A Machine Learning Approach to Predict rTMS Therapy Response in Major Depressive Disorder

Mohammad Shams

***Abstract*—Machine learning techniques have been utilized to predict the outcome of repetitive transcranial magnetic stimulation (rTMS) treatment in depression, e.g., through classifying the responders (R) and non-responders (NR) to rTMS treatment for major depression disorder (MDD) patients. MDD is among the leading causes of disability in the world with affecting more than 260 million people, and a major contributor to the overall global burden of disease. In this study, the outputs of the Local Subset Feature Selection (LSFS) method were used by an SVM classifier to evaluate the capability of the proposed method in the prediction of rTMS treatment response in depression cases. A Leave-One-Out cross-validation method is applied to the input data to evaluate the performance of the response classification. The achieved accuracy, sensitivity, and specificity were 89.5%, 90%, and 87%, respectively. The main restriction of this study that would limit its usage in clinical applications is the small sample size.**

***Index Terms*—Major depressive disorder (MDD), repetitive transcranial magnetic stimulation (rTMS), electroencephalography (EEG), classification**

I. Introduction

MAJOR depressive disorder (MDD), also simply known as depression, is a debilitating mental disease characterized by at least one discrete depressive episode lasting at least two weeks. MDD is a leading cause of disability worldwide with affecting more than 260 million people and is a major contributor to the overall global burden of disease with genetic, biological, and environmental risk factors [1]. Although both psychotherapy and psychopharmacology are effective in treating MDD [2], nonetheless the development of new therapeutic procedures is still a need to reduce adverse side effects of pharmacological treatments [3]. Evidence-based guidelines on the therapeutic use of Repetitive transcranial magnetic stimulation (rTMS) found that this procedure is effective in treating psychiatric disorders like depression and schizophrenia [4]. rTMS is a noninvasive method that stimulates brain nerve cells by applying magnetic fields for treating major depression. This noninvasive treatment is safe and pain-free with minimal side effects [5], [6]. In 2008, the US Food and Drug Administration (FDA) approved rTMS as a therapy for mildly treatment-resistant depression. Since the initial clearance by the FDA, rTMS is progressively being applied into clinical practice [7].

By considering the applicability of rTMS, prediction the clinical response to rTMS is a need to improve the treatments for depression by reducing the costly ineffective pharmacological and psychotherapeutic treatments. Several previous studies have used machine learning techniques to predict response to rTMS. Bailey et al. predicted responses to rTMS treatment for depression by classification on EEG data recorded during working memory by using weighted phase lag index (wPLI) in gamma frequency band, and achieved classification accuracy of 91% [8]. In another study, they used the features of EEG power and weighted phase lag index (wPLI) only in alpha and theta frequency bands, alpha peak frequency (iAPF) and frontal theta cordance. Their classification accuracy of the combination of mood and EEG features by LSVM classifier was 86.6% [9]. Khodayari-Rostamabad et al. used power spectral, mutual information (MI), and coherence features through applying machine learning methods to predict MDD treatment response [10]. Also, Mumtaz et al. employed wavelet coefficients and coherence features to study the results in the same subject [11]. Their achieved results illustrated as accuracy, sensitivity, and specificity are shown in Table I. Furthermore, since EEG is a signal with nonlinear dynamics, it seems applying the Time-Frequency (TF) features of EEG may be informative for predicting the response to treatment. However, as mentioned before, most of the previous studies in the prediction of MDD treatment outcome only have focused on linear, nonlinear, spectral, bispectral and cordance features. The novel contribution of this proposed approach is that, as far as we are aware, no study has analyzed TF features of EEG for Prediction of rTMS therapy Response in MDD. The flowchart of this study is shown in Figure 1 on page 2.

II. Methods

A. Participants

In this study, 36 patients with MDD in the age range of 16–71 years participated. The patients were referred to the Brain and Cognition Clinic, Tehran, Iran. MDD diagnosis was performed by three experienced psychiatrist based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) published by the American Psychiatric Association (APA). Participants also assessed by Hamilton Rating Scale for Depression (HRSD), and Beck Depression Inventory (BDI-II), and all have the HRSD score $\geq$12 and BDI-II score $\geq$15. In this study, 18 patients were under psychiatric medication treatment (antidepressant, mood stabilizer or antipsychotic) fixedly from more than four weeks before the treatment. All patients provided written informed consent. The design and all procedures adhered to the latest version of the Declaration of the Brain and Cognition Clinic guidelines.

B. Treatment and Clinical Assessment

Before EEG recording and receiving rTMS treatment, the participants were washed out at least 5 days of antidepressant, antipsychotics and mood stabilizer medications. The rTMS was applied over the left dorsolateral prefrontal cortex

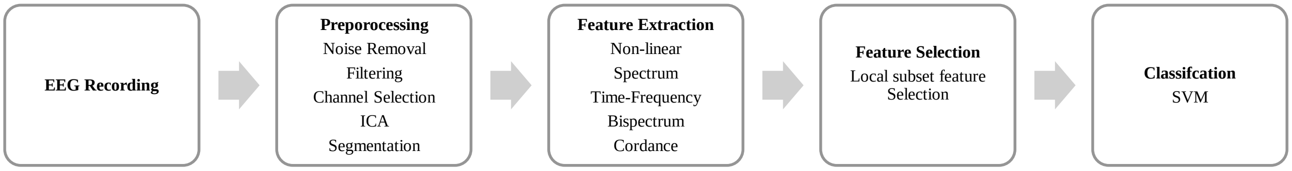

Fig. 1: The flowchart of the study process

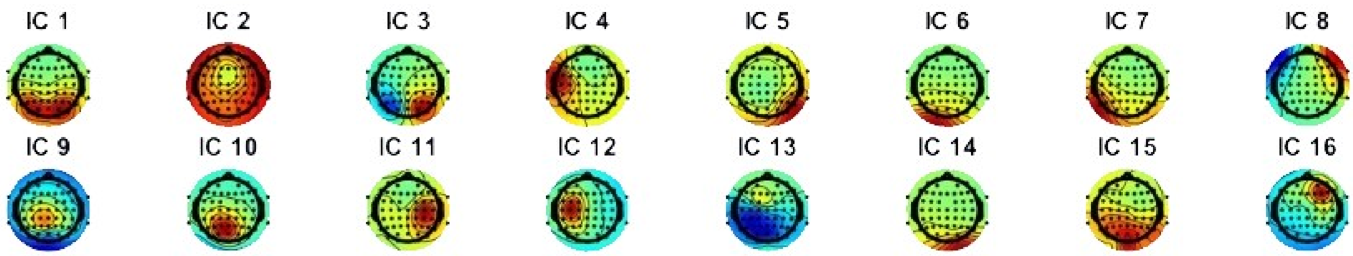

Fig. 2: Extracted Independent Components (ICs) from EEG signals

(DLPFC) at a stimulation site defined by a point 5 cm anterior in a parasagittal line to the motor threshold location, with the coil held tangentially to the scalp and its handle pointing back and away from the midline at 45° [12]. The rTMS procedure was delivered by a TAMAS (REMED, Daejeon, Korea) with a figure-of-eight shaped coil (field. strength $\sim$3 Tesla).

C. EEG recording and preprocessing

The EEG signal was recorded by a g.tec (Guger Technologies OG, Austria) for 10 minutes with eyes closed using a cap with 64 scalp electrodes, placed according to the 10–20 electrode position system with a sampling rate of 500 Hz. The preprocessing of EEG signals was accomplished using the EEGLAB toolbox to remove artifacts caused by neck and shoulder movements, eye blinks, and hotspots. The baseline drift containing the low-frequency components is removed through a Butterworth Band-pass filter with low and high cut-off frequencies of 1 and 50 Hz, respectively. Next, Independent Component Analysis (ICA) technique is applied to the EEG signal to remove irrelevant components. The extracted ICs for EEG signals are depicted in Figure 2. The noisy Independent Components (IC) were labeled by applying a multiple artifact rejection algorithm (MARA) through EEGLAB [5]. Then, the noisy ICs were removed by inspecting the power spectrum of them and considering the labels. After all, the pruned EEG data was reconstructed. For each data, a length of 300 s of EEG signal was retained to equalize the length of all preprocessed data. Also, resting EEG was recorded at baseline at end of the treatment, and the participants were requested to avoid falling asleep.

D. Feature extraction

A total number of 26 features are categorized into five groups, including nonlinear, spectral, bispectral, cordance, and time-frequency. These measures are extracted from the baseline EEG of both groups of responders (R) and non-resposnders (NR). Each feature (except cordance measures) is computed for all EEG channels. The studied measures are described in the following [5]. Nonlinear features applied to the EEG signals includes LZC, KFD, and CD. The power spectrum indicates the power of the signal in its frequency components, so the power of EEG signals was estimated in delta, theta, alpha, and beta frequency bands. The Bispectrum analysis of a signal returns a 2D mapping of the level of interaction between all frequency pairs in the desired band of signal. Quantitative features related to moments and entropy must be extracted to apply bispectrum, so as to characterize and compare time series [5], [13]. Moreover, cordance is a measure of regional brain activity, which is computed using QEEG measures of brain wave patterns in an algorithm developed at the UCLA Laboratory of Brain, Behavior, and Pharmacology [14]. Lastly, another approach to analyze non-stationary signals is Time-frequuncy (TF) analysis. TF methods are mainly categorized into three main categories: (1) Nonparametric linear TF methods (based on linear filtering, including the short-time Fourier transform, and the wavelet transform), (2) Nonparametric quadratic TF representations (such as the Wigner-Ville distribution and its filtered versions), and (3) Parametric time-varying methods based on autoregressive models with time-varying coefficients [15].

E. Local subset feature selection (LSFS)

A crucial step in machine learning approaches has been feature selection process. Filtering out the redundant features with low capability for discriminating, especially When dealing with large number of features, could feasibly decrease the computational time, and increase the accuracy of the system subsequently. Most of the previous studies have executed global feature selection methods, which is not necessarily served as an optimal strategy when a single subset of features were selected to apply over all regions of the sample space [16]. There would be circumstances that some features are more informative for classification in various parts of the sample space [17]. In this study, the local subset feature selection is utilized with defining a sets of selected features in subsets. Besides increasing the prediction accuracy, this method emphasize the type of dominant features. Basically, the proposed method is based on the turning feature subset selection procedure into a sequential decision-making problem. In order to create a unified model for the proposed method, the concept of decision tree is developed to a unified model, known as the feature tree. Consequently, by applying the concept of feature tree, the sample space is divided into localities and corresponding features are assigned to them. To form the feature tree, three types of nodes are designed, named splitting nodes, leaf nodes, and features nodes. The former contains a feature and a threshold, representing a split in the sample space and so having two children. Differently, the leaf nodes assigned to only one locality. Additionally, a feature nodes represents a feature that is attributed to all of its decedent localities, and may have one child at maximum. In this model, the concept of a compound locality refers to each sub-tree corresponding to a set of neighbor localities. So, when neighbour localities tend to share the features, the model mentioned above simplifies the selection of identical features for them. Notably, the mutual features of neighbor localities are factored together in the parent feature node [18]. The process of developing LSFS based on feature tree concept was designed referring to previous works [17], [19]. Notwithstanding, local feature selection may increase the risk of overfitting. As the features are selected on a local basis from a limited number of samples, a noise feature would get higher possibility to appear in the selected combination compared to to global feature selection process. This issue is more likely in case of large number of features.

F. Feature classification

To classify the groups of R and NR, a Support Vector Machine (SVM) as a well-known supervised learning model, was utilized. The methodology of SVM is based upon finding a hyper plane that separates the data related to both classes with the maximum possible margin. The regression is implemented in order to to determine the best model from a set of models (Estimating Functions) to approximate the future values accurately. The generic support vector regression estimating function is [19],

$$f(x) = (w.\Phi(x)) + b \tag{1}$$

Where $w \subset R_n$ and $b \subset R$. Φ is a nonlinear function that maps x into a higher dimensional space. W and b are the weight vector and bias, respectively. The weight vector (w) is written as,

$$w = \sum_{i=1}^{L}(\alpha_i - \alpha_i^*) \tag{2}$$

By substituting eq. (1) into eq. (2), the generic equation can be rewritten as,

$$f(x) = \sum_{i=1}^{L}(\alpha_i - \alpha_i^*)(\Phi(x_i).\Phi(x)) + b \tag{3}$$

$$f(x) = \sum_{i=1}^{L}(\alpha_i - \alpha_i^*)\ k(x_i.x) + b \tag{4}$$

In eq. (3), the kernel function $k(x_i.x) = (\Phi(x_i).\Phi(x))$ is replaced with the dot product. $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_l)$ is the vector of non-negative Lagrange multipliers [19]. It is noteworthy that a Leave-One-Out cross-validation method is applied on account of the input data limitation so as to evaluate the performance of the classification. Accordingly, at each iteration, one of the subjects is selected for testing and the remaining (35 subjects) as the training data. This process repeated 36 times through all subjects, so that each subject was chosen as the test subject once. The overall capability of this classification procedure is the average of those 36 repetitions, meanwhile, the network error was calculated for each step. The response prediction was evaluated by its achieved accuracy, specificity, sensitivity. Accuracy indicates the ratio of correct predictions to the total predictions. Specificity is the ratio of number of correctly predicted as NR (true negatives) to total number of NRs (total negatives), and sensitivity refers to the ratio of number of correctly predicted as R (true positives) to total number of Rs (total positives).

III. Results and discussion

The ability of the proposed method was evaluated using accuracy (AC), sensitivity (SN), Specificity (SP). In the following equations (5) to (7), accuracy indicates the ratio of correct predictions to the total predictions. Specificity is the ratio of number of correctly predicted as NR (TN: true negatives) to total number of NRs (total negatives), and sensitivity refers to the ratio of number of correctly predicted as R (TP: true positives) to total number of Rs (total positives). FN refers to false negatives (incorrectly predicted NR) and FP refers to false positives (incorrectly predicted R).
Accuracy (AC): The ratio of correct predictions to the total predictions.

$$AC = \frac{TP + TN}{TP + TN + FN + FP} \tag{5}$$

Sensitivity (SN): The ratio of true positives to the total positives.

$$SN = \frac{TP}{TP + FN} \tag{6}$$

Specificity (SP): The ratio of true negatives to the total negatives.

$$SP = \frac{TN}{TN + FP} \tag{7}$$

Consequently, the outputs of local subset feature selection (LSFS) method were fed into an SVM classifier in order to evaluate the capability of above-mentioned features in prediction of rTMS treatment response in depression cases. The achieved accuracy, sensitivity, and specificity are 89.5%, 90%, and 87% respectively. Furthermore, in order to compare the results of this studied method with previous studies that have applied various types of machine learning techniques for prediction of rTMS treatment response in MDD cases, their classification results are represented below in Table I. The reason that this proposed method could not outperform some of previous works is related to the assumption that a combination of computed measures as a feature set may decrease the efficiency of the system. Accordingly, searching for the most effective EEG features for prediction of rTMS treatment response should proceed in the future works. Moreover, a limitation of this study is due to small sample size; accordingly, this method cannot be applied confidently for the prediction of treatment in clinical applications, since the number of participants in this study was not sufficiently high.

TABLE I: Comparison of classification results of different studies that applied machine learning approaches for prediction of rTMS treatment response for MDD

Study	Accuracy (%)	Sensitivity (%)	Specificity(%)
Bailey et al. [9]	86.6	89	84
Bailey et al. [8]	91	92	91
Khodayari et al. [10]	88	81	95
Mumtaz et al. [11]	87.5	95	80
Proposed method	89.5	90	87

IV. Conclusion and future works

In this research, an algorithm for the prediction of rTMS therapy response in MDD patients is proposed. This approach is based on the local subset feature selection method by employing a combined set of features acquired from five groups of features, including nonlinear, spectral, bispectral, cordance, and time-frequency. The results of the proposed method to predict responding to rTMS therapy in major depressive disorder shows accuracy of 89.5%, sensitivity of 90%, and specificity of 87%. For further studies, other classifiers such as k-Nearest Neighbor (kNN), Multilayer Perceptron (MLP), and Mixture of Experts (ME) can be considered to investigate their outcome in this subject. Additionally, it would not be unlikely that a combination of measures may decrease the accuracy of the system; consequently, selecting specific sets of features from five groups of features mentioned previously rather than the combination of computed features could be taken into account in future studies.

References

[1] "Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017," *The Lancet*, vol. 392, no. 10159, p. 1789–1858, Nov 2018.

[2] C. Otte, S. M. Gold, B. W. Penninx, C. M. Pariante, A. Etkin, M. Fava, D. C. Mohr, and A. F. Schatzberg, "Major depressive disorder," *Nature Reviews Disease Primers*, vol. 2, no. 1, 2016.

[3] C. Baeken, A.-K. Brem, M. Arns, A. R. Brunoni, I. Filipčić, A. Ganho-Ávila, B. Langguth, F. Padberg, E. Poulet, F. Rachid, and et al., "Repetitive transcranial magnetic stimulation treatment for depressive disorders," *Current Opinion in Psychiatry*, vol. 32, no. 5, p. 409–415, 2019.

[4] J.-P. Lefaucheur, N. André-Obadia, A. Antal, S. S. Ayache, C. Baeken, D. H. Benninger, R. M. Cantello, M. Cincotta, M. D. Carvalho, D. D. Ridder, and et al., "Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rtms)," *Clinical Neurophysiology*, vol. 125, no. 11, p. 2150–2206, 2014.

[5] F. Hasanzadeh, M. Mohebbi, and R. Rostami, "Prediction of rtms treatment response in major depressive disorder using machine learning techniques and nonlinear features of eeg signal," *Journal of Affective Disorders*, vol. 256, p. 132–142, 2019.

[6] A. Iznak, E. Iznak, E. Damyanovich, I. Oleichik, P. Bologov, I. Kazachinskaya, and T. Medvedeva, "Transcranial magnetic stimulation in combined treatment of pharmacoresistant depression: Dynamics of clinical, psychological, and eeg parameters," *Human Physiology*, vol. 41, pp. 503–509, 09 2015.

[7] S. M. Mcclintock, I. M. Reti, L. L. Carpenter, W. M. Mcdonald, M. Dubin, S. F. Taylor, I. A. Cook, J. O'Reardon, M. M. Husain, C. Wall, and et al., "Consensus recommendations for the clinical application of repetitive transcranial magnetic stimulation (rtms) in the treatment of depression," *The Journal of Clinical Psychiatry*, vol. 79, no. 1, p. 35–48, 2018.

[8] N. Bailey, K. Hoy, N. Rogasch, R. Thomson, S. Mcqueen, D. Elliot, C. Sullivan, B. Fulcher, P. Fitzgerald, and et al., "Responders to rtms for depression show increased fronto-midline theta and theta connectivity compared to non-responders," *Brain Stimulation*, vol. 11, no. 1, p. 190–203, 2018.

[9] N. Bailey, K. Hoy, N. Rogasch, R. Thomson, S. Mcqueen, D. Elliot, C. Sullivan, B. Fulcher, Z. Daskalakis, P. Fitzgerald, and et al., "Differentiating responders and non-responders to rtms treatment for depression after one week using resting eeg connectivity measures," *Journal of Affective Disorders*, vol. 242, p. 68–79, 2019.

[10] A. Khodayari-Rostamabad, J. P. Reilly, G. M. Hasey, H. D. Bruin, and D. J. Maccrimmon, "A machine learning approach using eeg data to predict response to ssri treatment for major depressive disorder," *Clinical Neurophysiology*, vol. 124, no. 10, p. 1975–1985, 2013.

[11] W. Mumtaz, L. Xia, M. A. M. Yasin, S. S. A. Ali, and A. S. Malik, "A wavelet-based technique to predict treatment outcome for major depressive disorder," *Plos One*, vol. 12, no. 2, Feb 2017.

[12] P. Vlcek, M. Bares, T. Novak, and M. Brunovsky, "Baseline difference in quantitative electroencephalography variables between responders and non-responders to low-frequency repetitive transcranial magnetic stimulation in depression," *Frontiers in Psychiatry*, vol. 11, 2020.

[13] E. B. Assi, L. Gagliano, S. Rihana, D. K. Nguyen, and M. Sawan, "Bispectrum features and multilayer perceptron classifier to enhance seizure prediction," *Scientific Reports*, vol. 8, no. 1, 2018.

[14] A. F. Leuchter, I. A. Cook, R. B. Lufkin, J. Dunkin, T. F. Newton, J. L. Cummings, J. Mackey, and D. O. Walter, "Cordance: A new method for assessment of cerebral perfusion and metabolism using quantitative electroencephalography," *NeuroImage*, vol. 1, no. 3, p. 208–219, 1994.

[15] E. Ebrahimzadeh, M. Pooyan, and A. Bijar, "A novel approach to predict sudden cardiac death (scd) using nonlinear and time-frequency analyses from hrv signals," *PLoS ONE*, vol. 9, no. 2, Apr 2014.

[16] E. Ebrahimzadeh, M. S. Manuchehri, S. Amoozegar, B. N. Araabi, and H. Soltanian-Zadeh, "A time local subset feature selection for prediction of sudden cardiac death from ecg signal," *Medical Biological Engineering Computing*, vol. 56, no. 7, p. 1253–1270, 2017.

[17] M.-H. Z. Ashtiani, M. N. Ahmadabadi, and B. N. Araabi, "Bandit-based local feature subset selection," *Neurocomputing*, vol. 138, p. 371–382, 2014.

[18] M. Z. Soroush, K. Maghooli, S. K. Setarehdan, and A. M. Nasrabadi, "A novel approach to emotion recognition using local subset feature selection and modified dempster-shafer theory," *Behavioral and Brain Functions*, vol. 14, no. 1, 2018.

[19] E. Ebrahimzadeh, A. Foroutan, M. Shams, R. Baradaran, L. Rajabion, M. Joulani, and F. Fayaz, "An optimal strategy for prediction of sudden cardiac death through a pioneering feature-selection approach from hrv signal," *Computer methods and programs in biomedicine*, vol. 169, pp. 19–36, 2019.