International Journal of Information Technology and Knowledge Management
January June 2009, Volume 2, No. 1, pp. 209-215

A HYBRID MODEL IN PREDICTION OF ADHD USING ARTIFICIAL NEURAL NETWORKS

K. Arthi[*] & A. Tamilarasi[**]

In this paper, a hybrid artificial network model called DIAGADHD is proposed for the diagnosis of ADHD (Attention Deficit/ Hyperactivity Disorder) using neuro fuzzy technique. This model is a combination of unsupervised training algorithm using self organizing maps and supervised training algorithm using radial basis function. The linguistic values of suspected children are received from the parents or the teachers and then converted into fuzzy membership values. Those values are given as input to the hybrid model and trained for diagnosing ADHD. The approach proposed in this paper uses a hybrid neural network system consisting of Kohonen's self organizing maps followed by a radial basis function which uses fuzzy membership values as input. The model is trained in two phases on ADHD data. The trained hybrid model is tested for its effective performance and the experimental results are compared with the back propagation algorithm.

Keywords: ADHD, Neuro fuzzy, Kohonen's Self-organizing maps, Radial basis function hybrid model, DIAGADHD.

1. Introduction

ADHD (Attention Deficit/Hyperactivity Disorder) is one of the most common neurobehavioral disorders of childhood and can persist through adolescence and into adulthood. A person with ADHD has a chronic level of inattention, impulsive hyperactivity, or both such that daily functioning is compromised. The symptoms of the disorder will be present at levels that are higher than expected ones for a person's developmental stage, which interferes with the person's ability to function in different settings (e.g., in school and at home). A person with ADHD may faces lots of difficulties in certain part of life such as peer and family relationships, and school or work performance [1]. The proposed hybrid model will surely assist the special educators and psychologists who are the predictors of ADHD.

2. Previous Work on Diagnosis of ADHD

Previously an integrative model have been proposed which incorporates new neuro anatomical findings and emphasizes the interaction between parallel processing pathways as potential loci for dysfunction[18].Another study investigates, the pure time perception of Chinese children with ADHD by using a duration discrimination task[3].A recent study proposes a computerized continuous performance test (CPT) as a diagnostic tool for classifying attention-deficit hyperactivity disorder (ADHD) [11]. In another study of ADHD , a two-leveled hybrid system has been developed for brief theoretical analysis within Ginsberg's unified framework of multivalued logic[15].

2.1 Categories of ADHD

The two main criteria used to make a diagnosis of ADHD are attention symptoms and hyperactivity symptoms [1].

The key feature associated with symptoms of inattention includes:

- failing to give close attention to details and difficulty sustaining attention in tasks or play
- not listening when spoken to
- not following through on instructions and failure to finish tasks
- difficulty organizing tasks and activities
- avoiding, disliking or being reluctant to engage in tasks that require sustained mental effort
- losing things necessary for tasks or activities
- easily distracted

The key features associated with **symptoms of hyperactivity** (sometimes known as hyperactivity-impulsivity) include:

- fidgeting with hands or feet, squirming in seat
- leaving seat when remaining sitting is expected
- running about or climbing excessively
- difficulty playing or engaging in leisure activities and often 'on the go'

[*] Department of Computer Applications, Karpagam College of Engineering, Coimbatore-32, Tamilnadu, India. *Email: art_sek@rediffmail.com*

[**] Department of Computer Applications, Kongu Engineering College Perundurai, Erode, Tamilnadu, India. *Email: sek_art@rediffmail.com*

- talking excessively and blurting out answers before a question is completed
- interrupting others

3. ARTIFICIAL NEURAL NETWORK

Artificial neural network (ANN) are powerful computational systems consisting of many simple processing elements connected together to perform tasks analogously to biological brains. ANN can be effectively used as a tool in medical decision making [5]. The learning process of the artificial neural networks is similar to that of the learning function of the brain. During training, samples are given to the input layer that yields changes of the activation state of output processing elements. The network adjusts the weights depending upon the difference between the required and the calculated output values.

4. NEURAL NETWORKS AND FUZZY LOGIC

Neural networks have wide application in oncology, neurology, brain function and radiology [9]. It is one of the powerful AI techniques that has the capability to learn a set of data and constructs weight matrixes to represent the learning patterns[4][12]. But fuzzy logic provides an approach to approximate reasoning in which the rules of inference are approximate rather than exact [20]. A fuzzy logic system has a series of rules comprising of an antecedent and a consequent combined as if –then semantics. It deals with uncertainty in knowledge that simulates human reasoning in incomplete data or fuzzy data. Neuro fuzzy combines the advantages of fuzzy set theory and neural networks[10]. It can handle linguistic informations and efficiently mimics the human decision making process. Neural Networks was implemented as a hybrid with high textual description method to detect abnormalities within the same images with high accuracy and also in finding brain disorders [7] [13].

4.1 Kohonen's Self Organizing Map (KSOM)

In Self-organizing map, during the initial learning process of ordering phase, the learning rate parameter should be set close to unity and then gradually decreased. While in the convergence phase of the learning process, the learning rate parameter attains relatively small values for a long time[8]. In this unsupervised training algorithm, the process is continued for particular number of epochs or the learning rate reduces to a very small rate. KSOM has been used effectively in medical decision support system. A system made of KSOM and back propagation has been built as an aiding tool in the analysis of mammograms for the diagnosis of breast cancer [16]. Previously, a new approach have been proposed which uses a hybrid neural network system, consisting of a self-organizing map followed by back propagation network, to restrict the number of spatial grey level dependence matrices that need to be computed[2]. Recently, a hybrid artificial neural network model has been developed which consist of a self-organizing map followed by a back propagation neural network, that can automate between the normal subjects and the subjects with Parkinson's disease or the spino cerebeller degeneration patients [19].

4.2 Radial Basis Function (RBF)

A radial basis function neural network solves a nonlinear problem by casting input samples into a higher dimensional space in a non-linear way. It is based on supervised learning which is good in modeling nonlinear data and also helps in learning the given application quickly. RBF networks rapidly trains than that of back propagation networks where the weights are the centers of a set of basis function calculated using K means clustering. All hidden units in the RBF networks have the same width or degree of sensitivity to inputs. Calculation of individual width increases the performance of the RBF network. In a previous study, this network had been used successfully for structural classification of thyroid diseases [14]. In that work experimental results show that the trained RBFNN model outperforms the corresponding MLPNN model. Another classification technique based on RBFNN produces successful results and outperforms a number of classifiers which are based on the feed forward neural network architecture [6].

5. PROPOSED HYBRID MODEL

Since in medical decision making, the unsupervised technique Kohonen's SOM and the supervised technique RBF have been used successfully, a hybrid of both the networks can be used in diagnosing ADHD.

5.1 Fuzzy Membership Values

For the proposed model, the fuzzy rule base or bank of fuzzy associative memory rules can be specified by using the answers from the Questionnaire given in Appendix A. Fuzzy associations or rules (A,B) associate output fuzzy sets B of control values, with input fuzzy sets A of input variable values[17]. Fuzzy associations as antecedent-consequent pairs or IF-THEN statements for ADHD can be represented as

$$Rj : \text{IF } x_1 \text{ is } A_1^{\,j} \text{ and } x_2 \text{ is } A_2^{\,j} \ldots x_m \text{ is } A_m^{\,j}, \text{ then } y_1 \text{ is } \beta_1^{\,j}, \ldots y_M \text{ is } \beta_M^{\,j} \quad (1)$$

where $x_1, x_2 \ldots x_m$ represents input vectors of symptoms of ADHD

m – dimension of the input vectors

j - rule index ($j = 1, 2 \ldots k$)

M – Number of output neurons

y_1 – class of output ADHD is HIGH (H), MEDIUM (M), and LOW (L) or NOT ATALL (NAA)

For Example, high level of ADHD can be represented by using IF – THEN rules as

R1:IF Q1 is *AL* AND *Q*2 is *NVL* AND … AND Q18 is *CN* then ADHD is HIGH (2)

Similarly for MEDIUM (M), LOW (L) or NOT ATALL (NAA) can be specified using fuzzy rules.

A fuzzy set is an extension of classical set and can be defined as a set of ordered pairs which is represented as

$$A = \{x, \mu_A(x) / x \,\varepsilon\, X\} \quad (3)$$

X represents universe of discourse

x represents elements of *X*

where $\mu_A(x)$ membership function of *x* in *A* which take a values between 0 and 1 that can be calculated using symmetric gaussian function as

$$\mu_{Ai}(x) = e \, \frac{-(x-ai)^2}{2\sigma^2} \quad (4)$$

where a_i and σ define the shape of each membership function and can be depicted in figure 1.

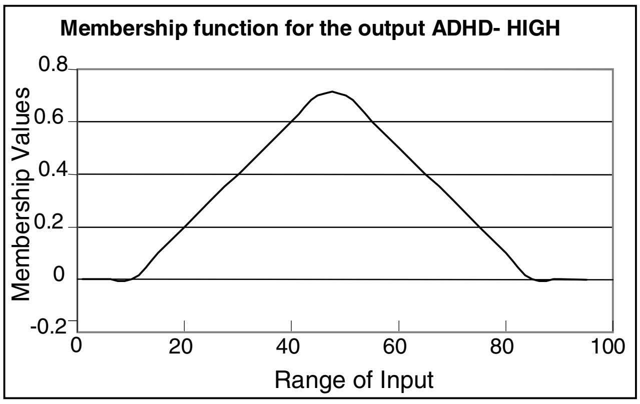

Figure 1: Membership Function for the Output ADHD-High

5.2 Architectural Representation of Neuro Fuzzy Model

The advantages of SOM and RBF algorithm are combined together to form a very efficient model in predicting ADHD disorder as shown in figure 2. The four-layer structure of the hybrid neural network model shown in figure 2 is a modified version, where the input layer accepts fuzzy membership values as an input which is trained using unsupervised training method of self organizing maps. The resultant weights of the KSOM layer are fed into RBF layer along with the given input of *xi*, to find the output of three classes *y ℓ* of ADHD using supervised learning algorithm called radial basis function.

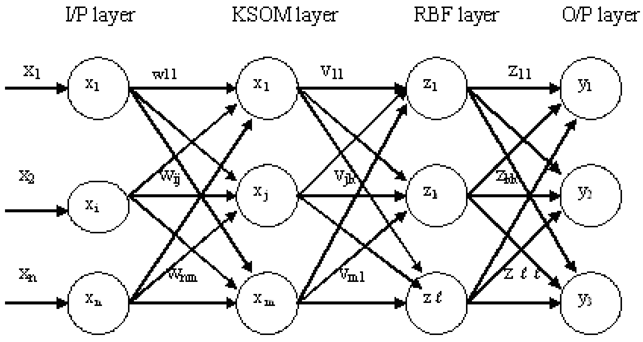

Figure 2: Architectural Representation of the Hybrid Model (DIAGADHD)

5.2 Flowchart

The representation of the hybrid model can be drawn as a flowchart as shown in figure 3

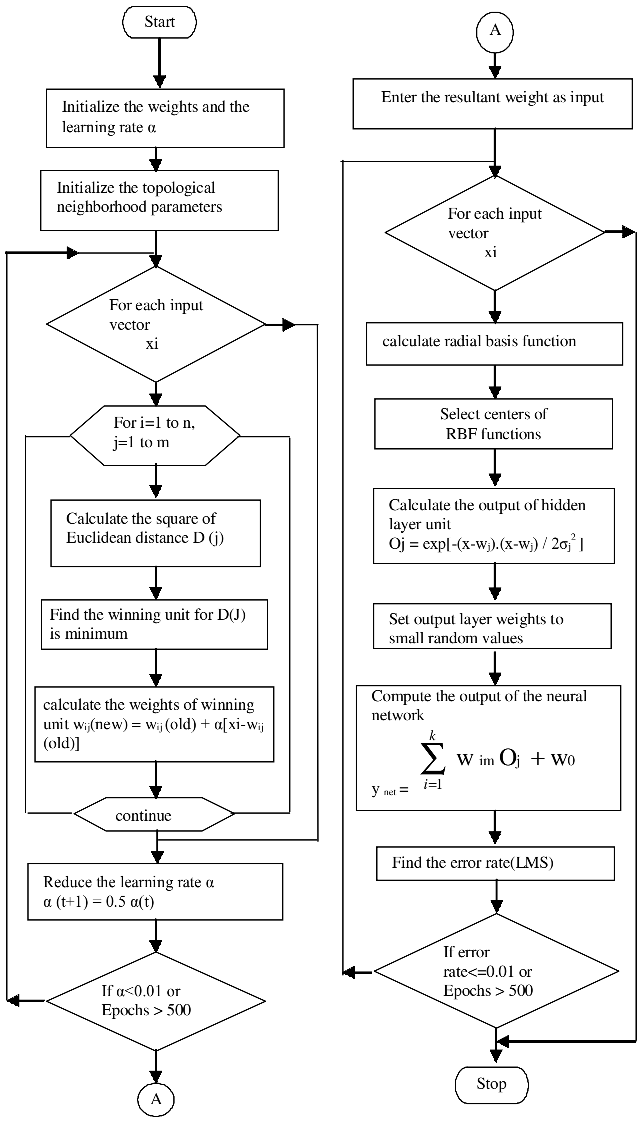

Figure 3: Flow Chart for the Training Process of DIAGADHD Algorithm

5.3 Algorithm

Algorithm: DIAGADHD ( )

Phase 1: Input layer to the KSOM layer.

This phase involves the following 10 steps.

Step 0: i) initialize the weight W_{ij} as random values.

ii) Set the radius of the neighborhood

iii) Initialize the learning rate α which is a slowly decreasing function of time.

Step 1: Conversion of given linguistic values into fuzzy membership values.

i) For each input values convert into fuzzy membership values using Gaussian membership function.

$$F(\mathrm{x};\, \sigma,\, c) = e^{\frac{-(x-c)^2}{2\sigma^2}}$$

where x and c represents the input values and centre respectively.

ii) Continue the steps 2 to 8 when stopping condition is false.

Step 2: Do the steps 3 to 5 for each input vector x of fuzzy membership value.

Step 3: Compute the square of the Euclidean distance (i.e.) for each j = 1 to m

$$D\,(j) = \sum_{i=1}^{n}(xi - wij)^2$$

Step 4: Find the winning unit index J, so that $D(j)$ is minimum.

Step 5: For all th units j within a specific neighbourhood of J and for all i,

Calculate the new weights as

$w_{ij}(\text{new}) = (1- \alpha)\, w_{ij}(\text{old}) + \alpha\, x_i$

Step 6: Update the learning rate α using the formula $\alpha(t+1) = 0.5\, \alpha(t)$

Step 7: Reduce the radius of topological neighborhood at specific times.

Step 8: Test for condition either number of epochs > 500 or the learning rate(α) reduces to a least value of 0.1.

Step 9: Now the resultant weight is given as input to the model for training with supervised learning algorithm of radial basis function.

Phase 2: From KSOM layer to RBF layer and then to the output layer.

Step 0: set the weights as the updated weights obtained from the previous algorithm.

Step 1: Perform steps 2 to 6 when the stopping condition is false.

Step 2: Perform steps 3 to 5 for each input.

Step 3: Each input unit (xi, for all i = 1 to n) receives input signals and transmits to the next layer.

Step 4: Calculation of activation function for the hidden layer.

i) Calculate activation level Oj for the hidden unit j.

$$\mathrm{O}j = \exp\,[-(x-w_j).(x-w_j) \,/\, 2\sigma_j^2\,]$$

where x represents input vector,

w_j represents weight vector associated with hidden unit j and σ_j^2 average distance between the cluster center and the training instances in that cluster for the hidden unit j and can be calculated as

$$\sigma_j^2 = \frac{1}{M}\,\sum\,(x\text{-}w_j)\cdot(x\text{-}w_j)$$

Where x represents training patterns in the cluster,

w_j indicates center of the cluster associated with the hidden unit j and

M represents number of training instances in that cluster.

Step 5: Calculation of activation function for the output layer.

$$y_{net} = \sum_{i=1}^{k} w_{im}\; O_j + w0$$

Where k represents the number of hidden layer nodes,

y_{net} gives the output value of m^{th} node in output layer,

w_{im} represents the weight between i^{th} RBF unit and the m^{th} output node

w_0 represents the biasing term at n^{th} output node.

Step 6: Calculation of LMS error and test for the stopping condition.

Here in this model, the iterations will be repeated until number of epochs reaches 500 or the error value deviates to 0.01.

6. Experimental Results and Discussion

A questionnaire is set based on the symptoms of inattention and hyperactivity of ADHD as listed in section 2.1. Fuzzy membership values are created for the given linguistic input values and stored for 165 records. The stored values are given as input to the unsupervised training algorithm of Kohonen's self organizing maps for training. During the first phase of training, nodes that are neighbours of the winning node are allowed to update their weights. The neighborhood of nodes that are allowed to update their weights decreases during the first phase. During the second phase, all weights are adjusted by small amounts until the network converges. In Kohonen's self-organizing map, the process is continued for particular number of epochs or the learning rate reduces to a very small rate. The resultant weights of the SOM algorithm act as centres of a set of Gaussian basis function along with the fuzzy input values for training the network to predict ADHD. In RBF networks, the weight which is set before the second layer of weights is adjusted. Here in this model, the resultant weights from KSOM is fed into the hidden layer unit instead of adjusting the weights. The network is trained until the error rate converges to 0.01 or the number of epochs reaches maximum of 600. Here the network learns quickly when it reaches the error rate of 0.01 for the epoch of 310 as shown in table 1 and figure 4. Likewise the network is trained using back propagation algorithm with the same datasets and the result is compared with the hybrid model result. Back Propagation network is a multilayer feed forward network, with differentiable activation function units are used to learn a training set of input-output examples. While training the network and for each iteration, the error value is calculated. The training is stopped when the error begins to rise, later it had been decreasing steadily as the net begins to memorize and lose its generalization property. Here in this model, the weight could be assigned between 0 to 1 with a bias of (–1) to the hidden unit and the output unit. The training vector starts with a $n \times n$ matrix, which contains values between 0 to 1 and ends with the error rate of 0.1. The weight is updated and propagated back to the net till the error rate is reduced to a least value or until a maximum number of epochs is reached. Here in this model the training is stopped when it reaches 600 epochs of error rate 0.01 as shown in table 2 and figure 5. The major difference between the hybrid neural network model and back propagation network lies in the

Table 1
Represents the Error Rate with Epochs using Hybrid Model (DIAGADHD)

Sno	*Learning Rate*	*Epochs*	*Average Error Rate*
1	0.5	52	0.45
2	0.5	75	0.8
3	0.4	104	1.01
4	0.3	126	0.7
5	0.3	152	0.34
6	0.2	180	0.2
7	0.2	200	0.18
8	0.2	220	0.127
9	0.2	258	0.069
10	0.2	278	0.038
11	0.2	298	0.02
12	0.1	310	0.01
13	0.1	335	0.01
14	0.1	400	0.01

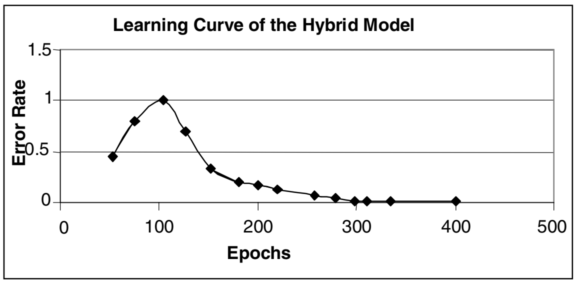

Figure 4: Learning Curve of the Hybrid Model

behaviour of the single hidden layer. In this paper back propagation uses sigmoidal activation function but the proposed neural network model uses gaussian function.

Table 2
Represents the Error Rate with Epochs using Back Propagation Network

Sno	*Nodes in Hidden Layer*	*Epochs*	*Learning Rate*	*Error Rate*
1	1-25	0.3	30	0.192
2	1-25	0.2	50	0.078
3	1-25	0.2	100	0.58
4	1-25	0.2	150	0.52
5	2-50	0.3	200	0.34
6	2-50	0.3	300	0.29
7	3-75	0.2	400	0.2
8	3-75	0.1	500	0.1
9	1-25	0.1	600	0.1
10	1-25	0.1	600	0.1

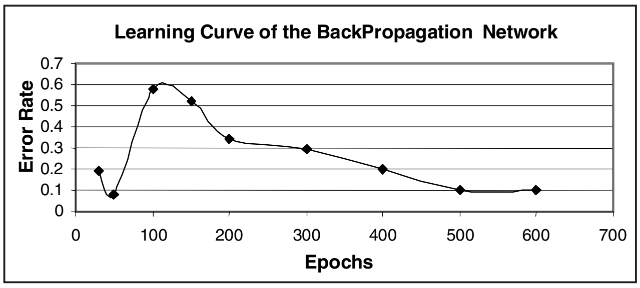

Figure 5: Learning Curve of the Back Propagation Network

7. CONCLUSION

It is observed that the hybrid model of SOM and RBF networks have better accuracy when compared to back propagation neural networks. Even they are less vulnerable to problems with non-stationary inputs due to the behaviour of the radial basis function hidden units. The proposed model will be a valuable alternate predictive method in diagnosing ADHD for the pediatricians and special educators. In future the concept of fuzzy cognitive maps can be applied in diagnosing the neurological disorder of autism or ADHD.

ACKNOWLEDGEMENT

The authors would to like to thank Dr. S.Thejavathy for her kind support in collecting the datasets.

Appendix A

Name of the child:
Age
Sex
Relationship: (mother/father/teacher)

1. Does your child feel difficult while sustaining any task?
   (a) certainly not (CN)
   (b) No, I don't think so (NI)
   (c) always (AL)
   (d) At times (AT)
2. Will your child pay a close attention while playing?
   (a) Yes, Quite a lot (QL)
   (b) Yes, Sometimes (YS)
   (c) No,Very Little (NVL)
   (d) Don't Know (DK)
3. Does your child listen to your speech?
   (a) Yes, Often (YO) (c) not sure (NS)
   (b) Yes, I think so (Y) (d) Sometimes (S)
4. Will your child follow to the instructions you give?
   (a) Never (NR) (c) Usually does (UD)
   (b) Rarely (RY) (d) Not at all (NAA)
5. Will your child finish the task successfully?
   (a) Not Sure (NS) (c) Never (N)
   (b) Certainly not (CN) (d) Sometimes (S)
6. Does your child show any difficulty in organizing a task or an activity.
   (a) Usually does (UD) (c) At times (AT)
   (b) Yes of course (YO) (d) Don't Know(DK)
7. When your child is ignored will they feel mentally disturbed?
   (a) I think so (IT) (c) May be(MB)
   (b) Definitely yes (DY) (d) Not sure(NS)
8. Will your child sustain any mental effort when she/he is avoided or disliked?
   (a) Rarely (RY) (c) Sometimes (S)
   (b) No (N) (d) Always (AL)
9. Does your child loose any necessary thing for the sake of tasks and activities?
   (a) Yes, Always (YA)
   (b) No (N)
   (c) Not often always (NOA)
   (d) May be (MB)
10. Is the child easily distracted?
   (a) Certainly(CY)
   (b) No, I don't think so (NID)
   (c) Sometimes (S)
   (d) can't remember (CR)
11. Will the child fidget with hands or feet?
   (a) Yes, of course (YOC)
   (b) No (N)
   (c) Sometimes (S)
   (d) Don't Remember (DR)
12. Will your child squirm in seat?
   (a) at times (AT) (c) May be (MB)
   (b) Not truly (NT) (d) Yes (Y)
13. Will your child leave leave the place when she/he is expected to sit.
   (a) Yes (Y)
   (b) No (N)
   (c) I don't remember (IDR)
   (d) does sometimes(DS)
14. Does your child unnecessarily run around or climb excessively?
   (a) Yes (Y) (c) No (N)
   (b) Not sure (NS) (d) Don't Know (DK)
15. Does your child feel difficult while playing or engaging in leisure activities.
   (a) Once a while (OW) (c) Usually does (UD)
   (b) Often (O) (d) Yes (Y)
16. Does your child talk excessively?
   (a) No (N) (c) May be (MB)
   (b) At times (AT) (d) Slightly (S)
17. Will the child blurt out answers before your question is complete?
   (a) Always (A)
   (b) Usually does (UD)
   (c) Don't Remember (DR)
   (d) No (N)
18. Will the child interrupt others?
   (a) No (N)
   (b) For some reason (FR)
   (c) Certainly not (CN)
   (d) At times (AT)

References

[1] American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders-IV* Washington DC: American Psychiatric Association, (1994).

[2] Arrowsmith, M. J., Varley, M. R., Picton, P. D., Heys J. D. Hybrid Neural Network System for Texture Analysis. *Seventh International Conference on Image Processing and Its Applications. Conf. Publ.* **1**, (465), (1999), 339–343.

[3] Binrang Yang, Raymond C. K. Chan, Xiaobing Zou, Jin Jing, Jianning Mai and Jing Li. Time Perception Deficit in Children with ADHD. *Journal of Brain Research*, **1170**, (2007), 90–96.

[4] Botis, T. and Halkiotis, S. Neural Networks for the Prediction of Spirometric Reference Values *Med. Inform. Internet Med.* **28** (4), (2003), 299–309.

[5] Dorffner, G., Erich, P., Markus., Stefan, K., Paolo, P., Gerold, P., Heinz, S., Experiences with Neural Networks as a Diagnostic Tool in Medical Image Processing. *Artif.Intell.Med* 1996.

[6] Haralambos Sarimveis, Philip Doganis, Alex Alexandridis. A Classification Technique based on Radial basis Function Neural Network. *Advances in Engineering Software* **37**, (4), (2006), 218–221.

[7] Karkanis, S. A., Magoulas, G. D., Grigoriadou, M., and Schurr, M. Detecting Abnormalities in Colonoscopic Images by Textual Description and Neural Networks. *Machine Learning and Applications: Machine Learning in Medical Applications*. Chania, Greece, 59–62.

[8] T. Kohonen, The Self-Organizing Map, *IEEE Proceedings*, **78** (1990), 1464–80.

[9] Kornel, P., Bela, M., Rainer, S., Zalan, D., Zsolt, T. and Janos F. Application of Neural Networks in Medicine, *Diag. Med. Tech*. **4**(3), (1998), 538–546.

[10] Kosko B. Neural Network and Fuzzy System, *A Dynamical Systems Approach to Machine Intelligence*, Prentice Hall International Edition (First Edition), (1992).

[11] Lee H., Cho S., Shin M. Supporting Diagnosis of Attention-deficit Hyperactive Disorder with Novelty Detection. *Artificial Intelligence in Medicine*, **42**, (3), (2008), 199–212.

[12 Mizumoto, M Shi Y., A New Approach of Neurofuzzy Learning Algorithms. IN: Ruan, D ed. *Intelligent Hybrid Systems, Fuzzy Logic, Neural Networks*, *Genetic Algorithms*, Kluwer Academic Publishers, (1997), 109–129.

[13] Pranckeviciene, E., Finding Similarities Between an Activity of the Different EEG's by Means of a Single Layer Perceptron. *Machine Learning in Medical Applications. Chania*, Greece, (1999), 27–39.

[14] Rizyan Erol, Seyfettin Noyen Ogulata, Cenk Sahin, Z.Nazan Alparslan. A Radial Basis Function Neural Network (RBFNN) Approach for Structured Classification of Thyroid Diseases. *Journal of Medical Systems*, **32**, (3) (2008), 215–220.

[15] Sabina Munteanu. A Hybrid Model for Diagnosing Multiple Disorders. *Int. Journal of Hybrid Intelligent Systems*, **2**, (1), (2005), 35–55.

[16] Santos Andre, T. C. S; Da Silva, A. C. R. A Neural Network Made of a Kononen SOM Coupled to a Map Trained via Back Propagation for the Diagnosis of Malignant Breast Cancer from Digital Mammograms. *International Conference on Neural Networks*, **5**, (1999), 3647–3650.

[17] Wang, L. X., and Mendal, J. M. Generating Fuzzy Rules by Learning from Examples. *IEEE Transactions on Systems, Man and Cynernetics*, (Nov/Dec, 1992), 1414–1427.

[18] F. Xavier Castellanos, Edmund J. S Somya-Barce, Michael P. Milham and Rosemary Tannock Characterizing Cognition in ADHD: Beyond Executive Dysfunction.

[19] Yamaguchi Tsuyoshi, Igasaki Tomohiko, Hayashida Yuki, Murayama Nobuki. *Classification of Ataxia using a Self-Organizing Neural Network Model*. **105** (578) (2006), 5–8.

[20] Yuan,Y., and Suarga S. On the Integration of Neural Networks and Fuzzy Logic Systems. IN: *IEEE International Conference on Systems, Man and Cybernetics*, (1995), 452–7.