International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 271-285

https://doi.org/10.70517/ijhsa46319

Risk assessment of financial data based on logistic regression algorithm

Yufeng Wu[1,*]
[1] Taiyuan Tourism College, Taiyuan, Shanxi, 030009, China
Corresponding authors: (e-mail: 13643412072@163.com).

Abstract Benford's law is a commonly used method to test the quality of financial data, and introducing Benford's law into the logistic regression model for financial risk early warning can increase the number of effective variables representing the quality of financial data and improve the prediction accuracy of the early warning model. This paper applies Benford's law to test the quality of financial data, constructs modified Benford's factor, and combines it with financial variables to establish a Benford-Logistic model for financial risk warning. Taking Chinese A-share listed companies from 2006 to 2023 as samples, the Lasso method is used to screen the explanatory variables and determine the optimal model so as to realize the risk assessment of the financial data, and the validity of the model is verified by taking Company A as the target. The results of the study show that the introduction of the modified Benford quality factor into the logistic regression model can improve the accuracy of the early warning model for the risk of corporate financial data, and the model constructed is effective when applied to Company A. It is in line with the actual situation of Company A. The early warning model is of great significance in the prevention of the financial risk of the enterprise.

Index Terms logistic regression, modified Benford factor, Lasso method, financial data risk assessment

I. Introduction

The wide application of communication technology in the new era has led to the improvement of the production level of various industries, and also brought brand new challenges to the security of enterprises. In the form of the continuous development of Internet communication technology, the disadvantages of its existence are also increasingly presented, network security issues have become a hot topic [1], [2]. Due to the continuous improvement of the market economy system, the pressure of market competition is increasing day by day, the financial protection and risk assessment of enterprises become very important [3]. In today's context, there are many weaknesses in the financial risk assessment system of many small and medium-sized enterprises, and asset losses caused by uncontrollable factors can occur at any time, which seriously affects and restricts the development of enterprise information technology [4]-[6]. The existence of risk is unavoidable in the process of business and financial activities, the assessment and measurement of financial risk is the key content of the whole financial risk management research [7]-[9]. Therefore, how to identify the financial risks of enterprises, measure and assess the size of financial risks and use appropriate methods to reduce and avoid the occurrence of risks has always been an important research content in the field of enterprise financial risk management.

The traditional financial risk management mainly relies on qualitative analysis and empirical judgment, and identifies and controls risks through the establishment of risk management systems, improving internal control processes and other measures [10]. However, with the increasing complexity of the economic environment, the financial risks faced by enterprises have become increasingly diversified and hidden, and traditional methods have become difficult to capture and quantify these risks in a comprehensive and timely manner, and there are certain limitations [11]-[13]. For this reason, it is necessary to introduce advanced technical means to make up for the shortcomings of traditional financial risk management, and to enhance the science and effectiveness of risk management [14], [15].

In this paper, 15 financial indicators are selected from 5 aspects, such as profitability, solvency, growth, operation, cash flow, etc. ST listed companies are taken as samples and normal listed companies are taken as control samples. Construct the Modified Benford Factor based on the Modified Benford Law, add the Modified Benford Factor as a new explanatory variable into the traditional binary logistic regression model, and apply the Lasso method for the screening of explanatory variables. Thus, the binary logistic model based on modified Benford's distribution law is constructed, which realizes the dynamic assessment and instant warning of the risk of corporate financial data. On this basis, the model is applied to Company A as an example, and the financial risk of Company A and its reasons are explored to provide reference for the assessment and early warning of enterprise financial risk.

II. Logistic regression-based early warning modeling of financial data risks

In this chapter, based on the binary logistic model with modified Benford factors, we construct an early warning model of enterprise financial data risk to realize the dynamic assessment and immediate early warning of enterprise financial data risk.

II. A.Benford's Law

Benford's law [16] refers to the fact that in a large natural dataset, the frequency of occurrence of the numbers 1 to 9 in the first place has a certain regularity and a monotonically decreasing trend. According to Benford's law, assuming that the first digit is m, the probability of occurrence of m is:

$$p(m) = log_{10}(1+\frac{1}{m})m = (1,2,\ldots,9) \tag{1}$$

Due to China's unique delisting system, when a company suffers consecutive losses, there will be a “ST” warning or delisting to indicate the risk. If an enterprise is delisted after being *ST and the operation situation is still not improved, listed companies will commit financial fraud in order to prevent delisting. According to the existing theoretical research and empirical analysis, the first digit of a large number of financial data obey Benford's law, which is also known as the “first digit law”. Therefore, if for a certain indicator data, its first digit m in Benford's law of the theoretical value compared with the actual occurrence of the real value, the greater the difference between the two, the higher the possibility that the data have abnormalities or tampered with. This is because the data should follow the Benford's Law distribution in real life, and data that does not conform to Benford's Law may be the result of human intervention. Therefore, the lower the quality of financial data, the more likely the listed company is to engage in financial fraud. χ^2 Goodness-of-fit is a common method to test whether the probability distribution of the first digit obeys Benford's law. N is the sample size, f_m is the actual frequency of occurrence of the first digit m of the indicator data to be tested, $f_{B,m}$ is the theoretical frequency of m occurrence, and χ^2 the goodness-of-fit test statistic is:

$$x^2 = N\Sigma_{m=1}^{9}\left[(f_m - f_{B,m})^2 / f_{B,m}\right] \tag{2}$$

II. B.Benford factor construction method

Let $X_j\left\{j=1,2,3,\ldots,k\right\}$ denote the financial indicator variable. Use $r_m^{(j)}$ to denote the difference between the observed frequency of the first and last digits m of X_j and the theoretical frequency of Benford's law, and the expression for $r_m^{(j)}$ is:

$$r_m^{(j)} = f_m^{(j)} - f_{B,m}^{(j)} \quad j=(1,2,3,...,k) \tag{3}$$

where $f_m^{(j)}, f_{B,m}^{(j)}$ denotes the observed frequency of the first digit m of the financial indicator X_j and the theoretical frequency of Benford's law, respectively, and satisfies $\Sigma_m = 1^9 f_m^{(j)} = 1, \Sigma_{m=1}^{9} f_{B,m}^{(j)} = 1$, then we have:

$$\Sigma_{m=1}^{9} r_m^{(j)} = 0 \quad j=(1,2,3,...,k) \tag{4}$$

When the data are modified, the observed frequency of the first digit of Indicator X_j will differ from the theoretical frequency. The greater the absolute value of the difference $r_m^{(j)}$, the higher the risk of financial fraud at the sample point. In this paper, we focus on the first digit with the largest absolute value of the difference between the observed and theoretical frequencies. The first digit with the largest absolute value of the first digit frequency difference is $n^{(j)}$. Human intervention in the financial data results in a difference between the observed and theoretical frequencies of the first digit of each indicator X_j, which can be expressed as an absolute difference of $r_m^{(j)}$. The larger the absolute difference, the higher the risk of financial fraud in the sample. In this paper, the focus is on identifying the first digit with the largest absolute difference between the observed and theoretical frequencies, which is denoted by $n^{(j)}$, i.e.:

$$n^{(j)} = argmaxd\left|r_m^{(j)}\right| j=1,2,...,k \tag{5}$$

Construct the Benford factor for Indicator $X_j\{j=1,2,3,\ldots,k\}$, denoted B_j. B_j takes the value 1 if the first digit of Indicator $X_{i,j}$ at Sample Point i is $n^{(j)}$. otherwise, it takes the value 0. The following expression is available:

$$B_{i,j}=\begin{cases}1 & \text{The first digit of } X_{i,j} \text{ is } n^{(j)}\\ 0 & \text{other}\end{cases} \tag{6}$$

II. C.Logistic regression analysis

Logistic regression analysis [17] is a relatively effective method in solving the problem of 0~1 regression, and does not require that the sample must meet the normal distribution, which is more applicable to the prediction of corporate financial risk. Its basic regression model is:

$$P=\frac{e^{\beta 0}+(e^{\beta 1}+e^{\beta 2}+\cdots+\beta^{n})x^{n}}{1+e^{\beta 0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n} \tag{7}$$

The Logistic function is essentially a combination of a linear regression model with a Sigmoid function. The Sigmoid function is:

$$g(x)=\frac{1}{1+e^{-mx}} \tag{8}$$

The derivative of the Sigmoid function is:

$$g(x)=\frac{me^{-mx}}{1+e^{-mx}} \tag{9}$$

Bringing this into a linear regression function as shown in equation (10) yields the Logistic function as shown in equation (11):

$$h_\theta(x)=g(\theta^T x) \tag{10}$$

$$h_\theta(x^{(i)})=g(\theta^T x^{(i)})=\frac{1}{1+e^{-\theta^T(x^i)}} \tag{11}$$

Among them:

$$\theta^T x^{(i)}=\sum_{j=0}^{n}\theta_i x_j{}^{(i)} \tag{12}$$

In summary, the binary logistic regression model is:

$$P(y^{(i)}=1\,|\,x^{(i)};\theta)=h_\theta(x^{(i)}) \tag{13}$$

$$P(y^{(i)}=0\,|\,x^{(i)};\theta)=1-h_\theta(x^{(i)}) \tag{14}$$

The probability value of the likelihood ratio statistic obtained by the maximum likelihood method was used to eliminate the variables and the formula was calculated as:

$$L(\theta)=\prod_{i=1}^{n}P(y^{(i)}=1\,|\,x^{(i)};\theta)=\prod_{i=1}^{n}h_\theta(x^{(i)})^{y(i)}[1-h_\theta(x^{(i)})^{1-y(i)}] \tag{15}$$

II. D.Binary Logistic Model with Modified Benford Factors

The binary logistic model with modified Benford factors is an improvement of the traditional binary logistic model. The process of constructing the binary logistic model with modified Benford factors mainly consists of two parts, one is the setting of modified Benford factors and the other is the estimation of the binary logistic model with modified Benford factors.

II. D. 1) Modifying the Benford factor construction

For the problem of correctly classifying normal and abnormal listed companies, listed companies are categorized according to their financial indicators. While none of the listed companies' financial indicator data sets are unbounded data sets, the quality of listed companies' financial indicator data is to be assessed using the Modified Benford Law. If the financial indicator data are artificially tampered and manipulated, it will deviate from the Modified Benford Law. The test value of Modified Benford's Law is the basis for judging the abnormality of financial indicators.

Denote x_i as the financial indicator and d_j as the difference in absolute value between the observed frequency f_{x_i} of the first digit j of the dataset of financial indicator x_i and the theoretical frequency f_{AB,x_i} of the modified Benford law. Assume that the modified Benford test value for the data set of financial indicator x_i is higher than the critical value. The absolute value of the difference between the observed frequency f_{x_i} of the first digit and the theoretical frequency f_{AB,x_i} of the modified Benford law is:

$$d_j = \left| f_{x_i} - f_{AB,x_i} \right| (j = 1,2,3 \cdots 9) \tag{16}$$

The maximum value of d_j is denoted as $d_{\max}$, and $d_{\max}$ corresponds to the first digit of financial indicator x_i, denoted as a_i.

The modified Benford factor B_i is set according to the first digit of the dataset for financial indicator x_i. If the first digit of the dataset for financial indicator x_i is a_i, variable B_i takes the value of 1. Otherwise, variable B_i takes the value of 0. Note:

$$B_i = \begin{cases} 1 & D_p = a_i \\ 0 & \text{other} \end{cases} \tag{17}$$

where D_p is the first digit of the i rd observation of FSI x_i.

II. D. 2) Benford-Logistic Modeling

Assume that financial indicator x_i fails the Modified Benford test. Set the Modified Benford Factor B_i based on the first digit of the dataset for Financial Indicator x_i. Add Variable B_i as an explanatory variable to the dichotomous logistic model to get the dichotomous logistic model with Modified Benford Factor as:

$$\ln\left(\frac{P(Y=1 \mid X,B)}{1-P(Y=1 \mid X,B)}\right) = \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \sum_{j=1}^{q} \gamma_j B_j \tag{18}$$

The dependent variable Y is the variable of financial abnormality of listed companies, which is recorded as 1 for the observation of the dependent variable Y for financially abnormal listed companies and 0 for the observation of the dependent variable Y for financially normal listed companies. $X_p X_2 ... X_p$ are the P explanatory variables, which represent the financial indicators of listed companies. Among them, there are q explanatory variables whose first digit does not obey the Modified Benford Law. The Modified Benford factor is sequentially denoted $B_1, B_2, ..., B_q, \beta_0, \beta_1, ..., \beta_p, \gamma_1, \gamma_2, ..., \gamma_q$ as the unknown parameters of the model. The probability of financial anomalies of listed companies can be obtained from model (18) as:

$$P(Y=1 \mid X,B) = \frac{\exp\left(\beta_0 + \sum_{i=1}^{p} \beta_1 x_1 + \sum_{j=1}^{q} \beta_j B_j\right)}{1+\exp\left(\beta_0 + \sum_{i=1}^{p} \beta_1 x_1 + \sum_{j=1}^{q} \beta_j B_j\right)} \tag{19}$$

The greater the probability $P(Y=1 \mid X,B)$, the greater the risk of financial anomalies for the firm.

The steps to categorize listed companies using the binary logistic model with modified Benford factors are as follows:

Step1: Select the financial indicators to be tested of the classified listed companies, and carry out Benford test on these financial indicators in turn. According to the Benford test value, identify the financial indicators with low data quality.

Step2: Based on the low data quality indicators screened out in Step1, construct the modified Benford factor, calculate the observed frequency of the first digit of each observation of the financial indicators and the theoretical frequency of Benford's law, calculate the absolute difference between the observed frequency of the first digit of each observation of the financial indicators and the theoretical frequency of Benford's law, screen out the observation with the largest absolute difference and record the observation with the largest absolute difference. Record the first digit corresponding to the observation with the largest absolute difference. Determine whether the first digit of the sample point of the financial indicator is the same as the first digit of the sample point. If the first digit of the sample point is the same as the number, the Benford factor of the sample point is assigned to 1. Otherwise, the Benford factor of the sample point is assigned to 0.

Step3: Construct a new binary classification logistic model based on modified Benford's law, add the modified Benford factor constructed in Step2 as a new explanatory variable to the traditional binary classification logistic model, and construct the binary classification logistic model with the addition of modified Benford's factor to classify the listed companies.

The binary logistic model with modified Benford factor proposed in this paper integrates the advantages of Benford's law that can effectively assess the data quality and the high classification accuracy of the logistic model, overcomes the defects of using Benford's law alone that can not realize the correct classification of listed companies, and improves the classification correctness of the logistic model.

III. Early warning of corporate financial risk based on Benford-Logistic modeling

This chapter establishes Benford-Logistic financial risk early warning model based on the selection of variables and data. And take Company A as an example to apply the model.

III. A. Description of variables and data sources

III. A. 1) System of financial early warning indicators

In this paper, we utilize Benford's law to perform data quality tests and construct Benford factors for the independent variables in data set D. In particular, data set $D=\{(X_1,Y_1),(X_2,Y_2),...,(X_2,Y_2)\}$ contains N samples, each with k independent variables, i.e., $X_i=(X_{i,1},X_{i,2},\ldots,X_{i,k})\{i=1,2,3,\ldots,N\}$, and a category variable Y_i. The Benford factor is constructed as a new independent variable by calculating the first digit of the largest difference between the observed and theoretical frequencies for each independent variable X_i. For each independent variable X_i, the first digit with the largest difference is labeled as B_j, and the constructed Benford factor is added to the model as a new independent variable, noting the new variable as $X_i^B=(X_{i,1},X_{i,2},\ldots,X_{i,k},B_{i,1},B_{i,2},\ldots,B_{i,k})\{i=1,2,3,\ldots,N\}$, where $B_{i,j}$ denotes the Benford factor of the j th independent variable, resulting in the final data set $D^B=\{(X_1^B,Y_1),(X_2^B,Y_2),...,(X_N^B,Y_N)\}$.

With reference to the existing research literature, this paper selects 15 financial indicators from profitability, solvency, growth, operating capacity, cash flow, etc. to establish the enterprise financial risk early warning indicator system, and the financial early warning indicator system containing Benford factors is shown in Table 1. Among them, X1-X15 indicates the financial indicators, and B1-B15 indicates the Benford factors corresponding to the indicators.

The financial indicators of the enterprise mainly assess the financial status of the enterprise from different perspectives. Among them, profitability reflects the enterprise's power and source of sustainable reproduction in the future, and the main indicators include return on net assets (X1), return on assets (X2) and operating profit margin (X3). Solvency reflects the enterprise's ability to repay debts, which is an important guarantee for the stable operation and development of the enterprise, and the main indicators include equity ratio (X4), quick ratio (X5) and interest coverage multiple (X6). Growth capacity is to reflect the potential ability of the enterprise to further expand and develop under the current situation, the main indicators include the growth rate of total assets (X7) and operating profit growth rate (X8). Operating capacity reflects the efficiency and effectiveness of the enterprise's operating assets, and the main indicators include inventory turnover ratio (X9), accounts receivable turnover ratio (X10), fixed asset turnover ratio (X11) and total asset turnover ratio (X12). Cash flow is the basis for enterprises to maintain daily operation and production activities, including the cash inflow, cash outflow and their total amount generated by enterprises through certain economic activities in a certain accounting period, which is the key factor for the sustainable operation and development of enterprises. These indicators are interrelated and collectively reflect different aspects of an enterprise's financial position. Key indicators include cash flow from operating activities per share (X13), cash current liabilities ratio (X14) and cash flow from operating activities per share growth rate (X15).

III. A. 2) Sample selection and data sources

In this paper, listed companies are treated with special treatment (ST) as an indication of the emergence of financial risk. Listed companies that are ST in China's A-share market from 2006 to 2023 are selected, and then for each ST listed company, a listed company in normal operation is selected to be paired with it according to the principle of the same industry in the same year and the most similar asset size. The financial risk of the current year is predicted using the listed company's financial index data of the previous year. After deleting the listed companies in the financial industry and those with missing relevant data, 215 ST companies and normal companies were obtained. Whether a listed company is ST or not is the dependent variable, ST companies are assigned a value of 1 and normal companies are assigned a value of 0. The data are obtained from RESET Financial Database. The results of descriptive statistical analysis of financial indicators are shown in Table 2.

Table 1: Financial early warning indicator system containing Benford factor

Indicator type	Symbol	Indicator name
Profitability	X1	Return on net assets
	B1	
	X2	Return on assets
	B2	
	X3	Operating profit margin
	B3	
Debt-paying ability	X4	Equity ratio
	B4	
	X5	Quick ratio
	B5	
	X6	Interest coverage ratio
	B6	
Growth ability	X7	Growth rate of total assets
	B7	
	X8	Growth rate of operating profit
	B8	
Operational capacity	X9	Inventory turnover rate
	B9	
	X10	Accounts receivable turnover rate
	B10	
	X11	Fixed asset turnover rate
	B11	
	X12	Total asset turnover rate
	B12	
Cash flow	X13	Cash flow per share
	B13	
	X14	Cash flow liability ratio
	B14	
	X15	Cash flow per share growth rate
	B15	

The non-parametric Wilcox test is used to statistically test whether there is a significant difference in the distribution of financial indicators between the sample of ST and normal companies, and the results of the test are shown in the last column of Table 2. Among the 15 financial indicators selected in this paper, except for the non-parametric tests of operating profit growth rate (X8), inventory turnover (X9), accounts receivable turnover (X10) and cash flow per share growth rate (X15) which are not significant, the remaining 11 financial indicators can significantly differentiate between ST companies and normal companies. Based on the mean values of these significantly different financial indicators, it can be seen that ST companies are significantly lower than normal companies.

III. A. 3) Construction and assignment of Benford factors

Use Benford's law to evaluate the data quality of financial indicators of listed companies. Table 3 shows the observation frequency of the first digit of each financial indicator, the theoretical frequency under Benford's law, the difference between the two and the χ^2 goodness-of-fit test, the second column of the table (1) represents the theoretical frequency of the first digit under Benford's law, (2) represents the observation frequency of the first digit of the financial indicator, and (3) represents the difference between the observation frequency of the first digit of the financial indicator and the theoretical frequency under Benford's law, in the table: %. The χ^2-test cut-off value of 10% of the significance level is 20.09, while the χ^2-test value of equity ratio (X4), quick ratio (X5), inventory turnover (X9), accounts receivable turnover (X10), total asset turnover (X12), and cash flow growth rate per share (X15) is greater than 20.09, so it can be considered that there is a significant difference between the first-digit distribution of these indicators and Benford's law.

Table 2: Descriptive statistical analysis of financial indicators

Financial indicators	ST Company		Normal company		Wilcox tests the significance level
	Mean value	Standard deviation	Mean value	Standard deviation	
X1	-0.112	0.606	-0.005	0.459	0.001
X2	-0.029	0.138	0.046	0.113	0.000
X3	-0.315	0.594	-0.038	0.476	0.000
X4	0.978	0.761	1.604	1.145	0.000
X5	0.741	0.649	1.175	0.968	0.000
X6	0.667	14.925	13.863	26.102	0.000
X7	-0.016	0.304	0.172	0.274	0.000
X8	-0.483	2.796	-0.096	3.095	0.738
X9	5.884	7.042	4.862	5.508	0.314
X10	12.426	17.654	10.945	16.136	0.573
X11	4.279	10.381	7.104	12.457	0.000
X12	0.482	0.419	0.768	0.495	0.000
X13	0.115	0.507	0.223	0.694	0.017
X14	0.077	0.285	0.164	0.319	0.004
X15	1.281	5.672	0.808	4.245	0.382

The corresponding Benford factors were constructed based on the six financial indicators that were significant in the Benford law goodness-of-fit test. Financial indicators X4, X5, X9, X10, X12, X15 first digit frequency is higher than the theoretical frequency of the largest first digit are 1, 8, 3, 7, 8, 1, respectively, the corresponding difference of 11.01%, 3.59%, 9.02%, 4.50%, 3.70%, 8.15%, respectively, in accordance with the modified Benford factor constructing method of equation (17), respectively, constructed Benford factors B4, B5, B9, B10, B12, B15 and assigned values. Financial indicators X4, X5, X9, X10, X12, X15 first digit frequency is lower than the theoretical frequency of the largest first digit are 2, 2, 1, 1, 2, 2, the corresponding difference of -5.55%, -7.98%, -4.67%, -6.90%, -6.24%, -3.64%, respectively, and also in accordance with the formula (17), respectively, to construct the Benford factor B4′, B5′, B9′, B10′, B12′, B15′ and assigned values. A total of 12 Benford factors were obtained.

III. A. 4) Screening of explanatory variables

At the early stage of the establishment of corporate financial risk early warning model, more explanatory variables are generally introduced to avoid omitting important variables, but the Logistic model containing too many variables tends to cause some variables to be statistically insignificant due to multicollinearity, and the stability of the model deteriorates, which greatly reduces the estimation and prediction accuracy, so it is necessary to make variable selection of the explanatory variables. The stepwise regression method tends to retain some unimportant variables when the independent variables are correlated, while the Lasso method can more accurately screen out important variables, and this paper adopts the Lasso method to screen 15 financial indicator variables [18]. The model error (vertical axis) and the number of variables screened (numbers at the top of the graph) for different reconciliation parameters λ (horizontal axis) are shown in Figure 1. λ The value taken affects the degree of compression of the model variables; the value of λ corresponding to the left dashed line in Fig. 1 minimizes the model error, and the value of λ corresponding to the right dashed line makes the model relatively concise. Here, the $\log(Lambda)$ value of the right dotted line, i.e. $\lambda = e^{-3}$, is taken, and five explanatory variables are retained based on the Lasso method, namely equity ratio (X4), interest coverage multiple (X6), total asset growth rate (X7), inventory turnover ratio (X9) and total asset turnover ratio (X12), which indicates that these variables play a more important role in influencing the financial risk of the firm. Using this set of variables, a Benford-Logistic financial risk early warning model is developed.

Table 3: Comparison of observed frequency and theoretical frequency

Index	Frequency	1	2	3	4	5	6	7	8	9	χ^2
Benford	(1)	29.98	17.58	12.56	9.72	8.03	6.81	5.92	5.14	4.62	-
X1	(2)	30.15	24.59	12.05	7.73	5.52	5.82	4.17	4.91	6.03	19.35
	(3)	0.17	7.01	-0.51	-1.99	-2.51	-0.99	-1.75	-0.23	1.41	
X2	(2)	30.21	17.24	10.93	6.82	8.69	8.17	3.78	6.55	7.94	19.28
	(3)	0.23	-0.34	-1.63	-2.90	0.66	1.36	-2.14	1.41	3.32	
X3	(2)	31.45	20.08	11.74	10.56	8.53	5.19	5.19	3.76	3.58	6.74
	(3)	1.47	2.50	-0.82	0.84	0.50	-1.62	-0.73	-1.38	-1.04	
X4	(2)	40.99	12.03	7.92	5.77	8.17	7.18	5.24	5.24	7.93	41.87
	(3)	11.01	-5.55	-4.64	-3.95	0.14	0.37	-0.68	0.10	3.31	
X5	(2)	29.73	9.60	9.06	11.71	12.05	8.16	5.45	8.73	6.03	37.69
	(3)	-0.25	-7.98	-3.50	1.99	4.02	1.35	-0.47	3.59	1.41	
X6	(2)	37.64	14.74	10.93	8.17	10.58	5.52	4.95	4.17	3.79	17.24
	(3)	7.66	-2.84	-1.63	-1.55	2.55	-1.29	-0.97	-0.97	-0.83	
X7	(2)	30.52	19.91	13.27	7.59	6.55	7.56	4.38	3.35	7.15	14.16
	(3)	0.54	2.33	0.71	-2.13	-1.48	0.75	-1.54	-1.79	2.53	
X8	(2)	34.56	15.53	11.25	6.31	7.16	6.83	7.42	4.74	6.82	15.45
	(3)	4.58	-2.05	-1.31	-3.41	-0.87	0.02	1.50	-0.40	2.20	
X9	(2)	25.31	14.35	21.58	10.94	8.18	7.16	4.95	4.11	3.78	30.24
	(3)	-4.67	-3.23	9.02	1.22	0.15	0.35	-0.97	-1.03	-0.84	
X10	(2)	23.08	16.29	14.51	11.76	7.42	6.54	10.42	5.76	4.73	21.49
	(3)	-6.90	-1.29	1.95	2.04	-0.61	-0.27	4.50	0.62	0.11	
X11	(2)	30.68	18.18	10.14	10.14	7.93	7.93	4.71	5.47	5.18	4.64
	(3)	0.70	0.60	-2.42	0.42	-0.10	1.12	-1.21	0.33	0.56	
X12	(2)	28.73	11.34	11.25	8.39	10.28	8.45	6.83	8.84	6.02	26.08
	(3)	-1.25	-6.24	-1.31	-1.33	2.25	1.64	0.91	3.70	1.40	
X13	(2)	26.72	17.21	15.25	9.83	7.42	6.74	7.12	4.41	5.82	6.46
	(3)	-3.26	-0.37	2.69	0.11	-0.61	-0.07	1.20	-0.73	1.20	
X14	(2)	29.12	16.93	13.63	10.14	8.45	5.69	5.51	6.03	4.91	2.12
	(3)	-0.86	-0.65	1.07	0.42	0.42	-1.12	-0.41	0.89	0.29	
X15	(2)	38.13	13.94	9.82	10.42	5.19	5.69	4.70	5.56	7.14	23.65
	(3)	8.15	-3.64	-2.74	0.70	-2.84	-1.12	-1.22	0.42	2.52	

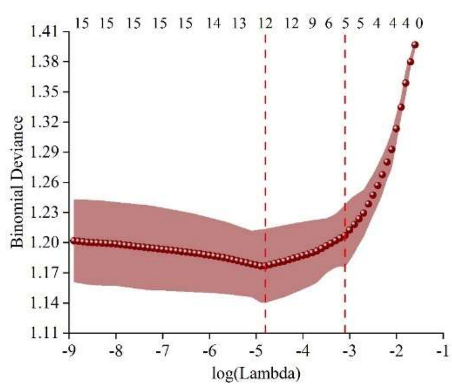

Figure 1: The trend of Lambda corresponds to the number of variables

III. B. Application of Financial Risk Early Warning Models - Company A as an Example

III. B. 1) Analysis of the current financial situation of Company A

Company A belongs to the industry of real estate development, and its main development products include residential houses, office buildings, stores, parking spaces and hotels, and involves in the business of primary development of land, property management and leasing of commercial properties. Therefore, this paper describes the financial status quo of Company A from the accounting elements of assets, liabilities, owner's equity, income and profit, in order to have a clearer perception of the company's financial status and operation, and to provide some reference for analyzing its financial risk performance and causes.

1) Basic situation of assets

The asset composition of Company A in 2022-2023 is shown in Table 4. It can be seen that the total assets of Company A in 2023 is 8145786.31 million yuan, down 8.84%. Among them, the current assets have 5074369.84 million yuan, accounting for 62.29% of the total assets, down 10.94% compared with 2022. And the non-current assets are 3,071,414,164,700 yuan, a year-on-year decrease of 5.14%, accounting for only 37.71% of the total assets.

Table 4: Asset composition of Company A from 2022 to 2023

Project Name	In 2023 (ten thousand yuan)	Proportion of assets /%	In 2022 (ten thousand yuan)	Proportion of assets /%
Assets	8145786.31	100.00	8935172.65	100.00
Current assets	5074369.84	62.29	5697481.59	63.77
Non-current assets	3071416.47	37.71	3237691.06	36.23

To understand the asset portfolio and asset liquidity of Company A, the composition of its current assets is further analyzed.The composition of Company A's current assets in 2022-2023 is shown in Table 5.In 2023, Company A's current assets consist of money funds, accounts receivable, prepayments, other receivables, inventories, contract assets, assets held for sale, and other current assets, with the highest percentage of inventories, as high as 91.12%. The proportion of more liquid assets such as money funds is not high and decreases year-on-year. in 2023, Company A's money funds only amounted to 1,315,245,100 yuan, accounting for 2.59% of current assets, indicating that the enterprise's asset structure is irrational, with a reduced solvency capacity and an increased risk of debt repayment.

Table 5: The composition of current assets of Company A from 2022 to 2023

Project Name	In 2023 (ten thousand yuan)	Proportion of assets /%	In 2022 (ten thousand yuan)	Proportion of assets /%
Monetary funds	131524.51	2.59	203142.51	3.56
Accounts receivable	13476.76	0.27	28357.83	0.50
Prepaid amount	24018.65	0.47	17732.65	0.31
Other receivables	249215.48	4.91	282631.47	4.96
Inventory	4623519.62	91.12	5151584.29	90.42
Contract assets	12058.85	0.24	-	-
Hold assets for sale	7314.58	0.14	-	-
Other current assets	13241.39	0.26	14032.84	0.25

2) Liabilities and owner's equity

The liabilities and owner's equity of Company A in 2022-2023 are shown in Table 6.The owner's equity of Company A in 2023 is 195,342,830,000 Yuan, which is 65.57% lower than that of 2022. The total liabilities were 7,950,443,480,000 Yuan, and the gearing ratio was 97.60%, which was 3.95% higher than the same period of the previous year, which indicated that the solvency of Company A decreased.In 2023, the current liabilities of Company A were 6,721,624,200 Yuan, and the amount of the non-current liabilities was 1,228,819,600 Yuan.

Table 6: Liabilities and owner's equity of Company A from 2022 to 2023

Project Name	In 2023 (ten thousand yuan)	Proportion of assets /%	In 2022 (ten thousand yuan)	Proportion of assets /%
Total liabilities and equity	8145786.31	100.00	8935172.65	100.00
Owner's equity	195342.83	2.40	567423.71	6.35
Liabilities	7950443.48	97.60	8367748.94	93.65
Current liabilities	6721624.32	82.52	5843183.26	65.40
Non-current liabilities	1228819.16	15.08	2524565.68	28.25

3) Costs, revenues, expenses and profits

The cost, income and profit of Company A in 2022-2023 are shown in Table 7.The operating profit of Company A in 2023 and 2022 are -2516,435,800 and -314,256,800 respectively.The net profit in 2023 is -310,465,300 and in

2022 is -359,136,200, which shows that the company's poor operating condition is is in loss, and measures should be taken as soon as possible to improve the company's operating condition. From the change of operating income, the operating income in 2023 is 4404,368,700 Yuan, decreasing by 29.63% year-on-year, and the decrease reaches -1854,795,600 Yuan, indicating that the sales of Company A are greatly reduced. Operating cost was 2316,585,900 Yuan, a decrease of 42.27% year-on-year.The financial expense of Company A in 2023 was 301,624,500 Yuan, an increase of 28.26% year-on-year.

Table 7: The costs, revenues and profits of Company A from 2022 to 2023

Project Name	In 2023 (ten thousand yuan)	Growth rate /%	In 2022 (ten thousand yuan)
Total revenue	440436.87	-29.63	625916.43
Business assembly	685273.64	-17.49	830495.64
Operating cost	231658.59	-42.27	401312.46
Sales cost	31722.43	-30.91	45913.51
Management fee	61754.82	-20.15	77342.60
Financial cost	301624.35	28.26	235174.95
Investment income	110372.53	-6113.51	-1835.41
Operating profit	-251643.58	-19.92	-314256.48
External income	5264.71	10.55	4762.19
Outside expenditure	21439.48	-9.25	23625.84
Net profit	-310465.73	-13.55	-359136.52

The details of the operating income of Company A in 2023 are shown in Table 8.The operating income of Company A is derived from the fields of real estate development, property management, hotel operation and commercial management, etc., of which the operating income obtained through real estate development accounted for the largest share of the total operating income, which was 63.78%, a year-on-year decrease of 3.91%. The share of operating income obtained through hotel operation and commercial management in the total operating income was 7.33% and 20.13% respectively, representing an increase of 1.65% and 13.32% year-on-year, whereas the share of property management income in the total operating income decreased by 11.06%, indicating that commercial management had become the second main business of Company A. Among the four major businesses, only the operating income obtained through property management is increasing, but its increase is not significant and it does not account for a high proportion of the total operating income, therefore, Company A's total operating income in 2023 decreased by 29.63% year-on-year.

Table 8: The operating income situation of Company A in 2023

Project Name	Operating income of the current period (ten thousand yuan)	Proportion of total income /%	Operating income of the last period (ten thousand yuan)	Proportion of total income /%	Year-on-year growth /%
Real estate development	280915.32	63.78	423674.03	67.69	-3.91
Property management	38603.15	8.76	124092.79	19.82	-11.06
Hotel operation	32272.86	7.33	35540.85	5.68	1.65
Business management	88645.54	20.13	42608.76	6.81	13.32

III. B. 2) Early warning results and analysis of financial risk in Company A

(1) Company A's financial risk early warning results

The Benford-Logistic corporate financial risk early warning model constructed in the previous paper is applied to Company A, and the data of Company A from 2018-2023 are used to predict the financial risk situation of Company A from 2020-2025.The results of the financial risk prediction of Company A from 2020-2025 are shown in Table 9.

It can be seen that according to the early warning model, Company A operates normally in 2020-2023, and a financial crisis will occur in 2024, which is consistent with the actual situation of the company, indicating that the financial risk early warning model constructed in this paper is effective and applicable to Company A.

Table 9: Financial risk forecast of Company A from 2020 to 2025

Data year	Predicted year	P	Prediction result	Actual result
2018	2020	0.3351	Operating normal	Operating normal
2019	2021	0.2407	Operating normal	Operating normal
2020	2022	0.4435	Operating normal	Operating normal
2021	2023	0.3627	Operating normal	Operating normal
2022	2024	0.6924	Financial crisis	Financial crisis
2023	2025	0.7241	Financial crisis	

(2) Analysis of Company A's financial risk early warning results

Combined with the financial risk early warning model, based on 15 relevant financial indicators, analyze the performance and reasons of Company A's financial crisis from five aspects, including solvency, profitability, operating ability, growth ability and cash flow.

(1) Solvency analysis

According to the financial risk early warning model, the solvency of Company A is analyzed mainly from the equity ratio, quick ratio and cash flow interest coverage multiple. During the period of 2018-2023, the indicators of Company A's solvency capacity are shown in Table 10, and the results of Company A's equity ratio compared with the industry are shown in Figure 2.

Combined with Table 10 and Figure 2, it can be seen that the equity ratio of Company A shows an increasing trend from 2018 to 2023 and has been higher than the average level of the real estate industry, and the increase is even more obvious in 2022 and 2023. It shows that Company A's source of funds is mainly creditors rather than investors, and its capital structure is unstable and weak in long-term debt service capacity. The fluctuation of the long-term debt to working capital ratio is relatively large, and the quick ratio keeps decreasing from 2018 to 2023, from 0.9217 in 2018 to 0.2219 in 2023, indicating that the solvency of Company A is weakened, which is due to the fact that the current liabilities of Company A are too high in 2021 and 2023, far exceeding the quick assets. During 2018-2023, there are four years in which Company A's cash flow interest coverage multiple is negative, indicating that Company A's operating activities have been in a state where cash inflow is less than cash outflow for a long period of time, and it is necessary to pay attention to the debt servicing risk brought about by cash shortage.

Table 10: The debt-paying ability indicators of Company A from 2018 to 2023

Data year	Equity ratio	Quick ratio	Cash flow interest coverage ratio
2018	7.1816	0.9217	-3.4023
2019	8.3642	0.8190	-2.3164
2020	8.0521	0.5405	2.4152
2021	8.5274	0.4003	-1.3724
2022	15.0183	0.3105	-1.3126
2023	43.0562	0.2219	0.1953

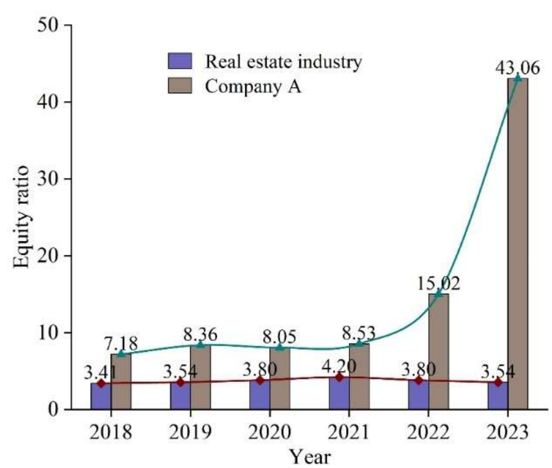

Figure 2: Comparison of the equity ratio of Company A with the industry from 2018 to 2023

(2) Profitability analysis

The profitability of Company A is analyzed in terms of return on net assets, return on assets and operating profit margin.The profitability indicators of Company A from 2018-2023 are shown in Table 11. Among them, the results of the comparison of Company A's return on net assets with the industry are shown in Figure 3.

Combined with Table 11 and Figure 3, it can be seen that the trends of Company A's return on net assets, return on assets and operating profit margin are basically the same from 2018-2023, with an upward trend and then a downward trend, and these indicators are less than 0 in 2022 and 2023.Company A's return on net assets is lower than the industry average, and the gap with the industry widens from 2022 onwards, which indicates that the Company's profitability is lower than the real estate industry level and is declining. Changes in profitability indicators are mainly due to declining revenues and profits.

Table 11: The profitability indicators of Company A from 2018 to 2023

Data year	Return on net assets	Return on assets	Operating profit margin
2018	0.0512	0.0065	0.0638
2019	0.0546	0.0058	0.0369
2020	0.0649	0.0071	0.0382
2021	0.0621	0.0068	0.0565
2022	-0.4886	-0.0412	-0.5681
2023	-0.8452	-0.0375	-0.7153

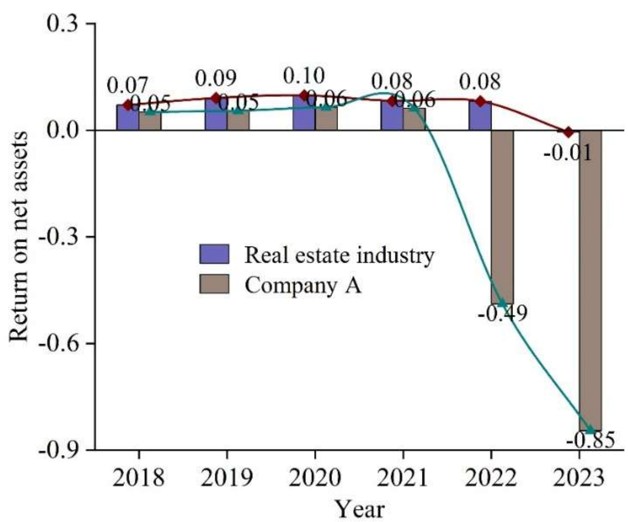

Figure 3: Comparison of return on net assets of Company A with the industry

(3) Analysis of operating capacity

This paper focuses on the accounts receivable turnover ratio to analyze the operating ability of Company A. The results of Company A's accounts receivable turnover ratio from 2018-2023 compared with the industry are shown in Figure 4. It can be seen that Company A's accounts receivable turnover ratio is increasing from 2.05 to 21.93 in recent years, and the rate of increase is accelerating. Combined with the declining trend of Company A's operating income from 2018 to 2023, it is not difficult to find that the average occupancy of Company A's accounts receivable decreases sharply in 2023, which indicates that Company A's operating ability has been improved to some extent. In addition, Company A's accounts receivable turnover ratio is much lower than the average level of the real estate industry, indicating that Company A's accounts receivable turnover ratio needs to be improved and its operating capacity needs to be enhanced.

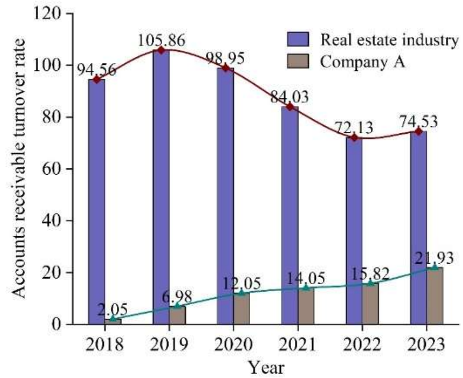

Figure 4: Comparison of accounts receivable turnover rate of Company A with the industry

(4) Analysis of growth capacity

Company A's growth indicators for 2018-2023 are shown in Table 12. It can be seen that the total asset growth rate and operating profit growth rate of Company A in 2018-2023 are fluctuating, but the overall trend is declining, and from the perspective of the combined two indicators, the financial performance in 2022 and 2023 is the worst, and the growth capacity of Company A is declining.

Table 12: The growth ability indicators of Company A from 2018 to 2023

Data year	Growth rate of total assets	Growth rate of operating profit
2018	0.3586	-1.1254
2019	0.4275	0.3768
2020	0.2394	0.7341
2021	0.0861	0.1539
2022	0.0582	-5.6251
2023	-0.1052	0.2854

(5) Cash flow analysis

This paper focuses on cash flow per share from operating activities to analyze the cash flow situation of Company A. Company A's cash flow per share from operating activities compared with the industry is shown in Figure 5.Cash flow per share from operating activities of Company A shows a decreasing trend from 2018 to 2023.Specifically, Company A's cash flow per share from operating activities in 2019 and 2020 is positive, and the rest of the years are negative. In 2021, Company A's negative cash flow from operating activities per share indicates that the company's cash inflow from operating activities is less than its outflow. In 2022 and 2023, the negative cash flow from operating activities per share is due to the negative net income, specifically due to the downturn in the real estate sales market, the difficulty in financing which leads to the restriction of investment in Lee's real estate projects, and a significant decline in the real estate segment's sales revenue. Compared with the real estate industry, Company A's cash flow from operating activities per share is highly volatile and lower than the average level of the real estate industry in 2021-2023 and 2019, which should be taken seriously by Company A's management.

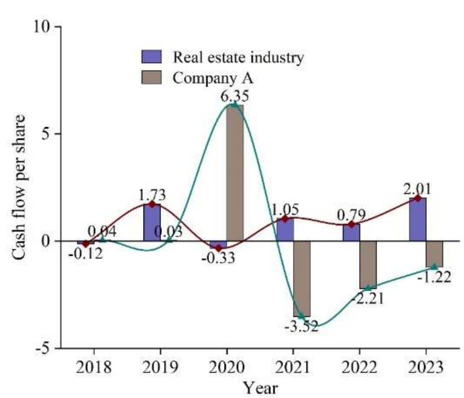

Figure 5: Comparison of cash flow per share of Company A with the industry

IV. Conclusion

In this paper, the modified Benford factors are added into the traditional logistic regression model to construct a Benford-Logistic based corporate financial risk early warning model, which realizes the assessment and early warning of financial risk.

The study selects 15 financial indicators from profitability, solvency, growth, operating ability, cash flow, etc. to establish the enterprise financial risk early warning indicator system, and takes 430 Chinese A-share listed companies from 2006 to 2023 as the research samples, and adopts the Lasso method to screen the explanatory variables, and selects the Benford-Logistic model with the optimal fitting effect and prediction effect. Applying the model to Company A, the model prediction results show that Company A operates normally from 2020 to 2023, and there is a greater chance of financial crisis in 2024, which is in line with the company's actual situation and proves the effectiveness of the financial risk early warning model built in this paper.

In terms of solvency, the equity ratio of Company A shows an increasing trend from 2018 to 2023 and has been higher than the industry average, indicating that the capital structure of Company A is unstable and the long-term solvency is weak. And the quick ratio decreases from 0.9217 in 2018 to 0.2219 in 2023, indicating that Company A's solvency is weakened. Meanwhile, during 2018-2023, Company A has a negative cash flow interest coverage multiple for four years, indicating that Company A needs to focus on the debt servicing risk caused by cash shortage. In terms of profitability, Company A's return on net assets is lower than the industry average and the gap with the industry is gradually widening, indicating that the company's profitability is lower than the level of the real estate industry and is declining. In terms of operating ability, Company A's accounts receivable turnover ratio accelerated from 2.05 to 21.93, but still far below the industry average, indicating that Company A's accounts receivable turnover ratio needs to be improved and its operating ability needs to be enhanced. In terms of growth capacity, Company A's total asset growth rate and operating profit growth rate both fluctuate more, but the overall trend is declining, synthesizing the two indicators shows that Company A's development capacity is declining. As for cash flow, Company A's cash flow per share from operating activities fluctuates greatly and is lower than the industry average in 2021-2023 and 2019, and Company A's management should pay attention to this.

Funding

2024 Shanxi Vocational Education Teaching Reform and Practical Research Project (No. 202403128): “Practice and Research on Innovation and Entrepreneurship Education in Vocational Colleges”.

"Dual-Qualified" Eminent Teacher Studio (No. M28): Shanxi Culture and Tourism Vocational University Preparatory Group.

References

[1] Li, Y. (2022). Security and Risk Analysis of Financial Industry Based on the Internet of Things. Wireless Communications and Mobile Computing, 2022(1), 6343468.

[2] Pandow, B. A., Bamhdi, A. M., & Masoodi, F. (2020). Internet of things: financial perspective and associated security concerns. International Journal of Computer Theory and Engineering, 12(5), 123-127.

[3] Sosnovska, O., & Zhytar, M. (2018). Financial architecture as the base of the financial safety of the enterprise. Baltic journal of economic studies, 4(4), 334-340.

[4] Kaczmarek, J., Alonso, S. L. N., Sokołowski, A., Fijorek, K., & Denkowska, S. (2021). Financial threat profiles of industrial enterprises in Poland. Oeconomia Copernicana, 12(2), 463-498.

[5] Sun, S. (2019). Financial risk assessment of listed enterprises in marine engineering equipment manufacturing industry. Journal of Coastal Research, 94(SI), 788-792.

[6] Wang, F., Ding, L., Yu, H., & Zhao, Y. (2020). Big data analytics on enterprise credit risk evaluation of e-Business platform. Information Systems and e-Business Management, 18(3), 311-350.

[7] Ding, Q. (2021). Risk early warning management and intelligent real-time system of financial enterprises based on fuzzy theory. Journal of Intelligent & Fuzzy Systems, 40(4), 6017-6027.

[8] Dokiienko, L. (2021). Financial security of the enterprise: an alternative approach to evaluation and management. Business, Management and Economics Engineering, 19(2), 303-336.

[9] Blakyta, G., & Ganushchak, T. (2018). Enterprise financial security as a component of the economic security of the state. Investment management and financial innovations, 15(2), 248-256.

[10] Srivastava, S. (2017). Qualitative Methodology Pertaining to Risk Assessment in Small Enterprises as per Indian Industrial Scenario. International Research Journal of Engineering and Technology (IRJET), 4(8), 1877-1886.

[11] Song, Y., & Wu, R. (2022). The impact of financial enterprises’ excessive financialization risk assessment for risk control based on data mining and machine learning. Computational Economics, 60(4), 1245-1267.

[12] Докієнко, Л. М., Гринюк, Н. А., Лапко, О. О., & Крамарев, Г. В. (2020). Financial security diagnostics tools for corporate enterprises. Financial and credit activity problems of theory and practice, 3(34), 184-194.

[13] Li, Z. (2024). Identification of enterprise financial risk based on logistics model. International Journal of Business Intelligence and Data Mining, 25(1), 106-127.

[14] Lyu, J. (2022). Construction of enterprise financial early warning model based on logistic regression and BP neural network. Computational Intelligence and Neuroscience, 2022(1), 2614226.

[15] Li, Z., Ying, Y., & Wang, Y. (2024). Identification of Financial Risks in Enterprises Based on Logistic Models. International Journal of Multiphysics, 18(3).
[16] Sherrie Wallace. (2022). A Probabilistic Interpretation of Benford's Law to Evaluate Financial Fraud. International Journal of Management Science Research,5(6).
[17] Irsyadi Yani, Ismail Thamrin, Dewi Puspitasari, Barlin, Yulia Resti & Dendy Adanta. (2025). Plastic-Type Classification for Sorting System Based on Digital Image using Multinomial Logistic Regression with k-Fold Cross Validation. Engineering Letters,33(4).
[18] Jiajia Chen, Cheng Shen, Haiyan Xue, Benyin Yuan, Bing Zheng, Lianglan Shen & Xingxing Fang. (2025). Development of an early prediction model for vomiting during hemodialysis using LASSO regression and Boruta feature selection. Scientific Reports,15(1),10434-10434.