

[56] Dennis Sweeney, David Anderson, and Thomas Williams. *Statistics for Business and Economics*. Thomson Learning EMEA, London, 7 edition, 2007.

[57] Anjan V Thakor. The Financial Crisis of 2007 – 2009 : Why Did It Happen and What Did We Learn ? *Review of Corporate Finance Studies Advance Access*, 4(2):1–51, 2015.

[58] Lyn C. Thomas, David B. Edelman, and Jonathan N. Crook. *Credit Scoring and Its Applications*. Society for Industrial and Applied Mathematics, jan 2002.

[59] Chih-Fong Tsai. Combining cluster analysis with classifier ensembles to predict financial distress. *Information Fusion*, 16:46–58, mar 2014.

[60] Gang Wang, Jinxing Hao, Jian Ma, and Hongbing Jiang. A comparative assessment of ensemble learning for credit scoring. *Expert Systems with Applications*, 38(1):223–230, jan 2011.

[61] David West, Scott Dellana, and Jingxia Qian. Neural network ensemble strategies for financial decision applications. *Computers & Operations Research*, 32(10):2543–2559, oct 2005.

[62] Dorota Witkowska, W Kaminski, Krzysztof Kompa, and Iwona Staniec. Neural networks as a supporting tool in credit granting procedure. *Information Technology for Economics & Management*, 2(1), 2004.

[63] Jiří Witzany. *Credit Risk Management Pricing, Measurement, and Modeling*. Springer, 2017.

[64] CL Wu and KW Chau. Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis. *Journal of Hydrology*, 399(3-4):394–409, 2011.

[65] Bee Wah Yap, Seng Huat Ong, and Nor Huselina Mohamed Husain. Using data mining to improve assessment of credit worthiness via credit scoring models. *Expert Systems with Applications*, 38(10):13274–13283, sep 2011.

[66] Zaher Mundher Yaseen, Sadeq Oleiwi Sulaiman, Ravinesh C. Deo, and Kwok Wing Chau. An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction. *Journal of Hydrology*, 569(August 2018):387–408, 2019.

[67] Jozef Zurada, Niki Kunene, and Jian Guan. The Classification Performance of Multiple Methods and Datasets: Cases from the Loan Credit Scoring Domain. *Journal of International Technology and Information Management*, 23(1), 2014.