Bank Customers (Credit) Rating System Based On Expert System and ANN

Project Review

Yingzhen Li

Abstract

The precise rating of customers has a decisive impact on loan business. We constructed the BP network, which requires financial statements as the input and predicted ratings that range from A to AAA. To make flexible prediction, we developed another fuzzy expert system, which requested the analyst to answer raised questions. Finally, we combined these two models and provided a final rating by trade-off.

1 Introduction

Loan is one of the main source of the bank's profits, especially the loan to enterprises is a large amount of bank's fund and income. Loan quality will directly related to the recovery of bank funds, and the precise rating of customers has a decisive impact on loan business. However, the approval of loan is a very complex and onerous process. While managers expect a software-based standardized procedure to deal with this issue, nowadays many of them rely on their experience for judgement. This leaves banks with greater risk of subjectivity, but it also present the chance of novel systems of credit rating.

Our project aims at developing a customer's credit rating system, which raises an opportunity to achieve a more standardized loan approval procedure. We rated the customers from qualitative and quantitative aspect and combine their results by trade-off. The qualitative analysis was based on fuzzy expert system, while the quantitative part utilized the BP network model.

This review of project is organized as listed. Section 2 introduces the background of customer rating. Qualitative analysis is described in Section 3, while quantitative part is discussed in Section 4. Section 5 shortly introduced the software development, Section 6 evaluated the performance of the system as well as discussed the future work, and the review is concluded in Section 6.

2 Background

2.1 Credit Rating System in China

So far in China, most credit rating system is by hand operation, which is vulnerable to mistakes and risks as well as in low efficiency. Another problem is the focus on big company when they develop the system, hence they pay less attention to medium and small companies, which are the potential customers.

2.2 Overseas Banks: Best Rating Systems Available Now

Many famous banks around the world use internal rating system for risk management, and that of Bank of America and Citi Bank are the best ones among them.

- Bank of America rates their customers with the same criteria when issuing company loans. However, they may consider other factors such as the financial conditions, scale and the ownership structure when rating their customers of retail loans. They consider their system a mixture of pure judgement and template matching, and now they are developing a rating score card to improve the existing system.
- Citigroup trains their advanced rating system with a large dataset. The rating system consists of customer rating, which applied statistical models with axillary of external rating, scoring model and manager judgement, and debt rating which combines the result of customer rating and other factors such as loan structure to generate the final grades.

The rating systems discussed above still have some disadvantages, such as the subjective judgement, to be improved. A good rating system should not only rely on various kinds of advanced mathematical and economical model, but also have large data sources and industry experience to support. These requests have presented opportunities for novel researches in rating systems.

3 Qualitative Analysis: the Fuzzy Expert System

In order to add flexibility to the rating system, we design a qualitative rating process for companies based on the fuzzy expert system. Users give gradings of the questions raised by the system according to their understanding of the company's financial condition, which then are converted into the input of the expert system. The logical framework of the expert system is showned as Figure 1.

We also introduce the fuzzy logic to the expert system. The fuzzy expert system works following the listed 4 steps:

- step 1: set the inputs and outputs
  - user rates the qualities with respect to the company in 3 levels: low, average and high[1].
  - the output is a real number ranged from 0 to 10, which corresponds to 3 different levels: low, normal and high.
- step 2: set membership function
  - the grade of membership illustrates the degree of the input score belonging to different levels of the output
  - we use the triangle-shaped membership functions
- step 3: set rules of the expert system (see the project report section 2.4.2 for detailed rules)
- step 4: defuzzify the outputs by centroid defuzzification (compute the center of area under the curve if the output satisfies more than one rules, i.e. corresponds to more than 1 level)

Finally, the scores generated by all the rules are summed up to a final grade of the qualitative part, which ranges from 0 to 80.

[1]see the project report Section 2.4.1 for detailed scoring table

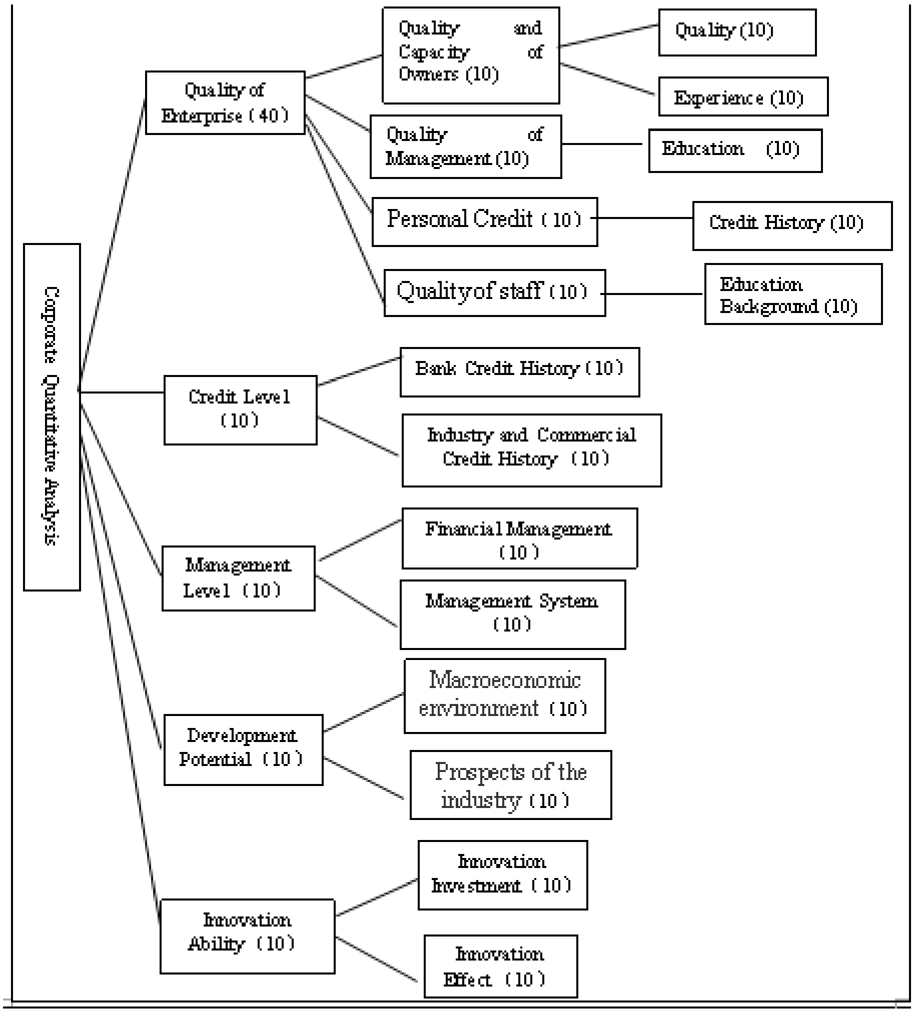

Figure 1: The logical framework of the expert system.

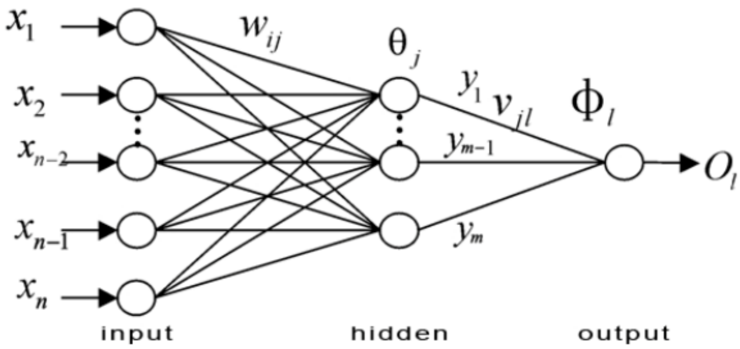

Figure 2: The structure of the BP networks. The activate function of each node is the sigmoid function

4 Quantitative Analysis: the ANN Model

4.1 Overview

Artificial neural network (ANN) is a mathematical model, which simulates the activities of neurons of biological neural network. It can encompass many models such as regression by simply tweaking with the activation functions and the network architecture. In the project we introduce a 3-layer Back Propagation networks to analyse the financial data samples and grade the testing samples.

Back Propagation (BP) network is one of the classical models of Feed Forward Networks, where connections between the units do not form a directed cycle. In this network, the information moves in only one direction, forward, from the input nodes, through the hidden nodes (if any) and to the output nodes, then return the error backward to the input layer to adjust the weights of connections. Figure 2 shows the a typical structure of 3-layer BP networks.

4.2 Our Approach

Input signal is collected from the financial statements, and after forward propagation the network generates the rating as outputs. Input data consists the features listed below[2]:

- debt paying ability
  - asset-liability ratio, liquidity ratio, total debt/EBITDA, all capitalization ratio, has-been interest multiples, quick Ratio, net cash Flows from Operating /total debt
- financial benefit
  - ROE, sales profit margins, ROA, cash inflows from operating activities / sales revenue, the ratio of profits to cost
- capital operation
  - total asset turnover, current assets turnover, inventory turnover, accounts receivable turnover, sales growth rate, capital accumulation rate, total assets growth rate, three-year average profit growth rate

[2]see the project report Section 2.4.3 for detailed computation

There are 14 nodes in the hidden layer. The output of the network is a grading ranged from 0 to 100. We convert the grading to the credit level ranged from C to AAA:

- $[90, 100]$: AAA, $[80, 90)$: AA, $[70, 80)$: A, $[60, 70)$: BBB, $[50, 60)$: BB, $[40, 50)$: B, $[0, 40)$: C.

4.3 Experiment

The network model was tested by importing 43 training samples (financial statements) and 43 test samples. We trained the network with resilient back propagation method (applying different learning rates when the sign of present gradient varies from the last one), and the computation converges after 353 iterations. This trained network had achieved the average confidence level of 0.907, which can be seen as a relatively high level.

5 Short Description of the Software

See the development document for details.

5.1 Technical Supports

- Database: MySQL
- Developing language: Java
- Graphical interface development: SWT/JFACE
- Simulation tests: Matlab

5.2 Interface Interaction Design

- qualitative analysis: a questionnaire for user scoring
- quantitative analysis: requesting the financial statement as input

5.3 Combining the Qualitative and Quantitative Results

The total score of qualitative part is 80, and that of quantitative part is 100. The final score is computed as

$$Final = Qualitative \times 0.375 + Quantitative \times 0.7,$$

and this real number score is converted into levels ranged from C to AAA as Section 4.2 indicated.

6 Evaluation and Discussion

6.1 Application Evaluation

The test results showed a high performance of our system. However, one should notice that the training set of the network was quiet small. The computational complexity of neural network, facing the big data, significantly impact on the efficiency of our system. Furthermore, the sample we collected for training and test could hardly represent most of the companies' financial condition because the credit grading of them are ranged from A to AAA. However, we notice the difficulty of getting small companies' financial statement, which limits the improvement of our system from this aspect.

The design of the rules in the expert system need to be more specific, and the scoring table should provide more details for users to refer to. Also the ratio of combination of the two gradings must be carefully chose, which requests of large dataset for testing.

6.2 Future work

As Section 2 indicated, the rating systems of domestic banks are in lower efficiency when compared with those of foreign banks. How ever, overseas banks' rating systems still have some subjective judgements which need refinement. Also different criteria are adopted by banks, which asks for the flexibility of the system's rules. The system is expected to provide the module which permits the modification of the rules by the user, as well as adapt the different types of financial statements as the input of the quantitative parts.

7 Conclusion

Precise rating of customers has a significant impact on loan business. In this project we developed a bank customer rating system to help in the loan approval procedure. The system consisted of fuzzy expert system as the qualitative part and the BP network as the quantitative part, then the results of these two models were combined by trade-off. The experiment showed a good performance of our system. Problems to be addressed are the computational complexity when facing big data, as well as the validation of the design of expert system's rules. We expect to add more flexibility of our system with balance of objective and subjective judgements, and hope the refined one to be a prototype of a standard software-based loan approval system.