49 Neural Methods for Analyzing Financial Information: How to Find the Enterprises with High Bankruptcy Risk?

Kimmo Kiviluoto, Erkki Oja and Jyrki Maaranen

Assessing the probability of bankruptcy of an enterprise is one of the key issues in a credit granting decision. Besides analyzing the strategy, personnel etc. of the firm, the financiers usually perform an analysis of the financial statements using some mathematical model. The standard approach has been to use a model based on Linear Discriminant Analysis, but a wide variety of other statistical techniques have also been proposed. Recently, models utilizing neural networks have been introduced and compared with the "traditional" techniques.

The importance of the problem has made it something of a benchmark test for different models. Usually, in these tests the problem has been reduced to a classification of companies into healthy and non-healthy ones. There are two characteristics common to many of the reported studies: they are based on fairly small data sets, and the proportion of the bankrupt firms is much higher in the data than in the total population, from which the sample is selected. This makes the results somewhat difficult to interpret – with small data sets, especially when the results are not cross-validated, the differences in classifier performance cannot be clearly distinguished from statistical noise, and with biased sample, one may also get an over-optimistic view of the classifier performance on the total population.

Another aspect is trying to analyze and understand the bankruptcy phenomenon: which factors contribute to an increased bankruptcy risk, or how does an increased risk of bankruptcy manifest itself?

The present study is conducted in co-operation with Finnvera Ltd., a service company that specializes in financing and development of small and medium-sized enterprises in Finland. The material consists of a certain segment of Finnvera Ltd.'s customer companies.

The study consists of two parts: qualitative analysis and classification. In the qualitative phase, the financial statements are analyzed with the Self-Organizing Map (SOM), which forms a "non-linear regression" from the input space into a plane. This makes it possible to visually examine the differences between firms that go bankrupt and those that do not (see figure 107) [1, 2, 3, 4, 5, 6]. New developments of the SOM are also discussed here, including three-dimensional SOM's, and a hierarchical model to analyze the year-to-year trajectories of an enterprise on the SOM [7, 8, 9, 10].

The classification of companies into healthy and non-healthy ones is done in two different ways: trying to minimize the total number of misclassifications, and using the Neyman-Pearson criterion, i.e. fixing the type I error (classifying a bankrupt company erroneously as a healthy company) to a suitable value, and with this constraint minimizing the type II error (classifying a healthy company erroneously as a bankrupt company). In practice, a classifier that is based on the Neyman-Pearson criterion would be the preferred one: type I error is much more costly than type II error, but because the proportion of non-bankrupt companies is higher, a classifier

that minimizes the total number of misclassifications would pay more attention on minimizing the type II errors.
The classifiers used in the quantitative study are the following: Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), k-Nearest-Neighbour Classifier (kNN), Learning Vector Quantization (LVQ), Self-Organizing Map (SOM), and SOM-based Radial Basis Function Network (RBF-SOM). The modification of the LVQ algorithm to incorporate the Neyman-Pearson criterion is an original contribution of this study; the other methods are used in a fairly standard manner. The classification results are presented in tables 15 and 16 [1, 2, 6].
The results show that the Self-Organizing Map is a valuable tool in the qualitative analysis of the financial statement data. In classification, the LVQ and kNN classifiers performed best, when the aim was to minimize the total number of misclassifications. With the Neyman-Pearson criterion, the LDA classifier reached the level of the LVQ and kNN, with the SOM classifiers coming close to these.
In the future, the focus will shift from annually given financial statements to more frequent time series. A first step into this direction has been to analyze parallel sales time series with Independent Component Analysis (ICA) [11].

References

[1] K. Kiviluoto, "Analyzing financial statements with the self-organizing map," Master's thesis, Helsinki University of Technology, Espoo, Finland, May 1996.

[2] K. Kiviluoto and P. Bergius, "Analyzing financial statements with the self-organizing map," in *Proceedings of the workshop on self-organizing maps (WSOM'97)*, (Espoo, Finland), pp. 362–367, Neural Networks Research Centre, Helsinki University of Technology, June 1997.

[3] K. Kiviluoto and P. Bergius, *Visual Explorations in Finance using Self-Organising Maps*, ch. Maps for analysing failures of small and medium-sized enterprises. Springer, 1998.

[4] K. Kiviluoto and P. Bergius, "Exploring corporate bankruptcy with two-level self-organizing maps," in *Proceedings of the Fifth International Conference on Computational Finance*, (Boston, Massachusetts, USA), pp. 373–380, London Business School, Kluwer Academic Publishers, Dec. 1998.

[5] K. Kiviluoto and P. Bergius, "Two-level self-organizing maps for analysis of financial statements," in *Proceedings of the 1998 IEEE International Joint Conference on Neural Networks (IJCNN'98)*, vol. 1, (Piscataway, New Jersey, USA), pp. 189–192, IEEE Neural Networks Council, May 1998.

[6] K. Kiviluoto, "Predicting bankruptcies with the self-organizing map," *Neurocomputing*, vol. 21, pp. 191–201, 1998.

[7] K. Kiviluoto, "Topology preservation in self-organizing maps," in *Proceedings of the International Conference on Neural Networks (ICNN'96)*, vol. 1, (Piscataway, New Jersey, USA), pp. 294–299, IEEE Neural Networks Council, June 1996.

Table 15: Classification results using Neyman-Pearson criterion with different error I values (per cent), based on financial statements given 2 . . . 0 years before bankruptcy

Classifier	error I target	total error	st.dev.	errorI	st.dev.	error II	st.dev.
LDA	0,20	19,0	(1,6)	21,0	(6,2)	18,8	(2,3)
	0,25	15,7	(1,0)	25,7	(5,4)	14,6	(1,5)
	0,30	14,1	(1,0)	29,5	(5,1)	12,5	(1,4)
LVQ	0,20	20,5	(2,0)	21,9	(4,1)	20,3	(2,4)
	0,25	15,9	(0,8)	25,7	(5,4)	14,9	(1,6)
	0,30	14,3	(1,0)	30,3	(4,5)	12,5	(1,5)
RBF-SOM	0,20	18,3	(1,2)	20,7	(5,6)	18,1	(1,7)
	0,25	15,8	(0,8)	26,4	(6,1)	14,7	(1,3)
	0,30	13,5	(1,0)	30,5	(6,4)	11,7	(1,6)
SOM	0,20	20,1	(1,9)	19,9	(6,3)	20,1	(2,6)
	0,25	16,6	(1,3)	25,4	(6,7)	15,7	(2,0)
	0,30	14,8	(0,4)	30,4	(6,8)	13,2	(0,9)

[8] K. Kiviluoto and E. Oja, "S-map: A network with a simple self-organization algorithm for generative topographic mappings," in *Advances in Neural Information Processing Systems* (M. I. Jordan, M. J. Kearns, and S. A. Solla, eds.), vol. 10, (Cambridge, Massachusetts, USA), pp. 549–555, MIT Press, 1998.

[9] K. Kiviluoto and E. Oja, "Softmax-network and S-Map – models for density-generating topographic mappings," in *Proceedings of the 1998 IEEE International Joint Conference on Neural Networks (IJCNN'98)*, vol. 3, (Piscataway, New Jersey, USA), pp. 2268–2272, IEEE Neural Networks Council, May 1998.

[10] K. Kiviluoto, "Comparing 2D and 3D self-organizing maps in financial data visualization," in *Methodologies for the Conception, Design and Application of Soft Computing – Proceedings of the 5th International Conference on Soft Computing and Information/Intelligent Systems (IIZUKA'98). Iizuka, Fukuoka, Japan.* (T. Yamakawa and G. Matsumoto, eds.), vol. 1, (Singapore), pp. 68–71, World Scientific, Oct. 1998.

[11] K. Kiviluoto and E. Oja, "Independent component analysis for parallel financial time series," in *ICONIP'98 Proceedings. Kitakyushu, Japan.* (S. Usui and T. Omori, eds.), vol. 2, (Tokyo, Japan.), pp. 895–898, APNNA, JNNS., Ohmsha, Oct. 1998.

Table 16: Classification results when minimizing the total number of misclassifications (per cent)

Classifier	total error	error I	error II
LVQ (e)	9,0	55,5	4,3
LVQ (p)	8,6	65,2	2,7
kNN (k=15)	8,5	75,2	1,5
LDA	10,5	47,1	6,6
QDA	11,1	55,9	6,5

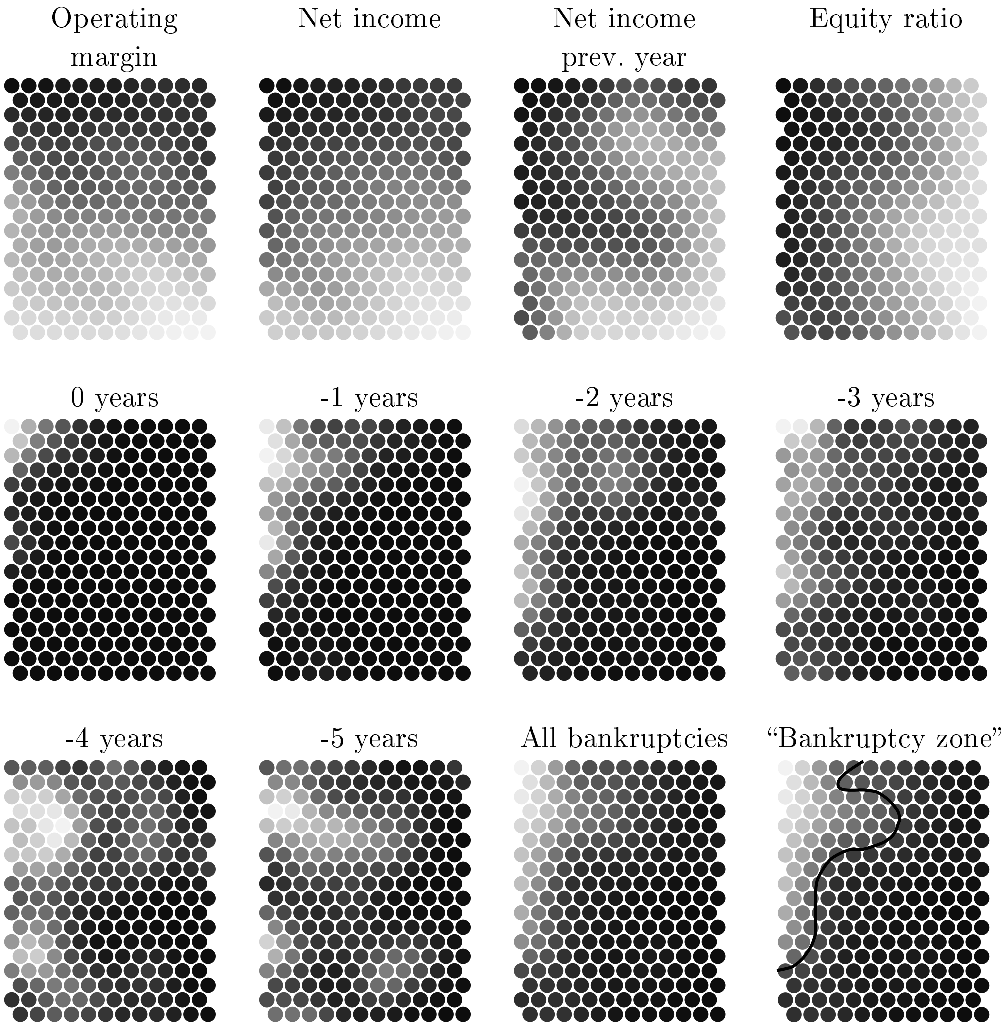

Figure 107: Financial indicators vs. bankruptcies, depicted 5 ... 0 years before the bankruptcy. On the upmost row, light color indicates good values of the financial indicators; on the two lower rows, light color indicates higher proportions of bankruptcy companies. In the lower right corner, a "bankruptcy zone" is drawn: more than one third of the companies that are projected on the left side of the line have gone bankrupt.