A Framework for Schizophrenia EEG Signal Classification With Nature Inspired Optimization Algorithms

SUNIL KUMAR PRABHAKAR, (Member, IEEE), HARIKUMAR RAJAGURU, (Fellow, IEEE), AND SEONG-WHAN LEE, (Fellow, IEEE)

1Department of Brain and Cognitive Engineering, Korea University, Seoul 02841, South Korea
2Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam, 638401, India
3Department of Artificial Intelligence, Korea University, Seoul 02841, South Korea

Corresponding author: Seong-Whan Lee (sw.lee@korea.ac.kr)

This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea Government (MSIT) (No. 2019-0-00079, Department of Artificial Intelligence, Korea University).

ABSTRACT

One of the severe and prolonged disorder of the human brain which disturbs the behavioral characteristics of an individual completely such as interruption in the thinking process and speech is schizophrenia. It is a manifestation of many symptoms such as hallucinations, functional deterioration, disorganized speech and hearing sounds and speeches that are non-existent. In this paper, a computerized approach based on optimization and classification is done to analyze the classification of schizophrenia from Electroencephalography (EEG) signals. As EEG can analyze a lot of brain disorders and is used to study the diseases of the brain in an in-depth manner, it can be used to analyze the schizophrenia EEG signals. In this paper, three feature extraction techniques are employed such as Partial Least Squares (PLS) Non linear Regression technique, Expectation Maximization based Principal Component Analysis (EM-PCA) technique and Isometric Mapping (Isomap) technique. The extracted features are further optimized with four optimization algorithms such as Flower Pollination algorithm, Eagle strategy using different evolution algorithm, Backtracking search optimization algorithm and Group search optimization algorithm. The optimized values are then classified with varied versions of both Adaboost classifier and Naïve Bayesian Classifier. The individual results show that for normal cases, Isomap features when optimized with Backtracking search optimization algorithm and classified with Modest Adaboost classifier, a classification accuracy of 98.77% is obtained. The individual results show that for schizophrenia case, when Isomap features are optimized with Flower Pollination optimization algorithm and classified with Real Adaboost classifier, a classification accuracy of 98.77% is obtained.

INDEX TERMS EEG, schizophrenia, regression, optimization, classifiers.

I. INTRODUCTION

One of the chronic psychiatric disorders which troubles human beings to a great extent is schizophrenia [1]. When a patient is affected with schizophrenia, the patient exhibits a high level of disturbance in thoughts and perceptions. When affected with schizophrenia, the patient feels extreme difficulty in dealing with relationships [2]. When a person is affected with schizophrenia, the day to day activities of the person is severely affected thereby affecting the employment, marriage and lifestyle [3]. There are both positive and negative symptoms in schizophrenia. The negative symptoms include lack of normal capabilities to perform routine tasks, lack of confidence and motivation to do even the simplest task accompanied by lack of speech. Auditory hallucinations and delusions along with thought disorders are considered to be positive symptoms. Thus, due to the disturbance in some brain functions, schizophrenia easily occurs and people in various sectors or professions have been troubled by this disorder. To the individual this disorder has caused damage in a microlevel and to the country, this disorder has caused a huge damage in macrolevel affecting the economic system.
of the country too. When a person is showing psychological signs of incapacity to perform the routine duties, then the patient is tested for this psychotic disorder [4]. When a person is affected with schizophrenia and is at the initial stages at the workplace the colleagues and teammates treat him with suspicion not knowing the natural and original condition of the person. A limb mutilation, serious muscle injury, infections due to medicine and allergy and even cancer can bring back the person to normal social life but if a person is affected with this psychological disorder, then it is very difficult for the person to return back to social life. For the diagnosis of mental disorders, EEG has emerged as a powerful tool because it can interpret the brain state so well [5]. EEG is the simple measurement of electrical activity with the help of electrodes placed on various locations on the scalp. Due to its efficiency as a quick recordable tool for analyzing the general cognitive activity, this low-resolution diagnosis tool is widely used everywhere. Thus, the dynamics of the brain can be analyzed by EEG as this provides a high spatio-temporal data [6]. EEG even shows very small intra-personal differentiation with large inter-personal differentiation. EEG is thus quite feasible to diagnose a lot of neurological disorders. Some of the important works done in the schizophrenia analysis such as classification and interpretation of this neurological disorder is given below.

The generalized application of EEG for analyzing schizophrenia patients were in the following fields such that it is was used to assess the functional connectivity and working memory [7], analyzing the Event Related Potentials during selective attention for single subject [8], combining odd ball and mismatch evoked potential paradigms for the single subject classification [9], alpha resting analysis based on nonlinear dynamic studies [10], the connectivity maps-based analysis [11], graph based analysis in brain connectivity [12], heat diffusion based dissimilarity analysis [13], resting EEG in the first episode and chronic schizophrenia stages [14], computation of the abnormal dynamics of EEG oscillations on multiple time scales [15] and the visual evocation of emotion through EEG entropy analysis in schizophrenia was reported in [16].

The non-linear complexity analysis of brain functional Magnetic Resonance Imaging (fMRI) schizophrenia signals was reported in [17] and the multimodal classification with both Magnetoencephalography (MEG) and Functional Magnetic Resonance Imaging (fMRI) data utilizing both static and dynamic connectivity measures was reported in [18]. For the nonlinear analysis [19] of schizophrenia EEG data, principal Lyapunov exponent was calculated in [20], the long-range correlations in choice sequences was performed in [21], the dimensional complexity was reported in [22], multifractal behavior of schizophrenia EEGs was reported in [23], signal complexity analysis for schizophrenia patients in [24], multiscale entropy analysis in [25] and entropy modulation deficit analysis was reported in [26].

Artificial Neural Networks (ANN) was used to classify the schizophrenia with the help of EEG by Li and Fan [27] and reported an average classification accuracy of 60% only. A multi-domain Convolutional Neural Networks (CNN) was proposed by Phang et al. for the classification of EEG based brain connectivity networks in schizophrenia and reported an average classification accuracy of 91.69% [28]. A holistic approach of classification of schizophrenia based on EEG signals was done by Boostani et al. and they reported an average classification accuracy of 87.51% [29]. The neural classifiers for schizophrenia diagnosis support and diffusion imaging data was done by Mayo et al. and reported a specificity of 0.967 and a sensitivity of 0.872 [30]. For the automated diagnosis of schizophrenia using EEG signals, a deep CNN model was proposed by Oh et al. and they obtained 98.07% for non-subject and 81.26% for subjects [31]. For both schizophrenia and healthy adults, machine learning techniques was utilized to identify the EEG features for prediction of the working memory by Chen et al. reporting a highest classification accuracy of 87% [32]. In this work, after feature extraction, optimization techniques are used and then it is classified with many classifiers for schizophrenia classification from EEG. The block diagram of the work is shown in Fig. 1. The paper is organized as follows. The materials and methods are discussed in section 2, followed by the usage of feature extraction techniques in section 3. It is followed by the optimization techniques for feature extraction in section 4. Section 5 explains the usage of classifiers for classification and section 6 gives the results and discussion and section 7 gives the conclusion.
II. MATERIALS AND METHODS
The EEG signals was collected from 14 patients suffering from schizophrenia and they were obtained from the Institute of Psychiatry and Neurology in Warsaw, Poland [12]. For the experiment, seven males and seven females with average ages of 27.9 + 3.3 years for male and 28.3 + 4.1 years for female was selected. 14 healthy subjects within the same group are also utilized for the study and they were also obtained from the Institute of Psychiatry and Neurology in Warsaw, Poland. When the participants remained with their eyes closed and in a relaxed state, at a sampling rate of 250 Hz, the EEG was collected for 15 minutes duration. The standard 10-20 International system was used to collect the data. The electrodes were kept at suitable positions and the acquired EEG signals was then divided into segments where the EEG signals are considered to be stationary.

A nineteen channel EEG signals are obtained per subject over the duration of fifteen minutes. Each channel of EEG signals totally consists of 2,25,000 samples which are divided into groups of five thousand sample segments. Therefore, totally each channel now represents the data of matrix [5000 × 45]. For pre-processing, Independent Component Analysis (ICA) was utilized. Then feature extraction is initiated using PLS based non linear regression, EM-PCA, and Isomap methods. The attained features matrix per method are in the form of [5000 × 10]. Then four types of optimization procedures like flower pollination optimization, eagle strategy using different evolution optimization, back tracking search optimization, and group search optimization is utilized to further extract a better represented feature column matrix as [5000 × 1]. This procedure is repeated for all the channels among the subjects.

III. FEATURE EXTRACTION
As the EEG signal is quite chaotic, to explain the most important features of the morphology of the EEG signals, nonlinear features are utilized. Instead of analyzing regular features like Detrend Fluctuation Analysis (DFA), Hurst exponent, Recurrence Quantification Analysis (RQA), Entropy, Fractal Dimension, Kolmogorov complexity, Hjorth component, Lempel-Ziv complexity, Auto Regressive (AR) Coefficients, Wavelet transform, Eigen vectors etc. [33], here in our paper PLS based Nonlinear regression features, EM-PCA features and Isomap features are extracted before optimization is proceeded. The feature extraction based on Isomap was implemented for applications like video manifold [34], semi-supervised local multi-manifold computation [35] and electromechanical equipment fault prediction [36]. The feature extraction based on EM-PCA was implemented for applications like face recognition [37] and classification in BCI [38]. The feature extraction based on PLS-NLR is useful in applications like development of Relevance Feature vector machine [39] and denoising application [40]. Owing to the versatility of these feature extraction techniques, it has been preferred in our works rather than the conventional techniques.

A. PLS BASED NON LINEAR REGRESSION FEATURES
Partial Least Square (PLS) method is quite advantageous to perform ordinary multiple linear regression as the collinearities in the predictor variables are assumed here which is nothing but the combination of the original input data in a linear manner [41]. Based on a covariance criteria, the input variable matrix is decomposed and it is mainly relied by the PLS. The latent variables or factors are found out by the PLS that are correlated with the output variables and descriptors of the input variables. Using the following linear equation, the concentration of the component is expressed as follows:

\[ Z = WA + G \]  \hspace{1cm} (1)

where \( W \) denotes an input matrix of wavelength signals, \( A \) represents a matrix containing regression coefficients and the bias vector is represented by \( G \). The matrix \( A \) has the following form:

\[ A = W^T V \left( V^T W W^T \right)^{-1} V^T Z \]  \hspace{1cm} (2)

where the linear combination of input and output variables are expressed by the latent variables \( T \) and \( V \).

By means of mapping the input data into a high dimensional space, the enhancement of the regression ability of a linear model can be done easily. A prediction in high-dimensional feature space is realized by the kernel technique without an explicit mapping of the original space. In the original space, the function of two elements is described by the kernel which is concerning to be the dot product of them in the feature space. For a linear algorithm, a kernel extension can be completed by means of replacement of the dot product calculation of elements. Therefore a non linear extension of PLS is kernel PLS. A non linear mapping \( \Phi : w \in R^N \rightarrow \Phi(w) \in G \) is utilized to transform the original data into a feature space. A non linear PLS is obtained for the original input data when the construction of a linear PLS regression is done. The calculation of the kernel gram matrix \( K \) can be done in the following form:

\[ K = \Phi \Phi^T \]  \hspace{1cm} (3)

The regression model with the component concentration is expressed as

\[ \hat{Z} = \Phi_i \Phi_i^T V \left( V^T K V \right)^{-1} V^T Z \]

\[ \hat{Z} = K_i V \left( V^T K V \right)^{-1} V^T Z \]  \hspace{1cm} (4)

where \( \hat{Z} \) and \( Z \) represent the output variables for both the validation set and the calibration set, the matrix representation of the validation variable feature space mapping is represented as \( \Phi_i \), \( K_i \) is the matrix comprising of \( K_{ij} = K (w_i, w_j) \), where for the validation set and the calibration set, \( w_i \) and \( w_j \) form as the input variables. When the kernel function is selected, the determination of the non linear regression is done.
The feature of high-dimensional space mapping is determined by the Kernel and the regression performance of it is affected. To develop a strong nonlinear regression model, different kernels are utilized generally such as Linear, Gaussian, Polynomial, Inverse Multi quadratic, semi local, exponential, rational, Knodel etc. In our work, Gaussian Kernel is used and is expressed as

\[ K(w_i, w_j) = \exp \left( -\frac{\|w_i - w_j\|^2}{2\sigma^2} \right) \]  

(5)

**B. EM-PCA FEATURES**

In this type of feature extraction, there are 4 substages: Derivation of EM step, orthogonalization step, data projection step and PCA Eigenspace generation step For preprocessing, estimation of mean vector \( \mu \) and mean subtraction to compute the difference between the input data and mean is done. From a large amount of high dimensional data, EM-PCA can easily extract a few Eigen vectors and Eigen values [42]. For a \( q \) dimensional variables, the covariance structure can be captured with the operation less than \( q(q + 1)/2 \) dimensions.

(1) Derivation of EM Step:
Assume the Eigen vector matrix \( V \). To estimate the input parameters \( V \) for orthogonalization, this process is used. E-step and M-step are the two important steps used here. Unless the differences between the variance is less than the threshold value or is equal to the threshold value, the EM step is carried out repeatedly. The eigen vectors will be selected as the input in a random manner before entering the step and is expressed as;

\[ \begin{align*}
Z &= \left( V^TV \right)^{-1} V^T \hat{Y} \\
V &= \tilde{Y}Z^T \left( ZZ^T \right)^{-1}
\end{align*} \]

(6)

(2) Orthogonalization:
The Gram-Schmidt orthonormalization process is carried out here. The first vector is normalized initially and then the remaining vectors is transformed into weighted normalized vectors after subsequent iterations.

(3) Projection of Data: Into the transpose of the M-step matrix, the input vector is projected and manipulated so that it can be easily multiplied to the result obtained from mean subtraction step.

(4) PCA eigen space generation: Using a standard PCA, the final stage is done here. Initially the estimation and subtraction of mean is done then the covariance and Singular Value Decomposition (SVD) is computed. Finally the edge values are selected which gives the reduced feature values. Depending on the selection method of the number of eigen vectors, the degree of accuracy of EM-PCA depends. The algorithm for the extraction of EM-PCA features is expressed in Algorithm 1.

---

**Algorithm 1 EM-PCA Feature Extraction**

Input: Matrix \( y \)
Output: Matrix Eigenvectors \( (V) \), eigen values \( (\lambda) \)

1. Mean vector estimation: \( \mu = \left( \frac{1}{N} \right) \sum_{i=1}^{N} y_i \)
2. The input data is centered around the mean and is represented as \( \hat{Y} = Y - \mu 1_N \)
3. To random values, the elements of \( V \in \Re^{M \times k} \) is set.
4. Repeat
5. E step: \( Z = (V^TV)^{-1} V^T \hat{Y} \)
6. M step: \( V = \tilde{Y}Z^T \left( ZZ^T \right)^{-1} \)
7. Until the change of \( \sigma^2 = \left( \frac{1}{(N-k)} \right) \left( \text{VAR}(\hat{Y}) - \text{VAR}(Z) \right) \leq \text{threshold} \)
8. \( V \) is orthogonalized
9. The input data is projected on \( V : Z = V^T \hat{Y} \)
10. On \( Z \), the PCA is performed. Obtain \( V' \) and \( \lambda' \)
11. Rotate \( V \) for \( V' : V' = VV' \)
12. The eigen values \( \lambda = \lambda' \) are determined.

---

**C. ISOMAP FEATURES**

When the data points are present close to a low dimensional non linear manifold embedded in a high dimensional space, and linear approximation cannot be used to adequately represent the non linear structure, the standard Multidimensional Scaling (MDS) cannot be used. Recovering the low-dimensional structure of a non linear manifold is quite difficult with MDS. Therefore, Isomap has been designed to discover the structure of high dimensional data and then trace its embedding in a low-dimensional Euclidean space making it as a class of non linear embedding schemes. In Isomap, instead of considering the Euclidean distance, the geodesic distances between the points are extracted [43]. By constructing a sparse graph, the computation of the geodesic distances are done in which every node is connected to its closest neighbours. Between each pair of node, the geodesic distance is considered to be the shortest path length in the graph so that the connections would be more feasible. Then to the classical MDS, these approximate geodesic distances are then utilized as input. The algorithm for the extraction of features through Isomap is expressed in Algorithm 2.

**D. PARAMETRIC ANALYSIS ON EXTRACTED FEATURES**

Fig. 2 explains the Histogram plot for average EM-PCA features in normal cases and Fig. 3 explains the histogram plot for average EM-PCA features among the schizophrenia cases.

From the Fig. 2 it is observed that the histogram plots of the EM-PCA extracted features for normal case shows the Gaussian normal behavior for the underlying EEG signals.

It is observed from the Fig. 3 that the histogram occupies the entire region with ups and downs as well as outlier
**Algorithm 2** Isomap Feature Extraction

Step 1: Neighborhood graph construction: On the manifold $M$, the points which are neighbours to each other is determined. Within a fixed radius $\varepsilon$ (Isomap), two simple techniques are used to connect every point to all the points or to all its $K$ nearest neighbours (K Isomap). Over the data point, the representation of these neighborhood relationships are done as a weighted graph $W$ with respective weight edges $e_y(i, j)$ between the neighboring points.

Step 2: Shortest path computation: On the manifold $M$, the geodesic distances $e_M(i, j)$ is estimated by Isomap between all pairs of points by means of computation of their respective shortest path distance $e_W(i, j)$ in the graph $W$.

Step 3: Construction of $d$-dimensional embedding: The classical MDS is applied to the matrix of graph distance $E_W = e_W(i, j)$ by means of construction of an embedding data in a low-dimensional Euclidean space $Z$ so that estimated intrinsic geometry of the manifold is preserved. K-Isomap is used here and so KNN algorithm is used in step 1 of this algorithm 2.

**FIGURE 2.** Histogram for EM-PCA features in normal case.

**FIGURE 3.** Histogram for EM-PCA features in schizophrenia case.

**FIGURE 4.** CDF plot for Isomap features for normal case.

**FIGURE 5.** CDF plot for Isomap features for schizophrenia case.

The statistical parameters such as mean, variance, skewness, kurtosis, geometric mean, harmonic mean, sample entropy and approximate entropy with different features.
for schizophrenia cases and normal cases are expressed in Table 1. In addition to that, the Canonical Correlation Analysis (CCA) with different features for schizophrenia and normal cases are expressed in Table 2.

The significance of the features like PLS non linear regression, EM-PCA, and Isomap among the schizophrenia and normal groups can be analyzed for non linearity and overlapping through the extraction of statistical features like mean, variance, skewness, kurtosis, Geometric mean, Harmonic mean and non linear features such as Sample entropy, Approximate entropy, Renyi entropy, Fuzzy entropy, Shannon entropy and Singular Value Decomposition (SVD). Table 1 demonstrates the average of parameters at different features for Schizophrenia and Normal Cases. It is observed from the Table 1 that all statistical parameters and entropies among the schizophrenia and normal cases are numerically overlapped and also exhibits the presence of non linearity as projected by higher values of variance and kurtosis. The sample entropy and approximate entropy indicates the peaked and trough regions of the features among the schizophrenia and normal groups. Hence it is worth to investigate on the correlation property of the features among the classes. One such
type of analysis is Canonical Correlation Analysis (CCA) which will act as a benchmark parameter for the features among the classes.

Table 2 shows the CCA of different features like PLS non-linear regression, EM-PCA and Isomap for Schizophrenia and Normal Cases. As observed in the Table 2, the average value of CCA is very low and therefore CCA exhibits the existence of no correlation among the features in the two classes. Hence it is wise to perform heuristic optimization techniques on the extracted numerical, non-linearly overlapped uncorrelated features to attain the segmentation of the features among the classes. Since EEG signals are complex and non-linear in nature, these signals will exhibit erratic and random response which can be conceptualized using chaotic theory. Chaos is a field of study which is well known as nonlinear dynamics. A nonlinear system is represented by the nonlinear time domain equations comprising the dynamic property of the variables in their non-linear form. Correlation Dimension (CD) is a famous study in chaos theory and it is calculated for EEG signals in this study. The Table 3 shows the average correlation dimension values for extracted features among the normal and schizophrenia cases. It is evident from the Table 3 that the correlation dimension values for both the classes are separated. The higher value of CD is due to the presence of randomness in the features. As for the Schizophrenia cases the low value of CD is exhibited by the existence of dampness in the features.

As reported in the Tables 1 and 3, the extracted features are uncorrelated as shown in the results attained from Correlation Dimension values but the features are non-linear and overlapped. The dynamics of the features are explored by calculating the entropy values and the Singular Value Decomposition (SVD) values. These calculated components indicate that the features are in need of optimization techniques for further processing. Therefore, a simple threshold may not be useful in the classification.

IV. NATURE INSPIRED OPTIMIZATION ALGORITHMS

The extracted features are then optimized through nature inspired optimization algorithms. The main purpose of any optimization algorithm is to search and locate the best solution to a particular problem. With the help of various agents, the search process is done generally which in turn evolves iteratively depending on certain rules of the algorithm involved. Therefore, in our work, four optimization algorithms are used specifically. The degrees of exploration with which the members of the group or search space can move across can be ascertained well with these four algorithms. In engineering, academia and industry, multi-objective optimization is quite challenging to solve and therefore sophisticated techniques were implemented to tackle it in an efficient manner. Applications of these four optimization algorithms have been used in various applications like engineering, medicine, astronomy, banking, financial risk management etc [44], but this is a new attempt to use them for schizophrenia EEG signal classification.

A. FLOWER POLLINATION ALGORITHM

This algorithm is developed and inspired from the flower pollination process of the flowering plants and it is usually extracted to multi-objective optimization [45]. The four rules are utilized in this algorithm for simplicity reasons.

1) Rule 1: The process of global pollination includes biotic and cross-pollination, and pollen carrying pollinators progress in a way that obeys Levy flights.
2) Rule 2: Self pollination and abiotic pollination are used for local pollination.
3) Rule 3: Flower constancy are developed by pollinators such as insects which is equivalent to a reproductive probability. The reproductive probability is proportional to the similarity of the two involved flowers here.
4) Rule 4: With the help of a switch probability \( p \in [0, 1] \), the controlling of interaction between the local pollination and global pollination can be done easily.

Pollinators such as insects carry the flower pollen gametes in the global pollination step and as a result the pollen can travel over very long distances because the insects can fly and cover a much longer distance range. Therefore, the mathematical representation such as rule 1 and rule 3 is expressed as

\[
\begin{align*}
\tilde{z}_j^{t+1} &= \tilde{z}_j^t + \gamma L(\lambda) \left( h_* - \tilde{z}_j^t \right) \\
\end{align*}
\]

where \( \tilde{z}_j^t \) is the pollen \( j \) or the solution vector \( z_j \) at iteration \( t \) and \( h_* \) is the current best solution found among all the solutions at the current iteration. To control the step size, the scaling factor \( \gamma \) is used. In reality the step size parameters is \( L(\lambda) \), to be more specific it is the Levy flight based step

| S.No | Feature Extraction Method | Normal Cases | Schizophrenia Cases |
|------|--------------------------|--------------|---------------------|
| 1    | PLS NON-LINEAR REGRESSION | 8.251        | 1.3479              |
| 2    | EM-PCA                   | 6.43         | 1.54                |
| 3    | ISOMAP                   | 7.52         | 1.962               |

TABLE 3. Average correlation dimension values.
size that indicates the pollination strength. With different distance steps, a long distance is covered by the insects and a Levy flight is utilized to mimic this specific characteristic more efficiently. Therefore, \( L > 0 \) is drawn from a Levy distribution and is represented as

\[
L \sim \frac{\lambda \Gamma (\lambda)}{\pi} \frac{1}{s^{1+\lambda}}, \quad (s >> s_0 > 0)
\]

The standard gamma function is denoted by \( \Gamma (\lambda) \) and for large steps, \( s > 0 \), this distribution is valid. Theoretically, it is required that \( |s_0| \gg 0 \), but in particular it can be as small as 0.1. The generation of the pseudo random step size is not important to object the Levy distribution. Generally, the Mantegna algorithm is utilized for drawing such random numbers by using Gaussian distributions \( X \) and \( Y \) by the following transfer function:

\[
s = \frac{X}{|Y|^{1/\lambda}} \quad \text{and} \quad X \sim N (0, \sigma^2) \quad \text{and} \quad Y \sim N (0, 1)
\]

Here \( X \sim (0, \sigma^2) \) implies that the samples are obtained from a Gaussian normal distribution with zero mean and a \( \sigma^2 \) variance. The representation of rule 2 and 3 for local pollination is done as

\[
z_j^{t+1} = z_j^t + \varepsilon (z_k^t - z_j^t)
\]

where \( z_k^t \) and \( z_j^t \) are the pollen obtained from different flowers of the same plant species. The flower constancy is easily mimicked in a limited neighborhood easily. If \( z_k^t \) and \( z_j^t \) come from the same species, this equality can become a local random walk if \( \varepsilon \) is drawn from a uniform distribution in \([0, 1]\).

The activities of the flower pollination can occur in both local and global scales. The flowers which are adjacent to each other in the not so far away neighborhood have high chances to be pollinated by local flower pollen than those far away. The switch probability (rule 4) is used to mimic this feature and can be used to switch between the global pollination to local pollination. Here \( p = 0.8 \) is used as an initial value. The procedure of the flower pollination algorithm is expressed in pseudocode 1.

### B. Eagle Strategy Using Different Evolution Algorithm

One of the famous metaheuristic strategy for optimization is Eagle strategy [46]. A combination of intensive local search and crude global search is used by the Eagle strategy which employs varies algorithms to match different applications. Utilizing a Levy flight random walk, the global search space is explored initially by this strategy. Once a promising answer is found out, then utilizing a local optimizer such as hill climbing or DE, an intensive local search is employed. Here in this work, DE is used. Then in a new region again this two stage method is initiated which starts with the exploration of a new global space followed by a local search. A balanced and good tradeoff between global search and fast local search can be obtained for different search stages and therefore any algorithm of our choice can be utilized here. Therefore to produce good results, the combination of various algorithms is efficiently utilized here. The switch between the local and the global search is controlled by the only parameter \( s_0 \). It serves the dual purpose of exploration and exploitation. It is a simple strategy or method and not an algorithm. To explore the search space in a much more effective and diverse manner, the algorithm utilized for the global exploration should have more randomness. As the system converges, the speed also rises. To utilize the local exploitation in an intensive manner, an efficient local optimizer is used. With a less number of functions, the main intention is to reach the local optimality as fast as possible. The procedure of the eagle strategy using different evolution is expressed in pseudocode 2.

### C. Backtracking Search Optimization Algorithm

The backtracking is one of the stochastic search technique [47] and is shown in Pseudocode 3. For solving different optimization problems, Backtracking search algorithm is used as it has a single control parameter and a very simple structure. It is a population-based technique and it has a memory where population from previous generations are present which helps in the analysis and generation of search direction matrix. Three basic genetic operators are present in this bio-inspired technique such as mutation, cross-over and selection.

---

**Pseudocode 1 Flower Pollination Algorithm**

Objective min or max \( f(z), z = (z_1, z_2, \ldots, z_d) \)

Initialize a population of \( q \) flowers with random solutions

Find the best solution \( h_s \) in the initial population

Define a switch probability \( p \in [0, 1] \)

While \( (t < \text{MaxGeneration}) \)

For \( j = 1 : q \) (all \( q \) flowers in the population)

If rand < \( p \)

Draw a (d-dimensional) step vector \( L \) from a Levy distribution

Global pollination via \( z_j^{t+1} = z_j^t + \gamma L (h_s - z_j^t) \)

Else

Draw \( \varepsilon \) from a uniform distribution in \([0, 1]\)

Do local pollination via \( z_j^{t+1} = z_j^t + \varepsilon (z_k^t - z_j^t) \)

End if

Evaluate new solutions

If new solutions are better, update them in the population

End for

Find the current best solution \( h_s \)

End while

Output the best solution found.
Pseudocode 2 Eagle Strategy Using Different Evolution Algorithm

| Begin |
|-------|
| Objective function $f_1(y), \ldots, f_N(y)$ |
| Population Initialization |
| While (stop criterion) |
| Global exploration by randomization |
| Objective evaluation to find promising solutions |
| If $s_r < \text{rand}$ |
| Switch to a local and intensive search |
| Efficient optimizer engagement to find the promising solutions by local search |
| If (better solutions are found), update the current best; end |
| End |
| End |

Pseudocode 3 Backtracking Search Optimization Algorithm

| Begin |
|-------|
| Initiate both population $A$ and historical population $oldA$ |
| While (stop condition is not met) |
| Perform the first type selection: $oldA = A$ in case $x < y$, where $x$ and $y$ are drawn uniformly in the range of 0 to 1. |
| Compute arbitrary changes in the position of $oldA$ |
| Generate the mutant |
| Generate the population $Z$ |
| Perform second type of selection: Population selection with best fitness from $Z$ and $A$ |
| Update the best solution |
| End while |
| Output the best solution |

Pseudocode 4 Group Search Optimizer Algorithm

| Begin |
|-------|
| Initialization |
| Evaluate the fitness of the current member |
| Find a producer |
| Producing |
| Choose 3 points randomly based on equation 13, 14, 15 |
| Best resource searching is done based on (16) |
| If best resource is not found, then stay in the current position only |
| If producer cant search a better area, then change the angle by (17) |
| Scrounging |
| Choose some of the group members randomly on the scrounger |
| Dispersion |
| Head angle generation using (16) |
| Obtain a random distance based on (18) and then move to a new point using (19) |
| Search for a better solution |
| Final optimum solution is found out |
| End |

A random mutation strategy is employed in this BSA that uses one direction individual for every target individual and is expressed as

$$\text{Mutant} = A + C \oplus (oldA - A)$$  \hspace{1cm} (12)

where $A$ represents the current population and $oldA$ represent the historical population. $C$ is a coefficient which helps to control the amplitude of the search direction matrix $(oldA - A)$. A complex cross over strategy which is non uniform in nature is used by BSA. In the cross over process, there are basically 2 steps, initially, a binary integer-values matrix of size $M \times Q$ is generated to specify the mutant individual which is to be explicitly manipulated by the usage of relevant individual. Here $M$ represents the population size and $Q$ represents the problem dimensions. Secondly, the relevant individuals are used to update the relevant dimensions of the mutant individuals. For BSA, there are two kinds of selection operator. For calculating the search direction, the historical population is selected which is employed by the first type selection operator. The current population is replaced by the historical population when the random number is comparatively smaller. The other type of selective operator is to determine the best individuals to get inside the next generation.

In Group search optimizer, there are 3 main evolution operations such as producing, scrounging and dispersion [48]. The procedure is given in the Pseudocode 4. In this algorithm, $A^m_j$ is the $j^{th}$ member at the $m^{th}$ iteration, $\psi^m_i$ is a head angle and $D^m_p$ is a unit vector. The scanning of three points expressed in the following equations are scanned by the producing operator as

$$A_z = A^m_p + r_1 q_{\max} D^m_p \left( \psi^m \right),$$  \hspace{1cm} (13)

$$A_r = A^m_p + r_1 q_{\max} D^m_p \left( \psi^m + \frac{r_2 \theta_{\max}}{2} \right)$$  \hspace{1cm} (14)

$$A_q = A^m_p + r_1 q_{\max} D^m_p \left( \psi^m + \frac{-r_2 \theta_{\max}}{2} \right)$$  \hspace{1cm} (15)

where the maximum search angle is expressed as $\theta_{\max}$; the maximum pursuit distance is experienced by $q_{\max}$. Under the normal distribution, $r_1$ and $r_2$ are the two parameters generated. $r_1$ and $r_2$ are the real numbers in the interval [0, 1]. If the changing of the current point is done, then the angle is represented as follows

$$\psi^{m+1} = \psi^m + r_2 \alpha_{\max}$$  \hspace{1cm} (16)

where $\alpha_{\max}$ represents the maximum turning angle. Otherwise the angle is fixed and is represented as

$$\psi^{m+b} = \psi^m,$$  \hspace{1cm} (17)

where $b$ is a constant. Around 75% of candidates for the remaining of the population is selected in the scrounger operator and in the dispersion operator, random walks are
utilized. The random distance is expressed as

$$ q_{ij} = b_i r_j q_{\text{max}} $$

(18)

The expression for the new point is written as

$$ A_j^{m+1} = A_j^{m} + q_j D_j^{m} \left( \varphi^{m+1} \right) $$

(19)

Algorithm 3 Real Adaboost Algorithm

1. The initial weights $w_{l,1} = \frac{1}{N}$, where $l = 1, 2, \ldots, N$

2. The following functions are performed for $t = 1, 2, \ldots, T$

   (a) To divide the training set $P$ into $Q$ partitions, based on the weighted instances, a weak classification is trained. A partition represents each leaf of the CART. For a partition $P_k^t$, where $k \in \{1, 2, \ldots, Q\}$, $W_{t+}^k$ and $W_{t-}^k$ are computed as follows:

$$ W_{t+}^k = \sum_{l: x_l \in P_t^k \land y_l=1} w_{l,t} $$

$$ W_{t-}^k = \sum_{l: x_l \in P_t^k \land y_l=-1} w_{l,t} $$

(21)

(b) The weak hypothesis for each partition $P_k^t$ is computed as

$$ f_t^k(x) = 0.5 \times \ln \left( \frac{W_{t+}^k}{W_{t-}^k} \right) $$

(22)

For every training instance $x_l$, the weak hypothesis $f_t(x_l)$ be $f_t^k(x_l)$, where $k$ is the index of partition that $x_l$ falls into.

(c) Weight updation is expressed as

$$ w_{l,t+1} = \frac{w_{l,t} \exp \left[-y_l f_t(x_l)\right]}{Z_t} $$

$$ Z_t = \sum_l w_{l,t} \exp \left[-y_l f_t(x_l)\right] $$

(23)

3. Assume $F_T(x_l) = \sum_{t=1}^T f_t(x_l)$ and the strong classifier output is expressed as

$$ G(x_l) = \text{sign} \left[ F_T(x_l) \right] $$

(24)

The value of the weak hypothesis in Adaboost are $\pm 1$. In real Adaboost, they are real numbers. The absolute value of $f_t^k(x)$ indicates a prediction confidence.

Algorithm 4 Modest Adaboost Algorithm

1. The initial weights are set as $w_{l,1} = \frac{1}{N}$, where $l = 1, 2, \ldots, N$

2. The following tasks are performed, $t = 1, 2, \ldots, T$

   (a) Based on the weighted instances, the weak classifier is trained and then $W_{t+}^k$ and $W_{t-}^k$ are calculated for every partition $P_k^t$

   (b) The inverted weight $\tilde{w}_{l,t}$ is calculated as

$$ \tilde{w}_{l,t} = \frac{1 - w_{l,t}}{M_t} $$

(25)

where $M_t = \sum_l (1 - w_{l,t})$. Then $\tilde{W}_{t+}^k$ and $\tilde{W}_{t-}^k$ are computed for each partition $P_k^t$ as follows:

$$ \tilde{W}_{t+}^k = \sum_{l: x_l \in P_k^t \land y_l=1} \tilde{w}_{l,t} $$

$$ \tilde{W}_{t-}^k = \sum_{l: x_l \in P_k^t \land y_l=-1} \tilde{w}_{l,t} $$

(26)

For every partition $P_k^t$, the weak hypothesis is represented as

$$ f_t^k(x) = W_{t+}^k \left( 1 - \tilde{W}_{t+}^k \right) - W_{t-}^k \left( 1 - \tilde{W}_{t-}^k \right) $$

(27)

The weak hypothesis $f_t(x_l)$ equals $f_t^k(x_l)$ for the instance $x_l$.

(c) The weights are then updated as

$$ w_{l,t+1} = \frac{w_{l,t} \exp \left[-y_l f_t(x_l)\right]}{Z_t} $$

$$ Z_t = \sum_l w_{l,t} \exp \left[-y_l f_t(x_l)\right] $$

(28)

3. Assume $F_T(x_l) = \sum_{t=1}^T f_t(x_l)$; Output the strong classifier as

$$ G(x_l) = \text{sign} \left[ F_T(x_l) \right] $$

(29)

To mitigate the contribution of weak hypothesis, inverted distributions are used by Modified Adaboost classifier.

Table 4 shows the average parameters for Optimization Techniques Like Flower Pollination Optimization, Eagle Strategy Using Different Evolution Optimization, Back Tracking Search Optimization and Group Search Optimization with features such as PLS Non Linear Regression, EM-PCA And Isomap for Schizophrenia Cases. From the table 4 it is observed that the removal of overlapping among the features after the optimization is prevalent. But the presence of non linearity still makes the optimized feature to devise better segmentation by a chosen group of classifiers. Table 5 shows the average parameters for Optimization Techniques Like Flower Pollination Optimization, Eagle Strategy Using Different Evolution Optimization, Back Tracking Search Optimization and Group Search Optimization with features such as PLS Non Linear Regression, EM-PCA And Isomap for normal Cases. From the table 5 it is observed that the removal of overlapping among the features after the optimization is prevalent. But the presence of non linearity again makes the optimized features to devise better segmentation by a chosen group of classifiers.
Table 4 shows the CCA for various Optimization techniques with different features among Schizophrenia and Normal Cases. If the CCA value is greater than 0.5, then it exhibits correlation among the classes. As indicated from the Table 6, except for PLS non-linear Regression with Flower pollination optimization and Eagle strategy using different evolution optimization, all the other methods preserve the non-correlative behavior of the extracted features.

### V. CLASSIFIERS
The optimized values are fed to the classifiers for schizophrenia classification. The classifiers used here are Adaboost classifier and its variants followed by Naïve Bayesian classifier and its variants.

#### A. ADABOOST CLASSIFIER
One of the techniques to attain an accurate classification is through means of ensemble mode-based classification such as Adaboost [49]. The learner’s instability is greatly attributed to the effectiveness of the boosting. A base classifier is employed sequentially for boosting based on a weighted version of the training sample set. The weak learner is referred to as the base classifier. To solve complicated behavior classification problems, boosting a set of weak learners can be done easily. As a weak learner, decision tree is applied. In Adaboost, equal weights are assigned to all the training samples. According to the weak learning error rate obtained in the previous iterations, the weights in each epoch are updated.

#### B. REAL ADABOOST
A most generalized version of Adaboost is called real Adaboost [50] and is explained in Algorithm 3 is follows:

#### C. MODEST ADABOOST
To suppress the generalization error of gentle Adaboost, modest Adaboost [51] was proposed and explained in Algorithm 4.
TABLE 5. Average parameters at various optimization techniques with different features for normal cases.

| Features                  | Optimization Methods                              | Mean   | Variance | Skewness | Kurtosis | Geometric Mean | Harmonic Mean | Sample Entropy | Approximate Entropy |
|---------------------------|--------------------------------------------------|--------|----------|----------|----------|----------------|---------------|-----------------|---------------------|
| PLS NONLINEAR REGRESSION  | FLOWER POLLMINATION OPTIMIZATION                | 0.099997 | 5.18E-05 | 5.408129 | 323.7425 | 0.0998        | 0.0995        | 9.247           | 1.7914              |
|                           | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 1.384843 | 34.55423  | 13.01092 | 218.2228 | 0.6874        | 0.262843      | 12.2032         | 1.7745              |
|                           | BACK TRACKING SEARCH OPTIMIZATION                | 0.280593  | 0.000369  | -1.32822 | 1.560711 | 0.0999        | 0.279131      | 10.2576         | 1.7812              |
|                           | GROUP SEARCH OPTIMIZATION                        | 0.016335  | 0.002959  | 9.750569  | 115.8116 | 0.0063        | 0.004237      | 12.2348         | 1.4881              |
| EM-PCA                    | FLOWER POLLMINATION OPTIMIZATION                | 0.09948  | 4.68E-08 | -0.48794 | 5.039481 | 0.099947    | 0.099947      | 8.2443          | 1.7917              |
|                           | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 0.215869  | 0.00274  | 0.306292  | 0.164018 | 0.2057        | 0.202538      | 12.0359         | 1.7846              |
|                           | BACK TRACKING SEARCH OPTIMIZATION                | 0.295363  | 7.43E-07 | -1.24832 | 2.494249 | 0.295362    | 0.295361      | 8.2083          | 1.787               |
|                           | GROUP SEARCH OPTIMIZATION                        | 0.537815  | 0.013352  | 0.324593  | 0.208435 | 0.52643     | 0.512255      | 12.1389         | 1.6575              |
| ISOMAP                    | FLOWER POLLMINATION OPTIMIZATION                | 0.099801  | 1.01E-06 | -0.39885  | 5.606084 | 0.2954       | 0.099791      | 10.0368         | 1.7916              |
|                           | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 0.141217  | 2.64E-06 | -0.05437  | 7.290381 | 0.5252       | 0.141199      | 8.8624          | 1.7462              |
|                           | BACK TRACKING SEARCH OPTIMIZATION                | 0.170444  | 3.05E-05 | 0.306341  | 1.810543 | 0.1704       | 0.170266      | 10.5577         | 1.7546              |
|                           | GROUP SEARCH OPTIMIZATION                        | 0.687871  | 0.000731  | 2.246077  | 31.84489 | 0.6872       | 0.686857      | 11.2655         | 1.5512              |

D. NAÏVE BAYESIAN CLASSIFIER

In machine learning modalities, one of the simplest and easiest to implement is NBC [52]. Based on the Bayes theorem, with assumptions of good independence among the features, it is expressed as

$$P(h|y) = \frac{P(y|h)P(h)}{P(y)}$$

(30)

where $P(h|y)$ indicates the feature probability of the target class of a given feature. The prior probability of the class is expressed as $P(h)$ and the likelihood of the probability of feature class is expressed as $P(y|h)$. The prior probability of the feature is expressed as $P(y)$.

E. GAUSSIAN NAÏVE BAYESIAN CLASSIFIER

Assumption of a known a priori and then reducing the probability of classification error is the essential principle in Bayes technique [53]. From the available training data set, the estimation of the class-conditional density function can be done. The updation of the training set conditioned density function is done during Bayesian estimation which allows the conversion to a posteriori density from a priori information. For the two class patterns $h_1$ and $h_2$, the Bayesian rule is implemented as

$$P(h_j|y) = \frac{p(y|h_j)P(h_j)}{\sum_{j=1}^{2} p(y|h_j)P(h_j)}$$

(31)

The 2 Bayes classification rules are as follows:

- If $p(h_1|y) > p(h_2|y)$, $y$ is assigned to $h_1$
- If $p(h_1|y) < p(h_2|y)$, $y$ is assigned to $h_2$

With $\mu_j$ as the mean value and $\Sigma_j$ as the covariance matrix, for the Gaussian probability distribution function [52], it makes it more feasible for analysis and is expressed as

$$p(y|h_j) = \frac{1}{(2\pi)^{n/2}|\Sigma_j|^{1/2}}\exp\left(-\frac{1}{2}(y - \mu_j)^T\Sigma_j^{-1}(y - \mu_j)\right)$$

(32)

A monotonic logarithmic discriminant function is chosen and is analyzed as

$$x_j(y) = -\frac{1}{2}(y - \mu_j)^T\Sigma_j^{-1}(y - \mu_j) + \ln P(h_j)$$

$$-\frac{n}{2}\ln 2\pi - \frac{1}{2}\ln |\Sigma_j|$$

(33)

j = 1, 2
TABLE 6. CCA at various optimization techniques with different features for schizophrenia and normal cases.

| Sl.no | Features                  | Optimization Methods                          | CCA  |
|-------|---------------------------|-----------------------------------------------|------|
| 1     | PLS NON LINEAR REGRESSION | Flower pollination optimization               | 0.6586 |
|       |                           | Eagle strategy using different evolution optimization | 0.5104 |
|       |                           | Back tracking search optimization              | 0.3655 |
|       |                           | Group search optimization                     | 0.2829 |
| 2     | EM-PCA                    | Flower pollination optimization               | 0.24857 |
|       |                           | Eagle strategy using different evolution optimization | 0.2264 |
|       |                           | Back tracking search optimization              | 0.2185 |
|       |                           | Group search optimization                     | 0.1822 |
| 3     | ISOMAP                    | Flower pollination optimization               | 0.1647 |
|       |                           | Eagle strategy using different evolution optimization | 0.1423 |
|       |                           | Back tracking search optimization              | 0.10428 |
|       |                           | Group search optimization                     | 0.09652 |

TABLE 7. Consolidated result analysis of PLS non linear regression with flower pollination optimization for normal cases.

| Classifiers       | PC(%)  | MC(%)  | FA(%)  | PI(%)  | Sensitivity(%) | Specificity(%) | Accuracy(%) | GDR(%)  |
|-------------------|--------|--------|--------|--------|----------------|----------------|-------------|---------|
| Adaboost Classifier | 89.32875 | 0      | 10.67125 | 88.045 | 89.32875       | 100            | 94.66438    | 89.32875 |
| Real Adaboost Classifier | 96.3525 | 0      | 3.64   | 96.205 | 96.3525       | 100            | 98.17625    | 96.35973 |
| Modest Adaboost Classifier | 84.38  | 0      | 15.62  | 81.4325 | 84.38         | 100            | 92.19       | 84.38   |
| NBC               | 52.44719 | 0      | 47.55281 | 9.27793 | 52.44719       | 100            | 76.22359    | 52.44719 |
| Gaussian NBC      | 71.3575 | 0      | 28.6455 | 59.85  | 71.3575       | 100            | 85.67875    | 71.35536 |
| Optimized NBC     | 91.67  | 0      | 8.33   | 91.58  | 91.67          | 100            | 95.835      | 91.67   |

TABLE 8. Consolidated result analysis of Isomap with group search optimization for schizophrenia.

| Classifiers       | PC(%)  | MC(%)  | FA(%)  | PI(%)  | Sensitivity(%) | Specificity(%) | Accuracy(%) | GDR(%)  |
|-------------------|--------|--------|--------|--------|----------------|----------------|-------------|---------|
| Adaboost Classifier | 82.94813 | 0      | 17.05188 | 79.43063 | 82.94813       | 100            | 91.47406    | 82.94813 |
| Real Adaboost Classifier | 68.16425 | 0      | 31.83875 | 53.27363 | 68.16425       | 100            | 84.08213    | 68.16221 |
| Modest Adaboost Classifier | 57.94813 | 0      | 42.05188 | 27.39563 | 57.94813       | 100            | 78.97406    | 57.94813 |
| NBC               | 53.85625 | 0      | 46.14375 | 14.28141 | 53.85625       | 100            | 76.92813    | 53.85625 |
| Gaussian NBC      | 97.52  | 0      | 2.47   | 97.45  | 97.52          | 100            | 98.43       | 97.52   |
| Optimized NBC     | 51.75  | 0      | 48.25  | 6.7475 | 51.75          | 100            | 75.875      | 51.75   |

For each class from the training data, the mean vector and covariance matrices of the discriminant function are calculated, and a hyperplane is used to separate the data.

F. OPTIMIZED NAIVE BAYESIAN CLASSIFIER

The bag-of-token model was utilized to optimize the N Bayesian standard algorithm here [54]. Based on the
TABLE 9. Consolidated results of accuracy (%) among the classifiers at various optimization techniques with different features for normal cases.

| Sl.No | Features          | Optimization Methods                  | Adaboost Classifier | Real Adaboost Classifier | Modest Adaboost Classifier | NBC  | Gaussian NBC | Optimized NBC |
|-------|-------------------|---------------------------------------|---------------------|--------------------------|---------------------------|------|--------------|---------------|
| 1     | PLS NON LINEAR REgression | FLOWER POLLENATION OPTIMIZATION       | 94.66438            | 98.17625                 | 92.19                     | 76.22359 | 85.67875     | 95.835        |
|       |                   | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 95.055             | 89.6                     | 76.6075                   | 95.705 | 79.69        | 89.85625      |
|       |                   | BACK TRACKING SEARCH OPTIMIZATION     | 93.555             | 77.40656                 | 98.16                     | 97.915 | 86.52344     | 77.92906      |
|       |                   | GROUP SEARCH OPTIMIZATION             | 82.616             | 77.21063                 | 94.66438                  | 78.90875 | 90.88625     | 89.6          |
| 2     | EM-PCA            | FLOWER POLLENATION OPTIMIZATION       | 85.22406            | 94.01125                 | 91.1457                   | 79.43  | 76.8775      | 89.6          |
|       |                   | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 78.07602            | 94.795                   | 79.82                     | 90.36875 | 91.93        | 83.105        |
|       |                   | BACK TRACKING SEARCH OPTIMIZATION     | 77.95355            | 85.54938                 | 93.23                     | 77.66781 | 76.49781     | 95.835        |
|       |                   | GROUP SEARCH OPTIMIZATION             | 77.08              | 76.37125                 | 85.74344                  | 77.47188 | 87.5         | 88.55         |
| 3     | ISOMAP            | FLOWER POLLENATION OPTIMIZATION       | 76.675             | 92.97                    | 78.25563                  | 85.67875 | 85.22406     | 95.315        |
|       |                   | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 96.355             | 94.66438                 | 76.8775                   | 95.835 | 84.375       | 94.53375      |
|       |                   | BACK TRACKING SEARCH OPTIMIZATION     | 91.93              | 81.25                    | 98.77                     | 98.17625 | 83.594       | 79.69         |
|       |                   | GROUP SEARCH OPTIMIZATION             | 77.43105           | 79.01992                 | 82.942                    | 91.93  | 92.97        | 96.355        |

non-negative number of occurrences of token $q$ in the observation, the value of each feature $q$ is calculated. The estimated probability is expressed as

$$ p(\text{token} = q | \text{class} = m) = \frac{1 + \beta_1}{N + \beta_2} $$

(34)

where in class $m$, the weighted number of occurrences of token $q$ is expressed as $\beta_1$ and in class $m$, the total weighted numbers of occurrences of all token $q$ is expressed as $\beta_2$, the total number of instances in the training set is expressed as $N$.

Based on the estimated posterior probability, the class label for each observation is predicted by the classifier. Therefore, with the maximum posterior probability each observation is assigned to the class.

VI. RESULTS AND DISCUSSION

The performance metrics parameters computed in this work is done by means of calculating the Perfect Classification (PC), Missed Classification (MC) and False Alarm (FA) respectively. The overall performance of the classification system...
TABLE 10. Consolidated results of accuracy (%) among the classifiers at various optimization techniques with different features for schizophrenia cases.

| SL.No | Features                  | Optimization Methods                          | Adaboost Classifier | Real Adaboost Classifier | Modest Adaboost Classifier | NBC    | Gaussian NBC | Optimized NBC |
|-------|---------------------------|-----------------------------------------------|---------------------|--------------------------|----------------------------|--------|--------------|----------------|
| 1     | PLS NON LINEAR REGRESSION | FLOWER POLLINATION OPTIMIZATION               | 80.86               | 95.835                   | 78.125                     | 79.43  | 76           | 98.17625       |
|       |                            | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 95.575             | 95.315                   | 95.705                     | 90.625 | 83.594       | 79.82          |
|       |                            | BACK TRACKING SEARCH OPTIMIZATION             | 93.75               | 83.78925                 | 96.875                     | 78.33727 | 90.49688   | 78.38625       |
|       |                            | GROUP SEARCH OPTIMIZATION                     | 77.6025             | 77.21063                 | 87.5                       | 84.8975 | 79.43        | 85.42          |
| 2     | EM-PCA                    | FLOWER POLLINATION OPTIMIZATION               | 78.90875            | 93.75                    | 77.6025                    | 77.08  | 81.9         | 85.74344       |
|       |                            | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 93.75               | 95.055                   | 76                        | 79.82  | 89.85625     | 79.43          |
|       |                            | BACK TRACKING SEARCH OPTIMIZATION             | 89.85625            | 90.49688                 | 95.315                     | 87.10938 | 97.655      | 77.83109       |
|       |                            | GROUP SEARCH OPTIMIZATION                     | 76.135              | 76.675                   | 77.37391                   | 75.875 | 77.47188 | 77.40656       |
| 3     | ISOMAP                    | FLOWER POLLINATION OPTIMIZATION               | 90.36875            | 98.77                    | 93.23                      | 78.38625 | 86.71875   | 86.52344       |
|       |                            | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 82.942             | 94.53375                 | 76.2025                    | 97.915 | 77.8066     | 85.22406       |
|       |                            | BACK TRACKING SEARCH OPTIMIZATION             | 92.71               | 86.32813                 | 96.875                     | 78.28828 | 78.10051   | 87.10938       |
|       |                            | GROUP SEARCH OPTIMIZATION                     | 91.47406            | 84.08213                 | 78.97406                   | 76.92813 | 98.43       | 75.875         |

The overall incorrect classification of the classification model is known as classification error and is calculated as

\[
\text{ClassificationError} = \frac{MC + FA}{PC + MC + FA} \times 100 \quad (36)
\]

The implication of diagnostic test is positive and if the subject has disease, it is termed as the sensitivity or the true positive rate is expressed as

\[
\text{Sensitivity} = \frac{PC}{PC + FA} \times 100 \quad (37)
\]

The implication that the diagnostic test is negative, and the person is healthy is termed by specificity and is
expressed as

$$Specificity = \frac{PC}{PC + MC} \times 100$$  \hspace{1cm} (38)

The expression for Performance Index is written as follows

$$PI = \left(\frac{PC - MC - FA}{PC}\right) \times 100$$  \hspace{1cm} (39)

Good Detection Rate (GDR): The impact of a successful detection is represented by GDR and its mathematical criteria is expressed as

$$GDR = \left(\frac{PC - MC}{PC + FA}\right) \times 100$$  \hspace{1cm} (40)

Mathematically, the Mean Square Error (MSE) is expressed as follows:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (O_i - T_j)^2$$  \hspace{1cm} (41)

where $O_i$ indicates the observed value at a specific time, $T_j$ denotes the target value at model $j$, $j = 1$ to 19, and $N$ is the total number of observations per patient in our case, it is 5000. The feature extracted values are then fully utilized by the testing and training classifiers in this research. The classifier MSE values are drastically reduced to a least level as the progress of the training section was done in a quite regressive manner. The training of the classifiers was performed with a zero-training error of MSE. In this work, the type of cross-validation methodology utilized was K-fold. Initially, the dataset is divided into ‘k’ equal size points. For the training of the classifiers, $k - 1$ groups are utilized for their performance assessment in every step, the remaining step is utilized. The validation repetition is done for $k$ number of times. The assessment of the classifier computation performance is done based on the $k$ results. In our work, the value of $k$ is chosen to be 10. Therefore 90% of the data was used for training and 10% of the data was used for testing. The process was repeated over 10 times for every fold of the methodology. Before the selection of the new sets both for both testing and training for the next and upcoming cycles, the random division of all the instances in the training and testing groups for the current cycle is done over the whole dataset. Ultimately at the final stage of the 10-fold process, the average values of the performance metrics are computed.

---

### TABLE 11. Average performance measures among the classifiers at various optimization techniques with different features for normal case.

| SL.No | Features                | Optimization Method                           | Performance Index (%) | Accuracy (%) | GDR(%) |
|-------|-------------------------|----------------------------------------------|-----------------------|--------------|--------|
| 1     | PLS NON LINEAR REGRESSION | FLOWER POLLINATION OPTIMIZATION             | 71.06507              | 90.46133     | 80.92351 |
|       |                         | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 63.5274              | 87.75229     | 70.73015 |
|       |                         | BACK TRACKING SEARCH OPTIMIZATION            | 63.36323              | 88.71484     | 71.19161 |
|       |                         | GROUP SEARCH OPTIMIZATION                    | 55.75721              | 85.64767     | 71.07847 |
| 2     | EM-PCA                  | FLOWER POLLINATION OPTIMIZATION             | 57.62146              | 86.04839     | 71.7779 |
|       |                         | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 58.24828              | 86.34913     | 63.34932 |
|       |                         | BACK TRACKING SEARCH OPTIMIZATION            | 47.79561              | 84.45559     | 62.77599 |
|       |                         | GROUP SEARCH OPTIMIZATION                    | 40.52984              | 82.11943     | 64.23609 |
| 3     | ISOMAP                  | FLOWER POLLINATION OPTIMIZATION             | 54.49636              | 85.68641     | 69.14376 |
|       |                         | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 71.37531              | 90.44011     | 78.06744 |
|       |                         | BACK TRACKING SEARCH OPTIMIZATION            | 66.41113              | 89.02004     | 78.01747 |
|       |                         | GROUP SEARCH OPTIMIZATION                    | 58.01986              | 86.77466     | 72.64787 |

---

### TABLE 12. Average performance measures among the classifiers at various optimization techniques with different features for schizophrenia case.

| SL.No | Features                | Optimization Method                           | Performance Index (%) | Accuracy (%) | GDR(%) |
|-------|-------------------------|----------------------------------------------|-----------------------|--------------|--------|
| 1     | PLS NON LINEAR REGRESSION | FLOWER POLLINATION OPTIMIZATION             | 47.59531              | 84.73771     | 69.47662 |
|       |                         | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 72.02977              | 90.10567     | 79.35475 |
|       |                         | BACK TRACKING SEARCH OPTIMIZATION            | 59.25582              | 86.93911     | 67.61314 |
|       |                         | GROUP SEARCH OPTIMIZATION                    | 41.19979              | 82.61011     | 57.04772 |
| 2     | EM-PCA                  | FLOWER POLLINATION OPTIMIZATION             | 41.69625              | 82.49745     | 63.09959 |
|       |                         | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 53.86156              | 85.65188     | 71.00238 |
|       |                         | BACK TRACKING SEARCH OPTIMIZATION            | 71.02531              | 89.7106      | 71.09572 |
|       |                         | GROUP SEARCH OPTIMIZATION                    | 13.43427              | 76.82289     | 39.58316 |
| 3     | ISOMAP                  | FLOWER POLLINATION OPTIMIZATION             | 68.54854              | 89.11787     | 76.57693 |
|       |                         | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 53.13418              | 85.77065     | 68.56712 |
|       |                         | BACK TRACKING SEARCH OPTIMIZATION            | 58.17742              | 86.56855     | 60.40064 |
|       |                         | GROUP SEARCH OPTIMIZATION                    | 46.6773               | 84.4689      | 68.93745 |
| SL No | Features | Optimization Methods                        | Adaboost Classifier | Real Adaboost Classifier | Modest Adaboost Classifier | NBC | Gaussian NBC | Optimized NBC |
|-------|----------|-------------------------------------------|---------------------|--------------------------|---------------------------|-----|--------------|---------------|
| 1     | PLS NON LINEAR REGRESSION | FLOWER POLLINATION OPTIMIZATION | 88.045              | 96.205                   | 81.4325                   | 9.27793 | 59.85        | 91.58         |
|       |          | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 89.18              | 78.93                    | 12.00438                  | 91.18 | 31.4425      | 78.4275       |
|       |          | BACK TRACKING SEARCH OPTIMIZATION      | 85.18063            | 17.50375                 | 97.87                     | 95.65 | 63.04125     | 20.93375      |
|       |          | GROUP SEARCH OPTIMIZATION            | 46.66075            | 16.2175                  | 88.045                    | 26.9975 | 77.6925     | 78.93         |
| 2     | EM-PCA   | FLOWER POLLINATION OPTIMIZATION      | 58.02563            | 86.37                    | 78.465                    | 30.01625 | 13.92188     | 78.93         |
|       |          | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 21.89844            | 88.38                    | 32.15563                  | 77.4225 | 80.72125     | 48.91188      |
|       |          | BACK TRACKING SEARCH OPTIMIZATION    | 21.09453            | 59.34                    | 84.315                    | 19.21875 | 11.22539     | 91.58         |
|       |          | GROUP SEARCH OPTIMIZATION            | 15.36               | 10.32656                 | 60.105                    | 17.9325 | 66.66        | 72.795        |
| 3     | ISOMAP   | FLOWER POLLINATION OPTIMIZATION      | 12.48375            | 83.6225                  | 23.01625                  | 59.85 | 58.02563     | 89.98         |
|       |          | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 92.455              | 88.045                   | 13.92188                  | 91.58 | 54.54        | 87.71         |
|       |          | BACK TRACKING SEARCH OPTIMIZATION    | 80.72125            | 40                       | 98.35                     | 96.205 | 51.163       | 31.4425       |
|       |          | GROUP SEARCH OPTIMIZATION            | 17.66453            | 25.49438                 | 48.1615                   | 80.72125 | 83.6225     | 92.455        |

Table 7 shows the consolidated Result Analysis of PLS Non-Linear Regression with Flower Pollination Optimization for Normal Cases. Table 8 shows the consolidated result analysis of Isomap with Group search optimization for schizophrenia cases. Table 9 gives the average Accuracy (%) among the Classifiers at Various Optimization Techniques with different features for Normal Cases. Table 10 gives the average Accuracy (%) among the Classifiers at Various Optimization Techniques with different features for schizophrenia cases. Table 11 gives the average performance measures among the classifiers at various optimization techniques with different features for normal case. Table 12 gives the average performance measures among the classifiers at various optimization techniques with different features for Schizophrenia cases. Table 13 gives the average performance index (%) among the classifiers at various optimization techniques with different features for normal cases. Table 14 gives the average performance index (%) among the classifiers at various optimization techniques with different features for schizophrenia cases. The average error rate among the classifiers at various optimization techniques with different features for normal cases is given in Table 15. The average error rate among the classifiers at various optimization techniques with different features for schizophrenia cases is given in Table 16. The average performance of parameters among the classifiers at various optimization techniques with different features for normal cases is given in Table 17. The average performance of parameters among the classifiers at various optimization techniques with different features for schizophrenia cases is given in Table 18.
TABLE 14. Average performance index (%) among the classifiers at various optimization techniques with different features for schizophrenia cases.

| SLN  | Features               | Optimization Methods                              | Adaboost Classifier | Real Adaboost Classifier | Modest Adaboost Classifier | NBC | Gaussian NBC | Optimized NBC |
|------|------------------------|----------------------------------------------------|---------------------|--------------------------|---------------------------|-----|-------------|---------------|
| 1    | PLS NON LINEAR REGRESSION | FLOWER POLLINATION OPTIMIZATION                   | 37.86053            | 91.58                    | 22.22                     | 30.01625 | 7.69        | 96.205        |
|      |                        | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 90.78               | 89.98                    | 91.18                     | 76.92 | 51.163      | 32.15563      |
|      |                        | BACK TRACKING SEARCH OPTIMIZATION                  | 85.7                | 52.00725                 | 93.33                     | 23.51391 | 77.17125    | 23.8125       |
|      |                        | GROUP SEARCH OPTIMIZATION                          | 18.79               | 16.2175                  | 66.66                     | 56.685 | 30.01625    | 58.83         |
| 2    | EM-PCA                 | FLOWER POLLINATION OPTIMIZATION                   | 26.9975             | 85.7                     | 18.79                     | 15.36 | 43.225      | 60.105        |
|      |                        | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 85.7                | 89.18                    | 7.69                      | 32.15563 | 78.4275     | 30.01625      |
|      |                        | BACK TRACKING SEARCH OPTIMIZATION                  | 78.4275             | 77.17125                 | 89.98                     | 65.2125 | 95.07       | 20.29063      |
|      |                        | GROUP SEARCH OPTIMIZATION                          | 8.64875             | 12.48375                 | 17.28938                  | 6.7475 | 17.93725    | 17.50375      |
| 3    | ISOMAP                 | FLOWER POLLINATION OPTIMIZATION                   | 77.4225             | 98.35                    | 84.315                    | 23.8125 | 63.765      | 63.04125      |
|      |                        | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 48.1615             | 87.71                    | 9.128125                  | 95.65  | 20.12984    | 58.02563      |
|      |                        | BACK TRACKING SEARCH OPTIMIZATION                  | 82.95               | 62.3175                  | 93.33                     | 23.21531 | 22.05922    | 65.2125       |
|      |                        | GROUP SEARCH OPTIMIZATION                          | 79.43063            | 53.27363                 | 27.39563                  | 14.28141 | 98.935      | 6.7475        |

Schizophrenia cases is tabulated in Table 8. Since the optimization is done through the nature inspired algorithms and the classifier’s output is plugged with False alarm in the range as low of 1.04% to higher value of 48.25%. Gaussian NBC Classifier attains higher parametric values like accuracy of 98.77%, GDR of 97.52% and PI of 97.45%. The Optimized NBC Classifier is ebbed to the lower parametric values like accuracy of 75.87%, GDR of 51.75% and PI of 6.747% and this lower performance indicates the presence of False alarm of 48.25%. As indicated in the Table 8 that all the classifiers have nil missed classification in this optimization technique and this makes the classifiers with low sensitivity and low threshold, and with high specificity one.

Table 9 depicts the consolidated results of Accuracy (%) among the Classifiers at Various Optimization Techniques with different features for normal cases. It is observed from the Table 5 that the optimized NBC classifier attains higher accuracy of 98.176% in PLS non linear regression feature extraction with flower pollination optimization method and the low accuracy of 76.22359% for NBC Classifier. Likewise for EM-PCA feature extraction with back tracking search optimization the classifier optimized NBC attained higher accuracy of 95.83% and low accuracy of 76.87% is reached for Gaussian NBC classifier in EM-PCA feature extraction with flower pollination optimization method. As in the case of Isomap feature Extraction Modest Adaboost classifier with back tracking search optimization attained higher accuracy of 98.77% and low accuracy value of 76.67% is reached in Adaboost classifier with flower pollination optimization method.

Table 10 depicts the Consolidated results of Accuracy (%) among the Classifiers at Various Optimization Techniques with different features for Schizophrenia Cases. It is observed from the Table 5 that the optimized NBC classifier attains higher accuracy of 98.176% in PLS non linear regression feature extraction with flower pollination optimization method and the low accuracy of 76% for Gaussian NBC Classifier. Likewise for EM-PCA feature extraction with back tracking search optimization the classifier Gaussian NBC attained higher accuracy of 97.65% and low accuracy of 76% is reached for Modest Adaboost classifier in EM-PCA feature extraction with eagle strategy using different evolution optimization method. As in the case of Isomap feature Extraction Real Adaboost classifier with flower pollination optimization method attained higher accuracy of 98.77% and low accuracy value of 75.87% is reached in Optimized NBC with Group search optimization method.

To study the effect of feature extraction and optimization techniques, irrespective of individual classifier...
TABLE 15. Average error rate (%) among the classifiers at various optimization techniques with different features for normal cases.

| Sl No | Features                  | Optimization Methods                                      | Adaboost Classifier | Real Adaboost Classifier | Modest Adaboost Classifier | NBC | Gaussian NBC | Optimized NBC |
|-------|---------------------------|-----------------------------------------------------------|---------------------|--------------------------|----------------------------|-----|--------------|---------------|
| 1     | PLS NON LINEAR REGRESSION | FLOWER POLLINATION OPTIMIZATION                          | 10.67125            | 3.640273                 | 15.62                      | 47.55281 | 28.646464   | 8.33          |
|       |                           | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION     | 9.89                | 20.83                    | 46.785                     | 8.59 | 40.62        | 20.31         |
|       |                           | BACK TRACKING SEARCH OPTIMIZATION                         | 12.89               | 45.18688                 | 2.08                       | 4.166167 | 26.95313    | 44.14188      |
|       |                           | GROUP SEARCH OPTIMIZATION                                 | 34.76441            | 45.57875                 | 10.67125                  | 42.1825 | 18.2275      | 20.83         |
| 2     | EM-PCA                    | FLOWER POLLINATION OPTIMIZATION                          | 29.55531            | 11.9775                  | 17.705                     | 41.14 | 46.245       | 20.83         |
|       |                           | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION     | 43.84797            | 10.41                    | 40.36                      | 19.27 | 16.14        | 33.79083      |
|       |                           | BACK TRACKING SEARCH OPTIMIZATION                         | 44.09289            | 28.90445                 | 13.54                      | 44.66438 | 47.00438    | 8.33          |
|       |                           | GROUP SEARCH OPTIMIZATION                                 | 45.84               | 47.2575                  | 28.51473                  | 45.05625 | 25          | 22.915        |
| 3     | ISOMAP                    | FLOWER POLLINATION OPTIMIZATION                          | 46.65               | 14.06                    | 43.48875                  | 28.64464 | 29.55531    | 9.37          |
|       |                           | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION     | 7.29                | 10.67125                 | 46.245                     | 8.33 | 31.25        | 10.9325       |
|       |                           | BACK TRACKING SEARCH OPTIMIZATION                         | 16.14               | 37.5                     | 1.04                       | 3.640273 | 32.81737    | 40.62         |
|       |                           | GROUP SEARCH OPTIMIZATION                                 | 45.13789            | 41.96297                 | 34.11534                  | 16.14 | 14.06        | 7.29          |

Performance an averaging of the parameters like accuracy and GDR is initiated among the six classifiers. The same is shown in Table 11 and Table 12. From Table 11, it is understood that a highest classification accuracy of 90.46% and a GDR of 80.92% is obtained if PLS non linear regression features are optimized with flower pollination algorithm for the normal cases. And also the lowest accuracy of 82.11% and GDR of 64.23% is attained in the EM-PCA feature extraction with group search optimization method. From Table 12, it is understood that a highest classification accuracy of 90.105% and a GDR of 79.35% is obtained if PLS non linear regression features are optimized with eagle strategy using different evolution optimization for schizophrenia cases. As shown in Table 12, EM-PCA feature extraction with group search optimization method reached the low value of accuracy 76.82% and GDR of 39.58% for schizophrenia cases. This is due to the accumulation of more false positive in the schizophrenia case.

The parameter, Performance Index (PI) indicates the influence of FN and FP in the classifiers. Table 13 and Table 14 identifies the better feature extraction, optimization methods and classifiers for normal and schizophrenia cases. From Table 13, it is analyzed for normal cases that for PLS non linear regression features and back tracking optimization algorithm, a PI of 97.87% is obtained if classified with Modest Adaboost classifier. Similarly when EM-PCA features and back tracking search optimization techniques are classified with Optimized NBC, a high PI of 91.58% is obtained. Also when Isomap features and back tracking search optimization techniques are classified with Modest Adaboost classifier, a high PI of 98.35% is obtained. This is the highest PI value for the normal cases.

From Table 14, it is analyzed for schizophrenia cases that for PLS non linear regression features and flower pollination algorithm, a PI of 96.205% is obtained if classified with Optimized NBC. Similarly when EM-PCA features and Back tracking optimization algorithm are classified with Gaussian NBC classifier, a high PI of 95.07% is obtained. As in EM-PCA feature extraction group search optimization fails to maintain good average PI among the all six classifiers. Also when Isomap features and Flower pollination algorithm are considered, then a high PI of 98.35% is obtained when classified with Real Adaboost classifier and is a higher PI value.

From Table 15, it is analyzed for normal cases that a least error rate of 2.08% is obtained when PLS non linear
TABLE 16. Average error rate (%) among the classifiers at various optimization techniques with different features for schizophrenia cases.

| Sl. No | Features and Optimization Methods | Adabost Classifier | Real Adabost Classifier | Modest Adabost Classifier | NBC | Gaussian NBC | Optimized NBC |
|--------|-----------------------------------|--------------------|-------------------------|---------------------------|-----|-------------|--------------|
| 1      | PLS NON LINEAR REGRESSION FLOWER POLLINATION OPTIMIZATION | 38.28              | 8.33                    | 43.75                     | 41.14 | 48          | 3.640273     |
|        | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 8.85               | 9.37                    | 8.59                      | 18.75 | 32.81737    | 40.36        |
|        | BACK TRACKING SEARCH OPTIMIZATION | 12.5               | 32.42555                | 6.25                      | 43.32547 | 19.01      | 43.2275      |
|        | GROUP SEARCH OPTIMIZATION | 44.795             | 45.57875                | 25                        | 30.20709 | 41.14      | 29.16425     |
| 2      | EM-PCA FLOWER POLLINATION OPTIMIZATION | 42.1825            | 12.5                    | 44.795                    | 45.84 | 36.19601    | 28.51474     |
|        | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 12.5               | 9.89                    | 48                        | 40.36 | 20.31       | 41.14        |
|        | BACK TRACKING SEARCH OPTIMIZATION | 20.31              | 19.01                   | 9.37                      | 25.78125 | 4.682851   | 44.33781     |
|        | GROUP SEARCH OPTIMIZATION | 47.73              | 46.65                   | 45.25219                  | 48.25 | 45.05625    | 45.18688     |
| 3      | ISOMAP FLOWER POLLINATION OPTIMIZATION | 19.27              | 1.04                    | 13.54                     | 43.2275 | 26.5625    | 26.95313     |
|        | EAGLE STRATEGY USING DIFFERENT EVOLUTION OPTIMIZATION | 34.11534           | 10.9325                 | 47.595                    | 4.166167 | 44.3868    | 29.55531     |
|        | BACK TRACKING SEARCH OPTIMIZATION | 14.58              | 27.34375                | 6.25                      | 43.42344 | 43.79898   | 25.78125     |
|        | GROUP SEARCH OPTIMIZATION | 85.01222           | 68.18853                | 88.01477                  | 59.46656 | 80.07167  | 57.40703     |

Features and back tracking search optimization is classified with Modest Adabost classifier. It is also analyzed that a least error rate of 8.33% is obtained when EM-PCA features and back tracking search optimization is classified with Optimized NBC. Also, a least error rate of 1.04% is obtained when Isomap features and back tracking search optimization is classified with Modest Adabost classifier.

From Table 16, it is analyzed for schizophrenia cases that a least error rate of 3.64% is obtained when PLS non linear features and flower pollination optimization is classified with optimized NBC classifier. It is also analyzed that a least error rate of 4.68% is obtained when EM-PCA features and back tracking search optimization is classified with Gaussian NBC classifier. Also, a least error rate of 1.04% is obtained when Isomap features and flower pollination optimization is classified with Real Adabost classifier.

To study the effect of feature extraction and Classifier performance irrespective of the optimization methods, an averaging of optimization results is initiated for the both Normal and schizophrenia Cases. The results are tabulated in the Table 17 and Table 18. From Table 17, it is analyzed for normal cases that when PLS non linear features are classified with classifiers, a high PI of 77.26% was obtained, a high classification accuracy of 91.47% was obtained, a high GDR of 82.94% along with an average error rate of 17.05% was obtained if classified with Adabost classifier. If EM-PCA features are classified with classifiers, then a high PI of 73.05%, along with a high classification accuracy of 89.27%, GDR of 72.66% and error rate of 21.46% was obtained if classified with Optimized NBC classifier. As in the same EM-PCA features, high GDR of 75.36% with the error rate of 24.63% is attained in the Real Adabost
TABLE 17. Average performance of parameters among the classifiers at various optimization techniques with different features for normal cases.

| SL.No | Features                      | Parameters(%)                     | Adaboost Classifier | Real Adaboost Classifier | Modest Adaboost Classifier | NBC     | Gaussian NBC | Optimized NBC |
|-------|-------------------------------|-----------------------------------|---------------------|--------------------------|---------------------------|---------|--------------|--------------|
| 1     | PLS NON LINEAR REGRESSION     | Performance Index                 | 77.26659            | 52.21406                 | 69.83797                  | 55.77636 | 58.00656     | 67.46781     |
|       |                               | Accuracy                          | 91.47259            | 85.59836                 | 90.60547                  | 87.18809 | 85.69461     | 88.30508     |
|       |                               | GDR                               | 82.94608            | 61.87823                 | 80.89224                  | 74.13005 | 64.44197     | 76.59703     |
|       |                               | Error Rate                        | 17.05392            | 28.80898                 | 18.78906                  | 25.62287 | 28.61132     | 23.40297     |
| 2     | EM-PCA                        | Performance Index                 | 29.09465            | 61.10414                 | 63.76016                  | 36.1475 | 43.13213     | 73.05422     |
|       |                               | Accuracy                          | 79.58341            | 87.68172                 | 87.48523                  | 81.23461 | 83.20133     | 89.27225     |
|       |                               | GDR                               | 50.60596            | 75.36264                 | 74.97007                  | 52.30466 | 66.40266     | 72.66278     |
|       |                               | Error Rate                        | 40.83404            | 24.65736                 | 25.02993                  | 37.53266 | 33.59734     | 21.46646     |
| 3     | ISOMAP                        | Performance Index                 | 50.83113            | 59.29047                 | 46.00866                  | 82.08906 | 61.83778     | 75.39688     |
|       |                               | Accuracy                          | 85.59776            | 86.97607                 | 84.38878                  | 92.905   | 86.54077     | 91.47344     |
|       |                               | GDR                               | 71.05222            | 73.95145                 | 68.77773                  | 84.81106 | 65.85312     | 82.36922     |
|       |                               | Error Rate                        | 28.80447            | 26.04855                 | 31.22227                  | 14.18873 | 26.92067     | 17.05313     |

TABLE 18. Average performance of parameters among the classifiers at various optimization techniques with different features for schizophrenia cases.

| SL.No | Features          | Parameters(%) | Adaboost Classifier | Real Adaboost Classifier | Modest Adaboost Classifier | NBC     | Gaussian NBC | Optimized NBC |
|-------|-------------------|---------------|---------------------|--------------------------|---------------------------|---------|--------------|--------------|
| 1     | PLS NON LINEAR REGRESSION | Performance Index | 58.28266            | 62.44619                 | 68.3475                   | 46.78379 | 41.51013     | 52.75078     |
|       |                   | Accuracy       | 86.94688            | 88.03747                 | 89.55125                  | 83.32244 | 82.38022     | 85.45063     |
|       |                   | GDR            | 73.89375            | 76.07392                 | 78.90069                  | 54.01398 | 56.45398     | 70.90199     |
|       |                   | Error Rate     | 26.10625            | 23.92608                 | 20.8975                   | 33.35564 | 35.24184     | 29.09801     |
| 2     | EM-PCA            | Performance Index | 49.94344            | 66.13375                 | 33.43734                  | 29.86891 | 58.66375     | 31.97891     |
|       |                   | Accuracy       | 84.6625             | 88.99422                 | 81.57285                  | 79.97109 | 86.72078     | 80.10277     |
|       |                   | GDR            | 57.97686            | 66.67414                 | 62.90352                  | 57.70329 | 73.38121     | 48.53224     |
|       |                   | Error Rate     | 30.68063            | 22.0125                  | 36.8543                   | 40.05781 | 26.56128     | 39.79486     |
| 3     | ISOMAP            | Performance Index | 71.98616            | 75.55903                 | 53.54219                  | 39.2398 | 51.22227     | 48.25672     |
|       |                   | Accuracy       | 89.3737             | 91.106                   | 86.32039                  | 82.87941 | 85.52646     | 83.68297     |
|       |                   | GDR            | 74.32943            | 82.21149                 | 72.64078                  | 57.38246 | 62.51949     | 62.63986     |
|       |                   | Error Rate     | 38.24439            | 26.87619                 | 38.84994                  | 37.57092 | 48.70499     | 34.92418     |

This peculiar situation is due to the averaging effect of optimization methods for EM-PCA feature extraction technique. If Isomap features are classified with classifiers, then a high PI of 82.08% and a high classification accuracy of 92.905%, GDR of 84.81% with Error Rate of 14.18% is obtained when classified with NBC classifier.
From Table 18, it is analyzed for schizophrenia cases that when PLS non linear features are classified with classifiers, a high PI of 68.34% was obtained, a high GDR of 89.55% was obtained, a high GDR of 78.9% along with an average error rate of 20.89% was obtained if classified with Modest Adaboost classifier. If EM-PCA features are classified with classifiers, then a high PI of 66.13% along with a high classification accuracy of 88.99% with low GDR of 66.67% and error rate of 22.01% was obtained if classified with Real Adaboost classifier. High GDR of 73.38% with the modest error rate of 26.56% is attained in the Gaussian NB classifier for EM-PCA feature extraction. The arrival of this situation is again attributed to the averaging effect of optimization methods for EM-PCA feature extraction.

If Isomap features are classified with classifiers, then a high PI of 75.55% and a high classification accuracy of 91.106%, GDR of 82.21% with Error rate of 26.87% is obtained when classified with Real Adaboost classifier. To select a better classifier a compromise strategy is needed among the classifier parameters like PI, Accuracy, GDR and Error rate. Generally, lesser the error rate means better accuracy among classifiers. As observed from the results of Table 17 and Table 18 the performance of classifier parameters are compressed due to the averaging effect of the optimization methods for the both cases of normal and schizophrenia. In other words, it implies that the optimization methods enhance the classifier performance to a higher level. The main reason why the combinations of Isomap-Backtracking search optimization-Modest Adaboost classification for the normal cases and the Isomap-Flower pollination optimization-Real Adaboost classification for schizophrenia cases gives the best result is because of the intrinsic property of Isomap algorithm which has been explored well by both Backtracking and Flower Pollination algorithms to optimize the features well with very less redundancy and so upon classification with the versions of Adaboost algorithm, it gives the best result.

VII. CONCLUSION AND FUTURE WORK

Characterized by abnormal behavior, decreased ability to understand reality, strange speech etc, schizophrenia is a very dangerous problem to the human community. People diagnosed with schizophrenia also have additional problems like depression, anxiety, lack of emotional expression and motivation. In this work, a comprehensive analysis of schizophrenia classification from EEG signals is done well with the help of feature extraction, optimization techniques and suitable classifiers. The methodology adopted here in this paper is quite promising and easy to implement. The average performance measures among the classifiers at various optimization techniques with different features for normal Cases and schizophrenia cases were explained in the work. The average PI and average error rate too was computed and presented among the classifiers at various optimization techniques with different features for normal cases and schizophrenia cases in this work. The individual results show that for normal cases, Isomap features when optimized with Backtracking search optimization algorithm and classified with Modest Adaboost classifier, a classification accuracy of 98.77% is obtained. For schizophrenia case, individual results show that when Isomap features are optimized with Flower Pollination optimization algorithm and classified with Real Adaboost classifier, a classification accuracy of 98.77% is obtained. Future works aim to work with different feature extraction techniques, optimization techniques and a plethora of other machine learning techniques to classify the schizophrenia from EEG signals.

REFERENCES

[1] C. R. Bowie and P. D. Harvey, “Cognitive deficits and functional outcome in schizophrenia,” Neuropsychiatric Disease Treatment, vol. 2, no. 4, pp. 531–536, Dec. 2006.
[2] E. M. Joyce and J. P. Raiser, “Cognitive heterogeneity in schizophrenia,” Current Opinion Psychiatry, vol. 20, no. 3, pp. 268–272, May 2007.
[3] X. Zhu, H.-I. Suk, S.-W. Lee, and D. Shen, “Canonical feature selection for joint regression and multi-class identification in Alzheimer’s disease diagnosis,” Brain Imag. Behav., vol. 10, no. 3, pp. 818–828, Aug. 2015.
[4] D. Freeman, “Studying and treating schizophrenia using virtual reality: A new paradigm,” Schizophrenia Bull., vol. 34, no. 4, pp. 605–610, Nov. 2007.
[5] Y. Chen, A. D. Attafu, I. Schlattner, W. T. Weldtsadik, M.-C. Roh, H. J. Kim, S.-W. Lee, B. Blankertz, and S. Fazli, “A high-security EEG-based login system with RSVP stimuli and dry electrodes,” IEEE Trans. Inf. Forensics Security, vol. 11, no. 12, pp. 2635–2647, Dec. 2016.
[6] J.-H. Kim, F. Biessmann, and S.-W. Lee, “Decoding three-dimensional trajectory of executed and imagined arm movements from electroencephalogram signals,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 23, no. 5, pp. 867–876, Sep. 2015.
[7] A. Peled, A. B. Geva, W. S. Kremen, H. M. Blankfeld, R. Esfandiarfar, and T. E. Nordahl, “Functional connectivity and working memory in schizophrenia: An EEG study,” Int. J. Neurosci., vol. 106, nos. 1–2, pp. 47–61, Jul. 2009.
[8] H. A. Neuhaus, F. C. Popescu, C. Grozea, E. Hahn, C. Hahn, C. Opgen-Rhein, C. Urbanek, and M. Dettling, “Single-subject classification of schizophrenia by event-related potentials during selective attention,” Neuroimage, vol. 55, no. 2, pp. 514–521, Mar. 2011, doi: 10.1016/j.neuroimage.2010.12.058.
[9] J. Laton, J. Van Schependom, J. Gielen, J. Decoster, T. Moons, J. De Keyser, M. De Hert, and G. Nagels, “Single-subject classification of schizophrenia patients based on a combination of oddball and mismatch evoked potential paradigms,” J. Neurol. Sci., vol. 347, no. 1–2, pp. 262–267, Dec. 2014.
[10] Q. Zhao, B. Hu, Y. Li, H. Peng, L. Li, Q. Liu, Y. Li, Q. Shi, and J. Feng, “An alpha resting EEG study on nonlinear dynamic analysis for schizophrenia,” in Proc. 6th Int. IEEE/EMBS Conf. Neural Eng. (NER), Nov. 2013, pp. 484–488, doi: 10.1109/NER.2013.6695977.
[11] Z. Dvey-Aharon, N. Fogelson, A. Peled, and N. Intrator, “Connectivity maps based analysis of EEG for the advanced diagnosis of schizophrenia attributes,” PLoS ONE, vol. 12, no. 10, Oct. 2017, Art. no. e0185852.
[12] E. Olejarczyk and W. Jernajczyk, “Graph-based analysis of brain connectivity in schizophrenia,” PLoS ONE, vol. 12, no. 11, Nov. 2017, Art. no. e0188629.
[13] A. Ulas, U. Castellani, V. Marino, M. Bellani, M. Tansella, and P. Brambilla, “Heat diffusion based dissimilarity analysis for schizophrenia classification,” in Proc. IAPR Int. Conf. Pattern Recognit. Bioinf., Berlin, Germany: Springer, 2011, pp. 306–317.
[14] S. R. Sponheim, B. A. Clementz, W. G. Iacono, and M. Beiser, “Resting EEG in first-episode and chronic schizophrenia,” Psychophysiology, vol. 31, no. 1, pp. 37–43, Jan. 1994.
[15] J. Sun, Y. Tang, K. O. Lim, J. Wang, S. Tong, H. Li, and B. He, “Abnormal dynamics of EEG oscillations in schizophrenia patients on multiple time scales,” IEEE Trans. Biomed. Eng., vol. 61, no. 6, pp. 1756–1764, Jun. 2014.
[16] W.-L. Chu, M.-W. Huang, B.-L. Jian, and K.-S. Cheng, “Analysis of EEG entropy during visual evocation of emotion in schizophrenia,” Ann. Gen. Psychiatry, vol. 16, no. 1, p. 34, Sep. 2017, doi: 10.1186/s12991-017-0157-z.
K. Rujirakul, C. So-In, and B. Arnonkijpanich, “PEM-PCA: A parallel expectation-maximization PCA face recognition architecture,” Sci. World J., vol. 2014, Apr. 2014, Art. no. 468176.

Y. Li and C. Guan, “An extended EM algorithm for joint feature extraction and classification in brain-computer interfaces,” Neural Comput., vol. 18, no. 11, pp. 2730–2761, Nov. 2006.

H. Cheng, H. Chen, G. Jiang, and K. Yoshihira, “Nonlinear feature selection by relevance feature vector machine,” in Proc. Int. Workshop Mach. Learn. Data Mining Pattern Recognit., 2007, pp. 144–159.

R. Rosipal, M. Girolami, L. J. Trejo, and A. Cichocki, “Kernel PCA for feature extraction and de-noising in nonlinear regression,” Neural Comput. Appl., vol. 10, no. 3, pp. 231–243, Dec. 2001.

Z. Wu, E. Xu, J. Long, F. Wang, X. Xu, Z. Jin, and A. Jiao, “Measurement of fermentation parameters of Chinese rice wine using Raman spectroscopy combined with linear and non-linear regression methods,” Food Control, vol. 56, pp. 95–102, Oct. 2015.

H. X. Wang, B. Luo, Q. B. Zhang, and S. Wei, “Estimation for the number of components in a mixture model using stepwise split-and-merge EM algorithm,” Pattern Recognit. Lett., vol. 25, no. 16, pp. 1799–1809, Dec. 2004.

H. Choi and S. Choi, “Robust kernel ISOMAP,” Pattern Recognit., vol. 40, no. 3, pp. 853–862, Mar. 2007.

X.-S. Yang, Nature-Inspired Optimization Algorithms, 1st ed. Amsterdam, The Netherlands: Elsevier, 2014, pp. 1–258.

X.-S. Yang, M. Karamanoglu, and X. He, “Flower pollination algorithm: A novel approach for multiobjective optimization,” Eng. Optim., vol. 46, no. 9, pp. 1222–1237, Oct. 2013.

H. Yacip and N. Çetinkaya, “An improved particle swarm optimization algorithm using eagle strategy for power loss minimization,” Math. Probl. Eng., vol. 2017, 2017, Art. no. 1063045.

W. K. Tian, S. Y. Liu, and J. X. Wang, “Study and improvement of backtracking search optimization algorithm based on differential evolution,” Appl. Res. Comput., vol. 32, no. 6, pp. 1653–1662, 2015.

C. Xiong, L. Hao, and D. Wang, “Group search optimizer with ranking-based differential operator,” in Proc. Int. Symp. Inf. Technol. Converg. (ISTCIT), Tianjin, China, vol. 4, Oct. 2015, pp. 3279–3285.

D. P. Young and J. M. Ferryman, “Faster learning via optimised AdaBoost,” in Proc. IEEE Conf. Adv. Video Signal-Based Serv. (AVSS), Sep. 2005, pp. 400–405.

S. Wu and H. Nagahashi, “Parameterized AdaBoost: Introducing a parameter to speed up the training of real AdaBoost,” IEEE Signal Process. Lett., vol. 21, no. 6, pp. 687–691, Jun. 2014.

A. Vezhevevets and V. Vezhevevets, “Modest AdaBoost; Teaching AdaBoost to generalize better,” Graphicon, vol. 12, no. 5, pp. 987–997, 2005.

P.-N. Tan, Introduction to Data Mining, London, U.K.: Pearson, 2006.

A. Meyer-Baese, Pattern Recognition for Medical Imaging, New York, NY, USA: Academic, 2004.

J. Friedman, T. Hastie, and R. Tibshirani, The Elements of Statistical Learning (Statistics), vol. 1, New York, NY, USA: Springer, 2001.

SUNIL KUMAR PRABHAKAR (Member, IEEE) received the B.E., M.E., and Ph.D. degrees from Anna University, Chennai, India, in 2012, 2014, and 2017, respectively. He is currently working as a Postdoctoral Research Fellow with the Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea. His research interests are signal processing, pattern recognition, and machine learning.

HARIKUMAR RAJAGURU received the B.E. degree in electronics and communication engineering from Bharathidasan University, Trichy, India, in 1988, the M.E. degree in applied electronics and the Ph.D. degree in information and communication engineering from Anna University in 1990 and 2009, respectively. His research interests are soft computing, pattern recognition, machine learning, bio signal processing, and image processing.

SEONG-WHAN LEE (Fellow, IEEE) received the B.S. degree in computer science and statistics from Seoul National University, Seoul, in 1984, and the M.S. and Ph.D. degrees in computer science from the Korea Advanced Institute of Science and Technology, Seoul, South Korea, in 1986 and 1989, respectively. He is currently the Hyundai-Kia Motor Chair Professor and the Head of the Department of Brain and Cognitive Engineering, Korea University. His research interests include pattern recognition, artificial intelligence, and brain engineering. He is a Fellow of IAPR and the Korea Academy of Science and Technology.