ISSN: 2472-3703 DOI: 10.6911/WSRJ.202401_10(1).0018

# Credit Decision Making for Micro, Small and Medium Enterprises Based on XGBoost Regression and Multi-Objective Planning

Yuyang Qi[1, †], Xiaolin Liu[2, †], Yueru Gong[2]

[1]Business School University of Jinan, University of Jinan, Jinan, China

[2]School of Electrical Engineering, University of Jinan, Jinan, China

† These authors also contributed equally to this work

## Abstract

**It is especially important for banks to establish a model to make a comprehensive assessment of the credit risk of micro, small and medium-sized enterprises (MSMEs) and adopt appropriate credit strategies when the strength and creditworthiness of the enterprises are known. This paper is based on machine learning and multi-objective planning to formulate credit decisions for MSMEs. First of all, the credit strategy for enterprises with credit records is developed, and a reliable and stable credit rating prediction training model is obtained through machine learning regression-XGBoost regression analysis. By establishing a multi-objective planning model with the objective function of maximizing the profit and minimizing the risk, we obtain the bank's credit strategy for enterprises with different credit ratings and sizes when the total amount of credit is fixed in a year. Then, the credit strategy of enterprises without credit records is formulated, and the credit ratings of 302 enterprises without credit records are obtained with the help of the XGBoost regression analysis training model. With the help of the multi-objective planning model of credit strategy, the bank's credit strategy for enterprises is obtained when the total annual credit amount is 100 million yuan. The results of the model are reliable and accurate, and can be further generalized to other similar areas of research.**

## Keywords

**Machine Learning, XGBoost, Regression Analysis, Multi-Objective Planning, Credit Strategy.**

## 1. INTRODUCTION

In today's rapid economic development, micro, small and medium-sized enterprises (MSMEs) play an increasingly important role in promoting national economic development, alleviating employment pressure, and promoting social stability. However, since MSMEs are limited by their scale and lack the necessary collateral assets for loans, it is particularly important for banks to model how to make a comprehensive assessment of the credit risk of MSMEs and adopt appropriate credit strategies when the strength and creditworthiness of the enterprises are known [1, 2].

In this context, this paper is based on machine learning and multi-objective planning to study the credit decision-making of micro, small and medium-sized enterprises. First of all, we study the credit strategy of enterprises with credit records, according to the enterprise information and transaction note information, extract the enterprise default, annual profit, daily profit and other characteristics, at the same time, the data and the characteristics of the results of the operation of the normalization process, and through the machine learning regression-XGBoost

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202401_10(1).0018

regression analysis, to obtain a reliable and stable credit rating prediction training model, through the establishment of the objective function for the profit is the maximum, By establishing a multi-objective planning model with the objective function of maximizing profit and minimizing risk, we obtain the credit strategies of banks for enterprises with different credit ratings and sizes when the total annual credit amount is fixed. Then, we study the credit strategy of enterprises without credit records, and obtain the credit ratings of 302 enterprises without credit records with the help of XGBoost regression analysis training model. With the help of the multi-objective planning model of credit strategy, we obtain the credit strategy of the bank for enterprises when the total annual credit amount is 100 million yuan [3, 4].

## 2. MODEL FORMULATION AND SOLVING

### 2.1. Research on credit strategies for companies with credit history

As a comprehensive index of enterprise credit risk assessment, credit rating is affected by many factors such as enterprise default, annual profit, daily profit, the number of enterprises that have transactions with credit enterprises, the rate of canceled invoices, the rate of negative invoices, the amount of tax, etc. We get the relationship between credit rating and each influential factor (enterprise default, annual profit, daily profit, etc.) through XGBoost regression, and can get the credit rating of 302 enterprises with no credit records on the basis of this [5]. The relationship between the credit rating and the influencing factors (enterprise default, annual profit, etc.) can be used as a basis to obtain the credit rating of 302 enterprises without credit records, in order to further get the credit strategy.

For the credit rating of each enterprise, the principle of "high rating, high score" is adopted, and the assignment method of "rating A-value 3, rating B-value 2, rating C-value 1, rating D-value 0" is used. According to the formula:

$$\alpha = \frac{B-A}{A} \times 100\% \tag{1}$$

Calculate to get the rate of change of annual profit for the year 18-19, where $\alpha$ denotes the rate of change of annual profit, B denotes the annual profit in year 19, andA denotes the annual profit in year 18.

In order to achieve consistency in the scale of the influencing factors, we used Eq.

$$X' = \frac{X-Xmin}{Xmax-Xmin} \tag{2}$$

The calculation results of the influencing factors $X2-X9$ were normalized as shown in Table 1.

**Table 1.** Results of partial normalization

| Enterprise code | E1 | E2 | …… | E122 | E123 |
|---|---|---|---|---|---|
| $X2$ | 0.779156199 | 1 | …… | 0.920705915 | 0.920735096 |
| $X3$ | 0.186638532 | 0.316547723 | …… | 0.278843296 | 0.360982154 |
| $X4$ | 0.104075402 | 0.690296031 | …… | 0.011549184 | 0.000398248 |
| $X5$ | 0.026942075 | 0.011972158 | …… | 0.060402685 | 0 |
| $X6$ | 0.06 | 0.02 | …… | 0.02 | 0.00 |
| $X7$ | 0.03 | 0.08 | …… | 0.14 | 0.49 |
| $X8$ | 0.76 | 0.32 | …… | 0.17 | 0.68 |
| $X9$ | 0.715415464 | 0.71627119 | …… | 0.711564389 | 0.715311369 |

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202401_10(1).0018

Based on the given enterprise information and transaction note information, we build XGBoost regression model through the training dataset, and calculate the importance of features through the built XGBoost model as shown in Figure 1, and apply the built XGBoost regression model to the training and testing data, and get the model evaluation results as shown in Table 2. With the help of this saved training model, the processed data are substituted into this training model to calculate the prediction to further get the credit strategy.

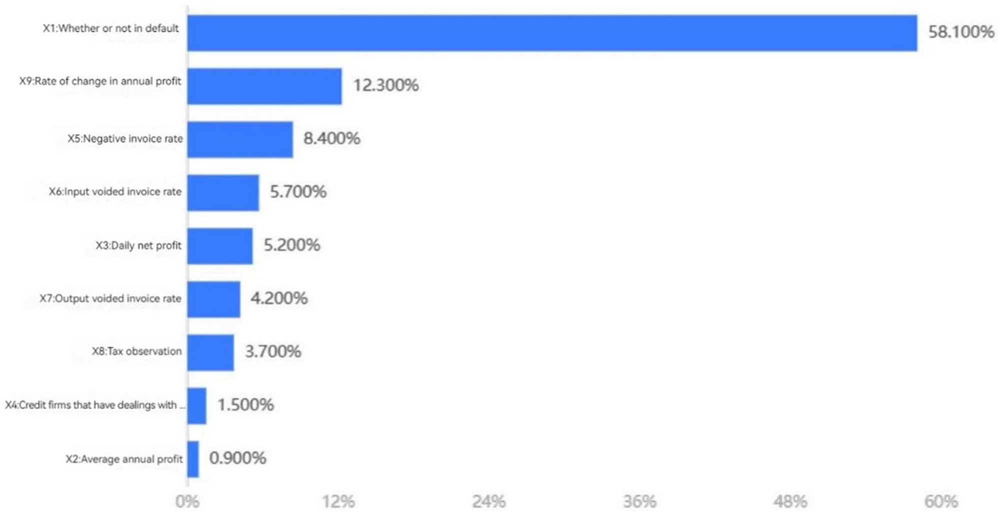


**Figure 1.** Ranking of feature importance

**Table 2.** Results of model evaluation

| | MSE | RMSE | MAE | MAPE | $R^2$ |
|---|---|---|---|---|---|
| training set | 0 | 0.002 | 0.001 | 3.534 | 1 |
| test set | 0.287 | 0.536 | 0.38 | 72.297 | 0.699 |

Using the significance of the features returned by the XGBoost regression, we find that a firm's defaults and annual profit change rate (firm's growth prospects) have the greatest impact on the firm's creditworthiness rating, which is consistent with the reality of the situation.

Meanwhile, the prediction evaluation indexes of the cross-validation set, training set and test set all affirmed the accuracy of the model to some extent. In order to further verify the fitting effect, the prediction graph of the test data is plotted as shown in Figure 2, which further reflects the accuracy of the model prediction results.

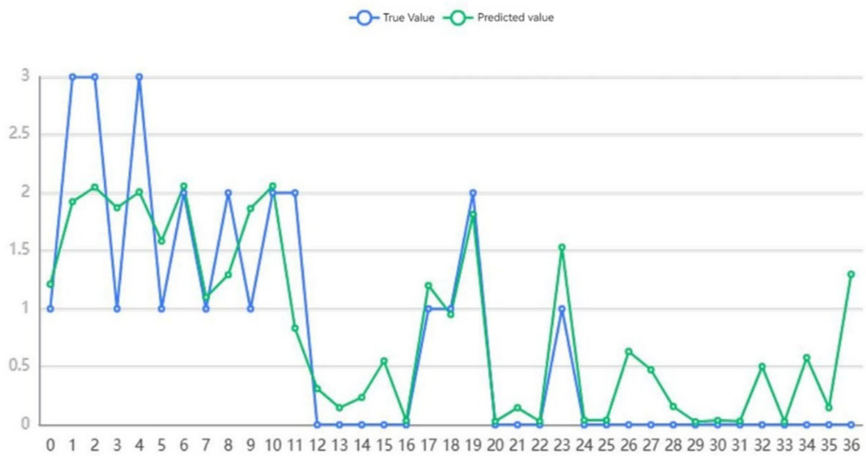


**Figure 2.** Test data prediction graph

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202401_10(1).0018

As one of the important financial institutions, the bank provides loans for SMEs not only as an important measure to respond to the call of the state to support SMEs, but also as a way to improve the utilization rate of surplus capital and to obtain income from investment. For the bank, under the constraints of fixed total annual credit amount, loan amount of the lending enterprise, annual interest rate and other constraints, how to recover the principal and interest of the loan within the stipulated period (i.e., maximize the profit and minimize the risk) is the ultimate goal of the bank to provide loans for MSMEs. This is a multi-objective planning problem.

Let the amount of loan provided by the bank for each enterprise be as follows $ai$ (unit: ten thousand dollars), and the interest rate of the loan is $ri, bi$ denote $0-1$ variable (1 for lending, 0 for not lending).) The interest rate is the variable (1 for lending, 0 for not lending). When banks provide loans to SMEs, they simultaneously pursue the two objectives of profit maximization and risk minimization, so the profit maximization objective function1 is shown below.

$$\max Z=\sum_{i=1}^{99} ai \times ri \times bi \tag{3}$$

Banks do not lend to enterprises with credit rating of D in principle, so in the 123 enterprises to remove the remaining 99 D-rated enterprises, and these 99 enterprises are ranked according to the credit rating, the first 27 enterprises with credit ratings for the first 27 enterprises, the 28th-65th enterprises with credit ratings for the first 27 enterprises, the 28th-65th enterprises with credit ratings for the second 99 enterprises. A The first 27 enterprises have credit ratings of D, the 28th-65th enterprises have credit ratings of D, and the 66th-99th enterprises have credit ratings of D. $B$ The first 27 are enterprises with credit ratings, the 28th-65th are enterprises with credit ratings, and the 66th-99th are enterprises with credit ratings. $C$ The first 27 enterprises were given credit ratings, 28-65 enterprises were given credit ratings and 66-99 enterprises were given credit ratings.

When the bank's total annual credit is fixed, the greater the proportion of the bank's lending to the total amount of lending for businesses with credit ratings of A The greater the proportion of the total amount of lending to the enterprises with credit rating of the bank, the bank's credit risk is minimized. Therefore, the risk minimization objective function 2 is shown below.

$$\max w=\frac{\sum_{i=1}^{27} ai \times bi}{\sum_{i=1}^{99} ai \times bi} \tag{4}$$

The total annual credit of the bank is fixed, and here it is assumed to be 100 million dollars, i.e., the actual amount lent by the bank shall not exceed 100 million dollars, so the constraint 1 is shown below.

$$\sum_{i=1}^{99} ai \times bi \leq 10000 \tag{5}$$

The bank's loan amount to the firm identified for lending is 100-100 million dollars, so Constraint 2 is shown below.

$$0 \leq (ai-10) \times bi \leq 90 \tag{6}$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202401_10(1).0018

The bank's interest rate on loans to firms identified as lenders is 4%-15%, so Constraint 3 is shown below.

$$0 \le (ri - 4\%) \times bi \le 11\% \tag{7}$$

The bank can only lend to a maximum of 99 firms, so constraint 4 is shown below.

$$\sum_{i=1}^{99} bi \le 99 \tag{8}$$

The relationship between lending rate and customer churn is shown in Figure 3, as the interest rate increases, the credit ratings of A and B. The credit ratings of the firms with A and B are increased, C of the credit rating of A, B, and the enterprise's customer turnover rate increase. But for the higher credit rating of the enterprise, the customer for the enterprise's higher degree of trust, and thus the requirements are higher, so with the increase in lending rates, at the same interest rate level, its customer turnover rate is relatively large. We believe that, for the credit rating of A for the enterprise, its customer turnover rate of 50% and above, the enterprise will suffer a significant loss, so the company with a credit rating of A should not be higher than 0.0745; similarly, for a company with a credit rating of B should not be higher than 0.1065. For a business with a credit rating of C For a firm with a credit rating of 0.1065, the interest rate on its loans should not exceed 15% of the interest rate ceiling set by the bank. Therefore, the interest rate constraints in conjunction with customer churn 5 are shown below.

$$\begin{aligned} &ri \le 0.0745\ i = 1,2 \dots\dots 27 \\ &ri \le 0.1065\ i = 28,29, \dots\dots 65 \\ &ri \le 0.15\ i = 66,67, \dots\dots 99 \end{aligned} \tag{9}$$

For the credit strategy model, we weaken the objective of profit as a constraint 6 as follows.

$$\sum_{i=1}^{99} ai \times ri \times bi \le 100 \times 99 \times 15\% \tag{10}$$

Borrowing the software solution yields a credit strategy for these firms when the bank's total annual credit is fixed.

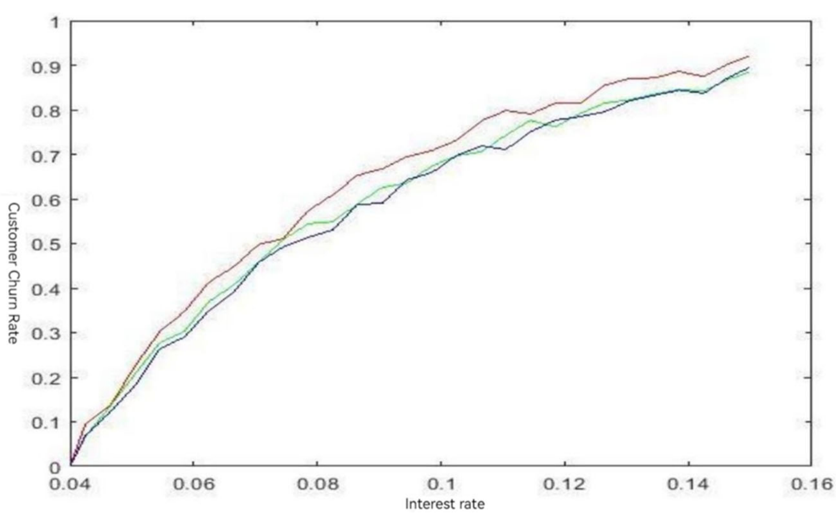


**Figure 3.** The relationship between customers and firms with a credit rating of $A$、$B$、$C$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202401_10(1).0018

## 2.2. Research on credit strategies for enterprises with no credit history

The given information of 302 enterprises lacks the relevant data of enterprise credit rating and whether they are in default or not, therefore, firstly, according to the prediction model preserved by machine learning-XGBoost regression, we should get the credit ratings of 302 enterprises with no credit records, and then we should get the credit strategy of the bank when the total amount of the annual credit of the bank is 100 million dollars by using the multi-objective planning model [6]. The credit ratings of the 302 enterprises with no credit records are shown in Table 3. The credit ratings of 302 enterprises without credit records are shown in Table 3.

**Table 3.** Projected creditworthiness ratings of 302 households with no credit history

| enterprise number | E124 | E125 | E126 | …… | E424 | E425 |
|---|---|---|---|---|---|---|
| predicted score | 2.037392 (2) | 2.528175 | 1.8651874 | …… | 1.3933897 | 1.3682768 |
| (3) | (2) | (1) | (1) | | | |
| credit level | B | A | B | …… | C | C |

# 3. MODEL EVALUATION AND DISSEMINATION

## 3.1. Model strengths and extensions

1) XGBoost regression is based on MSE, RMSE, MAE, MAPE, $R^2$ metrics to evaluate the model, and the training set $R^2 = 1$ and test set $R^2 = 0.699$ , indicating that the model is highly accurate and the data are well fitted.

2) As more data is supplied, machine learning can be updated with new codes and algorithms to improve output and produce more reliable and accurate results.

3) Widely used machine learning is a product of the development of science and technology, XGBoost regression model can not only be used in the financial field, respond to the national policy to support the development of small and medium-sized micro-enterprises, but also can be widely used in health care, engineering, aviation, aerospace technology, etc., in order to promote the development of social industries and progress.

## 3.2. Model shortcomings

1) Machine Learning has a huge demand for data, relying on rigorous logical relationships, large amounts of storage space and human behavioral specification options. In the face of some specific problems, can not make the best decision.

2) The multi-objective planning model is limited by constraints and may not be able to find the optimal solution of credit strategy.

# 4. CONCLUSION

In this paper, we study the credit decision-making of micro, small and medium-sized enterprises based on machine learning and multi-objective planning. First of all, we study the credit strategy of enterprises with credit records, according to the enterprise information and transaction note information, extract the enterprise default, annual profit, daily profit and other characteristics, at the same time, the data and the characteristics of the results of the normalization process, and through the machine learning regression-XGBoost regression analysis, to get a reliable and stable credibility rating prediction training model, through the establishment of the objective function for the profit is the maximum, By establishing a multi-objective planning model with the objective function of maximizing profit and minimizing risk,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202401_10(1).0018

we obtain the credit strategies of banks for enterprises with different credit ratings and sizes when the total annual credit amount is fixed. Then, we study the credit strategy of enterprises without credit records, and obtain the credit ratings of 302 enterprises without credit records with the help of XGBoost regression analysis training model. With the help of multi-objective planning model of credit strategy, we obtain the credit strategy of the bank for enterprises when the total annual credit amount is 100 million yuan. The results of the model are reliable and accurate, and can be further generalized to other similar areas of research.

## REFERENCES

[1] XU Miaoyao, GONG Zhaohua. Research on multi-objective credit decision-making for micro, small and medium-sized enterprises[J]. Sino-Arab Science and Technology Forum (in Chinese and English),2023(05):81-86.

[2] Yang Chao. Analysis of Credit Decision Making Based on Micro, Small and Medium Enterprises[J]. Old Brand Marketing,2023(08):151-153.

[3] CAI Wenyuan, ZHANG Yingying, YANG Xiaohui et al. Research on Credit Decision Making under Micro, Small and Medium Enterprises[J]. Commercial Observation,2022(33):41-44.

[4] ZHANG Erli, WANG Taixing,WANG Yulong. Research on Credit Decision Making of Micro, Small and Medium Enterprises[J]. Journal of Henan Institute of Education (Natural Science Edition), 2022, 31(02): 27-32+45.

[5] Mu Chengyu. Exploration of credit risk assessment model for micro and small enterprises based on Xgboost regression[J]. Business News,2022(15):119-122.

[6] Niu Yanqiu,Liu Shulin. Establishment and research on the mathematical model of credit decision-making of small and medium-sized microenterprises[J]. Heilongjiang Science,2021,12(24):118-119+122.