A Neuro Solution for Economic Diagnosis and Prediction

Nicolae Morariu
”Ştefan cel Mare” University of Suceava, Romania
nmorariu@usv.ro

Abstract:
The paper present a solution for the economic activity evolution diagnostic and prediction by means of a set of indicators. Starting from the indicators set, there is defined a measure on the patterns set, measure representing a scalar value that characterizes the activity analyzed at each time moment. A pattern is defined by the values of the indicators set at a given time. Over the classes set obtained by means of the classification and recognition techniques is defined a relation that allows the representation of the evolution from negative evolution towards positive evolution. For the diagnostic and prediction the following tools are used: pattern recognition and multilayer perceptron implemented in the REFORME software written by the author and the results of the experiment obtained with this software for macroeconomic diagnostic and prediction during the years 2005-2012 for diagnostic and 2013-2014 for prediction.

Keywords: pattern recognition, neural network, multilayer perceptron, indicators, diagnostic, prediction.

1. Introduction

The assessments of the development level for a specific activity can be carried out by using the analysis of the evolution of the indicators describing both the quantitative level as well as the qualitative mutations in time. The problems related to diagnostic and prediction are solved using pattern recognition techniques and the multilayer perceptron [14], implemented by the REFORME software [4].

The data corresponding to the evolution in time, for the activity considered, are processed using the methods already mentioned, methods assessing the overall evolution trend of the indicators. The output is a qualitative variable (classes or D(x)) representing the result of the assessment. The conceptual diagram for diagnostic, prediction and graphical representation is presented in figure 1.

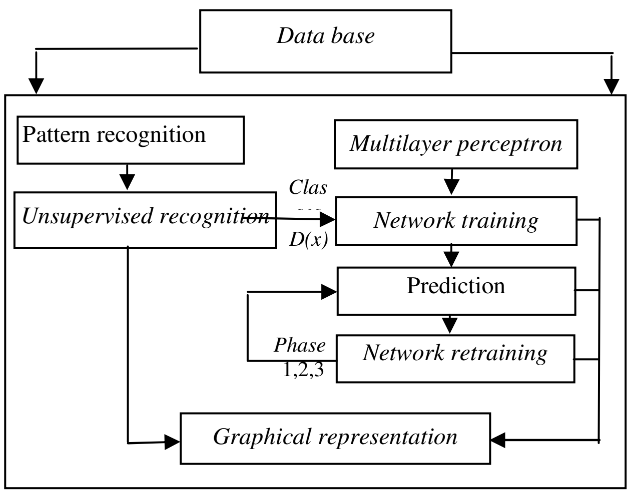

Figure. 1. The conceptual diagram for diagnostic and prediction

The database includes the data referring to the specific indicators and the achievements per indicators organized as tables. A minimal database with a high degree of flexibility includes the tables INDICATORS and VALUES having the structure described below:

INDICATORS		VALUES	
Codeind	indicator code	Codeind	indicator code
Nameind	indicator name	Date	report date
Umind	measurement unit	Value	indicator value

2. Pattern Recognition Techniques

If we note x_1, x_2, … , x_n the characteristics set and x is a form defined by these characteristics, then the x form can be considered as a vector $x(x_1, x_2, \dots, x_n)$. The n features of the vector x are subjected to different processing methods: data normalizing, linear and nonlinear transformations, reducing and pattern selection, [3],[6],[7].

A supervised recognition problem can be formulated as follows: Given a partition in classes of known forms, a new form must be included in one of these classes taking into account the existing classes.

The unsupervised recognition consists of dividing the space forms in classes using algorithms. The simplest grouping algorithm, named the threshold algorithm is based on the initial determination of a minimum distance between two forms, the threshold distance. According to the algorithm mentioned above, two forms belong to the same class if the distance between them is smaller than the threshold distance. Using the threshold algorithm, the value of the threshold establishes the number of classes and the classification process can be reapplied for different values of the threshold until the best classification is obtained.

If the values of the features are of different magnitude, then the features with high absolute values will have a greater influence over the classification results and the values of the features must have the same order of magnitude. The method frequently used is the domain adjusting method, the following transformation of the features values being applied:

$$x_{i,new} = \frac{\left(x_{i,old} - x_{i,min}\right)}{x_{i,max} - x_{i,min}} \qquad (1)$$

If P_1, P_2,…,P_m are the reference patterns (the prototypes) and C_1,C_2…,C_m are the corresponding classes, then the minimum distance classifier will assign the input pattern X to the class C_i if the distance $d = |X - P_i|$ is minimum.

The most used distances are derived from the general Minkovski distance:

$$d_{Minkovski} = \left[\sum_1^n (x_i - p_i)^k\right]^{1/k} \qquad (2)$$

For k=2, the Euclidean distance is obtained:

$$d_{Euclid} = \left[\sum_1^n (x_i - p_i)^2\right]^{1/2} \qquad (3)$$

For the determination of the class to which a pattern belongs, starting from a training set with known classification, one of the most frequently used method based on the minimum distance principle is the nearest neighbour method.

Let $F = \{f_1, f_2, \dots, f_n\}$ be a set of training patterns and C_1, C_2,…,C_p the classes in which the set F was divided using a classification algorithm. The rule of the nearest neighbour can be mathematically wrote as:

$$if\ d(f, f_a) \le d(f, f_k), k = 1,2 \dots n\ and\ f_a \in C_i\ then\ f \in C_i$$

The pattern recognition techniques are implemented in the REFORME software [4]. The module "pattern recognition" of the REFORME program has the following tasks:

- normalizes the inputs by means of the domain adjusting method;
- classifies unsupervised using the threshold algorithm with a classifier minimum distance based for the type of distance selected (Euclidean, Manhattan, Hamming) and specified threshold value. The classes resulted will be numbered from negative evolution toward positive evolution using an algorithm that induces the relation "<" over the classes set.
- determines to which class belong an unknown pattern and also determines the pattern at the minimum distance from the unknown pattern by means of the nearest neighbour rule.

3. Neural networks

An artificial neural network with an input layer and an output layer divides the input vectors in two semi plans. Solving complex problems implies the need of complicated decision regions, problem that can be solved by using networks with one or many extra layers between the input and the output layer [1], [10], [12].

Such a network can be treated as a generalization of the perceptron, being known as the multilayer perceptron (MLP) presented in Figure 2.

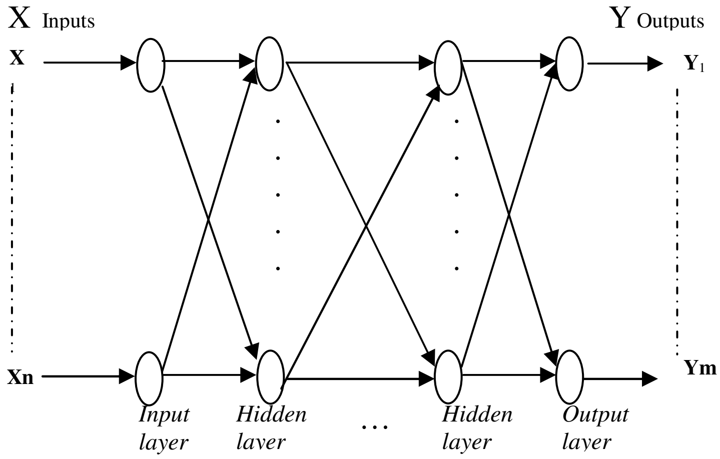

Figure 2. The multilayer perceptron architecture

For the architecture presented above, the input layer is made up from linear neurons whose goal is not to compute but to propagate the input signal to the corresponding neurons on the next layers. Excepting the neurons on the input layer, for each neuron is computed the activation (the sum of products between the inputs and weights) and then the activation function is applied obtaining the output of the neuron. The output of the previous neuron becomes input for the neuron from the next layer.

When chaotic time series are involved, prediction is a difficult problem and can be viewed as temporal pattern recognition task, for which purpose neural networks suit very well. The predicted value x(tk+1) of a variable x at a future time t k+1 is based on k previous values x(t1), x(t2), …, x(tk). Figures 3, 4 shows neural network structures for univariate and multivariate prediction [2]:

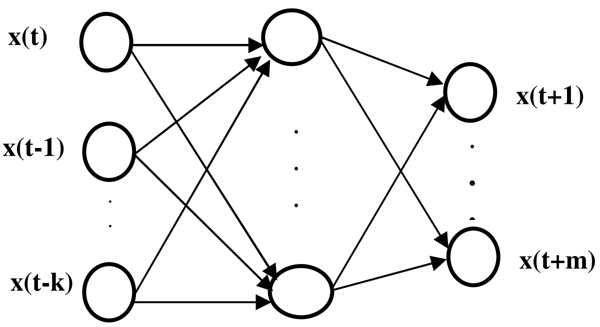

Figure 3. Connectionist univariate time-series

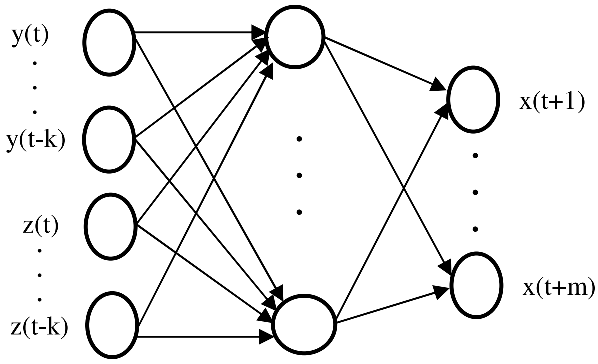

Figure 4. Connectionist multivariate time-series

The dependent variable subject to prediction can be different from the past data variables (independent variables), but both are on the time scale.

The multivariate prediction implies the prediction of both dependent variable x and the prediction of the independent variables y, z at the moment t+1 starting from their precedent values. In this case the prediction of the variables y and z can be achieved with a neural network designed for the univariate prediction.

The backpropagation method and the sigmoidal activation function are the most important and used method for multilayer feed forward neural networks training that minimizes the mean squared error using the gradient method.

4. Graphical representation

A graphical representation of the evolution of the activity in a coordinate system xOy can be obtained as follows: each pattern represents a point on a plane, the axis x being the time period corresponding to the pattern and y is the class that the pattern belongs to as shown in Figure 7.

The accuracy of the representation depends on the number of classes. This representation implies first to sort and then renumber the classes taking into consideration the evolution direction (from negative evolution towards positive evolution).

To order the classes, for each class c is assigned a number, M(c), that is calculated as presented below.

Let x(x1, x2,…,xn) be a pattern that belongs to the class c, with normalized parameters xi. For each parameter xi of the pattern x is assigned a weight pi representing the importance of the parameter xi (the weights can be computed for instance as the partial correlation coefficients or can be set up by the expert).

For the pattern x, D(x) is calculated using the equation:

$$D(x)=\sum_{i=1}^{n} p_i x_i \tag{4}$$

D(x) represents a measure of the activity described by the pattern x.

For each pattern belonging to the class c the similar method is carried out, M(c) being defined as:

$$M(c)=\left(\sum_{x\in c} D(x)\right)/\,p \tag{5}$$

where p is the number of forms belonging to class c.

The class c1 is in relation "<" toward the class c2 if M(c1)<M(c2).

An order relation over the class set has been defined. Renumbering the classes taking into account this order relation, a plot can be drawn representing the evolution of the activity analyzed as shown in Figure 7 and Figure 9.

Given two time intervals t1 and t2, t1<t2 and Dt1(x), Dt2(x):

- If Dt1(x)<Dt2(x) then the activity defined by the patterns x has a positive evolution at the moment t2 toward t1;

- If Dt1(x)=Dt2(x) then the activity defined by the patterns x is stationary at the moment t2 toward t1;

- If Dt1(x)>Dt2(x) then the activity defined by the patterns x has a negative evolution at the moment t2 toward t1.

A similar interpretation can be done considering M(c). A much accurate representation of the activity evolution in a coordinate system xOy can be obtained as follows: each pattern represents a point in plane, x being the time range that corresponds to the pattern and y is D(x) as shown in Figure 8.

Macroeconomic forecasting

We consider the next macroeconomic indicators:

CODE	**INDICATOR**	**U/M**
I1	**Gross Domestic Product**	**million RON**
I2	**Goods export (average)**	**million EURO**
I3	**Goods import (CIF)**	**million EURO**
I4	**Goods import (FOB)**	**million EURO**
I5	**Occupational populations (average)**	**thousands pers.**
I6	**Average number of employees**	**thousands pers.**
I7	**Gross Income (average)**	**RON**
I8	**Number of unemployment (the ending of the year)**	**thousands pers.**

and the coresponding values [9]:

Year	I1	I2	I3	I4	I5	I6	I7	I8
A-2005	288955	22255	32569	30061	8152.5	5921	968	523.0
A-2006	344651	25850	40746	37609	8245.1	6167	1146	460.5
A-2007	412762	29549	51322	47371	8447.2	6197	1396	367.8
A-2008	503959	33628	56337	52000	8476.0	6298	1730	403.4
A-2009	531251	28400	42040	38800	8300.0	6095	1820	620.0
A-2010	568500	29950	43230	39900	8370.0	6125	1885	590.0
A-2011	612000	31800	45290	41800	8445.0	6175	1970	550.0
A-2012	664700	34400	48290	44570	8540.0	6235	2075	520.0
A-2013	725100	37600	51870	47880	8640.0	6305	2190	500.0

Taking into consideration the specific problems regarding macroeconomic forecasting, the indicators represent the parameters of the pattern and a pattern X_t is defined by the values of the parameters for an year.

For diagnostic is considered the interval between 2005 and 2012, and for prediction the interval 2013-2014.

Each row of the datasheet represents a pattern and characterizes the degree of economical development for the year considered.

After the data normalization, classification (unsupervised recognition using the threshold algorithm) with REFORME software, the distance used being the Euclidean distance and the threshold value =1., results the division of the patterns into 4 classes, the ordering and the renumbering of the classes as well as the representation of the plot for the evolution using the two solutions (classes, D(x)). The results obtained are presented in the figures 5, 6,7.

B1 fx (A=Learning, T=Test evaluation, P=Recognition, AT=Learning+Test, X=Unutilis

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	ID_Pattern	(A	I1	I2	I3	I4	I5	I6	I7	I8	Class	D(x)	M©
2	A-2005		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.62	1	0.62	1.50
3	A-2006		0.13	0.23	0.34	0.34	0.19	0.64	0.15	0.37	1	2.39	
4	A-2007		0.28	0.48	0.79	0.79	0.60	0.72	0.35	0.00	2	4.01	4.82
5	A-2008		0.49	0.74	1.00	1.00	0.66	0.98	0.62	0.14	2	5.64	
6	A-2009		0.56	0.40	0.40	0.40	0.30	0.45	0.70	1.00	3	4.21	5.05
7	A-2010		0.64	0.50	0.45	0.45	0.45	0.53	0.75	0.88	3	4.65	
8	A-2011		0.74	0.62	0.54	0.54	0.60	0.66	0.82	0.72	3	5.24	
9	A-2012		0.86	0.79	0.66	0.66	0.79	0.82	0.91	0.60	3	6.10	
10	A-2013		1.00	1.00	0.81	0.81	1.00	1.00	1.00	0.52	4	7.15	7.15
11													
12	Minim		288955	22255.00	32569.00	30061.00	8152.50	5921.00	968.00	367.80			
13	Maxim		725100	37600	56337	52000	8640	6305	2190	620			
14													

Figure 5. The result after the data normalization, classification (unsupervised recognition using the threshold algorithm)

ID_Pattern	Class
A-2005	1
A-2006	1
A-2007	2
A-2008	2
A-2009	3
A-2010	3
A-2011	3
A-2012	3
A-2013	4

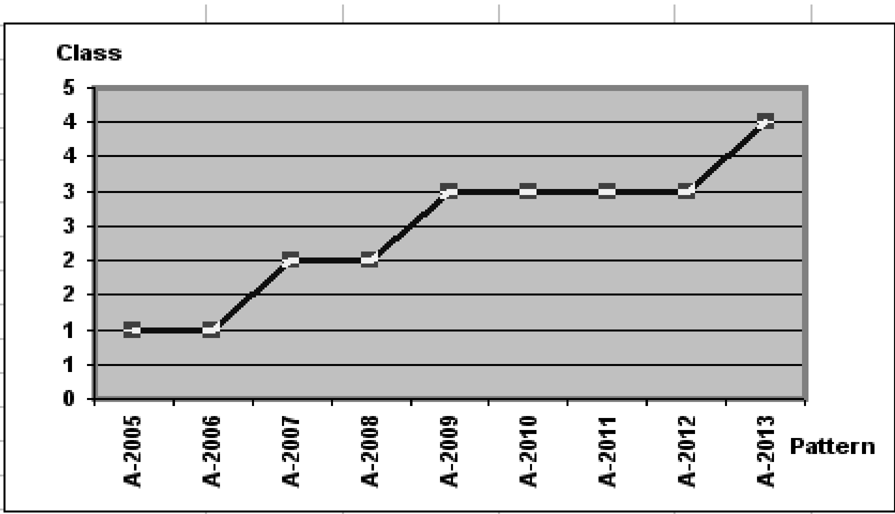

Figure 6. Graphical representation of analyzed activity using four classes
M(1)=1.50 M(2)=4.82 M(3)=5.05 M(4)=7.15

ID_Pattern	D(x)
A-2005	0.62
A-2006	2.39
A-2007	4.01
A-2008	5.64
A-2009	4.21
A-2010	4.65
A-2011	5.24
A-2012	6.10
A-2013	7.15

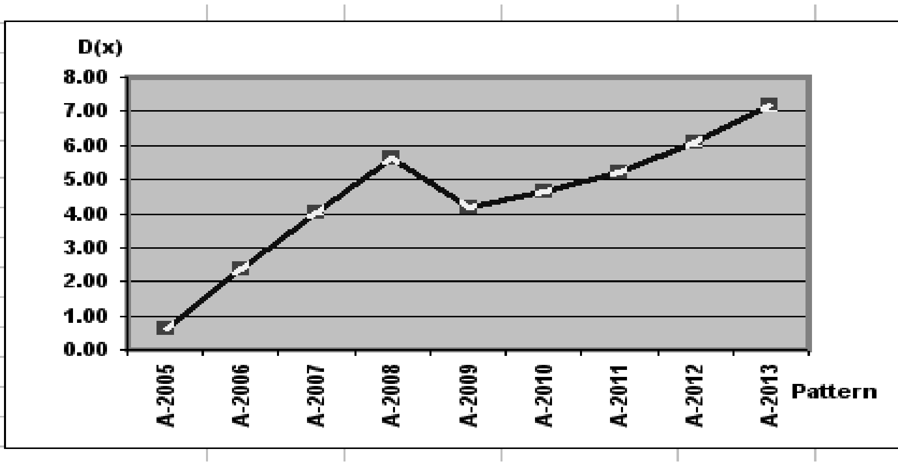

Figure 7. The graphical representation of analyzed activity using D(x)

For a better approximation one solution would be the design of specific models for nonlinear systems [2], [5], [8].

For the prediction of the values corresponding to the next
year the following neural architecture is defined: 8 neuron on the input layer, 8 neurons on the hidden layer and one neuron on the output layer. The network architecture is presented in Figure 10.

For training of the network uses the inputs for time t and the t+1 output is obtained. The results obtained for the training network (patterns A-2005 … A-2011) and for prediction (patterns A-2012, A-2013) for the next years (2013, 2014) are presented in the figure 11.

The training process is carried out in two phases as well presented in the figures 11, 12. The training procedure begins with the training for a number of years and then the prediction for the next years is carried out (figure 11).

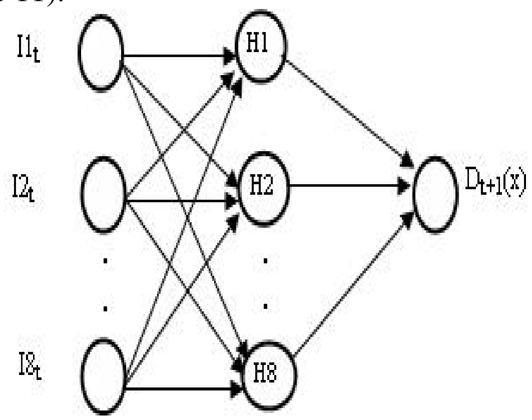

Figure 10. The network arhitecture for the next year prediction

The training is then restarted including in the training set the values predicted and the next prediction for the next interval or year is carried out (figure 12).

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	ID_Pattern	A=Le	I1	I2	I3	I4	I5	I6	I7	I8	Required D(x)t+1	Calculated D(x)t+1	Predicted D(X)t+1		
2	A-2005	A	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.62	2.39386	2.39426			
3	A-2006	A	0.13	0.23	0.34	0.34	0.19	0.64	0.15	0.37	4.01072	4.01142			
4	A-2007	A	0.28	0.48	0.79	0.79	0.60	0.72	0.35	0.00	5.64420	5.64628			
5	A-2008	A	0.49	0.74	1.00	1.00	0.66	0.98	0.62	0.14	4.20571	4.20756			
6	A-2009	A	0.56	0.40	0.40	0.40	0.30	0.45	0.70	1.00	4.64829	4.64798			
7	A-2010	A	0.64	0.50	0.45	0.45	0.45	0.53	0.75	0.88	5.23687	5.24011			
8	A-2011	A	0.74	0.62	0.54	0.54	0.60	0.66	0.82	0.72	6.09771	6.09848			
9	A-2012	P	0.86	0.79	0.66	0.66	0.79	0.82	0.91	0.60	7.14845		7.07005		
10	A-2013	P	1.00	1.00	0.81	0.81	1.00	1.00	1.00	0.52			7.71113		
11			Avg.error:, 9.41659519142369E-07, Learning rate = 0.3, No. Epochs = 6000, No. hidden neurons = 8												

Figure11. Result of the network training and prediction – phase 1

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	ID_Pattern	A=Le	I1	I2	I3	I4	I5	I6	I7	I8	Required D(x)t+1	Calculated D(X)t+1	Predicted D(X)t+1		
2	A-2005	A	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,62	2,39386	2,39327			
3	A-2006	A	0,13	0,23	0,34	0,34	0,19	0,64	0,15	0,37	4,01072	4,00894			
4	A-2007	A	0,28	0,48	0,79	0,79	0,60	0,72	0,35	0,00	5,64420	5,64182			
5	A-2008	A	0,49	0,74	1,00	1,00	0,66	0,98	0,62	0,14	4,20571	4,20410			
6	A-2009	A	0,56	0,40	0,40	0,40	0,30	0,45	0,70	1,00	4,64829	4,64452			
7	A-2010	A	0,64	0,50	0,45	0,45	0,45	0,53	0,75	0,88	5,23687	5,23764			
8	A-2011	A	0,74	0,62	0,54	0,54	0,60	0,66	0,82	0,72	6,09771	6,09253			
9	A-2012	A	0,86	0,79	0,66	0,66	0,79	0,82	0,91	0,60	7,14845	7,14641			
10	A-2013	P	1,00	1,00	0,81	0,81	1,00	1,00	1,00	0,52			8,07727		
11			Avg.error:.3.50290207903551E-06. Learning rate = 0.3. No. Epochs = 9000. No. hidden neurons = 8												

Figure 12. The results of network retraining and prediction - phase 2.

6. Conclusions

The paper describes a method that quantifies and represents the evolution of the economic activity. Starting from a set of indicators that characterizes the economic activity, are defined a scalar D(x) that allows to measure an activity described by the x patterns and an order relation "<" over the set of classes resulted using unsupervised classification techniques for the graphical representation of the activity evolution using classes.

One of the methods used for solving diagnostic and prediction problems is the regression analysis. In the case of linear systems the regression approach has practical relevance. Because of their ability to detect nonlinear dependences in the input data set, neural networks represent an efficient alternative to the existing methods.

The obtained results confirm that prediction for a shorten period of time (1 year) is exactly, but for a long period of time the results are approximately.

In order to achieve accurate long-term prediction, a short-term prediction is carried out first and then a multiple phase retraining of the neural network is done.

References

[1]. Abdi Herve, Les Reseaux de Neurones, Universite de Bourgogne,1993.

[2]. Nikola K. Kasabov, Foundations of Neural Networks, Fuzzy Systems, and Knowledge Engineering, A Bradford Book, The MIT Press, Cambridge, Massachusetts, London, England, 1998, pp 393-397

[3]. Menahem Friedman, Abraham Kandel Introduction to pattern recognition statistical, structural, neural and fuzzy logic approaches, World Scientific, 2000

[4]. *Nicolae Morariu, "REFORME – A software product for pattern classification and recognition by joint use of pattern recognition techniques and multi-layer perceptron", "The Proceedings of the Central and East European Conference in Business Information Systems", Cluj-Napoca, may 2004, Ed. Risoprint, ISBN 973-656-648-X, pag.100-105.*

[5]. Sorin Vlad, "Neural Networks Applications-Chaotic Time Series Prediction", Distributed Processing, Suceava University Publishing House, Romania, 2005.

[6]. Ştefan-Gheorghe Pentiuc, "Aplicaţii ale recunoaşterii formelor în diagnosticul automat", 158 p. ISBN 973-31-1096-5, Editura Tehnică, Bucureşti, 1997.

[7]. Romul Vancea, Ştefan Holban, Dan Ciubotariu, "Recunoaşterea Formelor Aplicaţii", Ed. Academiei R.S.R. 1989.

[8]. Iulian Năstac, Emilian Dobrescu, Elena Pelinescu, "Neuro-Adaptive Model for Financial Forecasting", Romanian Journal of Economic Forecasting – 3/2007

[9]. Proiecţia principalilor indicatori macroeconomici pentru anul 2009-2013, Comisia Naţională de Prognoză, 3 mai 2006., www.cnp.ro/user/repository/prognoza_2009-2013.pdf

[10]. Daniel Graupe, "Principles of Artificial Neural Networks, 2nd edition", World Scientific 2007
[11]. Nicolas Carnot, Vincent Koen, Bruno Tissot, "Economic Forecasting", Palgrave Macmillan 2005.
[12]. Alexander I. Galushkin, "Neural Networks Theory", Springer 2007.
[13]. Nicolae Morariu, Sorin Vlad, "Using Pattern Clasification and Recognition Techniques for Diagnostication and Prediction", Advances in Electrical and Computer Engineering, Suceava, 2007
[14]. Morariu, N; Iancu, E; Vlad, S, "A neural network model for time-series forecasting", Romanian Journal Of Economic Forecasting, Volume 12 Issue 4 dec./2009, ISSN: 1582/6163, Pages: 213-223.