CHAPTER 5

EEG SIGNAL CLASSIFICATION BASED ON NN WITH ICA AND STFT

5.1 OVERVIEW

A novel approach is proposed for Electroencephalogram signal classification using Artificial Neural Network based on Independent Component Analysis and Short Time Fourier Transform. The source EEG signals contain the electrical activity of the brain produced in the background by the cerebral cortex nerve cells. EEG is one of the most utilized methods for effective analysis of the brain functions. The accuracy of the EEG signal classification technique from the previous works is enhanced. FastICA is used as a preprocessing step, while STFT is used for adequate denoising of the EEG signal.

Feature extraction is performed based on three parameters, namely, Correlation Dimension, Lyapunov Exponent and Standard Deviation. The Artificial Neural Network is trained by integrating Levenberg-Marquardt training algorithm with Back Propagation Electrocardiogram algorithm. This results in higher classification accuracies for the three classification problems. This method is aimed at enhancing the clinical services of EEG recording and the decision making in epileptic seizure detection. This EEG signal classification method performs better than the EEG based on Adaptive Neuro Fuzzy Inference System classifier in terms of Accuracy, Specificity and Sensitivity.
5.2 OVERVIEW OF EEG SIGNAL CLASSIFICATION

EEG is an important technique for the recognition and classification of brain diseases. It is a non-invasive procedure for searching the human brain dynamics. Accurate detection of the neuron voltage can be achieved by using electrodes with low impedance. The EEG with epileptic seizures also includes an abnormal electrical activity which must be eliminated. The presence of epileptic seizures cannot be detected with EEG recording of short periods, hence the EEG must be continuously recorded and the whole EEG has to be analyzed to diagnose the epileptic seizures.

ICA is used for data transformation of the EEG signals. ICA algorithm is implemented as a preprocessing step to accelerate the epileptic seizure detection operation. The independent components that are overlapped under various conditions can be effectively differentiated using ICA. The EEG data are processed and the features corresponding to the various activities are determined for further analysis.

STFT is applied to regular time-frequency analysis of EEG signals. STFT maps a signal into a 2-dimensional function depending on time and frequency. The features have been extracted for the correct classification of the events in the EEG signal. The feature extraction is based on three parameters, namely, Standard Deviation, Correlation Dimension and Lyapunov Exponent.

LM BP training algorithm is used to address the complexity and the issue of non-linearity. This algorithm combines the best features of Gauss-Newton method (Wang 2012) for determining the minimum of a function and Steepest-descent algorithm (Gonzaga & Karas 2013) for approximation along the direction of the stationary phase. The results of the
classification are analyzed in terms of Accuracy, Specificity and Sensitivity with respect to the results obtained using ANFIS classifier.

5.3 NEED FOR PREPROCESSING AND DENOISING

The EEG data needs to be preprocessed to fasten the EEG signal classification process owing to the huge size of the EEG dataset. The necessity for data preprocessing relies on the noisy and inconsistent EEG data. Denoising is required to remove the unwanted noisy signals and artifacts.

5.4 EEG SIGNAL CLASSIFICATION METHODOLOGY

The flow of EEG signal classification based on ANN with ICA and STFT is shown in Figure 5.1. The EEG signal is a non-stable signal, therefore, adequate analysis is required to distinguish the epileptic seizures from the normal EEG signal. The EEG signal classification is performed based on two classifiers namely FeedForward BP Neural Network which is trained using LM algorithm and the ANFIS classifier.

The input EEG signal is subject to FastICA, followed by denoising and feature extraction. LM BP Neural Network training algorithm trains the Neural Network to classify the EEG signals as normal, seizure-free and seizure.

5.4.1 Independent Component Analysis (ICA)

ICA is a statistical and computational method for the estimation of latent variables occurring during data generation. Some of its applications are face recognition, optical imaging of neurons, facial recognition, and modeling receptive areas of main visual neurons (Hyvarinen et al 2010). It employs the Non-Gaussian statistical independence between the hidden variables to
disintegrate the multivariate observation data into a linear sum of statistically independent sections.

Figure 5.1 EEG Signal Classification Based on NN with ICA and STFT

Let $s(t)$ represent a source vector of dimensions $d \times n$, where $d$ is the dimension of data and $n$ represents the number of independent realizations of a random variable. ICA applies a linear model where the data vector $x(t)$ of dimensions $d \times n$, is generated according to Equation (5.1).
In Equation (5.1), \( \mathbf{A} \) is an unknown, non-singular \( d \times d \) mixing matrix. The linear ICA model computes a linear transformation based on statistically independent coefficients. The columns of the mixing matrix correspond to the basis vectors, whereas the entities of the source vector represent the basis coefficients. Both the basis coefficients and basis vectors are computed from the data vector. ICA is capable of recovering the unknown sources \( \{s_i(t)\} \) only when the data vector comprises of a linear mixture of the source vectors, under a finite number of data observations \( \{x(t)\} \).

### 5.4.2 EEG Signal Denoising

The denoising process consists of three sub-processes, namely, STFT computation, thresholding, and inverse STFT calculation. STFT is used for the denoising of time independent signals, where the STFT of a signal involves the Fourier Transform (FT) of crossing windowed signal blocks. FT is a method for input signal transformation of time-domain to frequency-domain. STFT consists of rectangular window for enhanced signal denoising. The spectrogram values below a particular range are set to zero and the process is known as thresholding. This enables the complete spectrogram reconstruction. The inverse STFT is calculated to obtain the denoised signal. The STFT is found by taking the Discrete-Time Fourier Transform (DTFT) over each windowed block. The STFT signal is given in Equation (5.2) and it is dependent on angular frequency \( \omega \) and delay parameter \( \tau \).

\[
F(\tau, \omega) = \text{STFT} \{s(n)\} \tag{5.2}
\]

The threshold STFT function \( F_d(\tau, \omega) \) is given in Equation (5.3), that is thresholding the STFT function according to Equation (5.4).
The threshold value is defined by the variable \('Threshold’\ as shown in Equation (5.5).

\[ F_d(\tau, \omega) = THR(F(\tau, \omega)) \]  
(5.3)

\[ THR(a) = \begin{cases} 
0, & |a| \leq \text{Threshold} \\
a, & |a| > \text{Threshold} 
\end{cases} \]  
(5.4)

\[ \text{Threshold} = \frac{\max(\text{orig}) - \sigma_{\text{orig}}}{|\min(\text{orig})|} \]  
(5.5)

In Equation (5.4), \(THR\ (a)\) represents the threshold function, \(a\) represents the data values, \(‘Threshold’\ represents the threshold value, and \(‘orig’\ represents the original signal. The threshold value is computed to be 0.6. The threshold function is equivalent to zero when the absolute data value is lesser than or equal to the \(‘Threshold’\ value, and it is equivalent to the data value when the absolute data value is greater than the \(‘Threshold’\ value. This type of thresholding is known as hard thresholding because the values greater than \(‘Threshold’\ value are not affected. The threshold STFT function is used in the computation of the inverse STFT function for obtaining the denoised signal as shown in Equation (5.6).

\[ x(n) = \text{STFT}^{-1}(F_d(\tau, \omega)) \]  
(5.6)

Figure 5.2 shows the block diagram of denoising based on STFT. The consequent signal is the reconstructed and denoised signal that would be applied in feature extraction.

\[ \text{Figure 5.2 Block Diagram for Denoising using STFT} \]
5.4.3 Feature Extraction

Feature extraction is a form of dimensionality reduction and a process of obtaining the relevant features from a group of signals that tends to the corresponding classification and diagnosis (Verma et al. 2010). The need for feature extraction lies on the large size of the input data to be processed. The redundant input data are transformed into a decreased set of features. The main constraint in the feature extraction process is the higher number of variables involved which may overfits the training sample.

Standard Deviation (SD) is used to represent the limit to which the peaks and troughs of a wave varies with an average from the mean denoised EEG voltage. When the SD is low, it denotes that the data values are close to the mean denoised EEG voltage and when the SD is high, it characterizes that the data values are dispersed over a large range. Few significant features can be extracted using SD after signal decomposition. The mean of the denoised EEG signal is computed for analyzing its statistical feature as shown in Equation (5.7).

\[
\mu = \frac{1}{N} \sum_{i=0}^{N-1} x_i
\]  

(5.7)

In Equation (5.7), \( \mu \) represents the mean of denoised EEG signal and \( x_i \) represents the denoised EEG signal component. The SD is calculated from the mean of the denoised EEG signal as shown in Equation (5.8) and \( N \) is the number of data points.

\[
\sigma = \sqrt{\frac{1}{N-1} \sum_{i=0}^{N-1} (x_i - \mu)^2}
\]  

(5.8)

The mean signal characterizes the measured condition, while the SD characterizes the noise and other interferences. The quality of the denoised EEG signal can be estimated using two parameters, namely, Signal-to-Noise Ratio (SNR_{dB}) and coefficient of variation \( (c_v) \), as shown in
Equations (5.9) and (5.10). Better quality signals possess lower $c_v$ and higher \( \text{SNR}_{\text{dB}} \).

\[
\text{SNR}_{\text{dB}} = 20 \log_{10} \left( \frac{\mu}{\sigma} \right) \tag{5.9}
\]

\[
c_v = (\sigma / \mu) \times 100 \tag{5.10}
\]

Correlation dimension is a quantity of space dimensionality applied by a group of random points. The number of independent variables is estimated by the correlation dimension required to describe the dynamics of the system. Correlation dimension characterizes the distribution of points in the phase dimension (Zurek et al 2012). Let $x_1, x_2, \ldots x_m$ be the $N$ points in the $m$-dimensional space, where $N = 1, 2, \ldots m$ and let $i, j$ be two points between which the absolute distance $|x_i - x_j|$ is calculated, where $i < j$. The estimate of the correlation integral for a positive number $\varepsilon$ is computed as given in Equation (5.11).

\[
\hat{C}(\varepsilon) = \frac{2}{N(N-1)} \sum_{i<j} \theta(\varepsilon - |x_i - x_j|) \tag{5.11}
\]

In Equation (5.11), the pair of points $(i, j)$ must be lesser than a predefined distance $\varepsilon$ and $\theta(.)$ represents the Heaviside step function, an unbiased estimation of the correlation integral. The correlation integral is computed as given in Equation (5.12).

\[
C(\varepsilon) = \iint \mu(x)\mu(y)\theta(\varepsilon - |x - y|) \, dx \, dy \tag{5.12}
\]

The correlation dimension $v$, of the denoised EEG signal is computed based on Grassberger-Proccacia algorithm (Kivimaki et al 2010). This parameter detects the deterministic behavior of signals. The correlation dimension of a mean signal is given by Equation (5.13).

\[
v = \lim_{\varepsilon \to 0} \frac{\log C(\varepsilon)}{\log \varepsilon} \tag{5.13}
\]
The Lyapunov Exponent of a dynamic system is a parameter that quantifies the separation rate of infinite proximate curves (Derya Ubeyli 2010). It differentiates the various orbits corresponding to the sensitive dependence on the initial constraints and determines the total predictability of a system. Conventionally, the Lyapunov exponents are obtained from the observed time series or motion of the dynamic system. In the present research, the Lyapunov exponents are computed based on the motion of the dynamic system to enable the calculation of the largest Lyapunov components.

Let \((x, y)\) be two nearest neighboring points in the phase dimension at a time interval \((0, n)\). The distances of the points in the \(i^{th}\) direction and \((x, y)\) are given as \(||\delta x_i(0)||\) and \(||\delta y_i(n)||\) respectively. The Lyapunov exponent is defined by average growth rate \(\lambda_i\) according to Equation (5.14).

\[
\frac{||\delta y_i(n)||}{||\delta x_i(0)||} = 2^{\lambda_i n} \quad (n \to \infty)
\]  

(5.14)

The Lyapunov exponent represents the average growth rate of the initial distance. The final states of the dynamic system are different from each other, even when the initial states are approximate to each other. The vector of all the Lyapunov exponents is known as Lyapunov spectra.

### 5.4.4 EEG Signal Classification

The extracted features are given into a FeedForward ANN comprising of \(N\) inputs, a hidden layer, and \(K\) outputs, where \(N\) denotes the size of the feature vector and \(K\) denotes the number of classes. ANN is a classifier constituting numerous simple and interconnected neurons each executing a computational function. LM algorithm is integrated with the BP algorithm to train the FeedForward ANN.
The ANN includes an input layer with the input variables to the problem and an output layer involving the solution to the problem. In the current research, the number of neurons in the hidden layer is chosen as 20, while the activation function in the hidden layer is a hyperbolic tangent sigmoid transfer function and that of the output layer of the ANN is a linear function. LM algorithm minimizes a non-linear function and the sum of squares based on the maximum neighborhood (Wilamowski & Hao 2010). It finds its applications in solving non-linear least squares by curve fitting. The advantages of LM algorithm are resistance to slow convergence problem and better cost function than the other training algorithms.

Let \( w \) denote the weight vector, the non-zero entities of the weight space and \( E(w) \) denote the error function to be minimized. The computation of the error function including the error terms \( e_i^2(w) \) for \( n \) individual errors is given in Equation (5.15).

\[
E(w) = \sum_{i=1}^{n} e_i^2(w) \tag{5.15}
\]

In Equation (5.15), the error term \( e_i^2(w) = (x_{di} - x_i)^2 \), where \( x_{di} \) denotes the desired value of the \( i \)th neuron and \( x_i \) denotes the actual output of the \( i \)th neuron. A new weight vector \( w_{(k+1)} \) is computed from the previous vectors using the LM algorithm as shown in Equation (5.16), where \( \delta W_k \) is equivalent to Equation (5.17).

\[
w_{(k+1)} = w_k + \delta W_k \tag{5.16}
\]

\[
\delta W_k = -(J_k^T f(w_k)) (J_k^T J_k + \lambda I)^{-1} \tag{5.17}
\]

In Equation (5.17), \( J_k \) denotes the Jacobian of the function \( f \) at \( w_k \), \( \lambda \) represents the Marquardt parameter, and \( I \) denotes the identity matrix. The outline of the LM training algorithm (Wilamowski & Hao 2010) is given as follows,
i. Computation of error functions according to Equation (5.15).

ii. The minimum value to $\lambda$ to be specified.

iii. Calculation of the weight vectors according to Equations (5.16) and (5.17).

iv. If $E(w_{(k+1)}) \geq E(w_k)$, increase the value of $\lambda$ and go to step (iii), else decrease the value of $\lambda$.

v. Update the weight vector according to $w_k$: $w_k \leftarrow w_{(k+1)}$ and go to step (iii).

The trained Neural Network classifies the test EEG signal into three divisions, namely, normal EEG, seizure-free EEG and seizure EEG.

5.4.5 Existing EEG Signal Classifier – ANFIS

The ANFIS based classifier is a FeedForward network and universal estimator to detect the presence of epileptic seizures. It combines the principles of Neural Networks and Fuzzy Logic and can approximate the non-linear functions (Abraham 2005). It aggregates the reasoning pattern of the fuzzy systems with the learning pattern of the Neural Networks. The ANFIS model correlates the inputs and outputs via their respective Membership Functions (MFs) of fuzzy sets and its associated parameters. This classifier is used for the adaptive construction of rule base and enables the extraction of the fuzzy rules from the numerical dataset. It learns the features in the dataset and adapts the system variables according to the error condition. ANFIS classifier is employed in biomedical applications for the modeling, classification, and controlling of real-time systems. It uses triangular MFs and the linguistic model of fuzzy if-then rules is based on a Sugeno type Fuzzy Inference System (FIS) (Abraham 2005). The MF parameters of the FIS are trained based on a combination of Least-squares and BP gradient descent methods.
Five layers are utilized to develop the Sugeno type FIS and the output signals from the previous layer nodes are assigned as the input signals in the current layer (Abraham 2005). The initial layer nodes produce the fuzzy membership grades of the corresponding fuzzy sets using MFs. The nodes in the second layer multiply the incoming signals and send the product signals. The firing strength of a nodal rule is determined by the output of that node. The third layer is a mid-level transmission layer and the nodes in the fourth layer are adaptive nodes. The output of a node in the fourth layer is equal to the product of the normalized firing strength and a first order polynomial. The final fifth layer comprises of a constant node which computes the overall output by combining all the incoming signals.

5.5 PERFORMANCE METRICS

The classifier performance is measured by three parameters, namely, Accuracy, Sensitivity and Specificity. These parameters are given according to the Equations (5.18), (5.19), and (5.20).

\[
\text{Accuracy} = \frac{(TP + TN)}{(TP + FN + FP + TN)} \quad (5.18)
\]

\[
\text{Sensitivity} = \frac{TP}{(TP + FN)} \quad (5.19)
\]

\[
\text{Specificity} = \frac{TN}{(TN + FP)} \quad (5.20)
\]

In Equations (5.18), (5.19), and (5.20), True Positive (TP) is the quantity of correctly classified epilepsy cases, False Positive (FP) is the quantity of incorrectly classified epilepsy cases, False Negative (FN) is the quantity of incorrectly classified healthy patients, and True Negative (TN) is the quantity of correctly classified healthy patients. Accuracy is the ratio of correct classification of epilepsy cases and healthy patients. Sensitivity (or) Recall (or) True Positive Rate (TPR) is the rate of the actual positive classes which are correctly identified. Specificity (or) True Negative Rate (TNR) is the rate of actual negative classes which are identified correctly.
Accuracy is the most important parameter during any classification process as it points out the correct outcomes of the classifier corresponding to the target. Higher sensitivity is used to denote the lower amount of incorrectly identified true cases, whereas higher specificity is used to denote the lower amount of incorrectly identified false cases (Schache et al 2011). A classifier should be both highly specific and highly sensitive so that the total incorrect classifications are avoided. Accuracy focuses on the testing of correct classifications, while sensitivity and specificity focuses on testing of incorrect classifications.

5.5.1 Dataset Information

The present research uses the EEG dataset organized by the professors in the Department of Epileptology, University of Bonn. Single channel EEGs were observed for individuals possessing various brain electric potentials at a sampling frequency of 173.61 Hz for 23.6 seconds. The acquired EEG data from the persons are ordered into the following three divisions,

(i) Healthy persons.

(ii) Persons with epileptic seizure during seizure-free interval (interictal).

(iii) Persons with epileptic seizure during seizure interval (ictal).

The EEG data are divided into five sets (Z, O, N, S, and F) each containing 100 single-channel EEG segments. The sets Z and O are obtained from the healthy individuals under opened and closed eyes corresponding to the external surface electrodes respectively. The set F has been acquired from the epileptogenic divisions of the brain with focal intellectual activity, while the set N has been taken from the hippocampal pattern of the brain that shows
non-focal interictal activity. The set S has been yielded from an epileptic individual during seizure interval. Each EEG segment contains 4096 samples. The EEG segments are observed for the subjects with the 128-channel amplifier that involves 12 Analog-to-Digital Convertors at a bit rate of 12 Kbit/s and a sampling rate of 173.61 Hz.

5.5.2 Preprocessing and Denoising

The EEG signal classification has been tested by EEG dataset in three patterns, namely, (Z, S), (Z, N, S), and (Z, O, N, S, F). The EEG signals are decomposed and preprocessed based on FastICA. The EEG signals and its consequence due to ICA is shown in the Figure 5.3.

![Figure 5.3 Original EEG Signal and EEG Signal after ICA](image)

The EEG signals are denoised by STFT. The spectrogram of the EEG signal after thresholding based on STFT is shown in Figure 5.4. The SNR values for the ZONSF dataset is plotted in the Figure 5.5. The classifier is trained with
LM training function for the required classification of EEG signals under normal, seizure-free and seizure classes. The EEG datasets are analyzed using the FeedForward BPNN in terms of Performance characteristics, Regression analysis, and Receiver Operating Characteristic (ROC).

Figure 5.4 Spectrogram of the Signal using STFT after Thresholding

Figure 5.5 SNR Values for the ZONSF Dataset
5.5.3 Performance Evaluation Parameters

The performance characteristics are analyzed by observing the Least Mean Squared Error (MSE) of the classifier respect to reference epochs. It is necessary to measure the difference between the classifications performed by the method and the actual present entities (Imbens & Kalyanaraman 2012). MSE shows the amount by which the classified quantity varies from the actual quantity. The mathematical expression of MSE is given in Equation (5.21). In Equation (5.21), \( \hat{Y} \) denotes the output vector of \( n \) predictions and \( Y \) denotes the vector of the actual values.

\[
\text{MSE} = \frac{1}{n} \sum_{x=1}^{n} (\hat{Y}_x - Y_x)^2 \tag{5.21}
\]

Regression analysis is a process of estimating the relationships among the variables. This part of statistical analysis is important to model the relationships between the variables and necessary to determine which parameters are important indicators (Dizikes 2010). ROC denotes the variation of sensitivity with respect to specificity. It is important to raise the sensitivity level to the maximum as soon as the specificity level increases. The ROC analysis is essential for decision making and for discriminating the good and poor classification results.

5.5.4 Analysis of Dataset – (Z, S)

The analysis has been performed with 200 EEG segments (Z, S) and the results in signal discrimination are shown in Figure 5.6. It shows the performance of the Neural Network classifier corresponding to the training, validation and test performances. The individual performances are analyzed and the best characteristic is estimated. The number of epochs needed for attaining the best performance validation is estimated, that is the Mean
Squared Error rate with the least value. The best validation performance obtained here is MSE equal to $5.9456 \times 10^{-8}$ at 21$^{st}$ epoch.

![Graph showing best validation performance](image)

**Figure 5.6 Best Validation Performance – Dataset (Z, S)**

The training data are repeatedly applied to calculate the weights of the candidate solutions. The validation data are employed iteratively to compute the performance error of the non-training entities and stop the training when the non-training validation error estimate ceases to decrease and avoids overfitting. The percentage of data used for training is 70%, test is 15% and validation is 15%. The training data are applied to train the classification method, the validation data are utilized to choose the data values based on the performances of the training data, and the test data are employed. These segmentation of data gives the overall performance characteristics.

The regression analysis is performed in the Neural Network based EEG signal classification for the dataset (Z, S) and is shown in Figure 5.7. It shows a unit R value, which exhibits an exact linear relationship between the outputs and targets.
Figure 5.7 Regression Graph – Dataset (Z, S)

Figure 5.8 Best Validation Performance – Dataset (Z, N, S)
5.5.5 Analysis of Datasets – (Z, N, S)

The analysis has been performed with 300 EEG segments (Z, N, S) and the results in signal discrimination are shown in Figure 5.8. It shows the best validation performance for the dataset (Z, N, S). The best validation performance obtained here is equal to 0.12593 at the 17th epoch. The regression analysis is performed in the Neural Network based EEG signal classification for the dataset (Z, N, S) and is shown in Figure 5.9. The obtained regression value is \( R = 0.998 \) which denotes that the output varies merely from the target.

![Regression Graph – Dataset (Z, N, S)](image)

**Figure 5.9 Regression Graph – Dataset (Z, N, S)**

5.5.6 Analysis of Dataset – (Z, O, N, S, F)

The analysis has been performed with 500 EEG segments (Z, O, N, S, F) and the results in signal discrimination are shown in the Figure 5.10. It indicates the best validation performance for the dataset (Z, O, N, S, F). It is noted that the best validation performance here is equal to 0.014512 at the
25\textsuperscript{th} epoch. The regression analysis is performed in the Neural Network based EEG signal classification for the dataset (Z, O, N, S, F) and is shown in Figure 5.11. The obtained Regression R-value is 0.95802 which denotes that the output varies more from the target than in the other combinations of datasets.

Figure 5.10 Best Validation Performance – Dataset (Z, O, N, S, F)

Figure 5.11 Regression Graph – Dataset (Z, O, N, S, F)
The ROC is examined as a comparison of true positive rate vs. the false positive rate for the dataset (Z, O, N, S, F) and is shown in the Figure 5.12. The aim of an ideal ROC plot is to achieve a unit TP rate for any FP rate. Hence, the ROC curve needs to occupy the top leftmost corner of the ROC plot. When a point on the ROC curve is closer to the ideal coordinate, the accuracy of the test would be higher. However, when the points are closer to the diagonal, the accuracy of the test would be closer. The Area Under ROC curve (AUC) is a metric of the accuracy of a diagnostic test. Higher the AUC, higher is the accuracy of the classification. The AUC for the test curve was obtained as 0.98, followed by the training curve with an AUC of 0.96, and the validation curve has an AUC of 0.91.

![Figure 5.12 ROC plot – Dataset (Z, O, N, S, F)](image)

5.5.7 Confusion Matrix Analysis

The EEG signal classification using the FeedForward BPNN is compared with ANFIS classifier (Abraham 2005) for the three classification problem of EEG datasets, namely, (Z, S), (Z, N, S), and (Z, O, N, S, F).
The confusion matrices for the classification of the datasets (Z, S), (Z, N, S), and (Z, O, N, S, F) using ANFIS classifier and FeedForward BPNN are given in the Table 5.1, Table 5.2 and Table 5.3.

**Table 5.1 Confusion Matrix for (Z, S) Dataset**

<table>
<thead>
<tr>
<th>Class</th>
<th>Z</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Z</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table 5.2 Confusion Matrix for (Z, N, S) Dataset**

<table>
<thead>
<tr>
<th>Class</th>
<th>Z</th>
<th>N</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>N</td>
<td>0</td>
<td>99</td>
<td>1</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>1</td>
<td>99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Z</th>
<th>N</th>
<th>S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>N</td>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table 5.3 Confusion Matrix for (Z, O, N, S, F) Dataset**

<table>
<thead>
<tr>
<th>Class</th>
<th>ANFIS Classifier</th>
<th>Feed Forward BPNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z</td>
<td>O</td>
<td>N</td>
</tr>
<tr>
<td>Z</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>O</td>
<td>0</td>
<td>97</td>
</tr>
<tr>
<td>N</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>F</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The values of Accuracy, Sensitivity, and Specificity are calculated from the Equations (5.18), (5.19), and (5.20) for the three sets of EEG datasets, namely, (Z, S), (Z, N, S), and (Z, O, N, S, F) using ANFIS classifier and FeedForward BPNN. The computed performance measures are enlisted in Table 5.4.
Table 5.4 Performance Measures Computed from the Confusion Tables

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ANFIS Classifier</th>
<th>Feed Forward BPNN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy (%)</td>
<td>100</td>
<td>99.33</td>
</tr>
<tr>
<td>Sensitivity (%)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The results shown in the Table 5.4 reveals the potential of the EEG signal classification using Feed Forward BPNN in comparison to that of ANFIS classifier. A comparison of the EEG signal classification made in the present research work is compared with the other existing EEG signal classification techniques that have utilized the same EEG test. Table 5.5 shows the comparison of classification accuracy with other techniques using the same dataset.

Table 5.5 Comparison of the Classification Accuracy

<table>
<thead>
<tr>
<th>Reference</th>
<th>Technique</th>
<th>Dataset</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kannathal et al (2005a)</td>
<td>Chaotic measures-Surrogate Data analysis</td>
<td>Z, S</td>
<td>~90.00</td>
</tr>
<tr>
<td>Polat &amp; Güneş (2007)</td>
<td>FFT-Decision tree</td>
<td>Z, S</td>
<td>98.72</td>
</tr>
<tr>
<td>Subasi (2007)</td>
<td>DWT Mixture</td>
<td>Z, S</td>
<td>95.00</td>
</tr>
<tr>
<td>Tzallas et al (2009)</td>
<td>Time Frequency analysis – ANN</td>
<td>Z, S</td>
<td>100.00</td>
</tr>
</tbody>
</table>
Table 5.5 (Continued)

<table>
<thead>
<tr>
<th>Reference</th>
<th>Technique</th>
<th>Dataset</th>
<th>Accuracy(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercy (2012)</td>
<td>DWT and ICA with NN</td>
<td>Z, S</td>
<td>98.00</td>
</tr>
<tr>
<td>Hosseini et al (2013)</td>
<td>Qualitative and Quantitative Evaluation</td>
<td>Z, S</td>
<td>96.90</td>
</tr>
<tr>
<td>Proposed Work</td>
<td>ICA, STFT and FeedForward BPNN</td>
<td>Z, S</td>
<td><strong>100.00</strong></td>
</tr>
<tr>
<td>Tzallas et al (2009)</td>
<td>Time Frequency analysis – ANN</td>
<td>Z, N, S</td>
<td>100.00</td>
</tr>
<tr>
<td>Proposed Work</td>
<td>ICA, STFT and FeedForward BPNN</td>
<td>Z, N, S</td>
<td><strong>100.00</strong></td>
</tr>
<tr>
<td>Guler &amp; Ubeyli (2005)</td>
<td>WT-ANFIS</td>
<td>Z, O, N, S, F</td>
<td>98.68</td>
</tr>
<tr>
<td>Proposed Work</td>
<td>ICA, STFT and FeedForward BPNN</td>
<td>Z, O, N, S, F</td>
<td>96.20</td>
</tr>
</tbody>
</table>
5.6 SUMMARY

This chapter classifies the EEG signals based on Artificial Neural Network with the consequence of ICA and STFT. ICA is applied in the preprocessing stage and STFT is used for signal denoising. The feature extraction process has been performed based on three quantities, namely, Standard Deviation, Correlation Dimension and Lyapunov Exponent. The relevant features in the EEG signals are excerpted and given to the trained NN for signal classification. The Neural Network is trained with a LM algorithm to yield the performance results within optimal epochs. This EEG signal classification based on FeedForward BPNN performs better than that of ANFIS classifier in terms of Accuracy, Sensitivity, and Specificity.

The EEG datasets were taken from the contributions of the professors in the Department of Epileptology, University of Bonn. Three sets of datasets, namely, (Z, S), (Z, N, S), and (Z, O, N, S, F), were taken to validate the performance of the EEG signal classification. This method resulted in a perfect classification of 100% for (Z, S) and (Z, N, S) datasets. The EEG classification accuracy has been found to be only 96.2% for the dataset (Z, O, N, S, F) and so the average EEG classification accuracy is computed to be 98.73%. The results inferred that both FeedForward BPNN and ANFIS classifiers resulted to acceptable classification accuracy, even though the influence of FeedForward BPNN was little more than that of the ANFIS classifier. This EEG signal classification method is a reliable computerized technique for epileptic seizure diagnosis in clinical practice. This epileptic seizure detection method needs to be further enhanced in terms of classification accuracy based on optimized feature selection.