International Journal for Housing Science and Its Applications
Publish August 4, 2025. Volume 46, Issue 3 Pages 3289-3297

https://doi.org/10.70517/ijhsa463277

Data Mining and Risk Analysis of Accounting Statements Based on Regression Analysis

Peiling Quan[1,*], Tianyue He[2] and Yinzhi Yu[3]
[1] School of Accounting, Anhui University of Finance and Economics, Bengbu, Anhui, 233030, China
[2] School of Economics, Anhui University of Finance and Economics, Bengbu, Anhui, 233030, China
[3] School of Business Administration, Anhui University of Finance and Economics, Bengbu, Anhui, 233030, China
Corresponding authors: (e-mail: peiling_quan@163.com).

Abstract The proliferation of corporate financial data makes it more and more difficult to mine from a large number of accounting statements that have value and identify potential risks of a company. This study collects the historical accounting statement data of Company J from 2020 to 2024, and screens out the key financial indicators through data mining techniques. Meanwhile, the financial risk status of Company J in recent years is analyzed by combining the accounting statement data. Then, using the probability of the company's financial risk as the dependent variable, the correlation between the financial indicators and the probability of risk emergence is analyzed, and a regression prediction model is established. The total assets of Company J in 2020~2024 increase year by year, and by the end of 2024, the total assets reach 367.41 billion yuan. The company's short-term solvency in 2020~2023 is weak, and the cash ratio, quick ratio, and current ratio are lower than the industry average. The study extracted six main factors affecting the company's financial risk, and its overall variance is 85.58%. Logistic regression analysis shows that a 1-unit increase in the values of the six main factors reduces the likelihood of the company's financial risk by 0.209~4.056. This study provides a quantifiable analytical method for accounting statement data mining, which can help enterprises to strengthen the control of financial risk and provide references for investment analyses. Provide reference.

Index Terms Data mining, regression prediction, accounting statement, risk analysis

I. Introduction

Enterprises face increasingly complex operational challenges in a competitive business environment. And accounting statements, as a centralized representation of an enterprise's financial condition and operating results, are an important tool for management and external stakeholders to understand the health of the enterprise [1]. Through in-depth analysis of accounting statements, enterprises can assess their financial health, identify potential business risks, and formulate reasonable development strategies, but there are some problems in actual operation [2]-[4]. Therefore, it is of great practical significance to discuss in depth the problems existing in the analysis of enterprise accounting statements and put forward corresponding countermeasures.

The collection of financial information has a certain complexity, the collection is difficult, the number of large, need to consume a lot of time [5]. And many enterprises presented to the management of the financial statements are mostly presented through a number of processing, organizing and preliminary calculations, so that the listed data are basically rough, once the technology is not mature, it will lead to a lack of accuracy in the analysis of financial data [6]-[8]. And with the continuous development of information technology, the survival posture of enterprises has also changed, the emergence of big data can quickly help the financial sector to establish analysis tools [9], [10]. Enterprises are paying more and more attention to the use of financial data in the context of economic globalization, and the use of data mining tools to extract the value behind the data is an important way to meet the competitive requirements of today's corporate financial shared services [11], [12]. Therefore, the construction of data mining and risk analysis framework construction based on accounting statements becomes the focus of enterprise financial management, and only the stable development of big data technology can ensure the security and stability of management accounting work in Chinese enterprises [13]-[16].

This paper mines the accounting statement data of Company J over the years, and determines the corporate financial risk impact factors centered on current ratio and other indicators through literature review and expert consultation. The collected financial data is used to fill in the missing values and remove the outliers by year, and the data is normalized by min-max standardization method to eliminate the influence of the scale. Combined with the above data, the asset composition as well as solvency risk of Company J is analyzed. Using the forced extraction method, the main influencing factors are extracted from the financial indicators, and a logistic regression model is constructed to predict the probability of financial risk under the perspective of accounting statement data.

II. Study of financial risk in Company J

II. A.Forms of organization

Established in June 1998, Company J's main business is solid waste treatment, with its core focus on two main aspects, including investment, construction and operation of domestic waste incineration power generation, and synergistic development of kitchen waste, sludge, medical waste, hazardous waste disposal and other businesses. The waste incineration power generation business is mainly operated under the BOT and other franchising modes. During the franchising period of the waste incineration power generation project, the company can obtain revenue from waste treatment services and electricity sales through waste incineration.

Company J adopts the “linear + matrix” integrated management mode, in line with the “streamlined and efficient organization, full workload, priority of benefits, taking into account the principle of fairness”, the orderly realization of waste incineration power generation project operation and maintenance of integrated management mode, in line with the rational setup of positions, In order to optimize the allocation of human resources and achieve the goal of maximizing the benefits of the company and the interests of the staff, the company has set up EHS in addition to the Production Department, Comprehensive Department and Finance Department.

II. B.Financial risk early warning management system

Having a good risk management structure is twice as effective as having a good risk management structure when running a business. Pay attention to risk management policies and clarify the basic principles and objectives of risk management. Be familiar with the risk management process and be able to describe in detail the steps of risk identification, assessment, control and monitoring. Be able to utilize good risk assessment tools, using qualitative or quantitative tools to assess the level of risk. Utilize the Risk Management Information System (RMIS), a system primarily used to collect, store and analyze risk management data, to encourage the active participation of employees in risk management as much as possible, and to ensure the effective implementation of risk management measures. Comprehensively analyze the causes of risks, prepare detailed risk management reports, comprehensively monitor and manage early warning reports, and formulate effective risk management strategies and implement them into the company's management process as soon as possible. While the risk management system established by Company J includes risk management organization structure, job responsibilities, risk classification, risk management assessment criteria, etc., there are still some limitations in the specific qualitative and quantitative assessment of the risk early warning system.

III. Accounting statement data mining and indicator selection

III. A. Selection of data and indicators

In this paper, based on the accounting statement data indicators of J Company's financial database, from 2020 to 2024, all the financial analysis factors are selected to total 243 indicators. They are 35 per share factors, 41 profitability factors, 13 quality of earnings factors, 12 cash flow factors, 19 capital structure factors, 24 solvency factors, 15 operational capacity factors, 11 growth category factors, 34 financial report derived factors (MRQ) and 39 financial report derived factors (MTT). Through reviewing the literature as well as consulting with experts, we finally selected: current ratio, liquidity, gearing ratio, degree of company specialization, price stability, enterprise scale, net profit growth rate, sales profit margin, sales revenue growth rate, industry competitiveness, degree of cooperation, return on net assets, and accounts receivable turnover, a total of 13 indexes as the key factors for the financial risk study, which are recorded in order as V1~ V13.

III. B. Data pre-processing

In order to analyze the link between accounting statement data indicators and financial risk, the collected financial statement data are preprocessed to reduce data noise interference. Due to the non-uniformity of the quantitative outline of each indicator as well as the existence of missing and abnormal data, it is necessary to first carry out data normalization, as well as to fill in the missing values and remove abnormal values.

The first is to remove outliers, first remove the industry with missing values, this paper uses the 3σ principle, the reason why the 3σ principle is used because the data after the normal distribution test is to obey the normal distribution. Here different years are taken into account, so the outliers are removed according to the upper and lower quartiles of different years.

Data normalization refers to the elimination of inconsistencies in the quantitative outline of each indicator, so that each indicator can be comparable, this paper, according to the characteristics of the accounting statement data, chose the min-max standardization method [17], and its specific conversion function is:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

where $\min(x)$ and $\max(x)$ denote the minimum and maximum values of this accounting statement data among all the samples, respectively, and x' denotes the value of the accounting statement data index of the sample after normalization, by which the original data is mapped between [0,1].

III. C. Financial risk analysis of J

III. C. 1) Asset composition risk

In recent years, the total operating income of Company J has continued to grow, and the asset structure of the enterprise is still dominated by non-current assets such as inventory assets, and the increase in cooperative development projects has led to a significant increase in the remaining accounts receivable and further enlargement. There is a risk that the company will be squeezed out of funds, and the size of restricted assets will continue to expand, which will negatively affect the refinancing ability and asset liquidity. According to the company's accounting statements for the fourth quarter of 2024, at the end of 2024, the company's total assets were 367.41 billion yuan, an increase of 7.21% from the end of 2023, of which current assets accounted for 85.68%. The asset composition of Company J in 2020~2024 is shown in Figure 1.

At present, the company's current assets are still growing and large. As of the end of 2024, the company's total current assets will be 314.813 billion yuan, and the proportion of current assets such as other monetary funds, inventory, and receivables is relatively high. As the number of development projects increases and the scale of the business continues to expand, so does the size of the company's inventory. According to statistics, the completed assets and unsettled assets formed by the company's project development costs, product development and construction contracts accounted for 85.7%, 8.2% and 3.1% respectively, the inventory limit ratio was 40.25%, and the total depreciation was 543 million yuan. In 2024, the company's inventory turnover will be 0.64 times, a year-on-year increase of 0.22 times. At the end of 2024, the company's capital volume continued to increase to 37.26 billion yuan, a year-on-year increase of 24.1%, mainly due to the increase in the company's borrowing scale and the large amount of liquidity to form a certain degree of support for cash liquidity. At the end of the same period, the restricted part of the monetary funds accounted for 33.8%. With the impact of the increase in cooperative development projects, other receivables have continued to grow in recent years, and the book amount of other receivables at the end of 2024 was 51.796 billion yuan, of which the funds receivable from other related parties, various guarantee deposits, government receivables and reserves, and withholding payments were 23.57 billion yuan, 4.32 billion yuan, 2.54 billion yuan and 460 million yuan respectively. In general, the number of other receivables of Company J is relatively large and the growth trend is obvious, and there is a situation where funds are occupied.

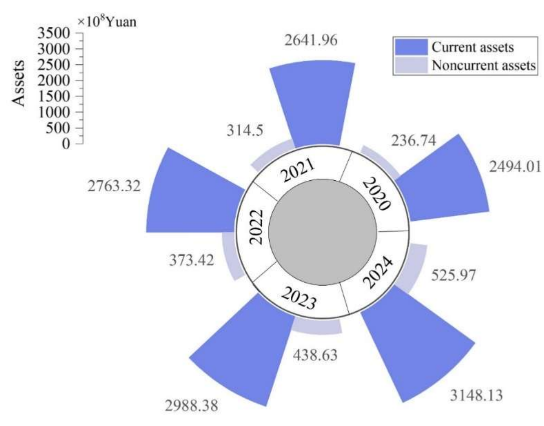

Figure 1: J company 2020 ~ 2024 annual capital production

III. C. 2) Solvency risk

A company's debt repayment capacity can be affected by factors such as the scale of financing and financing channels. Good debt repayment ability is the basis for guaranteeing the company's fulfillment of the debt repayment agreement, and it is also an important basis for evaluating whether the company is healthy and sustainable development and obtaining the possibility of further refinancing. In this paper, four financial indicators are selected to analyze the repayment ability of J Company's long-term debt and short-term debt. Table 1 shows the specific financial data of J Company's debt service ability from 2020 to 2024.

Short-term debt service is to repay the current debt, the resources used for repayment are current assets, it can reflect the relationship between the company's maturing debt and disposable current assets. In terms of measurement indexes, there are current ratio, quick ratio and so on. In this paper, we will analyze the three indicators of cash ratio, quick ratio and current ratio to assess the solvency of assets and immediate liquidity of Company J, and then to assess its short-term solvency.

The data in the table shows that, in terms of the trend of change, the trend of the current ratio in the last five years shows consistency with the trend of the quick ratio in the last five years. Over the five years, both the quick ratio as well as the current ratio reach their peaks in 2024. Compared with the industry mean line, both current ratio and quick ratio of Company J are below the industry average between 2020 and 2023, indicating that in terms of short-term debt repayment, Company J has a low capacity, especially in 2022 at the lowest level.After 2023, Company J has significantly improved its capacity in terms of short-term debt repayment with a positive trend. In financial terminology the ratio of funds that a company can liquidate in time to current liabilities is called cash ratio. It is a dynamic indicator of solvency as a company's repayment of current liabilities is accomplished by the direct use of cash, which intuitively reflects the company's ability to repay short-term liabilities. Although the higher the value of the cash ratio, the stronger the ability of the company to repay debts directly and the lower the risk of debt repayment, too high a value of the cash ratio also means that the flow rate of the company's funds is low, and the funds are ineffective. In the table, Company J's cash ratio peaks at 0.79 in 2024. In addition, the company's cash ratio is lower than the industry average throughout the five-year period, indicating that the company is relatively weak in terms of its ability to repay current debt.

Long-term solvency refers to the ability of a company to repay the principal and interest on all of its debt in the long run. Gearing ratio is equal to the amount of liabilities borne by each dollar of assets. Gearing ratio reflects the capital structure of the enterprise, and is an important indicator of the enterprise's long-term solvency. The gearing ratio shows how much of the enterprise's total source of funds is provided by creditors, so this indicator is chosen to analyze the long-term debt. As can be seen from the table, the overall balance sheet ratio of company J was stable before 2022, basically the same as the industry average, and the decline was larger after 2022, and the overall trend of decline and change between 2022 and 2021 was reversed in the direction of the average value of the peer group, which reflects the importance of company J's own long-term liabilities and the implementation of effective measures to reduce the debt ratio.2023 annual report shows that company J's The net profit of shareholders is 438 million yuan, and the owner's equity has been improved to some extent.After 2023, compared with the average level of the industry's gearing ratio, the level of J company's gearing ratio is always below it. It shows that the capital structure of the company tends to be rationalized, which reduces the pressure of long-term debt repayment, and it also reflects that the company has a more reliable protection for the owner's equity, and in the future, the difficulty of financing risk will be reduced.

Table 1: 2020~ 2024 the financial data of J Company’s solvency

		2020	2021	2022	2023	2024
Mobility ratio	L	1.78	1.99	1.75	2.28	3.63
	Mean	2.55	2.16	2.49	2.59	2.52
Quick motion ratio	L	1.54	1.38	1.28	1.47	2.65
	Mean	2.11	1.65	2.49	1.94	2.34
Cash ratio	L	0.16	0.06	0.32	0.30	0.79
	Mean	0.82	0.64	0.76	0.85	0.96
Assets liability ratio	L	39.96	41.62	42.39	30.08	20.75
	Mean	36.32	49.90	43.51	62.82	50.22

IV. Logistic regression-based corporate financial risk prediction study

IV. A. Logistic regression

IV. A. 1) Logistic distribution

Logistic distribution is a continuous distribution if continuous type random variable X, the distribution function of X is:

$$F(x) = P(X \le x) = \frac{1}{1+\exp(-(x-\mu)/\gamma)} \tag{2}$$

Then X is said to follow a logistic distribution with density function:

$$f(x) = F^{'}(x) = \frac{e-(x-\mu)/\gamma}{\gamma(1+e-(x-\mu)/\gamma)^2} \tag{3}$$

where μ is the location parameter that determines the center of the distribution function. γ is the scale parameter and $\gamma > 0$. The logistic distribution has no shape parameter, which makes the image of the logistic distribution probability density function always bell-shaped, and the shape approximates that of the image of the normal distribution probability density function.

IV. A. 2) Regression models

Logistic regression [18], also known as logistic regression, is a common classification model mainly used to solve binary classification problems. Let $X = (X_1, X_2, \cdots, X_p)$ be a p-dimensional covariate, Y be a response variable and Y takes the value of 0 or l, then the conditional probability of $Y=1$ is:

$$P = P\{y_i = 1 \mid x_{i1}, x_{i2}, \cdots, x_{ip}\} \tag{4}$$

where y_i is the value of the response variable for the i th observation sample, and the i th observation sample of X is denoted as $X_i = (x_{i1}, x_{i2}, \cdots, x_{ip}), i = 1, 2, \cdots, n$. n is the sample capacity.

Then the Logistic regression model is:

$$P = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip})}{1+\exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip})} \tag{5}$$

$$= \frac{1}{1+\exp(-(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}))} \tag{6}$$

where β_0 is the constant term and $\beta = (\beta_1, \beta_2, \cdots, \beta_p)^T$ is the regression coefficient.

The logistic regression model is derived from the sigmiod function [19], which is a special case of the logistic distribution function. When $\mu = 0$ and $\gamma = 1$, then $f(x) = \frac{1}{1+e^{-x}}$.

The sigmoid function has a function value between 0 and 1. The meaning of this function is to measure the probability that the input data belongs to 1. Usually 0.5 is chosen as the threshold value, when $sigmoid(x) > 0.5$, it means that the input data belongs to 1. When $sigmoid(x) < 0.5$, it means that the input data belongs to 0.

A logarithmic transformation of Eq. gives the logistic regression model the following form:

$$\ln\frac{P}{1-P} = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \tag{7}$$

IV. A. 3) Model parameter estimation

A great likelihood estimation of the parameters of the logistic regression model gives the likelihood function:

$$\begin{aligned} L(\beta_0, \beta \mid X_i, y_i) = \prod_{i=1}^{n} \left[P\left\{y_i = 1 \mid x_{i1}, x_{i2}, \cdots, x_{ip}\right\}\right]^{y_i} \\ \times\left[1-P\left\{y_i = 1 \mid x_{i1}, x_{i2}, \cdots, x_{ip}\right\}\right]^{1-y_i} \end{aligned} \tag{8}$$

Taking logarithms on both sides of the equation gives the log-likelihood function:

$$\ln L(\beta_0,\beta \mid X_i,y_i)=\sum_{i=1}^{n}\left[y_i\left(\beta_0+\sum_{j=1}^{p}\beta_j x_{ij}\right)-\ln\left(1+e^{\beta_0+\sum_{j=1}^{p}\beta_j x_y}\right)\right] \tag{9}$$

Denoted as $l(\beta_0,\beta)$, and taking the maximum value of $l(\beta_0,\beta)$, the estimates of β_0 and β can be obtained as $\hat{\beta}_0$ and $\hat{\beta}$. Thus, solving for the unknown parameters in logistic regression is transformed into a log-likelihood function solving problem.

Let the estimate of the parameter (β_0,β) be the current estimate of the parameter $(\hat{\beta}_0,\hat{\beta})$, then the objective function can be rewritten:

$$\begin{aligned}(\hat{\beta}_0,\hat{\beta}) &= \arg\min_{\beta_0,\beta}\left(-\frac{1}{2n}l(\beta_0,\beta)\right)\\ &= \arg\min_{\beta_0,\beta}\left(\frac{1}{2n}\sum_{i=1}^{n}w_i(z_i-\beta_0-X_i^T\beta)^2\right)\end{aligned} \tag{10}$$

where $w_i=p(X_i)^{(k)}(1-p(X_i)^{(k)})$, $z_i=\beta_0^{(k)}+X_i^T\beta^{(k)}+(y_i-p(X_i)^{(k)})/w_i$, $p(X_i)^{(k)}=\exp(\beta_0^{(k)}+X_i^T\beta^{(k)})/(1+\exp(\beta_0^{(k)}+X_i^T\beta^{(k)}))$.

IV. B. Screening of key financial risk factors

Before constructing the financial risk evaluation model based on logistic regression, the evaluation indexes were factor analyzed, and the public factors were identified and extracted to ensure the accuracy and effectiveness of the model.

For determining the number of factors, two methods are usually used: one is the method based on the characteristic root, i.e., the importance of the factors is judged according to the size of their characteristic root, so as to determine the number of factors. The second is the way through the overall variance explanation, that is, according to the degree of explanation of the factors to the overall variance to force the extraction of a certain number of factors. In this paper, we adopt the forced extraction method to extract 6 public factors from 13 financial risk factors, labeled as F1~F6. Table 2 shows the factor variance explanation of the factors after factor analysis in detail. As can be seen from the table, the overall variance explained by the first six public factors is as high as 85.58% (≥85%), which is sufficient to represent the vast majority of information in the data.

Table 2: Factor variance interpretation

Constituents	Eigenvalue	Percentage of variance	Cumulative percentage
1	4.213	32.41	32.41
2	2.398	18.45	50.86
3	1.691	13.01	63.87
4	1.206	9.28	73.15
5	0.983	7.56	80.71
6	0.633	4.87	85.58
7	0.454	3.49	89.07
8	0.425	3.27	92.34
9	0.367	2.82	95.16
10	0.265	2.04	97.2
11	0.182	1.40	98.6
12	0.115	0.88	99.48
13	0.068	0.52	100

Interpreting factor loadings is a critical step in clarifying the meaning of public factors during factor analysis. However, the initial factor loading matrix is usually difficult to directly reveal the actual meaning of the public factors because its information is too comprehensive. To solve this problem, the method of rotating the axes is used, aiming to make the factor loadings clearer and closer to the extremes of 1 or 0. The methods of rotating the coordinate axes include orthogonal and oblique rotations, while the variance maximization method in orthogonal rotation is the common method chosen. After seven rotations, the converged factor loading matrix was successfully obtained as shown in Table 3.

According to the data in the table, the following conclusions can be drawn: the loading of factor F1 is more significant in the indicators of current ratio, liquidity, gearing ratio and the degree of specialization of cross-border e-commerce platforms, and F1 is named as the factor of solvency of financing enterprises. Factor F2 loads more significantly on price stability and enterprise scale, so it is named as the financing enterprise asset quality factor. Factor F3 is loaded more significantly on net profit growth rate and sales profit margin, so we name it as financing enterprise profitability factor. Factor F4 is more prominent in loading on sales revenue growth rate, so it is named as financing enterprise development potential factor. Factor F5 has a larger loading on industry competitiveness, so we name it as industry competitiveness factor. Factor F6 loads significantly on years/degree of cooperation, naming it as supply chain quality factor. In the following section, logistic regression analysis of financial risk based on accounting statement data indicators will be conducted according to the conclusion of this section.

Table 3: The factor load matrix of convergence

	F1	F2	F3	F4	F5	F6
V1	0.970	0.077	0.049	0.036	0.085	-0.092
V2	0.953	0.145	0.070	-0.018	0.090	-0.148
V3	-0.835	0.119	-0.008	-0.141	0.109	0.023
V4	-0.539	-0.084	-0.070	0.376	0.323	0.396
V5	0.122	1.011	-0.053	0.091	0.046	0.043
V6	0.213	0.812	0.150	-0.180	0.174	-0.307
V7	-0.112	0.067	0.890	0.189	-0.048	0.009
V8	0.265	-0.006	0.807	-0.024	0.140	0.004
V9	0.086	-0.007	0.140	0.957	-0.040	-0.183
V10	0.010	0.134	0.063	-0.031	0.885	-0.097
V11	-0.187	-0.136	0.033	-0.239	-0.109	0.828
V12	0.083	0.035	0.270	0.005	-0.002	-0.016
V13	0.155	-0.167	0.175	-0.036	0.046	-0.084

IV. C. Financial risk regression prediction based on factor analysis

According to the results of factor analysis, 6 male factors were screened. The company's financial risk as the dependent variable, according to the national enterprise system will have the financial risk phenomenon of the enterprise as 0, no financial risk phenomenon of the enterprise as 1. The use of SPSS on the Logistic of the six male factors and corporate violation status regression analysis, SPSS to carry out financial risk regression analysis based on the data of the accounting statements the output of the output results are shown in Table 4 below.

According to the variables in the equation output from the regression model, it can be seen that the model indicators include F1~F6, six variables, and the relationship between these six variables and the probability of corporate default is negatively correlated. The expression of the financial risk assessment model of the company in the perspective of accounting statement data obtained from the above results is as follows:

$$\begin{aligned}\ln\left(\frac{p}{1-p}\right) &= -2.333-4.056F1-2.407F2 \\ &\quad -2.168F3-3.467F4-0.209F5-1.466F6\end{aligned} \tag{11}$$

This indicates that the financing firm solvency, financing firm asset quality, financing firm profitability, financing firm growth potential, industry competitiveness, and supply chain quality factors of the company under the accounting statement data perspective will have a significant impact on the financial risk profile of the company. Through the model, it can be seen that the above six factors have a significant negative correlation with the financial risk status of the company under the perspective of accounting statement data. When the value of the above six factors increases by one unit, the possibility of the company's financial risk will be reduced by 0.209~4.056, i.e., the lower the probability of the company's financial risk problems under the perspective of accounting statement data.

By analyzing the above results, it can be concluded that whether the government, banks and other financial institutions or the company itself in the process of its level of financial risk should be assessed from the company's solvency, asset quality and other aspects of the stability of the various aspects.

Table 4: Analysis of financial risk regression based on report data

		F1	F2	F3	F4	F5	F6	Constants
B		-4.056	-2.407	-2.168	-3.467	-0.209	-1.466	-2.333
S.E.		0.417	0.74	0.801	0.523	0.641	0.825	0.409
Wald		6.632	18.821	0.806	3.991	26.963	0.868	21.671
Freedom		1	1	1	1	1	1	
P		0.000	0.000	0.000	0.000	0.000	0.000	0.000
Exp(B)		0.226	0.158	0.072	0.499	0.204	0.052	0.305
95% C.I. of Exp(B)	Lower limit	0.109	0.064	0.021	0.259	0.087	0.009	0.124
	Upper limit	0.527	0.353	0.162	0.841	0.523	0.191	0.663

V. Conclusion

The article analyzes the financial risk management status of Company J and collects its accounting statement data from 2020 to 2024, from which it mines the key indicators that have an important impact on the company's financial risk. Based on the indicators of J Company's asset composition, current ratio and cash ratio, its financial risk situation in the past five years is analyzed. In addition, the factors that mainly affect the financial risk are extracted from the financial indicators, the financial risk prediction model is constructed, and the financial risk under the company's accounting statement data is analyzed based on regression analysis.

At the end of 2024, the total assets of J Company were 367.41 billion yuan, of which current assets accounted for more than 85%.

From 2020 to 2024, J Company's cash ratio has been lower than the industry average, and its current debt repayment ability is weak.

In this study, six factors, including the debt service ability factor of financing enterprises and the asset quality factor of financing enterprises, are used as the core for the regression prediction of the company's financial risk, and its cumulative variance explained ratio is 85.58%.

The regression analysis found that there is a significant negative correlation between the above six factors and the company's financial risk, i.e., enhancing the values of the six factors, the lower the probability of financial risk problems.

References

[1] Liu, W., Ren, X., & Chen, J. (2022). Problems and Countermeasures in the Analysis of Enterprise Accounting Statements. Scientific Journal of Economics and Management Research Volume, 4(6).

[2] Cohen, J., Krishnamoorthy, G., & Wright, A. (2017). Enterprise risk management and the financial reporting process: The experiences of audit committee members, CFO s, and external auditors. Contemporary Accounting Research, 34(2), 1178-1209.

[3] Trencheva, M. (2021). Preparation of Financial Statements and Financial Risk Analysis. In ICERI2021 Proceedings (pp. 1955-1963). IATED.

[4] Klamut, E. (2018). Accounting and enterprise risk management. International Journal of Contemporary Management, 17(4), 235-254.

[5] Chen, X., & Metawa, N. (2020). Enterprise financial management information system based on cloud computing in big data environment. Journal of Intelligent & Fuzzy Systems, 39(4), 5223-5232.

[6] Liu, J. (2023). The application of enterprise financial management information system in the big data environment. Advances in Economics and Management Research, 6(1), 151-151.

[7] Gao, J. (2022). Analysis of enterprise financial accounting information management from the perspective of big data. International Journal of Science and Research (IJSR), 11(5), 1272-1276.

[8] Liu, H., & Sun, G. (2021). Research on the application of big data technology in the integration of enterprise business and finance. Journal on Big Data, 3(4), 175.

[9] Jin, M., Wang, Y., & Zeng, Y. (2018). Application of data mining technology in financial risk analysis. Wireless Personal Communications, 102, 3699-3713.

[10] Zhang, X. (2021). Application of data mining and machine learning in management accounting information system. Journal of Applied Science and Engineering, 24(5), 813-820.

[11] Li, H. A. I. Y. A. N. (2024). Corporate Accounting Management Risks Integrating Improved Association Rules and Data Mining. WSEAS Transactions on Computer Research, 12, 348-358.

[12] Yao, L. (2019). Financial accounting intelligence management of internet of things enterprises based on data mining algorithm. Journal of Intelligent & Fuzzy Systems, 37(5), 5915-5923.

[13] Wei, L., Li, G., Zhu, X., & Li, J. (2019). Discovering bank risk factors from financial statements based on a new semi-supervised text mining algorithm. Accounting & Finance, 59(3), 1519-1552.

[14] Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. Expert Systems with Applications, 90, 374-393.

[15] Kwak, W., Shi, Y., & Lee, C. F. (2021). Data mining applications in accounting and finance context. In handbook of financial econometrics, mathematics, statistics, and machine learning (pp. 823-857).

[16] Li, B. (2025). Using Data Mining Technology to Improve the Reliability of Accounting Information Cloud Data. Informatica, 49(5).

[17] Mohammed Shantal,Zalinda Othman & Azuraliza Abu Bakar. (2025). Missing data imputation using correlation coefficient and min-max normalization weighting. Intelligent Data Analysis,29(2),372-384.
[18] Yongxin Ma,Jiaojiao Qi,Xusheng Zhang,Kejun Liu,Yimin Liu,Xuehai Yu... & Bendong Chen. (2025). Development and application of an early warning model for predicting early mortality following stent placement in malignant biliary obstruction: A comparative analysis of logistic regression and artificial neural network approaches.. Oncology letters,29(5),237.
[19] K. Narita,M. Goto & K. Takamura. (2005). A Case of Gastric Cancer and Sigmiod Colon Cancer Complicated by Liver Abscess. Nippon Daicho Komonbyo Gakkai Zasshi,58(2),74-79.