Predicting the Class of a Mentally Disabled Patient to Check the Level of Mental Retardation by using Feed Forward Back Propagation Neural Network

P.C.Gupta
Jaipur National University,
Jaipur, Rajasthan, India

Anjali-Mathur
Jaipur National University,
Jaipur, Rajasthan, India

ABSTRACT
Mental disorders have a large impact on individuals, families, and communities, and are one of the main causes worldwide of disability and distress. Correct diagnosis of mental disorders is essential in clinical practice, pharmacological research, and successful treatment. Patients with mental retardation often have multiple and sometimes complicated medical problems. In this paper we have proposed a feed forward back propagation neural network to classify the level of mental retardation by using Matlab software. There are six neurons in the input layer which represent the attribute of a patient. Output layer contains four neurons which represent four different levels of mental retardation in which each patient will be classified

Keywords
Electroencephalogram, Matlab, Artificial Neural Network, Feed Forward Back Propagation.

1. INTRODUCTION
Brain is the synchronized structure made of billions of neuronal network. Neurones and their synaptic as well as dendriatic connections join many other neurones. This makes an integrated unit. [1] According to the modern theories number of neurones are not so important for excellent brain model, rather their multiplicity of synaptic and dendriatic connections along with obscure role of glial cells. In case of damage these interconnections and their supporting cells, glial cells, suffer much, making their working inefficient to contribute to intelligence. It is a cause of mental retardation [1]. Mental Retardation is a condition in which there is delay or deficiency in all aspects of development, i.e. there is global and noticeable deficiency in the development of motor, cognitive, social, and language functions. This is the commonest form of developmental disability. In many ways, mental retardation is also representative of developmental disabilities in general, in its causation, nature, and care. Unfortunately, the etiology and pathogenesis of many mental disorders are still unknown. Psychiatrists must thus resort to classify disorders according to their symptoms. Clinical interviews with the patient (and sometimes family relatives), and observations of the patient's attitude and behavior are used to elicit symptoms and then establish a diagnosis. Mental retardation is a common condition. There is a shameful practice, very common in Asia region, is chained the patients to poles or their beds to cure them. In the Erwadi Tragedy in India in 2001, over 20 people with mental illness were burned to death after a fire swept through an institute in which they were chained to their beds. In surveys in the general population in India [2] among people of all ages, it has been found that around 2% have mental retardation. But if one estimates the problem only in children, (under 18 years of age) there will be about 3% of cases with mental retardation among all children under 18 years of age in the same village. Regarding learning disability, a study by UNICEF in Sri Lanka revealed that 12% of primary school children had learning disability. Another report from Sri Lanka estimated that 15% of school going children suffered from some form of disability. A study in children (aged 2-9 years) from Bangladesh found that around 7% had some form of disability. Mental retardation, the second most common form of disability, was seen in around 2% of children. Severe mental retardation in Bangladeshi children (2-9 years old) was estimated to be around 6 per 1000, in keeping with the reports from other countries. In 1999, the Planning Division, Department of Mental Health of Thailand conducted an epidemiological study on mental health problems countrywide and found that the rate of occurrence of mental retardation was 1.3% [2].

2. LITERATURE SURVEY
Different work has already been done for mental disabilities. [3-14] A research investigates an appropriate neural classifiers for the recognition of mental tasks from on-line spontaneous EEG signals. The classifiers are to be embedded in a portable brain-computer interface. One more research shows a brain computer interface design based on mental tasks with a zero false activation rate. The classification is performed through a radial basis function neural network. The order of autoregressive model varied from 2 to 20. In the year 2009 a research work was proposed for Classification of Five Mental Tasks from EEG Data by using principal component analysis of Neural Network. The principal component analysis (PCA) was used for feature extraction of the relevant frequency bands from raw electroencephalogram (EEG) signals. In the year 2011 another research work was proposed as Reliable identification of mental tasks using time-embedded EEG and sequential evidence accumulation. In this work Eleven channels of EEG were recorded from a subject performing four mental tasks. A time-embedded representation of the untransformed EEG samples were constructed. Classification of the time-embedded samples were performed by linear and quadratic discriminate analysis and by an artificial neural network. Some research work were based on Electroencephalogram in which EEG from one subject performed three mental tasks had been classified using Radial Basis Function and Support Vector Machines (SVM) to control over fitting. A method for EEG pre-processing based on Independent Component Analysis (ICA) was proposed and three different feature extraction techniques were compared:

Parametric Autoregressive (AR) modeling, AR spectral analysis and power differences between four frequency bands. In another work the neural network has been used to classify EEG data by using autoregressive with maximum likelihood pre-process for epileptic seizure detection. The purpose of the work was to investigate the use of autoregressive (AR) model by using maximum likelihood estimation (MLE) also interpretation and performance of this method to extract classifiable features from human electroencephalogram (EEG) by using Artificial Neural Networks. Since the voltages recorded on an electro-encephalograph are the result of many processes that occur simultaneously in the brain, only events that involve larger areas of the brain, such as epileptic seizures, can be readily identified on the EEG recording. For this reason, mental disorders generally cannot be diagnosed from the electro-encephalograph. In this paper back-propagation feed forward neural network have been used to identify the status of mental retarded patient [3-14].

3. METHODOLOGY

Mental Retardation can be classified in four major classes as mild retardation, moderate retardation, severe retardation and profound retardation. Mild mental retardation is much more common than severe mental retardation, accounting for 65 to 75% of all cases with mental retardation. Looked at in another way, in a village of 1000 people, of the 20 who will have mental retardation, about 15 will have mild mental retardation and about five will have more severe forms. However social criteria is determined by individual's capacity to meet the demands of surrounding people. The targets for individual here will be different from each other and will depend on life situation and social and cultural contexts. Majority of people with lower IQ, will still manage to meet up to the demands of the society.

3.1 Mild Mental Retardation

Many of the characteristics of Mild mental retardation correspond to those of Learning Disabilities. The intellectual development will be slow, however, these candidates have the potential to learn within the regular classroom given appropriate modifications and/or accommodations. Approximately 85% of the mentally retarded population is in the mildly retarded category. Their IQ score ranges from 50–70. They can become fairly self-sufficient and in some cases live independently, with community and social support.

3.2 Moderate mental retardation

About 10% of the mentally retarded population is considered moderately retarded. Moderately retarded persons have IQ scores ranging from 35–55. They can carry out work and self-care tasks with moderate supervision. They typically acquire communication skills in childhood and are able to live and function successfully within the community in supervised environments.

3.3 Severe mental retardation

About 3–4% of the mentally retarded population is severely retarded. Severely retarded persons have IQ scores of 20–40. They may master very basic self-care skills and some communication skills. Many severely retarded individuals are able to live in a group home.

3.4 Profound mental retardation

Only 1–2% of the mentally retarded population is classified as profoundly retarded. Profoundly retarded individuals have IQ scores under 20–25. They may be able to develop basic self-care and communication skills with appropriate support and training. Their retardation is often caused by an accompanying neurological disorder. Profoundly retarded people need a high level of structure and supervision.

This work is to predict the level of mental retardation in the patients. For this we have keenly observed 20 patients suffering from mental retardation. These patients belongs to different societies, different classes, different age groups and from different family backgrounds. For this purpose we have conducted a survey work at Jaipur district and choose the patients from urban and rural areas, from mental hospitals and NGOs. We haven taken help from neurologists, psychiatrist parents, nurses, care-takers and relatives of these patients. To check the level of Mental retardation we have also used the Diagnostic and Statistical Manual of Mental Disorders criteria –IV. The Diagnostic and Statistical Manual of Mental Disorders is an ubiquitous assessment tool for getting the symptoms of mental disorders, it was compiled by the American Psychiatric for about 300 mental disorders.

Table 1 : Developmental Characteristics Related to Level of Mental Retardation (DSM-IV Criteria)

Mild retardation	Moderate retardation	Severe retardation	Profound retardation
Function at one half to two thirds of chronological age (IQ: 50 to 70)	Function at one third to one half of chronological age (IQ: 35 to 49)	Function at one fifth to one third of chronological age (IQ: 20 to 34)	Function at < one fifth of chronological age (IQ: < 20)
Slow in all areas	Noticeable delays, especially in speech	Marked and obvious delays; may walk late	Marked delays in all areas
May have no unusual physical signs	May have some unusual physical signs	Little or no communication skills but may have some understanding of speech and show some response	Congenital abnormalities often present
Can acquire practical skills	Can learn simple communication	May be taught daily routines and repetitive activities	Need close supervision
Useful reading and math skills up to grades 3 to 6 level	Can learn elementary health and safety habits	May be trained in simple self-care	Often need attendant care
Can conform socially	Can participate in simple activities and self-care	Need direction and supervision	May respond to regular physical activity and social stimulation

To predict the level of mental retardation we have used Artificial Neural Network. Artificial Neural Network is currently a ‘hot’ research area in medicine and it is believed that this will receive extensive application to biomedical systems in the next few years. In Artificial Neural Network, the knowledge lies in the interconnection weights between neurons. Therefore, training process is an important characteristic of the ANN methodology, whereby representative examples of the knowledge are iteratively presented to the network, so that it can integrate this knowledge within its structure. No assumption is needed about the underlying data probability distribution when ANN is used for pattern classification. Once trained, it can be configured to perform adaptively to improve its performance over time.

In most applications of Multi Layer Propagation, the weights are determined by means of the Back propagation algorithm, which is based on searching an error surface (error as a function of ANN weights) using gradient descent for points with minimum error. During the training phase, the weights are successively adjusted based on a set of inputs and the corresponding set of desired output targets. Each iteration in back propagation constitutes two sweeps: forward activation to produce a solution, and a backward propagation of the computed error to modify the weights. The forward and backward sweeps are performed repeatedly until the ANN solution agrees with the desired value within a pre-specified tolerance. The back propagation algorithm provides the needed weight adjustments in the backward sweep. The back-propagation algorithm is a nonlinear procedure because of the nonlinear threshold element contained in each node, and its behavior is very complex because of the layered structure.

4. BACK PROPAGATION ALGORITHM

It is a supervised learning method, and is a generalization of the delta rule. It requires a teacher that knows, or can calculate, the desired output for any input in the training set. Backpropagation requires that [15] the activation function used by the artificial neurons be differentiable. As the algorithm's name implies, the errors propagate backwards from the output nodes to the inner nodes. Technically speaking, backpropagation calculates the gradient of the error of the network regarding the network's modifiable weights. This gradient is almost always used in a simple stochastic gradient descent algorithm to find weights that minimize the error. Often the term "backpropagation" is used in a more general sense, to refer to the entire procedure encompassing both the calculation of the gradient and its use in stochastic gradient descent. Backpropagation usually allows quick convergence on satisfactory local minima for error in the kind of networks to which it is suited. Backpropagation networks are necessarily multilayer perceptrons (usually with one input, one hidden, and one output layer). In order for the hidden layer to serve any useful function, multilayer networks must have non-linear activation functions for the multiple layers: a multilayer network using only linear activation functions is equivalent to some single layer, linear network.

4.1 Working With Back-Propagation

The application of the generalized delta rule thus involves two phases: During the first phase the input x is presented and propagated forward through the network to compute the output values y_o^p for each output unit. This output is compared with its desired value d_o^p, resulting in an error signal δ_o^p for each output unit. The second phase involves a backward pass through the network during which the error signal is passed to each unit in the network and appropriate weight changes are calculated. The training algorithm of back propagation involves four stages, these are:-

1. Initialization of weights
2. Feed Forward
3. Back Propagation for errors
4. Updating of the weights and biases.

During first stage some small random values are assigned to initialize the weights. In next stage each input unit (X_i) receives an input signal and transmits this signal to each of the hidden units $z_1 \ldots z_p$. Each hidden unit then calculates the activation function and sends its signal z_j to each output unit. The output unit calculates the activation function to form the response of the net for the given input pattern.

During back propagation of errors, each output unit compares its computed activation y_k with its target value t_k to determine the associated error for that pattern with that unit. Based on the error, the factor δ_k(k=1, ….,m) is computed and is used to distribute the error at output unit y_k back to all units in the previous layer. Similarly, the factor δ_j(j=1,…, p) is computed for each hidden unit z_j. During final stage, the weight and biases are updated using the δ factor and the activation.

4.2 Weight Adjustments With Sigmoid Activation Function

- The weight of a connection is adjusted by an amount proportional to the product of an error signal δ, on the unit k receiving the input and the output of the unit j sending this signal along the connection:

$$\Delta_p w_{jk} = \gamma \delta_k^p y_j^p$$

- If the unit is an output unit, the error signal is given by $\delta_o^p = (d_o^p - y_o^p) F(s_o^p)$, take as the activation function F the 'sigmoid' function as defined $y^p = F(s^p) = \frac{1}{1+e^{-sp}}$. In this case the derivative is equal to :

$$F(s^p) = \frac{\partial}{\partial s^p} \frac{1}{1+e^{-sp}}$$

$$= \frac{1}{(1+e^{-sp})^2}(-e^{-sp})$$

$$= \frac{1}{(1+e^{-sp})} \frac{e^{-sp}}{(1+e^{-sp})}$$

$$= y^p(1-y^p)$$

such that the error signal for an output unit can be written as:

$$\delta_0^p = (d_0^p - y_0^p) y_0^p (1-y_0^p)$$

- The error signal for a hidden unit is determined recursively in terms of error signals of the units to which it directly connects and the weights of those connections. For the sigmoid activation function:

$$\delta_h^p = F(s_h^p)\sum_{0=1}^{N_0}\delta_0^p wh_0 = y_h^p(1-y_h^p)\sum_{o=1}^{N_0}\delta_0^p wh_0$$

4.3 Learning Rate And Momentum

The learning procedure requires that the change in weight is proportional to

$$\partial E^p / \partial w.$$

The constant of proportionality is the learning rate. For practical purposes we choose a learning rate that is as large as possible without leading to oscillation. One way to avoid oscillation at large, is to make the change in weight dependent of the past weight change by adding a momentum term:

$$\Delta w_{jk}(t+1) = \gamma\delta_k^p y_j^p + \alpha\Delta w_{jk}(t),$$

where the indexes the presentation number and F is a constant which determines the effect of the previous weight change. Although, theoretically, the back-propagation algorithm performs gradient descent on the total error only if the weights are adjusted after the full set of learning patterns has been presented, more often than not the learning rule is applied to each pattern separately, i.e., a pattern p is applied, E^p is calculated, and the weights are adapted (p = 1, 2,..... P). There exists empirical indication that this results in faster convergence. Care has to be taken, however, with the order in which the patterns are taught. For example, when using the same sequence over and over again the network may become focused on the first few patterns. This problem can be overcome by using a permuted training method [15].

5. OBSERVATIONS, RESULTS AND DISCUSSIONS

In our survey work we keenly observe 20 patients. 700 samples of these 20 patients have been collected. We had taken 6 attributes. The attributes are based on special characteristics of patients based on different criteria. The attributes are general symptoms, which can be easily detected for a patient.

(a) IQ level at chronological age:- Every person has some IQ level which get changed according to age. A normal person has a specific range of IQ level at every state of age, but a mentally retardate patient has choked IQ level.

(b) Physical Activeness :- This is a feature which differs from person to person according to their physical growth and the age, but on an average every person shows some physical activeness. The patients of mentally retardation can show general activeness or hyper activeness or in some cases the patient could be completely dull. Therefore it becomes an important symptom to diagnose the patient.

(c) Congenital abnormalities:- Generally some patients have inborn abnormalities such as weak physical structure, problems with vision, hearing or speech, suffering from cerebral palsy, etc. Presence or absence of such abnormalities are very useful symptom to classify the patient.

(d) Self Manageable :- The level of mental retardation is based on the self managing ability of the patient. The patients of mild retardation can take care of themselves such as comb their hair, wear the dress and shoes, eat food, talk and walk easily, etc.

(e) Learning skills:- The ability to learn any new thing is a normal human tendency. Mental patients have loose this ability. Patients of mild retardation can learn some things but not all and they require more time to learn, but the patients of profound retardation can't learn anything.

(f) Social sense:- The civic sense shows the normality of a person. Extremely reserve nature, showing shyness or start shouting at public places are the symptoms of mental illness. Level of mental retardation can be classified on this basis also.

We graded these attributes in the range of 0 to 0.9 as the value of attributes. Presence or absence of the attribute is graded in this range i.e. between 0.0 to 0.9. This grading system is important because we are classifying the patients in four classes on the basis of the values of these attributes.

Table 2 : Attributes and its values

Sl. No.	Attributes	Attribute Identifier	Grade Range
1	IQ level at chronological age	A1	0.0– 0.9
2	Physical activeness	A2	0.0– 0.9
3	Congenital abnormalities	A3	0.0– 0.9
4	Self manageable	A4	0.0– 0.9
5	Learning skills	A5	0.0– 0.9
6	Social sense	A6	0.0– 0.9

For each patient the combination of these six values forms the basis of input in the neural network. The output of neural network shows the level of mental retardation. Each patient has been classified in either of these four classes.

Table 3 : Classes

Sl. No.	Class	Class Identifier	Category
1	Mild retardation	C1	0
2	Moderate retardation	C2	1
3	Severe retardation	C3	2
4	Profound retardation	C4	3

Each class in the above table shows the level of mental retardation.

Table 4 contains different classes and the range of values which could be contained in each class.

Table 4 : Distribution of graded values of attributes for different classes

Attributes	C1	C2	C3	C4
A1	0.7– 0.9	0.4– 0.6	0.1– 0.3	0.0
A2	0.7-0.9	0.4– 0.6	0.1– 0.3	0.0– 0.1
A3	0.0– 0.3	0.4– 0.6	0.6– 0.8	0.8– 0.9
A4	0.7- 0.9	0.4– 0.6	0.1– 0.3	0.0
A5	0.7– 0.9	0.4– 0.6	0.1– 0.3	0.1– 0.0
A6	0.7– 0.9	0.4– 0.6	0.1– 0.3	0.1– 0.0

To classify a sample value in any single class we require the graded range of all the attributes. Following table contains sample data of 20 patients and the combination of attributes for each patient.

Table 5 : Sample Data

Sl.No.	Class	ATTRIBUTES					
		A1	A2	A3	A4	A5	A6
1	0	0.7	0.7	0.3	0.8	0.6	0.6
2	1	0.5	0.8	0.4	0.5	0.4	0.5
3	2	0.4	0.5	0.6	0.4	0.5	0.3
4	3	0.2	0.3	0.3	0.2	0.2	0.2
5	0	0.5	0.6	0.7	0.7	0.5	0.4
6	1	0.6	0.5	0.6	0.4	0.4	0.4
7	2	0.3	0.7	0.2	0.5	0.4	0.2
8	3	0.2	0.2	0.5	0.3	0.4	0.4
9	0	0.6	0.3	0.2	0.6	0.5	0.7
10	1	0.7	0.6	0.4	0.5	0.4	0.4
11	2	0.5	0.6	0.7	0.4	0.4	0.5
12	3	0.3	0.4	0.7	0.4	0.3	0.1
13	0	0.6	0.8	0.3	0.7	0.6	0.3
14	1	0.6	0.2	0.4	0.6	0.5	0.6
15	2	0.4	0.8	0.7	0.5	0.3	0.6
16	3	0.0	0.2	0.3	0.0	0.1	0.0
17	0	0.6	0.3	0.6	0.7	0.5	0.4
18	1	0.5	0.8	0.6	0.4	0.5	0.5
19	2	0.3	0.5	0.7	0.3	0.2	0.3
20	3	0.2	0.1	0.5	0.3	0.4	0.1

Each input sample consists of six bits in which one bit represent one attribute while output contains four bits in which one bit represent one class of the output.

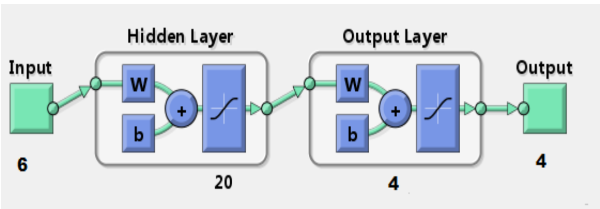

Fig 1 : Sigmoid Diagram

Different samples are applied on the network to perform three different activities:-

Training: It is used to adjust the weights on the neural network.

Validations: It is used to minimize over fitting i.e. for verifying that any increase in accuracy over the training data set.

Testing: It is used only for testing the final solution in order to confirm the actual predictive power of the network.

We have conducted several training sessions. The training uses Scaled Conjugate Gradient function and it measures the performance on the basis of Mean Squared Error. Mean Square Error (mse) is the average squared difference between outputs and targets. Lower values of mean square errors are considered as better ones while zero denotes no error.

Training Session 1 :

In the first training session we have taken 1000 epochs and the training was completed in 50 iterations. Figure 2 shows a brief description of first training session.

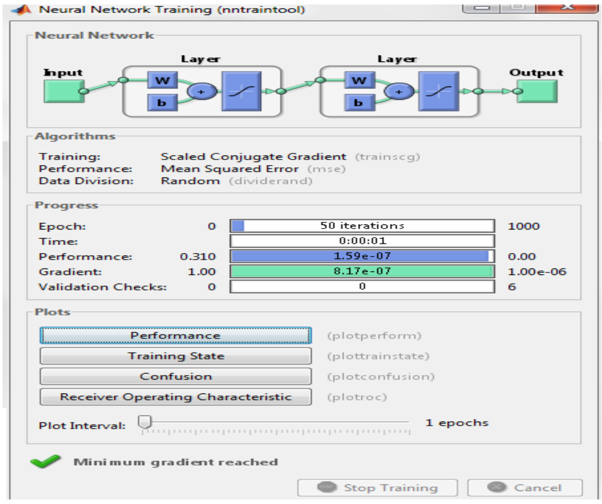

Fig 2 : Result 1

Fig.2 describes time taken for training, total number of epochs used in training, performance in terms of total errors and the value of Gradient (as the training function is Scaled conjugate Gradient) and validation checks. The more specific results are shown in the following figures:

Figure 3 shows a consolidated result of all three types of samples i.e. training, validation and testing. In our training session1 we get the Best validation performance 0.002428 at 50 epochs.

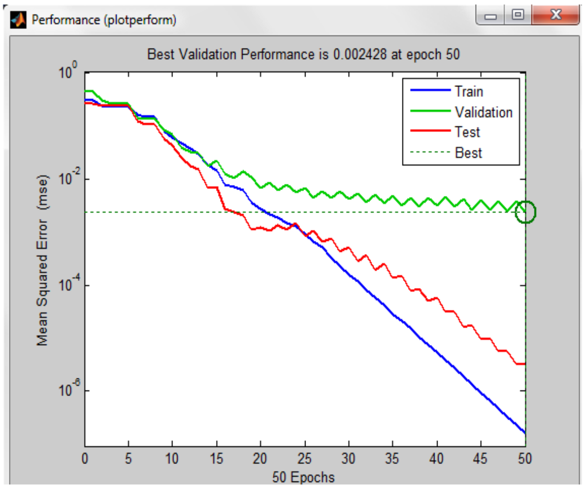

Fig 3 : Performance 1

During the time of training the value of gradient and validation checks are shown separately in the following figure. The training session takes 8.1669e-007 gradient value at 50 epochs.

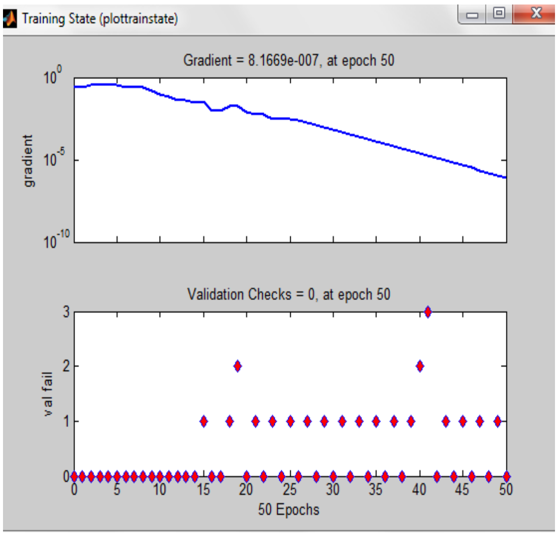

Fig 4 : Training1

The training is not completed until the percent error is zero. Percent error (%e) indicates the fraction of samples which are misclassified. A value of 0 means no misclassifications while 100 indicate maximum misclassifications. In this training session we get the following error results:

Results

	Samples	MSE	%E
Training:	630	1.45828e-1	43.54838e-0
Validation:	35	1.18428e-1	33.33333e-0
Testing:	35	1.00338e-1	29.62962e-0

Fig 5 : Error Report (session 1)

It shows that mean square errors and percent errors both are not zero that is the network is not trained therefore some more training sessions are required.

We made several training sessions on the same sample set to train the network. In our 7th session we get zero percent errors. This shows that the network is completely trained now.

Results

	Samples	MSE	%E
Training:	630	1.93402e-3	0
Validation:	35	6.07197e-3	0
Testing:	35	2.53236e-3	0

Fig 6 : Error Report (Session 10)

Training automatically stops when generalization stops improving, as indicated by an increase in the mean square error of the validation samples.

6. CONCLUSION

In this research work we have predicted the level of mental retardation of the patient on the basis of their symptoms and behaviors. Following points are derived after the complete work:

1. 700 samples of 20 patient have been collected from different sources.
2. Each sample possesses six attributes, which represent symptoms and behavior of each patient.
3. Patients are classified into any of the four classes on the basis of the level of mental retardation.
4. Feed forward back propagation neural network tool is used to train, test and validate the network.
5. 630 samples have been used for training of the network and 70 samples are used for testing.
6. Training of each sample is performed up to 10 times to minimize the error.
7. 70 unknown samples are applied in the network to find the result. Network correctly classifies 54 samples into their respective classes. Therefore network is around 80% accurate.

From this work we are able to identify level of mental retardation of a patient if its symptoms and behavior are known.

7. REFERENCES

[1] http://en.wikipedia.org/wiki/Backpropagation

[2] http://bjp.rcpsych.org/content/179/1/85.1.

[3] Millan, J.delR., Mourino, J., Cincotti, F., Varsta, M., Heikkonen, J., Topani, F., Marciani, M.G., Kaski, K., Babiloni, F., Joint Res, 2000,"Neural network for robust classification of mental tasks", Engineering in Medicine and Biology Society.

[4] Vijay Khare , Jayashree Santhosh , Sneh Anand , Manvir Bhatia, 2009,"Classification of five mental tasks from EEG Data using Neural Network based on principal component analysis", The IUP Journal of Science & Technology, Vol. 5.

[5] Charles Anderson, Elliott Forney, Douglas Hains and Annamalai Natarajan, 2011, "Reliable identification of mental tasks using time-embedded EEG and sequential evidence accumulation", Journal of Neural Engineering.

[6] José del R. Millán, Josep Mouriño, Marco Franzé, Febo Cincotti, Markus Varsta, Jukka Heikkonen, and Fabio Babiloni, "A Local Neural Classifier for the Recognition of EEG Patterns Associated to Mental Tasks".

[7] Dehariya, A.; Chaudhary, V.K.; Khan, I.; Karsoliya, S., 2011, "An effective approach for medical diagnosis preceded by artificial neural network ensemble", Electronics Computer Technology (ICECT).

[8] Djordjevic, V.; Reljin, N.; Gerla, V.; Lhotska, L.; Krajca, V., 2009,"Feature extraction and classification of EEG sleep recordings in newborns", Information Technology and Applications in Biomedicine.

[9] MR Narasingarao, R Manda, GR Sridhar, K Madhu, AA Rao, "A Clinical Decision Support System Using Multilayer Perceptron Neural Network to Assess Well Being in Diabetes".

[10] Sarah M. Hosni, Mahmoud E. Gadallah, Sayed F. Bahgat, Mohamed S. AbdelWahab, 2006, "Classification of EEG Signals Using DifferentFeature Extraction Techniques for Mental-Task BCI".

[11] Abdulhamit Subasia, M. Kemal Kiymika, Ahmet Alkana, Etem Koklukayab, 2005, "Neural Network classification of EEG signals by using AR with MLE preprocessing for Epileptic Seizure Detection", Mathematical and Computational Applications, Vol. 10.

[12] Abdulhamit Subasi, Ergun Erc¸elebi, 2004, "Classification of EEG signals using neural network and logistic regression", Computer Methods and Programs in Biomedicine .

[13] Vitaly Schetinin TheorieLabor, Friedrich-Schiller,2001, “Polynomial Neural Networks Learnt to Classify EEG Signals”.

[14] leslie Cromwell, Fred J. weibell, Erich A. Pfeirffer, 2007, “Biomedical Instrumentation and Measurements”, Pearson Prentice Hall.

[15] http://en.wikipedia.org/wiki/Backpropagation

[16] S. N. Sivanandam, Sumathi & Deepa, 2006, “Introduction to neural networks using MATLAB 6.0”,Tata McGraw-Hill Education.

[17] Prakash Chandra Gupta & Manu Pratap Singh, 2009, “Analysis of Performance Evaluation for the Classification of Protein through Solubility with Evolutionary Hybrid Algorithm in Feedforward Neural Network” International Journal of Bioinformatics.

[18] Simon Haykin, 2005, “Neural Networks”, Second edition by, Prentice Hall of India.

[19] Harry B.Burke, 1997, “Evaluating Artificial Neural Networks for Medical Applications”, New York Medical College, Deptt. Of Medicine, Valhalla, IEEE.

[20] The MathWorks Inc. PRO-MATLAB for Sun Workstations,1990, User's Guide,The MathWorks Inc.

[21] R. C. Eberhart and R. W. Dobbins, 1990, Neuml Network PC Tools, A Practical Guide. Academic Press, San Diego, California.