Transforming Passive Information from the Internet into Refined Information Using Self-Organising Maps

Jonas Karlsson[1], Tomas Eklund[1], Barbro Back[1], Hannu Vanharanta[2] & Ari Visa[3]

[1]Åbo Akademi University, Dept. of Information Systems, Turku, Finland

[2]Pori School of Technology and Economics, Pori, Finland

[3]Tampere University of Technology, Dept. of Information Technology, Finland

Jonas.Karlsson@abo.fi, Tomas.Eklund@abo.fi, Barbro.Back@abo.fi,

Hannu.Vanharanta@pori.tut.fi, Ari.Visa@tut.fi

Abstract. The Internet is a growing reservoir of mostly passive data and information. The challenge is to turn this data into knowledge. The aim of this paper is to show how passive information from the Internet can be transformed to refined information using self-organising maps. We illustrate the transformation using financial ratios from telecommunications and pulp and paper companies worldwide.

Keywords. Financial benchmarking, Self-organising maps, Neural networks, Financial performance, Telecommunication, Pulp and Paper

1 Introduction

There are many parties interested in the financial performance of a company. Investors want to find promising investments among the thousands of stocks available on the market today. Managers want to be able to compare the performance of the own company to that of others, in order to isolate areas in which the company could improve. Creditors want to analyse the company's long-term payment ability, and auditors want to assess the accuracy of a company's financial statements. Financial analysts want to compare the performance of a company to that of others, in order to find financial trends on the markets. A tool commonly used by these parties is financial competitor benchmarking. (Bendell, Boulter & Kelly, 1998)

Using ordinary spreadsheet programs, one can easily compare two to six companies at a time according to one ratio at a time. However, if one wants to obtain an overview of the competitors on the market, or want to take into account several ratios at the same time, spreadsheet programs are no longer of any use. Neural networks, in the form of self-organising maps, provide a new tool for clustering and visualisation of large amounts of information. This technique analyses the different characteristics of the input, and groups samples with similar characteristics together. In this report, we perform a financial benchmark for both telecommunications companies and pulp and paper companies using self-organising maps. Thus, in the financial comparison conducted in this report the self-organising map will analyse selected financial key ratios of companies, grouping companies with similar financial performance together.

The term self-organising map has become a very popular topic in today’s information technology society. Since its invention in 1981 over 4300 research papers have been written on the subject of self-organising maps (Kohonen, 2000). Some examples of more resent research papers include cloud classification (Ambroise, Seze, Badran & Thiria, 2000), image object classification (Becanovic, 2000), breast cancer diagnosis (Chen, Chang & Huang, 2000), classifying and clustering Internet traffic (Raivio, Riihijärvi & Mähönen, 2000), and extracting knowledge from text documents (Visa, Toivonen, Back & Vanharanta, 2000). Generally these reports can be divided into two groups, analyses and surveys. The analyses take a more technological approach towards the algorithm and function of the self-organising map, while surveys take the practical application of the self-organising map into consideration (Kaski, 1998).

However, although many papers on self-organizing maps have been published, very few studies have dealt with the use of self-organizing maps in financial benchmarking. An example of the application of neural networks for financial analysis is the study by Martín-del-Brío & Serrano-Cinca (1993). Martín-del-Brío et al. used self-organizing neural networks to study of the financial state of Spanish companies, and to predict bankruptcies among Spanish banks during the 1977-85 banking crisis.

This study builds upon previous studies by Back, Sere & Vanharanta (1997; 1998) and Back, Öström, Sere & Vanharanta (2000). In the study, Back et al. (1998) compared 120 companies in the international pulp and paper industry. The study was based on standardized financial statements for the years 1985-89. The companies used in the experiment were all based in one of three regions: North America, Northern Europe or Central Europe. The companies were clustered according to 9 different financial ratios: Operating profit, Profit after financial items, Return on Total Assets (ROTA), Return on Equity (ROE), Total Liabilities, Solidity, Current Ratio, Funds from Operations, and Investments. The ratios were chosen by interviewing a number of experts on which ratios they commonly used. The objective of the study was to investigate the potential of using self-organizing maps in the process of investigating large amounts financial data.

Back et al. (1997; 2000) are follow-up studies to the 1998 paper. The principle difference is that maps for the different years were trained separately in Back et al. (1998), while a single map was used in Back et al. (1997; 2000). Moreover, in Back et al. (2000) the data was from 1996-1997 and collected from the Internet. The results showed that a single map makes it easier to follow the companies' movements over years. The results of the studies also gave further evidence that self-organizing maps could be feasible tools for processing vast amounts of financial data.

The purpose of this study is to continue to assess the feasibility of using self-organizing maps for financial benchmarking purposes. In particular, in analysing the results, we will assess the discovered patterns by putting more emphasis on interpreting the results with existing domain knowledge. This paper is based on the findings of Eklund (2000) and Karlsson (2001).

The rest of the paper is organised as follows: Section 2 describes the methodology and the choice of financial ratios and companies. Section 3 presents the construction of the self-organising maps and Section 4 presents a detailed analysis of the maps. The conclusions of this paper are presented in Section 5.

2 Methodology

In this section we provide a description of the self-organising map and describe the choice of financial key ratios.

2.1 Self-Organising Maps

The self-organising map technique creates a two-dimensional map from the input data. This map resembles a landscape in which it is possible to identify borders that define different clusters (Kohonen, 1997). These clusters consist of input variables with similar characteristics, i.e. in this report of companies with similar financial performance.

The methodology used when applying the self-organising map is as follows (Back et al., 1998):

(1) Choose the data material. It is often advisable to pre-process the input data so that the learning task of the network becomes easier (Kohonen, 1997)

(2) Choose the *network topology*, *learning rate*, and *neighbourhood* width.

(3) Construct the network. The construction process takes place by showing the input data to the network iteratively using the same input vector many times, the so-called *training length*. The process ends when the *average quantisation error* is small enough.

(4) Choose the best map for further analysis. Identify the clusters using the *U*-matrix and interpret the clusters (give labels to them) using the feature *planes*. From the feature planes we can read per input variable per neuron the value of the variable associated with each neuron.

The network topology refers to the form of the lattice. There are two commonly used lattices, rectangular and hexagonal. In a rectangular lattice a node has four neighbours, while in a hexagonal lattice, it has six. This makes the hexagonal lattice preferable for visualization purposes (Kohonen, 1997). The learning rate refers to how much the winning input data vector affects the surrounding network. The neighbourhood width refers to how much of the surrounding network is affected. The average quantisation error indicates the average distance between the best matching units and the input data vectors. Generally speaking, a lower quantisation error indicates a better-trained map.

To visualise the final self-organising map we will use the *unified distance matrix method (U-matrix)*. The U-matrix method can be used to discover otherwise invisible relationships in a high-dimensional data space. It also makes it possible to classify data sets into clusters of similar values. The simplest U-matrix method is to calculate the distances between neighbouring neurons, and store them in a matrix, i.e. the output map, which then can be interpreted. If there are “walls” between the neurons, the neighbouring weights are distant, i.e. the values differ significantly. The distance values are also displayed in colour when the U-matrix is visualised. Hence, dark colours represent great distances while brighter colours indicate similarities amongst the neurons. (Ultsch, 1993)

By viewing the individual feature planes it is possible to visualise the values of a single vector column, i.e. in this research the maps for one financial key ratio. These feature planes can be analysed in order to discover how well the companies have been doing according to single financial ratios (Kohonen, 1996). Thus, with the feature planes, it is rather easy to see where the companies with good profitability are located on the map, and in the same fashion where the companies with poor profitability are located.

2.2 Choice of Companies

For both studies, we have selected the companies from five regions: Asia, Canada, Continental Europe, Northern Europe (the Nordic Countries) and the USA. Africa and South America were excluded due to lack of information. There is also an average of every region included as an additional “company”. The total number of telecommunications companies is 93 and the number of pulp and paper companies is 82. The averages will make a comparison between the different regions possible.

2.3 Choice of data and information

The starting point of this report was to use only the Internet as a source of financial data. Therefore, the data searching part of this research was executed by searching for financial statements on the homepages of the companies, as well as in different databases on the Internet.

Many of the companies did not have financial information for more than three years on their homepages. Therefore, in most cases, this was complemented with financial information from databases such as the U.S. Securities and Exchange Commission (http://www.sec.gov) for American companies, the System for Electronic Document Analysis and Retrieval (http://www.sedar.-com) for Canadian companies and Japan Financials (http://japanfinancials.com) for Japanese companies. Since no good database was found for the Continental European companies, the financial information has been complemented with annual reports received via regular mail.

2.4 Choice of Ratios

To conduct the financial benchmarking of the companies, the companies' financial statements have been used as the source of information. From these financial statements seven key ratios have been calculated for each company and used as input data when training the self-organising map. These key ratios will be briefly presented in the following section.

The selection of relevant key ratios was based on an empirical study by Lehtinen (1996) in which international accounting differences were analysed in more detail, especially concerning the reliability and validity of the ratios. Seven financial key ratios, which fulfilled the criteria of good validity and reliability, were selected and calculated for each of the companies. The key ratios can be divided into four different classes: *profitability ratios, liquidity ratios, solidity ratios* and *efficiency ratios*. In financial benchmarking it is common to choose ratios that measure different aspects of financial behaviour. In this financial benchmark, more emphasis was put on profitability since it can be regarded to be the driving force behind most public companies. Three profitability ratios were selected; *Operating Margin, Return on Total Assets and Return on Equity.* In the class liquidity only one ratio was selected, *Current Ratio* for telecommunications companies and *Quick Ratio* for pulp and paper companies. The solidity of a company was regarded to be nearly as important as profitability and therefore two ratios were selected, *Equity to Capital and Interest Coverage.* In the final class, efficiency, only one ratio was selected, the *Receivables Turnover* ratio.

3 Training the Maps

The software we will use when training and creating the self-organising map is called *The Self-Organising Map Program Package Version 3.1 (SOM_PAK),* and is based on the Kohonen self-organising algorithm. The software package has been developed by the SOM programming team at the Helsinki Univ. of Technology.

During the training process, several tests were carried out in order to determine suitable parameters. The hexagonal lattice type was preferred for the visualisation of the output map. Furthermore, the map ought to be rectangular, rather than square, in order to achieve a stable orientation in the data space (Kohonen, 1996). Commonly, the x-axis should be about 30 per cent greater than the y-axis, thus forming a rectangular output map. Another recommendation is that the training length of the second part should be at least 500 times the number of network units (Kohonen, 1997).

To ease the neural network's learning process and improve the quality of the map, the input was standardised. For example, if one of the selected key ratios has a range of 0 to 1, while another key ratio has a range of –100 to 100, the contribution of the second input will likely be given more weight than the first one. Because of this, it is essential to standardise the input data so that their value reflects their importance, or at least that the value is similar in relation to the other input data (Bishop, 1995). In this study, the standardisation has been done by scaling the input variables by the variance according to the following formula:

$$\tilde{x}_{in} = \frac{(x_{in} - \bar{x}_i)}{\sigma^2}.$$

During the training process the self-organising map was still placing too much weight on extreme values, even after standardisation. In order to receive an interpretable map, Johnson and Wichern (1997) suggest that the input data should be modified by limiting how great values the extreme observations were allowed to take. In this study the extreme values have been limited to –50 respectively 50.

The constructed map for each line of business was trained using input data for the years 1995-99, i.e. only one map was created and analysed per line of business. The reason creating only one map for the period 1995-99, and not one map for each year, is that now the same clusters appear for all years, and in the same places. If one map would be trained for each year, different clusters would probably appear, and would have to be analysed and interpreted separately.

The trained map for the telecommunications companies is of the size 9 x 6 neurons, while the map for the pulp and paper companies is of the size 7 x 5 neurons. On these maps (figures 2 and 3a), it is easy to define the different clusters by looking at the colour shades of the borders between the hexagons. The brighter colours of the hexagons imply similar characteristics, while darker colours represent greater distances. The coloured borders between the hexagons are of great value when trying to determine and interpret clusters. Furthermore, it is also possible to visualise company movements in an interpretable fashion on this size of map.

The software we have used to visualise the final constructed self-organising maps and the feature planes in this report is a program called *Nenet version 1.1a*. This software was developed by the Nenet team at the Helsinki University of Technology. Nenet is a user-friendly program designed to illustrate the use of self-organising maps, and provides an easy way to visualise the output maps with not only the U-matrix method but also the Interpolated 2D U-matrix method, and as parameter level maps.

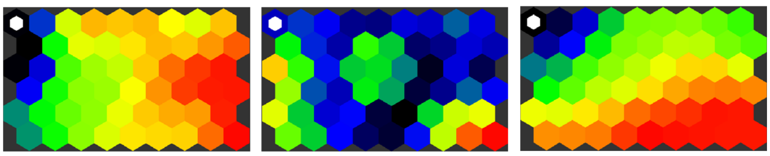

Figure 1. Examples of feature plane maps.

The feature planes in Figure 1 show a map for each of the financial key ratios, on which warmer colours, i.e. red, represent high values, which in our case imply good values, and darker colours show low values, which in our case implies poor values. A high value does not necessarily mean a good value, but it does in the case of our selected key ratios.

Several hundred maps were trained during the course of the experiment. The best maps, rated according to quantisation error and ease of readability, were then selected and used as a basis when training further maps. In order to achieve statistical accuracy (Kohonen, 1997), the initial phase in the pulp and paper study includes 1,750 steps and the final phase 17,500 steps. The learning rate factor was set to 0.5 in the first phase and 0.05 in the second, which are commonly used starting points. The neighbourhood radius was set to 12 for the first phase and 1.2 for the second. In the telecommunications study the initial phase includes 5,000 steps and the final phase 50,000 steps. The learning rate factor was set to 0.05 in the first phase and 0.02 in the second. These values may seem very small but provided the best results. The neighbourhood radius was set to 9 for the first phase and 1 for the second.

The initial network radius was very large in both studies, but seemed to provide for the overall best maps. Decreasing the radius only resulted in poorer maps. As Kohonen (1997) noted, the selection of parameters appears to make little difference in the outcome when training small maps. With different selections of parameters, the changes in the quantisation error were very small, usually as little as 0.001.

3.1 Defining the Clusters for the Telecommunications Companies

By analysing the output map more carefully, six major clusters of companies were identified. To identify the clusters we used both the U-matrix map and the feature planes. By analysing the colours of the borders between the hexagons, as well as the colour of the hexagon itself, it is possible to find similarities as well as differences. Furthermore, the values of the neurons have been evaluated in order to determine that the clusters are correct. The identified clusters are presented in Figure 2, in the form of a U-matrix map:

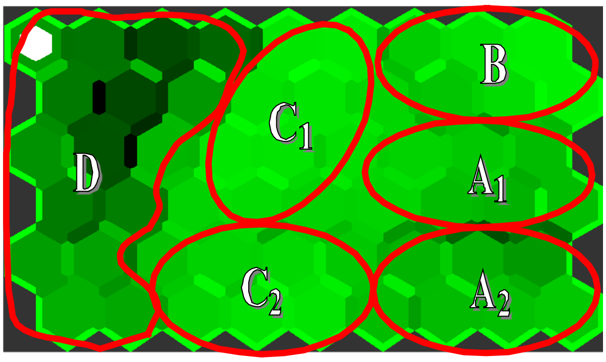

Figure 2. The identified clusters on the self-organised map.

In Figure 2 six different clusters of companies were identified. These clusters have been labelled: A_1, A_2, B, C_1, C_2 and D. The conclusions of the interpretations is as follows:

Group A_1 and Group A_2 represent the best in class companies. For the companies situated in subgroup A_1, profitability is very good, with very high values in the financial ratios Operating Margin, ROTA, and ROE. Solidity is decent, i.e. the values of the Equity to Capital ratio and the Interest Coverage ratio vary from good to average.

Group A_2 is the second subgroup of the best in class group. The companies situated in this group are characterised by slightly lower profitability than Group A_1, but instead liquidity and solidity are much better. These companies generally have the best values in Current Ratio on the map.

Group B is where the companies with slightly poorer performance than those in Group A_1 and A_2 are situated. These companies are distinguished by good profitability, and especially the ROE ratio is excellent. These companies also have somewhat poorer liquidity and solidity than the companies in Group A.

Group C_1 is the better of two subgroups in Group C. Here the companies possess decent profitability, good liquidity, and also good values in the Equity to Capital ratio.

Group C_2 is the slightly poorer of the two middle groups. These companies have decent profitability, but poor liquidity. Interest Coverage and Receivables Turnover are also poor, but Equity to Capital, on the other hand, is very good.

Group D is the poorest group. The companies with poor financial performance can be found in this group. Distinguishing features are commonly poor profitability and solidity. Liquidity is average and Receivables Turnover varies from very good to poor. Generally this group contains service providers from Europe and the USA, but also some Japanese companies, mostly for the years 1998-99.

3.2 Defining the Clusters for the Pulp and Paper Companies

By studying the final U-matrix map (Figure 3 a), and the underlying feature planes of the map, a number of clusters of companies, and the characteristics of these clusters, were identified (Figure 3 b).

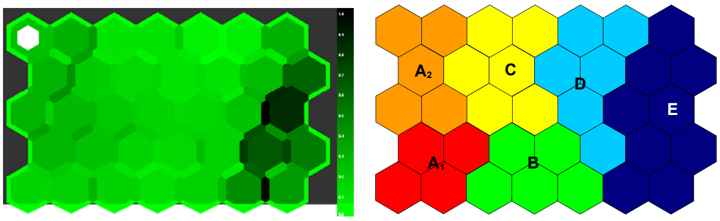

Figure 3. (a) The final U-matrix map and (b) identified clusters on the map.

The groups and their characteristics are presented below.

Group A consists of the best performing companies. Group A is divided into two subgroups: A_1 and A_2. The companies in subgroup A_1 are the best performing of all companies, especially according to profitability ratios. These companies have very high profitability, solidity, and efficiency, and medium liquidity. Subgroup A_2 consists of well performing companies with high profitability (especially in Return on Equity ratios), and average solidity and liquidity.

Group B is an average group, performing decently according to all ratios. The companies in Group B have low profitability but high solidity.

Group C can be classed as a slightly above average group. Group C has lower Equity to Capital ratios than Group B. However, Group C has higher profitability, notably in Return on Equity ratios. In addition, Group C contains the companies that have the highest liquidity. Group C has average to high profitability, average solidity, and very high liquidity.

Group D is best classed as slightly below average. The group has average solidity, but low profitability and very low liquidity. This group also contains the companies with the lowest efficiency ratios.

Group E is the poorest performing group. The group contains companies that are performing poorly according to almost all ratios, especially profitability ratios.

4 Benchmark Analysis of the Companies Over Time

In section 3, the clusters of companies on the self-organising output maps were identified and analysed. In this section a more detailed analysis is conducted concerning company movement during the years 1995-99. Furthermore, the competing companies in the specific markets will be benchmarked against each other. The companies on the analysed self-organising maps will be labelled numerically. Section 4.2 presents the top 5 companies in the telecommunications market and section 4.2 presents the top 5 companies in the pulp and paper market. In the original studies (Eklund, 2000; Karlsson, 2001) several benchmarks were performed, including country averages, best and worst performers, largest companies, regional benchmarks, and merger analysis. However, only the top 5 benchmarks are illustrated in this study.

4.1 Top 5 Telecommunications Companies

In Figure 4, a visualisation of the largest manufacturers in this study is presented. This benchmark is international, i.e. companies from different countries are benchmarked against each other. This means that differences in accounting practises might influence the resulting output map.

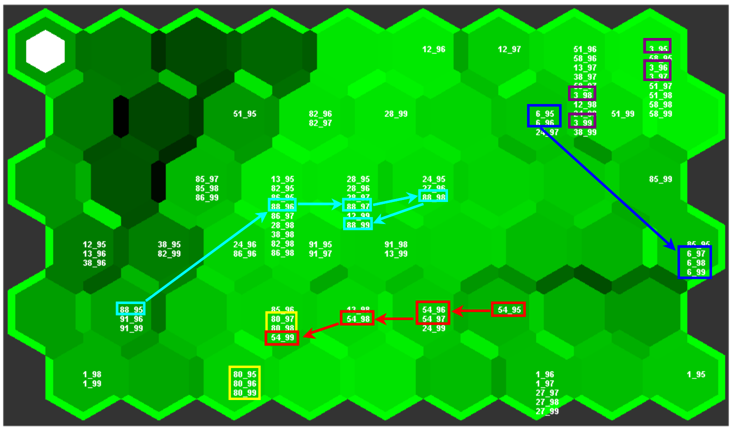

Figure 4. The movement over the years of largest manufacturers.

As Figure 4 shows, Nokia (No. 6) is the only mobile phone manufacturer that experiences a steady increase in financial performance. Nokia is situated in Group B during the years 1995-96, and took a leap into Group A_1 in 1997-99 (blue arrow). The reason for the success is a constant increase in all of the selected key ratios. Judging by the feature planes, Nokia was experiencing excellent profitability in 1999, and very good liquidity and solidity. Only the value of Receivables Turnover was somewhat lower than for the previous years.

Motorola (No. 54), on the other hand, shows an almost steady decrease in performance. In 1995 the company was situated in Group A_2, but has since then moved into the slightly poorer of the middle groups (red arrows). This is probably due to increasing competition on the telecommunications market. Examining the financial statements of Motorola reveals, for example, that net income has been decreasing steadily since 1995. Studying the selected key ratios shows that Motorola exhibits very good Equity to Capital but, on the other hand, its profitability has decreased during the last four years.

The third of the major mobile phone manufacturers, Ericsson (No. 3), is firmly situated in Group B during the years 1995-99 (purple squares). Ericsson is one of few of the selected companies that display high Receivables Turnover ratios. Furthermore, the Operating Margin and ROE ratios are excellent. The values in liquidity and solidity are slightly poorer.

Sony (No. 88), has experienced a slight improvement in their performance. In 1995, the company was situated in the poorest group, but eventually ended up in the slightly better middle group, C_1, in 1999 (turquoise arrows). What is interesting about Sony's performance is that they do not seem to experience any effects of the Asian crisis, except for a slight backtracking in 1999. One reason for why Sony is not experiencing any greater effects of the crisis could be the fact that they are a large international company, thus the Asian market comprises only a small part of their operations. Overall, Sony is performing much like the average Asian company.

Matsushita (No. 80) is the second largest manufacturer of electronic equipment in the world, more known for their brand Panasonic. Matsushita does not experience any greater effects from the Asian crisis, even though they backtrack slightly in 1999. Overall, Matsushita is situated in the same two neurons for all five years, in Group C2 (yellow squares). This company shows rather similar values in all of the selected key ratios, except for the Equity to Capital, they constantly show great values. The reason why Matsushita is not so much affected by the crisis is probably that they are a large, international company like Sony.

Analysing this map reveals that most of the manufacturing companies are situated either in the middle groups or in Group B. Only a few companies manage to place themselves in Group A_1 or A_2. Similarly, only a few companies have been placed in the poorest group. Most of the Asian companies are situated close to or in Group D.

4.2 Top 5 Pulp and Paper Companies

In the following figure (Figure 5), the Top 5 pulp and paper manufacturing companies according to Pulp and Paper International (Matussek, Janssens, Kenny & Rhiannon, 1999) are benchmarked against each other. The movements from year to year are illustrated with arrows, blue for International Paper, yellow for Stora Enso, orange for Kimberly-Clark, black for Oji Paper, and red for UPM-Kymmene.

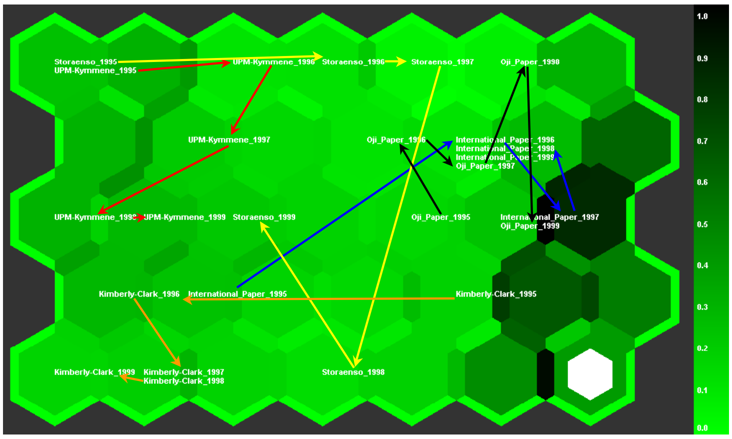

Figure 5. Movements of the top 5 pulp & paper companies (1995-1999).

An interesting note is that with the exception of 1995, International Paper, the largest pulp and paper manufacturer in the world, is consistently found in one of the poorly performing groups. In 1995, International Paper is located in the A_1 group, but falls into the D and E groups for the remainder of the time period.

The best performing company in the Top 5 is without doubt Kimberly-Clark (third largest), which stays in Group A1 during the years 1996-99, after a dramatic climb from the poor end of the map. Kimberly-Clark explains the poor performance of 1995 with the merger between Kimberly-Clark and Scott Paper Company. The merger required the sale of several profitable businesses in order to satisfy US and European competition authorities. The merger also caused a substantial one-time charge of 1,440 million USD, decreasing profitability further. However, profitability was back up again in 1996.

The poorest performing company is Oji Paper, the largest Japanese company, and the fourth largest in the world. A likely reason for Oji Paper's worsening performance is the Asian financial crisis.

The performance of UPM-Kymmene (fifth largest) is slightly better than the Finnish average, remaining in either the A_2 or C groups. Stora Enso (second largest) on the other hand moves from very good in 1995 to downright poor performance in 1997. The substantial change in position on the map in 1998 was due to a combination of two factors. The first was decreased profitability, due to costs associated with the merger of Stora and Enso. The second factor was a strengthened capital structure, which of course further affected profitability ratios like ROE and ROTA. However, profitability improved again in1999. Both Stora Enso and UPM-Kymmene were performing excellently in 1995, when market pulp prices were high, but the profitability of both companies fell as market pulp prices dropped[1].

5 Conclusions

In this study, financial information for 88 companies in the international telecommunications industry and 76 companies in the international pulp and paper industry has been collected using the Internet as a source of information, and a financial database has been created. A number of financial ratios, chosen from a previously published empirical study, have been selected and calculated based on the information in the database. Then, a data-mining tool, the self-organizing map, has been used to perform a financial competitor benchmarking of these companies.

In the original studies (Eklund, 2000; Karlsson, 2001), a number of benchmarks were performed, of which the performance of the top 5 companies was illustrated here.

The results of the study provide further evidence that the self-organizing map is a feasible and effective tool for financial benchmarking, and more generally, for converting passive information into refined information. The results are easy to visualize and interpret, and provide a very practical way to compare the financial performance of different companies. The discovered patterns were confirmed with existing domain knowledge.

---

[1] United Paper Mill (UPM) and Kymmene merged in 1996, and Stora and Enso-Guzeit merged in 1998. However, in the experiment we have used consolidated reports for the entire experiment.

We will continue with this line of research, and our next step will be to insert the most recent quarterly financial data into the maps. This will make the maps more up to date, and the method will therefore be of practical interest for management, but also to other interested parties. Another interesting topic would be to compare the companies' movements on the map to their stock prices.

6 Acknowledgements

The financial support from TEKES (grant number 40943/99) and the Academy of Finland is gratefully acknowledged.

References

Ambroise, C., Seze, G., Badran, F., and Thiria, S. (2000), 'Hierarchical Clustering of Self-Organizing Maps for Cloud Classification'. Neurocomputing, Vol. 30, No. 1, pp. 47-52.

Back, B., K. Sere & H. Vanharanta (1997), Analyzing Financial Performance with Self-Organizing Maps. *In* Proc. of the Workshop on Self-Organizing Maps WSOM'97, Espoo, Finland, in Jun 4-6, pp, 356-361.

Back, B., Sere, K. and Vanharanta, H. (1998), Managing Complexity in Large Data Bases Using Self-Organizing Maps. Accounting Management and Information Technologies. No. 8, 1998, pp. 191-210.

Back, B., K. Öström, K. Sere & H. Vanharanta (2000), Analyzing Company Performance Using Internet Data. In Proc. of the 11th Meeting of the Euro Working Group on DSS, Ed. By Zaraté, Toulouse, France, pp. 52-56.

Becanovic, V. (2000), 'Image Object Classification Using Saccadic Search, Spatio-Temporal Pattern Encoding and Self-Organization'. Pattern Recognition Letters, Vol. 21, No. 3, pp. 253-263.

Bendell, T., Boulter, L. and Kelly, J. (1998), Benchmarking for Competitive Advantage. Pitman Publishing. London.

Bishop, C. M. (1995), Neural Networks for Pattern Recognition. Oxford Univ. Press. Avon. pp. 296-300

Chen, D., Chang, R., and Huang, Y. (2000), 'Breast Cancer Diagnosis Using Self-Organizing Map for Sonography'. Ultrasound in Medicine and Biology, No. 26, pp. 405-411.

Eklund, T. (2000), On the Application of Self-Organizing Maps in Benchmarking – As Applied to the Pulp and Paper Industry. Unpublished Masters Thesis at the Dept. of Information Systems at Åbo Akademi University, Turku

Johnson, R. A. and Wichern, D. W. (1997), Business Statistics: Decision Making with Data. John Wiley & Sons, Inc. USA. pp. 63

Karlsson, J. (2001), Financial Benchmarking of Telecommunications Companies. Unpublished Masters Thesis at the Dept. of Information Systems at Åbo Akademi University, Turku.

Kaski, S., Kangas, J., and Kohonen, T. (1998), ‘Bibliography of Self-Organizing Map (SOM) Papers 1981—1997’. Neural Computing Surveys, 1, pp. 102-350. <http://www.icsi.berkeley.edu/~jagota/NCS/> (read Jul. 24th 2000).

Kohonen, T., Hynninen, J., Kangas, J. and Laaksonen, J. (1996), SOM_PAK: The Self-Organizing Map Program Package. Helsinki Univ. of Technology.

Kohonen, T. (1997), Self-Organizing Maps. 2nd edition. Springer-Verlag. Heidelberg.

Kohonen, T. (2000), 4311 Works That Have Been Based on the Self-Organizing Map (SOM) Method Developed by Kohonen. <http://www.cis.hut.fi/nnrc/refs/> (read Oct. 24th 2000).

Lehtinen, J. (1996), Financial Ratios in an International Comparison. Acta Wasaensia No. 49. Vasa.

Martín-del-Brío, B. & C. Serrano-Cinca (1993), Self-organizing Neural Networks for the Analysis and Representation of Data: Some Financial Cases, Neural Computing and Applications, 1993, No. 1, pp. 193-206.

Matussek, H., I. Janssens, J. Kenny & J. Riannon (1999), The Top 150: A Tale of Two Halves, *Pulp and Paper International* 1999: Sept, 27-39.

Raivio, O., Riihijärvi, J., and Mähönen, P. (2000), ‘Classifying and Clustering the Internet Traffic by Kohonen Network’. Proc. of the ICSC Symposia on Neural Computation (NC’2000) May 23-26, 2000 in Berlin, Germany. ICSC Academic Press.

Ultsch, A. (1993), ‘Self organized feature planes for monitoring and knowledge acquisition of a chemical process’, Proc. of the Int'l Conf. on Artificial Neural Networks, pp. 864-867. Springer-Verlag. London.

Visa, A., Toivonen, J., Back, B. and Vanharanta, H. (2000), A New Methodology for Knowledge Retrieval from Text Documents. Proceedings of TOOLMET2000 Symposium – Tool Environments and Development Methods for Intelligent Systems, pp. 147-151.