Basic and Clinical
NEUROSCIENCE
November & December 2024, Vol 15, No. 6

Research Paper

Brain Activity Flow and Machine Learning for Predicting Drug Response in Patients With Major Depressive Disorder

Seyed Morteza Mirjebreili[1], Reza Shalbaf [1], Ahmad Shalbaf [2*]

1. Institute for Cognitive Science Studies, Tehran, Iran.
2. Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Use your device to scan and read the article online

Citation Mirjebreili, S. M., Shalbaf, R., & Shalbaf, A. (2024). Brain Activity Flow and Machine Learning for Predicting Drug Response in Patients With Major Depressive Disorder. *Basic and Clinical Neuroscience, 15*(6), 775-794. http://dx.doi.org/10.32598/bcn.2024.2034.6

http://dx.doi.org/10.32598/bcn.2024.2034.6

Article info:
Received: 20 Apr 2024
First Revision: 15 Oct 2024
Accepted: 26 Oct 2024
Available Online: 01 Nov 2024

Keywords:
Electroencephalogram (EEG), Effective connectivity, Major depressive disorder (MDD), Machine learning (ML)

ABSTRACT

Introduction: A major challenge today is personalizing the treatment for patients with major depressive disorder (MDD) to make it more efficient. To address this issue, we have proposed a novel approach based on machine learning (ML) models that utilize neural activity flow prior to treatment with selective serotonin reuptake inhibitor (SSRI) medication.

Methods: The electroencephalogram signals of 30 patients were used to calculate the neural activity flow of each patient using the direct directed transfer function (dDTF). Then, based on the area under the curve (AUC) values, 30 important connections were identified for the delta, theta, alpha, beta, and gamma bands. To select the most critical neural activity flow, these neural activity flows were combined, and forward features, mRMR, and ReliefF methods were applied. Support vector machines (SVMs), decision tree, and random forest models are trained using selected neural activity flows.

Results: Results showed that most connections originated from F8, Pz, T5, and P4, mainly from the frontal and parietal lobes. In addition, the SVM model showed 98% accuracy in classification using forward feature selection, where most of the neural activity flows were selected from alpha and beta. Finally, results indicate that patients who responded to treatment differed in their patterns of frontoparietal neural activity flows, implying that the frontoparietal network (FPN) is primarily involved in treatment response at alpha and beta frequencies.

Conclusion: Therefore, the proposed method can accurately detect responders in MDD patients. It can reduce costs for both patients and medical facilities.

*** Corresponding Author:**
Ahmad Shalbaf, Associate Professor.
Address: *Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran.*
E-mail: *shalbaf@sbmu.ac.ir*

Highlights

- The support vector machines (SVM) model achieved 98% accuracy in predicting response to SSRI medication in major depressive disorder (MDD) patients;
- Alpha and beta band neural activities were different between MDD patients who responded to selective serotonin reuptake inhibitor (SSRI) medication and non-responders;
- Frontal and parietal lobes dominate critical neural activity for MDD treatment prediction;
- Frontoparietal network connectivity at alpha and beta bands predicts the response to SSRI medication in MDD patients;
- Neural activity flow and feature selection improve EEG-based prediction of SSRI medication outcome for MDD.

Plain Language Summary

Depression is one of the most common psychological disorders, and finding the right treatment for it can be challenging. The use of antidepressant medications may not be effective for everyone, and it often takes weeks or months of trial and error to determine the right approach. This delay can lead to frustration, increased suffering, and higher healthcare costs. Our study aims to improve this process by using brain activity data to predict whether a depressed patient will respond to SSRIs. We recorded brain signals from patients using the EEG (electroencephalography) before treatment. We then used a machine learning approach to analyze the brain activity patterns and found key differences between those who responded well to SSRIs and those who did not. We discovered that the frontal and parietal regions (areas related to thinking and attention) had different activity patterns in patients who benefit from SSRIs. Our predictive model was highly accurate, correctly identified responders to treatment by 98%. This model can help doctors personalize treatments for depressed patients more quickly, reducing the time spent on ineffective medications. By predicting who is likely to respond to specific treatments, we can improve patient outcomes, reduce suffering, and lower healthcare costs.

1. Introduction

Major depressive disorder (MDD) is the most commonly diagnosed psychiatric disorder worldwide, affecting more than 300 million individuals (Organization, 2017). Symptoms of MDD include changes in mood, interests, pleasure, cognitive functions, and vegetative symptoms. Furthermore, MDD increases the risk of developing conditions such as diabetes mellitus, heart disease, and stroke. MDD has also been associated with suicide, which is approximately the reason for half of the 800000 suicides worldwide (Otte et al., 2016).

In the present day, several antidepressants that act on neurotransmitter receptors are being used to treat depression. Almost all drugs act on two or more neurotransmitter receptors, i.e. two serotonin receptors, two noradrenergic receptors, or both. Also, several treatments are being investigated, including estrogen replacement therapy, mifepristone (RU-486 or C-1073), as well as antagonists, such as corticotropin-releasing factor, neurokinins, and injectable pentapeptides (Sambunaris et al., 1997; Stahl & Grady, 2003). MDD is a highly heterogeneous disorder, meaning that only a few people find antidepressants effective. Several pretreatment variables have been found to moderate the treatment response, including depression severity and neuroticism, older age, less impairment in cognitive control, and employment (Cohen & DeRubeis, 2018; Webb et al., 2019).

In general, 40% of people suffering from MDD have treatment-resistant depression (TRD) since treatment of MDD requires a trial-and-error sequential treatment strategy, and first-line therapies do not meet their needs (Arteaga-Henríquez et al., 2019; Berlim et al., 2008; Leuchter et al., 2009). Often, MDD patients suffer from delayed treatment response, functional impairment, increased suicide risk, and high medical costs due to the inability to predict which treatment will work. Consequently, more effective treatment strategies for patients with MDD are urgently needed (Bremer et al., 2018; Goldman et al., 1999; Schwartz et al., 2021).

Because of advances in neuroimaging techniques, biomarkers from neuroimaging studies are of great value for achieving precision medicine for many psychiatric disorders (Kang & Cho, 2020). In recent years, neuroimaging studies have been published utilizing a variety of methods, including electroencephalogram (EEG), brain volumetric magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), and diffusion tensor imaging to identify biomarkers for treatment response to antidepressants (Kemp et al., 2008). EEGs can be an effective and relatively inexpensive method for studying developmental changes in brain-behavior relationships, and their high temporal resolution makes them particularly useful for examining neural activity flow in the nervous system (Bell & Cuevas, 2012; Smit et al., 2008).

Recently, many studies have focused on EEG to predict how a patient will respond to antidepressant medication (de la Salle et al., 2020; Jaworska et al., 2019; Khodayari-Rostamabad et al., 2013; Shahabi & Shalbaf, 2022). For example, patients who respond to treatment demonstrate improved absolute alpha power at baseline, which can be used as a biomarker to predict treatment response (Baskaran et al., 2018). Also, the interhemispheric neural activity flow in the temporal lobe exhibits 99.61% classification capability using only four EEG channels (Zhang et al., 2022). A study conducted by Mumtaz et al. (2017) involved extracting time-frequency features from different frequency bands of EEG signals and classifying them by using three time-frequency decomposition techniques, including wavelet transforms, short-time Fourier transforms, and empirical modes of decompositions, to predict treatment-outcome for MDD patients. Combining the best features from the decomposition methods described above provides a classification accuracy of 91.6% (Mumtaz et al., 2017). Likewise, Jaworska et al. utilized demographical features with EEG data to improve the classification results (Jaworska et al., 2019). As demonstrated by Salle et al., changes in theta cordance of the prefrontal and midline right frontal in the first week of treatment can provide a predictive indicator of the response to antidepressants (de la Salle et al., 2020). Additionally, Kautzky et al. used a random forest approach to correctly identify 25% of patients with TRD based on clinical variables and three polymorphisms (Kautzky et al., 2015). Moreover, Patel et al. predicted an 89% treatment response using various biometrics, including demographic information and structural and functional imaging features (Patel et al., 2015).

This research significantly contributes to predicting treatment outcomes in MDD through the innovative use of neural activity flow based on the direct directed transfer function (dDTF). Firstly, we demonstrate that neural activity flow features, particularly those derived from the dDTF, can accurately predict antidepressant response in MDD patients. It provides insight into differentiating between individuals who positively respond to selective serotonin reuptake inhibitors (SSRIs) and those who do not. Secondly, our work achieves new state-of-the-art accuracy in EEG-based prediction for MDD treatment by incorporating neural activity flow as a feature in machine learning (ML) models, including support vector machines (SVMs), linear discriminant analysis (LDA), decision trees (DT), and random forests (RF), surpassing existing benchmarks and enhancing the potential clinical applicability of our findings. Lastly, our analysis identifies specific brain regions and networks that indicate treatment failure in MDD, enriching our understanding of the neural underpinnings of treatment outcomes and offering critical insights for developing targeted interventions. In summary, our work presents a novel and comprehensive approach to predicting treatment outcomes in MDD, leveraging neural activity flow and ML models. Our contributions include the accurate prediction of antidepressant response, achieving state-of-the-art accuracy in EEG-based prediction, and identifying specific neural correlates of treatment failure, collectively representing a significant step forward in the field and providing valuable insights and tools for advancing personalized treatment strategies for individuals with MDD.

2. Materials and Methods

Dataset

The EEG signal dataset used in this study was provided by Mumtaz et al. (2017). The data sets included 34 MDD patients and 30 healthy individuals. Among the 34 MDD patients, 17 men and 17 women have a mean age of 40.3±12.9 years. For the eyes closed condition, only 30 EEG segments, of 19 channels each, were available, which were used in this study. MDD patients were diagnosed using the diagnostic and statistical manual of mental disorders, fourth edition (DSM-IV) criteria (GUZE, 1995). An MDD patient was treated for four weeks with SSRI antidepressants. If there is a 50% improvement from pre- to post-treatment, the MDD patient is considered a responder; otherwise, the subject is considered a nonresponder. Based on the Beck depression inventory (BDI), 12 patients responded to treatment, while 18 patients showed no significant improvement (Table 1).

The 10-20 electrode placement system records the EEG for 5 minutes using a 19-electrode EEG cap with linked-ear references. Five different brain regions are represented by electrodes: The frontal lobe containing Fp1, F3, F7, Fz, Fp2, F4, and F8; the parietal lobe containing P3, Pz, and P4; the occipital lobe containing O1 and O2; the left and right temporal lobe containing T3, T4, T5, T6 electrodes; and finally the central lobe with C3, C4, and Cz electrodes.

EEG preprocessing

To prevent erroneous subsequent analysis and ensure that the underlying neuronal activity is accurately reflected in the data, the preprocessing steps have been carried out using the EEGLAB open-source toolbox. To remove baseline drift, a 1-Hz high-pass filter is first applied. Then, the CleanLine open-source plugin is used to remove line noise. Lastly, 3 minutes of data are used for further analysis.

Effective connectivity

The concept of effective connectivity or neural activity flow refers to the influence a node has over another based on a model of neuronal integration, which identifies neuronal coupling mechanisms (Liu et al., 2017). Among the first models used to establish causality between two time series is Granger causality, which was introduced in economics. As explained by Granger causality, a time series of X_1 causes a time series X_2, if knowledge of X_1 helps to make predictions of X_2 more accurate (Granger, 1969). A measure of brain activity associated with Granger causality is the DTF. This function represents a linear combination of causal influences along all causal pathways, direct and indirect, beginning at one site and ending at another (Korzeniewska et al., 2008). To distinguish direct from indirect flows, a dDTF is proposed (Korzeniewska et al., 2003). The dDTF method determines the strength and direction of the direct flow of neural activity using DTF combined with partial coherence. To calculate the dDTF, the source information flow toolbox (SIFT) is used (Delorme et al., 2011; Mullen, 2014). Through SIFT, each subject's EEG data was divided into 18 segments, each lasting 10 seconds. A multivariate autoregressive (MVAR) model of order 20 was then fitted to the data, satisfying two criteria of stability and consistency. This process indicates that the model produces data with the same correlation structure as the actual EEG data and is stable/stationary. This step is crucial in ensuring the accuracy and reliability of subsequent analyses. Then, dDTF values were calculated for each data segment across all frequency ranges, and since we had 19 electrodes, a matrix with the shape of 19*19* frequency was obtained. At the end of the process, the delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-45 Hz) bands were extracted.

To provide an overview of the dDTF method, the EEG signal is first fitted with an MVAR model. Then, to model a k-channel process, X(t) is modeled as follows (Equation 1):

1. $X(t)=(X_1(t), X_2(t),..., X_k(t))$

This equation would lead to the following expression for the MVAR model:

2. $X(t)=\sum_{j=1}^{p} A(j)X(t-j)+E(t)$

In the above Equation, X(t) represents the data vector in time t, E(t) represents the white noise vector, A(i) represents the model coefficients, and p represents the order of the model. After that, as a result of converting the model Equation into a frequency domain, we obtain:

3. $X(f)=A^{-1}(f)E(f)=H(f)E(f)$

In the above Equation, X(f) is the input signal, E(f) is white noise, and the H(f) matrix is referred to as the transfer matrix of the system, in which f denotes the frequency of the input signal. The DTF can be defined as follows in accordance with the transfer function of MVAR:

4. $DTF^2_{j\to i}(f)=\frac{|H_{ij}(f)|^2}{\sum_{m=1}^{k}|H_{im}(f)|^2}$

For the dDTF formula, the DTF must be modified with partial coherence as follows:

5. S(f)=H(f)VH*(f)

6. $pCoh^2_{ij}(f)=\frac{M^2_{ij}(f)}{M_{jj}(f)M_{ii}(f)}$

where S(f) is power spectra, V is the variance of the noise E(f) and M_{ij} is spectral matrix S by removing the ith row and jth column. Finally, dDTF is defined by the given formula:

7. $dDTF_{ij}(f)=DTF_{ij}(f)\ pCoh_{ij}(f)$

Feature selection

To select the best features for discriminating between responder and nonresponder groups, one-seventh of the data was set aside for testing, and then LDA was used

to calculate the area under a curve (AUC) of every neural activity flow in each band. This study uses the AUC since the area under the receiver operating characteristic curve (AUC-ROC) is equivalent to the Mann–Whitney U statistic (Mason & Graham, 2002). ROC curves were calculated by comparing the model's false-positive rate against its true-positive rate across a range of thresholds. AUC-ROC value was obtained based on the mean values of all CV sets. The mean AUCs are a valid measure of the model's performance in a generalized setting in which the model was trained, given that each of the analyzed learners received a unique training set and a unique model-external validation dataset during training. Next, the top 30 connections from each band with the highest AUC are selected; then, the feature selection algorithms are applied.

This paper uses three feature selection methods. The first is based on the area AUC-ROC. In this method, a subset of features is assessed empirically by measuring the prediction accuracy of the feature subset selected by our method. In other words, the forward selection is an iterative process in which we start without any features in the model. We continue to add new features to our model in each iteration, and then we select the subset of features with the highest accuracy out of all the others (Mamitsuka, 2006). Second, the minimum-redundancy maximum-relevance (mRMR) algorithm has been used to rank features to minimize redundancy while maximizing relevance. The mRMR algorithm uses mutual information to compute similarity scores between features and labels of a subset, aiming to minimize the mean mutual information between two features and to maximize the mean mutual information between each feature and the specific label (Amini et al., 2023; Ding & Peng, 2005; Şen et al., 2014). As for the last method, ReliefF is used. Like k-nearest neighbors, ReliefF assigns weights to each feature based on its ability to separate class labels. If the squared Euclidean distance between a feature and its nearest instances of the same class is greater than the distance between the two instances of the other class, the weight of the feature decreases. Based on the Manhattan distance, ReliefF calculates both negative and positive weights for each feature (Al-Nafjan, 2022; Peker et al., 2015).

Classification

In artificial intelligence, supervised learning refers to a subcategory of ML that uses labeled datasets to train algorithms capable of classifying data or predicting outcomes. SVMs are supervised learning algorithms used to classify two data groups. The algorithms draw lines (hyperplanes) to separate groups based on their patterns. An SVM builds a learning model that assigns new examples to one group or another. As a result of these functions, SVMs are called non-probabilistic binary linear classifiers. This paper also uses a random forest composed of many individual decision trees. Trees in the forest generate a class prediction, and the class with the most votes determines the class prediction for our model. In addition, the LDA classifier is used to classify two groups using a linear combination of features.

Statistical analysis

The AUC value was calculated for each neural activity flow to evaluate the importance of each neural activity flow. Afterward, the 30 top connections with the highest value were selected. Also, to evaluate any ML model's performance, we need to test it on some unseen data. Based on the model's performance on unseen data, we could say whether our model is under-fitting, over-fitting, or well-generalized. A cross-validation (CV) procedure is used to assess the effectiveness of ML models; it can also be used to evaluate a model if we have insufficient data. For a CV to be performed, some training data must be kept aside for evaluation later. In this paper, the k-fold method was used for CV. During k-fold cross-validation, the original sample is divided into k subsamples of equal size. A single subsample of the k subsamples was retained as the validation data for testing the model, while the remaining k-1 subsamples were used as training data for training the model. In the process of trial and error, it was determined that 7 is the optimal value for k. Further analysis was conducted based on the results of the 7-fold CV.

Overview of the proposed method

Figure 1 illustrates the proposed method. In the first step, raw EEG data were preprocessed using EEG-Lab, an open-source toolbox. A high-pass filter with a 1-Hz frequency and CleanLine noise were applied as part of the preprocessing steps. Afterward, the signals were divided into 18 segments, each lasting 10 seconds. The neural activity flow was calculated from each segment, and a matrix of 19-channels * 19-channels * 45-frequency steps was obtained. Then, the delta, theta, alpha, beta, and gamma bands are extracted by averaging over the frequency ranges of 1-4 Hz, 4-8 Hz, 8-13 Hz, 13-30 Hz, and 30-45 Hz. Following this, AUC-ROC forward feature selection, mRMR, and Relief-F algorithms were used to find the best features from all frequency bands. The selected features trained SVM, LDA, RF, and DT classifiers.

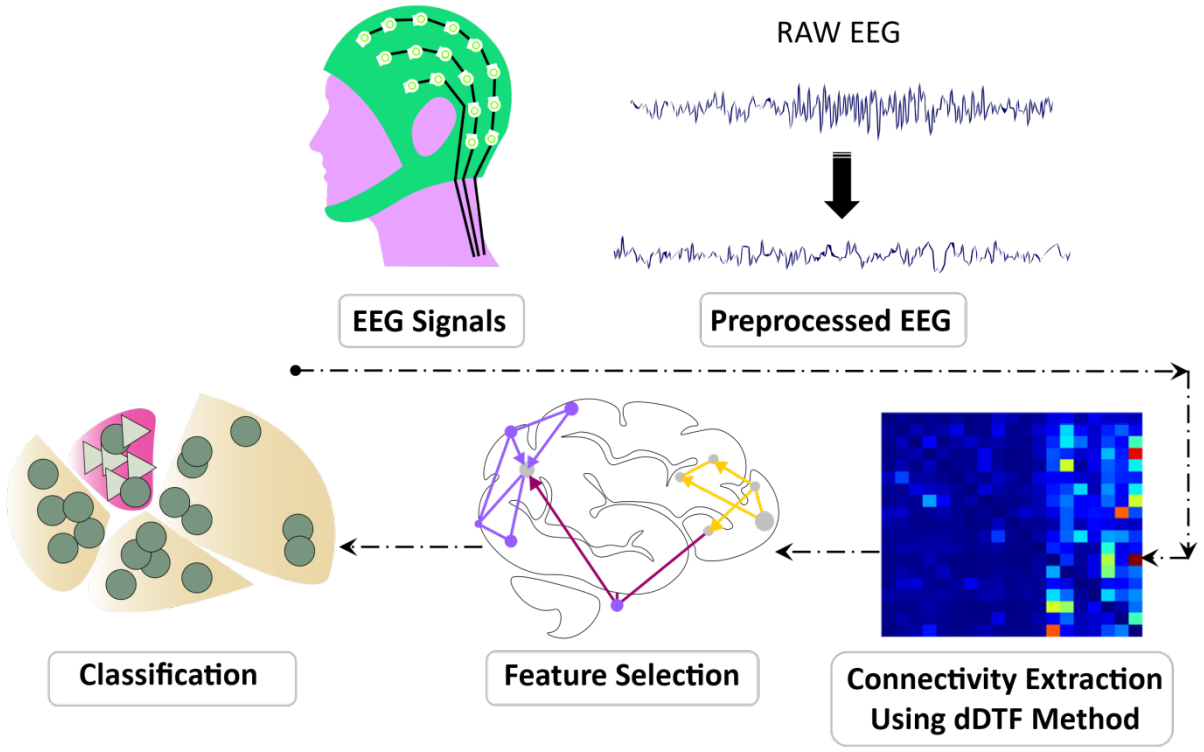
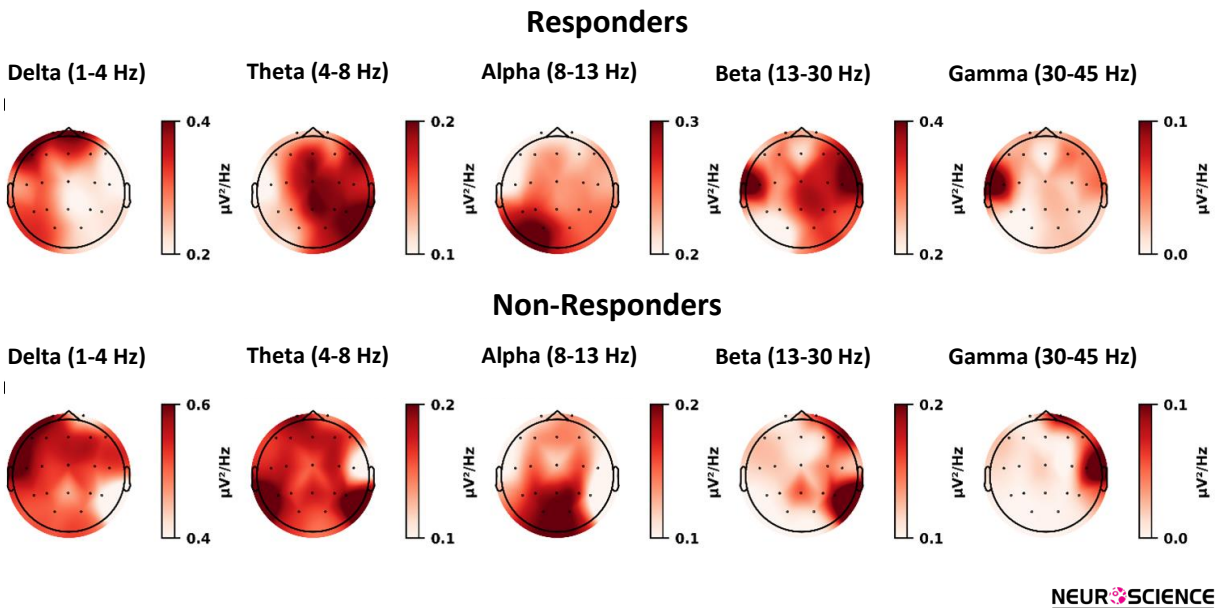


Figure 1. An EEG signal recorded in a 10-20 system and preprocessed

NEUROSCIENCE

EEG: Electroencephalogram; dDTF: Direct directed transfer function.

Notes: The dDTF matrix for the delta, theta, alpha, beta, and gamma frequency bands is calculated in the following step. Afterward, several feature selection methods are used to select the best neural activity flows. In the final step, the selected features are used to perform classification.



NEUROSCIENCE

Figure 2. The mean power spectral density (PSD) of responders and nonresponders in the delta, theta, alpha, beta and gamma bands

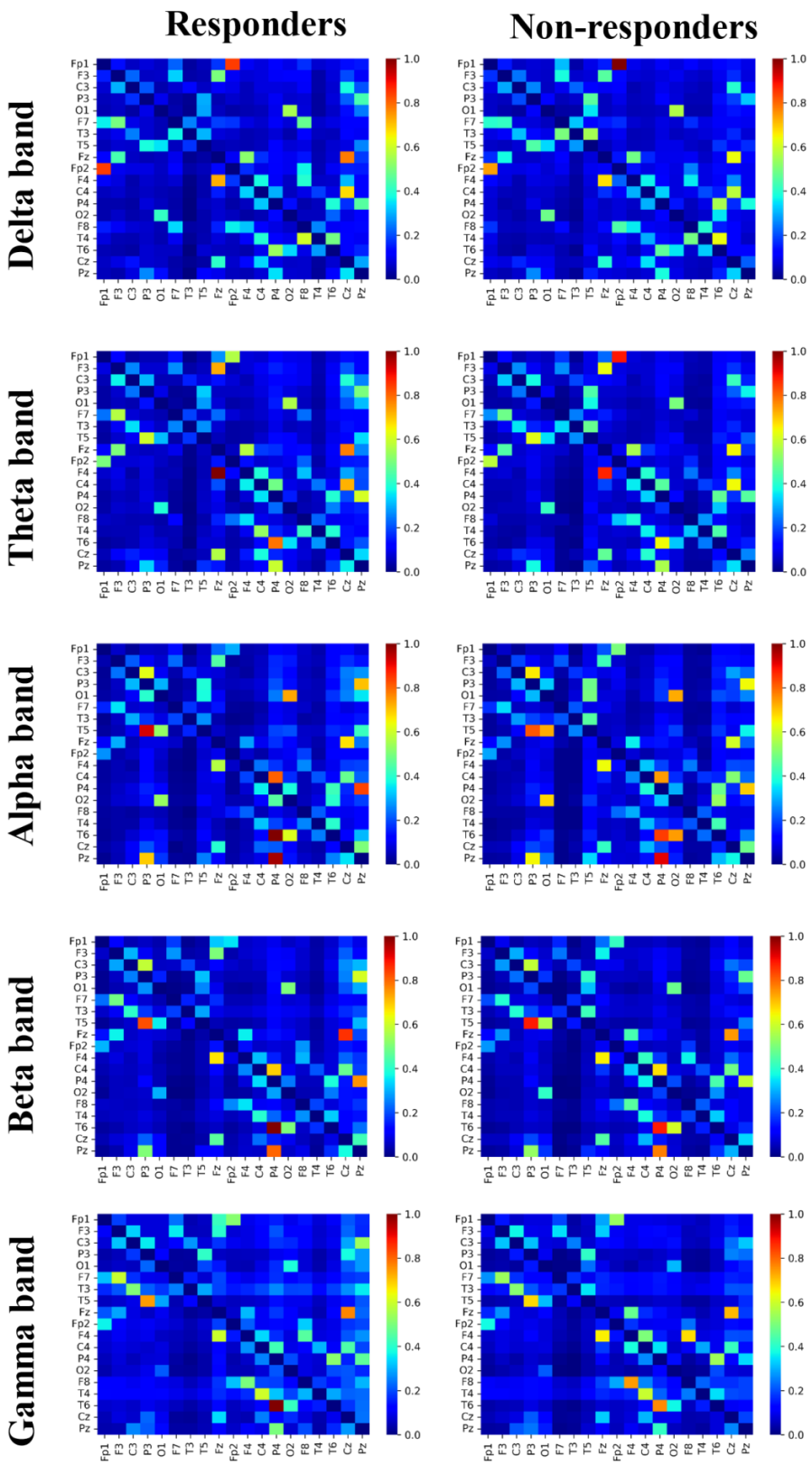


Figure 3. Normalized neural activity flow (dDTF) in responders vs nonresponders across bands

MDD: Major depressive disorder; dDTF: Direct directed transfer function.

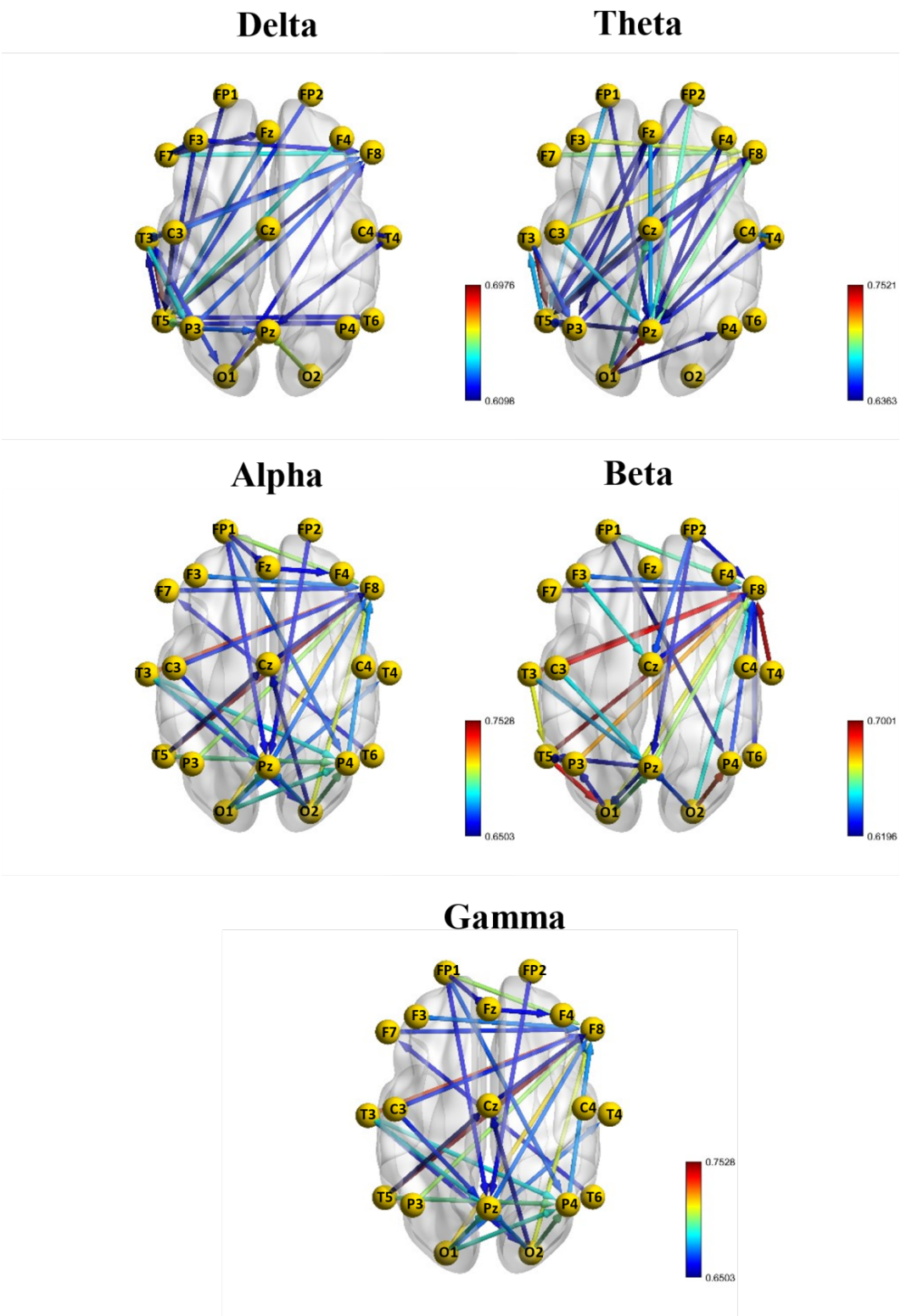


Figure 4. Based on AUC values, the top 30 neural activity flows showing differences in propagation between responders and nonresponders

AUC: Area under curve.

Notes: Nodes represent electors in the 10-20 system, and edges represent AUC values.

Table 1. Summary of MDD patients' clinical characteristics

Information	Mean±SD		
	Responder	Nonresponder	Total
Age (y)	40.7±13	41.1±12.5	40.3±12.9
Gender (female/male)	8/8	9/9	17/17
Pretreatment BDI-II	18.4±7. 4	22.8±12.5	20.6±8.6
Post-treatment BDI-II	9.1±6.3	22.1±3.3	15.6±4.5

BDI: Beck depression inventory.

NEUROSCIENCE

3. Results

Topographic maps

Figure 2 shows topographic maps for MDD patients in the five frequency ranges of delta, theta, alpha, beta, and gamma. It is clear from this figure that the responder group shows lower delta power compared to the nonresponder group, whereas the responders show higher beta power. In the theta band, the most significant difference is observed in the left temporoparietal lobe, whereas in the alpha band, the most significant difference is observed in the central areas. As can be observed in the beta band, respondents generally showed higher power, particularly in the left temporal lobe, which is also observed in gamma.

Regional differences in responders versus nonresponders

After calculating each segment's neural activity flow matrix, the mean values in responders and nonresponder groups were calculated (Figure 3). Then, based on the dDTF values corresponding to each directed connection,

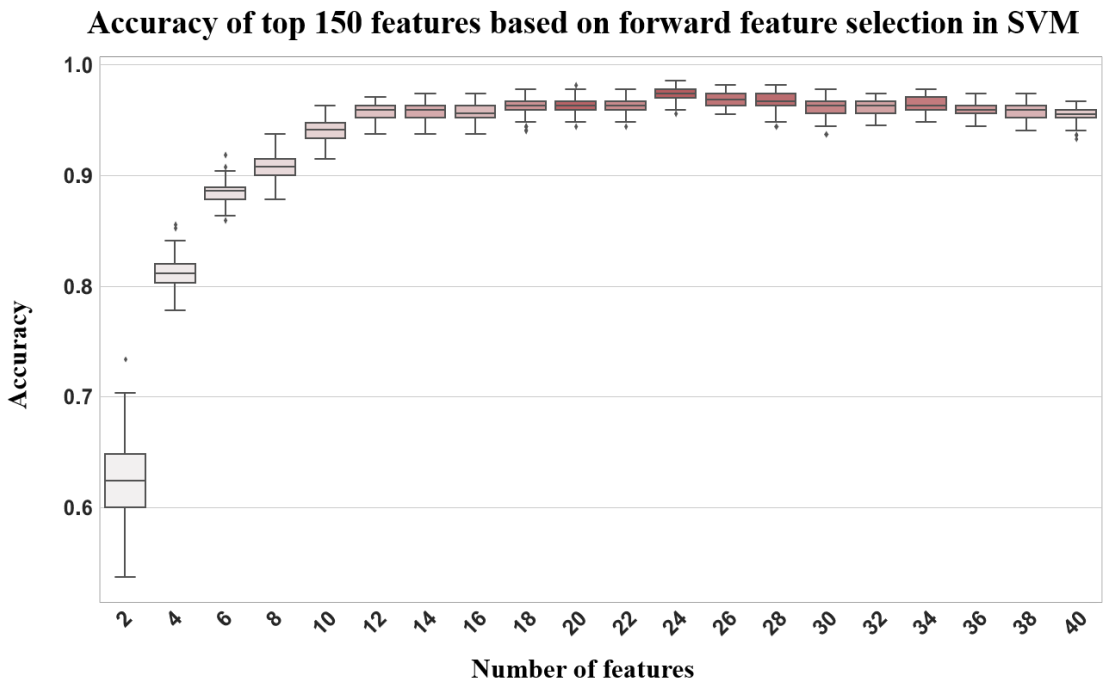


Figure 5. SVM forward feature selection based

NEUROSCIENCE

SVM: Support vector machine.

NOTES: The x-axis indicates the number of features used for classification, and the y-axis indicates the model's accuracy. With 23 features, the best accuracy was obtained, and since there was no improvement in accuracy, the x-axis is restricted to 40 features.

Table 2. Top 30 AUC values of neural activity flows for responders vs nonresponders by band

Delta			Theta			Alpha			Beta			Gamma		
From	To	AUC	From	To	AUC	From	To	AUC	From	To	AUC	From	To	AUC
T3	T5	0.698	T3	T5	0.752	T3	T5	0.763	T5	F8	0.7	T5	F8	0.753
O1	Pz	0.668	O1	Pz	0.739	T5	T3	0.68	T4	F8	0.698	T3	F8	0.733
O2	Pz	0.659	C3	F8	0.708	T6	F8	0.654	T3	F8	0.697	O1	F8	0.716
P3	T5	0.653	F3	F8	0.703	F8	C4	0.648	C3	F8	0.692	O2	F8	0.711
Cz	T5	0.652	F7	F8	0.694	Pz	F8	0.648	T5	O1	0.689	P3	F8	0.705
F4	T5	0.644	F8	Pz	0.691	O2	P4	0.647	O2	P4	0.685	Fp1	F8	0.704
F7	F8	0.643	Fp2	Pz	0.687	T3	O1	0.646	P3	F8	0.677	O2	P4	0.699
T3	P3	0.64	O1	Cz	0.686	T5	O1	0.642	T3	T5	0.668	T5	P4	0.697
Fz	T5	0.632	C3	Pz	0.672	F7	F8	0.64	O1	F8	0.668	O1	P4	0.691
C3	F8	0.627	T5	T3	0.67	O1	Cz	0.639	Pz	F8	0.662	O1	Pz	0.691
C3	T3	0.627	Fp1	T5	0.67	F3	F8	0.637	O1	Pz	0.658	T3	P4	0.688
P3	Pz	0.627	Fz	Pz	0.667	Fp1	F8	0.636	Fp1	F8	0.656	T3	Pz	0.686
P3	F8	0.625	T4	C4	0.666	T6	P4	0.636	O2	F8	0.651	P4	F8	0.676
F7	F3	0.624	F4	T5	0.663	O1	T5	0.635	F3	Cz	0.65	O2	Pz	0.674
T3	O1	0.622	T3	P3	0.655	F4	C4	0.633	C3	Pz	0.647	F3	F8	0.672
Fp2	T5	0.622	P3	T5	0.653	Fz	F4	0.630	T3	Pz	0.643	Pz	F8	0.671
P3	T3	0.621	T4	Pz	0.651	O1	P4	0.628	F3	F8	0.637	O1	T4	0.671
F3	T5	0.621	P3	F8	0.648	F8	F4	0.626	O2	Pz	0.633	Fp1	P4	0.669
T4	C4	0.618	Fp2	T5	0.645	C4	F4	0.626	Fp2	Cz	0.633	C3	F8	0.665
O1	F8	0.618	O1	P4	0.645	Cz	C4	0.625	Pz	O1	0.633	F7	F8	0.665
F7	Fz	0.618	Fz	T5	0.643	Fp2	F8	0.625	P4	F8	0.633	C3	Pz	0.665
T5	T6	0.617	O1	F8	0.64	P3	F8	0.625	Pz	T5	0.629	Fp1	Pz	0.662
F3	F8	0.617	T5	Pz	0.64	O2	T5	0.624	O1	P3	0.627	Fz	F4	0.662
T6	T5	0.616	Fp1	Pz	0.639	O2	T6	0.623	Cz	F8	0.626	Fp1	Fz	0.662
T3	F8	0.615	Cz	F8	0.639	P4	Pz	0.622	Fp2	F8	0.626	T3	O2	0.661
T4	Pz	0.614	F8	T5	0.639	T5	Cz	0.622	T6	F8	0.626	T6	F7	0.659
T5	T3	0.612	F4	Pz	0.639	Cz	F4	0.621	F7	F8	0.624	Fp2	Pz	0.659
T5	F8	0.611	Cz	T5	0.637	F7	T3	0.621	P3	T5	0.624	O2	Cz	0.659
Fp1	T5	0.611	Fz	P3	0.637	Fp1	T4	0.621	Fp2	Pz	0.623	T5	Cz	0.655
C3	T5	0.61	C4	Pz	0.636	Cz	T3	0.62	Fp1	P4	0.62	Cz	F8	0.65

Table 3. An overview of electrodes in all frequency bands

Electrode	Number of Connections From	Electrode	Number of Connections to
O1	16	F8	47
T3	15	Pz	26
T5	13	T5	24
O2	11	P4	11
Fp1	11	Cz	7
P3	10	T3	7
C3	9	C4	5
Cz	8	F4	5
F7	8	O1	5
Fp2	8	P3	4
F3	7	Fz	2
Fz	6	T4	2
Pz	5	T6	2
T4	5	F3	1
T6	5	F7	1
F8	4	O2	1
F4	4		
P4	3		
C4	2		

NEUROSCIENCE

Notes: On the left side, there are electrode names and the number of connections that originated from the electrode, and on the right side, there are electrode names and the number of connections from which connections end.

the AUC values were calculated for each independently. Afterward, they were ranked according to their top 30 AUC values (Table 2 and Figure 4). Table 3 summarizes each region in terms of frequency to identify the most critical regions. According to this Table, the frontal lobe has the highest number of neural activity flows, followed by the temporal and parietal lobes. As a result, the frontal and parietal lobes were the dominant regions for most connections. Moreover, an overview of all electrodes from the top 30 bands shows that most connections end in specific regions, particularly in the electrodes F8, Pz, T5, and P4 (Table 4).

Classification responder based on the neural activity flow

Table 5 presents the classification results of ML models for each frequency band and a combination of features from all bands, consisting of 150 connections (30 connections from each band). The top 150 features had the highest accuracy, specificity, sensitivity, and F1-measure in every model. Following that, the beta and alpha bands yielded the highest results. As seen, most models had a higher specificity than sensitivity, indicating that the models were more capable of correctly identifying patients who would respond to the treatment.

Table 4. A summary of all frequency bands' connections

Region	From						To					
C (C3, C4, Cz)	D	T	A	B	G	Sum	D	T	A	B	G	Sum
	4	5	4	3	3	19	1	2	5	2	2	12
F (Fp1, Fp2, F3, F4, F7, F8, Fz)	D	T	A	B	G	Sum	D	T	A	B	G	Sum
	9	13	10	8	8	48	9	6	11	15	15	56
O (O1, O2)	D	T	A	B	G	Sum	D	T	A	B	G	Sum
	3	4	6	6	8	27	1	0	2	2	1	6
P (P3, P4, Pz)	D	T	A	B	G	Sum	D	T	A	B	G	Sum
	4	2	3	6	3	18	5	13	4	8	11	41
T (T3, T4, T5, T6)	D	T	A	B	G	Sum	D	T	A	B	G	Sum
	10	6	7	7	8	38	14	9	8	3	1	35

NEUROSCIENCE

Notes: Rows represent the number of connections within each region, with C, F, O, P, and T reflecting the central, frontal, occipital, parietal, and temporal lobes of the brain and columns D, T, A, B, and G correspond to delta theta, alpha, beta, and gamma respectively. In this Table, the "from" columns show the number of connections originating from the specified region, and the "to" column shows the number of features ending in the specified region.

One of the main problems with ML models is the curse of dimensionality, which means that the error becomes larger as the number of features increases. To overcome this problem, different feature selection algorithms have been used, including mRMR, ReliefF, and forward feature selection algorithms. Based on Table 6, forward feature selection achieved the highest accuracy. Figure 5 shows how the highest performance can be obtained using only 23 features, and that accuracy decreases afterward, and the selected features are shown in Table 7. Thus, forward feature selection based on the ROC-AUC algorithm improved classification accuracy when using the best subset of features.

4. Discussion

In this study, we demonstrated an ML approach using EEG-derived neural activity flows that accurately predict antidepressant response and provide neuroscientific insights into mechanisms of treatment outcomes. Precisely, we extracted dDTF effective connectivity biomarkers in MDD patients to capture differences between treatment responders and nonresponders across brain regions and frequencies. Our findings indicate that frontoparietal network (FPN) connectivity at alpha and beta bands underlies response failures, aligning with cognitive theories implicating this circuitry. By combining dDTF neural activity flow features with SVM classifiers, our model significantly improved predictive performance over previous state-of-the-art EEG methods, achieving over 98% accuracy. This method establishes functional connectivity as an informative biomarker for guiding antidepressant selection while elucidating network deficits linked to treatment resistance.

Based on the findings in Table 3, F8, Pz, T5, and P4 are the most important regions that differ between respondents and nonresponders. Further, neural activity flows were used to predict treatment outcomes and different feature selection algorithms were applied to improve classification results. As a result, by using forward feature selection across all frequency bands, the best accuracy of 98.52% was achieved by SMV (Table 6). Although SVMs perform better in very high-dimensional spaces and SVM models have generalizability in practice, the risk of overfitting is lower in SVMs. Also, for its parameters, a unique global optimum can be easily determined (Garcia et al., 2003; Subasi & Gursoy, 2010). Additionally, SVM is based on its kernel, and selecting the appropriate kernel function can resolve any complex problem. In this study, the best result was achieved by using the radial basic function. Also, there have recently been some studies that have criticized the reproducibility of AI methods because the evaluation methods may be incorrect, and many of them may suffer from data leakage or overfitting (Gibney, 2022). As discussed, we have

Table 5. A comparison of SVM, LDA, RF, and DT classification results in the delta, theta, alpha, beta, and gamma bands

Band	Model	Mean±SD			
		Model Performance Metrics (%)			
		Accuracy	Specificity	Sensitivity	F1-measure
Delta	SVM	84.81±7.33	87.43±8.81	73.25±14.16	79.04±10.45
	LDA	81.46±8.65	81.43±6.4	69.09±21.11	73.26±15.38
	RF	80.37±8	79.87±8.92	66.51±16.99	71.84±13.49
	DT	75.55±6.26	75.51±9.07	59.32±16.00	65.08±10.36
Theta	SVM	84.82±4.22	89.34±9.3	71.84±8.09	79.04±5.55
	LDA	81.84±3.72	82.43±5.79	70.87±11.33	75.38±5.25
	RF	82.96±2.35	91.01±8.5	64.88±8.84	74.92±4.44
	DT	74.09±4.27	66.92±5.71	68.19±9.83	67.34±6.86
Alpha	SVM	86.26±4.42	91.87±7.7	73.57±10.17	80.96±5.82
	LDA	78.85±4.22	81.27±6.95	62.32±7.42	70.09±5.09
	RF	87.06±7.68	97.86±3.49	69.66±18.57	79.88±13.38
	DT	77±8.98	74.18±13.12	68.78±10.72	70.69±10.05
Beta	SVM	87.04±3.87	89.85±4.77	77.38±8.89	82.71±4.41
	LDA	82.96±6.27	84.59±8.28	71±9.57	76.94±8.22
	RF	87.04±6.42	92.4±7.27	74.5±11.03	82.14±8.66
	DT	76.24±4.97	74.41±9.19	64.27±13.12	67.77±7.83
Gamma	SVM	77.45±7.23	76.68±11.79	62.4±12.63	68.29±11.3
	LDA	70.4±6.47	67.06±6.91	53.73±15.17	58.21±11
	RF	80.81±8.03	80.57±13.68	68.51±12.38	73.49±11.95
	DT	66.64±5.9	60.67±9.12	55.6±13.74	56.38±8.19
Top 150	SVM	93.35±5.96	96.43±6.56	87.84±12.68	91.22±8.33
	LDA	82.63±6.83	79.39±12.33	76.29±12.6	77.11±9.88
	RF	88.56±8.4	95.76±6.71	74.81±16.92	83.05±14.11
	DT	80.35±5.54	83.15±6.85	65.15±9.88	72.55±6.79

NEUROSCIENCE

Abbreviations: SVM: Support vector machine; LDA: Linear discriminant analysis; RF: Random forests; DT: Decision trees.

used CV to train and evaluate the model on the entire dataset to overcome these problems. By fitting the model in every step and estimating its performance independently in each fold of the CV procedure, we can identify problems such as overfitting or selection bias and learn how the model will generalize to an independent dataset since it gives an almost unbiased performance. Overfitting was also overcome by using different feature selection methods.

Table 3 outlines the starting and ending points of the top 30 activity flows across all frequency bands; only

Table 6. SVM, LDA, RF, and DT results using mRMR, ReliefF, and forward feature selection

Feature Selection Method	Model	Mean±SD			
		Model Performance Metrics (%)			
		Accuracy	Specificity	Sensitivity	F1-measure
mRMR	SVM	94.1±4.67	95.57±6.43	90.26±10.69	92.25±6.5
	LDA	88.19±4.88	85.65±6.78	85.74±9.86	85.31±6.2
	RF	91.84±5.09	97.04±3.45	82.5±12.98	88.55±8.24
	DT	84.07±5.22	83.16±7.25	76.53±12.36	79.08±7.21
ReliefF	SVM	94.83±4.73	97.47±6.17	90.48±11.16	93.22±6.57
	LDA	85.2±5.23	83.03±7.89	79.98±12.06	80.73±7.8
	RF	88.92±7.55	96.84±5.49	75.83±18.15	83.39±13.42
	DT	84.84±3.65	85.59±10.19	76.22±4.95	80.14±4.2
Forward feature selection	SVM	98.52±1.29	99.25±1.84	97.3±3.14	98.22±1.57
	LDA	90.78±3.78	91.94±8.07	85.29±9.01	87.92±5.04
	RF	96.66±1.18	100±0	91.58±3.05	95.58±1.69
	DT	89.98±4.47	95.02±4.97	79.95±10.67	86.33±5.8

NEUROSCIENCE

Abbreviations: SVM: Support vector machine; LDA: Linear discriminant analysis; RF: Random forests; DT: Decision trees.

a few areas have the greatest impact on treatment outcomes. First and foremost, the frontal lobe region, especially the F8, is dominant for most neural activity flows. Following that, the parietal (Pz and P4) and temporal (T5) lobes were the dominant regions. Additionally, the findings indicated that this pattern represents a valuable brain biomarker that could be used to assess the treatment response of MDD patients before they begin their treatments, thereby reducing costs and reducing the time spent on patients and medical centers. For example, F8 also has a high classification capability (Hasanzadeh et al., 2021). Also, it may be possible to predict the efficacy of SSRIs by analyzing frontal EEG recordings collected during the first week of treatment (Iosifescu et al., 2009).

Table 7 illustrates the SVM-selected features for obtaining the best accuracy. With only 23 features, 98.5% accuracy has been achieved, and most of the features were selected from beta and alpha bands. Further, beta and alpha bands produced the best results when each frequency was used separately to predict treatment outcomes. Whether we use the neural activity flows of each frequency band separately or combine all features in all frequency bands, the beta and alpha bands are better discriminators for predicting a person's response to treatment.

A flexible and coordinated modulation of cognitive and emotional processes is enabled by the FPN, composed of lateral prefrontal and posterior parietal cortices (Martens et al., 2021). It has been shown that the FPN is activated during externally focused attention and goal-oriented task performance. A defining characteristic of MDD is the deficiency in concentration, and cognitive theories suggest that impaired top-down regulation of aberrant emotional processing perpetuates a bias toward negative effects (Fischer et al., 2016). It has been shown that depression symptoms are associated with decreased neural activity flows between the FPN and other parts of the brain (Schultz et al., 2018). The study results show that the most significant difference between these groups can be traced to the frontal and parietal lobes (Table 4). Study results supporting our findings suggest that the beta frequency of default mode network-FPN might serve as a neural marker for reoccurring illness (Pizzagalli, 2011; Whitton et al., 2018). Further studies have shown that EEG beta power correlates with cortisol secretion and attentional processing, as seen in Figure 2, where higher

Table 7. The selected features using the forward feature selection algorithm using the SVM model

Band	AUC	To	From
Alpha	0.763	T5	T3
Alpha	0.636	P4	T6
Beta	0.689	O1	T5
Gamma	0.699	P4	O2
Beta	0.7	F8	T5
Delta	0.627	Pz	P3
Alpha	0.642	O1	T5
Beta	0.65	Cz	F3
Alpha	0.633	C4	F4
Beta	0.627	P3	O1
Alpha	0.646	O1	T3
Delta	0.622	O1	T3
Alpha	0.626	F4	C4
Beta	0.662	F8	Pz
Alpha	0.621	T3	F7
Beta	0.633	F8	P4
Theta	0.666	C4	T4
Beta	0.624	T5	P3
Beta	0.633	Cz	Fp2
Delta	0.611	F8	T5
Beta	0.677	F8	P3
Gamma	0.662	F4	Fz
Gamma	0.665	F8	F7

Notes: The features are neural activity flows based on the dDTF method in all frequency bands.

NEUROSCIENCE

beta power is observed in the responding group, mainly in regions related to FPN. Moreover, Baskarana et al. reported that changes in beta asymmetry observed at 2 weeks post-treatment in the responding group may reflect differences in arousal induced by antidepressants (Baskaran et al., 2018). Finally, our findings suggest that treatment failure results from alpha and beta-frequency frontoparietal networks at the network level.

Although this simple approach to dDTF analysis enabled the classification of the treatment responses with 98% accuracy, it might not allow inference about the neural activity flow of particular frequency bands (delta, theta, alpha, beta, gamma, etc.) within the neural system. To ensure a good representation of low-amplitude (higher-frequency) rhythms in MVAR modeling, the analysis should be done step by step by filtering out high-amplitude (low-frequency) rhythm(s) (by high-pass filter with a gentle slope, not steep) and by fitting a separate MVAR model to the filtered signals.

Table 8. Comparing our work with other existing works on the prediction of antidepressant treatment outcome from eeg signals

Authors, Year	Feature	Classifier	Accuracy
Mumtaz et al., 2017	Combination of wavelets + STFT + EMD features	Logistic regression	91.6%
Jaworska et al.,2019	Demographic data, EEG power features, source localized current density	RF	88%
Zhdanov et al., 2020	Power spectral, spatiotemporal complexity	SVM	82.4%
de la Salle et al., 2020	Change in Montgomery Åsberg depression rating scale, change in prefrontal cordance	Logistic regression	85%
Khodayari-Rostamabad et al., 2013	Power spectral density, magnitude squared spectral coherence, mutual information, the log ratio of left-to-right hemisphere powers, and anterior/posterior log power ratios	Mixture of factor analysis classifier	87.9%
Patel et al., 2015	Demographical, cognitive ability, functional imaging, structural Imaging	AdTree	89.47%
Van der Vinne et al., 2019	EEG Abnormality (diffuse slowing, focal slowing, paroxysmal and nonparoxysmal activity)	ANOVA	Response rate=74%
Our work (2022)	Effective connectivity	SVM	98.52%

NEUROSCIENCE

Abbreviations: SVM: support vector machine; ANOVA: Analysis of variance; RF: Random forests; EEG: Electroencephalogram.

The proposed model achieved higher accuracy than other studies in predicting antidepressant treatment outcomes based on EEG signals by combining neural activity flow and forward feature selection (Table 8). This result indicates that, in addition to showing causality and assisting in understanding the leading cause of treatment outcomes, it can also improve classification results. Also, several limitations should be considered, including the small number of patients from one location that may affect the generalizability of our findings. Moreover, this study used only neural activity flow on channels, but future work may calculate neural activity flow on the brain source localization or combine different types of features. Also, our paper has only used ML models. However, with the advancements in deep learning models, it would be beneficial to use them to improve prediction results.

5. Conclusion

This study investigated a novel method for classifying treatment responses in MDD patients based on neural activity flows. Using neural activity flows, altered brain activity can be identified as causing TRD. Based on the findings of this study, it was demonstrated that the most important neural flows that differ between responders and nonresponders are related to the frontal and parietal lobes at beta frequency, suggesting that the FPN is mainly involved in treatment response. Also, using this kind of neural flow as an input feature in an SVM model and forward feature selection alongside, we could classify responders and nonresponders accurately with an accuracy of 98%. The results of this study suggest that ML models can help predict an individual's response to antidepressants at the beginning of a treatment program.

Ethical Considerations

Compliance with ethical guidelines

This study was approved by the Human Ethics Committee of Shahid Beheshti University of Medical Sciences, Tehran, Iran (Code:IR.SBMU.MSP.REC.1403.141).

Funding

This research was financially supported by Shahid Beheshti University of Medical Sciences, Tehran, Iran (Grant No.: 43009626).

Authors' contributions

Conceptualization, supervision, funding acquisition and resources: Ahmad shalbaf and Reza Shalbaf; Methodology: Ahmad shalbaf; Data collection: Morteza Mirjebreili and Ahmad shalbaf; Data analysis: Seyed Morteza Mirjebreili and Ahmad shalbaf; Investigation, and writing: All authors.

Conflict of interest

The authors declared no conflict of interest.

Acknowledgments

The authors thank Shahid Beheshti University of Medical Sciences, Tehran, Iran, for financial support.

References

Al-Nafjan, A. (2022). Feature selection of EEG signals in neuromarketing. *PeerJ Computer Science, 8,* e944. [DOI:10.7717/PEERJ-CS.944] [PMID]

Amini, N., Mahdavi, M., Choubdar, H., Abedini, A., Shalbaf, A., & Lashgari, R. (2023). Automated prediction of COVID-19 mortality outcome using clinical and laboratory data based on hierarchical feature selection and random forest classifier. *Computer Methods in Biomechanics and Biomedical Engineering, 26*(2), 160-173. [DOI:10.1080/10255842.2022.2050906] [PMID]

Arteaga-Henríquez, G., Simon, M. S., Burger, B., Weidinger, E., Wijkhuijs, A., & Arolt, V., et al. (2019). Low-grade inflammation as a predictor of antidepressant and anti-inflammatory therapy response in MDD patients: A systematic review of the literature in combination with an analysis of experimental data collected in the EU-MOODINFLAME Consortium. *Frontiers in Psychiatry, 10,* 458. [DOI:10.3389/fpsyt.2019.00458] [PMID]

Baskaran, A., Farzan, F., Milev, R., Brenner, C. A., Alturi, S., & Pat McAndrews, M., et al. (2018). The comparative effectiveness of electroencephalographic indices in predicting response to escitalopram therapy in depression: A pilot study. *Journal of Affective Disorders, 227,* 542-549. [DOI:10.1016/J.JAD.2017.10.028] [PMID]

Bell, M. A., & Cuevas, K. (2012). Using EEG to study cognitive development: Issues and practices. J*ournal of Cognition and Development, 13*(3), 281-294. [DOI:10.1080/15248372.2012.691143] [PMID]

Berlim, M. T., Fleck, M. P., & Turecki, G. (2008). Current trends in the assessment and somatic treatment of resistant/refractory major depression: An overview. *Annals of Medicine, 40*(2), 149-159. [DOI:10.1080/07853890701769728] [PMID]

Bremer, V., Becker, D., Kolovos, S., Funk, B., van Breda, W., & Hoogendoorn, M., et al. (2018). Predicting therapy success and costs for personalized treatment recommendations using baseline characteristics: Data-driven analysis. *Journal of Medical Internet Research, 20*(8), e10275. [DOI:10.2196/10275] [PMID]

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology, 14,* 209–236. [DOI:10.1146/annurev-clinpsy-050817-084746] [PMID]

de la Salle, S., Jaworska, N., Blier, P., Smith, D., & Knott, V. (2020). Using prefrontal and midline right frontal EEG-derived theta cordance and depressive symptoms to predict the differential response or remission to antidepressant treatment in major depressive disorder. *Psychiatry Research. Neuroimaging, 302,* 111109. [DOI:10.1016/J.PSCYCHRESNS.2020.111109] [PMID]

Delorme, A., Mullen, T., Kothe, C., Akalin Acar, Z., Bigdely-Shamlo, N., & Vankov, A., et al. (2011). EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing. *Computational Intelligence and Neuroscience, 2011,* 130714. [DOI:10.1155/2011/130714] [PMID]

Ding, C., & Peng, H. (2005). Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology, 3*(2), 185-205. [DOI:10.1142/S0219720005001004] [PMID]

Fischer, A. S., Keller, C. J., & Etkin, A. (2016). The clinical applicability of functional connectivity in depression: Pathways toward more targeted intervention. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging, 1*(3), 262–270. [DOI:10.1016/j.bpsc.2016.02.004] [PMID]

Garcia, G. N., Ebrahimi, T., & Vesin, J. M. (2003). *Support vector EEG classification in the Fourier and time-frequency correlation domains.* Paper presented at: First International IEEE EMBS Conference on Neural Engineering, 2003. Conference Proceedings, Capri, Italy, 20-22 March 2003. [DOI:10.1109/CNE.2003.1196897]

Gibney, E. (2022). Could machine learning fuel a reproducibility crisis in science? *Nature, 608*(7922), 250–251. [DOI:10.1038/d41586-022-02035-w] [PMID]

Goldman, L. S., Nielsen, N. H., & Champion, H. C. (1999). Awareness, diagnosis, and treatment of depression. *Journal of General Internal Medicine, 14*(9), 569-580. [DOI:10.1046/j.1525-1497.1999.03478.x] [PMID]

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424-438. [DOI:10.2307/1912791]

Guze, S. B. (1995). Diagnostic and statistical manual of mental disorders, 4th ed. (DSM-IV). *American Journal of Psychiatry, 152*(8), 1228-1228. [DOI:10.1176/ajp.152.8.1228]

Hasanzadeh, F., Mohebbi, M., & Rostami, R. (2021). Single channel EEG classification: A case study on prediction of major depressive disorder treatment outcome. *IEEE Access, 9,* 3417-3427. [DOI:10.1109/ACCESS.2020.3046993]

Iosifescu, D. V., Greenwald, S., Devlin, P., Mischoulon, D., Denninger, J. W., & Alpert, J. E., et al. (2009). Frontal EEG predictors of treatment outcome in major depressive disorder. *European Neuropsychopharmacology, 19*(11), 772-777. [DOI:10.1016/j.euroneuro.2009.06.001] [PMID]

Jaworska, N., de la Salle, S., Ibrahim, M. H., Blier, P., & Knott, V. (2019). Leveraging machine learning approaches for predicting antidepressant treatment response using electroencephalography (EEG) and clinical data. *Frontiers in Psychiatry, 9,* 768. [DOI:10.3389/FPSYT.2018.00768] [PMID]

Kang, S. G., & Cho, S. E. (2020). Neuroimaging biomarkers for predicting treatment response and recurrence of major depressive disorder. *International Journal of Molecular Sciences, 21*(6), 2148. [DOI:10.3390/ijms21062148] [PMID]

Kautzky, A., Baldinger, P., Souery, D., Montgomery, S., Mendlewicz, J., & Zohar, J., et al. (2015). The combined effect of genetic polymorphisms and clinical parameters on treatment outcome in treatment-resistant depression. *European Neuropsychopharmacology, 25*(4), 441-453. [DOI:10.1016/J.EURONEURO.2015.01.001] [PMID]

Kemp, A. H., Gordon, E., Rush, A. J., & Williams, L. M. (2008). Improving the prediction of treatment response in depression: Integration of clinical, cognitive, psychophysiological, neuroimaging, and genetic measures. *CNS spectrums, 13*(12), 1066–1088. [DOI:10.1017/S1092852900017120] [PMID]

Khodayari-Rostamabad, A., Reilly, J. P., Hasey, G. M., de Bruin, H., & Maccrimmon, D. J. (2013). A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder. *Clinical Neurophysiology, 124*(10), 1975-1985. [DOI:10.1016/j.clinph.2013.04.010] [PMID]

Korzeniewska, A., Crainiceanu, C. M., Kuś, R., Franaszczuk, P. J., & Crone, N. E. (2008). Dynamics of event-related causality in brain electrical activity. *Human Brain Mapping, 29*(10), 1170–1192. [DOI:10.1002/hbm.20458] [PMID]

Korzeniewska, A., Mańczak, M., Kamiński, M., Blinowska, K. J., & Kasicki, S. (2003). Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. *Journal of Neuroscience Methods, 125*(1-2), 195-207. [DOI:10.1016/S0165-0270(03)00052-9] [PMID]

Leuchter, A. F., Cook, I. A., Hunter, A. M., & Korb, A. S. (2009). A new paradigm for the prediction of antidepressant treatment response. *Dialogues in Clinical Neuroscience, 11*(4), 435–446. [DOI:10.31887/DCNS.2009.11.4/afleuchter] [PMID]

Liu, Z., Zhang, M., Xu, G., Huo, C., Tan, Q., & Li, Z., et al. (2017). Effective connectivity analysis of the brain network in drivers during actual driving using near-infrared spectroscopy. *Frontiers in Behavioral Neuroscience, 11,* 211. [DOI:10.3389/fnbeh.2017.00211] [PMID]

Mamitsuka, H. (2006). Selecting features in microarray classification using ROC curves. *Pattern Recognition, 39*(12), 2393-2404. [DOI:10.1016/J.PATCOG.2006.07.010]

Martens, M. A. G., Filippini, N., Harmer, C. J., & Godlewska, B. R. (2022). Resting state functional connectivity patterns as biomarkers of treatment response to escitalopram in patients with major depressive disorder. *Psychopharmacology, 239*(11), 3447–3460. [DOI:10.1007/s00213-021-05915-7] [PMID]

Mason, S. J., & Graham, N. E. (2002). Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society, 128*(584), 2145-2166. [DOI:10.1256/003590002320603584]

Mullen, T. R. (2014). The dynamic brain: Modeling neural dynamics and interactions from human electrophysiological recordings [PhD dissertation]. San Diego: University of California. [Link]

Mumtaz, W., Xia, L., Mohd Yasin, M. A., Azhar Ali, S. S., & Malik, A. S. (2017). A wavelet-based technique to predict treatment outcome for Major Depressive Disorder. *Plos One, 12*(2), e0171409. [DOI:10.1371/journal.pone.0171409] [PMID]

Otte, C., Gold, S. M., Penninx, B. W., Pariante, C. M., Etkin, A., Fava, M., & Mohr, D. C., et al. (2016). Major depressive disorder. *Nature reviews. Disease Primers, 2,* 16065. [DOI:10.1038/nrdp.2016.65] [PMID]

Patel, M. J., Andreescu, C., Price, J. C., Edelman, K. L., Reynolds, C. F., 3rd, & Aizenstein, H. J. (2015). Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. *International Journal of Geriatric Psychiatry, 30*(10), 1056-1067. [DOI:10.1002/GPS.4262] [PMID]

Peker, M., Arslan, A., Sen, B., Celebi, F. V., & But, A. (2015). *A novel hybrid method for determining the depth of anesthesia level: Combining ReliefF feature selection and random forest algorithm (ReliefF+RF).* Paper presented at: 2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA), Madrid, Spain, 02-04 September 2015. [DOI:10.1109/INISTA.2015.7276737]

Pizzagalli, D. A. (2011). Frontocingulate dysfunction in depression: Toward biomarkers of treatment response. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology, 36*(1), 183-206. [DOI:10.1038/NPP.2010.166] [PMID]

Sambunaris, A., Hesselink, J. K., Pinder, R., Panagides, J., & Stahl, S. M. (1997). Development of new antidepressants. *The Journal of Clinical Psychiatry, 58 (*Suppl 6), 40–53. [PMID]

Schultz, D. H., Ito, T., Solomyak, L. I., Chen, R. H., Mill, R. D., & Anticevic, A., et al. (2018). Global connectivity of the frontoparietal cognitive control network is related to depression symptoms in the general population. *Network Neuroscience, 3*(1), 107-123. [DOI:10.1162/NETN_A_00056] [PMID]

Schwartz, B., Cohen, Z. D., Rubel, J. A., Zimmermann, D., Wittmann, W. W., & Lutz, W. (2021). Personalized treatment selection in routine care: Integrating machine learning and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy. *Psychotherapy Research, 31*(1), 33-51. [DOI:10.1080/10503307.2020.1769219] [PMID]

Şen, B., Peker, M., Çavuşoğlu, A., & Çelebi, F. V. (2014). A comparative study on classification of sleep stage based on eeg signals using feature selection and classification algorithms. *Journal of Medical Systems, 38*(3), 18. [DOI:10.1007/S10916-014-0018-0] [PMID]

Shahabi, M. S., & Shalbaf, A. (2022). Prediction of treatment outcome in major depressive disorder using ensemble of hybrid transfer learning and long short term memory based on EEG Signal. *IEEE Transactions on Cognitive and Developmental Systems, 15*(3), 1279 - 1288. [DOI:10.1109/TCDS.2022.3207350]

Smit, D. J., Stam, C. J., Posthuma, D., Boomsma, D. I., & de Geus, E. J. (2008). Heritability of "small-world" networks in the brain: A graph theoretical analysis of resting-state EEG functional connectivity. *Human Brain Mapping, 29*(12), 1368-1378. [DOI:10.1002/hbm.20468] [PMID]

Stahl, S. M., & Grady, M. M. (2003). Differences in mechanism of action between current and future antidepressants. *The Journal of Clinical Psychiatry, 64* (Suppl 13), 13–17. [PMID]

Subasi, A., & Gursoy, M. I. (2010). EEG signal classification using PCA, ICA, LDA and support vector machines. *Expert Systems with Applications, 37*(12), 8659-8666. [DOI:10.1016/J.ESWA.2010.06.065]

van der Vinne, N., Vollebregt, M. A., Boutros, N. N., Fallahpour, K., van Putten, M. J. A. M., & Arns, M. (2019). Normalization of EEG in depression after antidepressant treatment with sertraline? A preliminary report. *Journal of Affective Disorders, 259,* 67-72. [DOI:10.1016/J.JAD.2019.08.016] [PMID]

Webb, C. A., Trivedi, M. H., Cohen, Z. D., Dillon, D. G., Fournier, J. C., & Goer, F., et al. (2019). Personalized prediction of antidepressant v. placebo response: Evidence from the EMBARC study. *Psychological Medicine, 49*(07), 1118-1127. [DOI:10.1017/S0033291718001708] [PMID]

Whitton, A. E., Deccy, S., Ironside, M. L., Kumar, P., Beltzer, M., & Pizzagalli, D. A. (2018). Electroencephalography source functional connectivity reveals abnormal high-frequency communication among large-scale functional networks in depression. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 3*(1), 50-58. [DOI:10.1016/j.bpsc.2017.07.001] [PMID]

WHO. (2017). *Depression and other common mental disorders: Global health estimates*. Geneva: World Health Organization. [Link]

Zhang, Y., Wang, K., Wei, Y., Guo, X., Wen, J., & Luo, Y. (2022).Minimal EEG channel selection for depression detection with connectivity features during sleep. *Computers in Biology and Medicine, 147,* 105690. [DOI:10.1016/j.compbiomed.2022.105690] [PMID]

Zhdanov, A., Atluri, S., Wong, W., Vaghei, Y., Daskalakis, Z. J., & Blumberger, D. M., et al. (2020). Use of machine learning for predicting escitalopram treatment outcome from electroencephalography recordings in adult patients with depression. *JAMA Network Open, 3*(1), e1918377.[DOI:10.1001/JAMANETWORKOPEN.2019.18377] [PMID]

This Page Intentionally Left Blank