Using Previous Knowledge for Stock Market Prediction Based on Fundamentalist Analysis with Fuzzy-Neural Networks

RENATO DE C. T. RAPOSO, ADRIANO J. DE O. CRUZ AND SUELI MENDES
Núcleo de Computação Eletrônica, Instituto de Matemática,
Federal University of Rio de Janeiro
PO BOX 2324, Rio de Janeiro, RJ, CEP 20001-970
BRAZIL

Abstract: - The purposes of this paper are: i) show how to use the previous knowledge of any economic analyst and represent this knowledge using frames [7]; ii) discuss the application of a combination of Neural Networks and Fuzzy Logic to predict the evolution of stock prices of Brazilian companies traded on the São Paulo Stock Exchange; iii) present the obtained results. The network indicates if a trader would have to keep, sell or buy a stock using a combination of information extracted from balance sheets (released every three months) and market indicators. The results show that the network combining the previous knowledge of the economic analyst deliver good results depending on the quality of the available data and others factors.

Key Words: - Previous Knowledge, Fuzzy Neural Networks, Stock Market Prediction, Fundamentalist Analysis

1 Introduction

One possible way to obtain a high level profile of an investor is the use of frames to model and evaluate the previous knowledge gathered by the specialist about the investor. In this paper we use this previous knowledge to extract several basic knowledge about the investor's profile. We also use a database to record all the extracted knowledge.

Since we inevitably will obtain too many variables, we also use operators from the widely utilized methods from statistic analysis, in particular Principal Component Analysis (PCA) in order to reduce the number of variables. After this step the resulting PCA components are used to train a Fuzzy Neural Network, see (figure 1) and [1,8].

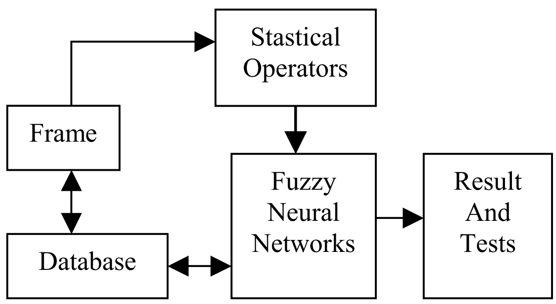

Figure 1: System Components

2 Frames

In the field of IA "Frame" refers to a special form of representing stereotyped concepts and situations. Attached to each frame there are different types of information. Some of this information is about how to use the frame. Some are about what we can expect if the information contained in the frame is confirmed. Other is about what to do if what we expect about the frames is not confirmed.

The basic structure of a frame is:

Title: name of the object represented by the frame.

Properties: features or attributes that describe an object. They can be static or dynamic.

Values: each property has a "slot" for the input of a specific value (Boolean, numerical or alphanumeric). The properties and its values form a list of declaration of the type object-attribute-value that represent an object.

Class: a value that is a name of other related object.

We use frames in our case because the previous knowledge obtained by an economist analyst is very frequently a stereotyped object and varies very little.

Personal Information
Age: 36 C. S.: married House: yes Car: 2 ...

Financial Goals
Capital Preservation: Yes Profit Level: Medium Short term goals: Investment time:

Investment Term
Long Term: Yes Medium Term: No Short Term: No

Market Knowledge:
International Markets: good Stock Market: good …

Figure 2: Examples of frames representing the Investor Profile

Frames were created by Minski [7] for recognizing automatically stereotyped objects. The problem tackled by frames is the same tackled by semantic networks. The problem that they try to solve is the problem of concept formation. Concept formation is directly linked to the problem of decision-making and information. As appointed by Suppes [10], if we examine the structure of decision theory we find that there is no place for the formation of new concepts by the decision-maker. The important thing we want to emphasize is that the theory of the decision-maker provides no place for the decision-maker to acquire new concepts on the basis of new information received. The theory is static in the sense that it is assumed that the decision-maker has a fixed conceptual apparatus available to him through out the time.

Frames and semantics networks fit perfectly the theory of the decision-maker. The concepts that define the objects are completely formal. The problem for the frames is to identify these objects, what means completely define them. The human being normally identifies the objects, excepts when they are new for them. But, they are capable of learning about them. The use of automatic means of forming concepts i.e. identifying objects is a extremely complicated problem. Minsky used to say that his frustration was not being capable to distinguish a dog from a cat. Frames do not permit to do so.

Frames work well in a micro world. A micro world is exactly what we have to tackle in this paper. Using frames to define an investor and enterprise profiles is perfectly appropriated. We consider that the world of investors and enterprises is fixed, static. Once we define them in terms of frames, our problem is to put them in the memory of the machine and from that in a database. Since all the frames have a label, recover them from memory or a database is not a difficulty matter.

Since we do not have an open world we don't have to distinguish a cat from a dog, for example.

There are, it seems to us, two important ways in which concept formation enter in the making of actual decisions. The first kind of modification in the decision structure, which may be introduced by concept formation, is a relatively straightforward refinement, or at least a modification of the initial partition of possible states of a micro-world by the consideration of additional concepts. The consideration of these additional concepts is almost always brought about by the reception of a cue or stimulus resulting from a new observation, for example, an investor with attributes that are not expected. The essential thing however in this kind of modification is that the concepts newly introduced are already a part of the conceptual apparatus of the decision-maker.

The second way in which the concept formation modifies the decision structure is the genuine case of concept formation properly. In this instance the decision-maker actually forms a concept he did not previously have in his repertory. But in our case this actually does not happen.

Taking in consideration this argumentation we consider the use of frames adequate to the problem we have in hand here.

3 Input Data and Sector Choice

In this work, we decided to select a short-term production sector. The cycle of manufacturing/ selling on short-term sectors is around six months. Considering the number of semesters available at our database, a short-term sector has more cycles of high-low prices. Some of the short-term sectors are: Textile, Food, Car Parts, Drink-Tobacco, Electric-Electronic and Commerce. In this work we have not trained the fuzzy neural network in all these sectors, nor tested companies from other sectors. However, we tested the resulting fuzzy neural network [8]against companies not included in the training set.

All the information about the companies will be represented by frame allowing the automatic selection of the sector.

At this stage of the research the selection was on an ad-hoc basis. After searching among all the short cycle sectors available on the database, we chose one that has more companies, more data and data that are more consistent. Following these criteria, textile sector was selected. This sector has 28 companies listed on our database. Table 1 shows the sectors analyzed in order of importance.

Sector	***Nº of Companies***	***Data since 1986***	***Performance on the stock market (%)*** [1]	***Appreciation rank among all 14 sectors***
Textile	28	Almost Complete	63,99	2
Food	21	Almost Complete	18,01	7
Electric-Electronic	8	Almost Complete	27,07	5
Commerce	13	Incomp.	33,00	4
Car-Parts	13	Incomp.	15,66	8
Drink-tobacco	2	Incomp.	15,45	9

(1)Sirotsky & Associados (http://www.sirotsky.com.br/news.html)

Table 1. Rank of Sectors

4 Economic Indicators by using previous knowledge of economic analylists

After choosing the sector we selected which economic indicators would be used as input to the fuzzy-neural network. The database has 52 economic indicators generated automatically from the balance sheets of the companies traded on the São Paulo stock market. The database also has indicators such as Brazilian inflation rate, reference interest rate, exchange rate, etc. So we decided to reduce this number in order to speed up the training process. In order to select the economic indicators we used the PCA. This selection would result from a statistical analysis of the database, therefore creating an automatic system from scratch.

5 Utilizing PCA in Case Study

This approach used the PCA to select the economic indicators that reduced to ten the number of indicators. This method transforms a set of correlated variables to a new set of uncorrelated variables. The PCA [5,11] finds components that are close to the original variables but arranged in decreasing order of variance. In order to illustrate the method, consider $\boldsymbol{X}^T = [X_1, X_2, \ldots, X_p]$ a p-dimensional random variable with mean μ a covariance matrix Σ. PCA transforms $\boldsymbol{X}^T$ in $\boldsymbol{Y}^T = [Y_1, Y_2, \ldots, Y_p]$ where each Y_j is a linear combination of the X's, so that

$$Y_j = a_{1j}X_1 + a_{2j}X_2 + \ldots + a_{pj}X_p$$

After applying PCA to the data, we identified ten components. These ten components retained almost 65% of the total variance of the sample as is indicated in the Figure 3. The PCA was applied to data from twenty companies from the textile sector, which has twenty-eight companies. These twenty were the companies that have more consistent data.

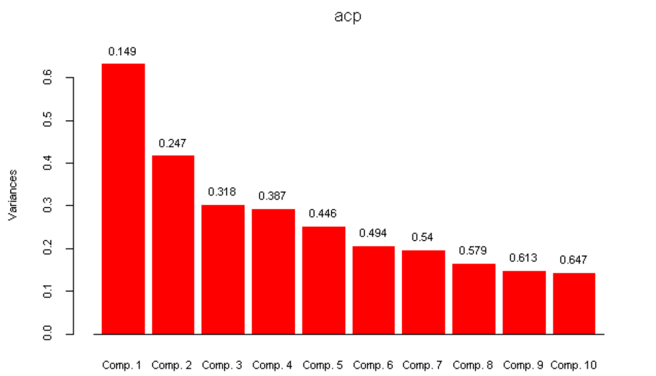

Figure 3. PCA results

6 Fuzzy-Neural Model

The fuzzy-neural model used is a feed-forward architecture with five layers of neurons as indicated in the Introduction section. It maps a fuzzy system [6] to a neural network [2,3,4] that will simulate the inference process executed in the fuzzy system. The first layer of the fuzzy neural system receives input values and feeds them to the second level, so it has ten inputs. The second layer determines the degree of membership of each variable to the fuzzy sets to which it belongs. The third layer represents the fuzzy rules that will combine the input variables using rules of the type **if-then**. In the next layer, each node will represent one fuzzy set from the consequent elements of the rules, the output variables. The last layer executes the process of defuzzification, yielding an exact value for each output variable. Several tests were realized to find the best architecture and the resulting numbers of neurons in each layer are 9 or 10, depending on the number of inputs, 3, 20, 3 and 3. The three outputs are: keep, buy and sell.

7 Training and Testing

In order to train the network, we divided the data into two sets, one training set and one test (recall) set. It is important to notice that the problem has two characteristics. First it behaves like a temporal series, since we have data, collected every trimester since 1986, from balance sheets of 28 companies. Second we want to classify companies into three groups: keep, buy and sell companies. Taking into account these characteristics, at the beginning of the training process, we separated the data into 35 trimesters to training and the last 8 to testing. Some trimesters were left out because not all data were available.

In order to check if the network was producing correct answers the target was defined according to following criteria. Whenever the stock price increased more than 5% from one trimester to another, the network had to indicate a buy option at the beginning of this trimester. If the price went down the same percentage then the answer had to be sell. A stable price indicated that the stock had to be kept.

The goal of the research is to produce a network that will give an indication of the best option until the next balance is released in three months time. After this period the network is retrained including the new released data. Retraining of the network is not a problem due to the tree month interval in between samples. Another test made was to check if the network was able to give meaningful answers two trimesters in advance.

We used a windowing like system [9]. At first a window of 35 trimesters with a step of one trimester was moved through the data. So the network was trained using the first 35 trimesters, after that an evaluation of the results and the window moved to the right one trimester after a retraining. This process went through the next four trimesters.

One interesting outcome of the training process was the possibility of using a window with fewer trimesters. A window of 14 trimesters gave the same performance, with the additional advantage of reducing the training time. This reduction shows that economic information older than three and half years was not relevant to our solution. This may be due to the very unstable situation of the Brazilian economy in the last two decades. It would be interesting to test on data from other countries.

The data set was composed of data from 20 textile companies spawning a period 43 trimesters. Some companies were left out on purpose so that a test on companies not trained could be performed later. The percentage of each target in the data set is shown in Table 2. Each line shows the number of targets at each trimester window. It should be noted that the figures for the three targets are not evenly distributed and this will show up at the network performance as it will be showed later. Remember that we are using real data extracted from balance sheets of companies listed at São Paulo stock market and they show the state of the economy during the period considered. First, it is important to note that buy targets are more frequent, showing that companies were growing most of the time. Another characteristic is the low frequency of the keep targets showing few periods of stability.

Window last trimester	Buy	Buy %	Sell	Sell %	Keep	Keep %	Total
35º (06/96)	160	59,3	79	29,3	31	11,5	270
36º (09/96)	148	54,2	91	33,3	34	12,5	273
37º (12/96)	141	51,5	97	35,4	36	13,1	274
38º (03/97)	134	48,7	103	37,5	38	13,8	275
39º (06/97)	120	43,5	113	40,9	43	15,6	276

Table 2. Percentage of targets at each training stage.

Table 3 shows the results obtained by the network generated from the Principal Component Analysis and this will be referenced as PCA network. The last column (Not Class) is the percentage of not classified inputs. These numbers were obtained after testing the 20 companies during 5 trimesters. Note that the PCA network obtained good results. Another contribution to these results may be due to the fact that the PCA method takes into account all economic indicators when creating the linear combination. Table 3 shows that the results for the buy and sell targets were reasonable. As for the keep target the networks did not perform well due to the low percentage of samples of this kind at the beginning of the considered period.

Partial results showed that when only the last two time windows were considered, this percentage is higher and the performance improves. Table 4 shows the results for the tests of 39th and 40th trimesters respectively. Another conclusion is that the network has not enough information to predict two semesters ahead.

Prediction (%)	Total	Buy	Sell	Keep	Not Class
Last trained trimester	100	100	100	100	0
One trimester ahead	75.0	77.4	77.5	58.3	6.1
Two trimesters ahead	53.6	72.7	59.5	31.2	2.0

Table 3. Average prediction capabilities of the PCA network.

Prediction (%)	Total	Buy	Sell	Keep	Not Class
39th	75.0	75.0	80.0	66.7	15.0
40th	73.7	100.0	68.8	100.0	5.0

Table 4. PCA network results for the 39th and 40th trimesters.

Table 5 shows the results of the tests applied to 2 companies not in the set of trained companies. The results were similar to obtained from trained companies indicating that both networks have good generalization capacity.

Prediction (%)	Total	Buy	Sell	Keep	Not Class
One trimester ahead	77.8	66.6	100.0	50.0	10.0

Table 5. PCA network results for 2 not trained companies.

8 Conclusions

The introduction of the concept of previous knowledge and frames in this case study shows that it is possible to build an environment where the investor and company profiles are generated semi-automatically. It is also possible to create automatically the investor portfolio. Frames can contribute to find the best allocation of investment.

_This study gave rise to important points regarding the suitability of applying fuzzy neural networks for long-term prediction about stock prices using indicators from fundamentalist analysis. First, it indicated that it is possible to expect good performance, especially after the observation that their results were considerably improved by the incorporation of data that have the numbers of targets evenly distributed. The second one is that, it is rather implausible to expect that the networks could provide acceptable predictions of two semesters in advance. Nevertheless, this point is under consideration and further investigations are under way.

Another point is that the experiments showed that the network performance was not reduced when the window of time was reduced drastically, indicating that old information lost significance as the economic factors changes.

The network presented similar results when trained and not trained companies are considered showing capacity of generalization, as shown in [1].

The concepts of Previous Knowledge and frames were introduced in order to improve the results of the completed model.

References

[1] A. O. Cruz, R. Raposo. *Stock Market Prediction Based on Fundamentalist Analysis with Fuzzy-Neural Networks,* 3rd WSES FSFS, Interlaken, Switzerland, 2002.

[2] E. M. Azoff. *Neural Network Time Series Forecasting of Financial Markets*, John Wiley & Sons Chichester, England, 1994

[3] G.M. Papadourakis, G. Spanoudakis and A. Gotsias. Application of Neural Networks in Short-Term Stock Price Forecasting, *Proceedings of 1st International Workshop Neural Networks in the Capital Markets*, London, November 1993.

[4] I. Matsuda. Application of Neural Sequential Association to Long-Term Stock Price Prediction. *Proceedings of IJCNN'91*, Singapore, 1991

[5] K. Sam. *Multivariate Statistical Analysis*, Radius Press, pp. 236-260.

[6] L. Zadeh. *Fuzzy Sets, Inf. Control*, vol. 8, pp. 338-353.

[7] M. Minsky. *A Framework for Representing Knowledge, in The Psychology of Computer Vision,* P. Winston, McGraw-Hill, New York, USA, 1975

[8] N. Kasabov. *Foundations of Neural Networks, Fuzzy Systems and Knowledge Engineering*, MIT Press, Cambridge, MA, USA, 1996

[9] R. R. Trippi, E. Turban. *Neural Networks in Finance and Investing*, Probus Publishing Company, Chicago, Illinois, USA, 1993.
[10] P. Suppes, *Concept Formation and Bayesian Decisions in Aspects of Inductive Logic*, ed. J. Hintikka and P. Suppes, North Holland Publishing Company Amsterdam, 1966
[11] W. R. Dillon, M. Goldstein. *Multivariate Analysis: Methods and Applications*, John Wiley & Sons, New York, USA, 1984.