CHAPTER 3

IDENTIFICATION OF AUTISM USING NEURO FUZZY MODEL

3.1 NEURO FUZZY MODEL

This chapter deals with a neuro fuzzy based model which is designed for identification or diagnosis of autism. The problem areas are gathered for every individual and the related linguistic inputs are converted into fuzzy input values which are in turn given as input to feed forward multilayer neural network. The network is trained using back propagation training algorithm and tested for its performance with the expertise (Arthi et al 2008).

The proposed model will assist not only the medical practitioners but also supports psychologist, special educators and occupational therapist. ANN have found many successful applications in learning for example (Kohonen, 1982). However, ANN is not able to perform logic-like rules because, the distribution of connection weights in the network is almost impossible to be interpreted in terms of If-Then rules (Kornel Papik et al 1998).

3.2 DESIGN OF THE PROPOSED MODEL

The proposed model helps in identifying autism in children. It considers the original autistic data converted into suitable fuzzy membership values and these are fed as input to neural network architecture, where further training takes place. The basic steps involved in the proposed model for diagnosing autistic disorder are as follows:
Step 1. Data Acquisition

Step 2. Conversion of input into fuzzy membership values

Step 3. Design of network structure

Step 4. Selection of a learning algorithm

Step 5. The organization of the training and test sets

Step 6. Network training with the data sets

Step 7. Network testing and performance evaluation.

The above steps could be depicted in the Figure 3.1. Initially the linguistic input values are converted into appropriate fuzzy membership values and given to back propagation network for training along with weight and bias. The net input and output is calculated for both the hidden and output layers. Error rate is calculated and the network is trained until it reaches the stopping condition of minimum error or maximum number of epochs.

To characterize the autistic disorder the experts like clinical psychologist, occupational therapist and a special educator are asked to describe the conceptual methodology that they use in every day clinical practice to assess the severity of autism. Experts stated that they usually utilize only those six problematic areas to judge the severity of autism as specified in Appendix 1. Thus the fuzzy model consist of only those six problematic areas like pragmatic language, theory of mind, eye contact, range of interest, interest in making friends and symbolic play problems which are the key characteristics of autism.
Invoke neural network Architecture: BP Network

Initialise weight and Bias

Calculate net input and output from hidden layer neurons

Calculate net input and output from output layer neurons

Compute error with the calculated & desired output

Back propagate the Error

Update the weights

If stopping condition reached i)min error ii) no of epochs

Obtain the output

Stop

Figure 3.1 Steps to identify Autism
3.2.1 Conversion of Input Data

A questionnaire in Appendix 1 is prepared based on core problem areas of autistic disorder as listed below and Yes or No options of the MCHAT are replaced by various linguistic options as given in the questionnaire below. All the linguistic values which are obtained from the parents of the child are stored in a database. The problem areas are represented as fuzzy membership values using an intuition technique which involves contextual and semantic knowledge. As an example, consider the membership function for the fuzzy variable like pragmatic language and Figure 3.2 shows various shapes on the universe of pragmatic language which is measured in numerical units as assigned by the experts like clinical psychologist, occupational therapist or a special educator.

As per the options selected for an individual, the values of linguistic variables (very well, fairly well) are converted into appropriate numerical value using the defuzzification method of center of gravity. Each curve is a membership function corresponds to various fuzzy variables such as very well, fairly well, little and no understanding. Those curves are a function of context which are developed by experts. Here the PAI, the pragmatic language referred to the expert comfort to get a set of curves. The triangular membership function of a vector \( x \) in Equation (2.2) corresponds to all the problem areas of autism. The components of the input vector consists of membership values to the linguistic properties such as very low, low, average and high understanding which can be represented as

\[
PAI = VW \text{ (very well), FW (Fairly well), LI (Little), NU (No understanding)}.
\]

For example, the membership value of VW (very well) and LI (Little) of PAI can be written as
\[ \mu(VW) = \left\{ \begin{array}{c} 0 \\ 0.2 \\ 0 \\ 0 \end{array} \right\} \quad \left( \begin{array}{c} 0 \\ 0 \\ 0 \\ 10 \end{array} \right) \left( \begin{array}{c} 0 \\ 20 \\ 25 \\ 30 \end{array} \right) \] (3.1) 

\[ \mu(LI) = \left\{ \begin{array}{c} 0 \\ 0.2 \\ 0 \\ 0 \end{array} \right\} \quad \left( \begin{array}{c} 0 \\ 30 \\ 40 \\ 50 \end{array} \right) \left( \begin{array}{c} 0 \\ 55 \\ 60 \end{array} \right) \] (3.2)

Similarly the membership function for fuzzy variables which corresponds to other problematic areas are also represented in Figures 3.2 to 3.8.

List of major Problem Areas of Autistic Disorder for diagnosing and evaluation

2. Early verbal language delays
3. Pragmatic language.
4. Theory of mind, relatedness and empathy
5. Eye contact problems.
6. Restricted/ stereotyped range of interests.
7. Appears uninteresting in making friends
8. Pretend or symbolic play problems.
10. Fascination with written words.

A fuzzy rule base defined in the Equation (2.5) can be written for the core problematic areas of autism is shown in Equation (3.3).

\[ IF \ PA1 \ is \ NU \ AND \ PA2 \ is \ NA \ AND \ PA3 \ is \ NR \ AND \ PA4 \ is \ H \ AND \ PA5 \ is \ H \ AND \ PA6 \ is \ NI \ THEN \ AD \ is \ H \] (3.3)
where $PA_1$ to $PA_6$ represents problem areas and $AD$ represent output Autistic disorder which are represented in Figures 3.2 to 3.8.

Further the converted fuzzy membership values are given as input to back propagation algorithm for training and testing.

$$PA_1 = VW \text{ (very well)}, FW \text{ (Fairly well)}, LI \text{ (Little)}, NU \text{ (No understanding)}$$

Figure 3.2 Fuzzy Membership Values for Pragmatic Language

$$PA_2 = VL \text{ (Very little)}, LI \text{ (Little)}, NA \text{ (Not at all)}$$

Figure 3.3 Fuzzy Membership Values for Theory of Mind
PA3 = NRY (Never or rarely), OP (only with parents), UD (Usually Does)

![Membership Values for PA3](image1)

Figure 3.4 Fuzzy Membership Values for Eye Contact

A4 = VL (Very Low), L(Low), A(Average), H(High)

![Membership Values for PA4](image2)

Figure 3.5 Fuzzy Membership Values for Range of Interest

A5 = VL (Very Low), L(Low), A(Average), H(High)
Figure 3.6 Fuzzy Membership Values for Interest in making friends

$A_6 = NI$ (Not interested), $VL$ (Very Little), $QL$ (Yes, Quite a Lot)

Figure 3.7 Fuzzy Membership Values for Symbolic Play Problems

Output = $VL$ (Very Low), $L$ (Low), $A$ (Average), $H$ (High)
3.2.2 Construction of Algorithm for Neuro fuzzy model

ANN has the ability to learn how to do tasks based on the given data for training or initial experience. In ANN model, the weight is given as \( W = (w_{ij}) \) in a matrix form. The net input to output unit \( Y_j \) is given as the dot product of the input vectors \( X_i = (x_1, x_2, x_3, \ldots x_i, \ldots x_n) \) and \( W_j \) (\( j^{th} \) column of the weight vector matrix)

\[
Y_{inj} = \sum x_i w_{ij} \quad \text{where } i = 1 \text{ to } n.
\]

The nonlinear activation function, where binary sigmoidal functions are mostly used in multilayer nets which ranges between 0 to 1 and can be represented as \( F(x) = 1 / (1 + \exp(-\sigma x)) \).

The algorithm identify ( ) uses initial original values, converted fuzzy membership values, initial weights and bias calculated from Nguyen-Widrow approach for identification of Autism (Nguyen, et al 1990). The learning rate (\( \alpha \)) is used for updating the weight and biases. The training continues until the epochs reaches a maximum value of 500 or the error rate less than or equal to 0.01. In the proposed algorithm the following steps are executed.
• In the feed forward stage of algorithm, the activation function for hidden and the output layer is calculated.

• In the next step, back propagation of errors are calculated

• Weights and bias are updated.

Algorithm: identify ( )

Step 1: //To initialize the Weights and Bias // For every hidden unit, calculate \( \beta = 0.7 \frac{(noh)}{noi} \)

Set the bias \( v_{oj} \) and \( w_{ok} \) as random number between \( -\beta \) and \( +\beta \)

Initialize the value of weights \( v_{ij} \) and \( w_{jk} \) between -1 and 1

where \( noh \) – number of hidden units and \( noi \) represents number of input units.

Step 2: //Specify the stopping condition//

If \( noep \leq 500 \) or \( err > 0.01 \) then Continue with Step 3.

Step 3: //Conversion of original linguistic input into fuzzy membership values//

For each original value, Calculate

\( trimem (i) = \frac{(range \ between \ 0 \ - \ 1)}{(range \ between \ the \ orgval)} \)

where \( i \) varies between 1 to \( m \) of problem areas

Step 4: //Feed Forward Stage – Calculation of Activation Function for hidden layers//

Sum of weighted input values for hidden layer can be calculated as

\[ z_{inj} = v_{oj} + \sum_{i=1}^{n} x_i v_{ij} \quad \text{where} \quad j = 1,2,\ldots,m. \]
Step 5: // Feed Forward Stage – Calculation of Activation Function for output layer/

\[ y_{ink} = w_{ok} + \sum_{j=1}^{m} (z_j w_{jk}) \text{ where } k = 1,2, \ldots , s. \]

Step 6: // Back Propagation of Errors/

Error term is calculated as

\[ \delta_k = (t_k - y_k) f(y_{ink}) \]

\[ \delta_{inj} = \sum_{k=1}^{n} \delta_j w_{jk} \text{ and } \delta_j = \delta_{inj} f(z_{inj}) \]

Step 7: // Updation of Weights and Biases/

(i) For Each Hidden Unit:

Weight Correction term: \[ \Delta v_{ij} = \alpha \delta_j x_i \]

Bias Correction term: \[ \Delta v_{oj} = \alpha \delta_j \]

\[ v_{ij}(new) = v_{ij(old)} + \Delta v_{ij} \]

\[ v_{oj}(new) = v_{oj(old)} + \Delta v_{oj} \]

(ii) For Each Output Unit:

Weight Correction term: \[ \Delta w_{jk} = \alpha \delta_k z_j \]

Bias Correction term: \[ \Delta w_{ok} = \alpha \delta_k \]

\[ w_{jk}(new) = w_{jk(old)} + \Delta w_{jk} \]

\[ w_{ok}(new) = w_{ok(old)} + \Delta w_{ok} \]

Step 8: // Test for Stopping Condition/

Test the values of noep and err to stop the process.
3.2.3 Creation of the proposed model

Feed Forward net is advantageous over single layer net in the sense that, it can be used to solve more complicated problems. Back Propagation provides a computationally efficient method for changing weights in a feed forward network, with differentiable activation function units, to learn a training set of input-output examples. Here the network is trained by back propagation algorithm, for each iteration the error value is calculated. The training is stopped when the error begins to rise, after 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 and the output unit or it may be calculated using Nguyen-Widrow initialization. The training vector starts with a $n \times n$ matrix of which it contains values between 0 to 1 and ends with the error rate of 0.01. 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. After creating a diagnostic model as in Figure 3.9, the performance can be evaluated with a set of inputs by a medical expert.

![Figure 3.9 Block diagram representing a model to diagnose Autism](image-url)
3.3 EXPERIMENTAL RESULTS AND DISCUSSION

In this model, the input data for the six problem areas which are considered as core area for predicting autism are stored in the database and then converted into fuzzy membership values. The output will be the Autistic disorder. For this model, 40 samples are collected for both training and testing. In each training epochs, the connection weights are adjusted to minimize the total mean square error in the output. Parameters like learning rate, number of hidden layers and bias improves the performance of the network. The Table 3.1 and Table 3.2 gives the parameters like hidden layers, nodes in hidden layers, epochs and learning rate to train and to test this neuro-fuzzy model. In this neuro-fuzzy model, network learns quickly and gives the output error rate of 0.01 and remains constant after 400 epochs as shown in Figure 3.10. The overall performance of this model is 85-90%.

<table>
<thead>
<tr>
<th>S.No</th>
<th>Nodes in the hidden layer</th>
<th>Epochs</th>
<th>Learning Rate</th>
<th>Total Sum of Squared Error</th>
<th>Average Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-25</td>
<td>30</td>
<td>0.1</td>
<td>4.223</td>
<td>0.39</td>
</tr>
<tr>
<td>2</td>
<td>3-75</td>
<td>50</td>
<td>0.3</td>
<td>3.468</td>
<td>0.317</td>
</tr>
<tr>
<td>3</td>
<td>2-50</td>
<td>100</td>
<td>0.2</td>
<td>4.344</td>
<td>0.124</td>
</tr>
<tr>
<td>4</td>
<td>1-25</td>
<td>150</td>
<td>0.1</td>
<td>4.890</td>
<td>0.0101</td>
</tr>
<tr>
<td>5</td>
<td>1-25</td>
<td>200</td>
<td>0.1</td>
<td>4.998</td>
<td>0.003</td>
</tr>
<tr>
<td>6</td>
<td>1-25</td>
<td>300</td>
<td>0.1</td>
<td>4.589</td>
<td>-0.004</td>
</tr>
<tr>
<td>7</td>
<td>2-50</td>
<td>400</td>
<td>0.2</td>
<td>5.511</td>
<td>-0.006</td>
</tr>
<tr>
<td>8</td>
<td>2-50</td>
<td>500</td>
<td>0.2</td>
<td>5.540</td>
<td>-0.004</td>
</tr>
<tr>
<td>9</td>
<td>2-50</td>
<td>600</td>
<td>0.2</td>
<td>5.554</td>
<td>-0.0033</td>
</tr>
<tr>
<td>10</td>
<td>3-75</td>
<td>600</td>
<td>0.2</td>
<td>5.913</td>
<td>0.008</td>
</tr>
</tbody>
</table>
Table 3.2 Test data scores (Test Data 50 \%)

<table>
<thead>
<tr>
<th>S.No</th>
<th>Nodes in the hidden layer</th>
<th>Epochs</th>
<th>Learning Rate</th>
<th>Total Sum of Squared Error</th>
<th>Average Error Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-25</td>
<td>30</td>
<td>0.1</td>
<td>5.065</td>
<td>0.192</td>
</tr>
<tr>
<td>2</td>
<td>3-75</td>
<td>50</td>
<td>0.3</td>
<td>4.85</td>
<td>0.078</td>
</tr>
<tr>
<td>3</td>
<td>2-50</td>
<td>100</td>
<td>0.2</td>
<td>3.283</td>
<td>0.58</td>
</tr>
<tr>
<td>4</td>
<td>1-25</td>
<td>150</td>
<td>0.1</td>
<td>2.855</td>
<td>-0.052</td>
</tr>
<tr>
<td>5</td>
<td>1-25</td>
<td>200</td>
<td>0.1</td>
<td>2.409</td>
<td>-0.054</td>
</tr>
<tr>
<td>6</td>
<td>1-25</td>
<td>300</td>
<td>0.1</td>
<td>0.961</td>
<td>-0.036</td>
</tr>
<tr>
<td>7</td>
<td>2-50</td>
<td>400</td>
<td>0.2</td>
<td>0.019</td>
<td>-0.002</td>
</tr>
<tr>
<td>8</td>
<td>2-50</td>
<td>500</td>
<td>0.2</td>
<td>0.012</td>
<td>-0.001</td>
</tr>
<tr>
<td>9</td>
<td>2-50</td>
<td>600</td>
<td>0.2</td>
<td>0.008</td>
<td>-0.001</td>
</tr>
<tr>
<td>10</td>
<td>3-75</td>
<td>600</td>
<td>0.2</td>
<td>0.012</td>
<td>-0.001</td>
</tr>
</tbody>
</table>

Figure 3.10 Learning curve for training data
3.4 SUMMARY

The reliability of the proposed neuro fuzzy model depends on data collected from the patients. This model provides a way for better classification or identification. Also it provides supportive tools for the medical experts, special educators and psychologist. It can be used to generate models for different medical applications in diagnosing autism.