Financial Internet Quarterly 2024, vol. 20 / no. 2, p. 62-76
DOI: https://dx.doi.org/10.2478/fiqf-2024-0012

Financial Internet Quarterly

MODELLING FINANCIAL VARIABLES USING NEURAL NETWORKING TO ACCESS CREDITWORTHINESS

Prashant Ubarhande[1], Arti Chandani[2], Mohit Pathak[3], Reena Agrawal[4], Sonali Bagade[5]

Abstract

This study examines the existing credit rating methodology proposed in the literature to explore the development of a new credit rating model based on the financial variables of the enterprise. The focus is on the period after the financial crisis of 2018. This study aims to develop a credit rating model using neural networking and tests the same for its accuracy. The goal of this study is to address the issue brought up by previous research on subjectivity in the data used to determine creditworthiness. The database for the study includes financial data up to July 2022 from December 2018. A model is created to assess an enterprise's creditworthiness using neural networking. This study first evaluated the existing credit rating models proposed in the literature. Next, based on financial data and neural networking, a model is developed. It was evident that the model developed in this study has an accuracy of 85.16% and 76.47% on train and test data respectively. There exist several models to determine the creditworthiness of borrowers but all failed to address the concern of subjectivity in the data. The model created in this study made use of cutting-edge technology such as neural networking and financial data. This paper's unique approach and model construction based on a comparison of existing models is rare in the literature and justify the originality of this paper with a practical value at the global level.

JEL classification: C45, E51
Keywords: Creditworthiness, Credit Rating Model, Neural Networking, Test Data, Accuracy Model testing

Received: 16.12.2023 Accepted: 15.05.2024

Cite this:
Ubarhande P., Chandani A., Pathak M., Agrawal R. & Bagade S. (2024). Modelling financial variables using neural networking to access creditworthiness. Financial Internet Quarterly 20(2), pp. 62-76.

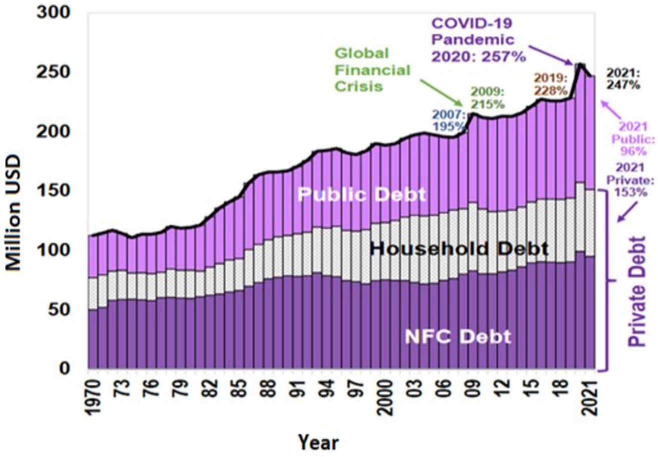
© 2024 Prashant Ubarhande et al., published by Sciendo. This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 3.0 License.

[1] Symbiosis Centre for Distance Learning, Pune, India, e-mail: ubarhande.prashant@gmail.com, ORCID: https://orcid.org/0000-0001-5070-0177.
[2] Jaipuria Institute of Management, Lucknow, India, e-mail: arti.chandani@jaipuria.ac.in, ORCID: https://orcid.org/0000-0001-7662-7967.
[3] International Management Institute Kolkata, India, e-mail: mohitpathak418@gmail.com, ORCID: https://orcid.org/0000-0002-2109-7759.
[4] Jaipuria Institute of Management, Lucknow, India, e-mail: dr.agrawal.reena@gmail.com, ORCID: https://orcid.org/0000-0002-6680-8251.
[5] Symbiosis Institute of Business Management, Hydrabad, India, e-mail: bagade.sonali21@gmail.com, ORCID: https://orcid.org/0000-0002-6269-5034.

Introduction

Debt is preferred for financing by businesses and organizations for several reasons such as interest payments are tax-deductible, debt financing helps to retain ownership and control, and it has a fixed repayment term along with interest rate certainty. These facts make debt a popular source of finance (Berman & Knight, 2009). In addition to the advantages of debt, Figure 1 shows an increasing trend in international debt volume (in trillion US dollars). It is evident that the amount of debt is rising. The International Monetary Fund (IMF, 2022) reports that since 1970, there has been an increase in the amount of private debt, loans, and debt securities.

Figure 1: Global Public and Private Debt



Source: IMF Global Debt Database, 2022.

Banks have the maximum share of debt and companies also rely on banks as a major source of capital (Judge, 2022). Banks, while approving debt to the borrowers, take due care in assessing the debt applications because if their borrowers are not able to repay the debt, those will be categorized as default. This leads to a rise in Non-Performing Assets (NPAs) for the lenders. To ensure that the risk of default is minimal, banks should have well-defined lending rules regarding the creditworthiness of the borrower (Khandelwal & Modi, 2021).

Credit ratings assigned by credit rating agencies (CRAs) indicate creditworthiness. Furthermore, a thorough evaluation of the borrowers' creditworthiness is necessary to keep such debt from becoming defaulted upon or turning into non-performing assets (NPAs). CRAs can help with this by giving borrowers credit rates through their independent evaluation of creditworthiness (Ali & Javid, 2015). Through the use of a specified mechanism and the credit rating model, CRAs assess borrowers' creditworthiness. In the finance sector, creating a model for credit rating is a complicated and vital task. Credit rating models are used to assess the creditworthiness of individuals, businesses, or financial instruments, and they significantly contribute to decision-making for lenders, investors, and other financial institutions.

Literature review

When Moody started evaluating bonds in the United States in 1909, the rating system got its start. Four rating agencies presently control the global scene: Moody's, Standard & Poor's, Duff & Phelps, and Fitch Ratings. These credit rating agencies (CRAs) identify three key factors that influence credit rates: political risk, cash flow redirection risk, and the capacity to attract private investment. Fitch Ratings only entered the sovereign rating business in 1975. Credit rating models aid the process of debt by determining the creditworthiness of borrowers. These models are commonly developed using statistical techniques (Thomas, 2000). Statistical techniques are synergistic in linear systems. The accuracy of such models decreases if there is a nonlinear relationship between variables considered to construct a credit rating system (Huang et al., 2004).

A structured framework based on financial variables was proposed in the late 20th century by Black and Scholes (Black & Scholes, 1973) and Merton (Merton, 1974). Credit rating models cannot predict defaults with 100% accuracy; however, models based on financial data are strong and irreplaceable (Agarwal & Taffler, 2008). The performance of these models can be enhanced by employing novel methods and implanting relevant data. After this, literature in the credit rating domain has been enriched in terms of many innovative methodologies to calculate credit rates.

According to Abiyev (2014), Artificial Neural Networks (ANN) can create more sophisticated neural networks and increase model accuracy. Creditworthiness is frequently predicted using Support Vector Machines (SVM), yet these models may have monotonicity issues. A monotonicity-constrained model was created by Chen and Li (2014) to solve this issue and enhance the model's performance over conventional SVM. Developing a sound loan pricing system with the use of a credit rating model is one of the most crucial components of credit management (Zhang & Chi, 2018). Lenders utilize the rates allotted to various organizations as a means of defining the borrower's risk (Dror, 2022). CRAs only make a little contribution to structured financial products. Therefore, the dependency on CRAs to fetch credibility must be defined by lenders (Tanja et al., 2019). Most studies show that credit rates and the tendency to default are inversely proportional (Basil & Elgammal, 2013). To determine creditworthiness, distance-to-default (DD) and probability-of-default (PD) can be compared with total assets (Edmund et al., 2013). Further, in subsequent years, researchers across the globe contributed to this domain and suggested the gaps in existing credit rating methodologies adopted by various CRAs. Seminal Work on Credit Rating Models and Methodologies highlighted the historical development of the credit rating domain and suggested gaps in the existing system. Allen et al. (2006) suggested requirement of a supportive credit ratings model with an objective to provide solution to the companies that defaulted, despite having good credit ratings. Use of financial data in creating a credit rating model is supported by Brocardo (2017). They suggested that creditworthiness can be accurately predicted if the credit rate is based on financial data because the financial data ensures the objectivity. Further, in subsequent years, various researchers have advocated the requirement of determining credit default based on advanced techniques and objective data. For this study, the credit rating models proposed post-2018 financial crises are considered. After any financial crisis, the expectations from credit rating agencies rise. On the other hand, credit rating agencies shall be held responsible for the crisis, which occurred due to debt market transactions or failure. 11 credit rating models were proposed in the literature post-2018 financial crisis across the globe. Table 1 represents detailed information and a critical analysis of these models.

Table 1: Credit rating models proposed in post-2018 financial crisis literature

Name of the Model	Author(s)	Country	Details of Model
Humen Area Textile and Garment Enterprises Model	(Guanzhi et al., 2018)	China	Limited to the textile industry and considers the PEST analysis of the textile industry only. This model is suitable only for textile companies in Humen, China.
Customer Number Bell-shaped Distribution Model	(Zhang & Chi, 2018)	China	Based on the weight assignment to the variables as per lenders, perception leaves scope for subjectivity.
Linear Discriminant Model	(Habachi & Benbachir, 2019)	Morocco	The model only suggested adding more variables to the existing system. The reliability of information is not addressed.
Fuzzy Inference System Environment Model	(Yazdi et al., 2019)	Iran	Moody’s model is considered as a base and the opinion on variables considered by Moody’s model is selected. This model proposed a refinement of Moody’s methodology.
Loss Given Default (LGD) Model	(Baofenga et al., 2019)	China	Calculates the loss to the banks due to default. Suitable to calculate the effectiveness of the existing model but with post default recognition.
Hybrid Model	(Li et al., 2019)	China	A robust model is not proposed, a combination of two separate models is proposed. One of the models failed to address a biased data concern.
ELECTRE TRI- Model	(Angilella & Mazzu, 2019)	Italy	The model is based on unrealistic assumptions. Models neglected the discordance phase of the borrower.
CDS Model	(Abid et al., 2020)	Tunisia	The model is developed by assuming that the real-world probabilities are less than the risk-free environment. This assumption is unrealistic. Also, the model is limited to the Thomson Reuters StarMine rating process.
Optimal Discriminant	(Zhanjiang	China	Supports small enterprises by comparing the industry

Genetic Backpropagation Neural Network Model	(Peng, 2021)	China	Supports the neural networking methodology. The major limitation is the subjectivity of the variables considered.
A Machine Learning Approach	(Wu et al., 2021)	Hong Kong	Studied the supply chain information only.

Source: Author's representation.

Research gaps

Recent literature revealed that the credit rating model addressing the constraints of subjectivity in the data used to develop it is rare. In this study, an attempt is made to develop a credit rating model using the financial performance of the companies to address subjectivity concerns. A study of existing credit rating models proposed in credit rating literature suggests the need for such a robust model based on advanced technologies and objective data. Credit markets' importance in modern economic systems has expanded gradually, notably in the last two decades. Similarly, the necessity of accurate credit risk assessment in financial markets has grown over time, and hence the role and performance of CRAs - as the primary suppliers of credit risk assessment - have been scrutinized more than ever before. After studying the literature, the following research gaps were identified:

1. There is scope to address the concern of subjectivity in the data while using it to construct a credit rating model,
2. A model based on purely objective inputs and techniques like neural networks has not been proposed.

The limitations of various models proposed in the literature enumerated in Table 1 were understood as the scope for the development of a model using advanced techniques such as neural networking and the financial variables as enumerated in Table 2.

The findings of the study suggest that using neural networks (NN) the model is capable of predicting the creditworthiness of the borrowers with an accuracy of 76.47%. We also used the data of Credit Suisse which was declared bankrupt and the model shows an 88% default probability for Credit Suisse for year 2021 and 2022. The findings and models can be used by banks, financial institutions, and lenders to determine the creditworthiness of borrowers. This can be used by credit rating agencies before assigning ratings to the companies.

Methodology

identified various financial variables from the literature which can be considered for the model development. These variables include Earnings after Tax, Earnings per Share, Fixed Assets Total Equity, Amount of Debt, Leverage Ratio (Total Liability/Total Assets), Current Ratio (Current Assets/Current Liabilities), Return on Sales (Earning After Tax/Net Sales), Total Asset Turnover (Total assets/Net Sales) and Firm Size (log Total Assets).

Earnings after tax (EAT) and the earnings per share (EPS) represent the share of the owners and play an important role in determining the financial health of an organization (Li et al., 2019). An organization having higher EAT tends to pay its debts properly and the probability of default is remote (Balikcioglu & Yilmaz, 2019; Camanho et al., 2020). Distributing dividends by the organization represents paying the cost of debt and the same is done as a post-interest payment. A positive EPS represents the ability to pay debts (Czarnitzki & Kraft, 2004). Assets owned by an organization and the liabilities of an organization can be studied to show creditworthiness as assets and liabilities represent the capacity of an organization to earn economic benefits in the future. These assets can be studied in various forms such as fixed assets (Hung et al., 2013), leverage ratios which represent the relationship between total liability and total assets (Raghunathan & Varma, 1992), current ratio (McCue et al., 1990), total asset turnover which represents the value of assets per unit net sales (Bhattacharya & Sharma, 2019) and firm size which is the algorithm of total assets (Li et al., 2019). The amount of debt and equity represents the capital structure of the organization and the risk associated with shareholders as well as lenders (Gupta, 2021). Total equity (Darren, 2006) and debt (Lee, 1993) are crucial variables in determining the creditworthiness of an organization (Chen & Chang, 2020). Return on sales is a financial ratio representing the efficiency of an organization to generate profits from its investments and contributes significantly to creditworthiness (Camanho et al., 2020). The subjectivity of the variables considered in determining the creditworthiness of the borrower is a persistent concern to lenders. The use of financial variables addresses the subjectivity constraint and results in reliable information. This subjective concern lowers the accuracy of the credit rating model. Therefore, in this study, a credit rating model based on financial variables and derivatives of these financial variables such as financial ratios are considered. Factors identified in Table 3 can be used as input for determining default probabilities vis-à-vis the creditworthiness of borrowers (Tahmoorespour et al., 2022).

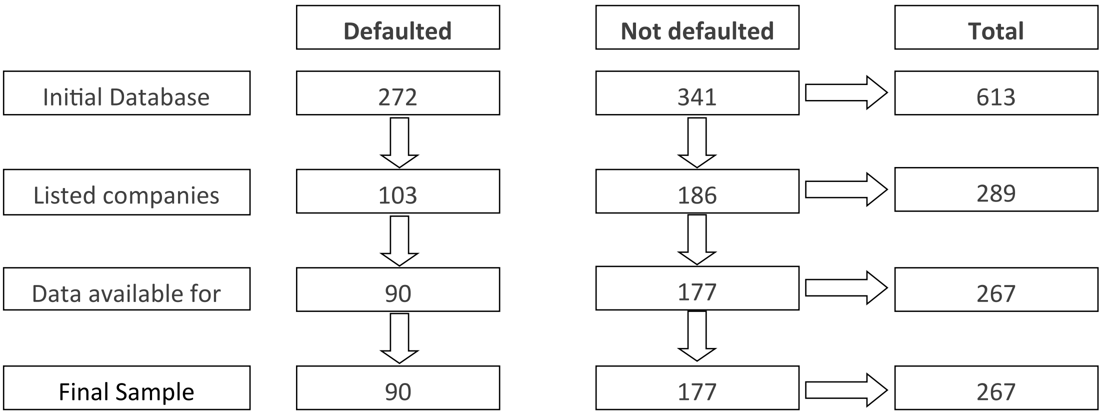
The credit rating model developed in this study shall be used to predict the default probabilities of companies. This model is developed using data collected for the defaulted and non-defaulted companies rated by seven credit rating agencies in India from the year 2018. The following steps enumerate the flow of activities in developing this credit rating model.

Sampling technique

Companies that have both defaulted and not defaulted over the study period make up the sample for this research. Companies that are rated by Indian-registered CRAs are taken into account when creating samples, and 613 of such companies were found. A sample of 90 defaulted and 177 non-defaulted enterprises was created using purposive sampling. Two predetermined criteria were used to identify companies: listing status and instrument size (Uzir et al., 2021).

From 613 total companies identified as population, 272 have defaulted, and 341 did not default. These companies were further filtered using two criteria: listing status (Chandera & Setia-Atmaja, 2020)and non-zero instrument size (Uzir et al., 2021). Of 613 companies, 289 were listed on stock exchanges in India. 22 such listed companies were not included in the final sample as the complete financial data required for this study was not available for these companies. All-in-all, a final sample of 267 companies was filtered out using purposive sampling as represented in Figure 2.

Figure 2: Sample construction



Source: Author's creation.

Class imbalance is typically defined as having a majority of excellent credit samples and a minority of bad credit samples. The imbalance ratio (IR), which is the ratio between the majority and minority samples, is typically more than 1 in popular credit datasets. It is challenging for classification models to learn the characteristics of minority classes due to a lack of minority class samples, and the prediction results are typically biased toward the majority classes (Brown & Mues, 2012). For the present study the imbalance ratio is 1.96.

Software used in the study

Microsoft Excel was used to arrive at a sample of this study by filtering the data of companies retrieved from various CRAs. The financial data of these companies was collected and recorded in Microsoft Excel. Analysis of the data collected was done using SPSS (IBM SPSS Statistics 28.0.0.0) and R Studio. The basic technique used in this study was neural networking which is a data analysis tool. Neural networking is based on artificial intelligence and trains computers to analyze data as per the human brain and involves interconnected neurons or nodes in a framework containing different layers to imitate the human brain (Giovanni & Michele, 2021).

Data analysis

Data used for the study

Financial data collected from secondary sources including the website of the National Stock Exchange for companies were divided into two classes viz. defaulted and not defaulted during the period of study. The derivatives such as financial ratios were considered input variables in developing the said model to improve the performance of the model, instead of using merely financial figures. Data was tested for normality by calculating skewness and kurtosis.

Number of hidden layers and neurons

Literature on the development of predictive models using neural networking suggests that one hidden layer is potentially adequate to meet the requirement even in the mapping of any complex function (Peng, 2021; Youjun & Liu, 2010; Du, 2018). Therefore, in this study, a single hidden layer was considered to construct a neural network model to determine the probabilities of credit default.

The number of neurons to be used in the proposed model was decided as per the technique suggested by (Sheela & Deepa, 2013) based on the number of input neurons (n).

In the present study, the input neurons (n) are 10.

Number of neurons in the hidden layer:

$$\frac{(4n^2+3)}{(n^2-8)} \tag{1}$$

Using this equation (1), the number of neurons in the hidden layer is:

$$\frac{403}{92}=4$$

Using an equation suggested in the literature by (Sheela & Deepa, 2013), the number of neurons in the hidden layer is determined as 4.

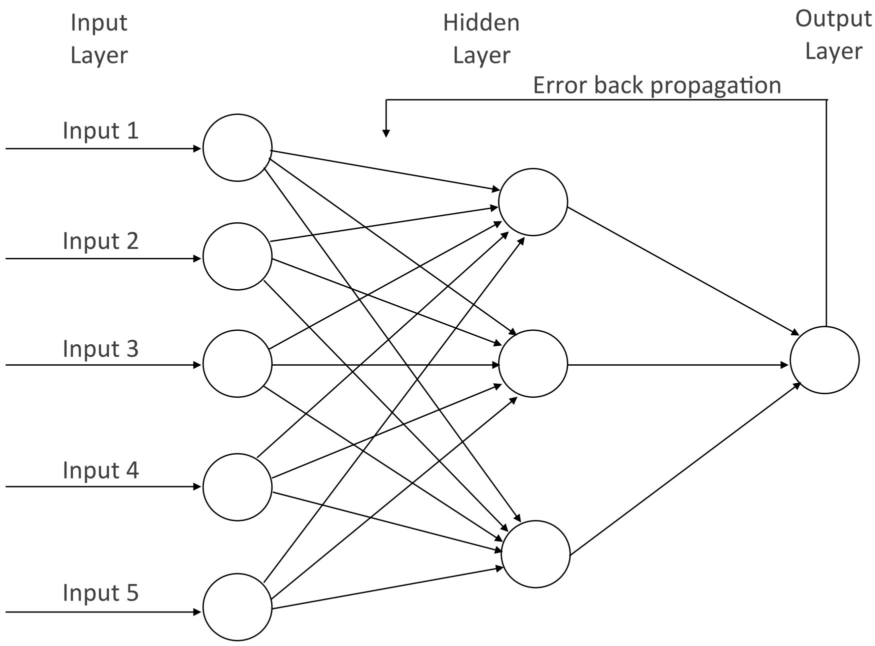
Designing the output layer

The default probabilities are to be ascertained using the suggested model. Financial data and other financial ratios make up the input data in this case. The probability of default is known for both the defaulted and non-defaulted companies in the data set; that is, 100% or 1 for the defaulted companies and 0% or 0 for the non-defaulted companies. The actual probabilities of default have been entered into the proposed model as expected output in order to train it. The model is supplied with financial data and default probabilities in order to produce the desired result. The same ranges from 0 to 1 since the output expected is the likelihood of default. The model is trained on the expected output as the actual probability of default (0 - 1). Therefore, the output layer is expected to be the imitation of the provided datasets. Hence, the output layer is selected to generate the probabilities between zero and one. Mathematically it can be expressed as:

$$Output \in \{0,1\} \tag{2}$$

The above expression suggests that the output expected will be between zero and one or 0% and 100%. Figure 3 represents a typical structure of neural networks (Yan et al., 2020).

Figure 3: Typical Neural Networking Model



Source: IEEE Access.

In the above neural networking model, the first layer is the input layer, the middle layer is the hidden layer and the last layer is the output layer. In this study, one hidden layer is considered and the output default probabilities range from zero to one.

Learning the model

Construction of data frame

Secondary data was collected from the online portals of respective companies and the National Stock Exchange, India. The ratios based on this financial data were calculated and used along with original financial figures as input. These variables and the ratios were provided as input to the model using R Studio as represented in Figure 4.

Figure 4: Data Frame with Default Rows

```
> head(df)
  ı..EAT   EPS   FA     TE  Debt         LR        CR          RoS         AT       FS Default
1  -3695 -39.90 8156 -13043 22965 -1.7607146 0.5028098 -0.017916480  7.4791199 4.221571       0
2     10   3.16  154    623     0  0.0000000 1.0486726  0.042616318  5.2730951 2.592177       0
3    100  10.96  839   9081   579  0.0637595 1.2198614  0.002497522  0.7930088 3.541579       0
4    131  17.86  260   3625  1560  0.4303448 1.6579123  0.007242498  5.1484185 4.103667       0
5    985  10.48  318   6261 18068  2.8858010 2.0779231  0.001032614  2.5353237 4.410457       0
6    106  19.60  194   2110     0  0.0000000 1.1030754  0.015806452 10.1935484 4.101747       0
```

Source: Output of R Studio.

EAT, EPS, FA, Debt, LR, CR, RoS, AT, and FS are the input variables and Default is the output provided to the neural-net package as default probability.

Partition of the data

The sample used for this study comprised 267 companies, of which 90 companies had defaulted during the study period and 177 companies had not defaulted. The entire sample was divided into two parts, viz. train data and test data. Normally, while dividing the data into such sets, the larger portion should be used for training the model and the smaller for testing the same. Therefore, in this study, the train data consists of 70% of the sample, and the test data consists of the remaining 30% of the sample (Sanusi et al., 2020). However, the model considered 182 and 85 observations as train and test data respectively.

Calculation of default probabilities

The reasonableness of weights assigned to financial indicators is a key basis of the credit rating model's reliability. An untrustworthy model may result in inappropriate decisions regarding creditworthiness, loan approvals, and other pertinent issues. In reliable systems, weights can be combined in a variety of ways, and determining reasonable weights for a single variable should be practiced. In this study, a credit rating model that accurately determines the default probabilities by assigning optimal weights to each financial variable is developed. Assigning such a weight is an ongoing procedure that ends when the intended outcome is achieved. The neural network model is predicated on the linear relationship between all the variables and starts with zero weights for each variable. After that, the weights are modified and altered several times until the intended output probability is attained. This procedure is repeated by the model for every company that is being studied, and the output is produced in the form of a credit rating model based on these weights.

$$W = \{W_1, W_2, W_3, \ldots\} \tag{3}$$

In this model, the input variables are 10 and can be represented as:

$$X = \{X1, X2, X3, \ldots\} \tag{4}$$

The hidden layer of the proposed model contains four nodes and it can be represented as:

$$Y = \{Y1, Y2, Y3, Y4\} \tag{5}$$

Using equations (3) to (5), the output layer can be expressed as:

$$\begin{aligned} 0 = {} & [(W1X1 + W2X2 + \ldots)Y1] + [(W1X1 + W2X2 \\ & + \ldots)Y2] + [(W1X1 + W2X2 + \ldots)Y3] + [W1X1 \\ & + W2X2 + \ldots)Y3] + [W1X1 + W2X2 + \ldots)Y4] \end{aligned} \tag{6}$$

The model developed in this study considered parameters provided as input and correlated them with default probabilities provided as expected output (O). Such iterations were repeated to assign weights to different parameters. These weights were improved recurrently to achieve the pre-provided result. Figure 8 represents the NN model with these iterations and the digit on each arrow specifies the weights assigned to the corresponding variable (Souza et al., 2018).

Default probabilities

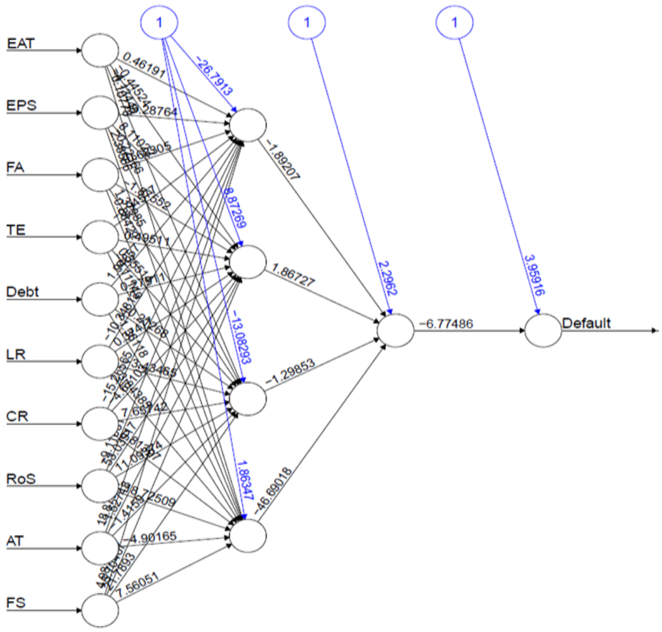
In this study, the default probabilities were calculated and the methodology suggested by (Peng, 2021) and (Huang et al., 2018) to categorize the company as default-prone was chosen. The median was taken as a base for such categorization and hence the threshold of 50% was selected to categorize a company as defaulted or non-defaulted. For instance, a company having a probability of default of more than 50% was categorized as a defaulted company. It was observed during the study that ratings assigned above the median rating, i.e., BB, have significant differences and the least probability of default. The neural network model developed in this study was trained on the similar logic that if the default probability is less than 0.5, the company is categorized as non-default and vice versa. Results for the first 10 companies in both training and test data sets are shown in Figure 5.

Figure 5: Default Probabilities

Train Data		Test Data	
2	0.07940651	1	0.88006794
3	0.07940651	5	0.10136751
4	0.07940651	6	0.07940651
7	0.07940651	9	0.06227245
8	0.88006794	10	0.27080262
11	0.07963946	12	0.27080262
13	0.27741924	23	0.10136751
14	0.47412440	24	0.47412440
15	0.27741924	26	0.27080263
16	0.07940651	28	0.27080262

Source: Author's Representation with R Studio.

Figure 6: The Neural Network Model



Source: Author's Representation with R Studio.

Testing the model and validation

Estimating accuracy of the model

The accuracy of the model was calculated first for train data and then for test data. The model was trained using test data and then the same algorithm was applied to test data. To calculate the accuracy of the model trained and tested, a confusion matrix was used (Hu & Davis, 2006).

A confusion matrix can be used to calculate the accuracy of programmed classifications (Hu & Davis, 2006) and recommends the pattern in which the model understands the provided data (Zeng, 2020; Totsch & Hoffmann, 2021).

Accuracy for Train Data

Using train data of 182 companies, a confusion matrix was generated as per the code developed. Figure 7 represents this confusion matrix.

Figure 7: Confusion Matrix for Train Data

	Actually Positive (0)	Actually Negative (1)
Predicted Positive (0)	True Positive (TP) [113]	False Positive (FP) [23]
Predicted Negative (1)	False Negative (FN) [04]	True Negative (TN) [42]

Source: Author's Representation.

The accuracy of the model can be estimated using the formula given below (Totsch & Hoffmann, 2021; Brownlee, 2013):

$$Accuracy = \left(\frac{Correct\ \ predictions}{Total\ \ predictions}\right)*100$$

Accuracy = 85.16%

The model developed using R Studio (Metin & Tepe, 2021) can predict creditworthiness with 85.16% accuracy.

Accuracy for test data

The model further used test data of 85 remaining companies and produced the following confusion matrix as in Figure 8.

Figure 8: Confusion Matrix for Test Data

	Positive (0)	Negative (1)
Predicted Positive (0)	True Positive (TP) [54]	False Positive (FP) [14]
Predicted Negative (1)	False Negative (FN) [06]	True Negative (TN) [11]

Source: Author's representation.

Accuracy = 76.47%

The model developed using R Studio (Metin & Tepe, 2021) can predict creditworthiness with 76.47% accuracy.

According to this study, if a company's chance of default is less than 50%, the model is taught to classify it as non-default and vice versa. See Figure 7 for an example of how Company 1 and Company 8 will be classified as default in the test and train data, respectively. When significant experimental data is available and used to establish certain predictive correlations, this model performs best (Sridhar, 2008).

Model testing

The model was tested using the case of Credit Suisse for its financial data for the years 2021 and 2022. The model predicted a high probability of default i.e., 88% for Credit Suisse for both years. This suggests that financial distress recorded due to default on debts by companies can be identified and prevented by taking appropriate measures using the proposed model.

Discussions and conclusions

The stakeholders of credit rating such as credit rating agencies, regulators, and lenders play a very important role in maintaining a healthy debt market in a country. Debt default leads to financial distress and affects the entire economy. Granting loans to credible borrowers prevents debt default and subsequent financial crisis. Credit rating agencies in India adopted the traditional methods of evaluation. CRAs use Monte Carlo simulation, CRAMEL model, ratio analysis, and weighted average techniques to analyze the data of the company and based on this analysis a credit rate is assigned. In the modern era of technological advancements, CRAs should use the advanced techniques of artificial intelligence and machine learning. In this study, neural networking, an advanced technique of machine learning, is used to develop a credit-rating model.

The accuracy of the model is determined based on default probabilities generated. The accuracy is calculated using a confusion matrix and it was observed that this model has an accuracy of 85.16% and 76.47% on train and test data respectively. In this study, a model to predict the default probabilities was developed using neural networking. By assigning and re-adjusting the weights to the 10 input variables, a relationship between input variables as financial data and output as default status was established using neural networking. To construct a model, the data was divided into two sets, viz. test data (70%) and train data (30%).

Train data was used to train the model for the behavior of financial variables and ratios calculated based on the financial data collected from the websites of different CRAs, companies, NSE, and BSE. The model developed in this study generated the default probabilities after studying the test data set. In this study, a threshold of 50% (Peng, 2021) was selected to categorize the company as in default. Based on all the default probabilities obtained for the train data, the accuracy of the model was estimated using a confusion matrix. The model was tested using test data and the accuracy was calculated. It was evident that the model so developed has an accuracy of 85.16% and 76.47% on train and test data respectively.

The focal point of the present study is the applications of credit rating in debt processing to minimize debt default. Various stakeholders in the process of debt approval include lenders, credit rating agencies, borrowers, and regulators. The present study will help all these stakeholders in deciding an optimal proposal to avoid debt default by ensuring the creditworthiness of borrowers. Managers of the lender banks will benefit from the proposed credit rating model and the analysis of credit rating agencies. Managers of the lending banks can scrutinize the financial variables and predict the creditworthiness using this model. Selection of credit rating agencies can also be done using the findings of this study. Managers of credit rating agencies or credit managers will benefit from the model proposed in this research by having an alternate tool to evaluate the creditworthiness of issuer companies based on financial data which is purely objective. This will empower credit managers with an impartial source of information in the form of audited financial statements.

Borrowers can use this model for self-appraisal concerning creditworthiness. By providing the financial input to the model, the borrower can calculate their default probability or tentative rate with the help of the proposed model. Borrowers can also use the analysis of credit rating agencies given in this study to select the credit rating agency for getting themselves rated. Regulators play the role of ombudsmen in the domain of credit rating. They can use the findings and the methodology proposed to develop the credit rating model in this study to strengthen the existing credit rating models. Regulators can encourage the use of advanced techniques and limit the use of objective information in the determination of credit rates

According to Murphy (2008) and Adegbite (2018), controlling mechanisms developed within the company can be a primary solution to recovering from a default crisis. The present study proposed a credit rating model by considering the critics of existing credit rating models and addressed the limitations of the existing credit rating system. Analysis of CRAs and the CR model proposed in this study will benefit lending organizations, regulators, and borrowers to access creditworthiness and contribute to developing a sound debt market where the probability of default is inconsequential.

Contributions of the study

This study substantially contributes to the existing body of literature by its results which are very novel and new in the Indian context. In India, no credit rating model has been proposed after the financial crisis of 2018. This study enriched the existing body of literature with an additional credit rating model based on the advanced technique of machine learning. The accuracy of credit rating will increase if this model is adopted by the lenders to either verify the rates assigned by CRAs or to assign credit rates by themselves. As far as the usefulness of the proposed credit rating model is concerned, it will not be an exaggeration if we state that the lenders can adopt this model to rate their debt borrowers for creditworthiness because the accuracy of this model is high and the input variables are subjective. Existing credit rating models are based on traditional techniques. A genetic backpropagation model understands the behavior of provided inputs and relates it to the output. The system repeats this process for each set of inputs and trains the model to behave by the values provided. The credit rating model proposed in this study is trained using 182 observations and the accuracy obtained is 85.16%. All in all, the model proposed in this study shall contribute towards the development of the internal credit rating mechanism by lenders, shall be used by borrowers to assess their creditworthiness, and shall be used by policymakers to compare the performance of existing credit rating models with it to improve the credit rating system further.

Limitation and future research agenda

The study has been carried out with literature support and uses advanced techniques such as machine learning. However, there are certain limitations within which this study is performed. First, the qualitative variables are out of the scope of this study and the critical analysis carried out in this study is limited to all seven CRAs of India. Availability of the period at disposal is one of the important limitations of this study as the present study considers the financial events post-2018. Sampling techniques used in this study such as purposive sampling and census sampling have their limitations as in purposive sampling a sample is derived as per requirement of the study.

The proposed critical analysis of credit rating agencies provides insight into gaps in the existing methodologies. The present research study attempts to develop a credit rating model by addressing the constraint of subjectivity in the data by considering financial variables. Researchers can extend the development and use of the model by considering the qualitative factors after rigorous checking for the existence of subjectivity in the data to remove it. This model is developed using the data of companies rated by seven credit rating agencies exclusively within India. Researchers further can add international credit rating agencies to the list. A comparison of Indian and foreign credit rating agencies can be undertaken to enrich the findings of this study. In the future, the researchers can increase the timeframe to collect the financial data of different companies. The proposed model is constructed by considering the four neurons in the hidden layer and 10 input variables. This study can be extended by adding more input variables and trying different numbers of neurons.

References

Abid, A., Abid, F. & Kaffel, B. (2020). CDS-based implied probability of default estimation. The Journal of Risk Finance, 21(4), 399-422.

Abiyev, R.H. (2014). Credit Rating Using Type-2 Fuzzy Neural Networks. Mathematical Problems in Engineering, 2014, 1-8.

Adegbite, G. (2018). 2008 Global Financial Crisis-Ten Years After; Is Another Crisis 'Resonating'? SSRN Electronic Journal, 1-11.

Agarwal, V. & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. Journal of Banking and Finance, 32(8), 1541-1551.

Ali, S. & Javid, A. (2015). Relationship between credit rating, capital structure and earning management behaviour: Evidence from Pakistani listed firms. PIDE Working Papers, 1(121), 1-44.

Allen, D.E., Powell, R. & Singh, A. (2006). A dynamic credit ratings model. Journal of Risk Model Validation, 10(1), 59-80.

Anderegg, W.R., Callaway, E.S., Boykoff, M.T., Yohe, G. & Root, T.Y.L. (2014). Awareness of both type 1 and 2 errors in climate science and assessment. Bulletin of the American Meteorological Society, 95(9), 1445-1451.

Angilella, S. & Mazzu, S. (2019). A credit risk model with an automatic override for innovative Small and Medium-sized Enterprises. Journal of the Operational Research Society, 70(10), 1784-1800.

Balikcioglu, E. & Yilmaz, H. (2019). How Fiscal Policies Affect Credit Rates: Probit Analysis of Three Main Credit Rating Agencies' Sovereign Credit Notes. Transylvanian Review of Administrative Sciences, 56, 5-22.

Baofenga, S., Xuea, Z., Bi, W. & Yizhe, D. (2019). Credit rating and microfinance lending decisions based on loss given default (LGD). Finance Research Letters, 30, 124–129.

Basil, A.N. & Elgammal, M.M. (2013). Innovation and Credit Ratings, does it matter? UK evidence. Applied Economics Letters, 20(5), 428-431.

Berman, K. & Knight, J. (2009). Finance And Investing. Retrieved from Harvard Business Review: https://hbr.org/2009/07/when-is-debt-good (Accessed 16.12.2023).

Bhattacharya, S. & Sharma, D. (2019). Do environment, social and governance performance impact credit ratings: a study from India. International Journal of Ethics and Systems, 35(3), 466-484.

Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities. Journal of Political Economy, 81(3), 637-659.

Brocardo, M.L., Rolt, C.R., Dias, J.D., Custodio, R.F. & Traore, I. (2017). Privacy information in a positive credit system. International Journal of Grid and Utility Computing, 8(1), 61-69.

Brown, I. & Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. Expert Systems with Applications, 39(3), 3446-3453.

Brownlee, J. (2013). What is a Confusion Matrix in Machine Learning. Retrieved from Machine Learning Mastery: https://machinelearningmastery.com/confusion-matrix-machine-learning/ (Accessed 16.12.2023).

Camanho, N., Deb, P. & Liu, Z. (2022). Credit rating and competition. International Journal of Finance & Economics, 27(3), 2873-2897.

CARE. (2021). CARE's Default and Transition Study. CARE, Switzerland.

Chandera, Y. & Setia-Atmaja, L. (2020). Lending relationships, credit ratings and bank loan spreads: evidence from Indonesian listed companies. International Journal of Managerial Finance, 16(4), 455-479.

Chen, C.C. & Li, S.T. (2014). Credit rating with a monotonicity-constrained support vector machine model. Expert Systems with Applications, 41(16), 7235-7247.

Chen, M.C. & Chang, C.W. (2022). The effect of the accounting expertise of chief financial officers on corporate credit ratings. Asia-Pacific Journal of Accounting & Economics, 29(2), 451-471.

Chikoore, R., Kogeda, O. & Ojo, S. (2020). Recent Approaches to Drift Effects in Credit Rating Models. (pp. 237-253). Springer.

Czarnitzki, D. & Kraft, K. (2004). Innovation indicators and corporate credit ratings: evidence from German firms. Economics Letters, 82(3), 377-384.

Darren, K. (2006). Credit Ratings and Capital Structure. The Journal of Finance, 61(3), 1035-1072.

Dror, P. (2022). Corporate risk-taking with long-term, short-term and subordinate credit ratings. International Journal of Managerial Finance, 18(2), 398-423.

Du, Y. (2018). Enterprise credit rating based on genetic neural network. In MATEC Web of Conferences (Vol. 227, p. 02011). EDP Sciences, Les Ulis.

Edmund, A.D., Powell, R. & Singh, A.K. (2013). A dynamic credit ratings model. 20th International Congress on Modelling and Simulation, MODSIM (pp. 1291-1297). Elsevier, Australia.

Egbe, I.S., Onuora, U.R., Iteh, A.A. & Onyeanu, E.O. (2021). Effect of Economic Variables on the Financial Performance of Listed Firms Manufacturing Consumers Goods in Nigeria. Universal Journal of Accounting and Finance, 9(6), 1235-1246.

Guanzhi, L., Yang, X., Jun, W. & Tao, Y. (2018). A Theoretical Credit Reporting System based on Big Data Concept: A Case study of Humen Textile Garment Enterprises. (pp. 22-26). ICBDE, New York.

Chi, G. & Zhang, Z. (2017). Multi criteria credit rating model for small enterprise using a nonparametric method. Sustainability, 9(10), 1834.

Gupta, R. (2023). Financial determinants of corporate credit ratings: An Indian evidence. International Journal of Finance & Economics, 28(2), 1622-1637.

Habachi, M. & Benbachir, S. (2019). Combination of linear discriminant analysis and expert opinion for the construction of credit rating models: The case of SMEs. Cogent Business & Management, 6(1), 1685926.

Hardjo, K. (2023). Corporate risk-taking after changes in credit rating. International Journal of Managerial Finance, 19(1), 48-62.

Hu, Q. & Davis, C. (2006). Accurate automatic quantification of taxa-specific plankton abundance using dual classification with correction. Marine Ecology Progress Series, 306, 51-61.

Huang, X., Liu, X. & Ren, Y. (2018). Enterprise credit risk evaluation based on neural network algorithm. Cognitive Systems Research, 52, 317-324.

Huang, Z., Chen, H., Hsu, C., Chen, W. & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks:a market comparative study. Decision Support Systems, 37(4), 543–558.

Hung, K., Cheng, H. W., Chen, S.S. & Huang, Y.C. (2013). Factors that affect credit rating: An application of ordered probit models. Journal of Economic Forecasting, 16, 94-108.

IMF. (2022). IMF Global Debt Database. International Monetary Fund, Washington.

Judge, A. (2022). Do Credit Ratings Determine Capital Structure? International Journal of the Economics of Business, 29(1), 89-118.

Krystyniak, K. & Staneva, V. (2023). Is this rating worth it? The benefits of credit ratings in the dynamic tradeoff model. International Journal of Managerial Finance, 19(5), 1147-1177.

Khandelwal, V. & Modi, S. (2021). India's NPA Situation with an Emphasis on Public and Private Banks. http://dx.doi.org/10.2139/ssrn.3857444.

Kisgen, D.J. (2006). Credit Ratings and Capital Structure. The Journal of Finance, 61(3), 1035-1072.

Kraus, S., Breier, M. & Dasi-Rodriguez, S. (2020). The art of crafting a systematic literature review in entrepreneurship research. International Entrepreneurship and Management Journal, 16, 1023–1042.

Lee, S.H. (1993). Are the credit ratings assigned by bankers based on the willingness of LDC borrowers to repay? Journal of Development Economics, 40(2), 349-359.

Li, B., Li, H. & Gong, M. (2019). Establishment of a Mathematical Model for Enterprise Credit Risk Recognition and Rating Based on Hybrid Learning Algorithms. IOP Conference Series: Materials Science and Engineering, 563, p. 52036.

Li, B., Li, H. & Gong, M. (2019). Establishment of a Mathematical Model for Enterprise Credit Risk Recognition and Rating Based on Hybrid Learning Algorithms. IOP Publishing, Bristol.

McCue, M.J., Renn, S.C. & Pillari, G.D. (1990). Factors affecting credit rating downgrades of hospital revenue bonds. Inquiry, 27(3), 242-254.

Merton, R.C. (1974). On the pricing of corporate debt: the risk structure of interest rates. Journal of Finance, 29(2), 449-470.

Metin, I. & Tepe, G. (2021). Gravity model: A bibliometric analysis and detailed overview. International Journal of Business and Society, 22(1), 365 - 381.

Murphy, A. (2008). An Analysis of the Financial Crisis of 2008: Causes and Solutions. SSRN Electronic Journal, 1-28.

Orlova, S., Harper, J. & Sun, L. (2020). Determinants of capital structure complexity. Journal of Economics and Business, 110, 1-16.

Pasricha, P., Selvamuthu, D. & Arunachalam, V. (2017). Markov regenerative credit rating model. The Journal of Risk Finance, 18(3), 311-325.

Peng, H. (2021). Research on credit evaluation of financial enterprises based on the genetic backpropagation neural network. Scientific Programming, 2021, 1-8.

Pervaiz, K., Virglerova, Z., Khan, M., Akbar, U. & Popp, J. (2021). Sovereign credit ratings and asian financial markets. E&M Economics and Management, 24(1), 165-181.

Raghunathan, V. & Varma, J. (1992). CRISIL Ratings: when does AAA means B? Vikalpa, 17(2), 35-42.

Reddy, Y. V. (2000). Credit Rating: Changing Perspectives. RBI Bulletin, 54(5), 1-15.

S&P. (2022). Global Bank Country-By-Country Outlook 2023 - Greater Divergence Ahead. S&P Global, New York.

Sanusi, N.A., Moosin, A.F. & Kusairi, S. (2020). Neural Network Analysis in Forecasting the Malaysian GDP. The Journal of Asian Finance, Economics and Business, 7(12), 109-114.

Sheela, K.G. & Deepa, S.N. (2013). Review on Methods to Fix Number of Hidden Neurons in Neural Networks. Mathematical Problems in Engineering, 2013, 1-11.

Souza, P.V., Araujo, V.S., Guimaraes, A.J., Araujo, V.J. & Rezende, T.S. (2018). Method of pruning the hidden layer of the extreme learning machine based on correlation coefficient. IEEE Latin American Conference on Computational Intelligence LA-CCI (pp. 1-7). IEEE, Guadalajara- Jalisco.

Sridhar, N. (2008). Models for predicting corrosion. In: N. Sridhar, & L. Yang (Ed.), Techniques for Corrosion Monitoring (pp. 499-538). Woodhead Publishing, Sawston.

Tahmoorespour, R., Ariff, M., Abdul Karim, Y.F., Lee, K.T. & Anthony, S.D. (2023). A multi-country study of factors and threshold values affecting sovereign debt-taking behavior. International Journal of Managerial Finance, 19(3), 645-669.

Tanja, V., Riaan, D.J. & Simon, G. (2019). A motivation for banks in emerging economies to adapt agency ratings when assessing corporate credit. South African Journal of Economic and Management Sciences, 22(1), 1-11.

Thomas, L.C. (2000). A survey of credit and behavioural scoring:forecasting financial risk of lending to consumers. International Journal of Forecasting, 16(2), 149-172.

Totsch, N. & Hoffmann, D. (2021). Classifier uncertainty: evidence, potential impact, and probabilistic treatment. PeerJ Computer Science, 7(398), 1-15.

Uzir, M.U., Halbusi, H.A., Lim, R., Jerin, I., Hamid, A.B., Ramayah, T. & Haque, A. (2021). Applied Artificial Intelligence and user satisfaction: Smartwatch usage for healthcare in Bangladesh during COVID-19. Technology in Society, 67, 1-15.

Vashakmadze, T., Martirosyan, E. & Sergeeva, A. (2017). Shareholder Dilemma: What Stake to Offer to Financial Investors? Proposals Based on Financial Principles. GLOBAL AND NATIONAL BUSINESS THEORIES AND PRACTICE: BRIDGING THE PAST WITH THE FUTURE (pp. 1824-1832). Euromed Press, Rome.

Wu, H.C. & Wu, Y.T. (2016). Evaluating credit rating prediction by using the KMV model and random forest. Kybernetes -The International Journal of Cybernetics, Systems and Management Sciences, 45(10), 1637-1651.

Wu, J., Zhang, Z. & Zhou, S. X. (2022). Credit rating prediction through supply chains: A machine learning approach. Production and Operations Management, 31(4), 1613-1629.

Yan, B., Yao, H.P., Nakamura, M., Li, Z. F. & Wang, D. (2020). A case study for software quality evaluation based on SCT model with BP neural network. IEEE Access, 8, 56403-56414.

Yazdi, A.K., Hanne, T., Wang, Y.J. & Wee, H.M. (2019). A Credit Rating Model in a Fuzzy Inference System Environment. Algorithms, 12(139), 1-15.

Yoshino, N. & Taghizadeh-Hesary, F. (2015). Analysis of Credit Ratings for Small and Medium-Sized Enterprises: Evidence from Asia. Asian Development Review, 32(2), 18-37.

Youjun, L.I. & Liu, H. (2010). Credit evaluation for real estate development enterprise based on BP neural network system. International Institute of Statistics and Management Engineering Symposium. Northeast Petroleum University, Qinhuangdao City.

Yung, P. & Yip, B. (2010). Construction quality in China during transition: A review of literature and empirical examination. International Journal of Project Management, 28(1), 79-91.

Zeng, G. (2020). On the confusion matrix in credit scoring and its analytical properties. Communications in Statistics - Theory and Methods, 49(9), 2080-2093.

Zhang, Y. & Chi, G. (2018, May 14). A credit rating model based on a customer number bell-shaped distribution. Management Decision, 56(5), 987-1007.

Li, Z. & Guo, L. (2021). Credit Rating Model of Small Enterprises Based on Optimal Discriminant Ability and Its Empirical Study. Complexity, 2021, 1-12.