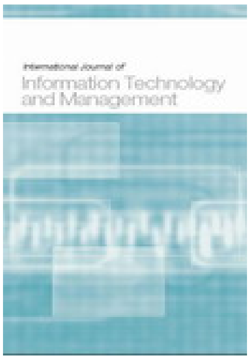


International Journal of Information Technology and Management

ISSN online: 1741-5179 - ISSN print: 1461-4111
https://www.inderscience.com/ijitm

An enterprise financial credit risk measurement method based on differential evolution algorithm

Lixia Du, Xin An

DOI: 10.1504/IJITM.2022.10051754

Article History:

Received:	14 June 2022
Accepted:	23 September 2022
Published online:	28 January 2025

An enterprise financial credit risk measurement method based on differential evolution algorithm

Lixia Du*

Department of Economic Management,
Xinzhou Teachers University,
Xinzhou, 034000, Shanxi, China
Email: lixiad@mls.sinanet.com
*Corresponding author

Xin An

College of Economics and Management,
Handan University,
Handan, 056005, Hebei, China
Email: 6554673@qq.com

Abstract: In order to reduce the time cost and risk misjudgement rate of financial information risk measurement, this paper proposes a new enterprise financial credit risk measurement method based on differential evolution algorithm. Firstly, after preprocessing the enterprise financial credit risk data and determining the location of the clustering centre, a differential evolution automatic clustering model is constructed. Secondly, according to the clustering results, the differential evolution algorithm is used to measure the basic process of enterprise financial credit risk. Finally, the improved differential evolution algorithm is used for iterative measurement to achieve enterprise financial credit risk data measurement. The experimental results show that the time cost of the proposed method for enterprise financial credit risk measurement can be controlled within 0.4 s, and the error rate is not more than 1% under the condition of 1,000 data.

Keywords: differential evolution algorithm; corporate finance; credit risks; measurement method.

Reference to this paper should be made as follows: Du, L. and An, X. (2025) ‘An enterprise financial credit risk measurement method based on differential evolution algorithm’, *Int. J. Information Technology and Management*, Vol. 24, Nos. 1/2, pp.67–77.

Biographical notes: Lixia Du received her Doctor’s degree from the Jeonju University, P.R. Korea. Currently, she teaches in the Economics and Management Department of Xinzhou Teachers University. Her research interests include financial management and fintech.

Xin An received her Doctor’s degree from the Woosuk University, P.R. Korea. Currently, she teaches in the Economics and Management College of Handan University. Her research interests include regional economics, commercial bank management and online finance.

1 Introduction

With the rapid development of information technology and the arrival of the era of knowledge economy, the business activities of enterprises are very important in the daily operation of modern companies (Hu, 2021; Dimitrova-Grajzl et al., 2022; Agarwal et al., 2020), among which the management and operation of enterprise finance is the key (Zhang, 2022). There are financing risks, safety risks, market risks, social responsibility risks, environmental risks and corporate credit risks in the independent operation, self-financing and self-reliance of enterprises (Stevenson et al., 2021; Antonelli et al., 2021). In this case, the company's credit crisis has become a very prominent problem, which has had a profound impact on the company's operation, related to the confidence of investors, borrowers, partners and customers in the company, affected the trust of the enterprise, and thus had a significant negative impact on the company's development (Sadler, 2020). Therefore, credit risk assessment and management is increasingly concerned by people, and it is the main component of financial risk management.

Tang et al. (2021) proposed a robust credit risk measurement method, using the cost-sensitive kernel method as data-driven, designed a credit risk measurement decision paradigm, solved the linear and nonlinear cost-sensitive kernels respectively, and analysed the method. The experiment found that this method has better measurement results for credit risk measurement datasets, but it does not collect data and apply real data, and does not consider issues such as measurement timeliness. Caruso et al. (2021) puts forward a new method based on mixed data clustering technology, which is used in financial credit risk measurement. It studies the correlation of financial credit risk data in both quantitative and qualitative aspects, which can effectively solve the problems of different credit risk data characteristics and large data volume. However, the accuracy and measurement time of financial credit risk measurement still need to be further studied and explored. Zhu (2020) proposes an information risk measurement method based on analytic hierarchy process (AHP) fuzzy comprehensive evaluation method. This method takes six evaluation dimensions as the starting point, designs an index system for credit risk measurement, and combines AHP and fuzzy comprehensive evaluation method to build a credit risk measurement model. However, the measurement error rate of this method is high.

In order to solve the problems of long-time measurement and high misjudgement rate in the above traditional methods, an enterprise financial information risk measurement method based on differential evolution algorithm is proposed. The overall research technical route of this method is as follows:

1. Preprocess enterprise financial credit risk data, including default data, debt data, default level, credit quality and other data. The value range of each eigenvalue is limited to [0, 1], which provides a basis for subsequent clustering operations; divide the financial and credit risk data of enterprises to clarify the positioning results of clustering centres. Define the clustering index, and realise the construction of differential evolution automatic clustering model.
2. Based on the differential evolution automatic clustering model, the differential evolution algorithm is designed to measure the financial credit risk of enterprises. On this basis, the improved differential evolution algorithm is optimised and the excellent data measurement results of enterprise financial credit risk are obtained.

3 Experimental verification, taking the time cost and false positive rate of credit measurement as experimental comparison indicators, this method is compared with traditional methods.

2 Construction of differential evolution automatic clustering model

In order to improve the effectiveness of risk measurement, a differential evolution automatic clustering model is constructed to cluster the financial data of enterprises, including the net assets, tradable shares, long-term and short-term debt and non-tradable shares.

Differential evolution automatic clustering mainly expresses the clustering quantity and clustering quality centre through a single individual, then constructs multiple individuals into groups, and then uses differential evolution method to optimise and obtain the optimal solution (Liu and Zeng, 2021; Clements et al., 2020) to improve the accuracy of credit wind power measurement. Among them, the data sample preprocessing can facilitate the subsequent clustering operation, so that the data has a relative value and research significance. Therefore, the range of each eigenvalue is limited to between, and the specific data sample preprocessing method is:

$$X(i, j) = \frac{X^*(i, j) - X_{\min}(j)}{X_{\max}(j) - X_{\min}(j)} \tag{1}$$

Among them, $X^*(i, j)$ is the original dataset of the collected credit risk data, and $X_{\max}(j)$ and $X_{\min}(j)$ are the maximum and minimum values of each column of the credit risk data.

In the differential evolution automatic clustering algorithm, chromosomes are represented by the arrangement of coding regions to judge the chromosomes and their positions, which can be used to solve the problem of clustering centre location (Duarte et al., 2019). Therefore, differential evolution automatic clustering algorithm is applied to divide the data of enterprise financial credit risk. D is used to represent the spatial dimension of the data points of enterprise financial credit risk, $K_{\max}$ is used to represent the most likely number of clusters, and m_i ($i = 1, 2, \cdots, K_{\max}$) is used to represent the i cluster centre. Then, the division result of a single individual V of the data of enterprise financial credit risk can be expressed as follows:

$$V = \left(T_1, T_2, \cdots, m_1, m_2, \cdots, m_{K_{\max}}\right) \tag{2}$$

Among them, $m_i = (m_{i,1}, m_{i,2}, \cdots, m_{i,D})$ and m_i can be used to represent the main structure of enterprise financial credit risk data, and T_i ($i = 1, 2, \cdots, K_{\max}$) is the symbol of successful selection of cluster centroid.

In the design of the differential evolution automatic clustering model, in order to ensure the clustering effect, the method of quantitative testing is used to evaluate the excellent degree of the clustering results, and the clustering results are evaluated mainly by transforming the clustering validity indicators into the form of corresponding statistical mathematical functions. The determined clustering indicators generally include two objectives (Assef et al., 2019), namely, determining the best partition and data cluster (Kinateder et al., 2021). Therefore, the design clustering index is:

$$S_{i,q} = \left[\frac{1}{N} \sum\nolimits_{X \in C_i} \| X + m_i \|_2^q \right]^{1/q} \tag{3}$$

$$d_{ij,t} = \left\{ \sum\nolimits_{p=1}^{d} \left| m_{i,p} - m_{j,p} \right|^t \right\}^{1/t} = \| m_i - m_j \| \tag{4}$$

$$R_{i,qt} = \max\nolimits_{j \in K, j \neq i} \left\{ \frac{S_{i,q} + S_{j,q}}{d_{ij,t}} \right\} \tag{5}$$

$$DB(K) = \frac{1}{K} \sum\nolimits_{i=1}^{K} R_{i,qt} \tag{6}$$

Among them, X refers to the collection of all the data of the enterprise's financial credit risk, *mi* refers to the centroid of the i data clustering, q can only be taken as integer, and $t \geq 1$, q and t are not related, C_i refers to all the data types of the dataset, and K represents the total number of clusters.

Therefore, the differential evolution automatic clustering efficiency model is established as follows:

$$m_i = \frac{1}{N_i \sum\nolimits_{x_j \in C_i} x_j}. \tag{7}$$

3 Measurement of financial credit risk based on improved differential evolution algorithm

3.1 Basic process of differential evolution algorithm for enterprise financial credit risk measurement

The differential evolution algorithm flowchart is shown in Figure 1.

Differential evolution algorithm is a global heuristic search algorithm. The specific process of differential evolution algorithm is:

Step 1 Dividing the enterprise financial credit risk data into two parts, including the data training set to indicate the training degree and training behaviour of the feature data, and the test set to analyse the effective behaviour of the data, so as to determine the overall effect of the data. Unify the result data after training, confirm the correlation degree between different data, and construct the initial confidence rule base.

Step 2 Train the parameters according to the cluster validity model in Section 2.3. The steps are as follows:

1 Coding and initialisation

Assume that the definition parameter vector individual is as follows:

$$X_i^g = \left[x_{i,1}^g, x_{i,2}^g, \cdots, x_{i,D}^g \right], i = 1, 2, \cdots, NP \tag{8}$$

Formula (8) is a g iterative i individual. When a vector individual has random variation, it is necessary to expand the population size to NP and set D components to satisfy the operation of random feasible region. The initial random group is constructed as follows [equation (9)]:

$$x_{i,j}^{o} = x_{\min,j} + rand_{i,j}(0,1)\left(x_{\max,j} - x_{\min,j}\right) \tag{9}$$

Among them, $x_{\max,j}$ and $x_{\min,j}$ are the upper and lower bounds of the j component in x_i individual, and $rand_{i,j}(0, 1)$ is the random number of [0, 1] interval, which is used to represent the random change of vector.

2 Differential variation operation

Under the condition of satisfying $NP \geq 4$, vector x_i^g will produce three values of r_1, r_2, r_3 that distinguish individuals when performing differential mutation operation. The formula for variation is as follows:

$$v_{ij}^{g} = x_{r_1 j}^{g} + 0.5\left(x_{r_2 j}^{g} - x_{r_3 j}^{g}\right) \tag{10}$$

In the expression, r_1, r_2, r_3 is a randomly generated constant and satisfies $r_1 \neq r_2 \neq r_3 \neq i$.

3 Cross-operation

The cross-formula is as follows:

$$u_i(j) = \begin{cases} v_i(j), & \text{if } rand \leq CR \\ x_i(j), & \text{otherwise} \end{cases} \tag{11}$$

4 Selective operation

Select the formula below:

$$x_i'(j) = \begin{cases} u_i, & \text{if } f(u_i) \leq f(x_i) \\ x_i, & \text{otherwise} \end{cases} \tag{12}$$

In the calculation of differential evolution, the average error ME can be used as the target index of the training set:

$$ME = \frac{1}{T}\sum_{t=1}^{T} E(t) \tag{13}$$

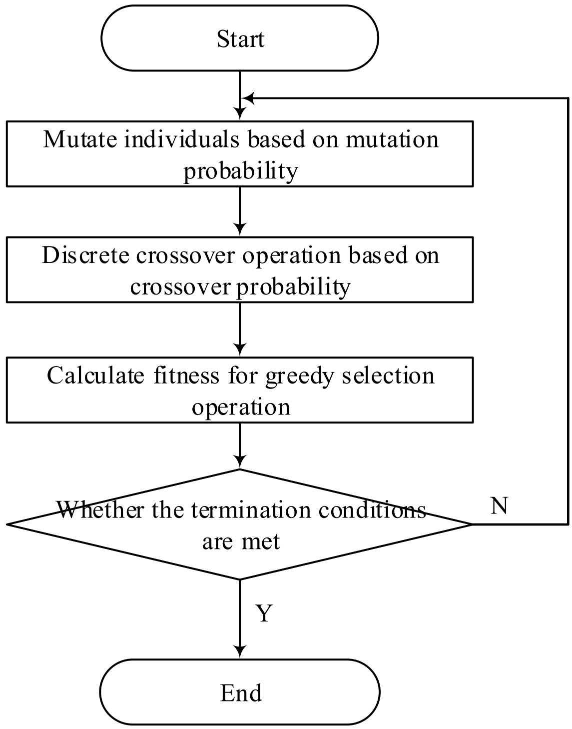
$$E = \begin{cases} 1 & \text{if } \hat{n} \neq n \\ 0 & \text{if } \hat{n} = n \end{cases} \tag{14}$$

In the formula, n is the real class of the test dataset, and n is the reasoning classification result.

In the whole training process, the training algorithm is finished until the evolution algebra of the enterprise financial credit risk data is the same as the original one.

Step 3 Enter the test dataset to validate the data content in the execution database.

Figure 1 Basic differential evolution algorithm flowchart



3.2 Improved differential evolution algorithm

At the beginning of evolution, the probability of selecting these two strategies is $p_1' = p_2' = 0.5$. In the process of evolution, after each generation K evolution, the probability p_1', p_2' of success of these two mutation strategies is calculated respectively, and then the probability of selection of generation K strategies is determined based on this probability, as follows:

$$p_1 = p_1' / \left(p_1' + p_2'\right) \tag{15}$$

$$p_2 = p_2' / \left(p_1' + p_2'\right) \tag{16}$$

In order to improve the lack of dynamics in the later stage of evolution and enhance the diversity of the population, an improved mutation algorithm is constructed based on the differential evolution algorithm. The mutation operation is changed to:

$$t_i^{gen} = \begin{cases} x_i^{gen} + F * \left(x_p^{gen} - x_q^{gen}\right), & rand < 0.5 \\ x_i^{gen} + rand * \left(x_p^{gen} - x_q^{gen}\right), & \text{other} \end{cases} \tag{17}$$

Evolution can improve the effect of data operation and enhance the transformation of data results. So, we can select the best individual behaviour to search for the best individual x_{best}.

$$x_{best} = (1+\rho) \cdot x_{best} \tag{18}$$

In the formula, ρ is a randomly distributed number between –0.5 and 0.5. If *fitness*[x'_{best}] < *fitness*[x_{best}], then $x_{best} = x'_{best}$, and vice versa, then x'_{best} replaces the worst fit.

The steps of the improved differential evolution algorithm are as follows:

Step 1 Initialise the population, and randomly select a fixed number of cluster centroid and activation threshold.

Step 2 Set the activation threshold to 0.5, determine the threshold size, when the threshold conditions can be set to m_i as the cluster centre of effective corporate financial credit risk data.

Step 3 Each loop, do.

1. A data vector X_p is generated when the data is processed once, and the distance $d(X_p, m_{i,j})$ of the effective cluster centroid can be obtained by calculating the individual volume V_i of X_p.
2. Based on $d(X_p, m_{i,j})$, allocate X_p to the minimum distance point, namely:

$$d\left(X_p, m_{i,j}\right) = \min_{\forall b \in \{1,2,\cdots,K\}} \left\{d\left(X_p, m_{i,j}\right)\right\} \tag{19}$$

3. On the basis of the traditional differential evolution method, the adaptive value of individual data is obtained by operating the enterprise's financial credit risk data by means of variation, crossover and judgement.

Step 4 Determine the individual fitness values obtained. If the fitness does not converge at the maximum number of iterations, repeat Step 3, otherwise switch to Step 5.

Step 5 Finishes optimisation and outputs metrics. Thus, the enterprise financial credit risk measurement is realised.

4 Experimental section

4.1 Experimental protocol and data

Take the small and medium-sized enterprises as the research object, carries on the sample selection in the enterprise, the different enterprise's profession distribution situation is shown in Table 1. In order to verify the validity of the algorithm, the sample data include net assets, tradable shares, short-term and long-term debts and non-tradable shares, and all the data have not been processed before the experiment. Each enterprise metric contains 2,000 pieces of data. The training rounds were set to 1,000, the iterations were 3,000, and the population size of differential evolution algorithm was set to 90. In order to verify the effectiveness of this method, the proposed method is compared with Tang et al. (2021) method and Caruso et al. (2021) method.

4.2 Performance plans and indicators

The time cost and misjudgement rate of credit risk measurement are used as the main indexes to evaluate the performance of the method. The time overhead can be used to judge the running speed of the algorithm, and the error rate can be used to test the accuracy of the algorithm to verify the performance of the proposed method.

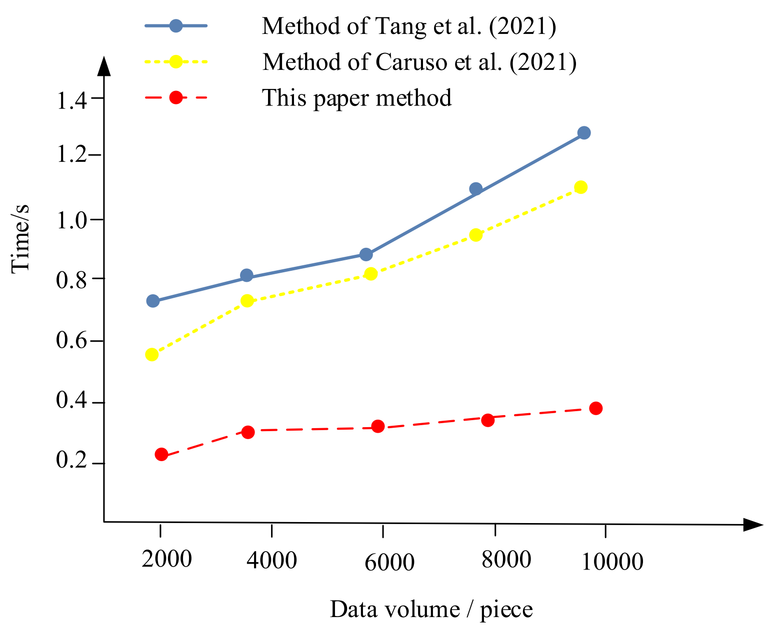
Table 1 Industry distribution of different enterprises

Industry	*Number of enterprises*	*Percentage / %*
Information technology	5	20.0
Public environmental protection	3	12.0
Construction industry	3	12.0
Manufacturing	12	48.0
Real estate	2	8.0
Total	25	100

4.3 Time cost comparison results

In order to verify the effectiveness of the proposed model, the methods of Tang et al. (2021) and Caruso et al. (2021) and the measurement method of this paper are used to measure the enterprise financial credit risk data. The test time is shown in Figure 2.

Figure 2 Time cost comparison graph (see online version for colours)

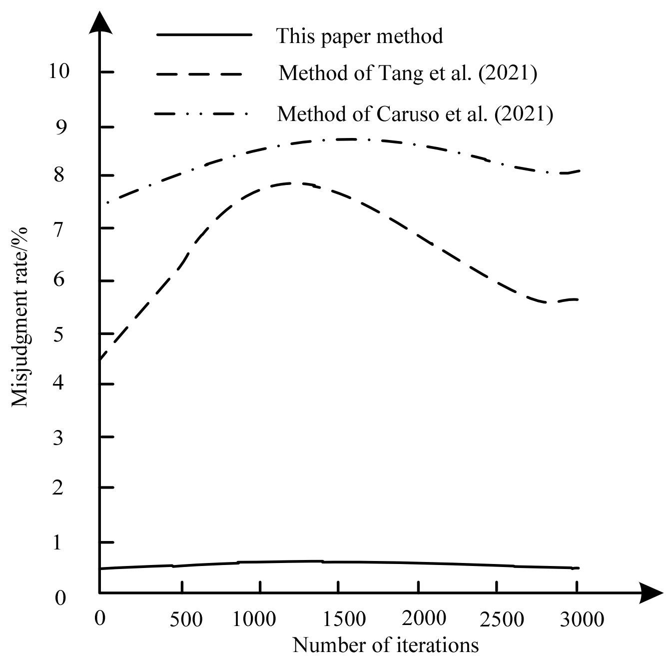


It can be seen from the time cost results that, with the increasing amount of data, the time cost of the method in this paper is always lower than the two methods in Tang et al. (2021) and Caruso et al. (2021), and the time cost is always maintained below 0.4 s, and the increasing trend of the time consumed is also significantly lower than the methods in Tang et al. (2021) and Caruso et al. (2021). This shows that the designed algorithm has high running speed and good real-time performance in practical applications. This is because this paper adopts the improved differential evolution algorithm, which can realise fast optimisation calculation, so it shortens the time cost of risk measurement.

4.4 Error rate comparison results

Under the condition of 3,000 iterations, the error rate is used as the test index of the risk measurement method to test the accuracy of the risk measurement of different methods. The specific experimental results are shown in Figure 3.

Figure 3 Comparison of error rate results



It can be seen from Figure 3 that the error rate of the designed method is significantly lower than the two methods in the literature. From the overall trend, under the condition of 3,000 iterations, the maximum error rate of the designed method is no more than 1%, while the maximum error rate of the methods in Tang et al. (2021) and Caruso et al. (2021) is 8%~9%. Therefore, it shows that this method can accurately measure the risk of enterprise financial information. The reason why this method has a low misjudgement rate is that it constructs an automatic clustering model of differential evolution to comprehensively cluster the financial related data of enterprises, and improves the differential evolution algorithm. The improved differential evolution algorithm is used to measure the financial risk of enterprises, so as to reduce the misjudgement rate of risk.

5 Conclusions

In view of the high accuracy of enterprise financial credit risk measurement and the difficulty of large-scale enterprise financial credit risk measurement, this paper proposes an enterprise financial credit risk measurement method based on differential evolution algorithm. The performance of the method is verified from both theoretical and experimental aspects. The method has low time cost and misjudgement rate when measuring financial credit risk. Specifically, under the condition of 10,000 pieces of data, the time cost of this method can be kept within 0.4 s; within 3,000 iterations, the maximum error rate of this method is no more than 1%. Therefore, it fully shows that the proposed measurement method based on differential evolution algorithm can better meet the requirements of financial information risk measurement.

Acknowledgements

This work was supported by Projects of Social Science Development in Hebei Province (No. 20210201221).

References

Agarwal, B., Jadwal, P.K. and Jain, S. (2020) 'Financial credit risk evaluation model using machine learning-based approach', *World Review of Entrepreneurship Management and Sustainable Development*, Vol. 16, No. 6, pp.576–581.

Antonelli, F., Ramponi, A. and Scarlatti, S. (2021) 'CVA and vulnerable options pricing by correlation expansions', *Annals of Operations Research*, Vol. 299, No. 1, pp.401–427.

Assef, F., Steiner, M.T., Neto, P. et al. (2019) 'Classification algorithms in financial application: credit risk analysis on legal entities', *IEEE Latin America Transactions*, Vol. 17, No. 10, pp.1733–1740.

Caruso, G., Gattone, S.A., Fortuna, F., Battista, T.D. and Parker, B.R. (2021) 'Cluster analysis for mixed data: an application to credit risk evaluation', *Socio-Economic Planning Sciences*, Vol. 73, No. 3, pp.397–404.

Clements, J.M., Xu, D., Yousefi, N. et al. (2020) 'Sequential deep learning for credit risk monitoring with tabular financial data', *Papers*, Vol. 565, No. 3, pp.278–305.

Dimitrova-Grajzl, V., Grajzl, P., Guse, A.J. et al. (2022) 'Community development financial institutions and individuals' credit risk in Indian country', *Journal of Economic Issues*, Vol. 31, No. 4, pp.713–723.

Duarte, F.D., Gama, A. and Gulamhussen, M.A. (2019) 'Credit risk, owner liability, and bank loan maturities during the global financial crisis', *European Financial Management*, Vol. 12, No. 6, pp.5797–5806.

Hu, Q. (2021) 'Credit risk prediction model of financial companies based on machine learning algorithm', *Electronic Technology & Software Engineering*, Vol. 16, No. 10, pp.146–147.

Kinateder, H., Choudhury, T., Zaman, R. et al. (2021) 'Does boardroom gender diversity decrease credit risk in the financial sector? Worldwide evidence', *Journal of International Financial Markets Institutions and Money*, Vol. 14, No. 20, pp.1–18.

Liu, P. and Zeng, Z. (2021) 'Evaluation of financial credit risk on pharmaceutical supply chain based on support vector machine', *Journal of Physics: Conference Series*, Vol. 17, No. 1, pp.12–18.

Sadler, H. (2020) 'ER2C SDMLC: enterprise release risk-centric systems development and maintenance life cycle', *Software Quality Journal*, Vol. 28, No. 3, pp.1755–1787.

Stevenson, M.P., Mues, C. and Bravo, C. (2021) 'The value of text for small business default prediction: a deep learning approach', *European Journal of Operational Research*, Vol. 295, No. 2, pp.758–771.

Tang, J., Li, J., Xu, W., Tian, Y., Ju, X. and Zhang, J. (2021) 'Robust cost-sensitive kernel method with Blinex loss and its applications in credit risk evaluation', *Neural Networks*, Vol. 143, No. 3, pp.327–344.

Zhang, A. (2022) 'The application of virtual teams in the improvement of enterprise management capability from the perspective of knowledge transfer', *PLOS One*, Vol. 17, No. 3, pp.1–14.

Zhu, H. (2020) 'Research on assessment model of commercial banks credit risk based on AHP-fuzzy comprehensive evaluation method', *Intelligent Computer and Applications*, Vol. 10, No. 4, pp.144–147.