EXPLORING CORPORATE BANKRUPTCY WITH TWO-LEVEL SELF-ORGANIZING MAP

K. KIVILUOTO

Laboratory of Computer and Information Science
Helsinki University of Technology
P.O. Box 2200, FIN-02015 Espoo, Finland
E-mail: Kimmo.Kiviluoto@hut.fi

P. BERGIUS

Kera Ltd.
P.O. Box 559, FIN-33101 Tampere, Finland
E-mail: Pentti.Bergius@kera.fi

The Self-Organizing Map is used in the analysis of the financial statements, aiming at the extraction of models for corporate bankruptcy. Using data from one year only often seems to be insufficient, but straightforward methods that utilize data from several consecutive years typically suffer from problems with rule extraction and interpretation. We propose a combination of two Self-Organizing Maps in a hierarchy to solve the problem. The results obtained with our method are easy to interpret, and offer much more information of the state of the company than would be available if data from one year only were used. Using our method, three different types of corporate behaviour associated with high risk of bankruptcy can be recognized, together with some characteristic features of enterprises that have a very low bankruptcy risk.

1 Introduction

The Self-Organizing Map (SOM) of Kohonen (1995) has been found to be a valuable tool for analyzing financial statements. In a number of earlier studies (Martín-del-Brío and Serrano-Cinca 1993; Back et al. 1994; Shumsky and Yarovoy 1997; Kiviluoto and Bergius 1997), the SOM has been used for data visualization, and in some cases also for classification of companies into healthy and bankruptcy-prone ones.

A common feature to the studies referenced above is that they are based on data from financial statements given either for a single year or for two consecutive years. However, the practice that has long been preferred by the analysts of Kera Ltd., a Finnish financing company, is to use data from several consecutive years – it has been found that single year data is simply not enough to give a reliable idea of the state of an enterprise. A straigthforward application of this philosophy would be to concatenate financial ratios from several years into a single input vector, which then could be visualized using SOM. The problem with this approach is that the map thus obtained is difficult to interpret: there are no simple explications for the different areas of the map.

The solution that we propose in this paper is to proceed in two phases. First, the instantaneous or single-year state of the enterprise is found using the SOM in a standard manner (see below); the instantaneous state is hence encoded as a position on the map. Then, the change in the corporate state can be analyzed using another SOM that is trained with vectors obtained by concatenating the position vectors on the first SOM from several consecutive years. Now each point on this second-level SOM corresponds to a *trajectory* on the first-level SOM, while the first-level SOM is well-suited for further analysis because its "coordinate axes" turn out to have very natural interpretations.

2 Self-Organizing Map for "semi-supervised" learning

Basically, the SOM associates each unit of a (usually two-dimensional) map with a point in the data space, and the SOM algorithm is capable to find these "links" between the map units and the data space in such a manner, that map units located near each other have also their associated data space points near each other (see figure 1). The mapping from data space onto the map is defined as follows: a data point is mapped to that map unit which has its link closer to the data point than any other map unit.

The SOM algorithm is usually used in the context of unsupervised learning, in which there are no known target variables – the data is first mapped to the SOM and then analyzed using e.g. some of the visualization methods discussed by Kohonen (1995). However, here we have also an output variable: the possible bankruptcy of the company. The solution is to train the SOM in a "semi-supervised" manner, used e.g. by Heikkonen et al. (1993): only the information that can be found from the financial statements is used for determining the shape of the map, other attributes of interest are just carried along with the weight vectors so that they can be later used for analysis.

Specifically, denote the set of financial indicators with vector $\mathbf{x}^{(f)}$ and other attributes of interest, such as a binary bankruptcy indicator, with $\mathbf{x}^{(a)}$; concatenating these, vector $\mathbf{x} \equiv [\mathbf{x}^{(f)T}\ \mathbf{x}^{(a)T}]^T$ is obtained. The weight vector $\mathbf{m}_j$ associated with each map unit j correspondingly has two parts: $\mathbf{m}_j = [\mathbf{m}_j^{(f)T}\ \mathbf{m}_j^{(a)T}]^T$.

The learning algorithm consists of first finding the *best-matching unit* index c for the present data vector using the rule

$$c = \operatorname*{argmin}_j ||\mathbf{x}^{(f)} - \mathbf{m}_j^{(f)}|| \qquad (1)$$

and then updating weight vectors using the rule

$$\mathbf{m}_j := \mathbf{m}_j + \alpha h(j,c)(\mathbf{x} - \mathbf{m}_j), \quad \forall j \qquad (2)$$

where α is the learning rate, and $h(j,c)$ is a neighborhood function which here has the form of a Gaussian.

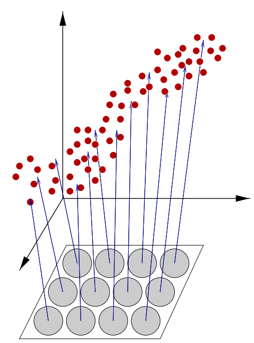

Figure 1: The SOM algorithm finds links between the data space and the map so that the topological relations are preserved as well as possible

The next phase is to teach a second-level SOM with data vectors obtained by concatenating annual positions of enterprises on the first-level SOM; this is shown schematically in figure 2.

Now the relative values of the first- or second-level SOM unit components are easily visualized as gray-level pictures for each component separately. In such a picture, the gray-level value of a map unit reflects the smoothed average value of the component in question of the data vectors mapped to that particular map unit or near it. For instance, if in such a picture some unit is almost black, the data vectors that are mapped to that map unit or to its neighbors have, on the average, very low relative values of the corresponding component.

3 Material

The material used in the present study consists of Finnish small and medium-sized enterprises, using the line of business, age, and size as the selection criteria. It was also required that the history and state of the enterprise be known well enough: if there were no data available for a longer period than two years before the bankruptcy, or if the last known financial statements were very poor,

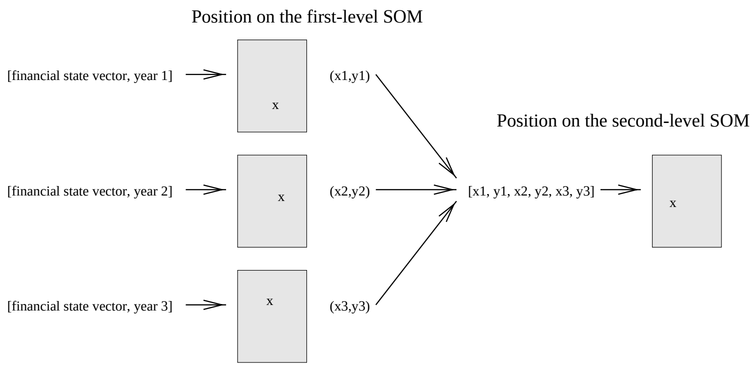

Figure 2: The second-level SOM input vectors consist of a company's position vectors on the first-level SOM during several consecutive years.

the company was rejected from the sample. However, in excess to these criteria, no data were rejected from the original population because it was "atypical", or looked like an outlier. In the final sample, there were 8 484 financial statements given by 2 116 companies, of which 568 have gone bankrupt.

The financial indicators used for training the first-level SOM are three commonly used ratios that measure the profitability and solidity of a company: (i) operating margin and rents, (ii) net income, and (iii) equity ratio. In addition to these, (iv) net income of the previous year is also included; together with the net income of the present year, it reflects the change in profitability.

Before training, the indicators were first preprocessed using histogram equalization separately for each indicator.

4 Results

The first-level SOM is shown in figure 3. In each subfigure, one component of the weight vectors associated with the map units is displayed. The gray-level coloring shows the relative values of that component in different parts of the map, black corresponding to "bad" values – either low values of financial ratios, or high proportion of bankruptcies. Note how the profitability generally increases when going downwards, while solidity generally increases to the right. The bankruptcies are concentrated in the upper left-hand corner of the map, where both profitability and solidity are low.

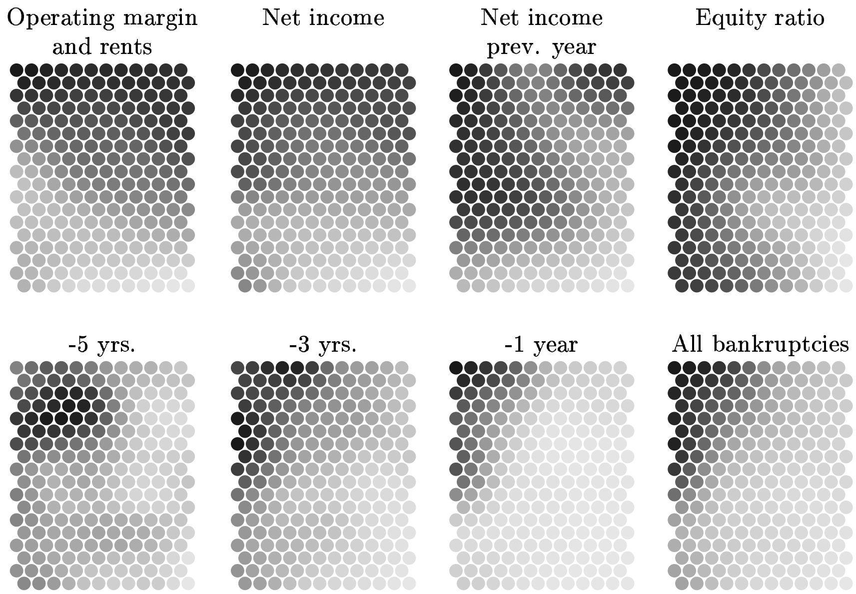

Figure 3: The first-level SOM. Shown are the relative values of the financial indicators, proportion of enterprises with varying number of years to bankruptcy, and proportion of all enterprises that went bankrupt within five years after giving the financial statement.

Examples of trajectories formed by an enterprise during several consecutive years are depicted in figure 4. Generally, the trajectories tend to move counterclockwise: a decrease in profitability, which shows as an upward movement, eventually results in a decrease in solidity, thus producing a leftward movement as well. Exceptions to this rule indicate abnormalities, such as sudden changes in the capital structure of the enterprise.

The second-level SOM is shown in figure 5, together with examples of the trajectories on the first-level SOM that correspond to different parts of the second-level SOM. There seem to be at least three different types of enterprise behavior that are associated with an increased risk of bankruptcy. The *first* of these is mapped to the upper right part of the second-level SOM: the enterprise is constantly in the low profitability, low solidity area. The *second* is in the upper middle part of the second-level SOM. These enterprises seem to be poorly

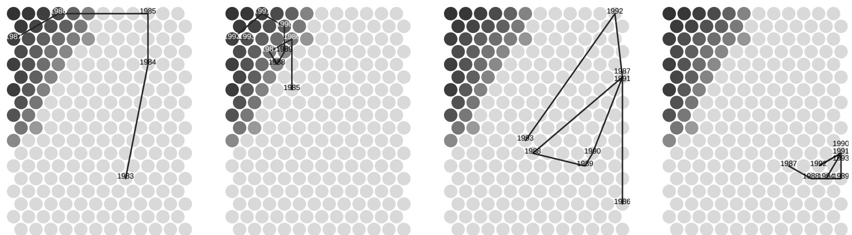

Figure 4: Trajectories of four companies on the first-level SOM – the two companies on the left went eventually bankrupt. The year the enterprise was mapped to each trajectory point is plotted next to the trajectory; the area with a high bankruptcy risk is marked with a (thresholded) darker shade.

controlled by the management, and so they experience large changes – long jumps on the first-level SOM – from year to year. This kind of behaviour becomes risky especially when it coincides with a decrease in solidity.

The *third* type of enterprise behaviour that is associated with an increased risk of bankruptcy can be found in the bottom-right corner of the second-level SOM. Almost all of the bankruptcies that fall into this group took place during the recession of the early nineties, and a common trait of the collapsed corporations here was that they typically had large loans in foreign currencies. First, the devaluation of the Finnish currency caused the solidity of these enterprises to weaken even though their profitability was very good; then the sales plummeted with the overall economic slump, which proved fatal for those firms that could not react fast enough and were already plagued by a relatively low solidity.

On the other hand, in the middle and on the lower left of the second-level SOM, the bankruptcy risk is very low. The first-level SOM trajectories here can be characterized as either having very small changes from year to year, or, if there are longer jumps, letting none of these notably weaken the solidity.

5 Discussion

The two-level self-organizing map can be applied to the qualitative analysis of financial statements with success. Its main advantage is that it is a very effective way to visualize higher-dimensional input data: among other things, it can be used to detect regions of increased risk of failure, to view the time

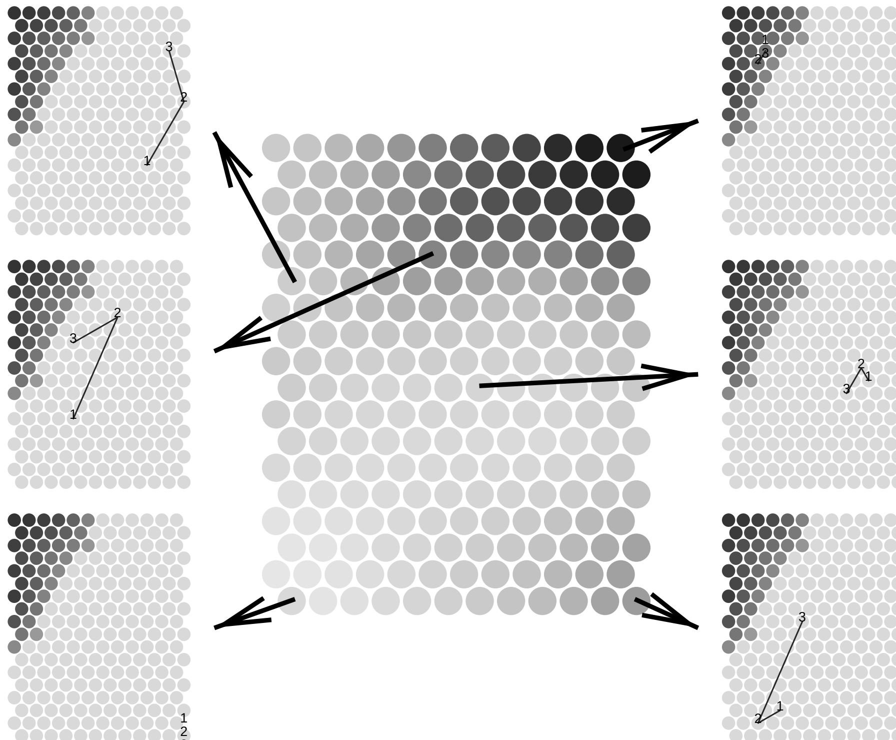

Figure 5: The second-level SOM and some first-level SOM trajectories that correspond to the indicated second-level SOM units.

evolution of the state of an enterprise, or to explore some typical corporate trajectories.

It seems plausible that this methodology could be extended also to quantitative analysis. The first step could be the classical "bankruptcy prediction problem", in which one tries to answer the question whether or not some company will go bankrupt, given its financial statements. This has become something of a benchmark problem and has been widely studied, so it could be used to get an idea of the performance level of the methodology introduced here; in the preliminary tests, the method has performed roughly at the same level as the popular Linear Discriminant Analysis (see Altman 1968).

Further steps in the quantitative analysis, and perhaps more interesting ones than the blind classification of bankruptcy prediction, might be estimating the probability of bankruptcy and testing different bankruptcy models or hypotheses. An interesting application of the latter would be contrasting the "failure trajectories" discussed above to those introduced by Argenti (1976); an initial examination hints that Argenti's bankruptcy trajectories might indeed have their representation also on the second-level SOM.

References

Altman, E. I. (1968, September). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *The Journal of Finance* (4), 589–609.

Argenti, J. (1976). *Corporate collapse – the causes and symptoms.* McGraw-Hill.

Back, B., G. Oosterom, K. Sere, and M. van Wezel (1994). A comparative study of neural networks in bankruptcy prediction. In *Multiple Paradigms for Artificial Intelligence (SteP94)*. Finnish Artificial Intelligence Society.

Heikkonen, J., P. Koikkalainen, and E. Oja (1993, July). Self-organizing maps for collision-free navigation. In *World Congress on Neural Networks*, Volume III of *International Neural Network Society Annual Meeting*, pp. 141–144. Neural Network Society.

Kiviluoto, K. and P. Bergius (1997, June). 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.

Kohonen, T. (1995). *Self-Organizing Maps.* Springer Series in Information Sciences 30. Berlin Heidelberg New York: Springer.

Martín-del-Brío, B. and C. Serrano-Cinca (1993). Self-organizing neural networks for the analysis and representation of data: Some financial cases. *Neural Computing & Applications 1*, 193–206.

Shumsky, S. A. and A. Yarovoy (1997, June). Neural network analysis of Russian banks. In *Proceedings of the workshop on self-organizing maps (WSOM'97)*, Espoo, Finland. Neural Networks Research Centre, Helsinki University of Technology.