

# An Expedient Study on Back-Propagation (BPN) Neural Networks for Modeling Automated Evaluation of the Answers and Progress of Deaf Students' That Possess Basic Knowledge of the English Language and Computer Skills

John Vrettaros<sup>1</sup>, George Vouros<sup>2</sup>, and Athanasios S. Drigas<sup>1</sup>

<sup>1</sup>NCSR DEMOKRITOS,  
Institute of Informatics and Telecommunications  
Net Media Lab

Ag. Paraskevi, 15310, Athens, Greece

<sup>2</sup>Aegean University, Info and Communication Systems Eng,  
83200, Karlovassi, Samos, Greece

{dr,jvr}@iit.demokritos.gr, georgev@aegean.gr

**Abstract.** This article studies the expediency of using neural networks technology and the development of back-propagation networks (BPN) models for modeling automated evaluation of the answers and progress of deaf students' that possess basic knowledge of the English language and computer skills, within a virtual e-learning environment. The performance of the developed neural models is evaluated with the correlation factor between the neural networks' response values and the real value data as well as the percentage measurement of the error between the neural networks' estimate values and the real value data during its training process and afterwards with unknown data that weren't used in the training process.

**Keywords:** Neural Networks, Back-propagation Network models, diagnosis.

## 1 Introduction

One of the most promising branches of current educational research is student diagnosis which according to VanLehn ([8]) is the process of inferring students' internal characteristics from their recognizable behavior. Student diagnosis is considered more than necessary as according to learning theories adaptive learning is a fruitful one if one wishes to obtain effective tutoring. That is the reason why one of the latest trends in educational software development is attempting to implement a simulation of the way that a human teacher adapts his tutoring to the individualization of each student [5]. In order for this kind of intelligence to be achieved, researchers have adopted many Artificial Intelligence methods. The most famous among them are neural networks, fuzzy logic as well as several search methods such as genetic algorithms.

Neural networks are on the top of the researchers' choice since they provide a system the ability to recognize patterns, to derive meaning from vague data and to identify matching in similar cases ([6]). Fuzzy set theory is widely used since it can deal in a reliable way with human uncertainty and it obtains smooth modelling of human decision – making. Genetic algorithms are ideal for efficient expert knowledge representation. Finally, Neuro – Fuzzy synergism is getting more and more popular in this area since it seems to overcome obstacles that come up when each of the methods involved is solely applied ([7]). Below we present several typical examples of the application of these methods in student's diagnosis.

Weon and Kim ([3]), implemented a system aiming at individual learning evaluation by making use of fuzzy linguistic variables representing each question. Finally, assessment came up taking into consideration the membership degree of uncertainty factors.

Vrettaros et al. ([4]) introduced a diagnostic system of taxonomies using fuzzy logic. Filled up questionnaires were processed in order for students' classification into one of the predefined knowledge levels to be achieved according to the given answers.

Stathakopoulou et al. ([5]) attempted to infer students' individual characteristics and use them in order not only to create but also to update the student model. This neural network - based fuzzy model is consisted of a fuzzy component and a neural network trained through actual students' profiles.

Hommsi et al. ([2]), presented an Adaptive and Intelligent Web – Based Educational System (AIWBES) based on Fuzzy – ART2 neural network and Hidden Markov Model (HMM), which is a stochastic method. The goal of the system described by the authors was to assess the learners' knowledge level and to obtain students' classification in one of six different levels taking into account several parameters.

In general, the aim of the e-learning environment that is studied is the realization of an e-educational system for the evaluation of the teaching of the English language to deaf people. The crucial point in the definition of the theoretical and technical aspects of the correlation between the e-learning models' subsystems is the definition of the configuration and the characteristics of their corresponding interconnections. According to the structural and operational details of the e-learning procedure the performance factor of the expert system and the collaboration of the e-learning model is the encoding and the contents of the inputs and outputs of the expert system as well as the structure, the standardization and the content of the database questions that require further attention and skilful handling.

The evaluation procedure for the teaching of English to deaf students relies on the attainment of the ESOL (English for Speakers of Other Languages) models (levels 1 and 2). These levels comprise five sections, which in ascending order are [A], [B], [C], [D] and [E]. Section [A] represents the Letter Recognition and their Alphabetical Order, section [B] represents the Spelling and the Vocabulary, section [C] represents the Grammar and the Sentence Structure, section [D] represents the Reading and section [E] represents the Writing.

According to the ESOL e-learning environment specifications the input / output parameters of an expert system can be determined without doubt, while simultaneously their translation is simple and quite straightforward.

As far as the input is concerned, there are five pairs of parameters in total and per question, which are:  $a = \{a_{val}, a_{rel}\}$ ,  $b = \{b_{val}, b_{rel}\}$ ,  $c = \{c_{val}, c_{rel}\}$ ,  $d = \{d_{val}, d_{rel}\}$  and  $e = \{e_{val}, e_{rel}\}$ . In other words, every pair corresponds to a language section of a certain level. Parameter  $a$  describes the recognition of the letters and the alphabetical order of section  $[A]$ , parameter  $b$  relates to the spelling/vocabulary of section  $[B]$ , parameter  $c$  represents the grammar/sentence structure for section  $[C]$ , the corresponding parameter for reading in section  $[D]$  is  $d$ , while the writing skill of section  $[E]$  is quantified with parameter  $e$ . The abbreviation index *val* (value) represents the evaluation of the specific section based on a particular response, while the abbreviation index *rel* (relevance) recognizes the level of relevance /weight of a specific question with respect to the contents of a section.

The evaluation values of the input parameters  $a_{val}, b_{val}, c_{val}, d_{val}$  and  $e_{val}$  originate from the range  $S = \{-1\} \cup [0, 1]$ . When a section is not examined by a question of the corresponding parameter, then the range is characterized by the value  $-1$ . A wrong answer according to a certain section leads to a corresponding value of zero (0), while the value of a sections' parameter is (1) if the chosen answer is correct according to that section. Accordingly, answers that are partially correct based on some sections have their values placed in between.

On the other hand, someone can argue that the relevance parameters  $a_{rel}, b_{rel}, c_{rel}, d_{rel}$  and  $e_{rel}$  characterize the question itself rather than its possible answers. Although this is true, the negotiation with relative parameters as part of a given answer is handy from an evaluation viewpoint. As a result, the relevance/weight is considered range between  $[0, 1]$ , where the value zero (0) or values close to zero mean low relevance, while the value one (1) or values close to one mean high relevance. Accordingly, the other weight values range somewhere in between. It must be pointed out, that the relevance parameters are the same for all the answers of a specific question.

The most skilful method for the provision of information in an expert system in relation with these sections is the sequence in array format of the ten values for the parameters of the input pairs:

$a_{val}$	$a_{rel}$	$b_{val}$	$b_{rel}$	$c_{val}$	$c_{rel}$	$d_{val}$	$d_{rel}$	$e_{val}$	$e_{rel}$
-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

As an example, let's consider that a question shows low relevance in section  $A$ , high relevance in section  $C$  and medium relevance in section  $B$ . Moreover, let's assume that the question under consideration does not include information about sections  $D$  and  $E$ . Now let's consider an answer to the previous question, which is

correct according to section A, partially correct according to section C, wrong according to section B. Obviously, it does not contain information relating to sections D and E. Such an answer ends in a ten value sequence, which comprises elements that are defined in the range  $S = \{-1\} \cup [0, 1]$ . In addition, it is obvious that the aforementioned ten value sequence can be directly encoded as an arithmetic string similar to the one below:

$a_{val}$	$a_{rel}$	$b_{val}$	$b_{rel}$	$c_{val}$	$c_{rel}$	$d_{val}$	$d_{rel}$	$e_{val}$	$e_{rel}$
1	0.1	0	0.5	0.7	0.9	-1	0	-1	0

This way, this particular arithmetic string can easily be used as an input to an expert system.

As far as the output is concerned, from observations and/or monitoring of the operational and correlating characteristics of the expert system, the conclusion that is derived is that that output parameters of the system are six (6) namely  $y_1, y_2, y_3, y_4, y_5$  and  $y_6$ . The first five parameters are the evaluation/estimation of the lingual skills per section, while the sixth parameter represents the total estimation of the user for the sum of all of the lingual skills, as follows:

$y_1$  = letter recognition and alphabetical order skills

$y_2$  = spelling / vocabulary skills

$y_3$  = grammar / sentence structure skills

$y_4$  = reading skills

$y_5$  = writing skills

$y_6$  = total lingual skills (**essentially, it is a weighted average of  $y_1 - y_5$ , representing a general estimation of the linguistic level of the student, as a professional pedagogic would have determined in a real case scenario**).

It is obvious that the output parameters are continuous. Due to the fact that the expert system outputs represent a certain estimation in connection with a specific lingual section, the evaluation is considered to be normalized in the range  $[0, 1]$ . The translation of the final arithmetic values is simple: zero means no lingual skills, one means perfect lingual skills, while all the other lingual skills levels may be estimated with similar arithmetic insertions. The output values, which are already arithmetically encoded, may be inserted in the e-learning environment as a six value arrayed sequence:

$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$
-------	-------	-------	-------	-------	-------

Let's assume that the next estimation is real for a particular student:

- 0.6 = letter recognition and alphabetical order skills
- 0.4 = spelling / vocabulary skills
- 0.2 = grammar / sentence structure skills
- 0.5 = reading skills
- 0.3 = writing skills
- 0.4 = total lingual skills

This particular six value arrayed sequence which comprises continuous elements could be directly encoded as an arithmetic string similar to the one below:

$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$
0.6	0.4	0.2	0.5	0.3	0.4

This way the final outputs are available straight away to rest of the e-learning environment.

The above (according to the ESOL specifications) skilful encoding and inputs and outputs content as well as the structure, standardization and content of the database questions is based on the use of neural networks technology for the e-learning expert system modeling for the automated estimation of the evaluation values of the teaching of English to deaf people.

The neural networks technology has been applied with success to many estimation problems with similar input/output characteristics. This study examines the expediency of applying this technology for modeling the automated evaluation of the deaf students' answers in the form of questions divided in five sections within the e-learning environment of the expert system.

In this study, section 2 provides a description of the neural networks characteristics and details about the development of a typical BPN model. In section 3, various Back-Propagation Networks (BPN) type neural networks models are developed for the automated evaluation of the deaf students' education and progress and their results are presented, determining hence, the expediency of the use of neural networks technology in such problems. In section 4, general neural networks technological issues are presented about generalization and over-fitting for the successful development and realization of effective models. Finally, in section 5 conclusions are presented and ideas are suggested for further research.

## 2 Development of Back-Propagation Network (BPN) Type Neural Networks for the Automated Evaluation of Deaf Students' Progress

In this study, the general schema of input/output data for supervised learning and their connection with neural networks results from the definition of ten (10) variables namely,  $v, v', w, w', x, x', y, y', z$  and  $z'$ , which belong in the range  $S = \{-1\} \cup [0, 1]$  for the input and with the definition of six (6) variables, namely,

$h, i, j, k, l$  and  $n$  that belong in the range  $[0, 1]$ , for the output. Hence, each input/output model from the total education data could be similar to the next model (where the symbol  $\emptyset$  denotes the blank value):

Ten (10) Input Values	$a$	$b$	$c$	$d$	$e$	
	$(v, v')$	$(w, w')$	$(x, x')$	$(y, y')$	$(z, z')$	
Six (6) Output Values	$y_1$	$y_2$	$y_3$	$y_4$	$y_5$	$y_6$
	$h$ if $v \neq -1$ else $\emptyset$	$i$ if $w \neq -1$ else $\emptyset$	$j$ if $x \neq -1$ else $\emptyset$	$k$ if $y \neq -1$ else $\emptyset$	$l$ if $z \neq -1$ else $\emptyset$	$n$

A sample of real values for seven (7) models from the sum of the education data is depicted in the following Table 1.

**Table 1.** A sample of supervised learning education data

pattern (#)		1	2	3	4	5	6	7
Input Variables	$a$ ----- $a_{val}$	0.6	0.1	1.0	0.9	0.6	0.1	0.1
	$a$ ----- $a_{rel}$	0.5	0.5	0.5	0.7	0.7	0.7	0.2
	$b$ ----- $b_{val}$	0.8	0.7	0.1	-1	-1	0.8	-1
	$b$ ----- $b_{rel}$	0.8	0.8	0.8	0	0	0.6	0
	$c$ ----- $c_{val}$	0.6	0.8	0.2	-1	-1	0.9	0.1
	$c$ ----- $c_{rel}$	0.1	0.1	0.1	0	0	0.5	0.8
	$d$ ----- $d_{val}$	-1	-1	-1	0.9	0.1	0.6	-1
	$d$ ----- $d_{rel}$	0	0	0	0.5	0.5	0.4	0
Output Variables	$e$ ----- $e_{val}$	-1	-1	-1	0.6	0.8	0.2	0.9
	$e$ ----- $e_{rel}$	0	0	0	0.3	0.3	0.1	0.3
	$y_1$	0.9	0.2	0.3	0.5	0.8	0.4	0.7
	$y_2$	0.7	0.8	0.2	$\emptyset$	$\emptyset$	0.1	$\emptyset$
	$y_3$	0.4	0.5	0.2	$\emptyset$	$\emptyset$	0.1	0.5
	$y_4$	$\emptyset$	$\emptyset$	$\emptyset$	0.9	0.0	0.1	$\emptyset$
	$y_5$	$\emptyset$	$\emptyset$	$\emptyset$	0.6	0.2	0.6	0.8
	$y_6$	0.1	0.5	0.7	0.1	0.2	0.9	0.7

It is obvious that the information that is contained in such education data, should be gathered and processed by a professional pedagogist, due to the fact that such a person seems to be the most suitable and competent person to create the aforementioned content.

For the development of typical back-propagation network (BPN) neural networks models, this study used the modeling and simulation software package Predict by NeuralWare, which is described in reference [11]. This software package provides the necessary mechanisms for the automated modifications and executes the scaling and limiting procedures to the users' input/output data. The basic idea for the modification of data, for the better identification of the statistical distribution of the input variables for the neural networks' training and for the maximization of its performance, is described in references [12, 13].

The software package Predict by NeuralWare demands that the presentation of the variables of the input vector,  $x$ , to be of the following pattern:

$$x = [a_{val} \quad a_{rel} \quad b_{val} \quad b_{rel} \quad c_{val} \quad c_{rel} \quad d_{val} \quad d_{rel} \quad e_{val} \quad e_{rel}]$$

with values for the evaluation parameters  $a_{val}$ ,  $b_{val}$ ,  $c_{val}$ ,  $d_{val}$  and  $e_{val}$  that originate from the aforementioned range  $S = \{-1\} \cup [0, 1]$  and with values for the relevance parameters  $a_{rel}$ ,  $b_{rel}$ ,  $c_{rel}$ ,  $d_{rel}$  and  $e_{rel}$  that originate from the aforementioned range  $[0, 1]$ . For the variables of output vector,  $y$ , the software package Predict by NeuralWare demands that the presentation is of the following pattern:

$$y = [y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6]$$

with values that originate from the aforementioned normalized range  $[0, 1]$ .

The developed BPN neural network used the first 28 of the 36 available data from Table 7 for its training and the remaining 8 were used for its evaluation trials. Its training was based on the maximization of the correlation with the adaptive gradient-descent technique and its final architecture is of this pattern: 26-2-22/0.9622. This pattern means that the final trained neural network comprises 26 neurons in the input (added inputs due to the modifications), 2 neurons in the hidden layer and 22 neurons in the output (added outputs due to the modifications) with a correlation factor of 0.9622.

The performance of the BPN network Architecture #3 is evaluated by comparing the results of the trained neural network for all of the input/output data, using the recall phase. Table 9 presents the real values of the 36 input/output data, the corresponding values of the BPN#3 neural network and the corresponding indication of the correct or wrong answer (estimation) of the BPN#3 neural network.

From the above Table 2 one can derive that the BPN network Architecture #3 (with verbally encoded outputs) truly learned 27 out of a total of 28 model inputs/outputs, hence a percentage of  $27/28 = 96.43\%$ , which is comparable with the correlation factor percentage during its training, which is 96.22%. In addition, from the above Table 2 one can derive that the BPN network Architecture #3 estimates correctly almost 8 of the 8 output values, hence, a  $8/8 = 100.0\%$  percentage.

From the performances of the three aforementioned BPN architectures namely, BPN network Architecture #1, BPN network Architecture #2 και BPN network Architecture #3, it is obvious that the neural networks technology can be used in the development of

a model for the automated evaluation of the answers and progress of deaf students that possess basic knowledge of the English language as well as computer skills within a virtual e-learning environment. However, this study focused on the development of typical BPN models based only on a small sum of data. It is regarded as worthwhile to develop more BPN models, which are trained for a larger sum of data and to determine the level of their best performance using all the performance improvement techniques of the neural networks, which are presented in the next section. In general though, the acceptance of a model depends on its performance, which may be evaluated based on statistical methods with the only reserve being that the remaining errors between the models' output values and the real values follow approximately the normal distribution.

**Table 2.** BPN network Architecture #3 – Recall Phase Results - Single Select Questions

Table 9: BPN network Architecture #3 – Recall Phase Results - Single Select Questions														
	Test#,	Real Input Values (verbally encoded)						Output Values BPN#3 (verbally encoded)						BPN#3 Correct or Wrong Indication
		Y1	Y2	Y3	Y4	Y5	Y6	Y1	Y2	Y3	Y4	Y5	Y6	
1	t1,q2	zero	N	N	zero	N	zero	zero	n	n	zero	n	zero	Correct
2	t1,q2	one	N	N	one	N	one	one	n	n	one	n	one	Correct
3	t1,q2	zero	N	N	zero	N	zero	zero	n	n	zero	n	zero	Correct
4	t1,q2	three	N	N	three	N	three	three	n	n	three	n	three	Correct
5	t1,q3	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	Correct
6	t1,q3	zero	zero	five	zero	zero	two	zero	zero	zero	zero	zero	zero	Correct
7	t1,q3	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	Correct
8	t1,q3	one	one	one	one	one	one	one	one	one	one	one	one	Correct
9	t1,q4	one	one	one	one	one	one	one	one	one	one	one	one	Correct
10	t1,q4	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	Correct
11	t1,q4	zero	zero	five	zero	zero	two	zero	zero	five	zero	zero	two	Correct
12	t1,q4	zero	zero	five	zero	zero	two	zero	zero	five	zero	zero	two	Correct
13	t1,q5	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	Correct
14	t1,q5	one	one	one	one	one	one	one	one	one	one	one	one	Correct
15	t1,q5	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	Correct
16	t1,q5	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	zero	Correct
17	t3,q1	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
18	t3,q1	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
19	t3,q1	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
20	t3,q1	N	one	one	one	N	one	n	zero	zero	zero	n	zero	Wrong
21	t3,q2	N	five	zero	zero	N	two	n	five	zero	zero	n	two	Correct
22	t3,q2	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
23	t3,q2	N	one	one	one	N	one	n	one	one	one	n	one	Correct
24	t3,q2	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
25	t3,q3	N	zero	five	zero	N	two	n	zero	five	zero	n	two	Correct
26	t3,q3	N	one	one	one	N	one	n	one	one	one	n	one	Correct
27	t3,q3	N	five	zero	zero	N	two	n	five	zero	zero	n	two	Correct
28	t3,q3	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
29	t3,q4	N	one	one	one	N	one	n	one	one	one	n	one	Correct
30	t3,q4	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
31	t3,q4	N	five	zero	zero	N	two	n	five	zero	zero	n	two	Correct
32	t3,q4	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
33	t3,q5	N	zero	five	zero	N	two	n	zero	five	zero	n	two	Correct
34	t3,q5	N	one	one	one	N	one	n	one	one	one	n	one	Correct
35	t3,q5	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct
36	t3,q5	N	zero	zero	zero	N	zero	n	zero	zero	zero	n	zero	Correct



### 3 Conclusions

In this study, using the ESOL specifications for the skilful encoding and the inputs and outputs content and also the structure, standardization and content of the database questions of an e-learning expert system, three BPN type models were developed for the estimation of the expediency of applying neural network technology towards the automated estimation of the values, for the evaluation and progress of teaching English to deaf people.

Based on the performances of the three type BPN architectures that were developed, namely, BPN network Architecture #1, BPN network Architecture #2 and BPN network Architecture #3, it was derived that the neural networks technology can be used towards the development of such an automated model for the evaluation of the answers and the progress of deaf students that possess basic knowledge of the English language and also basic computer skills, within a virtual e-learning environment.

However, it is regarded that this estimate would be more accurate and would have more certitude if the available input/output data were more. Additionally, in this study only BPN type neural network models were tested and it is thought of as necessary to develop additional models for this particular problem and to determine the level for their best performance using a larger sum of available input/output data as well as all the performance improvement techniques for the neural networks.

### References

1. Rumelhart, D.E., McClelland, J.L. (eds.): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Foundations*, vol. I. MIT Press, Cambridge (1986)
2. Rumelhart, D.E., McClelland, J.L., The PDP Research Group: *Learning and Relearning in Boltzman Machines*. In: Hinton, G.E., Sejnowski, T.J. (eds.) *Parallel Distributed Processing. Foundations*, ch. 7, vol. 1. MIT Press, Cambridge (1985)
3. Carpenter, G.A., Grossberg, S.: *The ART of Adaptive Patern Recognition by a Self-Organizing Neural Network*. *Computer*, 77–88 (March 1988)
4. Carpenter, G.A., Grossberg, S.: *A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine*. *Computer Vision, Graphics and Image Processing* 37, 54–115 (1987)
5. Carpenter, G.A., Grossberg, S.: *ART 2: Self-Organization of Stable Category Recognition Codes for Analog Input Patterns*. *Applied Optics* 26, 4919–4930 (1987)
6. Kohonen, T., et al.: *Statistical Pattern Recognition with Neural Networks: Benchmark Studies*. In: *Proceedings of the Second Annual IEEE International Conference on Neural Networks*, vol. 1 (1988)
7. Kohonen, T.: *Self-Organization and Associative Memory*, 3rd edn. Springer, New York (1989)
8. Murray, S.: *Neural Networks for Statistical Modeling*. Springer, Van Nostrand Reinhold (1993)
9. *NeuralWorks manual, Neural Computing: A Technology Handbook for Professional II/Plus and NeuralWorks Explorer, and NeuralWorks Advanced Reference Guide: Software reference for Professional II/Plus and NeuralWorks Explorer*, NeuralWare, Inc., USA (1995)

10. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning Internal Representations by Error Propagation. In: Parallel Distributed Processing, vol. I, pp. 318–364. MIT Press, Cambridge (1987)
  11. NeuralWorks, Predict: Computer Software package for applications of Neural Networks, NeuralWare, Inc., USA
  12. Tukey, J.W.: Exploratory Data Analysis. Addison-Wesley, Reading (1977)
  13. Chatterjee, S., Price, B.: Regression Analysis by Example. John Wiley & Sons, Chichester (1991)
- 
- [1] Stathakopoulou, R., Magoulas, G., Grigoriadou, M., Samarakou, M.: Neuro –Fuzzy knowledge processing in intelligent learning environments for improved student diagnosis (2005)
  - [2] Homsı, M., Lutfi, R., Carro, R.M., Ghias, B.: Student modeling using NN – HMM for EFL course
  - [3] Weon, S., Kim, J.: Learning achievement evaluation strategy using fuzzy membership function (2001)
  - [4] Vrettaros, J., Vouros, G., Drigas, A.: Development of a diagnostic system of taxonomies using fuzzy logic – case SOLO (useful for e-learning system) (2004)
  - [5] Vrettaros, J., Pavlopoulos, J., Vouros, G., Drigas, S.: The development of a self-assessment system for the learners answers with the use of GPNN (2008)