Draft of paper In : Soft Computing in Financial Engineering. Eds: Rita A. Ribeiro, R. Yager, H-J. Zimmermann, J. Kacprzyk. Physica-Verlag, Series on studies in Fuzzyness and Soft Computing, 1999, pp 320-338

Financial Analysis of Non-Financial Companies with Neural Networks

Rita Almeida Ribeiro and Fernando Moura-Pires

Departamento Informática
Faculdade Ciências Tecnologia
Universidade Nova de Lisboa
2825 Monte Caparica
Portugal
E-mail: {rr, fmp}@di.fct.unl.pt

Abstract

This paper presents a neural network approach to classify the 500 biggest Portuguese companies. The objective is to find relations and correlations between their relevant financial and economic attributes. Further, we want to elicit information according to the most important market players, such as banks, stockholders, managers and government. Thus, our proposal is a neural network analysis of financial and economic attributes using the most important market players perspectives.

Keywords: market players perspectives, financial and economic analysis, neural networks

1 Introduction

There are various perspectives on defining what a good company is. The government focuses on their contributions to the national economy; managers are concerned with efficiency, profits and productivity; bankers focus on financial aspects; and stockholders are primarily interested in profits.

For an individual investor the availability of all the above information can represent a complementary source of knowledge about the stock market. Further, it also provides the investor with measures for risk assessment, allowing an in-depth analysis of the companies performance.

Every year, the biggest non-financial 500 companies are ranked in the Portuguese magazine Exame [Exame, 1996]. In this paper we used 1995 data, where the companies are ordered by their net sales volume and the 500 selected (bigger) are those with net sales above 4,522 million escudos (about 27.4 million dollars). The magazine also presents a financial and economic ordering, per sector of activity, of the 10 best firms. This latter ordering uses a weighted average aggregation to select the best 10 companies per activity sector, with eight criteria: sales growth; net profits growth; net assets profitability; owners equity profitability; sales profitability; gross added value; solvency and general liquidity. Summarising, the classification used in the magazine [Exame, 96] for selecting the bigger and best NON-financial Portuguese companies is:

- Bigger: net sales greater than 4,522 million escudos
- Better: weighted average aggregation of grades for sales growth, net profits growth, net assets profitability, owners equity profitability, sales profitability, gross added value, solvency and general liquidity.

In this paper we used a neural network analysis approach because we want to automate the process of finding relations and correlation's between attributes, which a grade aggregation method cannot provide. Further, we also want to include information about financial and economic ratios [Aubert-Krier, 77], [Whalen, 86], according to the most important market players, such as banks, stockholders, managers and government. The most appropriate technique to meet our objectives is neural networks because it automates the classification by similar attributes (see examples in [Trippi, 1992]). Thus, we develop a neural network analysis of financial and economic attributes with four main market player perspectives, as follows:

A) <u>Government</u>, that is concerned with the contribution to the national economy. The indicators used are: a.1.) Gross added value (GAV) - sum of the net sales, production fluctuations, subsidies and net extraordinary profits; a.2.) Gross added value (GAV)/Net sales. This measures how much a company contributes to the national economy per escudo sold.

B) <u>Management</u>, that is concerned with the firms profitability, dynamism and efficiency. The attributes used are: b.1) Sales growth; it is given by the ratio Sales 95/Sales 94 ; b.2) Net profits growth; it is given by the ratio Net profits 95/Net profits 94, that measures the dynamism and the capacity to maintain or increase the market quota; b.3) Assets turnover; given by Net Sales/Assets, that represents the degree of efficiency of available resources; b.4) Productivity; it is the ratio gross added value (GAV)/Number of workers which measures the degree of efficiency of human resources.

C) <u>Stockholders</u>, that are mainly concerned with profitability. The indicators are: c.1) ROI (return on investment); it is the profit per unit of capital invested in the company; c.2) ROE (return on equity); ratio of Net profits/Owners equity. This ratio measures the profitability of the owners capital; c.3) Profit margin on sales; it is the ratio of profits after taxes/ sales; c.4) Sales profitability; measured with Current profits/ Sales.

D) <u>Banks,</u> that are concerned with the financial equilibrium of companies. The attributes used are: d.1) Indebtedness; given by the ratio Liabilities/Net Assets. It measures the capacity of the firm to contract loans (the bigger the worse); d.2) Solvency; given by the ratio Owners equity/Liabilities. It measures the long-term capacity to fulfil commitments; d.3) Financial autonomy; given by the ratio Owners equity/Net assets. It measures the participation of the owners equity in financing of the company activities (complement of Indebtedness); d.4)

General liquidity; given by the ratio Assets/Current liabilities. It measures the capacity to fulfil the short-term commitments; d.5) Cash flow. It measures the auto-financing capacity of the company.

The reason for this grouping is linked, as mentioned, with the different interests of managers, bankers, stockholders and government, as well as, with the idea of offering a general view, to new stock market investors, of the best Portuguese firms.

2 Classification Techniques

In order to reason one needs to classify one's knowledge. This process of classification can be defined as an ideal arrangement of the alike and unlike. According to Mirkin [Mirkin, 96] the purpose of the arrangement is:

"1) to shape and record knowledge;
2) to analyse the structure of phenomena; and
3) to relate different aspects of a phenomena in question to each other."

Classification has been used in many fields of knowledge, as for instance: mathematics, artificial intelligence, chemistry, physics, geology, and social sciences. In mathematics it comprises two kinds of activities: computation (find the exact or approximate solutions to various equations and optimisation problems) and deductions about properties of mathematical concepts (the art to construct, analyse and connect classifications of mathematical objects by means of logical tools). In artificial intelligence the process is similar, but sometimes the knowledge representation is different and it uses heuristics for optimisation problems. Addressing other fields is beyond the scope of this paper.

Here our focus is on neural network techniques within the artificial intelligence field. Three of the most important techniques for classification are, conceptual clustering [Genari, 89], clustering [Mirkin, 96] and neural networks [Haykin, 94]. The first technique is usually used for problems where the attributes are nominal while the others are used when the attributes are numeric. In our case, the attributes are numeric, hence, either clustering or neural networks can be used. These two techniques are equivalent, since they use a measure of similarity based in the Euclidean distance. We opted for a neural network approach for the classification because of software availability and prior experience of one of the authors [Nascimento, 1997].

In classification problems one of the best well-known neural network technique is denoted Self-Organizing Maps or Kohonen neural network [Kohonen, 85]. We selected this neural network for our classification approach. Since our objective is the classification of non-financial companies, only a brief overview of classification with Kohonen neural networks is presented.

Kohonen neural network uses competitive learning [Kohonen, 95]. The neurons are placed at the nodes of a lattice, usually a two-dimensional one. The neurons become selectively tuned to various inputs patterns in the course of a competitive learning process. The location of the winner neuron tends to become ordered, with respect to each other, in such a way that a meaningful coordinate system for different input spaces is created over the lattice. A Kohonen neural network attribute map is, therefore, characterised by a formation of a topographic map of input patterns, in which the spatial locations of the neurons in the lattice, correspond to the attributes of the input patterns.

The essential aspects of a Kohonen neural network, which are the same we used for the classification, are [Haykin, 94]:

1) the neurons compute a discriminate function (usually based on Euclidean distance);
2) there is a mechanism that compares these discriminate values and selects a winner;
3) there is another mechanism to activate the selected winner and its neighbours (function of a radius);
4) it contains an adaptive process that enables activated neurons to increase their discriminate function values, in relation to the input signal (it is dependent of a neighbourhood function and a learning rate).

The whole process is iterative, i.e. each example is classified one at a time, and then the algorithm is repeated until the error stabilises. The size of the error defines the stopping criterion for the iterative procedure.

3 Neural Network Analysis

In this section we describe the assumptions used for handling the data and the parameters of the Kohonen neural network. After we discuss the results obtained with this classification approach.

3.1 Data normalisation

As mentioned in the introduction, only a subset of all available attributes is used in this paper. Following to the four perspectives defined, each attribute, its range, its average and its standard deviation is shown in Table 1. As can be observed, the attribute scales are very different. This fact represents a problem for classification techniques which are based in dissimilarity measures using Euclidean distances [Kaufman, 90].

In order to compare the different attributes, it is usual to normalise the raw data. The most frequent methods of normalisation are:

Normalisation [0, 1], *i.e.*

$$y_j = \frac{x_j - a}{D}$$

where x_j is the value of the attribute of the example *j*, *a* is its minimum and *D* is the range of the attribute raw data (D= Max-Min).

Gaussian normalisation, *i.e.*

$$y_j = \frac{x_j - m}{\sigma}$$

m is the average of the attribute and is its standard deviation.

S normalisation, *i.e.*,

$$y_j = \frac{1}{n} \times \frac{x_j - m}{\sum_{i=1}^{n} |x_i - m|}$$

where *n* is the number of samples in the raw data.

		Raw data values			
Persp	Parameters	Min	Max	Mean	Std Dev.
A	Gross added value (GAV)	-1,683.0	424,732.0	6,606.2	31,400.6
	GAV / Net Sales	-35.5	196.5	24.5	21.2
B	Sales growth	-43.6	397.0	23.7	177.7
	Net profits growth	-30,025.7	167,360.4	434.6	8,438.5
	Assets turnover	0.1	53.9	2.2	3.2
	Productivity	-36.6	2,092.3	14.1	94.7
C	Return on investment (ROI)	-37.4	44.3	2.7	7.1
	Return on equity (ROE)	-585.7	290.9	6.0	45.8
	Profit margin on sales	-176.7	30.6	1.5	10.2
	Sales profitability	-52,534.0	131,395.0	877.5	8,003.5
D	Indebtedness	0	266.9	65.5	23.5
	Solvency	-0.6	11.7	0.8	1.1
	Financial autonomy	-1.7	1	0.3	0.2
	General liquidity	1.0	16	1.7	1.4
	Cash flow	-49,346.0	197,334.0	1,869.0	11,102.4

Table 1. Attributes characterisation

The choice of the normalisation method is a challenging problem and it is problem dependent. The Gaussian normalisation is good if the raw data follows a normal distribution, i.e., if the values are not very different. In our case, since the values are very different, a Gaussian normalisation yields a distribution where the values near the average are not very discriminating. The S normalisation seems more appropriate in our case because it provides a wider discrimination. In order to illustrate this point, Table 2 depicts the range of values for two attributes, without normalisation, with Gaussian normalisation and with S-normalisation. For the first attribute the Gaussian normalisation could have been used. For the second attribute the Gaussian normalisation should not be used because the range is to small (not discriminative).

Attributes	Normalisation	Min	Max
Financial Autonomy	without	-1.7	1
	Gaussian	-8.6	2.8
	S-Normal	-11.5	3.7
Sales profitability	without	-52,534.0	131,395.0
	Gaussian	-6.7	16.3
	S-Normal	-30.9	75.5

Table 2. Example of normalisation's.

Based on the data in Table 1 we have chosen, for our classification tests, the following two kind of analyses: no normalisation (raw data) and the S-normalisation.

3.2 Kohonen neural network parameters

For each market player perspective - government, managers, banks and stockholders/investors - we performed tests with different neural network sizes. To chose the best dimension for the neural network to be analysed, we performed tests for the following neural network dimensions: 5 x 15, 10 x 20, 30 x 40 and 40 x 60. We believe the dimension 30 x 40 is rich enough to discriminate our data. Other neural network sizes proved either too sparse (almost no companies were similar) or too concentrated (too many very different companies were grouped in the same group).

After some tests, the other parameters selected for the learning process of the neural network are : number of iterations greater than 1,000,000; learning rate varying in the interval [0.01, 0.1] and radius (neighbourhood function) varying in the interval [0, 40]. When the number of iteration increases, the learning rate and the radius decrease. Both the learning rate and the neighbourhood function belong to the adaptative process of the Kohonen neural network, as mentioned in section 2.

There are two tools to analyse the Kohonen neural network: U-mat maps and Sammon maps.

The U-mat maps are described in Ultsch [Ultsch, 93]. These maps are similar to the Kohonen neural network where the neurons are represented by a hexagon with a number (identity of the company, following the ordering of the magazine for big companies) or with a point ("•"). Each neuron is surrounded by other hexagons, with different grey levels representing the distance (in input space) between the neurons. The grey levels codification is the following: black - extremely distant; almost black - very distant; dark grey - quite distant; grey - distant; light grey - close; almost white - very close.

The Sammon maps [Sammon Jr, 69] represent the projection of the input space to a 2-dimensional space where the distances between the image points tend to approximate the Euclidean distances of the input vectors.

3.3 Analysis of the Results

In this section an additional explanation about the figures to be discussed is needed. We discuss the U-mat maps and their respective Sammon maps. The latter is a graphic representation of the former and it was used to improve readability. Both maps are generated by a program package [Kohonen, 96] which allows for the visualisation of distances, the U-mat using neighbouring distances with grey levels and the Sammon using Euclidean distances.

Two neural networks are generated for each of the four perspectives, one using raw data and other S-normalised data. Thus, we obtained a total of eight neural networks. For reasons of space only the normalised maps are shown. Further, for each neural network it was computed two outputs, a U-mat map and a Sammon map. These maps are used in the result analysis.

Fig. 1 - U-mat map example (perspective B)

In order to improve the visualisation of distinct groups, only partial regions of the entire U-mat maps are show. They correspond to the most distinctive groups here analysed. The distinctive groups are also manually underlined, both in the U-mat and Sammon maps, to further improve the visualisation. A reduced example of a complete U-mat is depicted in Fig. 1 to show the complete size of a 30x40 Kohonen neural network.

The focus here is to unveil how the classification process provides an important tool for a financial and economic analysis of similarities and dissimilarities between companies.

The groups allow for a detailed analysis (neuron by neuron) of which companies are included and the respective raw values show why they are similar. However, here we only analyse the groups extremely distant/distinct (surrounded by black neurons), very distant (surrounded by almost black) and distant (surrounded by dark grey and grey). Thus, we only analyse the groups including companies that are very distinguishable from the average. In addition, the discussion is not an in-depth financial and economic analysis of the results obtained for the indicators, but solely a discussion of the reasons why these groups are far away from others.

After introducing our assumptions, it is now possible to discuss the results obtained. In order to do so, we follow the four perspectives of the market players, as described in the introduction.

A) Government perspective - Contribution to the National Economy

The most interesting attributes which provide a measure of the contribution of companies to the national economy are the GAV and the GAV/Sales. Fig. 2 depicts the distinct groups in the U-mat map and the complete Sammon map.

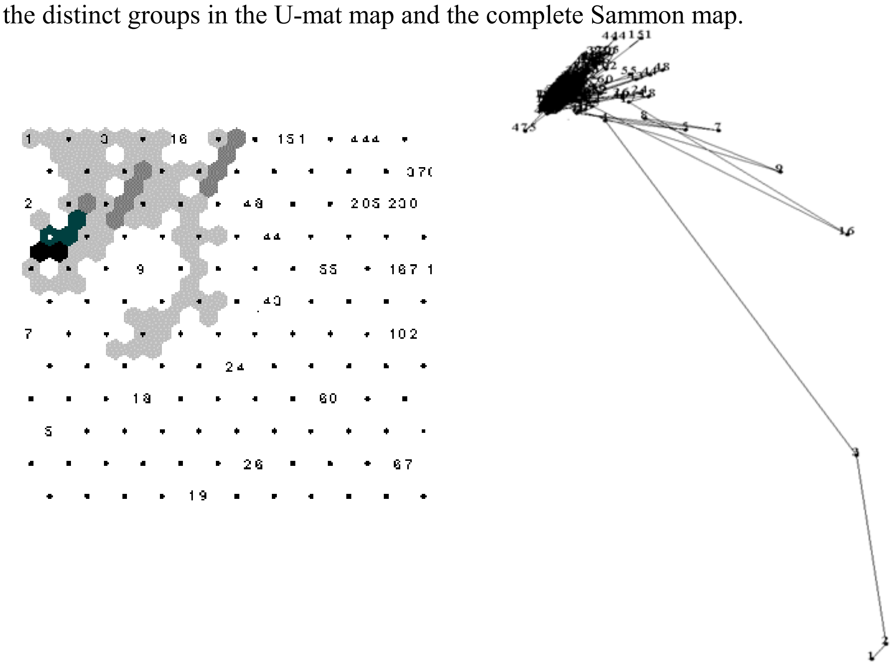


Fig. 2. Partial Umat and complete Sammon maps of contribution to National Economy

The results in the U-mat show two very distinct groups (surrounded by black, almost black and dark grey neurons) each containing respectively companies #1, #2 and #3. After, there are three quite distinct groups (surrounded by dark grey and grey neurons) containing companies #16, #9, #7. Then there are two other distinct groups (surrounded by grey neurons), one containing company #5 and another with companies #18, #24, #43, #55, #44, #48, #151, #444. The remaining companies in the figure are all considered very similar since they are surrounded by almost white neurons (very close). Finally, there is a far away group, in the lower right corner, containing company #475 (not shown to save space). This last company has negative values both for the GAV and the GAV/Sales, with values: -1,683 and -35.5, respectively. Hence, it is obvious why it was classified in an opposite group (extremely distant).

The reason for the first very distant (distinct) groups is due to big differences in their GAV values. Their GAV values are: #1 = 424,731 (thousand escudos); #2 = 416,080; #3 = 292,591. The GAV values for the next closest groups are: #16 = 153,712; #9 = 123,746; #7 = 75,721. The other distinct groups (surrounded by grey neurons) have values in the order of 30.000-50.000. Since all other companies have GAV values ranging from -134 to around 20.000 they are considered rather similar and grouped in big group, all within almost white neurons (not shown in the Fig. 2 for reasons of clarity).

Observing the Sammon map, we can now see how distant (Euclidean distance) the companies really are. We see that company #1 and #2 are more or less similar, company #3 is more distant and than the closest ones are company #16 and #9. Further up, we can still distinguish companies #7, #5, #48, #44, #8, #151, #444, #42 and #4. After these ones, all the other companies are almost identical (overlapping black area). Finally, we observe that company #475 lies in the opposite side of companies #1 and #2 and it is also distinct from the big group.

The results obtained using raw data have similarities with the ones in Fig. 2 but with less distinctive groups. For example, company #1 and #2 are considered identical, when in reality the GAV values are similar but the other attribute values, GAV/Sales, are quite different: #1 = 56.1 and #2 = 77.2. For the other groups the same problem arises. It should be pointed out that using normalised values generates much better comparative results because we are using distances as a similarity measure. When the scales are very different the values are not easily comparable.

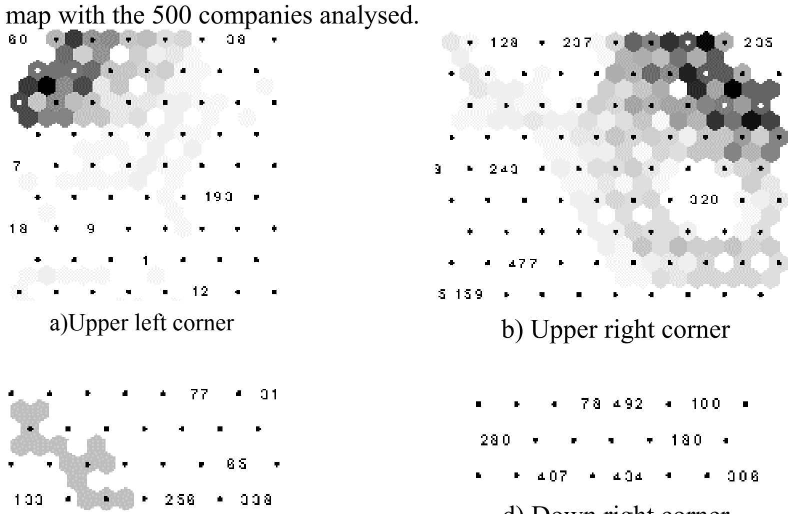
In summary, the U-mat unveils that after the small subset of companies located in the upper left hand corner, almost all the others are very similar in what concerns the GAV and GAV/ Sales attributes (all included in almost white neurons). In financial terms they contribute much less to the national economy. The higher contribution to the Portuguese national economy is due to the companies #1, #2, #3, #16, #7 and #5. One interesting point about these companies is that they are all state-owned companies, except number #7 (a petrol company, SHELL). Hence, in reality, the only private company with a significant national contribution is number #7 (SHELL).

B) Management perspective - Dynamism, Efficiency and Profitability

Although profitability is an issue of relevance to managers, we postpone its discussion till latter because it is the major concern of stockholders and investors. Also in this test we eliminated 90 companies from the set of examples, since there were no data available for more than one attribute. It should be stressed that when one attribute has no numerical value, the aggregated distance obtained will, necessarily, be smaller than the one of other companies with similar values in the other attributes. Non available data distorts groups because of the disparity of the aggregated values.

Draft of paper In : Soft Computing in Financial Engineering. Eds: Rita A. Ribeiro, R. Yager, H-J. Zimmermann, J. Kacprzyk. Physica-Verlag, Series on studies in Fuzzyness and Soft Computing, 1999, pp 320-338

The most relevant attributes for evaluating the efficiency and dynamism of companies are sales growth, net sales growth, assets turnover and productivity. Fig. 3 depicts the more distinctive groups in the U-mat maps and Fig. 4 shows the complete Sammon map with the 500 companies analysed.



a)Upper left corner

b) Upper right corner

c) Down left corner

d) Down right corner

Fig. 3 a),…,d) - Partial U-mat for Dynamism and Efficiency

In Fig. 3 there are four distinct sets of groups, located in the upper left corner, upper right corner, lower left corner and lower right corner. The upper left corner group includes company #60. The upper right corner includes three groups, the farthest with company #235 (surrounded by black and almost black neurons), then another surrounded by dark grey neurons with company #320 and the last (surrounded by grey neurons) includes companies #128 and #237. In the lower left corner (surrounded by dark grey and grey neurons) there is a group with company #133 and in the lower right corner (surrounded by grey and almost grey neurons) has a group with company #306.

The group with company #60 has the greatest value for productivity of the whole set, and this is the reason for being extremely distant (black surroundings). The value for productivity is #60 = 2,092.3 while the next greatest value is only #7 = 244.2 (in a light grey neuron). The reasons for the distinct groups in the upper right corner are quite different. The farthest group has the biggest value for Net sales growth, #235 = 167,360.4, but a low value for productivity #235 = 1.9. The next closer group, with company #320 has a lower value for Net sales growth, #320 = 29,218, but it compensates with a larger value for productivity, #320 = 10.6. The next closest group, with companies #128 and #237 have a much smaller values for Net sales growth, #128 = -75.6 and #237 = 2,839.8, but a good productivity level (#128 = 12.1 and #237 = 8.9). Further, these two companies share the best values for sales growth, #128 = 397 and #237 = 315.8. The distances between the groups just described are clearly visualised in the respective Sammon map (Fig. 4).

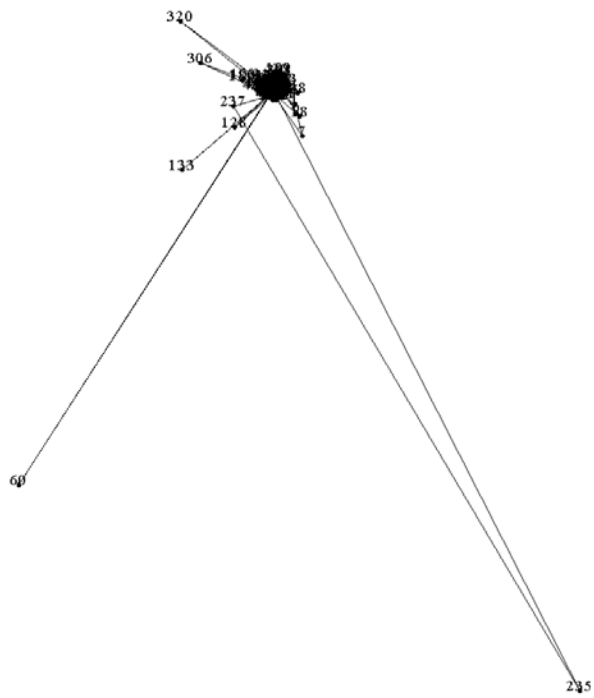


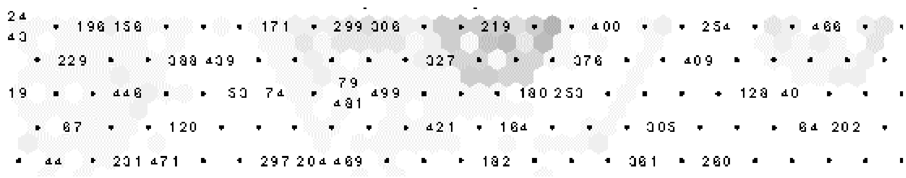
Fig. 4 Complete Sammon maps for Dynamism and Efficiency

Observing again Fig. 3 the opposite group in the lower side of the U-mat includes company #133, which has the worst bad result for Net sales growth, #133 = -30,025.7. Looking at the opposite group, which includes company #306, the value for Net sales growth is not bad #306 = 219.5, but it belongs to a distinct group because it has small values for all other attributes, sales growth #306 = -26, assets turnover #306 = 53.9 and productivity #306 = 2.1.

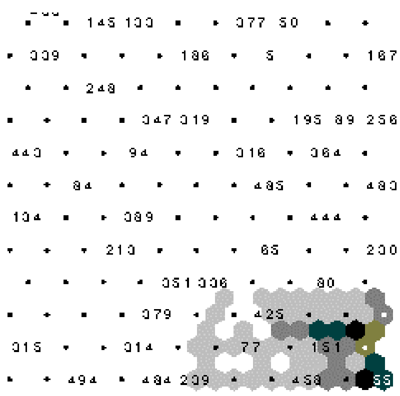
The situation in terms of efficiency and dynamism for all the companies located in the upper side of the U-mat is very good while the lower side of the U-mat contains the worst companies. It should also be noted that the attribute, Assets turnover, has small value variations for all companies and, hence, it is not important in the differentiation of the groups. The Sammon map (Fig. 4) shows that companies #60 and #235 are really very distinct form the others and that companies #320, #306, #237, #128 and #138 are distinct, but closer to all the more average ones (black area).

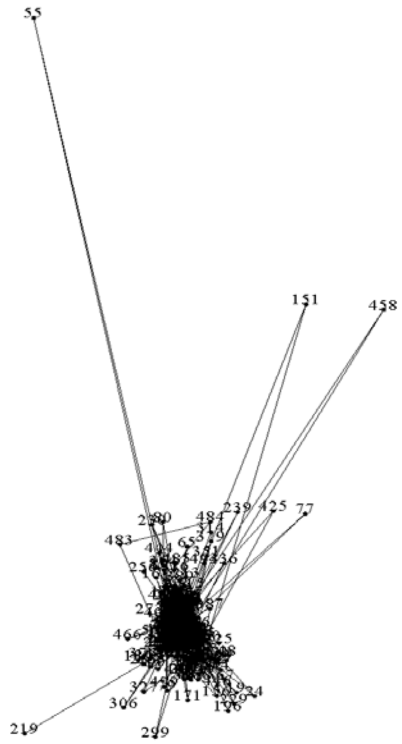
C) Stock holder perspective - Profitability

In this perspective the most important indicators to discuss are return on investment (ROI), return on equity (ROE), profit margin on sales and sales profitability. Fig. 5 depicts the partial U-mat maps with the most distinctive groups and the complete Sammon map.



a) Upper side





b) Down right corner c) Complete Sammon map

Fig. 5. Partial U-mat and Complete Sammon maps for Profitability

In Fig 5.a) the left upper corner has a group including companies: #24, #43, #196, #156, #229, #446, #67, #19, #44 (also light grey). In the upper middle section there are three groups, one with company #219, another with company #299, another with company #306 and another with company #466. In the U-mat lower right corner (Fig. 5 b) there are four distinct groups, one with company #55, another with companies #151 and #458, another with company #77, another with companies #239, #425 and quite a huge one including companies: #494, #484, #314, #379, #351, #336, #65, #80, #230, #444, #485, #483, #316, #364, #256, #89, #195, #167 (light grey - close).

Looking at the Sammon map, Fig.5. c), we verify that the left upper corner group is not very significant since it is quite close to the other companies (in the U-mat they were light grey) and that the same problem applies to the big group in the lower right hand corner. These two groups are not discussed in detail here because, as previously mentioned, we focus in distinctive groups (from black to grey).

Observing the groups located in the upper middle U-mat map (Fig. 5 a), we have the group with company #299 (dark grey) and then three others surrounded by grey. The raw values for these are depicted in table 3. The values are almost all positive for all the profitability attributes (except in company #466). In summary, these companies are quite profitable in financial terms.

Company	ROI	ROE	Profit margin on sales	Sales profitability
#219	16.2	290.9	1.3	123
#306	38.9	NA	0.7	37
#299	44.3	108.9	17.4	1,740
#466	14.3	23.1	10.1	-811

Table 3– Raw results comparison - stock holder perspectives

An interest point to highlight is given by analysing the raw data of the upper left corner group (even though it is light grey). This is the furthest group in relation to the almost black one and, hence, it is the group with the most profitable companies in the whole set tested. It includes the companies with the maximum values for all the attributes (table 1). The values are not detailed because it is a light grey group.

Observing now the attribute values for company #55 (Fig 5. b), located in the black and almost black group (lower right hand corner) we see that it has the worst values (see its values in table 1) for all attributes except for the ROI. It is clear that it is disastrous in profitability terms. It should also be noted that company #55 is the Portuguese public railroad and usually public railroad companies are not profitable. Companies #458 and #151, located in the closest group, also have quite bad values for ROI, ROE, the profit margin on sales and sales profitability (table 4). The next closest group (company #77) also shows bad values for all attributes as can be observed in table 4. The last group also has bad values for all attributes (except for ROE which has no available data).

Another interesting conclusion to elicit from table 4 data, is the effect of using an aggregated normalised distance for generating the groups. Companies #458 and #151 are in the same group, though they have very different values for Sales profitability because the overall value is similar. These distortions can only be studied by looking at the raw data.

Company	ROI	ROE	Profit margin on sales	Sales profitability
#458	-26.3	-568.2	-15.2	-671
#151	-20.5	NA	-41.4	-5,664
#77	-3.9	-233.2	-2.2	-562
#425	-13.1	NA	-10.3	-565
#239	-34.1	NA	-8.1	-388

Table 4 – Raw results comparison - stockholder perspectives

In addition, the results obtained in the Kohonen neural network for the raw data (not shown in the paper) were slightly different from the ones obtained with the normalised values because they are less discriminating. I.e. some companies are included in a group that has rather different values in the comparable attributes. For reasons of space these differences are not discussed further in this paper.

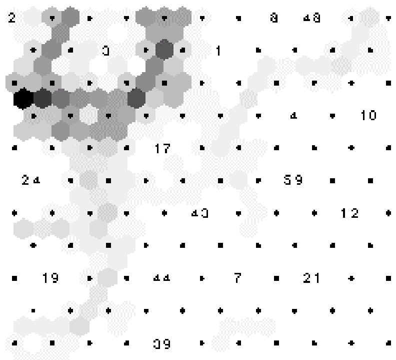
D) Banks perspective - Financial equilibrium

In this section we will discuss the most import indicators to assess the financial equilibrium of the 500 biggest Portuguese companies. The relevant attributes are, indebtedness, solvency, financial autonomy, general liquidity and cash flow. Fig. 6 depicts the U-mat map with the most dissimilar groups and the Sammon map.

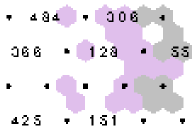
In the upper left corner of the U-mat (Fig. 6 a) and Sammon map (Fig. 7) we observe that there are two very distinct groups containing companies #2 and #3. Then there are four other distinct groups (surrounded by grey neurons), one including companies #1, #8, #48, another including company #17, another including company #24 and finally another with company #19. In the upper right corner (Fig 6 b) there are two other groups, one with company #55 and another with companies #128 and #306. Finally, there are other groups in the lower left corner (Fig. 6 c), but since they are surrounded by light grey neurons (close) we will not discuss them here. However, these light grey groups are quite distinct in the Sammon map (companies #365, #422, #67, #159 and #24) as can be observed. It should be highlighted that company #24 is included in a different group in the U-mat due to the different distances algorithm used (it is the only case in all tests performed).

Looking at the companies raw values we observe that company #2 is extremely different from the others. The main reason is that, comparatively, it has a very high cash-flow: #2 = 197,334. The next closest group in Fig 6 a (surrounded by black and dark grey neurons) includes company #3, which has a smaller cash flow, #3 = 124,550. Then, the closest four groups (surrounded by grey neurons) include companies #8, #48, #1; #17; #19; #24. The values of the cash flow for the first group are #8 = 22,912, #48 = 22,406, #1 = 26,924; for the second group the value is #17 = 25,840; for the third group the value is #19 = 24,993 and for the fourth group the value is #24 = 26,670. The distinct four groups are due to the values of the other attributes. In the first group the Indebtedness and Solvency ratios are similar (respectively #8 = 80.9; #48 = 85.2; #1 = 72 and #8 = 0.24; #48 = 0.17; #1 = 0.39).

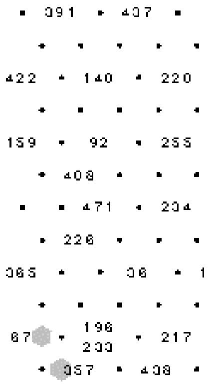
Draft of paper In : Soft Computing in Financial Engineering. Eds: Rita A. Ribeiro, R. Yager, H-J. Zimmermann, J. Kacprzyk. Physica-Verlag, Series on studies in Fuzzyness and Soft Computing, 1999, pp 320-338



a) Upper left corner



b) Upper right corner



c) Down left corner

Fig. 6. Partial Umat map for Financial Equilibrium

In the second group the values for Indebtedness and Solvency are #17 = 44.6 and #17 = 1.24. In the third group the values for Indebtedness and Solvency are #19 = 17.7 and #19 = 4.6. In the fourth group the values for Indebtedness and Solvency are, #24 = 9.4 and #24 = 9.68. After looking at the values it is quite obvious why these companies belong to different groups and why they are very distinct from companies #2 and #3.

We will now discuss the two groups located in the upper right corner (Fig 6 b) one including company #55 and another including companies #306 and #128. The raw data for their financial health are shown in table 5.

Company	Indebtedness	Solvency	Financial autonomy	General liquidity	Cash flow
#55	97.7	0.02	0.02	1	-49,346
#128	192.6	-0.48	-0.92	6	1,414
#306	266.9	-0.63	-1.60	NA	57

Table 5 - Results comparison - banks perspectives

It is very clear that company #55 is considered at the opposite extreme of companies #2 and #3 since its cash flow is rather negative (the minimum found for this attribute). Regarding the next closest group to company #55, we observe that the indebtedness, solvency and financial autonomy are similar and quite different from company #55. However, they are also quite different from the values of the companies in the groups in the opposite direction. An interesting aspect in this last set, is that the classification process grouped together almost bankrupt companies (when the indebtedness ratio is bigger than 100, firms are technically bankrupt).

Further, it should be noted that all the companies near this area have bad indebtedness and solvency indicators.

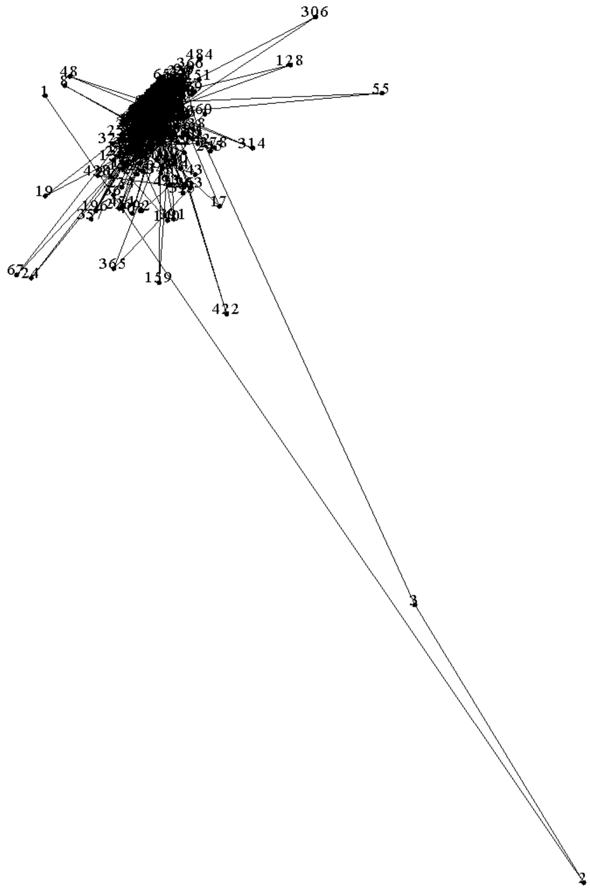


Fig. 7. Complete Sammon map for Financial Equilibrium

Much more information can be extract from the analysis of the groups and individual neurons, both in figure 6 and figure 7, but here we just focused in the most distinctive aspects. In addition, this classification process can facilitate the risk assessment done by banks, since it immediately separates the bad companies from good ones. Further, in this point we do not address the results obtained with raw data because those were more or less similar, though less discriminated. As previously mentioned this is due to using distances as a similarity measure.

Conclusion

We introduced a neural network method for classification of the most important financial and economic attributes of the 500 biggest non-financial Portuguese companies. Our objective was to detect which are the best and worst companies in financial and economic terms according to the most important market players.

The four market perspectives used are: government, stockholders, managers and banks. Market players have different interests when evaluating companies, hence, they look at different indicators. Our group analysis detected, for each market perspective, which are the distinctive companies both in good and bad terms. The information obtained can also be of use for individual investors, which can assess companies following their preferred criteria.

An interesting conclusion this analysis unveiled, is that the majority of the companies are quite similar, within each perspective, and only a handful present some outstanding feature(s). This fact is due to the relatively few distinctive groups detected. Another important aspect unveiled is that the comparison between companies should have been split into two groups, private and public companies, because their results are quite different. While a public company can present very bad results a private company would be out of business, hence in future works the splitting will be considered.

In summary, we used a neural network analysis for performing a financial and economic analysis of non-financial companies. With this approach, relevant financial information about similar and dissimilar companies can be extracted. Another advantage of this approach is that the automated isolation of groups could also have been used for a further in-depth financial and economic analysis.

References

Aubert,-Krier, J., (1977), Os Meios de Informação e Decisão, Editorial Presença.

Genari, J. H. et al, (1990), Models of Incremental Concept Formation, in Machine Learning - Paradigms and Methods, Jaime Carbonell (ed). MIT Press/Elvesier

Exame, Edição especial No. 1, As melhores e maiores (1996), Out/Nov

Haykin, S. (1994), Neural Networks - A comprehensive Foundation. Macmillan College Publishing Company.

Kaufman, L. and P. J. Rousseeum, (1990), Finding Groups in Data, A Wiley-Interscience Publication, New York

Kohonen, T. (1989), Self-Organization and Associative Memory. Springer-Verlag, Berlin, 3rd edition.

Kohonen, T. (1995), Self-Organizing Maps. Springer-Verlag, Heidelberg.

Kohonen, T. et al. (1996), SOM_PAK: The Self-Organizing Map Program Package, Report A31, January 1996, Helsinki University of Technology, Faculty of Information Technology, Laboratory of Computer and Information Science, Finland.

Mirkin, B. (1996), Mathematical Classification and Clustering. Kluwer Academic Publishers.

Nascimento, S. and F. Moura-Pires (1997). A genetic approach to fuzzy clustering with a validity measure fitness function. 2nd International Symposium on Intelligent data Analysis, IDA-97, London, UK, Springer-Verlag.

Sammon Jr., J.W. (1969), A nonlinear mapping for data structure analysis. IEEE Transactions on Computers, C-18(5):401-409, May 1969

Ultsch, A. (1993), Self organized feature maps for monitoring and knowledge acquisition of a chemical process. S. Gielen, B. Kappen (editors), Proceedings of the International Conference on Artificial Neural Networks (ICANN93), 846-867, London. Springer-Verlag.

Whalen, T. and B. Schott, (1986), Financial Ratio Analysis, in: Fuzzy Logic in Knowledge Engineering, (ed) H. Prade and C.V. Negoita, Verlag TUV Rheinland.

Trippi, R. R. and E. Turban, Eds. (1992). Neural networks in Finance and Investing, Irwin. Publishers.