A study of bank credit strategy based on logistic model

Siyuan Chen^{1,*}

¹School of Mathematics and Science, Beijing University of Chemical Technology, Beijing, China *Corresponding author: 2019130204@buct.edu.cn

Keywords: Logistic; quantitative analysis; credit strategy; multivariate linear programming

Abstract: This thesis is based on logistic model and multivariate linear programming, using MATLAB to study the lending strategies of banks to MSMEs. We first analyzed the factors that can affect the credit risk of enterprises, and studied the influence of four factors on the credit risk of enterprises, namely, profit, development potential, supply and demand, and credit rating, and then used the data to measure the above four factors, and predicted the default risk index of each enterprise through logistic model, and to make the results more intuitive, we introduced the credit score $A=100 \times (1 - \text{default risk index})$, thus providing a quantitative analysis of the credit risk of the enterprise. Then, we determined the interest rates of bank loans to enterprises with different credit ratings, and finally, we determined the credit strategy by considering the credit score. At the same time, we considered the case where there is no credit rating of enterprises, and to develop credit strategies for different enterprises, we first need to derive the credit score of each enterprise using the logistic model above, and then adjust the credit strategy by dividing the rating of each enterprise according to the ratio of each credit rating and the credit score above. The loans given to the enterprises of grades A, B, and C are taken as unknowns, and the total number of 100 million and each enterprise can only lend 100~1 million yuan are taken as restrictions, and the maximum value of the profit is solved by using multivariate linear programming to derive the loan amount for each grade of enterprises. Finally, the credit strategy is determined by ranking the enterprises in terms of their development potential and credit score, and lending the loans to each enterprise in turn.

1. Introduction

In today's era of mass entrepreneurship and innovation, there are countless small, medium and micro enterprises, however, the development of enterprises are inseparable from capital, and in the early stage of entrepreneurship, loans from banks become the main source for most enterprises to obtain capital. The development of enterprises is related to the development of the country, and banks have limited funds, so how to lend to enterprises becomes a crucial issue [3].

In order to develop a suitable credit strategy for the enterprise, we first need to quantify the credit risk of the enterprise. We consider the strength of the enterprise, the stability of supply and demand, and the assessment of the enterprise's creditworthiness as the main factors influencing credit risk from these three aspects, and finally the influencing factors are considered together through a weighted assignment method to finally give a credit strategy. In this thesis, the strength of the enterprise is measured by the amount of profit of each enterprise in the past three years and its

future development potential, where the development potential is measured by the annual profit growth rate of the enterprise; the supply and demand relationship is measured by the standard deviation of the ratio of total sales price and tax to total input price and tax, and the creditworthiness assessment is measured by the credit rating and invoice status, and finally the above three factors are weighted by the logistic model. Finally, the three factors are weighted by a logistic model to quantify the credit risk of the firm [4].

The relationship between the customer churn rate and the annual interest rate of the loan for enterprises with different credit ratings is presented in a table, which takes into account the annual interest rate of the loan and the customer churn rate, and calculates different interest rates for enterprises with different credit ratings in order to maximize the bank's revenue. In order to keep the bank's credit risk as low as possible, we give priority to lending to enterprises with high creditworthiness and strength in the case of unknown loan amounts, and then lend to enterprises with high creditworthiness and weakness but good development potential and give appropriate interest rate concessions.

2. Quantitative analysis of credit risk

2.1. Data pre-processing

2.1.1. Profits

In general, the strength of an enterprise is positively correlated with the size of its profit. The size of profit can be obtained by subtracting the total amount of input from the total amount of sales in recent years. In this regard, we first import the input invoice information and the output invoice information into the same worksheet, and use subtraction to calculate the total profit of all the enterprises in the annex in recent years. The total profit of some enterprises is shown in Figure 1.

2.1.2. Corporate reputation

Since the firms in Data 1 are all firms with credit records, for these firms, their corporate creditworthiness can be judged by both credit rating and invoice status. For Data 2, the enterprises in it are all enterprises without credit records. Therefore, for the enterprises in Data 2, the invoice status is an important factor in evaluating their corporate creditworthiness.

In the data, a negative invoice is one in which the purchaser cancels the transaction for any reason after invoicing for the transaction activity. A voided invoice is a transaction that is cancelled by either party for any reason after the invoice is issued for the transaction activity. Thus, we assume that the seller and the buyer are each responsible for half of all voided invoices in the input and output invoices of a business. Therefore, for each business, the invoice status reputation is: (input negative invoices + input voided invoices/2 + output voided invoices/2)/ (total number of input invoices + total number of output invoices). The result of the solution is partially shown in Figure 2.

2.1.3. Supply and Demand

For an enterprise, its supply and demand relationship directly determines whether the enterprise can develop stably. We first find out the ratio of sales amount to input amount for each quarter, and then find the standard deviation of the ratio to calculate whether the supply and demand relationship of an enterprise is stable. If the standard deviation of all the enterprises in Data 1 and Data 2 is smaller, the more stable the supply and demand relationship is, otherwise, the more unstable.

2.1.4. Development Potential

For a company that is in the early stages of development, it may not have an advantage in terms of profits compared to some long-established companies, but it has a good potential for growth. For these companies, it is not accurate to use profits to measure their strength, so we introduce the growth potential indicator.

For the growth potential, for all the companies in Data 1 and Data 2, we use the pivot table in Excel to process the data, calculate the total sales price tax created by the company in each year, and then find the growth rate of the total sales price tax for each year, and then find the average growth rate of the company in recent years. The average growth rate is used as the main indicator to represent the growth potential of the company. The final growth potential of some enterprises for Data 1 and Data 2 is shown in Figure 3 and Figure 4.

ise code	Total profit(million yuan)	Enterprise code	Total profit (million yuan)	Enterprise Code	Invoice information reputation	
E1	-181103.37	E124	-3646.98	E1	0.0422	
E2	49083.77	E125	194.75	E2	0.0428	
E3	53359.61	E126	51150.72	E3	0.0211	
E4	168301.33	E127	69509.71	E4	0.0828	
E5	886.05	E128	24538.06	E5	0.0489	
E6	8917.14	E129	27466.94	E6	0.0599	
E7	46311.96	E130	3513.22	E7	0.0481	
E8	23298.51	E131	10608.80	E8	0.0595	
E9	30458.30	E132	15490.76	E9	0.0262	
E10	36291.76	E133	9883.00	E10	0.0692	
E11	906.71	E134	12774.28	E11	0.0534	
E12	13856.51	E135	13122.84	E12	0.0356	
E13	14204.59	E136	901.00	E13	0.0676	
E14	223.84	E137	4313.67	E14	0.0758	
E15	22155.69	F138	9006 12	E15	0.0607	

Figure 1: Total profit

Firm Code	Supply and domand variance	Firm Code	Supply and demand variance
E1	0.2784	E124	0.1353
E2	0.0686	E125	0.2251
E3	0.0442	E126	0.0724
E4	0.3538	E127	0.0023
E5	0.1476	E128	0.1487
E6	0.2116	E129	0.1015
E7	0.3127	E130	0.5104
ES	0.0000	E131	0.1624
E9	0.0083	E132	0.0852
E10	0.0254	E133	1.1805
E11	0.2857	E134	0.0181
E12	0.1961	E135	0.1253
E13	0.2071	E136	0.3492
E14	0.0000	E137	1.4878
E15	0.0252	E138	0.1382

Figure 2: Reputation

Enterprise code	Development Potential	Company Code	Development Potential
E1	-0.0102	E124	4.3526
E2	0.0287	E125	2.6067
E3	-0.4815	E126	1.8122
E4	-0.6792	E127	0.6845
E5	7.3636	E128	1.9468
E6	5.5410	E129	0.5210
E7	-0.6121	E130	20.1238
E8	-0.2163	E131	0.8019
E9	-0.2316	E132	2.0644
E10	1.9991	E133	5.7536
E11	0.2889	E134	5.8672
E12	-0.5014	E135	4.7921
E13	-0.2663	E136	5.6386
E14	-0.0136	E137	3.4026
E15	0.4026	E138	1.6486

Figure 3: Supply and demand Figure 4: Development potential

2.2. Model Building

2.2.1. Model ideas

We need to quantify the credit risk of the 123 companies in Data 1. From the information such as credit rating, input amount, output amount, and invoice status in Data 1, we evaluate the credit risk of this enterprise. We can express the information such as creditworthiness rating in terms of parameters, and finally we can get the expressions of credit risk indicators in relation to these parameters and find the credit risk indicators of these 123 enterprises for analysis. Since the total annual credit amount is fixed and unknown, we regard it as a large amount and sufficient to give loans to all enterprises except those with a creditworthiness rating of D. Therefore, we quantify the credit risk of the enterprises and obtain the credit score.

2.2.2. Logistic model

The logistic regression algorithm is used to analyze existing data, establish a regression formula, and use the regression formula to calculate the function value of the new sample data to predict the new sample. Logistic regression algorithm mainly solves the dichotomous classification problem,

the function receives the predicted data and performs the prediction classification. In particular, when dealing with dichotomous problems, the regression function outputs 0 or 1 to determine the category. The 0 and 1 values allow us to predict which firms are likely to default and to reduce or deny loans to these firms [1].

2.2.3. Quantitative analysis methods

The expression of the Sigmoid function used in the logistic regression model is as follows [2].

$$f(x) = \frac{e^x}{1 + e^x} = \frac{1}{1 + e^{-x}} \tag{1}$$

We calculate the default risk index by setting the weights, and the default risk index is P.

$$\hat{P} = \frac{e^{B_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n}}{1 + e^{B_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_n x_n}}$$
(2)

Since the main influencing factors of enterprise credit risk are profit, supply and demand index, development potential index, and creditworthiness index, we set these four factors to be x_1, x_2, x_3, x_4 and the weights to be $\beta_i (i = 1, 2, 3, 4)$ respectively.

We obtain the enterprise credit risk index P.

$$\hat{P} = \frac{e^{B_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4}}{1 + e^{B_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4}}$$
(3)

By making the enterprise credit score, we can quantify the credit risk analysis of the enterprise based on the enterprise credit score.

2.3. Model solving

(1) Solution of the logistic model

We import the profit, supply and demand indicators, development potential indicators, and reputation indicators of 123 enterprises in Data 1 into MATLAB software, and solve the logistic model through MATLAB, we get the regression coefficients and evaluation results of each indicator, and the MATLAB partial results are shown as follows Figure 5.

Figure 5: Mtalab solution results

(2) Quantitative Analysis

We use the logistic model to calculate the regression coefficients of the four indicators, after considering these four regression coefficients as the impact weights of the four indicators on the default risk index of the enterprise β_i , the resulting weights are substituted into the Sigmod function to obtain the default risk index of the enterprise, i.e., the credit risk index P. The higher the value of P, the more likely the enterprise is to default, and in order to quantitatively analyze the

credit risk, the credit score A is introduced. $A = 100 \times (1 - P)$, from the credit score we rate these enterprises, and then the credit strategy is decided by the rating.

The regression coefficients of each factor, i.e., the weights, are calculated using MATLAB software: -1.161, -6.0246e-07, 1.2857e-05, -0.00082527, 2.6703. (where -1.161 is a constant term, and the rest are the weights of profit, supply and demand indicators, development potential indicators, and creditworthiness indicators, in that order).

The corporate credit scores corresponding to the p-values calculated by substituting each weight into the Sigmod function are shown in the following table 1 (only some of them are shown).

Company Code	Profits	variance	Development potential	Credibility	Default or not	P	Score A
E1	-181103.37	0.28	-0.01	0.04	0.00	0.28	71.96
E2	49083.77	0.07	0.03	0.04	0.00	0.25	74.55
ES	53359.61	0.04	-0.48	0.02	0.00	0.24	75.66
E4	168301.33	0.35	-0.68	0.08	0.00	0.26	73.81
E5	886.05	0.15	7.36	0.05	0.00	0.26	73.81
E6	8917.14	0.21	5.54	0.06	0.00	0.27	73.30
E7	46311.96	0.31	-0.61	0.05	0.00	0.26	74.23
E8	23298.51	0.00	-0.22	0.06	0.00	0.27	73.39
E9	30458.30	0.01	-0.23	0.03	0.00	0.25	75.17
E10	36291.76	0.03	2.00	0.07	0.00	0.27	73.07
E11	906.71	0.29	0.29	0.05	0.00	0.27	73.47
E12	13856.51	0.20	-0.50	0.04	0.00	0.25	74.51
E13	14204.59	0.21	-0.27	0.07	0.00	0.27	72.87
E14	223.84	0.00	-0.01	0.08	0.00	0.28	72.27
E15	22155.69	0.03	0.40	0.06	0.00	0.27	73.33
E16	23325.42	0.00	1.53	0.07	0.00	0.27	72.87
E17	4297.95	0.24	3.45	0.08	0.00	0.28	72.34
E18	6900.22	0.27	3.80	0.06	0.00	0.27	73.18
E19	807.95	0.28	1.15	0.07	0.00	0.27	72.64
E20	1008.12	0.25	1.22	0.07	0.00	0.27	72.65
E21	-847.78	0.41	4.51	0.06	0.00	0.27	73.18
E22	5570.75	0.29	5.24	0.10	0.00	0.29	71.27
E23	1094.84	0.14	0.97	0.04	0.00	0.26	74.10
E24	7570.46	0.12	0.40	0.05	0.00	0.26	73.56

Table 1: C	Corporate	Credit	Score
------------	-----------	--------	-------

3. Analysis of lending strategies for companies with credit history

Bank R's credit strategy is considered in terms of both the interest rate and the amount of the loan.

(1) Determination of loan interest rate

Let the bank's loan profit be Z, the loan interest rate be q, the customer churn rate be f, the number of enterprises of each credit level be m_i , the amount of loans to enterprises of each level be k_i , and the total amount of loans be D, where $Z = q \times (1 - f) \times D$, $D = \sum_{i=1}^{3} m_i \times k_i$.

When the loan interest rate q rises, the customer churn rate f rises. With the total loan amount D unchanged, the bank's profit is maximized when $q \times (1-f)$ is maximum, and the resulting maximum value and the determined interest rate are as follows Figure 6.

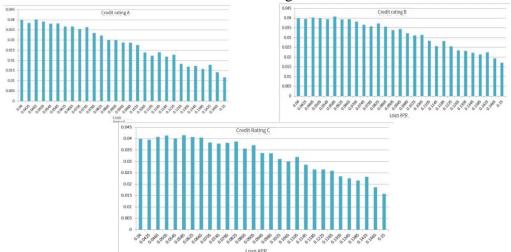


Figure 6: Diagram of annual interest rate for A, B and C grades

(2) Determination of loan amount

Since the bank's capital is unknown, it is assumed here that the bank has sufficient capital to maximize its profit by avoiding firms with small credit scores and giving priority to firms with high credit scores. We set up a system that refuses to lend to a company when its credit score is below 50, and for companies with a score above 50, we rank them according to their growth potential, starting with the top-ranked companies.

4. Analysis of lending strategies for companies with no credit history

Based on the previous section, Data 2 has no credit rating of the enterprises and the total credit amount of the bank is 100 million yuan, and we need to allocate 100 million yuan to these enterprises when each enterprise has a limit on the amount of loan. We first perform a quantitative credit risk analysis on all enterprises according to the method of the first question, and then rate the enterprises according to the results of the quantitative analysis. From the rating results we set different interest rates for different credit ratings and give loans to different credit ratings of k_i . We develop a credit strategy by linear programming these variables to find the maximum profit the bank gets and the loans to each credit rating when the maximum is obtained.

First, we quantitatively assess the credit risk of enterprises, and count the number of enterprises of grades A, B, and Cm_i . The interest rate is q, the customer attrition rate is f, the number of enterprises of each grade is m_i , the loans to enterprises of each grade are, and the expression of bank profit is $\begin{cases} Z = q \times (1 - f) \times D \\ D = \sum_{i=1}^{3} (m_i \times k_i) \end{cases}$ The maximum value of q×(1-f) is known from the first problem, and then linear programming is performed using MATLAB according to the above equation k_1, k_2, k_3 . The amount of loans to enterprises with credit grades A, B, and C when the bank profit is maximized, respectively k_1, k_2, k_3 . From this, we then rank the enterprises according to their development potential in each of the three grades A, B, and C, and lend to each enterprise in turn.

Using the logistic model to quantitatively assess the credit risk of the enterprises, the scores and ratings of the enterprises are as follows Table 2 (only part of the results are shown).

ode.	- Profit -	variance	 Development Potencial Creation 	dibility -	P	 Score - 	exp	 Grade (
E124	-3646.98	0.14	4.35 (. 14	0,38	62, 23	0.61	D
E125	194.75	0.23	2.61 0	.14	0.38	62.36	0.60	D
E126	51150.72	0.07	1.81 0	.13	0.35	65.24	0.53	D
£127	69509.71	0.00	0.68 0	. 03	0.26	74.05	0.35	A
E128	24538.06	0.15	1.95 (. 04	0.29	71.29	0.40	A
£129	27466.94	0.10	0.52 0	. 09	0.32	67.98	0.47	C
E130	3513.22	0,51	20.12 0	. 04	0.29	71.20	0.40	A
E131	10608.80	0.16	0.80 0	. 09	0.33	67.29	0.49	C
E132	15490.76	0.09	2.06 0	. 04	0.28	71,50	0.40	A
E133	9883.00	1.18	5.75 0	. 04	0.29	71.00	0.41	A
E134	12774.28	0.02	5.87 0	. 06	0.30	70.02	0.43	8
E135	13122.84	0.13	4.79 0	. 03	0.28	72.14	0.39	A
E136	901.00	0.35	5.64 0	. 03	0.28	71.80	0.39	A
E137	4313.67	1.49	3.40 0	. 06	0.30	69.64	0.44	B
£138	9006.12	0.14	1.65 0	. 04	0.29	70.83	0.41	8
E139	20322.02	0.01	0.62 0	. 05	0.29	70.85	0.41	B
E140	21771.77	0.03	1.08 0	. 06	0.30	69.61	0.44	B
E141	12706.57	0,08	4.09 0	, 06	0.30	70.09	0.43	B
E142	5252.99	0.15	9.71 0	. 07	0.31	69.27	0.44	B
E143	6040.29	1.73	6.53 0	. 05	0.29	70.55	0.42	8
E144	10533.41	0.04	25.26 0	. 04	0.28	71.68	0,40	A
E145	1593.08	0.14	4.29 0	. 07	0.32	68.45	0.46	C
E146	1561.86	0.34	23.76 0	. 06	0.30	70.23	0.42	B
E147	3497.71	0.79	23.96 0	. 06	0.30	70.12	0.43	B

Table 2: Enterprise score and rating results

There are 66 firms with grade A, 93 firms with grade B, 76 firms with grade C, and 67 firms with grade D. We give loans to firms with grades A, B, and C with a total loan amount of k_1, k_2, k_3 each.

We perform linear programming.

$$\begin{pmatrix} k_1 + k_2 + k_3 = 10000 \\ 660 \le k1 \le 6600 \\ 930 \le k2 \le 9300 \\ 760 \le k3 \le 7600 \end{cases}$$

$$(4)$$

To maximize the value of interest rate $q \times (1 - customer churn rate f)$, we should take the interest rate for grade A to be 0.0465, for grade B to be 0.0585 and for grade C to be 0.0585, we can obtain the bank's profit.

$$z = 0.0465 \times (1 - f_1) \times k_1 + 0.0585 \times (1 - f_2) \times k_2 + 0.0585 \times (1 - f_3) \times k_3$$
⁽⁵⁾

Through linear programming, we obtain that Z achieves a maximum value of \$4.15 million when x_1 =660, x_2 =1740, and x_3 =7600, and the results of ranking the development potential of different types of firms are as follows Table 3 (only part of the graph is shown).

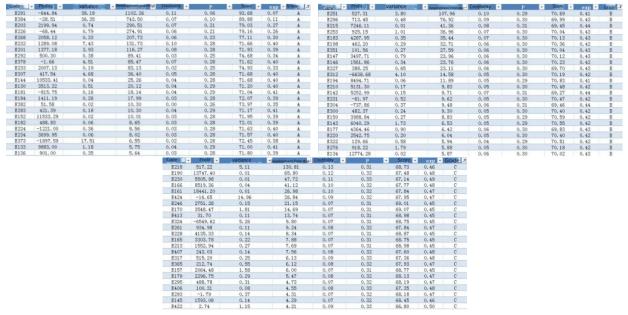


Table 3: Ranking of development potential of A, B and C grade companies

Since the interest rate is determined according to the credit rating, for the bank, as long as no enterprise defaults, the bank's profit is only related to the amount of the loan allocated to the different interest rates, not to the actual enterprise, so after dividing the loans into different interest rates, only the bank least likely to default is chosen to make the loan in order to maximize the mathematical expectation.

Thus, according to the above diagram, we combine the growth potential of the enterprise and the loan amount for different credit ratings, and now divide the loans into three parts, A, B and C, and allocate them in the order of their growth potential.

5. Summary

Some samples were taken, and whether or not to default was taken as the dependent variable, and enterprise profit index, development potential index, creditworthiness index, and supply and demand index were used to predict the rest of the samples in Data 1 using logistic function, and the accuracy rate was only 74%, indicating that there is room for model improvement.

Due to the limitation of known data, we only measure the credit risk of enterprises from four

aspects: profit, development potential, creditworthiness indicator, and supply and demand relationship. However, in real life, for most enterprises, the stability of the enterprise's capital chain, the size of the enterprise, and the availability of assets to secure loans are all important factors in measuring credit risk. After taking all these factors into consideration, we believe that the results will be more accurate if we use the logistics model to make predictions.

After improving our model, we can analyze the credit score of micro, small, medium and large enterprises and calculate the credit score of the enterprises. It can even be used as an important reference indicator for evaluating the overall strength of enterprises. The model in the third question can be used to analyze the impact of unexpected situations on each enterprise, especially to assess the impact on the supply and demand of enterprises, so that it can help banks to adopt different credit strategies in different situations.

References

[1] She Chaobing. Study on the application of logistic regression on bank personal credit risk assessment [J]. Technology and Innovation, 2018(19):113-114,118-119.

[2] Liang Han. Research on credit risk identification and prevention strategies for small and micro enterprises of ICBC in M city [D]. Sichuan: Southwest University of Science and Technology, 2019.

[3] Du Qian. Exploring the problems and strategies of credit risk management of small and micro enterprises in local commercial banks [J]. Enterprise Reform and Management, 2023(2):100-102.

[4] Xia Menglin. Research on the construction of credit decision model of commercial banks for small and micro enterprises under big data [D]. Chongqing University of Technology and Industry, 2022.