

ASSESSMENT OF EPILEPSY CLASSIFICATION USING TECHNIQUES SUCH AS SINGULAR VALUE DECOMPOSITION, APPROXIMATE ENTROPY, AND WEIGHTED K-NEAREST NEIGHBORS MEASURES

HARIKUMAR RAJAGURU, SUNIL KUMAR PRABHAKAR*

Department of ECE, Bannari Amman Institute of Technology, Erode, Tamil Nadu, India. Email: sunilprabhakar22@gmail.com

Received: 12 April 2016, Revised and Accepted: 18 April 2016

ABSTRACT

Objective: The main aim of this research is to reduce the dimension of the epileptic Electroencephalography (EEG) signals and then classify it using various post classifiers. For the evaluation and easy treatment of neurological diseases, EEG signals are used. The reflection of the electrical activities of the human brain is obtained by the measurement of potentials in EEG. To study and explore the brain functions in an exhaustive manner, EEG is used by both physicians and scientists. The study of the electrical activity of the brain which is done through EEG recording is a vital tool for the diagnosis of many neurological diseases which include epilepsy, sleep disorders, injuries in head, dementia etc. One of the most commonly occurring and prevalent neurological disorders is epilepsy and it is easily characterized by recurrent seizures.

Methods: This paper employs the concept of dimensionality reduction concepts like Fuzzy Mutual Information (FMI), Independent Component Analysis (ICA), Linear Graph Embedding (LGE), Linear Discriminant Analysis (LDA) and finally Variational Bayesian Matrix Factorization (VBMF). The epilepsy risk levels are also classified using post classifiers like Singular Value Decomposition (SVD), Approximate Entropy (ApEn) and Weighted KNN (W-KNN) classifiers.

Results: The highest accuracy is obtained when LDA is combined with Weighted KNN (W-KNN) Classifiers and it is of 97.18%.

Conclusion: Thus the EEG signals not only represent the brain function but also the status of the whole body. The best result obtained was when LDA is engaged as a dimensionality reduction technique followed by the usage of the W-KNN as post classifier for the classification of epilepsy risk levels from EEG signals. Future work may incorporate the possible usage of different dimensionality reduction techniques with various other types of classifiers for the perfect classification of epilepsy risk levels from EEG signals.

Keywords: FMI, ICA, LGE, LDA, W-KNN, EEG

© 2016 The Authors. Published by Innovare Academic Sciences Pvt Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>) DOI: <http://dx.doi.org/10.22159/ajpcr.2016.v9i5.12196>

INTRODUCTION

Seizures are nothing but the resulting consequences due to the abrupt changes in the basic electrical functioning of the brain, which results in altered behaviors such as memory loss, jerky movements, and temporary loss of health [1]. Due to some abnormalities in the wiring of the brain epilepsy occurs, it may also occur due to the imbalance of neurotransmitters, which are the most important nerve signaling chemicals [2]. Between neurons, electrochemical impulses are always generated, and therefore, thoughts, feelings, and actions are controlled by it only. Therefore, if epilepsy occurs, then the normal pattern of the activities of the neuron becomes heavily disturbed which causes strange behaviors accompanied by the loss of consciousness [3]. The neurons can become unstable because of genetic defects or due to metabolic abnormalities. Sometimes, because of a severe head injury or a brain tumor, the abnormal discharge occurs from the localized areas of the brain [4]. When a seizure occurs, the neurons present in the brain can fire as many as 500 times a second, much faster than the normal (1-100 μ V). The seizures can happen either occasionally, or it can happen up to more than hundreds of times a day also [5]. To understand epilepsy, the most valuable information is obtained from EEG. The abnormalities can be shown clearly on a standard EEG due to this epileptic activity [6]. For the diagnosis and treatment of epilepsy, the detection of seizures which is occurring in the EEGs forms a vital component.

Many versatile mathematical techniques have been implemented and analyzed for the processing of the EEG data [7]. The main aim of such mathematical techniques is to extract the significant features, to reduce the unwanted dimensions and also to gain insight about the

spatiotemporal nature of the signals. To differentiate between the EEG signals of the diseased patients with that of the normal ones, such mathematical techniques are useful. The EEG recordings always have a very huge amount of data and to process the whole EEG recorded signals is quite a hectic task. So, it is important to reduce the dimensions of the EEG recorded data, and then, this data can be provided as an input to the classifiers for the classification of epilepsy risk levels from EEG signals. Several dimensionality techniques for the processing of the epileptic EEG data have been discussed in Harikumar and Kumar [8]. Dimensionality reduction is achieved by the projection of the data into a lower dimensional space. By selecting the appropriate channels in a given data, dimensionality can be easily reduced [9]. By the projection of all the EEG data into a particular time domain signal which has a single dimension, the dimension of the data is easily reduced [8]. After reducing the dimension of the EEG data, it has to be classified to assess the epileptic risk levels.

The classification of the EEG signals plays a significant role in the plethora of biomedical research [10]. For the clear diagnosis of the diseases related to the brain, especially epilepsy and to gain a better insight about the cognitive processes, classification of the EEG signals play a key role. The EEG segments can be distinguished clearly with the help of any efficient classification techniques. Classification seems to be very important as it occurs in our daily lives also, which means that good decisions can be manipulated based on the currently available information. In the field of pattern recognition and machine learning, the process of classification always refers to a certain algorithm for assigning an input data into a particular number of categories which is defined. Thus, classification always assigns the class labels to the extracted features from the given observations in a particular data.

With the advent of neural networks being used for the detection of epileptic seizures, a great breakthrough in the field of biomedical signal processing was evolved [11,12]. Therefore, in this paper, fuzzy mutual information (FMI), independent component analysis (ICA), linear graph embedding (LGE), linear discriminant analysis (LDA), and variational Bayesian matrix factorization (VBMF) are used as dimensionality reduction techniques followed by the singular value decomposition (SVD), approximate entropy (ApEn), and W-KNN as post classifier for the classification of epilepsy risk levels from EEG signals. The organization of the paper is as follows: In section 2, the methods and materials are discussed followed by the dimensionality reduction techniques in section 3. Section 4 introduces the concept of SVD, ApEn, and W-KNN as post classifiers for the classification of epilepsy risk levels from EEG signals, and the results and discussion are given in section 5 followed by conclusion.

METHODS

For the performance assessment of the epilepsy risk levels using the FMI, ICA, LGE, LDA, and VBMF as dimensionality reduction technique followed by SVD, ApEn, and W-KNN as post classifiers, the raw EEG data of 20 epileptic patients who were under treatment in the Neurology Department of Sri Ramakrishna Hospital, Coimbatore in European Data Format are taken for study. The EEG is recorded by placing electrodes on the scalp according to the international 10-20 system. 16 channels of EEG are recorded simultaneously for both referential montages, where all electrodes are referenced to a common potential-like ear, and bipolar montages, where each electrode is referenced to an adjacent electrode. Recordings are made while the patient is fully awake but in resting condition and include periods of eyes open, eyes closed, hyperventilation and photonic stimulation. Amplification is provided by an EEG-machine (Siemens Minograph Universal). Before placing the electrodes, the scalp is cleaned, lightly abraded and electrode paste is applied between the electrode and the skin. Using this application of electrode paste, the contact impedance is <10 ΩW. In general, disks such as surface electrodes are used. In some cases, needle electrodes are used to pick up the EEG signals. The signals are recorded with the speed of 30 mm/s.

The pre-processing stage of the EEG signals is given more attention because it is vital to use the best available technique in literature to extract all the useful information embedded in the non-stationary biomedical signals [13]. The EEG recordings were done for about 30 minutes, and each of them was divided into epochs of 2-second duration. In general, a 2-second epoch is long enough to avoid unnecessary redundancy in the signal, and it is long enough to detect any significant changes in activity and to detect the presence of artifacts in the signal [13]. For each and every patient, the total number of channels is 16 and it is over three epochs. The frequency is considered to be 50 Hz and the sampling frequency is considered to be about 200 Hz. Each and every sample corresponds to the instantaneous amplitude values of the signal which totals to 400 values for an epoch. The total number of artifacts present in the data is four. Chewing artifact, motion artifact, eye blink, and electromyography are the four numbers of artifacts present and approximately the percentage of data which are artifacts is 1%. No attempts were made to select certain number of artifacts which are of more specific nature. The main objective to include artifacts is to differentiate the spike categories of waveforms from non-spike categories. The Fig. 1 shows the block diagram of the procedure.

Dimensionality reduction techniques

Reducing the dimensions is a basic pre-processing step, and it is important because the EEG data recorded is too huge to process, and therefore, the dimensions of the EEG data have to be reduced. The dimensionality reduction techniques employed here are FMI, ICA, LGE, LDA and VBMF, respectively.

LGE

A model training sample set is represented in a matrix form as $P=[p_1, p_2, \dots, p_N]$, where N represents the total number of samples. If $p_i \in R^q$ is

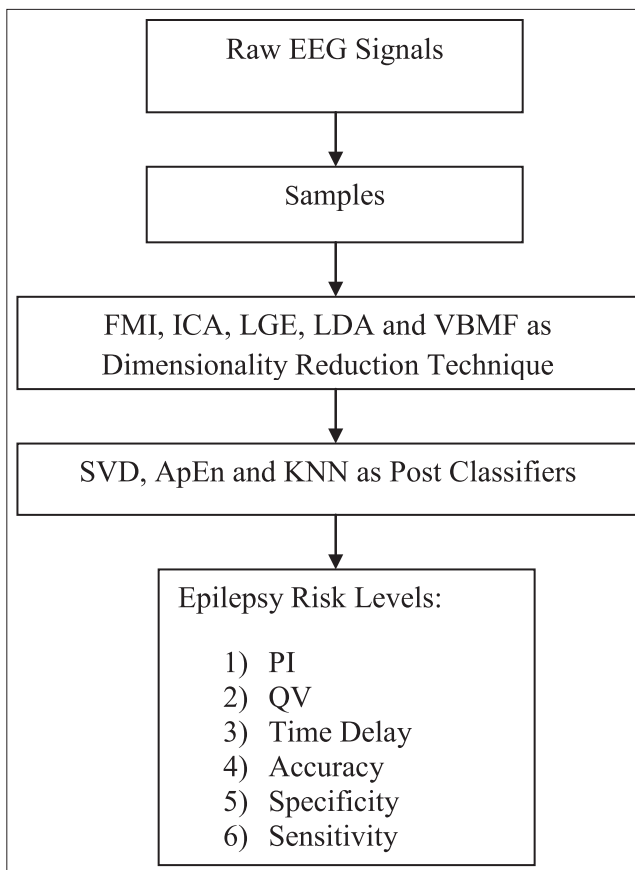


Fig. 1: Block diagram of the procedure

assumed, where q represents the feature dimensions, the value of this feature dimensions is too large, and therefore, high dimensional data are converted into a low dimensional data [1]. Hence, a mapping is done easily as follows:

$$F : p \rightarrow \hat{v}$$

The main function always transforms $p \in R^q$ into a low-dimensional space representation [14] as $\hat{v} \in R^{q'}$, when $q \gg q'$.

Therefore, it is represented as follows:

$$\hat{y} = \frac{F}{p}$$

FMI

Fuzzifying the data for the computation of mutual information is a very good technique [15]. The B-spline functions are generally used here for FMI process. Between any two variables, the FMI is generally calculated, and it is bounded by the minimum entropy levels. The FMI is given as follows:

$$FMI(C, f) = \frac{I(C, f)}{\min\{H(c), H(f)\}}$$

Where, C denotes the fuzzy entropy and $H(c)$ is represented as follows:

$$H(c) = - \sum_{i=1}^n P_i \log_2 P_i$$

ICA

Assuming that there are totally 'n' linear mixtures as x_1, \dots, x_n , where 'n' represents the independent components, it can be written mathematically as follows:

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \dots + a_{jn}s_n \text{ for all } j.$$

The vector-matrix notation is utilized completely, and the above equation can be written as follows:

$$x = As$$

Where, A denotes the matrix with particular elements a_{ij} , x is the random row vector of $x_1 \dots x_n$ or sometimes x^T is used which denotes the transpose of the row vector, s is also the random row vector of $s_1 \dots s_n$. Emphasizing the importance of columns of matrix A , the model can be written as follows:

$$x = \sum_{i=1}^n a_i s_i, \text{ where, } a_i \text{ denotes the columns of matrix } A. \text{ It is considered}$$

as a generative model where an observed data are described clearly. If the matrix A is estimated, then the computation of its inverse, say P , is obtained easily and then the independent component is obtained as follows:

$$S = Px$$

LDA

LDA is basically a supervised subspace learning method based on the Fisher criterion [16]. A linear transformation is aimed such that $V \in R^{d \times m}$ maps x_i in the a dimensional space to a b dimensional space. In such cases, the maximization is between the class scatters and the minimization is within class scatters such that:

$$\arg \max_w ((V^T Q_c V)^{-1} (V^T Q_v V))$$

Where, Q_c and Q_v are the corresponding between-class scatter matrix and within class matrix and is defined as follows:

$$Q_c = \sum_{k=1}^c n_k (\mu_k - \mu)(\mu_k - \mu)^T$$

$$Q_v = \sum_{k=1}^c \sum_{i \in C_k} (x_i - \mu_k)(x_i - \mu_k)^T$$

Where, the index set is represented by C_k , the mean vector is represented by μ_k , and n_k is the size of the k^{th} class, respectively. Therefore, the original mean vector of the data is represented as follows:

$$\mu = \sum_{k=1}^c n_k \mu_k$$

VBMF

The main goal of VBMF [17] is to approximate a particular unknown target $E \in B^{D \times M}$ from its 'n' observations such that:

$$A^n = \{A^{(i)} \in B^{D \times M}\}_{i=1}^n$$

If the assumption is made such that $D \leq M$, a simple redefinition of the transpose E^T is made as E so that $D \leq M$ holds. The most vital assumption of the matrix factorization is that E should always be a low-rank matrix. Assume that $C(\leq D)$ be the respective rank of E . The matrix E can be decomposed as the product of $F \in B^{M \times C}$ and $G \in B^{D \times C}$ as follows:

$$E = GF^T$$

Bayesian inference is very important in matrix factorization models, which includes the case of conjugate Gaussian likelihood potentials.

Post classifiers used here

The post classifiers used here are the ApEn, SVD, and W-KNN classifiers for the classification of epilepsy risk levels from EEG signals.

ApEn as post classifier

The complexity and irregularity of the signal can be easily quantified and measured [18]. The algorithm is as follows:

The data points are assumed as L and so the sequence is considered as follows:

$$D = d(1), d(2), d(3), \dots, d(L)$$

Assuming $d(i)$ is a subsequence of D such that $d(i) = [d(i), d(i+1) \dots d(i+m-1)]$, where m is representing the samples in between the subsequence. The distance $q(d_i, d_j)$ between the two corresponding vectors is measured. If the pre-defined threshold is considered as t , then the similarity of patterns [19] is computed as follows:

$$C_t^m(i) = \frac{N^m(i)}{N - m + 1}$$

Where, $N^m(i)$ is the count. The natural logarithm of $C_t^m(i)$ is found out and averaged for all values of i and is expressed mathematically as follows:

$$\phi^m(t) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_t^m(i)$$

Finally, the ApEn is calculated as follows [20]:

$$ApEn = \phi^m(r) - \phi^{m-1}(r)$$

SVD as a post classifier

The SVD is used here as a post classifier for the classification of epilepsy risk levels from EEG signals. The matrix present here can be decomposed into individual several component matrices so that the interesting properties of the original matrix are exposed [21]. To determine the principal components of a multi-dimensional signal, we can use the method of SVD easily. Consider a real $M \times N$ matrix X with many observations which may be decomposed as follows [21];

$$X = USV^T$$

Where, S is an M non-square matrix with zero entries anywhere, except on the leading diagonal with elements S_i arranged in descending order of magnitude. Each S_i is equal to $\sqrt{\lambda_i}$ the square root of the Eigen value of $C = X^T X$. A stem plot of these values against their index i is known as the singular spectrum. The smaller the Eigen values are the less energy along the corresponding Eigen vector there is. So, the smallest Eigen values are often considered to be due to noise. The columns of V are an $N \times N$ matrix of column vectors, which are the Eigen vectors C . Therefore, the $M \times M$ matrix U is the matrix of projections of X onto the Eigen vectors of C . If a truncated SVD of X is performed, then the truncated SVD is given by $Y = US_p V_p^T$ and the columns of $M \times N$ matrix Y are the noise reduced signal. SVD is advantageous since it combines two different uncertainty representations into a metric as total uncertainty and it also decomposes uncertainty measures (possibility, belief, probability, etc.) as a collection of vectors of different units, into a particular principle space. SVD is also used in various other techniques to reduce coupled non-linear behavior to uncoupled collections of linear behavior.

Weighted KNN (WKNN) classifier

The nearest neighbor rule (NN) is one of the simplest, oldest, and traditional classifier used in the field of pattern classification [22]. If a set of training samples and query are given, the aim of NN algorithm is to find a point which matches very closely to the query followed by the assignment of its class label to the query. In case of a W-KNN algorithm, based on the majority role of its k -nearest neighbors, the query can be labeled in a training set.

$$q_i = \begin{cases} \frac{d(p', p_k^{NN}) - d(p', p_1^{NN})}{d(p', p_k^{NN}) - d(p', p_1^{NN})} & \text{if } d(p', p_k^{NN}) \neq d(p', p_1^{NN}) \\ 1 & \text{if } d(p', p_k^{NN}) = d(p', p_1^{NN}) \end{cases}$$

Let $T = \{(p_i, l_i)\}_{i=1}^N$ denote the training set, where $p_i \in R^m$ is training vector and l_i is the corresponding class label. If a query x' is given, then its unknown class l' is assigned with the help of Euclidean distance functions and the prediction of the class labels of the query. In WKNN, the neighbors, which are closer to each other, are weighted more heavily; the neighbors, which are far from one another, are weighted less heavily. The weight w_i for the i^{th} NN of the given p' is defined as follows:

$$q_i = 1 \quad \text{if } d(p', p_k^{NN}) = d(p', p_1^{NN})$$

Decided by the majority weighted noting, the classification result of the query is made as follows:

$$l' = \underset{y}{\operatorname{argmax}} \sum_{(p_i^{NN}, l_i^{NN}) \in T} q_i p \delta(l = l_i^{NN})$$

RESULTS AND DISCUSSION

For FMI, ICA, LGE, LDA, and VBMF as dimensionality reduction techniques and SVD, ApEn, and W-KNN as post classifiers, based on the performance index (PI), quality values, time delay, and accuracy, the results are computed in Tables 1-3, respectively. The formulae for the PI, sensitivity, specificity, and accuracy are given as follows:

$$PI = \frac{PC - MC - FA}{PC} \times 100$$

Where, PC: Perfect classification, MC: Missed classification, FA: False alarm.

The sensitivity, specificity, and accuracy measures are stated by the following:

$$Sensitivity = \frac{PC}{PC + FA} \times 100$$

$$Specificity = \frac{PC}{PC + MC} \times 100$$

$$Accuracy = \frac{Sensitivity + Specificity}{2} \times 100$$

The time delay and the quality value measures are given by the following:

$$Timedelay = \left[2 * \frac{PC}{100} + 6 * \frac{MC}{100} \right]$$

$$Qualityvalues = \frac{10}{\left[\frac{FA}{100} + 0.2 \right] * Time \quad Delay}$$

On the careful examination of Table 1, it is evident that the LDA with ApEn combination gives the highest accuracy as of 94.89% and a high-quality value of about 19.08 also is obtained.

On the careful analysis of Table 2, it is inferred that the highest accuracy is provided in the FMI-SVD combination than the other combinations with SVD. Furthermore, the perfect classification (PC) of FMI-SVD combination was found to be about 90.83% with a higher quality value of 20.62.

On the careful analysis of Table 3, it is observed that the highest accuracy is obtained in LDA-WKNN combination when compared to the other combinations. The Figs. 2-5 provide the accuracy measures, quality value measures, time delay measures, and PI measures, respectively. If the accuracy measures are considered, it is WKNN classifier which outperforms the other two classifiers and with LDA as dimensionality reduction technique WKNN provides the highest

Table 1: Performance analysis of dimensionality reduction technique with ApEn classifiers

Dimensionality reduction techniques	PC	MC	FA	PI	Sensitivity	Specificity	Time	Quality	Accuracy
FMI+ApEn	87.29	3.47	9.23	84.81	90.76	96.52	1.95	18.73	93.64
ICA+ApEn	86.18	4.37	9.44	82.82	90.55	95.62	1.98	18.70	93.09
LDA+ApEn	89.79	0.48	9.72	88.15	90.27	99.51	1.82	19.08	94.89
LGE+ApEn	85.83	3.61	10.55	82.99	89.44	96.38	1.93	18.21	92.91
VBMF+ApEn	83.61	0.69	15.69	78.84	84.30	99.30	1.71	17.56	91.80

FMI: Fuzzy mutual information, ICA: Independent component analysis, LDA: Linear discriminant analysis, LGE: Linear graph embedding, VBMF: Variational Bayesian matrix factorization, ApEn: Approximate entropy, PC: Perfect classification, MC: Missed classification, FA: False alarm, PI: Performance index

Table 2: Performance analysis of dimensionality reduction technique with SVD classifiers

Dimensionality reduction techniques	PC	MC	FA	PI	Sensitivity	Specificity	Time	Quality	Accuracy
FMI+SVD	90.83	4.54	4.61	89.39	95.38	95.45	2.08	20.62	95.41
ICA+SVD	78.47	10.34	11.18	69.56	88.81	89.65	2.19	17.01	89.23
LDA+SVD	83.54	2.01	14.44	78.81	85.55	97.98	1.79	17.57	91.77
LGE+SVD	85.83	3.61	10.55	82.99	89.44	96.38	1.93	18.21	92.91
VBMF+SVD	81.04	0	18.95	75.42	81.04	100	1.62	16.42	90.52

FMI: Fuzzy mutual information, ICA: Independent component analysis, LDA: Linear discriminant analysis, LGE: Linear graph embedding, VBMF: Variational Bayesian matrix factorization, PC: Perfect classification, MC: Missed classification, FA: False alarm, PI: Performance index, SVD: Singular value decomposition

Table 3: Performance analysis of dimensionality reduction technique with K-NN classifiers

Dimensionality reduction techniques	PC	MC	FA	PI	Sensitivity	Specificity	Time	Quality	Accuracy
FMI+K-NN	94.37	5.62	0	93.97	100	94.37	2.22	22.50	97.18
ICA+K-NN	93.68	6.31	0	93.04	100	93.68	2.25	22.29	96.84
LDA+K-NN	94.37	5.62	0	93.78	100	94.37	2.22	22.57	97.18
LGE+K-NN	93.81	6.18	0	93.32	100	93.81	2.24	22.29	96.90
VBMF+K-NN	93.75	6.25	0	93.19	100	93.75	2.25	22.29	96.87

FMI: Fuzzy mutual information, ICA: Independent component analysis, LDA: Linear discriminant analysis, LGE: Linear graph embedding, VBMF: Variational Bayesian matrix factorization, PC: Perfect classification, MC: Missed classification, FA: False alarm, PI: Performance index, K-NN: K-nearest neighbors

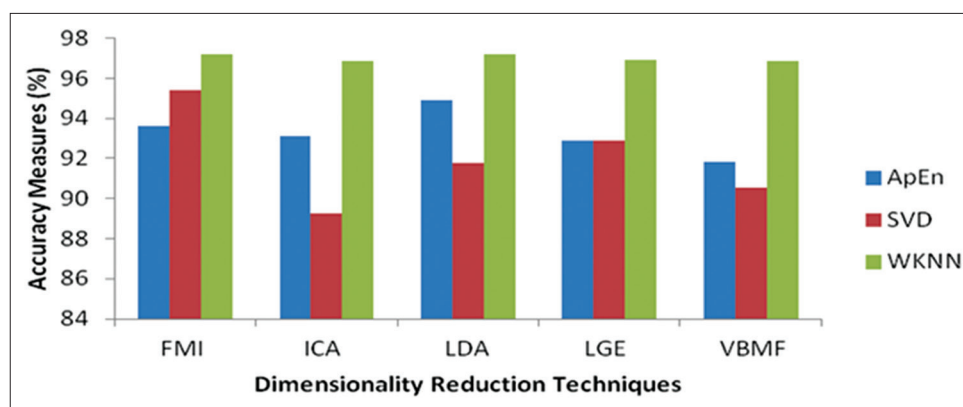


Fig. 2: Performance comparison of accuracy measures

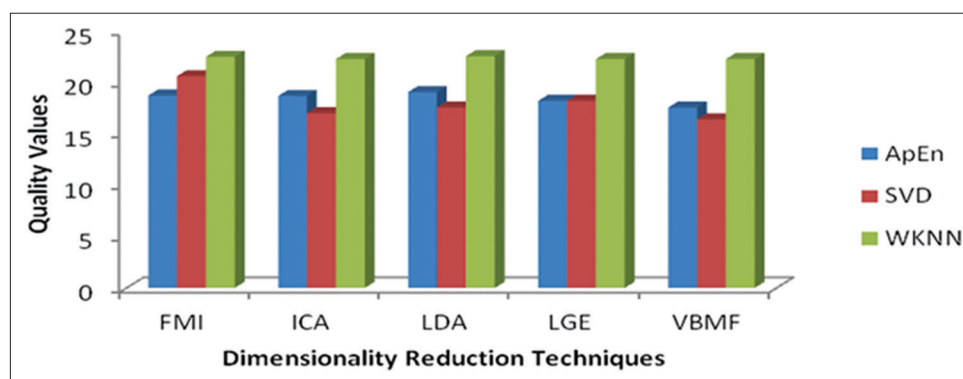


Fig. 3: Performance comparison of quality value measures

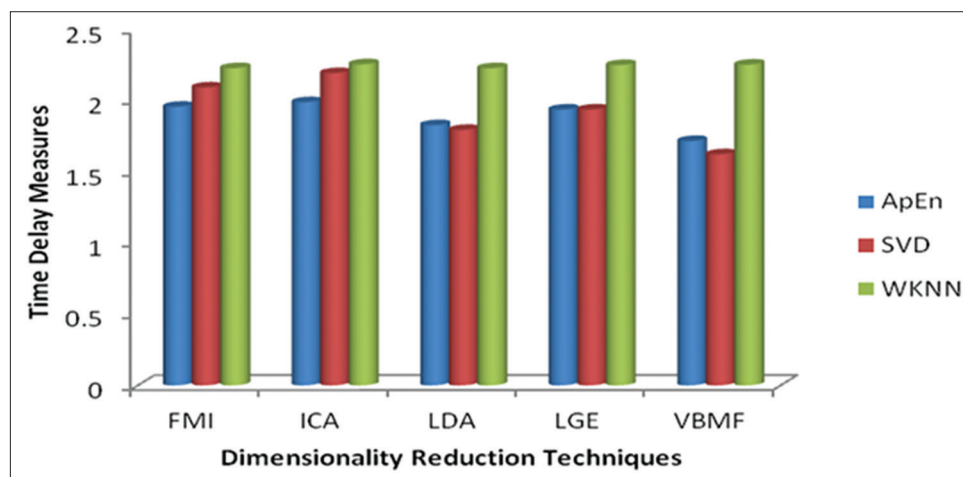


Fig. 4: Performance comparison of time delay measures

accuracy. Similarly, if the quality value measure is considered, FMI-SVD combination provides the highest quality value as of 20.62 with an overall accuracy of 95.41%, whereas the LDA-WKNN combination provides the second highest quality value of 22.57. In terms of time delay, LDA-ApEn provides a low value and in terms of PI WKNN classifier outperforms the other two classifiers and a high PI is found for the LDA-WKNN classifier.

CONCLUSION

Thus the EEG signals not only represent the brain function but also the status of the whole body, i.e., a simple action as blinking the eyes introduces oscillation in the EEG records. Then, the EEG is a direct way to measure neural activities, and it is important in the area of

biomedical research to understand and develop new processing techniques. EEG signal pre-processing and post-processing methods could be considered as a "pattern recognition system" with focus on the classification algorithms. In this research, the dimension of the EEG data was reduced using five different dimensionality reduction techniques, and then, it is classified using three types of classifiers such as SVD, ApEn, and FMI +W-KNN. Out of all the results, the best result obtained was when LDA is engaged as a dimensionality reduction technique followed by the usage of the FMI+W-KNN as post classifier for the classification of epilepsy risk levels from EEG signals. Future work may incorporate the possible usage of different dimensionality reduction techniques with various other types of classifiers for the PC of epilepsy risk levels from EEG signals.

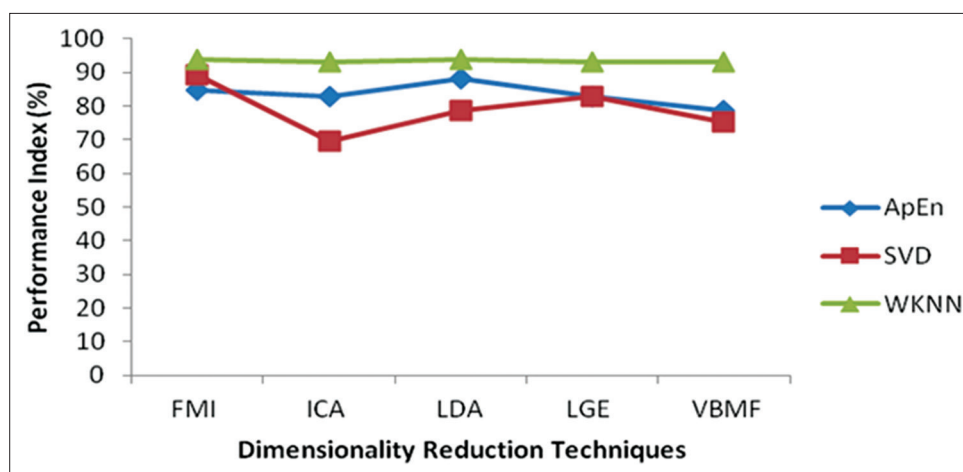


Fig. 5: Performance comparison of performance index measures

REFERENCES

- Prabhakar SK, Rajaguru H. Analysis of centre tendency mode chaotic modeling for electroencephalography signals obtained from an epileptic patient. *Adv Stud Theor Phys* 2015;9(4):171-7.
- Gotman J. Automatic recognition of epileptic seizures in the EEG. *Electroencephalogr Clin Neurophysiol* 1982;54(5):530-40.
- Finley KH, Dynes JB. Electroencephalographic studies in epilepsy: A critical analysis. *Brain* 1942;65:256-65.
- Liu A, Hahn JS, Heldt GP, Coen RW. Detection of neonatal seizures through computerized EEG analysis. *Electroencephalogr Clin Neurophysiol* 1992;82(1):30-7.
- van Gils M, Falck AR, White S, Piror P, Gade J. Signal processing in prolonged EEG recordings during intensive care. *IEEE EMB Mag* 1997;16(6):56-63.
- Kumar Y, Daniel ML, Anand RS. Feature Extraction of EEG Signals Using Approximate and Sample Entropy. 2012 IEEE Students Conference on Electrical, Electronics and Computer Science; 2012.
- Easwaramoorthy, Uthayakumar R. Analysis of biomedical EEG signals using wavelet transforms and multiultrafractal analysis. *IEEE EMB Mag* 2010;30:74-87.
- Harikumar R, Kumar PS. Dimensionality reduction techniques for processing epileptic encephalographic signals. *Biomed Pharmacol J* 2015;8(1):103-6.
- Xu L, Cheung C, Yang H, Amari S. Independent component analysis by the information-theoretic approach with mixture of densities. *IEEE Trans Neural Netw* 1997;5:1821-6.
- Bair E, Hastie T, Paul D, Tibshirani R. Prediction by supervised principal components. *J Am Stat Assoc* 2006;101(473):119-37.
- Harikumar R, Kumar PS. Classifiers for the epilepsy risk level classification from electroencephalographic signals. *Res J Pharm Biol Chem Sci* 2015;6(4):469.
- Hazarika N, Chen JZ, Tsoi AC, Sergejew A. Classification of EEG signals using the wavelet transform. *Signal Process* 1997;59(1):61-72.
- Harikumar. R, Kumar PS. Analysis of singular value decomposition as a dimensionality reduction technique and SRC as a post classifier for the classification of epilepsy risk levels from EEG signals. *J Chem Pharm Sci* 2015;8(2):191-4.
- Yan S, Xu D, Zhang B, Zhang HJ. Graph Embedding: A General Framework for Dimensionality Reduction. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2005.
- Sanchez L. A fuzzy definition of mutual information with application to the design of genetic fuzzy classifiers. Tozeur, Tunisia: International Conference on Machine Intelligence; 2005. p. 5-7.
- Gu Q, Li Z, Han J. Linear Discriminant Dimensionality Reduction, Machine Learning and Knowledge Discovery in Databases. Vol. 6911. Berlin: Springer; 2011. p. 549-64.
- Seeger M, Bouchard G. Fast Variational Bayesian Inference for Non-conjugate Matrix Factorization Models *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics (AISTATS)*; 2012. p. 1012-6.
- Guo L, Rivero D, Pazos A. Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks. *J Neurosci Methods* 2010;193(1):156-63.
- Pravin Kumar S, Sriraam N, Benakop PG, Jinaga BC. Entropies based detection of epileptic seizures with artificial neural network classifiers. *Exp Syst Appl* 2010;37:3284-91.
- Ocak H. Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Exp Syst Appl* 2009;36(2):2027-36.
- Prabhakar SK, Rajaguru H. Performance comparison of fuzzy mutual information as dimensionality reduction techniques and SRC, SVD and approximate entropy as post classifiers for the classification of epilepsy risk levels from EEG signals. *Proceedings of 2015 IEEE Student Symposium in Biomedical Engineering and Sciences (ISSBES)*. Malaysia: Universiti Teknologi Mara; 2015.
- Vivencio DP. Feature-Weighted K-Nearest Neighbour Classifier. *Proceedings of the IEEE Symposium on Foundations of Computational Intelligence*; 2007. p. 481-6.