Research Article

Enterprise Financial Risk Management Using Information Fusion Technology and Big Data Mining

Huabo Yue,1 Haojie Liao,1,2,3 Dong Li,2 and Ling Chen3

1College of Graduate Studies, Master of Management Program in Management (International Program) WALAILAK University, 222 Thaiburi, Thasala, Nakhon Si Thammarat 80160, Thailand
2Accounting and Audit School, Guangxi University of Finance and Economics, Nanning, Guangxi 530003, China
3College of Graduate Studies, Bansomdejchaopraya Rajabhat University, Bangkok10600, Thailand

Correspondence should be addressed to Haojie Liao; 2017110007@gxufe.edu.cn

Received 29 September 2021; Revised 2 November 2021; Accepted 5 November 2021; Published 15 December 2021

Academic Editor: Chin-Ling Chen

Copyright © 2021 Huabo Yue et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper aims to study enterprise Financial Risk Management (FRM) through Big Data Mining (BDM) and explore effective FRM solutions by introducing information fusion technology. Specifically, big data technology, Support Vector Machine (SVM), Logistic regression, and information fusion approaches are employed to study the enterprise financial risks in-depth. Among them, the selection of financial risk indexes has a great impact on the monitoring results of the SVM-based FRM model; the Logistic regression-based FRM model can efficiently classify financial risks; the information fusion-based FRM model uses a fusion algorithm to fuse different information sources. The results show that the SVM-based and Logistic regression-based FRM models can manage and classify enterprise financial risks effectively in practice, with a classification accuracy of 90.22% and 90.88%, respectively; by comparison, the information fusion-based FRM model presents a classification accuracy as high as 95.18%. Therefore, it is concluded that the information fusion-based FRM is better than the SVM-based and Logistic regression-based FRM models; it can integrate and calculate multiple enterprise financial risk data from different sources and obtain higher accuracy; besides, big data technology can provide important research methods for enterprise financial risk problems; SVM-based FRM model and Logistic regression-based FRM model can well classify enterprise financial risks, with relatively high accuracy.

1. Introduction

Today, the fast socio-economic development features a new technological revolution by Information Technology (IT), such as big data [1–3], cloud computing [4–6], and the Internet of Things (IoT), which is transforming people’s life, work, and the society towards informatization and intellectualization; moreover, thanks to IT [7–9] and the e-commerce industry based on it, the connection in between people and enterprises is getting ever closer despite their geographic distances and cultural or political barriers. In particular, statistics of the development status of e-commerce enterprises reveals that financial situation and financial risks [10–12] can determine how far and high an enterprise can develop; thus, the research of these factors has great practical significance. The Big Data Mining (BDM) approach might be just born to analyze enterprise financial data with its excellent identification effect, thereby being able to give early warning against enterprise financial risks [13–15].

To improve the level of enterprise Financial Risk Management (FRM) [16], scholars have conducted numerous studies and developed some effective theoretical methods. The financial risk of an entity can be defined as the entity’s possibility of money-losing in financial activities. In terms of financial risk theory analysis, there are large numbers of practical cases, some foreign scholars believe that the causes behind the enterprise financial risks are diverse and need specified analysis. Valaskova et al. (2018) [17] once argued that financial risks could be solved by regression analysis. Some experts pointed out that the economic situation was a critical enterprise financial risk factor [18, 19]. Thereupon,
some scholars put forward risk minimization, risk transfer, risk tolerance, and risk treatment process to prevent the occurrence of financial risk problems. Zhang et al. (2021) [20] confirmed that big data and fuzzy Analytic Hierarchy Process (AHP) could efficiently and accurately monitor the financial risks. After studying the enterprise financial risks [21], some scholars believe that enterprise financial risks can be effectively reduced by combining financial risk management with the mathematical model. Domestic scholars also put forward some strategic views and research results, such as using the dynamic portfolio to prevent financial risks. The monitoring and analysis of enterprise financial risk [22–24] ultimately aim to effectively manage the risks. Chinese scholars have established and improved the enterprise internal risk prevention and management system [25–27].

To sum up, the current research on enterprise financial risk has not involved the combination of information fusion and big data technology to study enterprise financial risk. Although there have been attempts on the application of DM to enterprise financial management, they have shown great deficiencies in terms of security and classification accuracy, so there is still much room for improvement. Given these shortcomings, this paper optimizes their deficiencies. Thereupon, three models are introduced into Big Data Mining (BDM) to delve into enterprise financial risks: Support Vector Machine (SVM, SVMs)-based FRM model, Logistic regression-based FRM model, and the information fusion-based FRM model. The innovation of this paper is to combine BDM with the SVM model, Logistic regression model, and information fusion technology, separately to study the enterprise financial risks. Consequently, the combination method can well classify enterprise financial risks, with very high accuracy. The contents provide a theoretical basis for the follow-up research, which is of great significance. It is imperative for enterprises to improve their risk-bearing abilities and promote enterprise development. The technical route of this paper reads 1. The proposal of the research direction, namely the enterprise FRM research; 2. Selection of research methods, namely BDM technology, SVM-based FRM model, Logistic regression-based FRM model, and information fusion-based FRM model; 3. Research results; 4. Conclusions analysis. Figure 1 shows the technical route.

2. Monitoring Models under Different Technologies

2.1. SVM-Based FRM Model. According to the current research on enterprise FRM, financial risk indexes have a great impact on the monitoring results of the SVM-based FRM model. Researchers prefer the business operational financial data as the input index of the FRM model, which shows that enterprise financial data are universal and widely desirable. Therefore, the following indexes are chosen for the proposed enterprise financial risk model: the financial structure of the enterprise, the ratio of retained earnings to total assets, the current debt to assets ratio, the ratio of asset management, and earnings management indexes [28–30]. Figure 2 shows the specific indexes of FRM.

The basic SVM model is a linear classifier with the largest interval in feature space, which can also be extended to a nonlinear classifier by the kernel function method. SVM can be used to classify two-dimensional (2D) patterns and find the 2D plane of decision-making in vector space compared with perceptron. Both models are classification models. SVM can optimize the hyperplane according to the interval maximization given correctly separable points; while perceptron strives to find a separation hyperplane that can completely and correctly separate the positive instance points from negative instance points in the training set. That is, under SVM multiple-pattern identification, the classifier must be built first.

SVM can well address both nonlinear classification and linear classification problems and deeply mine relevant data because SVM can segment the non-segmentable linear samples and transform them from low-dimensional space to high-dimensional space to obtain the best segmentation plane. If the training data set of the SVM algorithm is linear in the plane and can be segmented, then its decision function is calculated by Eq. (1):

\[ y_i = \omega x + b \]  

(1)

In Eq. (1), \( \omega \) means weight vector, and B is offset vector. If the training data set of the SVM algorithm is nonlinear in the plane and can also be segmented, then its decision function is shown in Eq. (2):

\[ \min \left( \frac{1}{2} \omega^T \omega + C \sum_{i=1}^{N} \delta_i \right) \]  

(2)

\[ y_i(\omega x + b) \geq 1 - \delta_i, \delta_i \geq 0, i = 1, 2, \cdots, N \]  

(3)

Eq. (3) is a supplement to Eq. (2). In Eq. (2), C is a penalty factor, and the minimum classification error and maximum spacing of positive and negative categories depend on C; \( \delta_i \) denotes a nonnegative relaxation variable. In the SVM-based enterprise FRM model, the enterprise financial data are used as the input vector, and the SVM algorithm model can well classify the enterprise financial risk and further monitor the risks.

2.2 Big Data Technology. Big data go far beyond the traditional database in terms of data acquisition, data storage, data management, and data analysis. Thus, big data have shown many advantages, such as abundant data information, fast data conversion, rich data types, and low-value density in the data processing. Considering these factors, this paper integrates BDM into enterprise financial analysis. The specific process of data generation, storage, analysis, and application of enterprise structured big data will go through several complicated steps, and the interrelationship between these processes constitutes the big data structure. Generally, algorithm prediction or document consultation processes all involve big data collection, data storage, data processing, and specific application.

Specific enterprise FRM process is divided into three parts: financial risk identification, evaluation, and
Financial risk management

SVM-based financial risk monitoring model
Big data technology
Logistic regression financial risk monitoring model
Information fusion-based financial risks monitoring model

Research methods

Research result

Analysis conclusion

Figure 1: Technical route.

Figure 2: FRM indexes.
management. Further, financial risk identification can be subdivided into the evaluation of the macroeconomic environment and the evaluation of enterprise credit rating; financial risk evaluation can be subdivided into default probability and credit risks. Figure 3 shows the enterpriseFRM process.

The selected enterprise operational general indexes are subdivided: first, the enterprise data can be divided into internal data and external data. Internal data contain channel data, financial data, and user data. External data include market data, government data, and industry data. Figure 4 shows the specific data type framework of the enterprise.

Flow data on the enterprise platform encompass the number of visitors, the length of visit, visitor profile, and visitor-pay conversion. The visitor profile can be subdivided into visitor age, visitor gender, and visitor category. Figure 5
illustrates the flow data framework of the enterprise platform.

Enterprise channel data can be divided into channel distribution data, basic channel data, the number of channels, and channel conversion data. Basic channel data can be subdivided into average visit time, import visits, and click volume. Channel conversion data can be subdivided into conversion rates, the number of orders, and payment amount. Figure 6 presents the channel data tree of the enterprise.

On-platform enterprise data can be divided into four parts: logistics data, visitor transaction data, revenue data, and user data. Figure 7 shows the data framework of on-platform enterprise data.

On-platform business-customer transaction data can be divided into order data, calculated data, and payment data. This division can completely analyze the on-platform customer transaction data and implement it throughout the whole transaction process. Order data are subdivided into order channel distribution, payment amount, and order amount. Calculated data are subdivided into conversion rates and payment rates. Payment data are subdivided into payment amount, payment types, and payment results. Figure 8 is a dendrogram of on-platform business-customer transaction data.

On-platform user data are divided into member user, order user, paying user, and user retention rate. Figure 9 displays the on-platform user data.

Today, many enterprises have established big data analysis platforms to monitor and manage financial risks through big data. The specific plate distribution of analysis model for enterprise big data platform is as follows: first, the model is stratified into data control layer, process scheduling layer, and internal and external structure data layer; second, in terms of local division, the data management platform is divided into data standard, data quality, metadata, and data security; the process scheduling platform is divided into process scheduling, monitoring and warning, and warning model; application data area is divided into user management and risk management; the big data area is divided into the big data to be processed and the big data processed; external enterprise users can be subdivided into business sand table exercise, exercise data area, subject data, and subject data area. The analysis model of enterprise big data platform is shown in Figure 10.

2.3. Logistic Regression-based FRM Model. Logistic regression and multiple linear regression are similar and both belong to the generalized linear model while differing in the dependent variables. The dependent variables of a binomial distribution, namely Logistic regression, can either be dichotomous or multi-classified, in which dichotomous is more commonly used and easier to explain. Thus, binary Logistic regression is most commonly used in practice.

The Logistic regression method can efficiently handle data classification problems and can be used for data monitoring and analysis. The linear regression model can be implemented by the maximum likelihood estimation method under the Logistic regression model so that the data set of binary variables can be classified [31–33]. Figure 11 illustrates the flow data framework of the enterprise platform.
demonstrates the specific financial indexes of the model input data.

The parameters of the Logistic regression model function are set as follows: $x_i$ means the interpretation of independent variables, and $y_i$ must obey the distribution of Eq. (4):

$$P(y_i = 1|x_i) = P[\epsilon_i \leq (\alpha + \beta x_i)] = \frac{1}{1 + e^{\alpha + \beta x_i}}$$

(4)

$$\beta = (\beta_0, \beta_1)$$

(5)

Eq. (5) is the vector $n \times 1$. $\beta_0$ means constant, and $\beta_1$ represents a coefficient vector. The enterprise FRM model implemented by the Logistic regression model is shown in Eq. (6):

$$y_i = \ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n$$

(6)

This Regression model can be transformed into Eq. (7) through calculation:

$$p_i = (y_i = 1|X) = \frac{1}{1 + \exp[-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n)]}$$

(7)

In Eq. (7), $y$ means monitoring and judgment, and $P$ refers to probability support for judgment results.

2.4. Information Fusion-Based FRM Model. The information fusion-based FRM model [34–36] uses a fusion algorithm to fuse different information sources. Dempster-Shafer’s (DS) evidence theory is a famous information fusion method with good practicability. DS evidence theory can be used to process data from different sources and finally, transform them into output results. DS evidence theory aims to gain trust. Based on probability, the fusion processing is conducted according to specific rules, and the trust function is particularly important in this process. Thereupon, the framework of the information fusion-based FRM model is built. Here, it can be assumed that $U$ is a nonempty set composed of multiple elements, which is the proposed framework. Then, the trust structure is set, namely, the Basic Probability Assignment (BPA) function, which can be simplified to an $m$ function. This probability distribution has the characteristics of Eq. (8) and Eq. (9).

$$U = \{U_1, U_2, \cdots, U_n\}$$

(8)

$$m(\emptyset) = 0$$

(9)

$$\sum_{A \subseteq U} m(A) = 1$$

(10)
<table>
<thead>
<tr>
<th>Data control layer</th>
<th>Data management and control platform</th>
<th>Process scheduling layer</th>
<th>Enterprise internal and external structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data standard</td>
<td>Data quality</td>
<td>Process scheduling</td>
<td>Internal users</td>
</tr>
<tr>
<td>Metadata</td>
<td>Data security</td>
<td>Platform</td>
<td>Business sand table exercise</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Exercise area</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Subject data area</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Subject data area</td>
</tr>
<tr>
<td>IT staff</td>
<td></td>
<td></td>
<td>E-commerce system</td>
</tr>
<tr>
<td>Real-time data query</td>
<td>Real-time data area</td>
<td>Big data to be processed</td>
<td>Internet data</td>
</tr>
<tr>
<td>Historical data query</td>
<td>Historical archive data area</td>
<td>Big data processed</td>
<td>Industry data</td>
</tr>
<tr>
<td>Financial analysis</td>
<td>Big data area</td>
<td>Database data exchange component</td>
<td>Enterprise data</td>
</tr>
<tr>
<td>Debt analysis</td>
<td></td>
<td>Data area data exchange component</td>
<td>Financial system</td>
</tr>
<tr>
<td>Investment analysis</td>
<td></td>
<td>Monitoring and warning</td>
<td>Big data exchanged components</td>
</tr>
<tr>
<td>Risk analysis</td>
<td>Application data area</td>
<td>Warning model</td>
<td>Big data processed</td>
</tr>
<tr>
<td>Project analysis</td>
<td>User management</td>
<td></td>
<td>Business sand table exercise</td>
</tr>
<tr>
<td></td>
<td>Financial management</td>
<td></td>
<td>Exercise area</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Subject data area</td>
</tr>
</tbody>
</table>

**Figure 10:** Analysis model for enterprise big data platform.

**Figure 11:** Model indexes.
Here, the subset of the U set is A, then \( m(A) \) is the probability distribution function. BeliefFunction \( (Bel) \) is calculated by Eq. (11):

\[
Bel(A) = \sum_{B \in A} m(B) \quad (11)
\]

\( Bel(A) \) is the expression of A’s belief, which represents the reliability function of the U set. In Eq. (12), \( m_1 \) and \( m_2 \) stand for the trust degree of the two reliability functions on different pieces of evidence. Equation Eq. (15) indicates that the probability under monitoring.

Here, the subset of the U set is A, then \( m(A) \) is the probability of \( A \) under monitoring. Equation Eq. (16) stands for the probability of the SVM-based FRM model.

\[
m(C) = \left\{ \begin{array}{ll}
\sum_{A \cap B \neq \emptyset} m_1(A_i) \ast m_2(B_j) \\
1 - \sum_{A \cap B = \emptyset} m_1(A_i) \ast m_2(B_j), A \cap B \neq \emptyset \\
0, A \cap B = \emptyset
\end{array} \right. \quad (12)
\]

\[
\sum_{A \cap B \neq \emptyset} m_1(A_i) \ast m_2(B_j) < 1 \quad (19)
\]

\[
\forall C \subseteq U, C \neq \emptyset, m(\emptyset) = 0 \quad (20)
\]

\[
m_1(F_i) = p_i, m_2(F_j) = p_j' \quad (21)
\]

Eqs. (19)-(21) are supplements to Eq. (18). To sum up, the trust degree of \( m(C) \) is obtained through the SVM and Logistic regression model. Then, the information fusion-based FRM model is implemented, which can effectively monitor the enterprise financial risk and improve the model reliability; this is of great significance for obtaining the specific situation of enterprise financial risk [37, 38].

3. Results and Discussion

3.1. Analysis of Enterprise Financial Risk. Following the investigation of deposits and loans of relevant banks, the relevant conditions of several banks are investigated to ensure the data authenticity, and the relevant data are summarized in Figure 12, in which the deposit loan ratio is 69.6% in 2018, 71.3% in 2019, and 72.3% in 2020; the ratio of liquidity in 2018, 2019, and 2020 are 34.2%, 34.6%, and 33.8%, respectively; the standard value in 2018, 2019, and 2020 is 25%. Obviously, the deposit loan ratio is increasing yearly from 2018 to 2020, showing that the deposit loan ratio has a great impact on liquidity and potential financial risks. The enterprise liquidity index is shown in Figure 12.

This section is divided into two parts according to the sample data. The first part is the test, and the second part is to verify the accuracy of the SVM-based FRM model and Logistic regression-based FRM model. In Figure 14, (a) presents the specific financial risk probability of five enterprises under the SVM-based FRM model: the probability of enterprise 1, 2, 3, 4, and 5 is 0.85, 0.95, 0.83, 0.88, and 0.93, respectively; (b) showcases specific financial risk probability of five enterprises under Logistic regression-based FRM model: the probability of enterprise 1, 2, 3, 4, and 5 is 0.95, 1, 0.98, 1, and 0.86, respectively; (c) displays specific financial risk probability of five enterprises under information fusion-based FRM model: the probability of enterprise 1, 2, 3, 4, and 5 is 0.99, 1, 0.99, 1, and 0.98, respectively. Figure 14 reveals the monitoring results of different enterprise FRM models.

Figure 15 signifies that the risk classification accuracy under the SVM-based, Logistic regression-based, and information fusion-based FRM models are 90.22%, 90.88%, and 95.18%, respectively; while the monitoring error rate is 9.68%, 9.56%, and 4.68%. Figure 15 shows the comparison of the accuracy and error rate of different enterprise FRM models.
Figure 12: Enterprise liquidity index.

Figure 13: Enterprise profitability index.
In Figure 15, the ordinate represents the index value. The specific index value can be obtained through the mapping software. Figures 14 and 15 suggest that the SVM-based FRM model and Logistic regression-based FRM model have shown excellent classification effect for enterprise financial risks with an accuracy of 90.22% and 90.88%, respectively; by comparison, the information fusion-based FRM model beats the SVM-based FRM model and Logistic regression-based FRM model with an accuracy as high as 95.18%. Thus, the information fusion-based FRM model outperforms the SVM-based FRM model and Logistic regression-based FRM model; it can process and calculate the enterprise financial risk data from different sources and obtain higher accuracy. The analysis of specific data shows that although this paper uses the real-time database to solve the uncertainty in information fusion, the fusion results from the combination of information fusion and BDM technology are still unsatisfactory; this will be the research focus in the future.

In summary, the analysis of specific sample data implies that the accuracy of the SVM-based FRM model and Logistic regression-based model in enterprise financial risk classification is relatively high, which are 90.22% and 90.88%, respectively; by contrast, the classification accuracy of information

Figure 14: Monitoring results of different FRM models (a: Result of SVM-based FRM model; b: Result of Logistic regression-based FRM model; c: Results of information fusion-based FRM model).
fusion-based FRM is significantly higher than SVM-based FRM model and Logistic regression-based FRM model with an accuracy of 95.18%. This shows that the information fusion-based FRM model is better than the SVM-based FRM model and Logistic regression-based FRM model.

3.3. Performance Comparison between Information Fusion-Based FRM Model and Similar Algorithms. This section further verifies the accuracy and classification effect of the proposed information fusion-based FRM model practical application by comparative analysis with K-nearest neighbor (KNN) and Classification And Regression Tree (CART) from Accuracy mean and F1 score. The experimental results are shown in Figure 16.

Figure 6 corroborates that the mean Accuracy and F1 value of the proposed information fusion-based FRM model is 0.9626 and 0.7958, respectively. Compared with KNN and CART models, the proposed information fusion-based FRM model has better practical applicability, a more prominent risk prediction effect, and can provide good algorithm support for enterprise FRM.

4. Conclusion
The advancement of science and technology, especially, state-of-art technologies, innovates people’s lives while bringing challenges at the same time. Likewise, the rapid development and growth of enterprises are also followed by various financial risks. Aiming at the current situation of enterprise financial risk, this paper makes an in-depth study on enterprise financial risk based on BDM, SVM, Logistic regression, and information fusion technology. The
following conclusions are drawn: big data technology can provide important research methods for enterprise financial risk problems. SVM-based FRM model and Logistic regression-based FRM model can well classify enterprise financial risks with high accuracy; the information fusion-based FRM model can further improve the classification accuracy of enterprise financial risks and shows high reliability and effectiveness; additionally, different enterprise risk indexes are analyzed, finding that there is a need for enterprises to strengthen FRM under big data, especially, the management of liquidity and profitability indexes. The shortcomings of this paper are summarized: although the proposed information fusion-based FRM model improves the classification accuracy of enterprise financial risks; but its accuracy might be able to get a higher level in future research. The follow-up research will further improve the enterprise FRM performance based on the drawn conclusions.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no competing interests.

Acknowledgments

Phase results of 2021 Guangxi Education Science Planning Funding Key Project (2021A038); Stage results of 2019 Guangxi Higher Education Undergraduate Teaching Reform Project Key Topic (2019GZ145); 2018 National Social Science Foundation of China (18BZY015); 2018 Academic Research Project of Master of Taxation (SWX20180010); Phase results of the Integration Innovation Project (2020GJA166) of Guangxi University for Nationalities in 2020; 2015 China-ASEAN Economic and Trade Development and South China Sea Strategic Collaborative Innovation Center Project (15YBB06).

References


