ICEB 2008 Proceedings

International Conference on Electronic Business (ICEB)

Fall 9-30-2008

# A Visual Map to Identify High Risk Banks - A Data Mining Application

Melody Y. Kiang

Robert T. Chi

# A Visual Map to Identify High Risk Banks - A Data Mining Application

Melody Y. Kiang, California State University, Long Beach, mkiang@csulb.edu
Robert T. Chi, California State University, Long Beach, rchi@csulb.edu

## Abstract

The increasing numbers of commercial bank failures have evolved into an economic crisis that has received much attention in recent years. The economic aftermath of large-scale bank failures is devastating. It triggers a domino effect that ripples across different sectors of the economy. It is therefore both desirable and warranted to explore new techniques and to provide early warnings to regulatory agencies. Regulatory agencies use a number of instruments to monitor their member banks. Those include quarterly and yearly reports, on-site examinations, and surveillance systems. The last instrument is of primary interest to our study. A surveillance system is an off-line system that identifies high-risk banks on the basis of their financial status. It provides early warnings to an agency to draw its attention to those banks that have a high likelihood of failure in the next one to two years. Numerous predictive models have been developed to identify problem banks for this purpose. Those models are closely related to research in financial distress spearheaded by Altman [1]. On the basis of Altman's framework, the detection of financial difficulties is a subject that has been particularly amenable to analysis with financial ratios. A bank bankruptcy prediction model should be able to make simultaneous consideration of those ratios that bear on the problem bank status.

Predictive models based on statistical discriminant techniques as well as neural networks have been widely used in the study of financial distress [4] [5] [6]. Although some techniques perform reasonably well in terms of rate of correctness in their predictions, all methods require complex and advanced analytical skill to explain and interpret the output from the prediction model. In this study, we introduce a special type of neural networks, the Self-Organizing Map (SOM) Network that can learn from complex, multi-dimensional data and transform them into visually decipherable clusters on an output map. The visual map can help an unsophisticated user to identify high-risk banks based on their locations on the map.

The Self-Organizing Map (SOM) network is a categorization network developed by Kohonen [3]. It was originally designed for solving problems that involve tasks, such as clustering, visualization, and abstraction. The main function of SOM networks is to map the input data from an n-dimensional space to a lower dimensional (usually one or two-dimensional) plot while maintaining the original topological relations. The physical locations of points on the map show the relative similarities between the points in the multi-dimensional space. In this research, we applied SOM to plot Texas bank bankruptcy data on a two-dimensional SOM map. The map can serve as an early warning system to help decision makers separate high risk banks from low risk ones.

The data used in this study is the same data set used in Tam and Kiang [6] that includes all banks failed in Texas during the period January 1985 through December 1987. The sample consists of 81 banks that failed between January 1985 and December 1987 and 81 matching non-failed banks, all from the state of Texas. Each bank is described by 19 ratios derived from its financial statements. Data were collected one year before the date of failure to see how well the model is capable of providing early warnings.

In this study, we first used a two-dimensional map to capture the relationships among the 162 Texas banks. The output map of SOM provides a graphical interface to help decision makers to visualize the differences in financial heath of the banks thus reduce the task from a multi-dimensional problem to a two-dimensional map. On the output map, each node may represent zero to many input data. The input data that are similar in higher dimension should be close to each other on the output map. One common approach implemented in previous studies was to consider each node on the output map as one cluster. However, a Kohonen network can have many nodes in the output layer; for example, a network of size 10x10 will have a total of 100 nodes in the output layer. When the number of nodes on the map is more than the number of clusters we desire, additional procedure to further group the nodes into fewer number of clusters is required. One way to do it is to manually group the points into clusters. Often

times it is hard to visually group the output from SOM especially when the map is highly populated. Hence, a more scientific approach that can help user to group the output from SOM network based on certain objective criteria is needed. In this study, we applied the extended SOM network method developed by [2] to automate the segmentation process to complement the usage of the Kohonen SOM networks. The method groups the output from SOM based on a *minimal variance criterion* to merge the neighboring nodes together. We then compared the result from SOM with that of discriminant analysis and logit models. The results show that the SOM method generates distinct groups as good as, if not better than, that of the discriminant analysis and logit approaches.

A possible future research is to use the map to determine cases to study for understanding bankruptcy. For example, we can study the banks that are close on the map or even fell on the same node but have opposite outcomes, one failed and one survived, a year after. The map tells us that the two banks were in very similar financial status a year prior to one bank went bankruptcy. We can run SOM model again to generate a new map using the new financial data and see if the two banks are now apart from each other. If the two banks still fell close to each other, it may suggest that the other bank maybe on the verge of bankruptcy. If the two banks have moved away from each other, we might want to take a closer look at the management of the two banks to study the reasons that make the two banks to have such different outcomes. In summary, SOM network is a valuable decision support tool that helps the decision maker visualizes the relationships among inputs and can be used as complementing tool with other classification methods such as discriminant analysis and logit models for data mining.

**Keywords:** SOM Kohonen Networks, bank bankruptcy prediction, clustering analysis

## References

[1] Altman, E., "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *Journal of Finance*, Vol. 4 No.4, 1968, pp.589-609.
[2] Kiang, M.Y., “Extending the Kohonen Self-Organizing Map Networks for Clustering Analysis,” *Journal of Computational Statistics and Data Analysis (CSDA)*, 38(2001) pp. 161-180.
[3] Kohonen, T., *Self-Organizing Maps*, Springer, 1995.
[4] Meyer, Paul A. and Howard W Pifer, “Prediction of Bank Failures,” *The Journal of Finance*, Vol. 25, No. 4 (Sep., 1970), pp. 853-868.
[5] Santomero, Anthony M. and Joseph D. Vinso, “Estimating the probability of failure for commercial banks and the banking system,” Journal of Banking & Finance, Vol. 1, No. 2, Oct. 1977, pp. 185-205.
[6] Tam, K.Y. and M. Y. Kiang, “Managerial Applications of Neural Networks: The Case of Bank Failure Predictions,” Management Science, Vol. 38, No. 7, pp. 926-947, July 1992.