

A STUDY ON THE ANN-BASED CREDIT RISK PREDICTION MODEL AND ITS APPLICATION

Yan PENG¹ Xuyan TU²

¹College of Information Engineering, Capital Normal University, Beijing 100037, China

²College of Information Engineering, University of Science and Technology Beijing, 100083, China

Abstract: Credit risk early warning system must provide precise support information to the bank managers for decision-making. So the risk prediction model is critical and regarded as the core part of the whole risk early warning system. Compared with other traditional prediction models, the neural network model has the advantages of self-studying, self-organizing as well as self-adapting. This dissertation puts forward an ANN-based risk prediction model, which combines the Self-Organizing Map Neural Network and the Probabilistic Neural Network together. Furthermore, the improved iteration ways in constructing and training the model are also presented, which includes the SOM boundary effect processing and the rare samples handling. The established model is trained with financial ratios for a specific credit risk early warning experiment. The preliminary experimental result demonstrates that the SOM-PNN model performs better than some traditional ones in the rates of prediction accuracy and efficiency.

Key words: Credit Risk; ANN; Prediction Model

1. INTRODUCTION

The Risk Early Warning is a complicated process. It discovers and distinguishes the risk sources, risk scopes, risk levels and risk trends in advance through a variety of techniques, and then generates the corresponding risk warning signals. This process could provide support information to the management for decision-making, which could in turn avoid or reduce the risk of loss over time. Obviously, the risk

which could in turn avoid or reduce the risk of loss over time. Obviously, the risk prediction model is very important in the early warning system, which is regarded as the core part of risk early warning activity.

However, there are a few technical difficulties in modeling the risk early-warning system with the traditional statistics-based methods. First of all, various properties or different regional possibility would require different models. Secondly, the supporting data for the model is generally very difficult to acquire and some information may be incomplete. Finally, the credibility of the model has to be verified through a long period of time. Practically, most of the processes depend largely on the subjective decisions.

Different from the traditional methods, the Artificial Neural Network (ANN) solves certain problems by training the network. Training an ANN is to input a series of samples that can generate the ideal output, which enables the ANN to learn the fundamental principle behind the "solution". The ANN is analogous to the central nervous system of the living creatures. The ANN has the advantages of self-studying, self-organizing as well as self-adapting. It is suitable for handling fuzzy data and is capable of reapplying its evaluation specialty through the learned experience and knowledge, which ensures the objectivity of the predicted result.

The ANN system avoids the complexity and difficulty existing in the traditional analysis processes. The ANN model is established in a non-linear and natural way, and does not require extracting the exact linear relationship, which makes it convenient for analyzing and model generating.

The main contribution of this paper is presented, i.e., establishing a SOM-PNN neural network model in the risk early warning system as well as the improvement of constructing and training the prediction model. The remainder of this paper is outlined as the following: the SOM-PNN based risk prediction model is generated after analyzing the SOM and PNN characteristics; each critical link of the model training and constructing process is finalized; the generated model is trained with financial ratios for a credit risk early warning experiment; the benefits of using this model in risk early warning are discussed; finally, the conclusion is made with an outlook of future works.

2. SOM&PNN NEURAL NETWORK

In this section, the Self Organizing feature Map Neural Network and the Probabilistic Neural Network are briefly introduced.

2.1 Kohonen Self -Organizing Model

T. Kohonen presented a neural network model called Self-Organizing feature Map(SOM) in 1993[1].

The SOM model is made up of two neural layers. The input layer has the same number of neurons as that of the variables, and its function is simply to capture the information. The output layer is organized in a rectangular pattern with x rows and y columns, which is called 'the map' [2]. The SOM projects the multidimensional input space into the output space. Through this process, the variables presenting a similar value are arranged close to one another in the input pattern on the created map. In order to achieve this target, the SOM uses a competitive algorithm known as "winner takes all". The minimum distance determines the superior match neuron,

$$\|X - W_c\| = \min_{j=1}^p \|X - W_j\| \quad (1)$$

where W_c is the synaptic weights of C ,

$\|X - W_j\|$ is the Euclidean distance,

$$\|X - W_j\| = \sqrt{\sum_{i=1}^n (X_i - W_{ij})^2} \quad (2)$$

where X_i is the input signal, and W_{ij} is the synaptic weight between the input neuron i and the output neuron j [3].

In this way, the essential topology of the input space is projected into the output space. This paper utilizes the standard non-directional SOM algorithm.

2.2 Probabilistic Neural Network (PNN)

The Probabilistic Neural Network[4] was introduced in 1990 by Specht. The benefit of using the PNN is that it can represent the architecture of a network in the form of *probability density functions*. In addition, most of the training algorithms for the PNN are easy to use.

The PNN network usually consists of four layers. The first layer is the input layer, which simply distributes the input signals to the pattern units, and its transmission function is linear. The second layer is the pattern unit layer, and each pattern unit is connected to the input layer with its weight W_{ij} . The activation function is used to construct the estimations of the population density function. The third layer is the summation unit layer, which has a linear summation function. The summation layer has one neuron for each class. Each summation neuron dedicated to a single class sums the pattern layer neurons corresponding to the number of that particular class. The fourth layer is the output layer. The output neuron is a threshold discriminator, which determines the maximum value of the summation unit outputs corresponding to its inputs.

3. PREDICTION MODEL BASED ON SOM- PNN

3.1 SOM-PNN Early-Warning System

It is innovative and worthwhile to develop the SOM-PNN model application in the early warning system, since there are few precedents in this field. The architecture of the SOM-PNN Early-Warning System is illustrated in Fig.1.

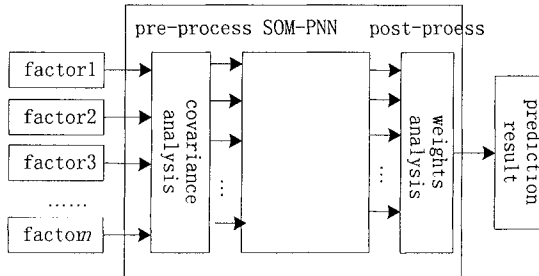


Fig.1 The Architecture of SOM-PNN Early Warning System

There are three primarily parts of the model:

- (1) Pre-process: the target of this part is to transform the risk data into the ratios that would be used as the network's inputs. Due to the limitation of the pages, this problem will not be addressed in detail.
- (2) SOM-PNN neural network: the core part of the model which addresses the prediction process.
- (3) Post-process: the task of this part is to analyse the predicting results. The weight analysis method is used to represent the influence of each factor.

The SOM model can yield high identification rate, however can't generate the classification result directly. The PNN is integrated with SOM to analyse the output of the SOM, then generates the classification result explicitly.

The SOM extracts the characteristic of the original data, i.e., 'compress' the data. Therefore this method increases the training efficiency of the PNN. The SOM is integrated with the PNN as the following [5]:

- (1) The SOM network is trained with the sample data first. Then the SOM MAP, which is acquitted through training and labelling, is used as the training sample of the PNN.
- (2) The testing sample set is used to test the whole network as shown in Fig.2.
- (3) The advantage of the integrated SOM-PNN model is that the trained SOM characteristic vector rejects the "noise" in the training samples. Therefore the precision and the efficiency of the PNN classification are increased significantly. Besides, the PNN generates the classification results explicitly.

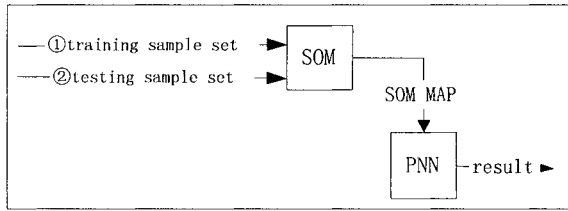


Fig.2 SOM- PNN model

3.2 Improvement of the Iteration Method

3.2.1 Handling the Rare Samples

In some cases, some risk data are either difficult to acquire or not under the estimated conditions. Then the specific input sample can not be trained sufficiently during the model training process. If the training frequency is very low, it can cause the classification result of the specific kind of input generating deviation, and decreasing the prediction precision. However, an increased overall training frequency could significantly increase the training time. Therefore, a specific method is used to identify the samples with less probability in order to improve the training process.

After the model has been trained for certain iterations, some input patterns innervate fewer neurons within the scope of its iteration radius in the preliminary MAP. Since the SOM MAP reflects the feature of the training samples, the following definition is introduced:

Definition: within the SOM MAP, if the number of innervated neurons in the iteration radius is less than the average value, this kind of sample is called the rare sample, i.e., the sample with the characteristic less needed than average. The expression is,

$$Z < \frac{S}{n \times m} \quad (3)$$

where S is the total sample number, $n \times m$ is the number of output neurons and Z is innervated neurons in the iteration radius scope of certain input sample A, .

Sample A is automatically selected in the program to strengthen the training frequency, so that this particular kind of samples can be trained sufficiently. In doing so, the model prediction precision can be improved and the training efficiency can be increased.

3.2.2 Processing the Boundary Effect

The Boundary Effect is an existing problem of SOM itself, because the density of the neurons at the boundary is higher than that in the other regions in MAP (as shown in Fig.3). Kohonen presented a solution in 1993, which assigns some special samples with special weights. The disadvantage of this method is that it needs some personal subjective experiences. Another kind of method is to increase the iteration frequency, which would increase the network training time.

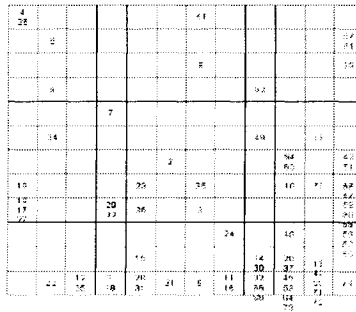


Fig.3 SOM MAP1

In this paper, the method of edge connection is proposed, which links the top and bottom edges as well as the left and right edges of the MAP. This makes the output similar to a sphere, thus every point in the MAP is adjacent to 8 neighbour points. Therefore, the boundary problem is eliminated, and the overall training efficiency of the network is improved.

4. CREDIT RISK PREDICTION EXPERIMENTS

The credit loan is the main income of a bank, which is also the most risky product[6]. In practice, the loans have been divided into *normal* and *concern* before they are granted. This kind of previous classification, in fact, is a risk analysis process. It provides the decision support for loaners. The Bank management can then use a series of methods to avoid or reduce the potential loan loss over time.

A particular characteristic of the credit early warning is that, the debtor's capability of repayment is the most important factor to be concerned about[7]. This paper primarily studies the early warning signals associated with certain financial situations, since the finance factors and its profit ability are closely related to the capabilities of making use of substance, capital structure as well as cash flows. This paper starts with the analysis of the financial ratios, and pays close attention to the debtor's current and future capability of repayment.

4.1 Sample Data

The sample data used in this paper are from 129 firms, among which 65 firms are bankrupt and the others are solvent. In order to satisfy both the modelling and verification practises, the sample data are randomly divided into two groups. The first group data are used for model training, which are from 74 firms and among which 36 firms are solvent and 38 firms are bankrupt. The second group data are used for model testing, which are from 55 firms and among which 28 firms are solvent and 27 firms are bankrupt.

The data contain four financial ratios taken from [8] for prediction experiments.

4.2 Constructing and Training of SOM and PNN

The SOM model has a two-dimensional topology, i.e., the model is made up of two neural layers. According to the finance ratios selected, the SOM network is constructed with 4 neurons at the input layer and 144 neurons at the output layer, and is arranged in a 12*12 square grid.

The training process of the SOM network is as the following:

The input financial ratios are located at the input layer once a single company is separated. Given that R_{ik} is the value of k company's ratio i and this ratio will be read by a neuron, the algorithm takes one neuron at a time at the output layer, and calculates the *Euclidean distance* as the similarity estimation. The "win neuron" is with the minimum $d(j, k)$ and is named neuron k^* . The algorithm now reduces the distance of $d(j, k^*)$ in order to change the Synaptic Weight W_{ij} (synaptic weights between the input neuron i and the output neuron j). This process is repeated until all trainings have been completed. The weights are fixed upon the training completion, then the network is ready to use.

The label method of the SOM Map is as the following:

Compare the *Euclidean distance* of the input vector $X(x_1, x_2, \dots, x_n)$ (in this case, it could be the financial information of any firm) with every neuron connection vector W_j , then take the nearest node and mark it for this firm. After carrying out certain SOM calculations, generate a 2- dimensional Map. Each node is now representing one firm.

The PNN training can be accomplished in a number of ways. Generally, two methods are frequently used. One approach relies on defining the network topology with a large training set in advance. Another approach is to introduce one neuron for each training pattern. Each class of the training samples is then used as a centre vector of the hidden nodes. However, this particular approach is not feasible for large datasets and small parameter changes could impact the network performance

significantly. The training method introduced in reference [9] is used in this paper, which automatically constructs the PNN network.

The SOM is trained initially, and then the MAP is derived (as shown in Fig.5). Two large zones can be recognized after this process. One corresponds to the solvent firms (1 to 36), and the other corresponds to the bankrupt firms (37 to 74) appeared with circular labels.

Upon the completion of the SOM, the PNN is then trained with each node of vector in Fig.4. Finally, the testing data set is used for the whole model testing.

The overall steps are as the following:

- (1) Import 55 testing samples into the SOM to form the input layer, then get the prediction result form the output layer of the PNN (*'1' indicates solvent and '0' indicates bankrupt*).
- (2) Use the PNN Model to predict solvent and bankrupt only.
- (3) Compare the results of the two models

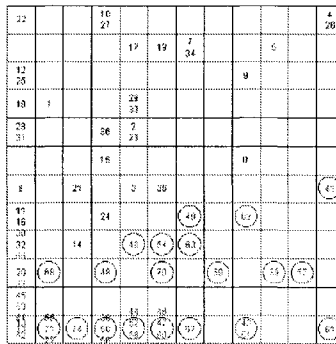


Fig.4 SOM MAP2

4.3 Predict Result Discussion

Table 1 compares the results with the SOM-PNN model and the PNN model only.

Table 1

No.	Samp le data	SOM- PNN Result	PNN Result	No.	Sample data	SOM- PNN Result	PNN Result
1	1	1	1	29	0	0	0
2	1	1	1	30	0	0	0
3	1	1	1	31	0	0	0
4	1	1	1	32	0	0	0
5	1	1	0	33	0	0	0
6	1	1	1	34	0	0	0

7	1	0	1	35	0	1	0
8	1	0	0	36	0	0	0
9	1	1	1	37	0	0	0
10	1	1	1	38	0	0	0
11	1	1	1	39	0	0	0
12	1	1	1	40	0	0	0
13	1	1	0	41	0	0	0
14	1	1	1	42	0	0	0
15	1	1	1	43	0	0	0
16	1	0	0	44	0	0	0
17	1	0	0	45	0	0	1
18	1	0	0	46	0	0	0
19	1	1	1	46	0	0	0
20	1	1	0	47	0	0	0
21	1	0	0	49	0	0	0
22	1	1	1	50	0	0	0
23	1	1	1	51	0	0	1
24	1	0	0	52	0	0	1
25	1	1	1	53	0	0	0
26	1	1	1	54	0	0	0
27	1	1	1	55	0	0	1
28	1	1	1				

The following conclusion can be presented from the experiments:

- (1) The precision by using the SOM- PNN model is 83.63%; and the precision by using the PNN model is 74.55%.
- (2) A visual result and a direct classification can be achieved by using the SOM-PNN model.
- (3) Reasons for the error classification:

Some bankrupt or solvent corporations are mistakenly classified in the experiment. The main reason is that the data used for classification in this experiment are raw finance data. However, the factors affecting a corporation’s condition could be well beyond this. Some corporation’s success or failure in the management could be determined by some special factors, such as a particular market condition, nature disasters and the termination of a certain manager, etc. These factors could all quickly change the condition of a corporation.

5. CONCLUSIONS

In this paper, a SOM-PNN model is introduced into the credit risk early warning system. The advantage of this model is that it combines the benefits of the SOM and the PNN, which in turn increases the accuracy and the efficiency of the risk

prediction. This model can be a useful tool for various kinds of applications, such as tax paying grade evaluation, customers grade evaluation, insurance, risk investment, etc. One potential improvement is to identify and synthesize the factors associated with the risk prediction, and makes this risk early warning model even more practical and accurate.

ACKNOWLEDGEMENT

This work is supported by Scientific Research Common Program of Beijing Municipal Commission of Education. Under contract NO. KM200410028012.

REFERENCES

- [1]. T. Kohonen, Automatic Formation of Topological Maps in Self-Organizing systems: Proceedings of the 2nd Scandinavian Conf. On Image Analysis, 1981:214 ~220
- [2]. Carpenter Gail A, Grossberg Stephen, The ART of Adaptive Pattern Recognition Self-Organizing Neural Network. Trans. IEEE on Computer, March 1988, 77~88.
- [3]. T. Kohonen, The Self Organizing Map, Proc. of the IEEE, (1990), Vol 78, 9; 1464-1480
- [4]. Sprech D F. Probabilistic Neural Network for Classification, Mapping and Associative Memory. IEEE ICNN San Dieg CA, 1988. I 525~532
- [5]. F. Ma, W. Wang, W.W. Tsang, Z. Tang, S. Xia and X. Tong, Probabilistic segmentation of volume data for visualization using SOM-PNN classifier, *Proceedings of the 1998 Symposium on Volume Visualization*, Pages 71-77
- [6]. Xi-wen Li, Zh-hong Zhao, Jian Wu. Establish the Risk Early-Warning System of Commercial Bank. Chinese financial, 2001,5: 37-38.
- [7]. Wen-Bo Pan. Research Papers for Commercial Bank Risk Classification. The Investment Studies, 2001,1:22-26
- [8]. Serrano, C. Self Organizing Neural Networks for Financial Diagnosis, Decision Support Systems, Vol 17, julio, 1996, 227-238,
- [9] Michael R. Berthold. Constructive Training of Probabilistic Neural Networks, *Neurocomputing* 19(1998) 167-183.