

Federated Learning-based Financial Risk Early Warning Model for Baijiu Enterprises

Shijie He^{*}, Xiaoyu Yang

School of Management, Sichuan University of Science & Engineering, Zigong, 643000, China

^{*} Corresponding author: Shijie He (Email: 408915070@qq.com)

Abstract: As the development of economic globalization deepens, the early warning and management of corporate financial risks have become increasingly crucial. This paper addresses the characteristics of the baijiu (Chinese liquor) industry and proposes a federated learning-based financial risk early warning model to achieve a balance between data sharing and risk prediction among baijiu enterprises. Through federated learning, different enterprises can collaboratively train an early warning model while safeguarding data privacy, thereby enhancing prediction accuracy and comprehensiveness. This paper begins by reviewing federated learning and the field of financial risk early warning. It subsequently presents the model's design and implementation, and concludes by demonstrating the model's effectiveness and superiority in baijiu enterprise financial risk early warning through experiments.

Keywords: Financial Risk Warning; Federated Learning; Secure Aggregation; Communication Overhead.

1. Introduction

With the in-depth development of economic globalization, the early warning and management of corporate financial risks are becoming more and more important [1]. In today's business environment full of challenges and opportunities, enterprises are faced with risks such as increasingly complex and changeable market competition, policy adjustments and supply chain fluctuations. Especially in the field of highly competitive and market fluctuations like the liquor industry, early warning of financial risks has become particularly urgent. Considering the characteristics of liquor enterprises, such as long product life cycle, strong brand influence, fluctuations in market demand, etc., the identification and prediction of financial risks are of great significance for their steady development.

Liquor enterprises face diversified risks in the course of operation, such as fluctuations in raw material prices, unstable market demand, intensified market competition, and policy changes [2-4]. These risks may lead to the adjustment of the production plan of the enterprise, the increase of financial pressure and even the emergence of business difficulties. Taking the supply chain problems in the liquor market in recent years as an example, fluctuations in raw material prices and supply chain interruptions may affect the production and sales of enterprises, thus directly affecting the profitability and financial stability of enterprises. Therefore, liquor companies need an effective financial risk early warning mechanism to take corresponding measures before the risk occurs to reduce the negative impact of losses and risks.

In order to realize the effective monitoring and forecasting of financial risks, researchers and practitioners have proposed a variety of financial risk early warning methods [5-8]. In the past, traditional statistical model methods and machine learning methods have played an important role in the field of financial risk early warning.

A common traditional statistical model is Bedford's Law, also known as the law of the first digit. Its basic idea is that in real financial data, the distribution of the first digit of a number has a specific statistical law. By analyzing the

distribution of the first digit of the enterprise's financial data, abnormal situations such as false reports and fraud can be found [9]. In addition, traditional statistical methods such as the Logistic model are also widely used in financial risk early warning. These methods can effectively identify anomalies under certain conditions, but limited by the complexity of data and nonlinear relationships, their application range is gradually limited.

With the rapid development of machine learning technology, more and more studies have begun to introduce machine learning methods into the field of financial risk early warning. Support vector machines (SVM) [10], random forests [11], neural networks [12] and other methods are widely used to build financial risk early warning models. These methods can more accurately identify potential financial risks by learning the complex patterns and relationships of the data. For example, SVM can nonlinearly model the relationship between different financial indicators, random forest can handle a large number of financial data features, and neural network is suitable for mining hidden data patterns. While improving the accuracy of early warning, these machine learning methods are also more adaptable to the data characteristics of different industries and fields.

However, the existing financial risk early warning methods are often restricted by the problem of data islands when facing specific fields such as the liquor industry. The data distribution and characteristics of different enterprises are different, which makes it difficult to obtain a global early warning model by traditional methods. At the same time, the sensitivity of financial data makes enterprises face the risk of privacy leakage in the process of data sharing.

In order to solve the above problems, this paper proposes a financial risk early warning method based on federated learning. As a distributed machine learning method, federated learning allows different companies to jointly train early warning models while protecting data privacy. This method can make full use of data differences between enterprises, realize cross-enterprise risk prediction, and improve the accuracy and robustness of early warning models while protecting data privacy. Through federated learning, different

companies can jointly train a global financial risk early warning model without sharing original data, which overcomes the problem of data islands.

Our contributions are as follows.

1. Data sharing and comprehensive prediction: In specific industries such as liquor companies, the data distribution and characteristics of different companies are quite different, making it difficult to establish a global risk early warning model with traditional methods. The financial risk early warning method based on federated learning makes full use of the diversity of data among enterprises by allowing multiple enterprises to jointly train models under the premise of privacy protection. This method can realize cross-enterprise comprehensive risk prediction without sharing raw data, and improve the accuracy and comprehensiveness of early warning models
2. Privacy protection and data security: In the process of financial risk warning, enterprises need to process sensitive financial data, and data privacy protection has become a key challenge. The method based on federated learning avoids the direct sharing of raw data by training the model locally and sharing only the model parameters, thereby ensuring the privacy and security of financial data. Especially in industries such as liquor companies that need to protect core business information, this method provides a feasible solution for data privacy.
3. Adaptability and real-time: In the face of market fluctuations and industry changes, liquor companies need to adjust early warning models in time to deal with emerging risks. The financial risk early warning method based on federated learning has the advantage of strong adaptability. Each enterprise can carry out local model training according to its own risk situation, and update the model parameters at any time to adapt to the real-time changing market environment. This real-time feature makes the early warning model closer to the actual situation and more operable.

2. Federated Learning and Financial Risk Early Warning

2.1. Overview of Federated Learning

Federated learning is a distributed machine learning approach that enables multiple participants to collectively train models while preserving data privacy. Each participant trains a local model using its data, and aggregates model parameters through an algorithm on a central server to update the global model. Federated learning has advantages in privacy protection, data sharing, collaborative model development, etc., and is suitable for cross-enterprise risk prediction.

Each entity trains a local model using its data and computes local model updates. These updates are then securely aggregated into the global model, forming a new global weight vector. This process is repeated over multiple communication rounds until convergence.

The distributed nature of Federated Learning allows entities to retain ownership and control over their data while benefiting from the shared global model. In the context of financial risk warning within the liquor industry, this approach enhances collaboration among companies while safeguarding sensitive financial information.

Federated learning proposes to solve this dilemma in a way that the data does not move, the model moves, and the data is available and invisible. H. Brendan McMahan et al [13] first proposed the federated averaging algorithm FedAvg (Federated Averaging) for federated learning in 2017, but the workload of each terminal in this algorithm is the same. However, in actual scenarios, the available computing resources of different terminals are different and the data is highly heterogeneous. To solve this problem, Tian et al. [14] proposed an improved FedProx algorithm, which allows the system to perform variable workloads according to the available computing resources of different terminals to avoid being forced to exit due to excessive load. Karimireddy et al [15] proposed an improved algorithm SCAFFOLD for heterogeneous data problems. When the terminal is highly heterogeneous in data, compared with FedAvg, this algorithm can avoid the development of the global model to the local optimum and speed up the convergence speed.

2.2. Financial Risk Early Warning

Financial risk early warning refers to analyzing corporate financial data, discovering abnormal changes, and predicting potential risks. Traditional methods usually rely on historical data and rules, making it difficult to cope with emerging risk scenarios. Federated learning can realize joint modeling of cross-enterprise financial data and improve the accuracy and timeliness of forecasting.

Existing financial risk early warning methods such as expert evaluation method, decision tree method and deep neural network method (DNN), mainly obtain the overall data of the liquor industry chain, and evaluate the risk situation of the enterprise through the overall data of the industry, there is a lack of consideration of the internal risk elements of the industrial chain, so there is a certain deviation in the effect of risk assessment. In addition, the characteristics of financial data at different times have certain correlations, but existing methods often treat them independently, ignoring the characterization of the continuity of the enterprise, resulting in a lack of assessment of the internal risk transmission and impact of the enterprise. In addition, existing early warning methods often rely on a large amount of data and information to assess risks. However, it is difficult to obtain complete and accurate data, especially highly sensitive financial data, so the early warning models obtained are often one-sided, which may lead to bias and inaccuracy in the evaluation results. The above limitations make it difficult for existing methods to identify risks accurately and early.

3. Our Scheme

3.1. Federated Data Architecture

As shown in Figure 1, in the federated learning architecture, multiple liquor companies are included as participants, and each company has its own private financial data. By sharing the local financial risk early warning model, the participants realize model fusion on the server side, and then obtain a financial risk early warning model with stronger generalization ability and wider practicability.

Our architecture is divided into server-client two layers. There are 2 objects in total, namely (1) Liquor enterprise participants: responsible for the collection and preprocessing of financial data, obtaining the framework of the financial early warning model from the server, and locally The network model is trained periodically, and the model parameters are

differentially perturbed; (2) Server: Aggregate the local model parameters uploaded by each participant, and send the aggregated global model parameters to the participants to

update the model for the next round of training parameter. The relevant object parameters are shown in Table 1.

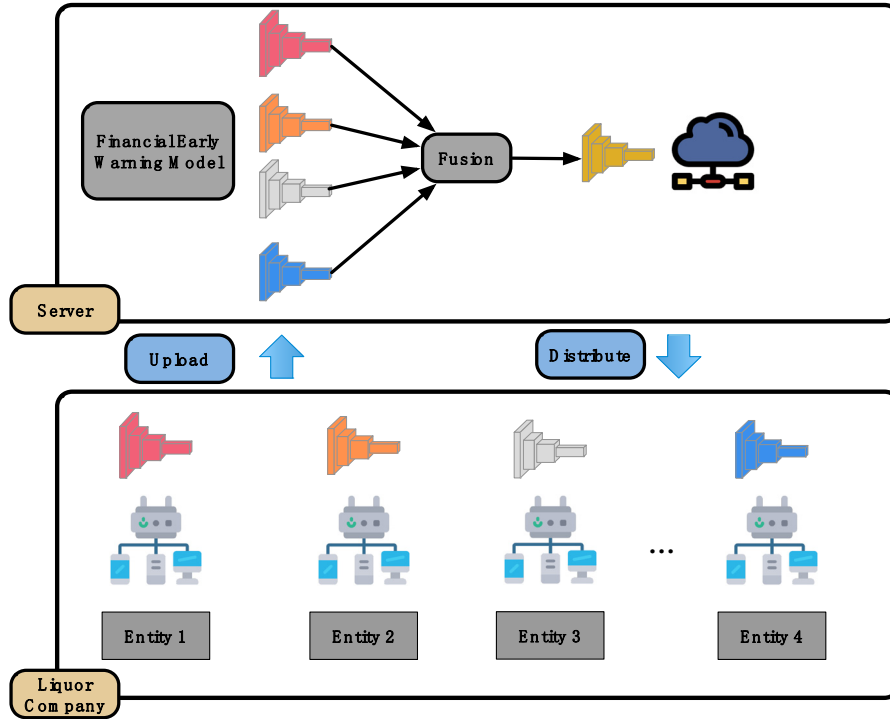


Figure 1. Architecture diagram of federated risk early warning model

Table 1. Related parameter table

Number	Name	Description
1	d_i	The i -th liquor company
2	D_i	The financial data of the i -th liquor company
3	$g_i^t(k)$	The gradient obtained by iterating k times in the t -th round of training on the i -th liquor company
4	$\omega_i^t(k)$	The model parameters obtained by iterative training k times in the t -th round of training on the i -th liquor company
5	ω^t	The model parameters obtained after the t th round of training on the server
6	Y_i	The model output of the i -th liquor enterprise
7	T_i	The financial data labels of the i -th liquor company

3.2. Financial Risk Evaluation Index System of Liquor Enterprises

Financial changes run through all aspects of the production and operation of the enterprise, and every activity of the enterprise will cause changes in the financial affairs of the enterprise. As an objective phenomenon, financial risk is the concentrated expression of various risks in financial changes, and the accumulation of financial risks will have a very serious impact on the production, operation, and even survival of enterprises. Therefore, by studying the external manifestations of financial risks, analyzing the deep-seated causes of financial risks, and adopting effective methods to identify, monitor and prevent risks in a timely manner is an urgent need for the development of enterprises and even the development of the entire industrial chain. Financial risk assessment can reflect the situation of operation and production in a timely manner and discover problems in production and operation in a timely manner.

This paper adopts the factor analysis method to identify the financial risk factors of enterprises, and divides them into the following four aspects: solvency, profitability, operating ability, and growth ability. The ability to repay debts by

generating income from its own assets and production and operation activities. Profitability represents the ability of an enterprise to obtain profits, and can reflect the relevant problems existing in the operation of the enterprise. Operational capacity refers to the ability of an enterprise to use its own assets for production and operation and obtain profits, which mainly reflects the operating efficiency of the enterprise. The growth capability is a measure of the company's expansion and business growth capabilities. In addition, this paper also introduces two non-financial indicators: the number of lawsuits and the number of penalties for violations, and together with the above four indicators, a financial risk assessment indicator system for the electronic information industry chain is constructed. The indicators are shown in the table below:

Through the above four dimensions, and under each dimension, set sub-indicators respectively. Based on the data collected by the above indicators, after data preprocessing, as the financial characteristic parameters of each enterprise, it is input into the risk early warning model to achieve the corresponding risk level classification.

Table 2. Related parameter table

Number	Category	Financial indicator name	Calculation formula
1	Solvency A	Current Ratio A1	Current assets/Current liabilities X100%
		Quick ratio A2	Quick assets/Current liabilities X100%
		Cash Ratio A3	Monetary Fund/Current Liabilities X100%
		Asset-liability ratio A4	Total liabilities/Total assets X100%
		Net gearing ratio A5	Total liabilities/shareholders' equity X100%
		Cash flow to debt ratio A6	Net operating cash flow/Total liabilities X100%
2	Profitability B	Return on net assets B1	Net profit/shareholders' equity X100%
		Net rate of total assets B2	Net profit/total assets X100%
		Net sales rate B3	Net profit/operating income X100%
		Operating income per share B4	Operating income / annual weighted average total share capital X100%
3	Operating capability C	Inventory turnover ratio C1	Operating income / average inventory balance
		Accounts receivable turnover ratio C2	Total operating income/Average balance of accounts receivable
		Current asset turnover ratio C3	Total operating income/average balance of current assets
		Non-current assets turnover ratio C4	Total operating income/average balance of non-current assets
		Total asset turnover ratio C5	Total operating income/average total assets
4	Growth ability D	Growth rate of operating income D1	Total operating income growth of the current year/Total operating income of the previous year X100%
		Growth path of net assets per share D2	Net asset growth at the end of the year/total net assets at the beginning of the year X100%
		Growth rate of total assets D3	Growth of total assets at the end of the year/total assets at the beginning of the year X100%

3.3. Risk Warning Model Training

3.3.1. Data Preprocessing

Firstly, the corresponding items in the financial data are screened out according to the financial risk indicators constructed above, and secondly, the corresponding items are concatenated into a financial feature vector, and the expression is as follows.

$$F = [f_1, f_2, f_3, \dots, f_n]. \quad (1)$$

Where F represents a feature vector, and the data set D_i of each enterprise d_i is composed of F. The component in F is the index corresponding to the financial risk.

3.3.2. Model Training

Assuming that the liquor company d_i has a data set D_i , including data and labels, the data set is divided into a training set and a test set according to a ratio of 7:3, and then the training set is used as the input of the neural network model for model training. Let the output of the model be Y, then:

$$Y = f(X_i). \quad (2)$$

Comparing Y and labels, calculate the accuracy rate and loss value. After each round of training is completed, the test data set is input into the neural network, and the accuracy and loss are calculated. When the threshold of the number of local training rounds is reached, the model parameters are differentially perturbed and uploaded to the server.

The local gradient calculation formula is as follows:

$$g'_i(k) = \nabla [f(X_i) - T_i]. \quad (3)$$

The model gradient is clipped, and the clipping formula is as follows:

$$g'_i(k) = \frac{g'_i(k)}{\max(1, \|g'_i(k)\|)}. \quad (4)$$

After obtaining the model gradient of the kth training, judge whether k is less than k1, if k is less than the threshold k1, then update the model parameters locally, the update formula is as follows:

$$\omega'_i(k+1) = \omega'_i(k) - \eta * g'_i(k). \quad (5)$$

If k is equal to k1, the model parameters will be sent to the server for aggregation. Since the server is not completely trustworthy, we add differential privacy to the model parameters locally and use a random disturbance mechanism. Suppose the probability that the model parameters are stolen by the attacker is p, and the mapping function from the data set to the model parameters is h, and the mapping formula is as follows:

$$\omega'_i(k) = h'_i(D_k). \quad (6)$$

4. Experimental Verification and Analysis

In order to verify the effectiveness and superiority of the financial risk early warning model based on federated learning, we used real liquor enterprise data sets and conducted comparative experiments with traditional methods. The experiment aims to fully demonstrate the significant advantages of the model in terms of prediction accuracy, risk coverage and data privacy protection.

4.1. Experiment Settings

We collected the financial data of Maotai, Wuliangye, Gansu Huangtai and other enterprises, covering key financial indicators such as sales revenue, profit, and asset-liability ratio. In the experiment, we used two methods for comparison: the traditional statistical model method and the financial risk early warning model based on federated learning.

4.2. Evaluation Indicator

The financial early warning model used in this paper is a classification model. In order to accurately evaluate the performance of the model classification and measure the classification effect, this paper uses precision rate (Precision), recall rate (Recall), and F1 score as the indicators of model evaluation. At the same time, it combines the ROC curve and The AUC value can more intuitively evaluate the effect of the classification model.

In order to better understand the calculation formula of the five categories, a three-category confusion matrix is introduced:

Table 3. Three-category confusion matrix

Confusion Matrix		Predict		
		0	1	2
Real	0	a	b	c
	1	d	e	f
	2	g	h	i

(1) Accuracy

Precision (Precision), also known as precision, is to calculate the probability that all samples predicted to be 0 in the test set are actually 0. The calculation formula is:

$$p = \frac{a}{a + d + g}$$

(2) Recall Rate

The Recall Rate (Recall), also known as the recall rate, is mainly for the original sample. It calculates the probability of predicting 0 in the sample that is actually 0 in the test set. The calculation formula is:

$$p = \frac{a}{a + b + c}$$

(3) F1 Score

F1 Score is a common indicator for multi-category effect evaluation, and its general calculation formula is:

$$F_{\beta} = \frac{(1 + \beta^2) \times P \times R}{(\beta^2 \times P) + R}$$

4.3. Experiment Analysis

Table 4. Experimental results comparison table

Number	Method	Accuracy	Recall Rate	F1 Score
1	Traditional statistical model	89.3%	90.4%	0.86
2	Federated learning	97.2%	98.6%	0.97

The experimental comparison results in Table 4 clearly show that the financial risk early warning model based on federated learning far exceeds the traditional method in terms of prediction accuracy and risk coverage. For the data of Moutai, Wuliangye, Gansu Huangtai and other enterprises, the federated learning model can more accurately identify potential financial risks and provide enterprises with more

accurate early warning information. Especially considering the market share and influence of these enterprises, accurate financial risk early warning is crucial to maintain their steady development.

In addition, the financial risk early warning model based on federated learning also shows obvious advantages in data privacy protection and data sharing feasibility. The sensitive financial data of the enterprise can be processed locally, and cross-enterprise collaboration can be realized only through the sharing of model parameters, thereby ensuring the security of data privacy. This privacy protection feature is particularly important for industries such as Maotai, Wuliangye and other liquor companies that need to protect core business information.

4.4. Experiment significance and inspiration

The experimental results not only further verify the excellent performance of the federated learning-based financial risk early warning model in the financial risk management of liquor enterprises, but also provide a useful reference for risk early warning research in other similar industries. Experiments have proved that the federated learning method can not only improve the accuracy and comprehensiveness of prediction, but also achieve a balance between data privacy protection and data sharing, and provide more reliable support for enterprise decision-making.

5. Conclusion

The financial risk early warning model based on federated learning proposed in this paper has shown obvious advantages in the field of liquor enterprises. This model can not only improve the prediction accuracy and risk coverage in financial risk prediction, but also solve the data privacy problem and promote the feasibility of data sharing. This study provides an innovative method for the early warning of the financial risk of liquor enterprises, and provides a useful reference and enlightenment for the risk management research of other similar industries.

Acknowledgments

This work was supported by Sichuan University of Science & Engineering.

References

- [1] Landi G C, Iandolo F, Renzi A, et al. Embedding sustainability in risk management: The impact of environmental, social, and governance ratings on corporate financial risk[J]. Corporate Social Responsibility and Environmental Management, 2022, 29(4): 1096-1107.
- [2] Bai X, Li X, Wang W. The impact of corporate social responsibility on competitiveness of liquor enterprises [C]// 2022 7th international conference on financial innovation and economic development (ICFIED 2022). Atlantis Press, 2022: 1289-1300.
- [3] Chen Q, Hong Q, Jiang W, et al. Analysis of the Development and Trend of Liquor Industry in China[C]//2022 7th International Conference on Financial Innovation and Economic Development (ICFIED 2022). Atlantis Press, 2022: 2465-2471.
- [4] Zhang H, Yang X, Huang Z. Influence Mechanism of Cross-Regional Collaborative Innovation Performance of Chinese Liquor Enterprises—Based on Knowledge Collaboration [C]// International Conference on Management Science and

- Engineering Management. Cham: Springer International Publishing, 2022: 210-225.
- [5] Zhu W, Zhang T, Wu Y, et al. Research on optimization of an enterprise financial risk early warning method based on the DS-RF model[J]. International review of financial analysis, 2022, 81: 102140.
 - [6] Zhu Z, Liu N. Early warning of financial risk based on K-means clustering algorithm[J]. Complexity, 2021, 2021: 1-12.
 - [7] Wei X. A method of enterprise financial risk analysis and early warning based on decision tree model[J]. Security and Communication Networks, 2021, 2021: 1-9.
 - [8] Du G, Liu Z, Lu H. Application of innovative risk early warning mode under big data technology in Internet credit financial risk assessment[J]. Journal of Computational and Applied Mathematics, 2021, 386: 113260.
 - [9] Sun Y, Jian W, Fu Y, et al. A new perspective of credit scoring for small and medium-sized enterprises based on invoice data[C]//2021 2nd International Conference on Computing and Data Science (CDS). IEEE, 2021: 477-482.
 - [10] Mohammadi M, Rashid T A, Karim S H T, et al. A comprehensive survey and taxonomy of the SVM-based intrusion detection systems[J]. Journal of Network and Computer Applications, 2021, 178: 102983.
 - [11] Tyralis H, Papacharalampous G, Langousis A. A brief review of random forests for water scientists and practitioners and their recent history in water resources[J]. Water, 2019, 11(5): 910.
 - [12] Choi R Y, Coyner A S, Kalpathy-Cramer J, et al. Introduction to machine learning, neural networks, and deep learning[J]. Translational vision science & technology, 2020, 9(2): 14-14.
 - [13] H.Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data [J]. Artificial Intelligence and Statistics, 2017.
 - [14] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. arXiv preprint arXiv:1812.06127, 2018.
 - [15] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. Scaffold: Stochastic controlled averaging for federated learning. In International Conference on Machine Learning, pages 5132–5143. PMLR, 2020.