Self-Organizing Map Approach for Identifying Mental Disorders

Mabruk. Ali Fekihal
Faculty of computing & IT
Sohar University
P.O.Box. 44, P.C. 311, Sohar, Sultanate of Oman

Jabar H. Yousif
Faculty of computing & IT
Sohar University
P.O.Box. 44, P.C. 311, Sohar, Sultanate of Oman

ABSTRACT
Classifications of mental illness such as schizophrenia are very broad; therefore, the proposed approach attains at practical and task-relevant diagnostic categories by use of clustering techniques. A Self-Organizing Feature Map (SOFM) approach was design and implemented for classifying transcribed speech samples and determines mental disorders. An unsupervised Artificial Neural Network was implemented using the NeuroSolution. The proposed classification system is used to determine whether a text or speech sample is generated by a person has mental illness or not. The proposed approach shows clearly that all the categories are identified and classified appropriately, with the proposed SOFM achieving a high accuracy of (97) in the classification phase for predicting the desired output.

General Terms
computer science, artificial intelligence, neural network, applications, soft computing.

Keywords
mental illness, self-organizing map, text clustering, text classification, unsupervised learning.

1. INTRODUCTION
Mental disorders or mental illnesses involve psychological or behavioral patterns that are usually associated with distress or disability, which are not part of the natural evolution of a person or culture. As such, mental disorders can be defined generally through a combination of features that reflect the feelings of a person or his actions and explain his thinking and perceptions. The rapid development in information technology and the significant growth in data transfer needs extensive studies to achieve better progress in classification, data analysis and data extraction [1].

The researchers face new challenges in order to overcome the rapid rise in health care costs faced by the community. They attempt to discover new and innovative techniques through the use of information technology that help to find efficient and inexpensive solutions. [2].

The implementation of artificial intelligence techniques in health care problems can help to diagnose diseases and assess treatment outcomes [3].

Soft computing (SC) is referring to the approaches of calculation that discover approximate solutions for satisfying real-world problems [4]. The SC relies on the use of the neural networks, fuzzy logic, and support vector machines.

The main objective of SC is simulating human capacities such as learning, thinking and problem solving, etc. Recent developments in the field of information technology and the use of soft computing techniques have helped to develop effective automated diagnostic systems [5].

In spite of the great challenges, the use of artificial intelligence techniques can help in developing effective solutions to many medical problems. The most qualified methods for learning from the rare data are using the artificial neural network. The purpose of this paper is to design and implement a text classification model for mental illness using a multilayer perception neural network.

Neural networks have specific features like massive parallelism, uniformity, generalization ability, learnability, trainability and adaptively. These features motivate researchers to approve neural networks based solution in solving problems. Neural approaches have been implemented effectively in many applications such as image processing, NLP, speech recognition, pattern recognition and classification tasks [6, 7].

Professional psychiatrists work hard consciously and unconsciously to investigate the language of their patients with the intention of identifying patterns, and use these patterns to assist in building an obvious clinical diagnosis [8]. Clustering is the process of separating data into groups that have similar features. Clustering is one of the significant themes of research that demonstrates important successes in many fields of research such as statistics, pattern recognition and machine learning [1].

Most machine learning techniques necessitate to labeled samples in the training phase. Nevertheless, obtaining of labeled samples is very limited [9].

2. MENTAL HEALTH BACKGROUND
According to Medilexicon's medical dictionary mental health is "emotional, behavioral, and social maturity or normality". It is all about how we think, feel and behave. Mental health care is important at every stage of life, from childhood and throughout adulthood. The term "mental illness" generally refers to a wide range of mental disorders that can be diagnosed by a health care professional [10].

Mental illness is to become one of the main problems in our society [11]. The World Health Organization anticipated that the world's leading disability will be mainly caused by the depression by 2020. Depression is "a mental state or chronic mental disorder characterized by feelings of sadness, loneliness, despair, low self-esteem, and self-reproach; accompanying signs include psychomotor retardation (or less frequently agitation), withdrawal from social contact, and vegetative states such as loss of appetite and insomnia."[12]

Mental illness can strike individuals in the prime of their lives, often during adolescence and young adulthood. All ages are susceptible to this illness, but the young and the old are especially vulnerable.

Services for people with mental illness problems are different from one country to another, but mainly through welfare programs funded by governments, public and private hospitals [13]. In-patient services are provided in general hospitals and separate psychiatric hospitals. There are also community-based services that provide a range of services such as clinic-based, mobile follow-up and treatment.

Despite significant achievement in mental health in some countries over the last decade there is still a lot be done to address the address the burden of mental ill health as identified in the World Health Organization's/World Bank's Report.

Research in the mental health field has increased resulting in the extensive collection of information and publication dealing with different features of mental health trying to find solutions that solve a wide range of problems [13, 14, 15].

3. NEURAL NETWORKS BACKGROUND

A neural network is a powerful data modeling tool that is able to capture and represent complex input/output relationships both linear and non-linear [1] .There are specific features that stimulate scientists to adopt neural network design themes in the different application fields. The main features are massive parallelism, uniformity, distribution representation and computation, learning ability, trainability, generalization ability, and adaptivity [16]. Neural networks are being successfully applied in a number of areas such as data classifications, resource allocation and scheduling, database mining, speech production and recognition and pattern recognition [1, 3].

The Self-Organizing Map (SOM) is considered as the most popular artificial neural network algorithm in the unsupervised learning category. Many industrial projects use the SOM as a tool for solving hard real-world problems. Many fields of science have adopted the SOM as a standard analytical tool: in statistics, signal processing, control theory, financial analyses, experimental physics, chemistry and medicine.

4. RELATED WORK

Imianvan a.a. & obi j.c. [17] proposed a traditional procedure for the medical diagnosis of Bipolar disorder employed by physician which is analyzed using Neuro-fuzzy inference procedure. Neuro-Fuzzy Logic explores approximation techniques from neural networks to find the parameter of a fuzzy system.

Jabar & Mabruk [5] designed and implemented a multilayered perceptron to classify transcribed speech samples and determine mental health problems. The proposed classification system is used to determine if a text or speech sample was generated by a person with mental health problem, such as schizophrenia or mania. As noted above classifications such as schizophrenia are very broad; therefore, they used clustering techniques to achieve a practical and task-relevant diagnosis. These demonstrated a high accuracy of (99%).

C. R. S. Lopes & et. al. [18] aimed to analyze common mental disorders (CMD) factors using Multilayer Percepton (MLP) trained with Simulated Annealing Algorithm. A Simulated Annealing is used for optimizing both architectures and weights simultaneously. It is aims to identified variables which occur with higher frequency with the mental disorder being studied.

The experiments using neural networks showed that the years of schooling were more often associated with CMD with 89.29%.

Mabruk and Joachim [19] implemented a text mining techniques to determine psychiatric problems using the clustering and classifications techniques. They aims to identify speech samples either was generated by person have mental illness or not. The research achieved high accuracy (92%) and precision (87%) and moderate to low recall.

Gal Kazas & Michael Margaliot [20] proposed an algorithmic approach for visualizing the topology of mental disorders using a self-organizing feature map (SOFM). They trained a SOFM to produce a two-dimensional map of 27 relatively well-known mental disorders. Each disorder is represented by an 82-dimensional input vector describing the symptoms associated with the disorder.

This map shows distinct clusters of mental disorders. Each cluster contains disorders that are similar to one another, and separate from those in other clusters. This provides a way to clearly visualize the topology of these mental disorders.

5. SOFMCONFIGURATION & DESIGN

The Kohonen network or Kohonen's Self-Organizing Map (SOM) or Kohonen's Self-Organizing (Feature) Map SO(F)M is one of the most popular network architectures. SOM provides data visualization techniques which help to understand high dimensional data by reducing the dimensions of data to a map. The main function of SOM networks is providing a way of representing multidimensional data in much lower dimensional spaces - usually one or two dimensions. This process, of reducing the dimensionality of vectors, is essentially a data compression technique known as vector quantization. In addition, the Kohonen technique creates a network that stores information in such a way that any topological relationships within the training set are maintained. SOM represents a clustering concept by grouping similar data together. Therefore it can be said that SOM reduces data dimensions and displays similarities among data. SOFMs are competitive neural networks in which the featured space is represented by organizing the neurons in a 2-dimensional grid (most simple case). According to the learning rule, vectors which are similar to each other in the multidimensional space will be similar in the 2-dimensional space. **SOFM**s are often used just to visualize an *n*-dimensional space, but its main application is data classification.

Teuvo Kohonen[20] writes "The SOM is a new, effective software tool for the visualization of high-dimensional data. It converts complex, nonlinear statistical relationships between high-dimensional data items into simple geometric relationships on a low-dimensional display [4]. As it thereby compresses information while preserving the most important topological and metric relationships of the primary data items on the display, it may also be thought to produce some abstractions."

5.1 Similarity of Data

To determine the Best Matching Unit all weight vectors are run through and distance from each weight to the sample vector is calculated. The weight with the shortest distance is the winner. The most common way to determine the distance is the Euclidean Distance and/or Cosine Distance.

The second layer's neurons have their own weight vector whose each neuron dimension is equal to the dimension of the input layer. The architecture of SOFM is as follows.

The **SOFM** Neural Network consists of two layers of neurons (Figure 1). The first layer which is not a neural layer receives the input data and then transfers it to the second layer.

Each neuron in the second layer has its own weights vector and a dimension equal to the dimension of the input layer. The neurons are connected to adjacent neurons by a neighborhood relation, which illustrates the topology or structure of the map. Such a neighborhood relation is assigned by a special function called a topological neighborhood as depicted in Figure 1.

5.2 Learning rule

Assume that the output nodes are connected in 1 or 2 dimensional array. And assume that all nodes in the input layer are connected to the nodes in output layer that means the network is fully connected.

- Initializing each node's in weights vector is initialized.
- All weights vectors of the second layer's neurons are set to random values.
- Randomly choose an input vector from the set of the training data. This calculates the differences between the input vector and all neurons vectors as follows:
-

$$D_{ij} = |X^l - W_{ij}| = \sqrt{(x_1 - w_{ij1})^2 + \cdots \ldots + (x_n - w_{ijn})^2}$$

- Calculate the most like weight for input vector. The winning node is usually known as the Best Matching Unit (BMU). NN chooses the winner-neuron, i.e., the neuron whose weights vector is the most similar to the input vector:
  $D(k_1,K_2) = \min_{i,j} (D_{i,j})$
  Where k1 and K2 are indices of the winner-neuron.
- Then, a correction of the weights vectors of the winner and all the adjacent neurons is performed. The neighborhood of a neuron is determined by a topological neighborhood function, as follows:

$$h(\rho,t) = \exp\left(\frac{\rho^2}{2\sigma^2(t)}\right)$$

Here, r is the distance to the winner-neuron:

$$\rho = \sqrt{(k_1 - i)^2 + (k_2 - j)^2}$$

ρ is a function dictating the space of the neighborhood. In the beginning of the functioning, it involves almost the whole space of the grid, but with time, the value of ρ decreases. As the attraction function is equal to 1, r equals zero. Shown here is the simplest form of a topological neighborhood function, but in real tasks, its modifications often use the Mexican hat (h (ρ, t)) or the French hat (h (ρ)). They are computed as follows:

$$h(\rho,t) = exp\left(-\frac{\rho^2}{\sigma^{2(t)}}\right)\left(1 - \frac{2}{\sigma^{2(t)}}\rho^2\right)$$

The value of (a) is constant.

$$h(p) = \begin{cases} 1, & |\rho| \le a, \\ -\frac{1}{3}, & a < |\rho| \le 3a \\ 0 & |\rho| < 3a, \end{cases}$$

The weights of all the neurons are updated as follows:

$$W_{ij}(t+1) = W_{ij}(t) + \alpha(t)h((\rho,t)(x^l(t) - W_{ij}(t))$$

α (t) is a learning rate function. The weight vector is updated only if the current neuron is the winner or adjacent to the winner. On each step, the NN determines the neuron whose weights vector is the most similar to the input vector. And then the current weights

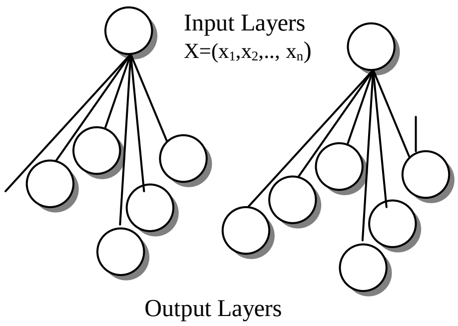

Figure 1. The architecture of an SOFM

vectors and its neighbors' are updated to make them closer to the input vector as depicted in Figure 2.

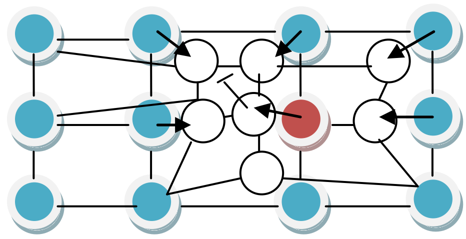

Figure 2: Updating the winner-neuron and its neighbors.

6. Experiments and Results

For the purpose of testing the proposed method , two hundred messages were used which they are randomly selected from both "sci.psychology psychotherapy" and "sci.philosophy". The header and signature information was removed from the text messages for the sake of anonymity [19, 1]. The two newsgroups are selected according to the prior discourse ideas in psychotherapy including diagnosis and the later may contain similar content but commonly must be dissimilar. The classification task is harder since the topic area of the two newsgroups is so similar that cross-posting can occur.

The experiments were performed using 200 text files comprising the original message plus their replies. Further a pre-processing technique is employed on the text samples to encode them into binary values. Data in its current form cannot be used directly as inputs to the neural network, so it needs to be converted into binary values. However, the inputs of the network are transcribed speech samples in binary forms. And, the output will determine whether the transcribed speech samples have been written by someone who has mental illness or not. This is done by searching the text for key symptoms of mental illness [19, 1, and 16].

The experiments used the clustering techniques in order to extract the task-relevant diagnostic group from psychiatric reports. In this paper the experiments were done to determine the nine clusters *symptoms of mental illness* (ADHD, attention deficit, disorder, hyperactivity, apprehensive, deranged, psychotherapy and none of them).

Typically the reports are comprised of the patient's biographic, referral information, as well as a description of indicators and a psychiatrist's proposition on treatment recommendations.

The data set is separated into three sets, first is the training data set which is used to train the network. It is used to produce the best neuron weight for generalizing the network. The experiments were used about 500 data sets for the training process. The second data set is the cross validation which is used usually to compute the error in a test data set at the same time that the network is being trained. The experiments used about 400 data sets for cross validation process. And the third set is testing data which is used to determine the best network performance. The experiments used about 71 data sets for the testing of network performance. The weights of best network are automatically saved during the training phase and will be loaded into the network before the testing process is come over.

The proposed system are performed using the NeuroSolutions software package. This is to specify the shape and size of the feature map generated by the unsupervised (Kohonen) layer. Besides, it determines the number of Hidden Layers in the network of the MLP connected to its output. One Layer was used as Hidden Layer. The size and dimension of the unsupervised output space is 5 Rows by 5 Columns of the 2D neural field. Increasing the number of Process Elements (PEs) will improve the resolution, but also amplify the training times. The experiments were used 12 and 9 PEs for input and output layers respectively. Basically, the size of the Kohonen layer must be maintained as small as possible because the SOFM trains slowly. The Starting and Final Radius are 2, 0 respectively. Every PE has an adaptive parameter which it must change according to some pre-specified procedure. We used the Back propagation learning technique which is consider as the most common approach of learning in neural network. The weights are changed based on their previous value and a correction term. The learning rule is the means by which the correction term is specified. The momentum component is used as learning rule. The step size is 1 and the momentum rate is 0.7. The Genetic optimization is used to find the best values for these parameters.

There are several ways to test the networks performance [21]. Usually, the mean squared error MSE is used. Easley, it is a function of two times the average cost. It is computed as follows:

$$MSE = \frac{\sum_{j=0}^{P}\sum_{i=0}^{N}(d_{ij} - y_{ij})^2}{NP}$$

Where, P is the number of output processing elements. N is the number of exemplars in the data set. yij is the network output for exemplar i at processing element j. dij the desired output for exemplar i at processing element j. The outputs of the network for three runs are summarized in Table 1, which records the minimum and final MSE. The Figure 3 illustrates the graph of the average MSE with Standard Deviation Boundaries for 3 Runs. The average of final MSE standard deviation for training data set is (0.008409), while the average of final MSE Standard deviation for cross validation data set is (0.008409). With the aim of testing the effectiveness of the proposed network, the experiments are conducted using about 1000 data sets in the testing phase which they extracted from the total data. In supervised learning a confusion matrix is used as a visualization tool which displays the information of the number of correct and incorrect predictions done by the system compared with the actual classifications in the test data. The size of this matrix is n-by-n, where n is the number of categories. The Performance of such systems is evaluated using the data in the matrix. Each column of the matrix represents a value of the output of network, while each row represents the desired values [3]. Table 2 shows a confusion matrix for the network testing.

The rows present the number of actual classifications in the test data. The columns present the number of predicted classifications done by the network. The table 2 clearly illustrates the number of correct and incorrect predictions of the network.

Table 1. Minimum and final MSE

Best Networks	**Training**	**Cross Validation**
Run #	2	2
Epoch #	999	999
Min. MSE	0.008409	0.008080
Final MSE	0.008409	0.008080

Table 2. The Confusion matrix of the network testing

Output Desired	O1	O2	O3	O4	O5	O6	O7	O8	O9
O1	3	0	0	0	0	0	0	0	0
O2	0	3	0	0	0	0	0	0	0
O3	0	0	3	0	0	0	0	0	~~1~~
O4	0	0	0	4	0	0	0	0	0
O5	0	0	0	0	1	0	0	~~1~~	0
O6	0	0	0	0	0	0	0	0	0
O7	0	0	0	0	0	0	1	0	0
O8	0	0	0	0	0	0	0	0	0
O9	0	0	0	0	0	0	0	0	54
Correct %	100	100	100	100	100	100	100	0	98

Table 3 summarizes the results of testing the network. It shows clearly that all the categories are identified and classified well. Consequently, the proposed SOFM achieves a high accuracy of (97) in classification task and predicting the desired output. Accuracy refers to the percentage of correct predictions done by the network when compared with the actual classifications in the test data.

Table 3. The result of testing the network

	O1	**O2**	**O3**	**O4**	**O5**	**O6**	**O7**	**O8**	**O9**
MSE	0.0002	0.0003	0.0126	0.0002	0.0071	0.0001	0.0000	0.0071	0.0127
NMSE	0.0062	0.0086	0.3121	0.0033	0.5111	#DIV/0	0.0026	0.5120	0.0727
MAE	0.0080	0.0129	0.0223	0.0066	0.0173	0.0039	0.0038	0.0167	0.0201
Min. Abs Error	0.0002	0.0036	0.0003	0.0003	0.0002	0.0000	0.0000	0.0000	0.0001
Max. Abs Error	0.0552	0.0554	0.9342	0.0547	0.5037	0.0551	0.0177	0.5520	0.9399
Correlation (r)	0.9972	0.9973	0.8584	0.9986	0.6992	#DIV/0	0.9989	0.7026	0.9638

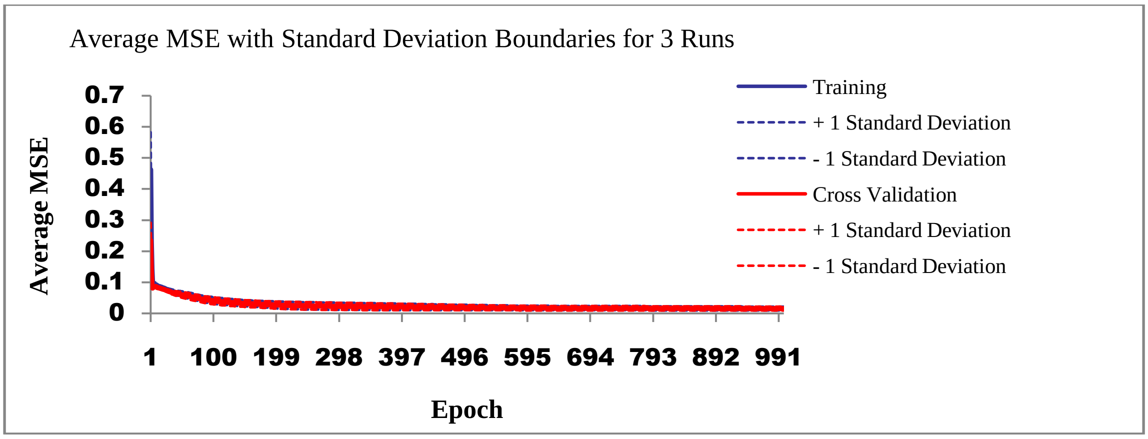

Figure.3 The Average MSE with Standard Deviation Boundaries for 3 Runs

7. COMPARISON AND CONCLUSION

7.1 Comparison with related work

The results of the comparison of our work with other researcher [19, 5, and 1] related mental health studies are summarized in table 4.

7.2 Conclusion

The high incidence of mental illness and associated costs to health care providers has led to increased research into factors that help reduce the risk of mental illness. Despite such efforts, the exact causes of many mental illnesses are still unclear. Research has however established that many specific mental illnesses correlate with a number of chronic physical health conditions such as high blood pressure, diabetes, and HIV/AIDS, resulting in higher cost to the health system [13, 10]

Table 4. Comparison results of proposed approach with other researchers

	MLP [13]	Joachim [14]	Mabruk [1]	Our SOFM
	NN	SVM	SVM	NN
Data size - training phase	20%	100%	100%	40%
Accuracy	99%	80%	92%	97%

The present study aims to identify mental illness through the analysis of texts, in a series of computer based experiments. The current work used the same data text set as used in related studies [19, 1]. This is only to test the experiments in same environment and under the same conditions. We designed and implemented SOFM network to analyze transcribed speech samples of patients, thereby determining if any of the newsgroup messages were written by people who had a mental illness or not.

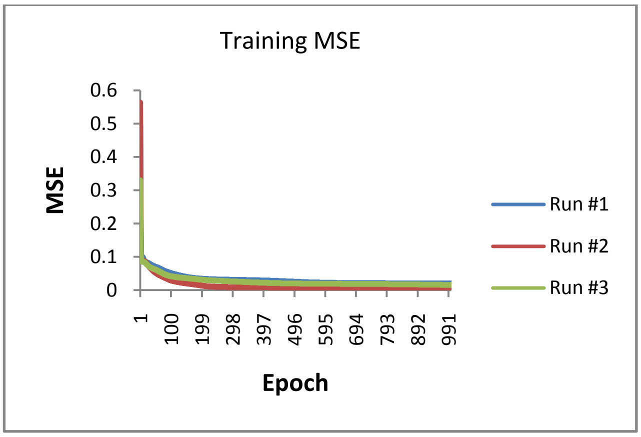

Figure.4 The result of network for training data through 3 runs

We analyzed and implemented nine sings of mental illness (O1 is ADHD , O2 is attention O3 is deficit , O4 is disorder, O5 is hyperactivity , O6 is apprehensive, O7 is deranged, O8 is psychotherapy and O9 is none of them).

Nevertheless, the robustness of the machine learning classification and clustering is important which was also implemented in this paper. It has been shown that the SOFM network is reliable and achieved with a high accuracy of (97%) in classification of the input text, and predicting if it is a type of mental illness or not. Besides, the proposed approach uses a small number of data to complete the classification process and train a network which is useful in experiments with rare data.

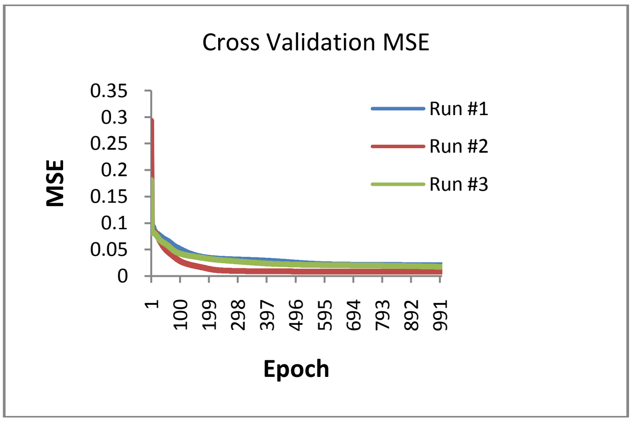

Figure.5 The result of network for cross validation through 3 runs

8. FUTURE WORK

The experiments proved that more attention must be devoted to follow the history record of patient. It may gives key features which identifying the illness's causes. Besides, the automatic pre-processing phase including the reading, encoding, and clustering of data is needed to reduce human errors. Furthermore, implementation of other soft computing techniques such as a hybrid of ANN and genetic algorithms to enhance the results is required.

9. REFERENCES

[1] Joachim Diederich , et.al, (2007). Ex-ray: Data mining and mental health, Applied Soft Computing 7 (2007) 923–928. doi:10.1016/j.asoc.2006.04.007.

[2] Lodhi, H.,Saunders, C.,Shawe-Taylor, J.,Cristianini, N., &Watkins, C. (2002). Text classification using string kernels. Journal of Machine Learning Research, 2, 419–444. Paris: EC2 & Cie. http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.130.2853

[3] Stehman, Stephen V. (1997). "Selecting and interpreting measures of thematic classification accuracy". Remote Sensing of Environment 62 (1): 77–89. doi:10.1016/S0034-4257(97)00083-7

[4] Zadeh, L.A. (2001). Applied Soft Computing. Applied Soft Computing 1, 1–2.

[5] Jabar H. Yousif, and Mabruk A. Fekihal," Neural Approach for Determining Mental Health Problems" , Journal of Computing, Volume 4, Issue 1, pp6-11 ISSN 2151-9617 ,NY, USA, January 2012. http://www.scribd.com/JournalofComputing/d/81059229-Neural-Approach-for-Determining-Mental-Health-Problems

[6] Akyol, D. E. (2004). Applications of neural networks to heuristic scheduling algorithms. Computers and Industrial Engineering, 46, 679-696. http://www.sciencedirect.com/science/article/pii/S036083520400066X

[7] Nazeran, H., &Behbehani, K. (2001). Neural networks in processing and analysis of biomedical signals. In M. Akay (Ed.), Nonlinear biomedical signal processing: Fuzzy logic, neural networks and new algorithms (pp. 69–97).8 Teuvo Kohonen, "Self-Organizing Maps" (3rd edition) Springer, ISBN 3540679219 , 2001.

[8] Teuvo Kohonen, "Self-Organizing Maps" (3rd edition) Springer, ISBN 3540679219 , 2001.

[9] Bonissone, P (2002). Hybrid Soft Computing for Classification and Prediction Applications. Conferencia Invitada. 1st International Conference on Computing in an Imperfect World (Soft-Ware 2002), Belfast. DOI: 10.1007/3-540-46019-5_28

[10] Hovenga, E. J. S. (2004). Globalisation of health and medical informatics education: What are the issues? International Journal of Medical Informatics, 73(2), 101–109. DOI: 10.1016/j.ijmedinf.2003.11.004

[11] Adelman, H. & Taylor, L. (2006). The current status of mental health in schools: A policy and practice analysis. Los Angeles, CA: Center for Mental Health in Schools at UCLA. http://smhp.psych.ucla.edu/currentstatusmh.htm

[12] MediLexicon International Ltd,Craythorne House, Burnside Mews, London Road, Bexhill-on-Sea, TN39 3LE.http://www.medilexicon.org

[13] Hadzic, M., Chen, M., & Dillon ,T. (2008) . Towards the mental health ontology , proceeding of the IEEE International Conf. on Bioinformatics and Biomedicine, USA. DOI: 10.1109/BIBM.2008.59

[14] Clarke, G., Lynch, F., Spofford, M., & DeBar , L. (2006). Trends influencing future delivery of mental health services in large healthcare systems. Clinical Psychology: Science and Practice, 13 (3) , 287-292. doi:10.1111/j.1468-2850.2006.00040.

[15] Patel, V., Prince, M. (2010). Global mental health - a new global health field comes of age. JAMA, 303, 1976-1977. doi: 10.1001/jama.2010.616DOI: 10.2105/AJPH.2010.192138.

[16] Lippman R. An introduction to computing with neural nets. IEEE Trans. ASSP Magazine 4, 4-22, 1987.

[17] Imianvan A.A. and Obi J.C.: Diagnostic evaluation of Hepatitis utilizing Fuzzy Clustering Means, World Journal of Applied Science and Technology, Vol.3. No.1 (2011). 23-3,http://wojast.com/fullpaper_vol3_1/23-30_Imianvan.pdf

[18] C. R. S. Lopes & et.al. T. B. Ludermir1, M. C. P. de Souto1 e A. B. Ludermir , "Neural Networks for the analysis of Common Mental Disorders Factors." Proceeding SBRN '02 Proceedings of the VII Brazilian Symposium on Neural Networks (SBRN'02), IEEE Computer Society Washington, DC, USA ©2002 , ISBN:0-7695-1709-9

[19] Abusaa, M., Diederich, J. and Al Ajmi, A. (2004). Web mining and mental health. In: IAWTIC 2004 Proceedings. International Conference on Intelligent Agents, Web Technologies and Internet Commerce, Gold Coast, Queensland. 12-14 July 2004. http://joachimdiederich.com/assets/IAWTIC04.pdf

[20] Gal Kazas & Michael Margaliot , "Visualizing the topology of mental disorders using selforganizing feature maps". http://www.eng.tau.ac.il/~michaelm/kazas.pdf

[21] Beyer, H.-G., &Schwefel, H.-P. (2002). Evolution strategies: A comprehensive introduction. Natural Computing: An International Journal, 1(1), 3–52. DOI: 10.1023/A:1015059928466.