Predicting Chain’s Manufacturing SME Credit Risk in Supply Chain Finance Based on Machine Learning Methods

Yu Xia 1, Ta Xu 1,*, Ming-Xia Wei 1,*, Zhen-Ke Wei 2,*, and Lian-Jie Tang 1

1 School of Management, Henan University of Technology, Zhengzhou 450001, China
2 Dyson School of Applied Economics and Management, Cornell University, New York, NY 14850, USA
* Correspondence: weimingxia@haut.edu.cn (M.-X.W.); zw638@cornell.edu (Z.-K.W.)

Abstract: Supply chain finance is an effective way to solve the financial problems of small and medium-sized manufacturing enterprises, and the assessment of credit risk is one of the key issues in supply chain financing. However, traditional credit risk assessment models cannot truly reflect the credit status of financing companies. In recent years, scholars working in this field have proposed using machine learning methods to predict the credit risk of supply chain enterprises, achieving good results. Nonetheless, there is no consensus on which approach is the most suitable for manufacturing companies. This study took small and medium-sized manufacturing enterprises as the research object, selected risk evaluation indicators according to the characteristics of the small and medium-sized manufacturing enterprises, and built a credit risk evaluation system. On this basis, we selected SMEs on China’s stock market from 2015 to 2020 as the sample data and evaluated corporate credit risk based on four commonly used machine learning algorithms. Then, combined with the evaluation results, a partial dependence plot method was used to visually analyze the important indicators. The results showed that a credit risk evaluation system for supply chain finance for manufacturing SMEs could be composed of the profile of the financing companies, the asset status of the financing companies, the profile of the core companies, and the operation of supply chains. The use of a random forest algorithm made it possible to more accurately assess the credit risk of manufacturing supply chain finance. Since the impacts of different indicators on the evaluation results were quite different, supply chain enterprises and financial service institutions should formulate corresponding strategies according to specific situations.

Keywords: manufacturing; supply chain finance; risk assessment; random forest; PDP

1. Introduction

Financing difficulties have always been an obstacle to the healthy development of China’s manufacturing SMEs. In recent years, the development of the supply chain finance model has greatly improved the financing efficiency of small and medium-sized enterprises [1,2]. The supply chain finance model has also attracted the attention of core enterprises in China’s manufacturing supply chain, small and medium-sized enterprises, and financial institutions. Supply chain finance can effectively improve the audit efficiency of financial institutions through the closed-loop management of capital flow, information flow, and business flow across the entire supply chain. At the same time, core enterprises in the supply chain can also reduce production costs and financing costs by integrating supply chain resources and providing credit guarantees to upstream and downstream enterprises, thereby enhancing their product competitiveness. Although supply chain finance can reduce the credit risk of enterprises in the supply chain compared to the traditional financing model, due to the more complex living environment faced by small and medium-sized manufacturing enterprises in China, such as significant operational risk, opaque financial information, and more intermediate links with core enterprises, credit risk is still the main obstacle for current supply chain finance businesses.
Corporate credit risk should be carefully assessed since its accuracy directly affects the future income levels of financial institutions [3]. In this context, academic and financial circles are actively exploring how to more accurately assess the credit risk in supply chain finance. For example, some scholars have used statistical methods, such as logistic regression and linear regression, to establish enterprise credit risk assessment models. Commercial banks in China usually use the method of credit rating to review the credit risk of financing enterprises. The above method is effective for large enterprises with transparent information, good financial systems, and credit records, but it is difficult to achieve the desired results in small and medium-sized manufacturing industries due to their special nature.

In recent years, the development of artificial intelligence technology has provided new ideas for the credit evaluation of manufacturing SMEs. Existing research shows that machine learning methods, such as the neural network, random forest, and support vector machine models, have higher accuracy than traditional credit evaluation methods, especially in the evaluation of nonlinear factors [4]. Machine learning algorithms can be trained with existing data to extract knowledge and predict future corporate credit risks based on the training model. It is worth noting that machine learning methods can produce relatively accurate predictions for a small amount of data based on a trained model. Although the use of machine learning models to predict credit risk has produced great achievements, there are still some shortcomings. For example, Li et al. [5] proposed a hybrid model combining logical regression and an artificial neural network to predict the credit risk of SMEs. However, their research objects were generally SMEs, and the research results could not reflect the real credit risk situation of a manufacturing industry. Fayyaz et al. [6] used a supervised learning algorithm to predict participant credit risk from the perspective of participant network attributes. Although these authors used the data for Iran’s automobile industry for analysis, they did not take into account the impact of the overall operation of the supply chain. Xuan [7] believes that a risk evaluation index based on a fuzzy preference relationship and risk-related theory can accurately reflect the credit risk status of a research enterprise. However, this author did not undertake a corresponding analysis of how the risk indicators affect the credit risk status of enterprises. According to the above analysis, although the application of machine learning methods to solve the problem of credit evaluation of SMEs has been recognized in academic and financial circles, current research on enterprise credit risk assessment cannot truly reflect the real situation of manufacturing SMEs.

Therefore, based on the real-life scenario of small and medium-sized manufacturing enterprises in China, this paper makes the following research contributions: (I) Taking small and medium-sized manufacturing enterprises in China as the research object and analyzing their special environment and characteristics, this paper determines the credit risk evaluation indicators suitable for small and medium-sized manufacturing enterprises in China in order to more truly reflect the status of manufacturing credit risk. (II) Most existing studies take a single SME as the risk assessment object but ignore the impact of upstream and downstream enterprises in the supply chain and the overall operation of the supply chain in the assessment enterprise. Additionally, today’s commercial operations are no longer the environment for the competitive development of a single enterprise and have changed from development of a single enterprise to supply chain development. The core enterprises in the supply chain and the operation status of the supply chain have increasingly important influences on the operation of SMEs. Therefore, this study introduces core enterprise and supply chain operations into the credit risk assessment of SMEs in order to more comprehensively reflect the financing environment of SMEs. (III) There are many existing machine learning methods, but few scholars have explored which risk assessment model is suitable for SMEs in China’s manufacturing industry. A reasonable risk assessment model can not only help banks carry out risk assessments but also help SMEs in the manufacturing industry obtain funds for development. Therefore, it is particularly important to explore which machine learning risk assessment model is
suitable for small and medium-sized manufacturing enterprises in China. By comparing the experimental results for current commonly used machine learning credit risk assessment models and using the partial dependence graph method to explore the impact of a single indicator on credit, this study aims to provide specific and feasible suggestions for supply chain managers.

The rest of this paper is organized as follows: Section 2 reviews the literature on credit risk assessment in supply chain finance. Section 3 introduces the selection of the financial credit risk evaluation indicators for the manufacturing supply chain. Section 4 uses four different machine learning models to evaluate the credit risk of the sample enterprises selected from the Chinese stock market and uses the partial dependence plot method to analyze the impact of the main indicators on the results. In Section 5, we summarize the overall research process, promote the idea of enlightened management, and describe future research prospects.

2. Literature Review

2.1. Traditional Credit Risk Assessment

Most of the existing research on supply chain financial risk assessment focuses on the selection of credit risk assessment indicators and the construction of assessment models.

For the construction of credit risk evaluation indicators, Zhao et al. [8] proposed a credit evaluation index system for SMEs with multiple theoretical perspectives, including taxable sales income and the frequency of paying value-added tax. Liu and Yan [9] constructed a credit risk evaluation system that comprehensively considered the risk ratings of farmers’ cooperatives, financing debt, and the macroeconomic environment. Zhang et al. [10] designed a risk prevention linkage mechanism for credit evaluation and risk measurement and proposed a retailer credit evaluation technology derived from the BBWM–cloud model. In addition, some scholars have used AHP and factor analysis to analyze the credit risk evaluation indicators for supply chain finance.

In terms of the credit risk assessment model, Crook et al. [11] and Teles et al. [12] used the logistic regression method to predict credit risk. The results showed that the method had good applicability. This is also the main reason why most commercial banks currently use logistic regression for credit risk assessment. Using the improved KMV model and the copula function, Zhang et al. [13] constructed a credit risk assessment model that can be used to evaluate a single financing enterprise and a combination of financing enterprises in the supply chain. Yang et al. [14] put forward a financial risk management model for the Internet supply chain and verified the accuracy of the model’s evaluation through a segmented sample regression analysis. In the field of agricultural supply chain finance, Bai et al. [15] used fuzzy rough set theory and the mean clustering method to evaluate the complex relationship between farmers’ characteristics, competitive environmental factors, and farmers’ credit levels.

2.2. Machine Learning Credit Risk Assessment

Although the above model can assess credit risk well, it requires relevant assumptions, such as the assumption of the normality of the independent variables, which deviates from the actual state of enterprises [16]. Moreover, most of the above risk assessment systems only consider the current status of SMEs and cannot reflect the impact of the complex environment in which SMEs are located, making it difficult for the assessment results to objectively and accurately reflect the true credit situation of the enterprises. In this context, relevant scholars have introduced machine learning methods into the evaluation of supply chain credit risk, combined them with enterprises’ existing financial and transaction data, and verified them based on a limited sample to make the evaluation more realistic [17].

There are many machine learning models currently available, such as support vector machines, logical regression, decision trees, neural networks, random forest algorithms, etc. Researchers select appropriate models according to their own research needs. Support vector machine (SVM) is a kind of generalized linear classifier that classifies data using a
supervised learning method. It uses the hinge loss function to calculate the empirical risk and adds a regularization term to the solution system to optimize the structural risk. It is a sparse and robust classifier. Support vector machine is one of the common kernel learning methods and can be used for nonlinear classification. Xie et al. [18] used a mixed model consisting of a support vector machine and logical regression to evaluate the credit risk of small, medium, and micro-enterprises. Logistic regression, a generalized linear regression method, is a traditional and widely used supervised learning method.

The training speed for logistic regression data is fast and the experimental results are easy to interpret, which are two of the reasons why most commercial banks use these results as a credit risk assessment [19]. However, their disadvantage lies in their ineffectiveness in solving nonlinear problems. Therefore, in some studies, relevant scholars have improved such data to make up for their shortcomings.

A decision tree is a very common classification method. It consists of a tree structure in which each internal node represents a test with an attribute, each branch represents a test output, and each leaf node represents a category. The decision tree model is highly efficient, and its results are easy to interpret. In addition, decision trees can also process irrelevant feature data and are widely used to mine nonlinear data [17]. This allows researchers to use fewer data samples to achieve an accurate analysis of research results. Wang et al. [20] used the C4.5 algorithm to predict the credit risk of enterprises with a small amount of sample data and achieved an average accuracy rate of 84.39%, proving that the algorithm can accurately predict the credit risk of enterprises. However, the decision tree model also has some shortcomings; in particular, when dealing with data that have strong feature correlations, the prediction effect is not satisfactory.

Artificial neural networks are also called neural networks or connection models. They are algorithmic mathematical models that imitate the behavior characteristics of animal neural networks and conduct distributed parallel information processing. This kind of network relies on the complexity of the system and processes information by adjusting the interconnections between large numbers of internal nodes [21]. In recent years, this kind of data has demonstrated explosive growth, and neural networks are also popular, with the public relying on their complex systems and excellent processing results. Additionally, neural networks can easily process noise and are not susceptible to damage [22].

The random forest algorithm is a classifier containing multiple decision trees, and its output categories are determined by the modes of the categories output by individual trees. Additionally, the application scenarios of random forest algorithms are very diverse, and they are widely used in risk assessment [23], marketing, and other areas. In addition, they can balance errors and obtain more accurate prediction results in a short time. Scholars have also used models such as LDA [24] and LRA [24] to predict SME risk, but through our literature review and analysis, we found that the application scenarios of the four supervised machine learning models—the random forest, support vector machine, neural network, and logical regression models—are more suitable for SME credit risk assessment.

3. Credit Risk Assessment System

3.1. Selection of Credit Risk Assessment Indicators

At present, most financial institutions use financial indicators to evaluate enterprise financing risks, mainly because financial data are easy to obtain and the business status of an enterprise can be intuitively understood through such data. However, with the development of the economy and information technology, enterprises are no longer independent individuals. They are also affected by industries, counterparties, and other internal factors. Therefore, solely relying on the financial indicators of an enterprise to reflect the real enterprise risk status is problematic. In particular, manufacturing enterprises have high labor costs, extensive inventories, and high working capital requirements for raw materials and product manufacturing [25,26]. In addition, the manufacturing industry is easily affected by market demand fluctuations and product prices in the supply chain, which require companies to maintain sufficient funds to ensure stable operations. In other words, manufacturing
enterprises rely more on the financial support of financial institutions and have a more complex financing environment. Therefore, the selection of manufacturing risk indicators should be combined with the special industry characteristics of manufacturing enterprises and the current operating conditions of the manufacturing supply chain. In addition, the operation of SMEs in the manufacturing industry mostly depends on the product demand of core enterprises in the supply chain, and financing requires credit guarantees from core enterprises. Thus, when selecting risk indicators, we should not only focus on financial aspects but also take non-financial indicators as important reference factors.

Many scholars have conducted extensive research on enterprise credit risk evaluation indicators. Zhu et al. [17] divided a set of risk assessment indicators into financial indicators and non-financial indicators. They found that the development of the industry and the degree of cooperation between SMEs and core enterprises are important indicators that affect the credit risk of enterprises. Kouvelis et al. [27] studied the impact of enterprise credit ratings on supply chain operations and specifically analyzed how credit ratings affect the choice of financing method. Wetzel et al. [28] concluded that the capital status and operation level of core enterprises play important roles in easing the financial constraints of supply chain partners and that a good inventory turnover can benefit the financing of enterprises. Zhao et al. [29] believe that the risk of enterprise management is a significant problem for the financing of enterprises and that incorporating the influence of enterprise management into a credit risk assessment can better reflect the real credit risk of enterprises. Based on the above analysis combined with the research by existing scholars on the evaluation indicators of supply chain financial credit risk, this study, from the perspective of manufacturing supply chain finance, selected risk indicators for manufacturing SMEs from four categories: the overview of the financing enterprise, the asset status of the financing enterprise, the overview of the core enterprise, and the operation of the supply chain.

3.1.1. Overview of the Financing Enterprise

The state of the financing enterprise is one of the foundations of supply chain finance, and it is an important component of credit risk assessment. It consists of financial indicators and non-financial indicators.

The financial status of financing enterprises is an important factor in credit risk evaluation. This is because manufacturing enterprises need to purchase raw materials and transport products. These make capital flow more frequent and thus the financial indicators can more intuitively reflect the overall situation of the enterprise. For instance, the higher the cash ratio of the financing enterprise, the stronger the liquidity of the enterprise, and good liquidity means that the enterprise can repay loans in time. In addition, the current ratio, total asset turnover rate, operating net profit rate, and other financial indicators of the enterprise reflect the status of the enterprise in terms of its operation, debt repayment, and profit. The better the performances of the above financial indicators of the financing enterprise are, the higher the probability of the operating compliance of the enterprise.

The non-financial indicators of financing enterprises mainly include the life cycle of the enterprise, the average age of executives, R&D investment, and the enterprise scale. Firstly, the life cycle of the enterprise corresponds to the different market sales environments that affect the future profitability of the enterprise’s products [30]. Since manufacturing enterprises mainly rely on the products that they produce to make profits, this paper divides the life cycle of enterprises according to the growth rate for the operating income. A growth rate for operating income of less than 5% represents a recession period, with a value of 0; a growth rate for operating income greater than 10% represents a growth period, with a value of 1; a growth rate for operating income between 5% and 10% represents a mature period, with a value of 2. Secondly, the average age of executives is an important indicator of corporate financing risk. Relevant research shows that middle-aged executives have a high level of education, strong business ability, and greater ability and willingness to repay corporate loans [31]. Financial institutions, out of concern for future repayment, are more willing to cooperate with mature and stable enterprise management. Thirdly, R&D
investment reflects the risks of enterprises to a certain extent. The ratio of R&D investment to operating income indicates the ratio of product R&D investment to operating income for a manufacturing enterprise and can reflect the strength of the research and development investment in an enterprise at a particular time. It is worth noting that manufacturing enterprises need to continuously develop and upgrade products to increase market share and revenue and maintain competitiveness [32]. Finally, the size of an enterprise is also one of the indicators used to measure its credit risk. Usually, most SMEs are at the stage where they are attempting to expand their market. For financial institutions, the larger the scale of financing enterprises, the higher their efficiency, which means that the probability of default is also lower.

Using the above analysis, the credit risk evaluation indicators for manufacturing SMEs were selected in terms of profitability, solvency, operating capacity, cash flow, R&D investment, life cycle, enterprise scale, and executive age. Specifically, the total asset turnover ratio, current ratio, quick ratio, cash ratio, operating net interest rate, net cash content of operating income, total assets, life cycle, average age of executives, total asset growth rate, and ratio of R&D investment to operating income of the financing enterprise were selected as specific indicators of the financing enterprise profile.

3.1.2. Asset Status of the Financing Enterprise

Inventory and accounts receivable, as important assets of enterprises, have become the basis of supply chain finance [33]. Since the main mortgage assets of manufacturing SMEs are raw materials and the product inventory, the quality of the inventory directly determines the quality of financing mortgage assets. Therefore, the value of mortgageable assets has a significant positive impact on obtaining commercial credit [34]. Furthermore, the inventory turnover ratio reflects the turnover speed of the inventory and whether the liquidity of the inventory and the occupancy of inventory funds are reasonable. Improvement in the inventory turnover rate can encourage enterprises to improve the efficiency of capital use and their performance while ensuring the continuity of production and operation [35]. Regarding the accounts receivable, the turnover rate and receivable turnover days are important indicators of an enterprise’s current assets and can reflect its short-term capital liquidity and short-term debt repayment ability.

In this context, we selected inventory turnover rate, accounts receivable turnover rate, and accounts receivable turnover days as the asset status indicators for financing enterprises.

3.1.3. Overview of the Core Enterprise

The core enterprise is an important component of the supply chain and has absolute control over the entire supply chain. In the process of carrying out supply chain financing business, core enterprises play a key role; in particular, when financing enterprises rely on the credit of core enterprises for financing, the guarantee provided by core enterprises can greatly reduce the risk of supply chain financing [36]. In view of the fact that core enterprises mainly play the role of credit endorsers in supply chain finance, the impact of core enterprises on supply chain finance is mainly reflected in credit evaluation, guarantee status, profitability, and solvency.

In terms of credit evaluation, current descriptions of credit risk generally adopt the method of credit rating. In this study, core enterprises with a rating of AA and below were defined as high-risk enterprises with a value of 1 and those with a rating of AA and above were defined as low-risk enterprises with a value of 0, in accordance with the credit evaluation status of the core enterprises. In terms of guarantees, the past guarantees of core enterprises are significant factors that financial institutions must consider. If the previous guarantees of core enterprises are upheld within the specified time limit, it is generally believed that core enterprises have better guarantee capabilities. The guarantee balance is the enterprise’s external guarantee balance after deducting the amount of the released guarantee responsibility and can reflect the guarantee status of the core enterprise within a certain period of time. In this part of the study, we assessed the guarantee situation of
core enterprises using their guarantee balance. Taking into account the financing needs of SMEs in the manufacturing supply chain and the financing situation of Chinese enterprises, it was determined that core enterprises with an annual guarantee balance of 0 have a good guarantee status and they were assigned the value of 1. Core enterprises with a guarantee balance greater than 0 and less than CNY 100 million have an average guarantee status and were assigned the value of 2. The guarantee status of core enterprises with a guarantee balance of more than CNY 100 million was considered poor and they were assigned the value of 3.

The operation status of core enterprises should also be considered. Good operation is the basis for supply chain development and implementation of supply chain finance. The asset–liability ratio, cash ratio, current ratio, and return on assets reflect the solvency and profitability of an enterprise from different perspectives.

In accordance with the above description, we selected the enterprise credit evaluation, guarantee status, asset–liability ratio, cash ratio, current ratio, and asset return ratio to represent the overview of the core enterprises.

3.1.4. Operation of the Supply Chain

The overall situation of the supply chain mainly reflects the environment in which the financing enterprise is located, including the strength of the supply chain relationship, the sustainable growth rate of the industry, and the growth rate of the industry’s operating profit. Among these factors, the strength of the supply chain relationship indicates the degree to which there is a business relationship between financing enterprises and core enterprises. The sustainable growth rate of the industry and the growth rate of the industry’s operating profit reflect the development of the industry in which the financing company operates. Firstly, the prerequisites for supply chain financing services are that SMEs have business transactions with core enterprises and that the product supply of the SMEs is important to the core enterprises; with these prerequisites met, the willingness of core enterprises to participate in supply chain finance will be higher. Consequently, this study selected the top five customers of SMEs as the consideration targets and used the ratio of core enterprises to the total sales of SMEs as the standard to measure the strength of the supply relationship between the two parties [37]. Specifically, when the ratio of core enterprises to SME sales was less than 5%, the relationship strength was weak and was assigned the value of 1; when the sales ratio was between 5% and 10%, the relationship strength was average and was assigned the value of 2; when the sales ratio was between 10% and 20%, the relationship strength was strong and was assigned the value of 3; when the sales ratio was more than 20%, the relationship was very strong and was assigned the value of 4. Secondly, in terms of the sustainable growth rate of the industry, the development of the industry in which the enterprise is located affects the future profitability and development capacity of the supply chain. The sustainable growth rate of an industry indicates the future development prospects. When the sustainable growth rate of an industry is high, the supply chain operation of the industry is more stable. Finally, the profit growth of the industry should also be considered. Considerable profits and significant industry operating profit margins are driving forces affecting enterprises’ ability to increase their productivity.

In light of the above discussion, this study selected the strength of the supply chain relationship, the industry sustainable growth rate, and the industry operating profit growth rate to assess the supply chain situation of financing enterprises.

3.2. Correlation Analysis of Credit Evaluation Indicators

To avoid the multi-linear problem caused by a high correlation between the indicators, we conducted correlation analysis with the abovementioned preliminary selected risk assessment indicators and eliminated the indicators that had too high a correlation. In SPSS 26 software, the Pearson correlation coefficient was used to analyze the correlation between the indicators, and the correlation analysis results of credit risk evaluation indicators are shown in Tables 1–3.
Table 1. Correlation analysis results (I).

<table>
<thead>
<tr>
<th></th>
<th>Cash Ratio</th>
<th>Current Ratio</th>
<th>Quick Ratio</th>
<th>Inventory Turnover Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cash ratio</td>
<td>1</td>
<td>0.719 **</td>
<td>0.814 **</td>
<td>0.162 *</td>
</tr>
<tr>
<td>Current ratio</td>
<td>0.719 **</td>
<td>1</td>
<td>0.948 **</td>
<td>−0.096</td>
</tr>
<tr>
<td>Quick ratio</td>
<td>0.814 **</td>
<td>0.948 **</td>
<td>1</td>
<td>−0.032</td>
</tr>
<tr>
<td>Inventory turnover rate</td>
<td>0.162 *</td>
<td>−0.096</td>
<td>−0.032</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: ** indicates that correlation was significant at the 0.01 level (two-tailed). * indicates that correlation was significant at the 0.05 level (two-tailed).

Table 2. Correlation analysis results (II).

<table>
<thead>
<tr>
<th></th>
<th>Current Ratio of Core Enterprise</th>
<th>Cash Ratio of Core Enterprise</th>
<th>Asset Liability Ratio of Core Enterprise</th>
<th>Return on Assets of Core Enterprise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current ratio of core enterprise</td>
<td>1</td>
<td>0.533 **</td>
<td>−0.523 **</td>
<td>0.320 **</td>
</tr>
<tr>
<td>Cash ratio of core enterprise</td>
<td>0.533 **</td>
<td>1</td>
<td>−0.476 **</td>
<td>0.322 **</td>
</tr>
<tr>
<td>Asset liability ratio of core enterprise</td>
<td>−0.523 **</td>
<td>−0.476 **</td>
<td>1</td>
<td>−0.443 **</td>
</tr>
<tr>
<td>Return on assets of core enterprise</td>
<td>0.320 **</td>
<td>−0.476 **</td>
<td>−0.443 **</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: ** indicates that correlation was significant at the 0.01 level (two-tailed).

Table 3. Correlation analysis results (III).

<table>
<thead>
<tr>
<th></th>
<th>Accounts Receivable Turnover Days</th>
<th>Operating Net Interest Rate</th>
<th>Total Asset Turnover Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounts receivable turnover days</td>
<td>1</td>
<td>−0.551 **</td>
<td>−0.505 **</td>
</tr>
<tr>
<td>Operating net interest rate</td>
<td>−0.551 **</td>
<td>1</td>
<td>0.218 **</td>
</tr>
<tr>
<td>Total asset turnover ratio</td>
<td>−0.505 **</td>
<td>0.218 **</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: ** indicates that correlation was significant at the 0.01 level (two-tailed).

It is generally believed that, when $|r|$ (Pearson correlation coefficient value) is greater than 0.5, this means that two variables are strongly correlated; when the $|r|$ value ranges from 0.30 to 0.49, it indicates that two variables are moderately correlated; when the $|r|$ value is less than 0.29, it indicates that two variables have a low correlation [38]. Considering that a high correlation between two indicators would affect the analysis results, the indicators with a strong correlation in the correlation test were excluded: the current ratio of the core enterprise, current ratio, quick ratio, operating net interest rate, and total asset turnover ratio. We thus built a manufacturing supply chain financial credit risk assessment index system, as shown in Table 4.

Table 4. Credit risk assessment index system.

<table>
<thead>
<tr>
<th>First-Level Indicator</th>
<th>Secondary Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of financing enterprise</td>
<td>Life cycle</td>
</tr>
<tr>
<td></td>
<td>Average age of executives</td>
</tr>
<tr>
<td></td>
<td>Total assets</td>
</tr>
<tr>
<td></td>
<td>Cash ratio</td>
</tr>
<tr>
<td></td>
<td>Total asset growth rate</td>
</tr>
<tr>
<td></td>
<td>Net cash content of operating income</td>
</tr>
<tr>
<td></td>
<td>Ratio of R&amp;D investment to operating income</td>
</tr>
<tr>
<td>Asset status of financing enterprise</td>
<td>Inventory turnover rate</td>
</tr>
<tr>
<td></td>
<td>Accounts receivable turnover days</td>
</tr>
<tr>
<td></td>
<td>Accounts receivable turnover rate</td>
</tr>
</tbody>
</table>
4. Model Evaluation Results and Analysis

4.1. Data Sources

Considering the slow development of China’s supply chain finance prior to 2015, reliable data to support research are scare. Nevertheless, we followed the current common practice in data selection: we selected the data for listed companies because these data are accurate, transparent, and easy to obtain. Therefore, the data source for this study was a selection of manufacturing SMEs on China’s stock market between 2015 and 2020. The sample of core enterprises was screened for the top five customers announced by the SMEs. Firstly, we judged whether the enterprise was a small or medium-sized manufacturing enterprise based on an analysis of the official website of the target enterprise. Secondly, we used the CAMAR database to query specific risk indicators according to the selected enterprise stock code. Finally, we defined a risky enterprise according to a comparison between the asset–liability ratio of the SME and the lower value from the Enterprise Performance Evaluation Standard Value publication. An enterprise with an asset–liability ratio higher than the lower value was defined as a risky enterprise and assigned the value of 1. Otherwise, it was defined as a non-risk enterprise and assigned the value of 0. During data screening, the companies with unlisted core enterprises or core enterprises with credit ratings that could not be queried were excluded, as well as SMEs with incomplete data. A total of 170 data samples were obtained through the data screening, including 140 samples for model training and 30 samples for model testing. All the data in this paper were derived from the official websites of the enterprises, their annual reports, or the CAMAR database.

We selected the data specifically through empirical observation so as to avoid the decline in the prediction function caused by overly concentrated tag values for data samples. As the selected indicators had different dimensions, which can have affected the research results, this study used SPSS 26 software to standardize the sample data based on the data standardization formula: $Z = \frac{x-u}{\sigma}$, where $u$ represents the mean value of the data and $\sigma$ represents the standard deviation of the data. The sample data processed through standardization conformed to the standard normal distribution, preventing any negative impact from different dimensions and feature result contributions on the model experiment. As there were only 170 data samples available, the following measures were proposed to solve the problem. Firstly, a support vector machine, logical regression, and other models with simpler calculation principles were selected. Secondly, this study used a tenfold cross-validation method to increase the amount of data that could be verified outside of the sample and thus ensure that the number of data samples in this study did not affect the experimental results.

4.2. Experimental Process

In this paper, support vector machine, random forest, multilayer perceptron, and traditional logistic regression models were used to compare and explore more suitable methods for manufacturing supply chain financial credit risk assessment, and then important characteristic indicators were selected for visual analysis. Firstly, we selected the original data based on the official websites of the enterprises and the CAMAR database. Secondly, SPSS

---

Table 4. Cont.

<table>
<thead>
<tr>
<th>First-Level Indicator</th>
<th>Secondary Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overview of core enterprise</td>
<td>Credit evaluation</td>
</tr>
<tr>
<td></td>
<td>Guarantee situation</td>
</tr>
<tr>
<td></td>
<td>Cash ratio</td>
</tr>
<tr>
<td></td>
<td>Asset–liability ratio</td>
</tr>
<tr>
<td></td>
<td>Return on assets</td>
</tr>
<tr>
<td>Supply chain operation</td>
<td>Strength of supply chain relationship</td>
</tr>
<tr>
<td></td>
<td>Industry operating profit growth rate</td>
</tr>
<tr>
<td></td>
<td>Industry sustainable growth rate</td>
</tr>
</tbody>
</table>
26 software was used to screen the data in combination with the real-life scenario. Thirdly, on the basis of the processed data, the data analysis software SPSS Modeler 18.0 was used to establish four models—a support vector machine, a random forest, a multilayer perceptron, and logical regression. The models were used to analyze the divided training set data, and the test set data were verified with the models after training. Finally, Python 3.7 was used to build a partial dependence plot model, and then the impact of characteristic indicators on the prediction results was visually analyzed.

4.3. Comparison of Model Results

4.3.1. Evaluation Index of Model Prediction Effect

Current machine learning model prediction effect evaluation includes the confusion matrix, accuracy, false negative rate, precision, and F-measure, which are specifically defined as follows:

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}
\]

\[
\text{False negative} = \frac{FN}{TP + FN}
\]

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

\[
F - \text{measure} = \frac{2TP}{2TP + FP + FN}
\]

where \(TP\), \(TN\), \(FP\), and \(FN\) refer to true positive, true negative, false positive, and false negative, respectively. The confusion matrix is a visual tool for supervised learning that is mainly used to compare classification results with the real information of the instance and is helpful for intuitively understanding the prediction of a model. Accuracy is the most commonly used indicator for a model, representing the proportion of the number of samples correctly predicted by the model compared to the total number of samples. The false negative rate represents the probability that the model will predict a risk-free enterprise to be a risky enterprise. If the false negative rate is too high, the bank will lose high-quality customers, which may affect future earnings. The accuracy rate indicates the proportion of the number of enterprises accurately predicted as risk-free compared to the total number of enterprises predicted as risk-free. This can also be understood as the precision rate. The F-measure is also an important indicator to measure the accuracy of binary classification problem models, and it takes into account the precision and recall of the classification model. The F-measure score can be regarded as a harmonic average of the model precision and recall, with a maximum value of 1 and a minimum value of 0. The above indicators are commonly used to reflect the performance of model evaluation. We used the above indicators to evaluate model performance.

4.3.2. Model Evaluation Results

Based on the selected evaluation model, the experimental data are analyzed, and the model performance is analyzed according to the model experimental results. The confusion matrix of the prediction results of the four models is obtained through experiments, as shown in Tables 5–8. The accuracy, false negative, precision and F-1 of different models are calculated according to the model confusion matrix, as shown in Table 9.

Table 5. Confusion matrix for random forest model experiment.

<table>
<thead>
<tr>
<th>Training Set</th>
<th>Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Truth</strong></td>
<td><strong>Forecast</strong></td>
</tr>
<tr>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>134</td>
</tr>
<tr>
<td>Negative</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 6. Confusion matrix for logistic regression model experiment.

<table>
<thead>
<tr>
<th></th>
<th>Training Set</th>
<th>Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Truth Forecast</td>
<td>Truth Forecast</td>
</tr>
<tr>
<td></td>
<td>Positive Negative</td>
<td>Positive Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>134</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 7. Confusion matrix for multilayer perceptron model experiment.

<table>
<thead>
<tr>
<th></th>
<th>Training Set</th>
<th>Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Truth Forecast</td>
<td>Truth Forecast</td>
</tr>
<tr>
<td></td>
<td>Positive Negative</td>
<td>Positive Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>134</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>1</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 8. Confusion matrix for support vector machine model experiment.

<table>
<thead>
<tr>
<th></th>
<th>Training Set</th>
<th>Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Truth Forecast</td>
<td>Truth Forecast</td>
</tr>
<tr>
<td></td>
<td>Positive Negative</td>
<td>Positive Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>134</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 9. Model evaluation results.

<table>
<thead>
<tr>
<th>Model</th>
<th>Training Set</th>
<th>Test Set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Accuracy</td>
<td>False Negative</td>
</tr>
<tr>
<td>RF</td>
<td>99.29%</td>
<td>0</td>
</tr>
<tr>
<td>LR</td>
<td>100%</td>
<td>0</td>
</tr>
<tr>
<td>MLP</td>
<td>99.29%</td>
<td>0</td>
</tr>
<tr>
<td>SVM</td>
<td>97.86%</td>
<td>0</td>
</tr>
</tbody>
</table>

It can be seen from the model experiment results that the accuracy rates for the different models with the training set were similar, but the models could not be directly evaluated with the training set. This was because the data related to all the influencing factors present in practice could not be obtained. In the validation set, the random forest model had the best accuracy, false negative rate, precision, and F-measure score, indicating that it could more effectively identify risky enterprises and was more suitable for credit risk assessment of SMEs in China’s manufacturing industry. The specific reasons for this were as follows. Firstly, the LR model is similar to the linear model and struggles to address nonlinear problems. However, there were nonlinear characteristic indicators in the real supply chain financial risk assessment, and logistic regression could not solve the problem of data imbalance. Therefore, although the accuracy of the logistic regression model in the test set was high, the accuracy of the verification results was low. Secondly, the performance of the SVM model mainly depends on the kernel function. However, the kernel functions of currently used SVM risk assessment models are generally artificially selected, have a certain level of randomness, and struggle to truly reflect the real-life scenario for the manufacturing supply chain. As the kernel function could not be accurately selected, it was difficult for the model verification results to reach a high level. Thirdly, the reasons for the low accuracy of the MLP test set may have been that the sample dimension was too high and the sample data were small. Considering the limited business cooperation between these two factors in real supply chains, it may not be possible to obtain sufficient data, thus making it impossible to accurately predict supply chain finance risk. Lastly,
there are various risk factors to consider in manufacturing supply chain finance, and more characteristic indicators should be selected when making risk predictions. As the random forest model can deal with high-dimensional data, it should show better performance in credit risk assessment. Moreover, the RF model can balance errors and has a strong generalization ability with good performance.

On the whole, the model test results corroborated the real situation of current manufacturing supply chain financial risk prediction. Additionally, the stochastic forest model was found to be more suitable for the risk assessment of manufacturing supply chain finance in terms of risk identification and accuracy.

4.4. Visual Analysis of Model Prediction Results

In order to further understand the specific impacts of different feature indicators on credit risk, the reader can refer to [39]. This section describes how Python 3.7 was used to build a partial dependence plot model on the basis of the above RF model and then presents a visual analysis of the effects of feature indicators on prediction results. Limited by space, we focus on the characteristic indicators that reflect the characteristics of the manufacturing industry and supply chain, such as the cash ratio, the average age of management, the asset-liability ratio of core enterprises, the ratio of R&D investment to operating income, the credit evaluation of core enterprises, and the strength of supply chain relationships.

4.4.1. Effect of Cash Ratio

Overall, as shown in Figure 1, the default probability of financing enterprises decreases with the increase in the cash ratio. When the cash ratio is greater than 0.8, the default probability remains at a low level. Since the cash ratio reflects the short-term solvency of an enterprise, the higher the cash ratio is, the stronger the liquidity of the enterprise and the stronger the short-term solvency. The financial sector generally believes that a corporate cash ratio above 0.2 is preferable. Nevertheless, an excessively high cash ratio also means that an enterprise’s current assets have not been properly utilized, resulting in the lower profitability of cash assets. In Figure 1, we can see that, when the cash ratio is close to 0.8, the profit of the financing enterprise’s cash assets decreases and the opportunity costs of the company increase, creating a certain risk for the enterprise.

![Figure 1. Effect of cash ratio.](image)

4.4.2. Effect of Average Age of Executives

Figure 2 shows the effect of the average management age on the default probability of financing enterprises. It is clear that the default probability of financing enterprises tends to decrease first and then increase with the average age of management. Specifically, when the average age of management is between 44 and 51 years old, the default probability of financing enterprises is at a low level. When the average age of the management reaches
51 years old, the learning and management decision-making abilities of managers decrease with their increasing age, and the business risk of the enterprise increases. This ultimately leads to an increase in the credit risk of the enterprise.

![Graph showing the effect of average age of executives on default probability.](image)

**Figure 2.** Effect of average age of executives.

### 4.4.3. Effect of Asset—Liability Ratio of Core Enterprises

Figure 3 shows the impact of the asset–liability ratio of core enterprises on the default probability of financing enterprises. In general, the credit default probability of financing enterprises increases with the increase in the asset–liability ratio of core enterprises. This result is consistent with the real-life scenario. In other words, when the asset–liability ratio of core enterprises is high, their solvency and development ability are reduced, which leads to an increase in operational risks for the entire supply chain and affects the development of SMEs. At the same time, the reduced solvency of core enterprises also means that, when SMEs default, core enterprises cannot provide financing guarantees for the SMEs in the supply chain.

![Graph showing the effect of asset–liability ratio of core enterprises on default probability.](image)

**Figure 3.** Effect of asset–liability ratio of core enterprises.

### 4.4.4. Effect of Ratio of R&D Investment to Operating Income

Generally, as shown in Figure 4, the impact of the ratio of R&D investment to operating income on the default probability of financing enterprises shows a trend of first decreasing and then increasing. Specifically, when the proportion of R&D investment increases from 0 to 4.75%, the default probability of enterprises significantly decreases, indicating that R&D innovation is very important for manufacturing enterprises. However, strong R&D also means high investment. When the proportion of R&D investment reaches a certain level, the probability of default begins to increase. As an explanation for this finding, it can
be suggested that, when an enterprise invests too much operating income in research and development, this may lead to other problems, such as a shortage of enterprise funds and unstable product sales, thereby increasing the operational risk of the enterprise. Therefore, most SMEs control the proportion of R&D investment in practice.

![Figure 4](image4.png)

**Figure 4.** Effect of ratio of R&D investment to operating income.

**4.4.5. Effect of Credit Evaluation on Core Enterprises**

Figure 5 reflects the impact of core enterprises’ credit rating on the default probability of financing enterprises. It is clear that there is a linear relationship between the credit evaluation of core enterprises and the default probability of financing enterprises. In other words, the better the credit rating of core enterprises is, the lower the default probability of financing enterprises. The main reason for this is that core enterprises serve as an important foundation for supply chain finance and significantly support the credit of the entire supply chain. The higher the credit rating of a core company is, the more likely it is to work with supply chain financing companies to avoid defaults. This also confirms the real-life scenario of risk avoidance, in which financial institutions preferentially select core enterprises with good credit ratings for cooperation with and financing of SMEs.

![Figure 5](image5.png)

**Figure 5.** Effect of credit evaluation on core enterprises.

**4.4.6. Effect of Strength of Supply Chain Relationships**

Figure 6 shows the effect of the strength of supply chain relationships on the default probability of financing enterprises. The figure shows that, when the relationship between financing enterprises and core enterprises remains strong or below a certain level, the default probability of financing enterprises is low. With a closer cooperative relationship
between suppliers and core enterprises in the supply chain, the core enterprises will be more willing to guarantee them in order to achieve win–win cooperation. The difference is that, as shown in the figure, after the strength of supply chain relationships reaches level three (when the relationships are strong), the default probability for financing enterprises significantly increases with further increases in the strength of the cooperation between the two sides. Generally, the closer the cooperation is between a financing enterprise and a core enterprise, the lower the credit risk of the financing enterprise, which seems to be inconsistent with common understandings of supply chain finance. In reality, this may be because the financing enterprise and the core enterprise are too closely related, causing SMEs to heavily rely on core enterprises. However, if the core enterprises encounter risks in their operation, due to the dependence of the financing enterprises on the core enterprises, the related risks of the core enterprises may directly affect the financing enterprises, resulting in an increase in their credit risk.

![Figure 6. Effect of strength of supply chain relationships.](image_url)

5. Concluding Remarks

On the basis of the characteristics of manufacturing enterprises, we preliminarily selected financial credit risk evaluation indicators for the supply chain of manufacturing SMEs. Then, we used SPSS 26 software to analyze the correlations between the primary indicators and eliminate the indicators with strong correlations in the correlation test. Finally, we obtained a financial credit risk evaluation system for the manufacturing supply chain that consisted of 18 evaluation indicators in four categories: the overview of the financing enterprises, the asset status of financing companies, the overview of the core enterprises, and the operation of supply chains.

Taking the manufacturing SMEs listed on China’s stock market as the data sample, we used SPSS Modeler 18 software to experimentally compare four common supervised learning algorithms: random forest, logistic regression, multilayer perceptron, and support vector machine. The results showed that the prediction accuracy and type I error rate for the random forest model were the best, indicating that the random forest model was more suitable for the credit risk assessment of China’s manufacturing SMEs.

Based on the random forest model, a PDP model was established using Python 3.7 software. Then, the credit risk prediction results were visually analyzed. The results revealed that the indicators for the credit rating and the asset–liability ratio of core enterprises were in line with the general perception of credit risk prediction results. However, the indicators for the strength of supply chain relationships and the cash ratio were different from the general understanding. In other words, the two indicators significantly increased the probability of default risk after reaching a certain level. In terms of R&D investment, the probability of default risk first decreased and then increased with the increase in R&D investment.

On the basis of the above research conclusions, the main suggestions are as follows.
The selection of credit risk evaluation indicators for manufacturing supply chain finance should take into account both financial indicators and non-financial indicators. In terms of financial indicators, financing enterprises should focus on strengthening the true records and disclosures of their finances. Non-financial indicators should be combined with the situations of specific enterprises. For example, in terms of R&D innovation for manufacturing products, attention should be paid to enterprises’ proportions of R&D investment. In addition, SMEs and core enterprises should maintain a stable and transparent trading environment, as this is more conducive to the acquisition of credit risk indicators.

In the financial risk assessment of the supply chain, each participant should make timely adjustments to their operation according to the credit risk assessment results. Manufacturing SMEs should make reasonable production and sales plans according to their operating conditions to avoid an increase in their credit risk due to improper operation. For example, financing enterprises should make reasonable use of their current assets to increase the profitability of cash flow assets and create reasonable production plans to prevent the risk of capital fracture. To ensure the stability of product supply, core enterprises should actively participate in supply chain finance-related services, efficiently adjust their relationships with SMEs, and provide credit support for SMEs in supply chain finance. In terms of the supply chain, although core enterprises are the core and entry point of supply chain finance, SMEs should properly handle their dependency relationships with core enterprises to avoid excessive reliance on them, which can result in additional risks to their development. In addition, when planning the production of products, enterprises in the supply chain should make arrangements according to the future popularity of products and the development prospects of the industry to avoid business risks caused by blindly expanding market share.

In terms of limitations, the research data samples in this paper were somewhat constrained, and all were listed enterprises. In future research, a large amount of raw data for SMEs should be collected, which would help improve the performance of the models and reflect the real prediction results. Secondly, in this paper, only four kinds of models were compared. In future research, different types of evaluation models could be included for a comparative analysis, and a more suitable model for manufacturing supply chain credit risk evaluation could be obtained.

**Author Contributions:** Data processing, T.X. and Z.-K.W.; Conceptual analysis, Y.X., M.-X.W. and L.-J.T.; Methodology, Y.X., T.X. and L.-J.T.; Project administration, M.-X.W.; Software, T.X. and Z.-K.W.; Writing—original draft, T.X. and Y.X.; Revisions, Y.X., Z.-K.W. and M.-X.W.; Writing—review and editing, M.-X.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Henan Provincial University Philosophy and Social Science Innovation Team and the Henan Provincial University Humanities (2019-CXTD-04) and Social Science Key Research Base (2015-JD-04).

**Acknowledgments:** The authors thank the School of Management, Henan University of Technology, for its support of this study.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


10. Zhang, Y.; Zhao, H.; Li, B.; Zhao, Y.; Qi, Z. Research on credit rating and risk measurement of electricity retailers based on Bayesian Best Worst Method-Cloud Model and improved Credit Metrics model in China’s power market. *Energy* 2022, 252, 124088. [CrossRef]


**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.