Comprehensive Analysis of Hierarchical Aggregation Functions Decision Trees, SVD, K-means Clustering, PCA and Rule Based AI Optimization in the Classification of Fuzzy based Epilepsy Risk Levels from EEG Signals

R.Harikumar¹, T.Vijayakumar²

¹ Bannari Amman Institute of Technology, Sathyamangalam, India.
e-mail: harikumarrajaguru@gmail.com
² Bannari Amman Institute of Technology, Sathyamangalam, India.
e-mail: vishal_16278@yahoo.co.in

Abstract—A comprehensive analysis for the performance of post classifiers such as Hierarchical Soft Decision Trees, Singular value decomposition(SVD), k-means clustering, Principal Component Analysis (PCA) and Rule based AI techniques in optimization of fuzzy outputs for the classification of epilepsy risk levels from EEG (Electroencephalogram) signals is presented in this paper. The fuzzy pre classifier is used to classify the risk levels of epilepsy based on extracted parameters like energy, variance, peaks, sharp and spike waves, duration, events and covariance from the EEG signals of the patient. Hierarchical Soft decision tree (post classifiers with max-min criteria) four types, SVD, K-means clustering, PCA and AI optimization are applied on the classified data to identify the optimized risk level (singleton) which characterizes the patient’s risk level. The efficacy of the above methods is compared and analyzed based on the bench mark parameters such as Performance Index (PI), and Quality Value (QV).

Keywords: EEG Signals, Epilepsy Risk Levels, Fuzzy Logic, Hierarchical Decision Trees, SVD, K-means clustering, PCA, AI Techniques

I. Introduction

Epileptic seizures are a principal brain dysfunction with important public health implications, as they affect 0.8% of humans [1]. Electroencephalograms (EEGs) are recordings of electrical potentials produced by the brain. Analysis of EEG activity has been achieved principally in clinical settings to identify pathologies and epilepsies since Hans Berger’s recording of rhythmic electrical activity from the human scalp [2]. In the past, interpretation of the EEG was limited to visual inspection by neurophysiologist, an individual trained to qualitatively make a distinction between normal EEG activity and abnormalities contained within EEG records. Epilepsy is a chronic disease characterized from recurrent seizures that cause sudden but reversible changes in the brain functions [3]. The Classification of epilepsy risk levels, according to international standard is difficult because individual laboratory findings and symptoms are often inconclusive [4]. Approximately 1% of the people in the world suffer from epilepsy. The electroencephalogram (EEG) signal is used for the purpose of epileptic detection as it is a condition related to the brain’s activity. A common form of EEG recording used for this purpose is an ambulatory recording which contains EEG data for a long duration of even up to a week [5]. It involves an expert’s effort in analyzing the entire data to detect traces of epilepsy. The traditional methods of analysis being tedious and time consuming, many automated epileptic EEG systems have emerged in recent years [3]. K.P.Adlassnig (1986) characterized the epilepsy disorder as sudden recurrent and transient disturbances of mental function and/or movements of body that results in excessive discharge group of brain cells [2]. The presence of Epileptiform activity in the EEG confirms the diagnosis of epilepsy, which sometimes confused with other disorders producing similar seizure like activity. Between seizures, the EEG of a patient with epilepsy may be characterized by occasional epileptic form transients–spikes and sharp waves.

A. General Techniques

Today, in the mass storage era, knowledge acquisition represents a major knowledge engineering bottle neck.
Computer programs extracting knowledge from data successfully attempt to alleviate this problem. Among such systems, inducing symbolic decision trees for decision making, or classification, are very popular [6]. The resulting knowledge, in the form of decision trees and inference procedures, has been praised for comprehensibility. This appeal to a wide range of users who are interested in domain understanding, classification capabilities, or symbolic rules that may extract from the tree and subsequently used in rule based decision system. Decision trees were popularized by Quinlan with the ID3 program as identified by Cezary.Z (1998) [7]. Perhaps, the most important feature of decision tree is it’s capability to break a complex decision making process into a collection of simpler decisions, thus providing a solution which is easier to interpret [6]. Alison (1993) noticed that the different types of epileptic seizures are characterized by different EEG waveform patterns [8]. With real-time monitoring to detect epileptic seizures gaining widespread recognition, the advent of computers has made it possible to effectively apply a host of methods to quantify the changes occurring based on the EEG signals. One of them is a classification of risk level of epilepsy by using Fuzzy Techniques examined in ref [1].

Many of these patients (20%) are resistant to treatment with drugs. The ability to anticipate the onset of seizures in such cases would permit clinical interventions [9]. Traditional signal analyses, such as the count of focal spike density, the frequency coherence or spectral analyses are not reliable predictors [10]. Many decisions are based on the determination of available alternatives enduring the relevant criteria. In these types of problems, the measurement of the satisfaction to the individual criteria is available [8]. The constructions of overall decision functions are complicated [11]. First, the construction of decision function requires a specification from the responsible decision maker of the relationship between the criteria for aggregation [12]. Once this specification of relationship and the criteria are obtained, the analyst is then facing with the problem of rendering this information into a form that can be evaluated in terms of the satisfaction to the individual criteria, which leads to the formulation of associated Multi Criteria Aggregation function [13]. This situation puts a premium of knowledge representation structures that allow for both a specification aggregation functions [14]. Based on the theory of fuzzy measures and the OWA operators, we introduce a hierarchical representation structures that allow for the construction of decision functions, which meets the above mentioned needs [15]. This paper addresses the application of hierarchical structured decision trees, SVD, K-means clustering, PCA and AI Techniques towards optimization of fuzzy outputs in the classification of epilepsy risk levels. We also present a comparison of these classifiers based on their Performance Indices and Quality values.

II. Materials and Methods

The EEG data used in the study were acquired from twenty epileptic patients who had been under the evaluation and treatment in the Neurology department of Sri Ramakrishna Hospital, Coimbatore, India. A paper record of 16 channel EEG data is acquired from a clinical EEG monitoring system through 10-20 international electrode placing method. With an EEG signal free of artifacts, a reasonably accurate detection of epilepsy is possible; however, difficulties arise with artifacts. This problem increases the number of false detection that commonly plagues all classification systems. With the help of neurologist, we had selected artifact free EEG records with distinct features. These records were scanned by Umax 6696 scanner with a resolution of 600dpi.

A. Acquisition of EEG Data

Since the EEG records are over a continuous duration of about thirty seconds, they are divided into epochs of two second duration each by scanning into a bitmap image of size 400x100 pixels. A two second epoch is long enough to detect any significant changes in activity and presence of artifacts and also short enough to avoid any repetition or redundancy in the signal [16]. The EEG signal has a maximum frequency of 50Hz and so, each epoch is sampled at a frequency of 200Hz. Each sample corresponds to the instantaneous amplitude values of the signal, totaling 400 values for an epoch.

B. Fuzzy System as a Pre Classifier

Fig 1 enumerates the overall epilepsy risk level (Fuzzy-Post optimization) classifier system. The motto of this research is to classify the epilepsy risk level of a patient from EEG signal parameters. This is accomplished as [17]. Fuzzy classification for epilepsy risk level at each channel from EEG signals and its parameters. Fuzzy classifier results from each channel are optimized using four types of HDT, SVD, K-means clustering, PCA, and hand rule optimization methods. A comprehensive performance of fuzzy classifier and post classifiers optimization methods are analyzed. The following parameters are extracted from EEG signals.

1. The energy in each two-second epoch is given by [1]

\[ E = \sum_{i=1}^{n} x_i^2 \]

(1)

Where \( x_i \) is signal sample value and \( n \) is number of samples. The scaled energy is taken by dividing the energy term by 1000.

2. The total number of positive and negative peaks exceeding a threshold is found.

3. Spikes are detected when the zero crossing duration of predominantly high amplitude peaks in the EEG waveform lies between 20 and 70 ms and sharp waves are detected when the duration lies between 70 and 200ms[4].

4. The total numbers of spike and sharp waves in an epoch are recorded as events.

![Figure 1. Fuzzy and Post Processing Classification System](image-url)
5. The variance is computed as $\sigma^2$ given by
\[
\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}
\]  
(2)

Where $\mu = \frac{\sum_{i=1}^{n}x_i}{n}$ is the average amplitude of the epoch.

6. The average duration is given by
\[
D = \frac{\sum_{i=1}^{p}t_i}{p}
\]  
(3)

Where $t_i$ is one peak to peak duration and $p$ is the number of such durations.

7. Covariance of Duration. The variation of the average duration is defined by
\[
CD = \frac{\sum_{i=1}^{p}(D - t_i)^2}{pD^2}
\]  
(4)

C. Fuzzy Membership functions

Energy is compared with the other six input features to give six outputs. Each input feature is classified into five fuzzy linguistic levels viz., very low, low, medium, high and very high [18]. The triangular membership functions are used for the linguistic levels of energy, peaks, variance events, spike and sharp waves, average duration and covariance of duration. The output risk level is classified into five linguistic levels namely normal, low, medium, high and very high [19].

Fuzzy Rule Set

Rules are framed in the format

**IF Energy is low AND Variance is low THEN Output Risk Level is low**

In this fuzzy system we have five linguistic levels of energy and five linguistic levels of other six features such as variance, peaks, events, spike and sharp waves, average duration and covariance of duration. Theoretically there may be $5^6$ (that is 15625) rules are possible but we had considered the fuzzy pre-classifier as a combination of six two inputs and one output (2×1) system. With energy being a constant one input the other input is selected in sequential manner. This two inputs one output (2×1) fuzzy system works with 25 rules. We obtain a total rule base of 150 rules based on six sets of 25 rules each.

D. Estimation of Risk Level in Fuzzy Outputs

The output of a fuzzy logic represents a wide space of risk levels. This is because there are sixteen different channels for input to the system at three epochs. This gives a total of forty-eight input output pairs. Since we deal with known cases of epileptic patients, it is necessary to find the exact level of risk the patient. This will also aid in the development of automated systems that can precisely classify the risk level of the epileptic patient under observation [15]. Hence an optimization of the outputs of the fuzzy system is necessary. This will improve the classification of the patient and can provide the EEGer with a clear picture [18]. A specific coding method processes the output fuzzy values as individual code. Since working on definite alphabets is easier than processing numbers with large decimal accuracy, we encode the outputs as a string of alphabets. The alphabetical representation of the five classifications of the outputs is shown in Table I.

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>U</td>
</tr>
<tr>
<td>Low</td>
<td>W</td>
</tr>
<tr>
<td>Medium</td>
<td>X</td>
</tr>
<tr>
<td>High</td>
<td>Y</td>
</tr>
<tr>
<td>Very High</td>
<td>Z</td>
</tr>
</tbody>
</table>

\[
PI = \frac{PC - MC - FA}{PC} \times 100
\]  
(5)

Where PC – Perfect Classification, MC – Missed Classification, FA – False Alarm, $PI= [(0.5-0.2-0.1)/0.5] \times 100 = 40\%$. The perfect classification represents when the physicians and fuzzy classifier agrees with the epilepsy risk level. Missed classification represents a true negative of fuzzy classifier in reference to the physician and shows Low level as Low level. False alarm represents a false positive of fuzzy classifier in reference to the physician and shows Low level as High level. The performance for Fuzzy classifier is as low as 40%.

- **Figure 2. Fuzzy Logic Output**

Let the fuzzy outputs as shown in Fig 2 is coded with appropriate numerical values. These numerical values are associated with the probability of each coded epilepsy risk level patterns. The five risk levels are encoded as Z>Y>X=W>U in binary strings of length five bits using weighted positional representation as shown in Table II. Encoding each output risk level of the fuzzy output gives us a string of six chromosomes, the value of which is calculated as the sum of probabilities of the individual genes. For example, if the output of an epoch is encoded as ZZYXXZ, its value would be 0.333331, [1]. Now the each input patterns are encoded in the numerical form of the range 0-1. The nonlinearities associated with fuzzy outputs in describing the epilepsy risk levels were identified by cross correlation. Thus the cross correlation function $r_{xy}(m)$ of the epochs $x(n)$ and $y(n)$ is defined by the equation (6) and assuming that both sequence have been measured from $n=0$ to $n=N-1$, in our case $n=1$ to 16,[17] The cross correlation $r_{xy}$ (m) plot obtained through the equation (6) is shown in the “Fig.3”,which emulates the occurrence of highly non periodic patterns in the fuzzy outputs. Therefore any closed solution will be failed for this purpose of optimization. Hence, it is advisable to prefer non linear techniques instead of linear one, such a one type is
HDT. Since, HDT is a common way to solve a wide variety of ill-posed problems which is not necessarily treated as hard constraint one.

Table II. Binary Representation of Risk

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>Code</th>
<th>Binary String</th>
<th>Weight</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very high</td>
<td>Z</td>
<td>10000</td>
<td>16/31=</td>
<td>0.086021</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.51612</td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>Y</td>
<td>01000</td>
<td>8/31=</td>
<td>0.043011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.25806</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>X</td>
<td>00100</td>
<td>4/31=</td>
<td>0.021505</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.12903</td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>W</td>
<td>00010</td>
<td>2/31=</td>
<td>0.010752</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.06451</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>U</td>
<td>00001</td>
<td>1/31=</td>
<td>0.005376</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.03225</td>
<td></td>
</tr>
</tbody>
</table>

$$r_p(m) = \sum_{n=0}^{N-1} \frac{1}{N} \sum_{n=0}^{N-1} x(n + m)y(n), \text{ for } 0 \leq m \leq N - 1$$

Where \( b_j \) is the jth largest of the y_i. Using these two structures we can express the decision function needed to solve the preceding situation [20]. Let \( D(x) \) be the overall alternative x letting \( N_i(x) = n_i \). We get \( D(x) = G(n_1, \ldots, n_8) \) where \( D(x) = E_{1}(F_{a}(n_1, n_2, n_3), \max(F_{a}(n_4, n_5), \min(F_{a}(n_6, n_7, n_8), \
\max(F_{a}(n_9, n_{10}, n_{11}), \min(F_{a}(n_{12}, n_{13}), F_{a}(n_{14}, n_{15}, n_{16})))))) \).

In our approach we consider a decision frame work in which we have a collection, \( N = \{N_1, N_2, N_3, \ldots, N_{16}\} \) of primary attributes. These first level concepts are decomposed into other concepts or primary attributes. We continue until we end up with all primary attributes.

B. Algorithm for HDT Optimization

The generic representation of HDT optimization is explained, let \( W = [P_{ij}] \) be the co-occurrence matrix with \( (i,j) \) elements which represents fuzzy based epilepsy risk level patterns of single epoch and 48 (16x3) patterns are available. Now the optimization is a two stage process through HDT, which is explained as below,

Figure 3. Cross Correlation Function plot for the Adjacent Epochs in fuzzy based Epilepsy Risk Level Outputs

III. Hierarchical Decision Trees for Optimization of Fuzzy Outputs

Our objective is to merge the epilepsy risk level representation, with approximate reasoning capabilities, and symbolic decision trees while preserving advantages of both: uncertainty handling and gradual processing of the former with the comprehensibility, popularity, and ease of application of the later. Hierarchical functions are non linear mapping from \((x_1, x_2, x_3, \ldots, x_n) \in R^k \) to \( y \in R \) and this nonlinear mapping is general enough to approximate any non linear function with arbitrary accuracy. In contrast to conventional single stage classifiers where each data sample is tested against all classes, thereby reducing efficiency, in a decision tree a sample is tested against only certain subsets of classes, therefore unnecessary computations are eliminated[7][13].
1. Deduce the 16x3 matrix epilepsy risk level into 16x1 viz row wise optimization through two types of optimization viz,
a) Hierarchical method of two level, and b) Maximum pattern in the particular row.
2. Deduce the 16x1 matrix into one optimum epilepsy risk level through HDT optimization with five levels.
Here also we have two decision methods at node level which are Max-min & Min-max combination. Therefore effectively we have four methods of HDT post classifier.

Stage I
1. The Hierarchical method converts the three column elements of i,j element into a single row element as N1=Max(N11,N12) & N2=Min(N11,N13) which is also depicted in the Fig. 4. And the other method is self explained in nature. Now the row of three elements is converted into single element. This is repeated for all the 16 rows and the matrix is reduced into 16x1 matrices.

Stage II: Group (16x1) elements as the leaf nodes of the tree N1 to N16. These leaves are aggregated by the rectangular nodes named as A1 to A6. This structure is a mixed averaging method, let N1, N2 ... N16 leaf nodes are available.

The aggregate weights of A nodes are as,

For A1, A3, A4, & A6 = 0.4 and A2 & A5 = 0.6 ; 0.4

The obtained singleton results are discussed in the following part of the paper.

IV. Singular Value Decomposition for Optimization of Fuzzy Outputs

The Singular Value Decomposition (SVD) is a well known approach that may be used for such tasks as dimensionality reduction, and determining the modes of a complex linear dynamical system [22]. SVD of a matrix has one or more columns that are identical, or that several groups of columns that are same which is useful in signal processing problems and applications. A SVD of an m × n matrix A=[a1, a2, a3, …, an] is the composition of A into the product of three matrices as follows

\[ A = U \Sigma V^T \]

where \( p = \min(m,n) \), \( U=[u_1, u_2, u_3, \ldots, u_m] \) is an \( m \times n \) ortho normal matrix, \( V=[v_1, v_2, v_3, \ldots, v_n] \) is an \( n \times m \) ortho normal matrix, and \( \Sigma \) is an \( n \times m \) matrix with elements \( \sigma_k \) along the diagonal and zeros everywhere else. Matrix \( U \) is called left singular matrix, \( V \) is called right singular matrix, and \( \Sigma \) is the singular value matrix [23]. If the singular values are ordered so that \( \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \), and if the matrix A has a rank \( r \), then the last \( p-r \) singular values are equal to zero, and SVD becomes

\[ A = \sum_{k=1}^{r} \sigma_k u_k v_k^T \]

SVD procedure takes vectors in one space and transforms them into another space. Advantages of using SVD are to combine two different uncertainty representations into a metric as total uncertainty. SVD decomposes uncertainty measures (possibility, belief, probability etc.), combined as a collection of vectors of different units, into a principle space. We need this feature since our uncertainty measures cannot be added directly, they contain different units (epilepsy risk level codes). SVD has been applied successfully in many other technical disciplines as a tool to reduce coupled non linear behavior to uncoupled collections of linear behavior [24]. The fuzzy outputs are (16x3 matrix) considered as matrix A and SVD is taken for that matrix. The highest Eigen value is considered as the pattern of the known patient’s epilepsy risk level. A group of twenty patients are analyzed in this study. The obtained singleton results are analyzed in this study. The obtained singleton results are discussed in the following part of the paper.
V. K-Means Clustering Algorithm

During last decades, growing attention has been put on data clustering as robust technique in data analysis. Clustering or data grouping describes important technique of unsupervised classification that arranges pattern data (most often vectors in multidimensional space) in the clusters (or groups). Patterns or vectors in the same cluster are similar according to predefined criteria, in contrast to distinct patterns from different clusters [25], [26]. K-means clustering algorithm proposed by Mac Queen in 1967 belongs to partitioning methods, which is widely used because of its simplicity and fast convergence. The primary process can be expressed as follows [27], [28].

1. Initialize K cluster centers chosen randomly.
2. Assign each xi to its nearest cluster center ck by Euclidean Distance (d).

\[KM(X,C) = \sum_{i=1}^{n} \min_{j \in [1..K]} \left\| x_i - c_j \right\|^2 \]  (10)

\[d(p,q) = \sqrt{\sum_{i=1}^{n} (q_i - p_i)^2} \]  (11)

3. Update each cluster center ck as the mean of all xi that belongs to it.
4. Repeat steps 2-4 until the cluster centers are stable.

K-means clustering is associated with computational expensiveness with respect to the cluster similarity distance measures. The results obtained through the K-Means clustering methods are discussed in the following section of the paper.

VI. Principal Components for Optimization of Fuzzy Outputs

All suboptimal transforms such as the DFT and DCT decomposes the signals into a set of coefficients, which do not necessarily represent in the constituent components of the signals. Moreover, as the transform kernel is independent of the data it is not efficient in terms of both de correlation of the samples and energy compaction. Therefore, separation of the signal and noise components is generally not available using these suboptimal transforms. Expansion of the data into a set of orthogonal components certainly achieves maximum de correlation of signals [29]. This can enable separation of the data into signal and noise subspace. PCA is widely used in data decomposition, classification, filtering, and whitening. Performing a Principal Component Analysis (PCA) is equivalent to performing an SVD on the covariance matrix. PCA uses the same concept as SVD and orthogonalization to decompose the data into constituent uncorrelated orthogonal components such that the auto correlation matrix is diagonalized [29]. Each Eigen vector represents a principle component and the individual eigen values are numerically related to the variance they capture in the direction of principle components. Principal Component Analysis (PCA) is a dimensionality -reduction technique that has been applied to many kinds of data. In fact, PCA is the optimal such linear transform-that is, for any choice for the number of dimensions, PCA returns the subspace that retains the highest variance [30]. In this section, we describe how to use PCA to optimize the code converters outputs.

PCA is a mathematical technique allows reducing the complex system of correlations in a smaller number of dimensions. χ being a table of P numeric variables (in columns) describing N individuals (in lines), w propose to seek a representation of N individuals (signals) e1, e2,...en in a subspace of initial space. In other words, we have to define K new variables, combination of P of initial space, which could make loss less possible information. These K variables will be called principal axes [31].

For N observations, we will have a matrix of NxP size which is given by

\[e = [e_1 \ e_2 \ e_3 \ .. \ e_n] \]  (12)

The average signal is defined by:

\[\psi = \frac{1}{N} \sum_{n=1}^{N} e_m \]  (13)

For each element the difference:

\[\delta_i = e_i - \psi \]  (14)

The computation of covariance matrix is:

\[C = \frac{1}{N} \sum_{n=1}^{N} \delta_m \delta_m^T = \frac{1}{N} A \times A^T \]  (15)

With: \[A= \{ \delta_1 \ \delta_2 \ ... \ \delta_N \} \]  (16)

However, the determination of the Eigen vectors of covariance matrix will require an excessive calculation [32]; the size of this matrix is (P)xP. If \[\psi\] is the Eigen vector of AXA^T its Eigen values are:

\[A^T \psi \psi = \mu \psi \]

Then the Eigen vectors of C are calculated by:

\[U= A \psi \]  (17)

Finally the principal component of each signal e_i is given by:

\[w_k = u_k^T \times (e_i - \delta) \]  (18)

The vector \[w_k\] represents the new parameters completely de correlated and optimized for classification. The results obtained through the PCA methods are discussed in the following section of the paper.

VII. Artificial Intelligence (A.I) Optimization

The AI optimization was a post classifier to classify risk levels of epilepsy based on a few simple rules. The proposed technique for optimizing the classified risk level of epilepsy is based on a few simple rules. The hand-rule optimization is a search process in which a maximized risk level output is obtained from a cluster of non-optimized fuzzy system outputs. In maximizing the risk levels by this optimization technique,

1) A simple weight age is given to the representations as follows, \[Z > Y > X > W > U\]
2) The variance is taken as a second parameter for optimization.
3) The repeated patterns with high risk levels are identified.
4) The occurrences of low risk level patterns are more than the high risked one then the weight age rule (1) is ignored. This process is carried out in three stages: Target stage,
Intermediate stage, and Output stage.
The target stage is the output epilepsy risk level obtained from the fuzzy data. In the target stage, the numerical values of the fuzzy logic outputs are converted into the corresponding encoded representations. Once each epoch is encoded as a string of 6 characters, each channel is divided into sets of three strings each. The intermediate stage is where each of the sets is optimized independently. A set of sixteen strings is obtained from the target stage by following row-wise optimization for every channel in each epoch. As an example, for the state WYYWYY, WYYWYY, WZYYWW the row wise optimized output is WYYXYX, which is obtained by using simple rules as mentioned above. This set of sixteen outputs is again divided into four sets of four strings each. A column-wise optimization procedure is followed wherein each of the set is optimized to a single string. The four strings obtained in this column-wise optimization are again grouped in sets of two and the output stage of operation is carried out. In the output stage, each group of two strings give one optimized string and another step in combining the two outputs yields the final optimized risk level of the epileptic patient.

VIII. Results and Discussion

To study the relative performance of the Fuzzy techniques, HTD systems (4 Types), SVD, K-means clustering, PCA, and AI optimization, we measure two parameters, the Performance Index and the Quality Value. These parameters are calculated for each set of twenty patients and compared.

A. Performance Index

A sample of Performance Index for a known epilepsy data set at average value is shown in table III. It is evident that the HTD Optimization with [MAX & h Max-min] method as well as [MAX & h Min-max] method gives a better performance than the AI optimization, fuzzy techniques and other two hierarchical techniques because of its lower missed classifications. SVD and PCA optimization methods are settled at PI of 95.88% and 95.12% respectively with low missed classification. K-means clustering method is pegged at low PI of 92.122% due to high missed classification of epilepsy risk levels.

B. Quality Value

In Order to compare different classifier we need a measure that reflects the overall quality of the classifier [13].Their quality is determined by three factors namely Classification rate, Classification delay, and False Alarm rate.

The Quality Value QV is defined by,

\[ Q_V = \frac{C}{(R_{fa} + 0.2)*(T_{dl} + P_{dcl} + 6*P_{msd})} \]  \tag{19} \]

Where, C is the scaling constant, \( R_{fa} \) is the number of false alarm per set, \( T_{dl} \) is the average delay of the on set classification in seconds, \( P_{dcl} \) is the percentage of perfect classification and \( P_{msd} \) is the percentage of perfect risk level missed. A constant C is empirically set to 10 because this scale is the value of QV to an easy reading range. The higher value of QV, the better the classifier among the different classifier, the classifier with the highest QV should be the best.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Perfect Classification</th>
<th>Missed Classification</th>
<th>False Alarm</th>
<th>Performance Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy logic</td>
<td>50</td>
<td>20</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>hier &amp; h max-min</td>
<td>95.42</td>
<td>3.33</td>
<td>1.25</td>
<td>95.2</td>
</tr>
<tr>
<td>hier &amp; h min-max</td>
<td>95.63</td>
<td>4.16</td>
<td>0.208</td>
<td>95.43</td>
</tr>
<tr>
<td>Max &amp; h max-min</td>
<td>96.84</td>
<td>0.416</td>
<td>2.17</td>
<td>96.77</td>
</tr>
<tr>
<td>Max &amp; h min-max</td>
<td>97.5</td>
<td>0.416</td>
<td>2.08</td>
<td>97.44</td>
</tr>
<tr>
<td>SVD Method</td>
<td>96.04</td>
<td>1.04</td>
<td>2.92</td>
<td>95.88</td>
</tr>
<tr>
<td>K-means clusters</td>
<td>92.79</td>
<td>4.33</td>
<td>2.875</td>
<td>92.122</td>
</tr>
<tr>
<td>PCA</td>
<td>95.8</td>
<td>1.87</td>
<td>2.25</td>
<td>95.12</td>
</tr>
<tr>
<td>AI optimization</td>
<td>80</td>
<td>5</td>
<td>10</td>
<td>81.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Methods</th>
<th>Weighted delay (s)</th>
<th>False-alarm rate/set</th>
<th>Performance Index %</th>
<th>Quality value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fuzzy logic</td>
<td>4</td>
<td>0.2</td>
<td>40</td>
<td>6.25</td>
</tr>
<tr>
<td>hier &amp; h max-min</td>
<td>2.108</td>
<td>1.25</td>
<td>95.2</td>
<td>22.3</td>
</tr>
<tr>
<td>hier &amp; h min-max</td>
<td>2.1662</td>
<td>0.208</td>
<td>95.43</td>
<td>20.9</td>
</tr>
<tr>
<td>Max &amp; h max-min</td>
<td>1.962</td>
<td>2.71</td>
<td>96.77</td>
<td>22.4</td>
</tr>
<tr>
<td>Max &amp; h min-max</td>
<td>1.975</td>
<td>2.08</td>
<td>97.44</td>
<td>22.9</td>
</tr>
<tr>
<td>SVD Method</td>
<td>1.9832</td>
<td>2.92</td>
<td>95.88</td>
<td>21.99</td>
</tr>
<tr>
<td>K-means clusters</td>
<td>2.1155</td>
<td>2.875</td>
<td>92.17</td>
<td>21.927</td>
</tr>
<tr>
<td>PCA</td>
<td>2.09</td>
<td>2.25</td>
<td>95.12</td>
<td>22.93</td>
</tr>
<tr>
<td>AI optimization</td>
<td>2.8</td>
<td>0.1</td>
<td>81.25</td>
<td>11.9</td>
</tr>
</tbody>
</table>
Table IV shows the Comparison of the fuzzy, HTD SVD, K-means clustering, PCA, and AI optimization techniques. It is observed from table IV, that HTD (Max& h max-min) (Max& h min-max), and PCA methods are performing well with the higher performance index and quality values. As such maximum pattern followed by decision trees are empowered with high false alarm rate and also low weighted delay. This indicates the lower threshold value of the classifiers. SVD and K-means clustering techniques are sustained with Quality value of 21.99 and 21.927 respectively.

IX. Conclusion

In this paper, we consider generic classification of the epilepsy risk level of epileptic patients from EEG signals and investigated the performance of six post classifiers in optimizing the epilepsy risk level of epileptic patients from EEG signals. The parameters derived from the EEG signal are compiled as data sets. Then the fuzzy logic is used to the risk level from each epoch at every EEG channel. HTD SVD, K-means clustering, PCA, and AI optimization techniques were chosen to optimize the risk level by incorporating the low false alarm and near nil missed classifications. HTD (max & hmin-max) has better performance index whereas HTD and PCA performs better than AI optimization techniques and Fuzzy Techniques with high Quality value and with moderate time delay. The performance of post classifiers such as SVD and K-means clustering methods are at the midway between the AI and PCA techniques. From this method we can infer the occurrence of High-risk level frequency and the possible medication to the patients.

The major limitation of this method is that if one channel has a high-risk level, then the entire group will be maximized to that risk level. This will affect the non-epilepsy spike region in the groups. A comparison of EM and SVM will be taken for further studies.

ACKNOWLEDGEMENTS

The authors express their sincere thanks to the Management and the Principal of Bannari Amman Institute of Technology, Sathyamangalam for providing the necessary facilities for the completion of this paper. This research is also funded by AICTE RPS.:F No 8023/BOR/RID/Rps-41/2009-10, dated 10th Dec 2010.

References

[23] Pamela McCauley-Bell and Adedeji B.Badiru, “Fuzzy Modeling and Analytic Hierarchy Processing to Quantify Risk levels Associated
with Occupational Injuries- Part I: The Development of Fuzzy- Linguistic Risk Levels”,
[21] Tomas Calvo, Radko Mesiar, and R.R.Yager,
“Quantitative Weights and Aggregation,” IEEE
Transactions on Fuzzy Systems, 12(1), pp. 62-69,
[22] Sankeun lee, Monson, “Properties of the Singular Value Decomposition for Efficient Data
Technique for noise Reduction in Electroencephalographic signals”, Elsevier Signal
version of the k-means clustering algorithm with a
distance based on cluster symmetry”, IEEE
Transactions on Pattern Analysis and Machine
Intelligence, 23 (6), pp. 674-680, June 2001.
for generalized linear mixed models.”, Journal of
[26] Stephen J Redmond and Conor Heneghan, “A
method for initializing the k-means clustering
algorithm using kD-trees”, Pattern Recognition,
algorithms”; IEEE Transactions on Neural
[28] Jain, M. Murty, P. Flynn, “Data Clustering: A
Review” ACM Computing Surveys, 31 ,
Prediction by supervised Principal Components,”
Journal of the American Statistical Association
Classification of Defects by the SVM Method and
the Principal Component Analysis(PCA)", World
academy of science, Engineering and Technology,
[31] A.Sophian,G.Tian, D.Taylor and J.Rudin, “A
Feature Extraction Technique Based on Principal
Component Analysis for Pulsed Eddy Current
[32] Alexandros. T. tzallas, Markos.G. Tsiopoulos and
Dimitrios. I.Fotiadis, “A Time- Frequency Based
Method for the Detection of Epileptic seizure in
EEG Recording.” Proc.12th IEEE International

Author Biographies

Dr. R. HariKumar received his B.E (ECE) Degree from
REC (NIT) Trichy1988.He obtained his M.E (Applied
Electronics) degree from College of Engineering, Guindy, and
Anna University Chennai in 1990. He was awarded Ph.D. in 1
& C Engg from Anna university Chennai in 2009. He has 22
years of teaching experience at college level. He worked as
faculty at Department of ECE, PSNA College of Engineering &
Technology, Dindigul. He was Assistant Professor in IT at PSG
College of Technology, Coimbatore. He also worked as
Assistant Professor in ECE at Amrita Institute of Technology,
Coimbatore. Currently he is working as Professor ECE at
Bannari Amman Institute of Technology, Sathyamangalam. He
is guiding eight PhD theses in these areas. He has published
forty papers in International and National Journals and also
published around seventy six papers in International and
National Conferences conducted both in India and abroad. His
area of interest is Bio signal Processing, Soft computing, VLSI
Design and Communication Engineering. He is life member of
IETE, IEEE and ISTE.

T.Vijayakumar received his B.E (ECE) degree from
Bharathiyar University in 2003 .He obtained his M.E
(Computer Communication)degree from Anna university
Chennai in 2007. He has five years of teaching experience at
college level. He is pursuing his doctoral programme as a
research scholar in Bio-Medical Engineering under guidance of
Dr. R.Harikumar at Department of ECE, Bannari Amman
Institute of Technology Sathyamangalam. Currently he is
working as Assistant Professor in Department of IT at Bannari
Amman Institute of Technology, Sathyamangalam. His area of
interest is Bio signal Processing, Soft computing, Embedded
systems and Communication Engineering. He is Member of
ISTE.