

# The research on prediction and evaluation of the enterprise's risk based on data warehouse

Chunhui Zhu

Software Research and Development Center  
Harbin Bank  
Harbin, Heilongjiang, China

Ting He

School of Management  
Harbin Institute of Technology  
Harbin, Heilongjiang, China

**Abstract**—In order to improve the ability of enterprises to avoid and resist risks, in this paper, we issued the strategic and operational risks of an enterprise as the research object, established the strategic/Operational risks model, and in which used the Back-Propagation neural network algorithm, Calculated the results of the model. Finally, we validated the model and algorithm is reasonable and effective in the experiments.

**Keywords**—strategic/operational risks; prediction and evaluation model; Back-Propagation neural network algorithm; Data warehouse

## I. INTRODUCTION

With the increasing economic globalization, companies are facing ever more complex and competitive market environment, and increasing numbers of factors related to the survival and development of enterprises. In recent years, Enron, World-Com<sup>[1,2]</sup> and China Aviation Oil event happened in the international market, which raised the alarm that enterprises need to strengthen risk management<sup>[3]</sup>. So the enterprise risk decision is particularly important! Strategic and operational management is regarded as constituting the core of the modern economy management by Western countries<sup>[4]</sup>. Risk management focus on the foundation and the development of enterprises, with the modern information technology, to make an accurate assessment and prediction of the enterprises future potential risks and avoid or reduce the loss<sup>[5]</sup>, which provide an important basis for risk decisions of the enterprises.

Western countries began to study risk prediction in World War II. Then Stones and Yates established the theoretical basis of risk decision<sup>[6]</sup>, the relationship between loss and profit optimization had become the core idea. Hall and Kutsch used expectancy theory in Enterprise Risk Management creativity, predicted the size of the expected benefits and risks<sup>[7]</sup>. In actuality, selection and optimization of samples, indicators and variables had been changed the prediction researches to the new direction<sup>[8]</sup>. In the continuous researches and practices, the effectiveness of risk prediction affect the risk management had become the consensus of the industry<sup>[9]</sup>. In recent years, with the data warehouse as the representative of computer application technology, we made considerable progress in the field of risk prediction and evaluation. However, for historical reasons, the research and application of risk prediction and evaluation is still in its infancy.

The above analysis shows that the index as an analysis of factors, while using the BP(Back-Propagation) neural network risk prediction and evaluation model is rare in the field of risk research. Based on these conditions, we treat the strategic and operational risk problem of the enterprise as the supporting point, after careful research, providing solutions to such problems.

## II. STRATEGIC /OPERATIONAL RISKS PREDICTION AND EVALUATION MODEL

At first, the ERM(enterprise-risk-management) Committee analyzed historical data of enterprise and set risk indicators and formula of the strategic/operational risks prediction and evaluation model. Then, according to the actual circumstances of an enterprise: TB-level data capacity, and distributed in various OLTP(On-Line Transaction Processing) database systems. Through the data warehouse conduct a series of ETL (extract, transform and load) operations. This model need to extract the risk information, the legal information, the financial information, the business information and other information of the enterprise. On this basis, the data warehouse conduct data preprocessed operations, to enhance data quality and ensure the correctness of the data mining results.

Data preprocessed operations include: removed the dirty data, mapped data, validated data, conversed data, merged data and so on, to achieve the purpose of the consistent data. The data warehouse conducted data preprocessed operations, shown in the figure 2-1:

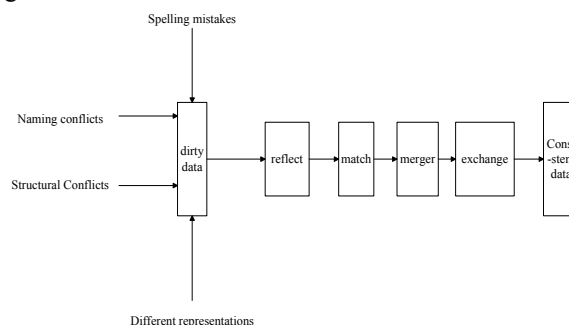


Fig.2-1 The data warehouse conducted data preprocessed operations

In the process of building data warehouse: at first, we need to create the functional metadata. Then based on the ETL and

preprocessed operations, to build the physical model tables. Through a series of processing for data-mining-width-table formed the fact table and supplementary table of the data mining. Finally, the results demonstrated by the data mining model. The process of data warehouse model, shown in the figure 2-2:

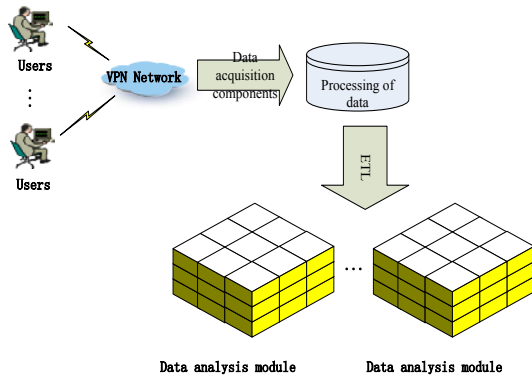


Fig.2-2 The process of data warehouse model

We mainly treated the risk indicators as factors in the field of prediction and evaluation of strategic risk research, in which the risk index value calculated by the risk formula of enterprise. Selected risk indicators are the main factor affecting the accuracy of risk prediction and evaluation, so based on business characteristics of the enterprise, risk indicators are divided into controllable and uncontrollable risk indicators.

After sorted, controllable risk indicators include: strategic management risk(U1), investment risk(U2), project management risk(U3), overseas investment risk(U4), human resource risk(U5); uncontrollable risk indicators include: price risk(U6), competitive risk(U7), industrial policy risk(U8), national policy risk(U9), country risk(U10), industry cycle risk(U11), technology investment ratio risk(U12).

Risk index value calculated by the risk index formula, for example: the value of strategic management risk(U1) is calculated by sales growth rate, return on equity, total asset turnover, the economy added rate these four index formula.

Strategic management risk(U1) index formula, shown in Table 3-1:

TABLE 3-1 STRATEGIC MANAGEMENT RISK(U1) INDEX FORMULA

Index (name)-	Index (formula)-
sales growth rate-	(The total current income - The total revenue over the same period last year) / The total revenue over the same period last year * 100%-
return on equity-	Net profit / Average net assets * 100%-
total asset turnover-	Net income from main operations / Average total assets-
(Frequency)-	
EVA rate-	EVA / Adjusted Capital-
EVA -	Net operating profit after tax - Cost of Capital = Net operating profit after tax - Adjusted Capital * The average cost of capital-
Adjusted Capital-	Average Equity + Average Total Liabilities - The average non-interest bearing current liabilities - Average in construction-

Similarly, in the field of operational risk prediction and evaluation research:

Controllable risk indicators include: accounts receivable and prepayments management risk(U1), inventory management risk(U2), budget risk(U3), financial reporting risk(U4), fixed assets and intangible assets management risk(U5), finance risk(U6), cash flow risk(U7), compliance risk(U8), product quality risk(U9), sales management risk(U10), governance risk(U11), purchasing and supply management risk(U12), production management risk(U13), equipment management risk(U14), public relations management risk(U15).

Uncontrollable risk indicators include: guarantee risk(U16), tax risk(U17), financial security risk(U18), maintenance of stability risk(U19), safety, health and environmental risk(U20), integrity risk(U21), reputational risk(U22), information risk(U23), legal risk(U24).

After collecting risk indicators, calculated by BP neural network algorithm. The algorithm referenced to the actual circumstances of an enterprise, by selecting samples, training samples, gaining the model weights, and shown the predicted results.

The model focuses on the strategic and operational risk of the enterprise, which using BP neural network algorithm to predict and evaluate the risk. The BP neural network algorithm is the classified analysis technology, the algorithm is divided into forward and reverse transmission. At first, we put the data to the input layer, through the N layer hidden nodes, transmit to the output layer. Second, we push back the results of the output layer to the input layer. During the time, adjusting the weights continuously to control the results of the input layer. At last, achieving the object of the error minimization. Because the characteristic of strategic and operational risk is nonlinear, so the node activation function is usually used S-type function, where Q is sigmoid parameter.

Equation 2-1:

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

In this equation, the output value is y, the adjacent nodes are i, j. Setting the output value of any node i is  $O_i$ , the number of samples  $(x_k, y_k)$  is N,  $(k=1,2,3,\dots,N)$ , the output result of the

layer i is  $O_{ik}$ , the input result of the node j is  $\sum_i W_{ij} O_{ik}$ , the weight is  $W_{ij}$ .

The error function, Equation 2-2:

$$E = \frac{1}{2} \sum_{k=1}^N (y_k - \hat{y}_k)^2$$

The weight correction, Equation 2-3:

$$W_{ij} = W_{ij} - \mu \frac{\partial E}{\partial W_{ij}}, \mu > 0.$$

The training process of BP neural network algorithm:

(1) Corrected weights:

If the error can't satisfy the requirements of the accuracy,

then making the process of error back propagation from the output layer to the input layer, correcting the weights layer by layer.

(2) Trained samples:

Randomly selecting from a sample of the overall samples, provided to the network, repeat until all samples completely.

(3) Controlled error:

Selected a sample and retrained until the error precision requirements in the end.

Since the data warehouse technology used later in our country, therefore, the analysis used the last decade data. In this model, we will issue a set of sample input-output problem to become a nonlinear optimization problem. We plan to select the initial weights, continue to modify the weights to achieve the results convergence.

Finally, according to results, the ERM committee recognized or the ERM committee recognized or fine-tuned the weights of model.

III. STRATEGIC/OPERATIONAL RISKS PREDICTION AND EVALUATION ALGORITHM FOR SOLVING THE MODEL

Prediction is scientific, timeliness, accuracy and other characteristics, which is the most important for the enterprise. However, the predicted characteristic often not been properly reflected in the enterprise. The large enterprises are the pillar industry of China, which should strengthen their risk resilience and increase the level of risk management. Therefore, risk predicted and evaluated fields have achieved good results by scientific prediction and used combination of quantitative and qualitative analysis, statistical analysis and calculation.

The algorithm is divided into forward and reverse transmission. In the forward process, the weights from the input layer, the hidden layer after layer by layer processed and transmitted to the output layer. The neurons in each layer only affect the next neurons. If the results of the output layer aren't expected, then transferred back propagation. Through the error signal is minimized by modified the each layer.

Setting the parameters of the algorithm:

(1) Setting the number of nodes:

The number of the input nodes is controlled/uncontrolled risk indicators of the enterprise. Generally, the number of hidden layer nodes is 75% of the input layer, the number of the output layer is 1 in this model. The number of hidden nodes settings will be also affected the accuracy of the prediction and evaluation, so we fluctuate 1 to 2 nodes in this model.

(2) Setting the initial weights:

The initial weights are set a group of unequal values generally. The random data set -0.5 to 0.5.

(3) Setting the minimum training rate:

The training rate is generally determined by the experience, the greater training rate, the greater weights

changed, the faster converged. But the training rate is too higher, it will cause the system oscillation. So without oscillations, the training rate the bigger is the better. The training rate set 0.9.

(4) Setting the dynamic parameters:

The dynamic parameters are generally determined by the experience. The dynamic parameters set 0.6 to 0.8.

(5) Setting the permissible error:

The permissible error set 0.00001 to 0.001.

(6) Setting the number of iterations:

The number of iterations is 1000.

(7) Setting the sigmoid parameter:

Sigmoid is the neuronal excitation parameters. The sigmoid parameter is 0.9.

IV. MODEL AND ALGORITHM VERIFICATION AND ANALYSIS

The strategic/operational risks prediction and evaluation model has been used in the risk decision support system of the enterprise. Risk control department of the enterprise consolidate and summarize the risk issues and divide into controllable and uncontrollable risk indicators. Through Training a series of processes, controlling the error precision and comparing the experimental data with real data constantly, in order to achieve the purpose of the strategic/operational risks prediction and evaluation for the enterprise.

Comparative actual and predicted values:

By comparing the actual and the predicted values of the strategic and operational risks in 2012, shown that the trend of actual and predicted lines are the same roughly. So the model can predict and evaluate the risk trend of the enterprise in a year generally.

The enterprise could focus on high-risks month, shown in the figure 5-1:

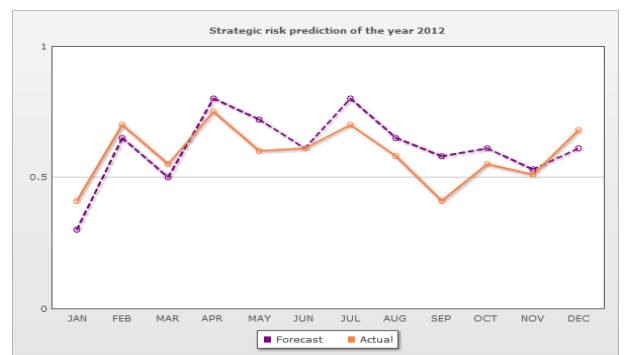


Fig.5-1 Comparative actual and predicted values(strategic risk)

The actual strategic risk value is based on the actual situation of enterprise risk criteria by the ERM committee calculated and summarized. The actual strategic risk value is between 0 and 1. The data are larger, the representatives of risk

factors are the higher. The predicted value is calculated by the algorithm model.

Comparative actual and predicted values(strategic risk), shown in Table 5-1:

TABLE 5-1 COMPARATIVE ACTUAL AND PREDICTED VALUES(STRATEGIC RISK)

month	actual	predicted
1	0.41	0.30
2	0.69	0.64
3	0.54	0.51
4	0.72	0.76
5	0.58	0.69
6	0.59	0.58
7	0.66	0.78
8	0.56	0.63
9	0.40	0.57
10	0.55	0.59
11	0.52	0.53
12	0.68	0.57

Comparative actual and predicted values(operational risk), shown in the figure 5-2:

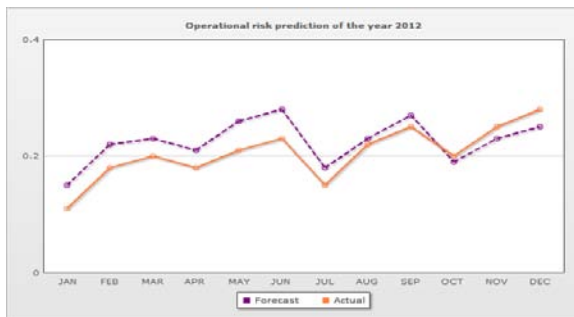


Fig.5-2 Comparative actual and predicted values(operational risk)

Comparative actual and predicted values(operational risk), shown in Table 5-2:

TABLE 5-1 COMPARATIVE ACTUAL AND PREDICTED VALUES(OPERATIONAL RISK)

month	actual	predicted
1	0.11	0.14
2	0.19	0.22
3	0.20	0.24
4	0.18	0.21
5	0.21	0.27
6	0.23	0.29
7	0.15	0.19
8	0.22	0.23
9	0.25	0.28
10	0.20	0.19
11	0.24	0.22
12	0.26	0.23

The actual operational risk value is based on the actual situation of enterprise risk criteria by the ERM committee calculated and summarized. The actual operational risk value is between 0 and 1. The data are larger, the representatives of risk factors are the higher. The predicted value is calculated by the algorithm model.

By comparison of figures and tables shown that the model could predict and evaluate the risk trend of the enterprise in the year generally. The model could play the important role in early warning and support for risk decisions better.

## V. CONCLUSION

This article is based on the risk characteristics and the actual needs of an enterprise, used the BP neural network algorithm and designed the strategic/operational risks prediction and evaluation model. In the design process of the model, the risk issues and risk indicators are associated. At last, we verified and analyzed the availability of the model. The algorithm model can make a rough prediction on the general direction of the trend of the enterprise strategic/operational risk. The significance of the model that realized the role of enterprise risk warning and increased risk control ability of the enterprise to better assist risk decisions.

## References

- [1] Ohlson. financial ratios and the probabilistic prediction of bankruptcy, 1980. 2-8
- [2] Jorge Ponce. The quality of credit ratings: A two-sided market perspective. Economic Systems, 2012:294-306.
- [3] Yates, J.F. & Stone, E. R. Risk decision-making Behavior. Chichester: John Wiley & sons Ltd, 1992:1-25.
- [4] E.Kutsch, M.Hall. Intervening conditions on the management of project risk: Dealing with uncertainty in information technology projects. International Journal of Project Management. 2005(23):591-599.
- [5] Nadine Gatzert, Michael Martin. Quantifying credit and market risk under Solvency II: Standard approach versus internal model. Insurance: Mathematics and Economics, 2012:649-666.
- [6] Wang X. T. Self- framing of risking choice. Journal of Behavioral Decision Making, 2004(17):1-16.
- [7] José M. Gascueña, Elena Navarro, Antonio Fernández-Caballero. Model-driven engineering techniques for the development of multi-agent systems. Engineering Applications of Artificial Intelligence, 2012: 159-173.
- [8] Jiawei Han, Micheline Kamber. Data Mining: Concepts and Techniques[M]. Morgan Kaufmann, 2001. 23-26
- [9] Sean N. Ghazavi, Thunshun W. Liao. Medical data mining by fuzzy modeling with selected features. Artificial Intelligence in Medicine, 2008. 195-206