Identifying Merger and Takeover Targets Using a Self-Organising Map

Emma Little
KPMG
Dublin
Ireland

Ruth Hickey
Ernst & Young
Dublin
Ireland

Anthony Brabazon
Natural Computing Research and Applications Group
University College Dublin
Ireland
Email: anthony.brabazon@ucd.ie
http://ncra.ucd.ie/

Abstract— **This study examines the potential of a self-organising map for classifying corporate takeover and merger targets using a range of financial information about companies. A sample of 200 US quoted companies, half which were merger or takeover targets and half which were not, are used to train and test the model. The best self-organised map correctly classified 94.8% of the firms in the training set one year prior to the takeover or merger event, and 95.2% in the out-of-sample validation set, averaged over five recuts of the data. The results provide support for a hypothesis that merger and takeover targets can be predicted, and that self-organised maps can be useful for this purpose.**

Keywords: Self organising map, prediction of takeover and merger targets.

I. Introduction

The objective of this study is to examines the potential of a self-organising map for classifying corporate takeover and merger targets using a range of financial information about companies.

The merger or acquisition of one company by another can take place in a number of ways [23]. The first involves the complete consolidation or merger of one company (the target) by another, known as the bidder or acquiring company. In the case of a merger the bidder takes over the entire stock of assets and liabilities of the target firm and the target no longer exists having been subsumed by the bidding firm's identity. With a consolidation, both firms join to form an entirely new entity, extinguishing both bidder and target's previous identities The second mechanism, often called a *tender offer*, involves the acquisition of voting shares in the target company by the bidding company. This can be done without consent from the target's management (a hostile takeover) as the bidder can make a tender offer directly to the shareholders of the target, offering to buy their shares.

A practical benefit of being accurately able to identify targets in advance of a merger or takeover bid, is that a premium is usually paid in order to acquire the shares in these companies at the time of merger / takeover. It is well-known that most gains which occur in mergers and takeovers accrue to the shareholders of the target firm [12]. Hence the ability to identify likely targets at an early stage could provide an investor with a trading edge and enable them to buy into potential target firms.

A. Prior Research

Due to the economic significance of being able to anticipate corporate mergers and acquisitions, the domain has attracted significant research attention over the past three decades. A wide a variety of methodologies have been applied in an attempt to uncover characteristics common to merger targets, and to forecast merger targets, including univariate analysis [24], MDA [4], [26], probit / logit analysis [21], [8], and multi-layer perceptrons (MLPs) [10]. The developed classification models have exhibited varying degrees of success ranging from below 50% to around 70% out of sample. Some of the best results were obtained by [23] who used binomial and multinomial models to predict merger targets. The resulting models produced a classification accuracy of 93%.

Most studies seeking to predict likely merger or takeover targets have relied heavily on the use of company accounting data, supplemented by market data such as share price, as modelling inputs. Such use of accounting data has a long providence in the related domains of corporate failure [2], [1] and prediction of corporate credit ratings [7].

B. Biologically Inspired Algorithms

Within the past decade, the array of computational technologies available to modellers has expanded considerably. Notably, a series of biologically-inspired methodologies have emerged and have been applied to a wide range of prediction and classification problems in business and finance [6]. These methods include MLPs, genetic algorithms [15], [20], genetic programming [19] and grammatical evolution [22]. As yet, few of these methodologies have been applied to the problem of merger / takeover target identification. This study applies a self-organizing map (SOM) for this purpose. To date one of the main applications of SOMs in finance has been in the area of corporate failure prediction [25], [16]. A number of additional business applications are described in [18].

No prior application of SOMs for the purposes of classifying companies as takeover / merger targets has been noted by the authors. This study aims therefore to contribute in two ways. First, it provides a description of the SOM in order to further disseminate knowledge of this methodology. Secondly, we examine the potential of a SOM for anticipating whether companies will be corporate merger and takeover targets.

C. Structure of paper

This contribution is organised as follows. Section 2 introduces the self-organising map (SOM). Section 3 describes the the data utilised. Section 4 provides the results of the constructed SOM models, followed by conclusions and suggestions for future work in Section 5.

II. Self-organising Maps

SOMs are a form of neural network which attempt to cluster data using an unsupervised learning algorithm [18]. As the SOM performs unsupervised learning, it is not necessary to associate an *output* with an input data vector. In certain applications, the outputs may not be known *a priori*. This may occur for example, when trying to segment a customer base. The clustering of the SOM serves to project (compress) the input data vectors into a low-dimensional space. Typically, the projection is onto a two-dimensional grid structure, thereby producing a visual representation of the (possibly) high-dimensional input data.

The SOM consists of two layers, the input layer (a holding point for the input data), and the *mapping* layer (see Figure 1). The input layer has as many nodes as there are input variables. The two layers are fully connected to each other and each of the nodes in the hidden layer has an associated weight vector, with one weight for each connection with the input layer.

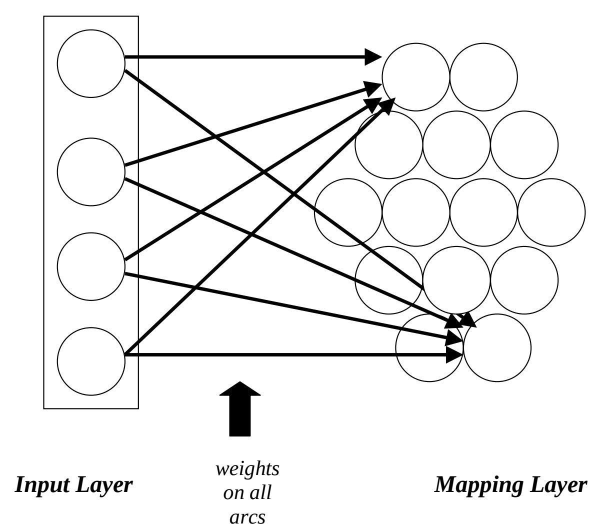

Fig. 1. A SOM with a 2-d mapping layer. On grounds of visual clarity, only the connections between the input layer and two of the mapping layer nodes are shown

The aim of the SOM is to group like input data vectors together on the mapping layer, therefore the method is *topology preserving* as items which are close in the input space are also close in the mapping space. During training the data vectors are presented to the SOM through the input layer one at a time. The nodes in the mapping layer *compete* for the input data vector. The winner is the mapping node whose vector of incoming connection weights most closely resembles the components of the input data vector. The winner has the values of its weight vector adjusted to move them towards the values of the input data vector, and the mapping layer nodes in the neighbourhood of the winning node also have their weight vectors altered to become more like the input data vector (a form of *co-operation* between the neighbouring nodes). As more input data vectors are passed through the network, the weight vectors of the mapping layer nodes will self-organise. By the end of the training process, different parts of the mapping layer will respond strongly to specific regions of input space. Once training of the network is complete, the clusters obtained can be examined in order to gain better insight into the underlying dataset (for example, what input items have been grouped together, what are the typical values for each input in a specific cluster).

A. Training an SOM

The general training algorithm for the SOM is as follows:

i. Initialise the weights between the input nodes and the mapping nodes.
ii. Present an input vector $\mathbf{x}$: $x_0, x_1, ..., x_{n-1}$.
iii. Calculate the distance between the input vector and the weight vector for each mapping layer node j

$$d_j = \sum_{i=0}^{n-1} (x_i - w_{ij})^2 \quad (1)$$

iv. Select the mapping node j^* that has the minimum value of d_j.
v. Update the weight vector for mapping node j^* and its neighbouring mapping nodes as follows

$$w(t+1)_{ij} = w(t)_{ij} + \eta(t)h(t)(x_i - w_{ij}) \quad (2)$$

where η is the learning rate of the map, and h defines a neighborhood function. Both the neighborhood size and the learning rate decay during the training run, in order to fine-tune the developing SOM.
vi. Repeat steps (ii)-(v) until the weights have stabilised.

In summary, SOMs can be considered as a non-linear, non-parametric regression technique that produces a topological, low-dimensional representation of data, which allows visualization of patterns / structures in the data.

III. Experimental Approach

The dataset consists of financial data on 200 quoted US firms, drawn from the period 1 January 2000 to 31 December 2002. All information was taken from the Compustat database. In order to reduce the problem of sectoral differences in financial data, the firms were all selected from the information technology sector (GICS Sector code of 45). Data was collected on 100 companies (group 1) who were successfully acquired (both merging firms and firms which were taken-over) between 2000 and 2002, and also on 100 firms (group 2) which were not merger or takeover targets during the same time period. Group 2 firms were matched to the group 1 firms based on GICS sector code and on year of acquisition of the merged firms, such that the group 2 firms were spread through the same sampling period as the merged firms. The dataset was recut five times to produce randomised training (150 companies) and test data (50 companies).

A. Selection of input variables

Based on an analysis of prior literature ([26], [5], [10], [8], [21], [24], [23], [11]) a total of seventeen variables were identified for initial evaluation. These variables were drawn from the following ratio categories.

i. Liquidity
ii. Debt
iii. Profitability
iv. Activity / Efficiency
v. Size
vi. Valuation
vii. Dividend payout
viii. Firm size

Following initial statistical tests which sought to eliminate variables which did not show a statistical difference in their mean across the two groups of companies, a total of ten ratios remained (see table I). Examining the cross-correlation coefficients of these ratios, two further ratios were removed (cash/total assets and the growth ratio) as they were found to be highly correlated with the working capital to total assets and the market/book value ratios respectively, and these two ratios were removed from the dataset.

B. SOM Model Construction

There are three main tasks in the development of a SOM classifier:

i. Training of the SOM
ii. Determining the clusters on the SOM
iii. Using the SOM to predict out-of-sample

Initially, as individual elements of the input data could have different magnitudes, each element was normalised. In constructing the SOM, several parameters of the map must be determined by the modeller, including the number of mapping layer nodes. The greater the number of nodes the greater the detail in the resulting map. Following a trial and error process, the size of the mapping layer was set at 1000 nodes. A variety of clustering algorithms can be used to create clusters on the feature map. We adopt a Ward-clustering algorithm. In this algorithm, each mapping node is initially designated as an individual cluster. At each step of the algorithm, the two nearest clusters are combined, thereby reducing the number of clusters by one at each step. The final number of clusters appearing on the feature map depends on the granularity required by the modeller and the choice of specific parameter values for the clustering algorithm. We selected these parameters in order to ensure that the number of final clusters was around 7. Following the training process the clusters were evaluated and classified or labelled as being Merged or Non-Merged by means of a simple voting mechanism using the state of the training data vectors in each cluster. In classifying the out of sample data vectors, each vector was assigned the class label of the cluster in which the closest mapping layer node resided.

IV. RESULTS

The results from our experiments are provided in Table II. An average classification accuracy in-sample (out-of-sample) of 94.80% (95.20)% is obtained. In order to assess the out-of-sample classification accuracy, Press's Q statistic (Hair, Anderson, Tatham and Black, 1998) was calculated, and the null hypothesis, that the out-of-sample classification accuracy is not significantly better than those that could occur by chance alone, was rejected at the 5% level.

Recut	Train (%)	Test (%)	Overall (%)
1	98.67	96.00	97.33
2	90.00	94.00	92.00
3	98.67	100.00	99.34
4	95.33	92.00	93.67
5	91.34	94.00	92.67
Average	94.80	95.20	95.00

TABLE II
CLASSIFICATION PERFORMANCE IN AND OUT OF SAMPLE FOR EACH OF THE FIVE RECUTS.

A sensitivity analysis for the results was undertaken by varying the number of nodes on the map (between 500 and 2000 nodes), and by also varying the number of clusters which the map was broken into (between five and nine clusters). The classification accuracies in and out-of-sample were not found to be sensitive to changes in parameter settings in these ranges.

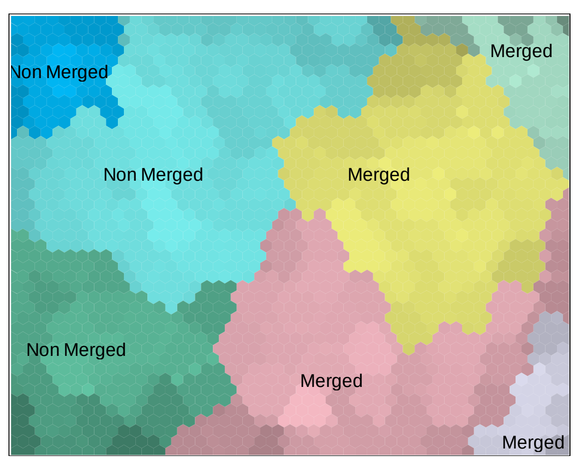

Fig. 2. A SOM mapping layer where the clusters are labelled merged vs non-merged.

	Merged	Non-merged	p-value
Working capital to Total assets	0.4682	0.3441	0.000
Working capital to Sales	1.1288	0.6090	0.021
Cash/Total assets	0.3961	0.3038	0.007
Long term debt / Total assets	0.0430	0.1008	0.000
Earnings before interest and tax / Sales	-0.2116	0.1279	0.000
Return on Equity	-26.304	11.148	0.024
Log(Net assets)	1.7317	3.0952	0.000
Net working capital	116.80	1353.19	0.000
Market price/book value per share	5.037	9.006	0.003
Growth ratio (3 year CAGR sales)	79.680	41.459	0.0594

TABLE I

MEANS OF RATIOS FOR MERGING VS NON-MERGING FIRMS.

In order to gain further insight into the nature of the clusters created by the SOM methodology, the 'typical' ratio values for firms in the various clusters of a number of the developed maps were examined (Figure 2 illustrates the final SOM corresponding to the third data recut). Table III provides the mean values of each model input for both the clusters corresponding to merged firms, and those corresponding to non-merged firms.

As might be expected given the high in and out of sample accuracy of the SOM on the third recut, the above cluster means are similar to those of the overall population. Based on both Table I and III, a few general comments can be made concerning the nature of 'typical' merger / takeover targets. Target firms tend to have higher levels of liquidity, lower debt levels, lower levels of profitability, and are smaller in size, than non-target firms.

V. CONCLUSIONS & FUTURE WORK

In this paper the objective was to assess the utility of an SOM methodology for classifying companies as merger or takeover targets, using a variety of financial information on the companies. The results indicate that SOMs can produce an accurate classifier for identifying targets, with accuracy levels in excess of 92% using data one year prior to merger / takeover. These results are highly competitive with results reported in prior literature.

Future work is suggested in several areas. First, a large dataset could be collected and analysed in order to further assess the generalisability of the findings of this study. In addition, the utility of SOMs for identification of likely target firms more than one year in advance could be examined.

REFERENCES

[1] Altman, E. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy, *Journal of Finance*, 23:589-609.

[2] Altman, E. (1993). *Corporate Financial Distress and Bankruptcy*, New York: John Wiley and Sons Inc.

[3] Barnes, P. (1990). The prediction of takeover targets in the U.K. by means of multiple discriminant analysis, *Journal of Business Finance & Accounting*, 17(1): 73-84.

[4] Barnes, P. (1998) Can takeover targets be identified by statistical techniques?: Some UK evidence, *The Statistician*, 47(4):573-591.

[5] Belkaoui, A. (1978). Financial Ratios as Predictors of Canadian Takeovers, *Journal of Business Finance & Accounting*, 5(5): 93-108.

[6] Brabazon, A. and O'Neill. M. (2006). *Biologically-inspired algorithms for financial modelling*, Berlin: Springer.

[7] Brabazon, A. and O'Neill, M. (2006). Credit Classification Using Grammatical Evolution, *Informatica*, (forthcoming).

[8] Castagna, A.D., & Matolcsy, Z.P. (1985) Accounting Ratios and Models of Takeover Target Screens: Some Empirical Evidence, *Australian Journal of Management*, 10(1):1-15.

[9] Clayton, R., & Fields, A., (1991) Prediction of acquisition candidates: Methodological Comparisons, *The Mid-Atlantic Journal of Business*, 27(3):233-250.

[10] Cheh, J.J., Weinberg, A., Yook, K.C., (1999) An Application Of an Artificial Neural Network Investment System To Predict Takeover Targets, *The Journal of Applied Business Research*, 15(4):33-44.

[11] Esphabodi, H., & Esphabodi, P. (2001). Predicting Corporate Takeovers, *The Journal of Business Forecasting Methods & Systems*, 20(3):25-28.

[12] Jensen, M., Ruback, R., (1983) The market for corporate control: the scientific evidence, *Journal of Financial Economics*, 11(1-4):5-50.

[13] Hair, J., Anderson, R., Tatham, R. and Black, W. (1998). *Multivariate Data Analysis*, Upper Saddle River, New Jersey: Prentice Hall.

[14] Harris, R., Stewart, J., Guilkey, D, & Carleton, W. (1982). Characteristics of Acquired Firms: Fixed and Random Coefficients Probit Analyses, *Southern Economic Journal*, 49(1):164-184.

[15] Holland, John H. (1975). *Adaptation in Natural and Artificial Systems*, Michigan: University of Michigan Press.

	Merged	Non-merged
Working capital to Total assets	0.470	0.348
Working capital to Sales	1.514	0.666
Long term debt / Total assets	0.0430	0.1030
Earnings before interest and tax / Sales	-0.602	0.126
Return on Equity	-83.25	12.00
Log(Net assets)	1.563	3.216
Net working capital	59.25	2041.0
Market price/book value per share	7.388	8.87

TABLE III

MEANS OF RATIOS FOR FIRMS IN MERGED CLUSTERS VS FIRMS IN NON-MERGING CLUSTERS, FOR DATA RECUT 3.

[16] Kiviluoto, K. and Bergius, P. (1998). Maps for analysing failures of small and medium-sized enterprises, *in Visual Explorations in Finance with self-organizing maps*, edited by Deboeck, G. and Kohonen, T., p. 59-71, Berlin: Springer-Verlag.

[17] Kohonen, T. (1982). Self-organized formation of topologically correct feature maps, *Biological Cybernetics*, 43:59-69.

[18] Kohonen, T. (1998). The SOM Methodology *in Visual Explorations in Finance with self-organizing maps*, edited by Deboeck, G. and Kohonen, T., p. 159-167, Berlin: Springer-Verlag.

[19] Koza, J. (1992). *Genetic Programming*, Massachusetts: MIT Press.

[20] Mitchell, M. (1996). *An Introduction to Genetic Algorithms*, Cambridge, Massachusetts: MIT Press.

[21] Meador, A., Church, P., & Rayburn, L., (1996). Development of Prediction Models for Horizontal and Vertical Mergers, *Journal of Financial And Strategic Decisions*, 9(1):11-23.

[22] O'Neill, M. and Ryan, C. (2003). *Grammatical Evolution: Evolutionary Automatic Programming in an Arbitrary Language*, Boston: Kluwer Academic Publishers.

[23] Powell, R. (1995). Takeover Prediction in the UK: A Multilogit Approach, *Working Papers in Accounting and Finance*, Accounting and Finance Division, The Queen's University of Belfast, 95(3).

[24] Rege, U., (1984). Accounting Ratios to Locate Take-Over Targets, *Journal of Business Finance and Accounting*, 11(3):301-311.

[25] Serrano-Cina, C. (1996). Self organizing neural networks for financial diagnosis, *Decision Support Systems*, 17:227-238.

[26] Stevens, D., (1973). Financial Characteristics of Merged Firms: A Multivariate Analysis, *Journal of Financial and Quantitative Analysis*, 8(2):149-159.