SIMULATION WITH ARTIFICIAL INTELLIGENCE TO FORECAST GDP DEPENDING ON LOGISTICS ELEMENTS

Margareta ILIE[1]
Norina POPOVICI[2]
Constantin ILIE[3]

ABSTRACT
The present paper as the main goal to establish a relatively new method to forecast gross domestic product depending on logistics elements using the artificial neural network simulation of the influence between the two types of data. In their research the authors determined the most appropriate form of neural network algorithm necessary for neural network training and they also analyzed the results obtained after training over the results of simulations. Also the authors intended to lay the foundations of using a non-linear methods that respond to the problems caused by rapid changes in the economic environment.
The research result was the building a feedforward type of artificial neural network with a 4-5-1 structure. The training has been tested and evaluated, yielding values of errors comparison between actual and simulated data by the neural network less than 0.5%.
The importance lies in the research results proving the efficiency of the method used in simulating the influence of elements of logistic activities over GDP values.

KEYWORDS: *neural network, GDP, simulation.*

JEL CLASSIFICATION: *C45*

1. INTRODUCTION

The ability to predict economic indicators influenced from a few economic activities is very appealing. Many risks, economic errors and financial crises of different sizes can be avoided or their negative effects can be minimized if they can be forecasted.
But seeing them generally involve complicated mathematical functions involving a large number of variables that cannot react with the same value change when sudden changes in size of economic indicators due to economic crises.
It would be advantageous and efficiently that one can predict the value of gross domestic product of a country based on some values that characterize logistics activity to that country. For examples one can chose can choose results of economic activities like transport or trade services or other indicators such as number of employees. Also must be determined and used a method that can provide conclusive, viable and reliable results for this kind of problem if possible lowering the limitations of classical mathematical functions.
This is what the authors of this paper suggests, by researching the possibility of forecasting the GDP from the influence of the following factors and using training and simulation offered by ANN.

[1] OVIDIUS University from Constanta, Faculty of Economic Sciences, Romania, ilie.marga@gmail.com
[2] OVIDIUS University from Constanta, Faculty of Economic Sciences, norinapopovici@yahoo.com
[3] OVIDIUS University from Constanta, Faculty of Mechanical, Industrial and Maritime Engineering, Romania, ilie.ctin@yahoo.com

2. METHOD

Considering the type of data, the heterogeneity of values and author's desire to use a method of modeling and simulation that does not directly involve computing mathematical functions between the data involved, and the best method was determined as the artificial intelligence branch called Artificial Neural Networks (ANN).
ANN have a long history of use from more than 70 years ago and are implemented in most human activities, from sports to NASA flights or chemical formulas and medical testing etc.
ANN method determine links between inputs and outputs, similar logical links are established by the human brain between different causal facts, situations or successive activities etc. It does not require mathematical functions defining influence between inputs and outputs, but it self organizes in learning and structure characteristics of situations, experiments and previous experiences.
ANN has a structure of an organization of neurons in the form of layers. Initial and final layer correspond respectively to input data and the output, while the layers between them are hidden layers. The hidden layers are network elements that acts like a black box (Udrescu M, 2009), in the sense that their structure is organized by ANN without being revealed to the user than at the end, after its finalization and without being altered by him in no time of processing.

3. RESEARCH

The data used are measured between 2004 and 2014 and represents statistical data of Netherlands and are the following:
Input data:
- ***Transportation and storage - Production value – TAS*** – Annual detailed enterprise statistics for services
- ***Trade balance [in mill. ECU/EURO] – TBA***
- ***International trade in services [Net balance at current prices (in % of GDP)] – ITS***
- ***Annual percentage change in total employed population [%] – APE*** – all persons who worked at least one hour for pay or profit during the reference week or were temporarily absent from such work, as presented by European Eurostat website.

Output data:
Gross domestic product at market prices [mill. Euro] – GDP – defined by the European Eurostat website is the final result of the production activity of resident producer units.
As you can see these data can be found in the position to directly or indirectly influence each other but there is not a mathematical function to calculate and accurately represent influences. The authors studied the simulation of influence of input (TAS, TBA, ITS, APE) on output data seeking to determine a way to be able to predict by simply reading the data input and using a trained ANN to determine the future values of GDP.
In order to perform modeling and simulation Alyuda NeuroIntelligence software was used and a laptop configuration was:
- ACER 16
- Processor: Intel® Core™ i5-4210M CPU @ 2.60Hz
- Installed memory: 3.98 GB
- System type: 64-bit Operating System, x64 based processor
- Operating System: Windows 8.1 Pre

The training and testing ANNs consist in several main steps: data analysis; data pre-processing; ANN design; testing and validation, as Hirose (2003) explained. Considering those steps the research was structured accordingly to them

3.1. Initial data and prerequisite

Data values used in research were registered against the European Union's Eurostat website and presented in Table 1 as initial data.

Table 1. Initial data values

Year	IN				OUT
	TAS	TBA	ITS	APE	GDP
2004		30,347	0.7	-1.2	523939.00
2005	58,320.0	34,201	1.1	0.7	545609.00
2006	64,326.3	37,271	1.4	2.2	579212.00
2007	69,275.7	42,458	1.6	2.9	613280.00
2008	72,406.7	38,742	1.5	1.6	639163.00
2009	63,574.4	39,244	1.0	-0.9	617540.00
2010	66,717.5	43,632	1.4	-0.7	631512.00
2011	71,424.2	48,898	1.5	0.9	642929.00
2012	73,283.0	53,274	1.3	-0.2	645164.00
2013	74,041.0	61,592	2.3	-0.9	650857.00
2014	74,500.0	63,635	2.2	-0.2	662770.00

Source: adapted from European Eurostat website http://ec.europa.eu/eurostat

3.2. Analysis

In order to determine characteristics to data network and future interaction with the processes which they will take place, the first step in using the network is developing a data analysis. The data analysis results are:

a) 5 columns and 11 rows analyzed;
b) 5 columns and 11 rows accepted for neural network training;
c) 5 numeric columns:
   - TAS;
   - TBA;
   - ITS;
   - APE;
   - GDP.
d) The output column: GDP

The data sets were divided into training, testing and validation. The data partition results are:

- 9 records to Training set (81.82%);
- 1 records to Validation set (9.09%);
- 1 records to Test set (9.09%).

3.3. Preprocessing

As concluded by Neagu and Ioniță (2004) in order to make data more suitable for neural network (for example, scaling and encoding categories into numeric values or binary) and improves the data quality (for example, filtering outliers and approximating missing values), they are pre-processed. The pre-processing used for the present research, considering the type of problem and the previous analysis, but also the researches, the pre-processing type used was numeric encoding. Numeric encoding means that a column with N distinct categories (values) is encoded into one numeric column, with one integer value assigned for each category. For example, as the Alyuda software help presents, for the Capacity column with values 'Low', 'Medium' and 'High', 'Low' will be represented as {1}, Medium as {2}, and High as {3}. The columns encoding and pre-processed data values are presented in table 2.

Table 2. Preprocessed columns of data encoding.

Parameters	TAS	TBA	ITS	APE	GDP
Column type	input	Input	input	input	output
Format	numerical	Numerical	numerical	numerical	numerical
Scaling range	[-1..1]	[-1..1]	[-1..1]	[-1..1]	[0..1]
Encoded into	1 columns	1 columns	1 columns	1 columns	1 columns
Min	58320	30347	0.7	-1.2	523939
Max	74041	63635	2.3	2.9	662770
Mean	68.152.088.889	44.844.909.091	1.38	0.381818	613.815.909.091
Std. deviation	4.526.652.034	10.346.892.915	0.383287	1.318.965	43.165.721.629
Scaling factor	0.000127	0.00006	1.25	0.487805	0.000007

Source: Our own simulation results.

3.4. Neural network structure

While the number of neurons in the input and output layers given actually by the number of input and output data, the number of neurons in the hidden layer is determined by various methods that are based on two versions: calculate the number of layers and the number of neurons in layer or by repeated attempts to determine most advantageous structure for training. Also, authors must define functions that characterize ANN neurons layers. These are presented in Table 3.
ANN type considered suitable for this type of problem is feedforward ANN with backwards data transmission in training - backpropagation training algorithm as Waszczyszyn (2000) defined.

Table 3. Network parameters.

Parameter	Value
Input activation FUNCTION	Hyperbolic tangent
Output name	Gross domestic product at market prices [mill. Euro]
Output error FUNCTION	Sum-of-squares
Output activation FUNCTION	Logistic

Source: Our own simulation parameters.

The final structure of the network is shown in Figure 1 as a three-layer 4 neurons in the input layer 5 neurons in the hidden layer neuron and one neuron in the output layer.

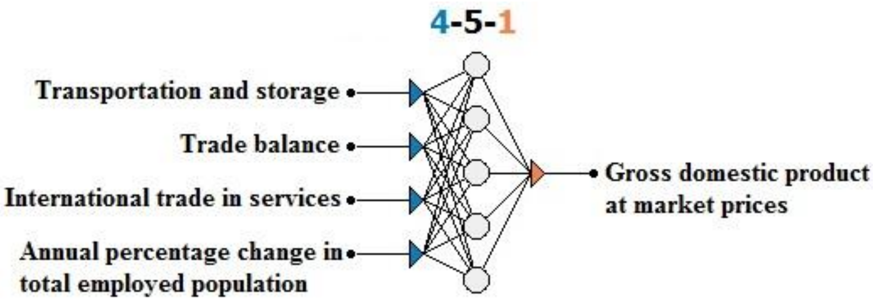

Figure 1. Neural network architecture
Source: Our own representation

3.5. Training

Training is the most complicated process of using ANN. The research authors choose the training algorithm and the speed and magnitude with which training is conducted by ANN. The algorithm for training the girl is Quick Propagation Algorithm.

Quick propagation is a heuristic modification of the back propagation algorithm invented by Scott Fahlman. This training algorithm treats the weights as if they were quasi-independent and attempts to use a simple quadratic model to approximate the error surface. In spite the fact that the algorithm has not theoretical foundation, it is proved to be much faster than standard back-propagation for many problems. The training report was:

- Training algorithm: Quick Propagation;
- Number of iterations: 13,580,626;
- Time passed: 01:02:09;
- Training stop reason: Desired error achieved.

The training results can be seen in the following figure.

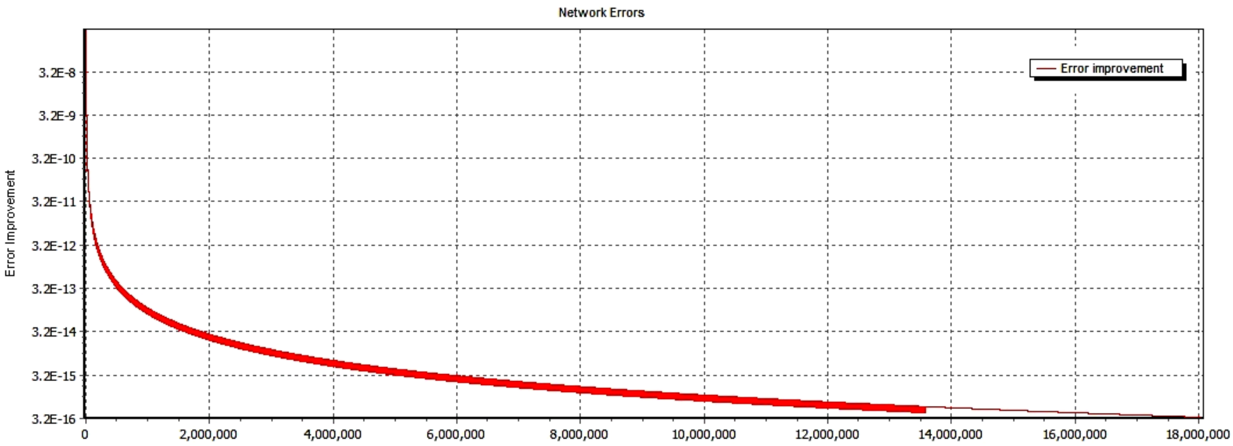

Figure 2. Network errors evolution in training process.
Source: Our own simulation with Alyuda NeuroIntelligence

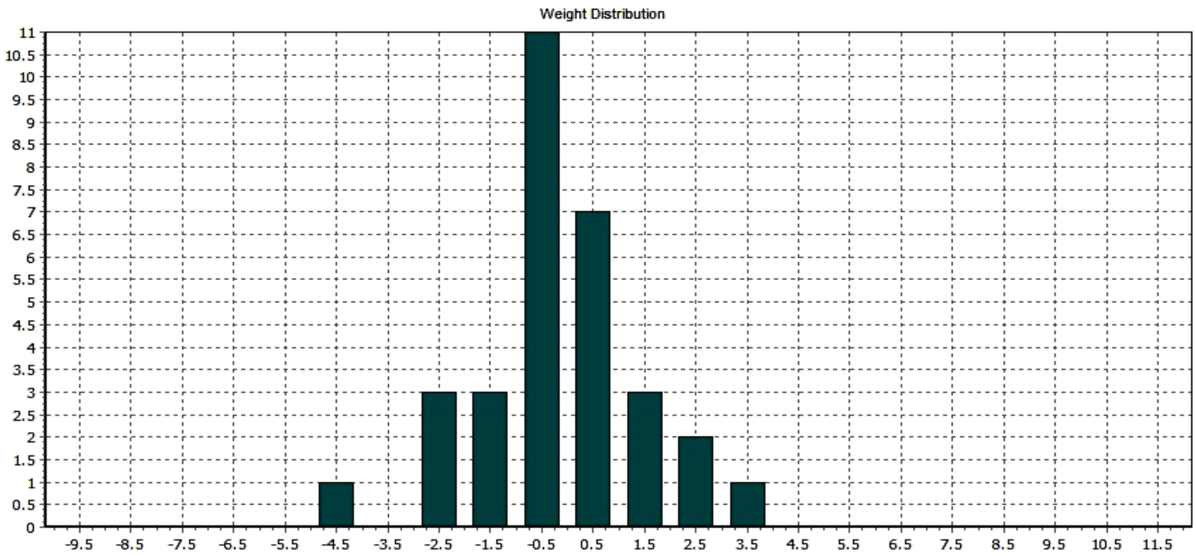

Figure 3. Weight distribution after training process.
Source: Our own simulation with Alyuda NeuroIntelligence

Another result of the training consists of how ANN determine the inputs with the most important influences on the output. So in Figure 4 and Table 4 is seen as the most influential in training process is ITS, while the smallest influence has TAS.

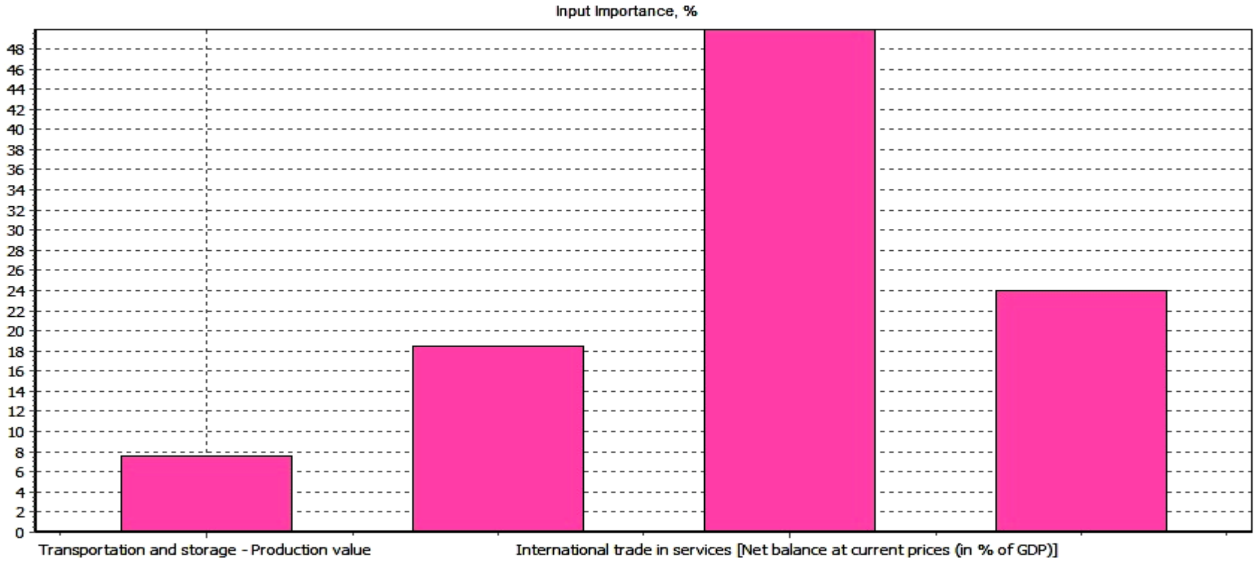

Figure 4. Input importance for the training results [%].
Source: Our own simulation with Alyuda NeuroIntelligence

Table 4. Column importance.

Input column name	Importance, %
Transportation and storage - Production value	7.590.857
Trade balance [in mill. ECU/EURO]	18.489.606
International trade in services [Net balance at current prices (in % of GDP)]	49.917.075
Annual percentage change in total employed population	24.002.462

Source: Our own simulation results.

3.6. Testing and Validation

Last ANN process before its final uses is testing and validating of training. This process determines the success of training by simple comparison of statistical data with data simulated by ANN.
The result of this comparison is shown in Table 5 and Figure 5.

Table 5. Testing summary.

Parameters	Target	Output	Absolute error	Absolute relative error
Mean:	613,815.90	617,729.73	3,918.18	0.006154
Std Dev:	43,165.72	46,140.04	9,259.95	0.01462
Min:	523,939	523,959.92	0.005485	8.88E-09
Max:	662,770	662,761.04	31,150.90	0.049327

Source: Our own simulation results.

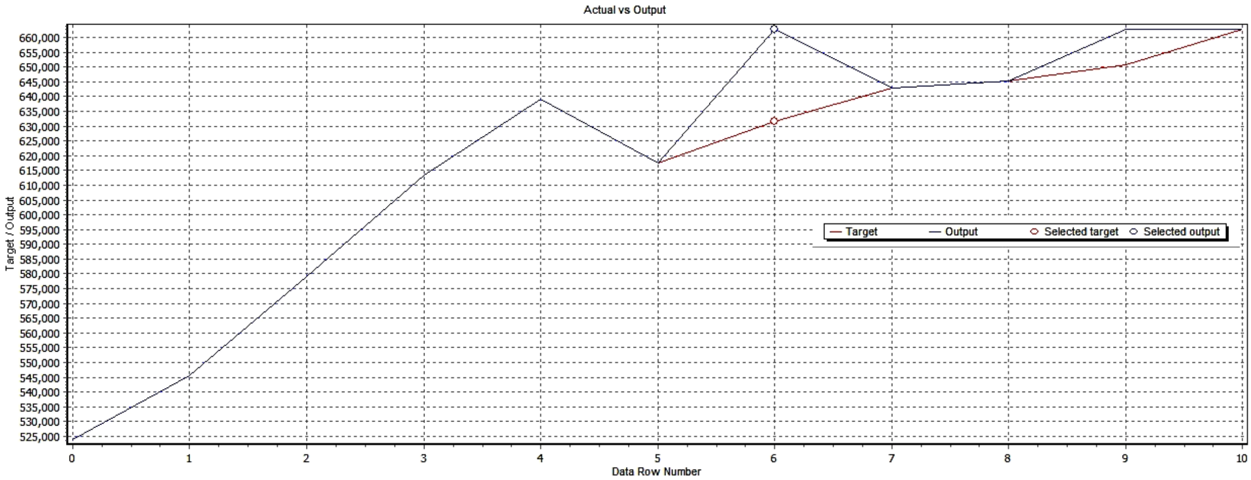

Figure 5. Actual values vs simulated ones.
Source: Our own simulation with Alyuda NeuroIntelligence

4. CONCLUSIONS

Training considering the test results it is obvious that steps taken in the research authors have chosen the best solutions in terms of the type ANN, structure and training algorithm. Such testing process resulted in a maximum error between actual values and simulated values by ANN less than 0.5%.

ACKNOWLEDGMENT

ANN training success and building can now be used for forecasting future GDP values from statistics or proposed for input elements of the network: TAS, TBA, ITS, APE.

Research can be expanded to include training in data from a longer period of time and increasing the number of data input in particular. It can be said that the authors of the research have paved the way for a possible more advanced research, our results establish the actual, solid and lucrative bases for future work.

REFERENCES

Aliuda (n.d). Retrieved June 3, 2015 from: http://www.alyuda.com/

Dumitru, I. N. (2000). *Rețele Neuronale Artificiale, Procesarea avansată a datelor*. Printech Publishing, Bucharest.

Hirose, A. (2003). *Complex-valued Neural Networks: Theories and Applications*. World Science Publishing Co Pte. Ltd., Singapore.

Irwin King. (2006). *Neural Networks, Fuzzy Inference Systems and Adaptive-Neuro Fuzzy Inference Systems for Financial Decision Making*. Neural information processing: 13th international conference, ICONIP 2006, Honk Kong, China, Proceedings, Part III, page 434.

Neagu, C. & Ioniță, C. (2004). *Rețele Neuronale. Teorie şi aplicații în modelarea proceselor şi sistemelor de producție*. Editura MatrixRom, Bucureşti.

Statistical data. Eurostat (n.d.). Retrieved 15 August 2015, from: http://ec.europa.eu/eurostat/cache/metadata/EN/nama_esms.htm

Udrescu, M. (2009). New Techniques Applied In Economics. Artificial Neural Network. ORADEA University Annals, Economic Sciences Series, Tom XVIII – Volume IV – Section: Management and Marketing, pg. 1080-1084, Oradea.

Waszczyszyn, Z. (2000). *Fundamentals of Artificial Neural Networks*. Institute of Computer Methods in Civil Engineering, Udine