Finding predictive EEG complexity features for classification of epileptic and psychogenic nonepileptic seizures using imperialist competitive algorithm

Please check the document version of this publication:

Finding Predictive EEG Complexity Features for Classification of Epileptic and Psychogenic Nonepileptic Seizures using Imperialist Competitive Algorithm

Negar Ahmadi*, Evelien Carrette**, Albert P. Aldenkamp***, Mykola Pechenizkiy*

*Department of Mathematics and Computer Science
Eindhoven University of Technology, Eindhoven, the Netherlands
**Neurology Department, Ghent University Hospital, Gent, Belgium
***Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands
Email: n.ahmadi@tue.nl , evelien.carrette@ugent.be , a.p.aldenkamp@tue.nl , m.pechenizkiy@tue.nl

Abstract—In this study, the imperialist competitive algorithm (ICA) is applied for classification of epileptic seizure and psychogenic nonepileptic seizure (PNES). For this purpose, after decomposing the EEG signal into five sub-bands and extracting some complexity features of EEG, the ICA is applied to find the predictive feature subset that maximizes the classification performance in the frequency spectrum. Results show that the spectral entropy and Renyi entropy are the most important EEG features as they are always appeared in the best feature subsets when applying different classifiers. Also, it is observed that the SVM-RBF and SVM-linear models are the best classifiers resulting in highest performance metrics compared to other classifiers. Our study shows that the reported algorithm is able to classify the epileptic seizure and PNES with a very high classification metrics.

Keywords-epileptic seizures, PNES, complexity, EEG, imperialist competitive algorithm.

I. INTRODUCTION

An epileptic seizure, also known as an epileptic fit or attack, is a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain [1]. The outward effect can vary from temporary confusion, loss of awareness and uncontrolled jerking movement to as subtle as a momentary loss of awareness or whole body convulsion. Patients are often unaware of the occurrence of seizure due to the random nature of them which may increase the risk of physical injury.

Psychogenic nonepileptic seizures (PNES) are attacks that may look like epileptic seizures, but caused purely by the emotions and not associated with abnormal electrical discharges in the brain. The symptoms of PNES usually reflect a psychological conflict that are inconsistent with a neurologic disease and it is often associated with distress, disability, and a poor prognosis [2]. PNES episodes are not purposely produced by the patient, and the patient is not aware that the seizures are non-epileptic, so the patient may become anxious over having these symptoms. The presentation of the differential diagnosis should be done early in the course of treatment for better patient acceptance, and treatment options should be presented early in the evaluation period [3].

Early diagnosis of epileptic seizure or PNES is critical. Due to the delay in early prediction of epileptic seizures, many patients may experience the attack, which could be avoided by the drug. Also, because of delay in early diagnosis, many patients experience significant morbidity from inappropriate treatment, including adverse effects of antiepileptic drugs and aggressive interventions, such as intubation for pseudostatus epilepticus [4]. However, PNES is commonly misdiagnosed as epileptic seizure or epilepsy, and patients are often treated for years with an incorrect diagnosis. The management of PNES as epileptic seizures can lead to very significant iatrogenic harm. Moreover, the failure to recognize the psychological cause of the disorder detracts physicians from addressing associated psychopathology, and enhances secondary somatization processes [5]. Last, the inappropriate treatment of PNES as epilepsy is costly.

In the current day practice, the intensive monitoring with electroencephalogram (EEG) and video over a long period is the standard way in differentiating PNES from epileptic seizure. It simultaneously records the patients brain electrical activity and captures corresponding behaviours on video. However, the long-term monitoring with EEG and video is expensive, time-consuming and can be very unpleasant for patients, and analysing large amounts of EEG/video-data is very labor intensive for medical personnel.

To overcome the above-mentioned issues, several scholars have focused on EEG signal analysis and process to aid in the diagnosis and treatment of brain disorders. Hence, various mathematical techniques were proposed in the literature for the detection of epileptic seizures and/or PNES in EEG signals. The first step in EEG signal analysis is to extract selected features by applying various time-domain, frequency-domain, time-frequency domain, or non-linear methods [6]. Then, the selected features should be considered as discriminative features for classification of these two groups by analysing different EEG signals. For this purpose, numerous classifiers such as (non-)linear classifiers and techniques based on neural networks are used for EEG classification [7].

In this paper, the imperialist competitive algorithm (ICA),

as a capable evolutionary algorithm based on the meta-heuristic of humans socio-political evolution [8], is applied for classification of epileptic seizure and PNES.

The ICA algorithm has been successfully applied to a variety of optimization problems [8], [9]. The key features of ICA are its fast convergent rate to reach global optimum, which has been proved in dealing with various optimization problems. The results reported in various studies [8], [9] confirm its competitiveness over other evolutionary algorithm such genetic algorithm. The ease of performing neighborhood movement, less dependency on initial solutions, and having a better convergence rate are other advantages of the ICA [8]. Hence, the advantages of ICA are beneficial to improvement of decision efficiency. In this algorithm, an individual of the population is called a country. The ICA divides its population into several groups, called empires, and allows these empires to evolve concurrently. In each empire, the best country is called imperialist and the others are called colonies. The ICA moves all colonies toward the imperialist through assimilation policy in each empire. The basic feature of the ICA is that it permits all empires to interact via imperialist competition policy. The competition policy simply moves a colony from the weakest empire to another empire. Some colonies may withstand absorption by the imperialists. These colonies make some improvements in their attributes, and this process is called revolution in the ICA. Revolution operation occurs after the assimilation process and causes unexpected random changes in one or more parameters of the problem. This operation increments exploration and prevents fast convergence of countries toward local minima. After decomposition of the EEG signals by the wavelet transform (WT), some selected signal complexity features are extracted at different frequency bands. Then, the ICA algorithm is applied to find feature subset that maximizes the classification performance in the frequency spectrums. given in the last section.

II. METHODOLOGY

A. Clinical Data

The experimental data used in this paper were obtained from the UZ Gent Hospital in Belgium. The EEG recordings were obtained from 20 epilepsy and 20 PNES patients and the recordings from each subject include 27 EEG recordings electrode and reference (G2) on the right mastoid bone plus the ground (G1) on the left mastoid bone. The sampling rate of all data channels is 256 Hz, and the duration of each trial is 3 hours. The 27 channels are: Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, Oz, O2, T9, T10, FT9, FT10, TP9, TP10. Fig. 1 shows the EEG recording positions on brain.

B. EEG Decomposition

EEG signals which are the records of the electrical activity going on inside the brain as taken from the scalp with the
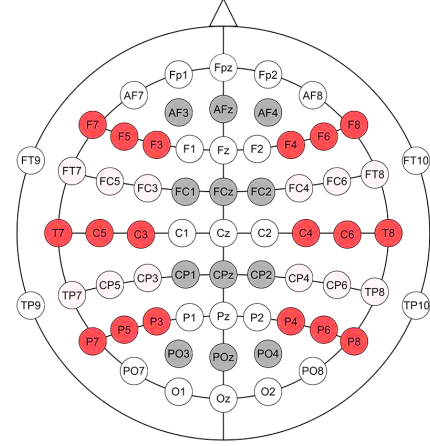
Figure 1: Electrode positions and labels on brain.

help of electrodes. Brain waves are measured in Hertz (Hz) cycles per second, and can change across a wide range of variables. The main frequency components of EEG signals are Delta (¡4 Hz), Theta (4-8 Hz), Alpha (9-13 Hz), Beta (14-30 Hz) and finally Gamma (above 30 Hz) (see Fig. 2). In this study, a wavelet-based time-frequency scheme [10] is used to decompose the EEG signals into the above-mentioned sub-bands. The wavelet decomposition is a smooth and quickly vanishing oscillating function with good localization in both frequency and time.
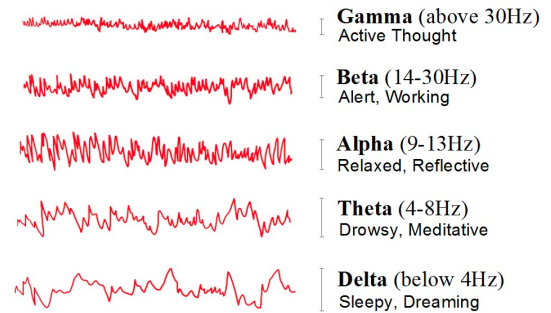
Figure 2: Sub-bands of EEG brain signal.

C. EEG Features

A feature represents a distinguishing property and a functional component obtained from a signal. Extracted features are meant to minimize the loss of important information embedded in the signal and simplify the amount of resources needed to describe a huge set of data accurately [11]. This is necessary to minimize the complexity of implementation, to reduce the cost of information processing, and to cancel the potential need to compress the information. In this paper, we use different features based on EEG signals.

1) Shannon entropy (ShE): Entropy is a way of measuring the degree of uncertainty or unpredictability of a random

variable. Shannon entropy is a non-linear measure quantifying the degree of complexity in a signal. Let X be a set of finite discrete random variables $X = \{x_1, x_2, , x_n\}; x_j \in R^d$. Now, the Shannon entropy, $H(X)$, is defined as [12]:

$$\text{ShE} = H(X) = -\sum_{j=1}^{n} p(x_j) \ln p(x_j) \tag{1}$$

Where $p(x_j)$ is probability of $x_j \in X$ satisfying $\sum_j p(x_j) = 1$. Entropy reflects how well one can predict the behaviour of each respective part of the trajectory from the other. Basically, higher entropy indicates more complex or chaotic systems, thus, less predictability.

2) Spectral entropy (SE): Spectral entropy (SE) computation uses Shannons entropy formula to represent the power spectral densities of the EEG signal as probabilities. For this purpose, fast Fourier's transformation (FFT) is used to obtain the spectrum. The SE corresponding to the frequency range $[f_1, f_2]$ is defined as [13]:

$$\text{SE}[f_1, f_2] = \frac{\sum_{f_2, f_1}^{f_1} P_n(f_i) \log(\frac{1}{P_n(f_i)})}{\log(N[f_1, f_2])} \tag{2}$$

where $N[f_1, f_2]$ equals the total number of frequency components in the frequency range and $P(f_i)$ is the power spectrum calculating from the FFT of signal X.

3) Renyi entropy (RE): Renyi entropy, as an index of dicersity, is generalizations of Shannon entropy that depend on a parameter. If $p(x_i)$ is a probability distribution on a finite set, its Renyi entropy of order α is defined as $RE = \frac{1}{1-\alpha} \ln \sum_{i=1}^{n} p(x_i)^\alpha$, where $0 < \alpha < \infty$. Renyi entropy approaches Shannon entropy as $\alpha \to 1$ [14].

4) Higuchi fractal dimension (HFD): Fractal dimension provides a measure of the complexity of EEG signals. HFD is a fast non-linear computational method for obtaining the fractal dimension of signals even when very few data points are available [15]. HFD is used to quantify the complexity and self-similarity of a signal. To compute the HFD, the data set is divided into a k-length sub-data set as $x_k^m : x_m, x_{m+k}, x_{m+2k}, ..., x_{m+(\frac{n-m}{k})k}$, where n is the total length of the data sequence, k is a constant and $m = 1, 2, ..., k$. The length $L_m(k)$ for each sub-data set is then computed as:

$$L_m(k) = \frac{\sum_{i=1}^{N-\frac{m}{k}} | x_{m+ik} - x_{m+(i-1)k} | (n-1)}{(\frac{n-m}{k})k} \tag{3}$$

Now, the mean of $L_m(k)$ for each k is computed to find the HFD as:

$$\text{HFD} = \frac{1}{k} \sum_{m=1}^{k} L_m(k) \tag{4}$$

5) 5) Katz fractal dimension (KFD): KFD is derived directly from the waveform, eliminating the preprocessing step of creating a binary sequence, can be defined as [16]:

$$\text{KFD} = \frac{\log_{10}(n)}{\log_{10}(\frac{d}{L}) + \log_{10}(n)} \tag{5}$$

where n is the number of steps in the curve, L is the total length of the signal, and d is the Euclidean distance between the first point in the series and the point that provides the furthest distance with respect to the first point.

D. Imperialist Competitive Algorithm

After extracting the entropy and fractal dimension features from the EEG signals, they are inputted to a classifier based on the imperialist competitive algorithm (ICA). The ICA is based on modelling of the attempts of countries to dominate other courtiers and like other evolutionary algorithms, starts with an initial population [8], [17]. In the ICA, populations are in two types: colonies and imperialists that the best countries in the population are selected to be the imperialist states and all the other countries form the colonies of these imperialists. Imperialist competition among these empires forms the basis of the ICA, as weak empires collapse and powerful ones take possession of their colonies. This competition and collapse mechanism will cause all the countries to converge to a state in which there exist just one empire in the world and all the other countries are its colonies.

The ICA algorithm starts by generating a set of candidate random solutions in the search space of the optimization problem. The generated random points are called the initial countries. In this study, we consider each country as a 1×30 array, where each element of array shows the existence of one complexity feature in one of the frequency bands and can take a zero or one as a value. During the initialization stage, an initial population $p_1, p_2, , p_N$ are randomly created, where each solution p_i is called a country and is a $1 \times n$ array and N (here: =100) denotes the number of countries in the population. A user-specified number of countries with the lowest cost in the population are chosen as imperialists, (N_{imp}), and the remaining countries are chosen as colonies, which all together form empires. For our purpose the cost function of each country with different feature subset can be calculated as the average of the misclassification rates of different classifiers. The initial number of colonies of an empire is convenience with their powers. To divide the colonies among imperialists proportionally, the power of an imperialist is defined as follows [8], [17]:

$$P_i = max_{i \leq j \leq N_{imp}}(c_j) - c_i \tag{6}$$

where P_i and c_i denote the power and cost of the imperialist of empire i, respectively. Therefore, the number of colonies assigned to empire I is defined as follows:

$$NC_i = round[| \frac{P_i}{\sum_{j=1}^{N_{imp}} P_j} | .(N_i - N_{imp})] \tag{7}$$

To divide the colonies, for each imperialist we randomly choose NC_i of the colonies and give them to it. After forming the initial empires and during the evolution step, the colonies start moving toward their relevant imperialist

country. Assimilation within each empire and competition among all empires occur in every generation until the termination condition (e.g., all countries have converged or a user-specified number of generations has been reached) is satisfied [18]. The colony moves toward the imperialist by x units, where x is a random variable with uniform or any proper distribution. The direction of the movement is the vector from colony to imperialist. In other words, given a colony p_c and its imperialist p_i, the assimilation operation moves p_c as follows [8], [17]:

$$p_c = p_c + \beta * \Delta * (p_i - p_c) \quad (8)$$

where β is a parameter greater than one, Δ is a $1 \times n$ array whose elements are random values between zero and one, and $*$ denotes element-by-element multiplication between two $1 \times n$ arrays. Note that a β greater than one, causes the colonies to get closer to the imperialist state from both sides. If during implementing the above equation, a greater value outside the search space happens, the out-of-bound value (e.g., x_i on the ith dimension) is replaced by its nearest boundary.

In the next step, the cost of all colonies is calculated after updating the position of all colonies through the assimilation process. Then, the cost of each colony is compared against the cost of their imperialists. If the cost of a colony be less than the cost of its imperialist, the colony and the imperialist swap roles to ensure that the imperialist of an empire is always the country with the lowest cost in the empire. Next, the cost and the power of each empire i are calculated using the cost of its imperialist and the average cost of the colonies in empire i as follows:

$$\phi_i = c_i + \zeta.mean(\text{cost(colonies of empire } i)) \quad (9)$$

$$E_i = [max_{1 \leq j \leq N_{imp}}(\phi_j)] - \phi_i \quad (10)$$

where ϕ_i and E_i are the cost and the power of each empires, respectively and ζ is a positive number with suggested value 0.1. A little value ζ for causes the total power of the empire to be determined by just the imperialist and increasing it will increase the role of the colonies in determining the total power of an empire. This study uses ζ=0.02. Competition among all empires is achieved by taking the weakest colony away from the weakest empire and giving it to a chosen empire, where the probability of empire i been chosen is calculated as $p_i = \frac{E_i}{max_{1 \leq j \leq N_{imp}}(E_j)}$.

In the ICA, imperialists try to attempt to achieve the colonies of other empires and control them. So during the competition the powerful imperialists will be increased in the power and the weak ones will be decreased in the power. When an empire loses all of its colonies, it is assumed to be collapsed and it's imperialist also becomes a colony of the latter empire. At the end the most powerful imperialist will remain in the world and all the countries are colonies of this unique empire. In this stage the imperialist and colonies have

the same position and power. After some iteration, only the most powerful empires will remains and all the other empires will collapse and their colonies will be under the control of this unique empire. The algorithm of the ICA is shown in Fig. 3.

---

1. *Initialization:* select some random points on the function and initialize the empires,
2. *Do:* for each empire i,
3. *Assimilation:* move colonies toward their relevant imperialist,
4. *Swap:* If the cost of a colony is less than the cost of its imperialist, exchange the positions of them,
5. *Cost Calculation:* compute the total cost of all empires,
6. *Imperialistic Competition:* pick the weakest colony from the weakest empire & give it to the empire that has the most likelihood to possess it,
7. *Elimination:* remove the powerless empire,
8. *Check:* if there is just one empire or the convergence criteria is satisfied, **STOP**, if not go to step **2**.

Figure 3: Algorithm of the imperialist competitive algorithm.

E. Classification

We use 80% instances (32 subjects) as the training data and the rest 20% (8 subjects) as the test data. Due to the limited number of subjects and in order to avoid over-fitting, we split the training and the test data randomly and repeat the split process 10 times. Our results are the means and variances of these 10 runs. To explore the importance of features and their combinations in the classification task, five widely used classifiers are applied. The selected classifiers are: support vector machine (SVM) classifier with linear and radial basis function (RBF) kernels [19], decision tree [20], random forest [21] and gradient boosting [22]. To evaluate the performance of these features in the classification task, three evaluation metrics, i.e. accuracy, precision and recall, are applied in the experiments.

III. RESULTS

In this study, five selected complexity measures should be examined in all sub-band frequencies of 27 EEG recording channels. Hence, there are more than 800 features for each subject. However, many of these features may measure related properties and so will be redundant. In order to have less computational complexity and to improve the accuracy of classification by reducing feature vectors size, the optimum feature subsets that contain and summarize all important data are obtained. The top five feature subsets for each classifier are presented in Table I. Here, each term shows the name of EEG feature (i.e. ShE, SE, RE, HFD and KFD) and the frequency sub band (i.e. Gamma, Alpha, Beta, Theta and Delta), where BB stands for broad band without frequency decomposition. These top five subsets are ranked based on the accuracy of the classification that is the number of correct predictions (or classification) made divided by the

total number of predictions made. The first subset is the winner of the ICA outputs that has the lowest cost function. The rest of subsets represent four other subsets (countries) with low cost functions among the whole subsets in the ICA. One can see that the winner and other ranked subsets for each classifier are different. The reason is that each classifier takes part in the ICA for calculating the cost function and generates feature subsets for itself. From the data in Table I it can be seen that spectral entropy (SE) and Renyi Entropy (RE) are the most important EEG features as they are always appeared in the best feature subsets.

Our analysis shows that the accuracy of the classification decreases significantly when SE and/or RE features are absent in a subset. The rest of features may be of same importance since by removing either of them the classification accuracy changes without any significant differences. Table I shows the degradation in the classification ranking and accuracy when the number of features in the subsets is reduced.

Accuracy, precision and recall of selected classifiers with the presented five best subsets are shown in Tables II to VI. The classification accuracy displays the correct classifications that maximize the total number of correct classifications. Precision is a measure of result relevancy, while recall is a measure of how many truly relevant results are returned. The presented data shows degradation in the performance measures when the rank is increases from 1st to 5th. Hence, the first subsets represents the feature subset trained by the ICA for classification of epileptic seizure and PNES as the highest classification metrics are achieved. Also, it can be observed that the SVM-RBF and SVM-linear are the best classifiers resulting in highest performance metrics compared to other classifiers.

IV. CONCLUSION

In this paper we reported the results on the constructed benchmark to investigate classification of epileptic seizure and PNES. We employed the imperialist competitive algorithm (ICA) to identify predictive features for classification and used state of the art classification techniques on signals including periods of seizures to see how accurately class labels can be predicted. The study demonstrated that the classification performance of the SVM-RBF and SVM-linear classifiers are the best when the ICA was employed. The reported algorithm showed a very high classification metrics for classification of the epileptic seizure and PNES and the results provide us with the new insights on feature importance. This study found spectral entropy and Renyi Entropy as the most important EEG features for classification of epileptic seizure and PNES.

REFERENCES

[1] R.S. Fisher, et al., Epileptic seizures and epilepsy: definitions proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE), *Epilepsia*, 46(4), 470-472, 2005.

[2] B.J. Smith, Closing the major gap in PNES research: finding a home for a borderland disorder, *Epilepsy Currents*, 14(2), 63-67 (2014).

[3] P.M. Gillig, Psychogenic nonepileptic seizures, *Innov. Clin. Neurosci.*, 10(11-12), 15-18, 2013.

[4] M. Reuber, et al. Diagnostic delay in psychogenic nonepileptic seizures, *Neurology*, 58(3), 493495, 2002.

[5] M. Reuber, C.E. Elger, Psychogenic nonepileptic seizures: review and update, *Epilepsy & Behavior*, 4(3), 205-216, 2003.

[6] U.R. Acharya, et al., Automated eeg analysis of epilepsy: a review. *Knowledge-Based Sys.*, 45(3), 147-165, 2013.

[7] F. Lotte, et al., A review of classification algorithms for eeg-based brain-computer interfaces. *J. Neural Eng.*, 4(2), R1-R13, 2007.

[8] E. Atashpaz-Gargari, F. Hashemzadeh, R. Rajabioun, C. Lucas, Colonial competitive algorithm: a novel approach for PID controller design in MIMO distillation column process. *Int. J. Intelligent Comput. Cyb.*, 1(3), 337355, 2008.

[9] S. Hosseini, A. Al Khaled, A survey on the imperialist competitive algorithm metaheuristic: implementation in engineering domain and directions for future research. *Appl. Soft Comput.*, 24, 1078-1094, 2014.

[10] H. Adeli, Z. Zhou, N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *J. Neurosci. Methods*, 123(1), 69-87, 2003.

[11] A.S. Al-Fahoum, A.A. Al-Fraihat, Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains, *ISRN Neuroscience*, 730218, 2014.

[12] N. Kannathal, M.L. Choo, U.R. Acharya, P.K. Sadasivan, Entropies for detection of epilepsy in eeg, *Comput. Methods Prog. Biomed.*, 80(3), 187-194, 2005.

[13] R.R. Nunes, M.P.D. Almeida, J.W. Sleigh, Spectral entropy: a new method for anesthetic adequacy, *Revista brasileira de anestesiologia*, 54(3), 404-422, 2004.

[14] C. Beck, F. Schlgl, *Thermodynamics of chaotic systems: an introduction.* Cambridge University Press, 1993.

[15] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, *Physica D*, 31, 277 283, 1988.

[16] M.J. Katz, Fractals and the analysis of waveforms, *Comput. Biol. Med.*, 18(3), 145-156, 1988.

[17] E. Atashpaz-Gargari, C. Lucas, Imperialist Competitive Algorithm: An algorithm for optimization inspired by imperialistic competition, *IEEE Congress on Evolutionary Computation*, 7. 46614666, 2007.

[18] J.L. Lin, Y.H Tsai, C.Y Yu, M-S Li, Interaction Enhanced Imperialist Competitive Algorithms, *Algorithms*, 5, 433-448, 2012.

[19] J.A. Suykens, J. Vandewalle, Least squares support vector machine classifiers, *Neural Process. Lett.*, 9(3), 293-300, 1999.

[20] S.R. Safavian, D. Landgrebe, A survey of decision tree classifier methodology, *IEEE T. Sys. Man Cyb.*, 21(3), 660-674, 1991.

[21] M. Pal, Random forest classifier for remote sensing classification, *Int. J. Remote Sens.*, 26(1), 217-222, 2005.

[22] R. Lawrence, A. Bunn, S. Powell, M. Zambon, M, Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis, *Remote Sens. Environ.*, 90(3), 331-336, 2004.

Table I: Best feature subsets for different classifiers.

rank	SVM-Linear
1st	SE-BB, RE-BB, KFD-Beta, ShE-Beta, SE-Alpha, SE-Beta, RE-Beta, RE-Alpha, KFD-Delta, RE-Theta
2nd	SE-BB, RE-BB, KFD-Beta, ShE-Beta, SE-Alpha, SE-Beta, RE-Beta, RE-Alpha, KFD-Theta, KFD-BB
3rd	SE-BB, RE-BB, KFD-Beta, ShE-Beta, SE-Alpha, SE-Beta, KFD-Delta, RE-Theta, ShE-Alpha, HFD-Delta
4th	SE-BB, RE-BB, KFD-Beta, ShE-Beta, SE-Alpha, KFD-Theta, RE-Theta, KFD-BB, ShE-Alpha, HFD-Theta
5th	SE-BB, RE-BB, KFD-Beta, ShE-Beta, RE-Beta, RE-Alpha, RE-Theta.

rank	SVM-RBF
1st	SE-BB, RE-BB, RE-Beta, KFD-Beta, RE-Theta, SE-Alpha, SE-Beta, ShE-Beta, KFD-BB
2nd	SE-BB, RE-BB, RE-Beta, KFD-Beta, RE-Theta, SE-Beta, KFD-Theta, ShE-Theta
3rd	SE-BB, RE-BB, RE-Beta, KFD-Beta, SE-Alpha, SE-Beta, HFD-Beta
4th	SE-BB, RE-BB, RE-Beta, KFD-Beta, ShE-Beta, ShE-Theta.
5th	SE-BB, RE-BB, RE-Beta, SE-Alpha, KFD-BB, RE-Alpha

rank	Gradient Boosting
1st	RE-BB, RE-Beta, ShE-Theta, SE-Alpha, SE-Beta
2nd	RE-BB, RE-Beta, SE-Beta, SE-BB, RE-Alpha, RE-Theta
3rd	RE-BB, RE-Beta, ShE-Theta, KFD-Beta, RE-Theta, SE-Alpha
4th	RE-BB, RE-Beta, SE-Beta, SE-BB, KFD-Beta, HFD-Theta
5th	RE-BB, RE-Beta, ShE-Theta, ShE-Beta

rank	Decision Tree
1st	SE-Beta, RE-Beta, KFD-Alpha, KFD-Beta, ShE-Theta, SE-Theta, RE-BB
2nd	SE-Beta, RE-Beta, KFD-Alpha, KFD-Beta, RE-Alpha, KFD-Theta
3rd	SE-Beta, RE-Beta, KFD-Beta, RE-BB, HFD-Alpha, HFD-Beta
4th	SE-Beta, RE-Beta, KFD-Alpha, KFD-Theta, RE-Theta, RE-BB
5th	SE-Beta, RE-Beta, SE-Theta, ShE-Theta, KFD-Alpha

rank	Random Forest
1st	RE-BB, RE-Beta, SE-Beta, ShE-Theta, RE-Alpha, SE-Alpha
2nd	RE-BB, RE-Beta, SE-Beta, ShE-Beta, KFD-Beta, KFD-BB
3rd	RE-BB, RE-Beta, RE-Alpha, SE-Alpha, HFD-Theta
4th	RE-BB, ShE-BB, SE-Theta, KFD-Theta
5th	RE-BB, ShE-BB, HFD-Alpha, RE-Theta

Table II: Performance metrics of SVM-Linear classifier

rank	Accuracy	Precision	Recall
1st	0.9489±0.0016	0.9411±0.0012	0.9392±0.0012
2nd	0.9355±0.0024	0.9363±0.0007	0.9391±0.0009
3rd	0.9231±0.0061	0.9319±0.0011	0.9265±0.0060
4th	0.9201±0.0032	0.9191±0.0010	0.9197±0.0031
5th	0.9169±0.0026	0.9068±0.0028	0.9094±0.0020

Table III: Performance metrics of SVM-RBF classifier

rank	Accuracy	Precision	Recall
1st	0.9503±0.0024	0.9609±0.0013	0.9523±0.0061
2nd	0.9469±0.0041	0.9455±0.0057	0.954±0.0075
3rd	0.9417±0.0015	0.9448±0.0061	0.9415±0.0063
4th	0.9363±0.0016	0.9283±0.0040	0.9303±0.0045
5th	0.9267±0.0058	0.9291±0.0048	0.9233±0.0023

Table IV: Performance metrics of gradient boosting classifier

rank	Accuracy	Precision	Recall
1st	0.9372±0.0016	0.9399±0.0015	0.9415±0.0054
2nd	0.9346±0.0016	0.9202±0.0013	0.9387±0.0008
3rd	0.9299±0.0016	0.9122±0.0029	0.9328±0.0012
4th	0.9125±0.0021	0.9184±0.0018	0.9073±0.0016
5th	0.9088±0.0044	0.8986±0.0041	0.9059±0.0132

Table V: Performance metrics of decision tree classifier

rank	Accuracy	Precision	Recall
1st	0.8711±0.0024	0.8669±0.0011	0.8779±0.0012
2nd	0.8635±0.0064	0.8611±0.0031	0.8544±0.0008
3rd	0.8515±0.0044	0.8569±0.0056	0.8591±0.0016
4th	0.8509±0.0061	0.8603±0.0032	0.8539±0.0062
5th	0.8466±0.0092	0.8314±0.0029	0.8442±0.0041

Table VI: Performance metrics of random forest classifier

rank	Accuracy	Precision	Recall
1st	0.8889±0.0009	0.8787±0.0012	0.8824±0.0012
2nd	0.8823±0.0009	0.8716±0.0008	0.8806±0.0023
3rd	0.8746±0.0015	0.8623±0.0054	0.8765±0.009
4th	0.8639±0.0026	0.8699±0.0049	0.8718±0.0057
5th	0.8625±0.0039	0.8512±0.0031	0.8651±0.0061