



Published in final edited form as:

Neurosci Inform. 2022 December ; 2(4): . doi:10.1016/j.neuri.2022.100110.

Gradient boosting decision-tree-based algorithm with neuroimaging for personalized treatment in depression

Farzana Z. Ali^{a,*}, Kenneth Wengler^{a,b}, Xiang He^{c,d,1}, Minh Hoai Nguyen^e, Ramin V. Parsey^f, Christine DeLorenzo^{a,b,f}

^aDepartment of Biomedical Engineering, Stony Brook University, Stony Brook, NY, USA

^bDepartment of Psychiatry, Columbia University and New York State Psychiatric Institute, New York, NY, USA

^cDepartment of Radiology, Stony Brook Medicine, Stony Brook, NY, USA

^dDepartment of Radiology, Northshore University Hospital, Manhasset, NY, USA

^eDepartment of Computer Science, Stony Brook University, Stony Brook, NY, USA

^fDepartment of Psychiatry, Renaissance School of Medicine at Stony Brook University, Stony Brook, NY, USA

Abstract

Introduction: Pretreatment positron emission tomography (PET) with 2-deoxy-2-[¹⁸F]fluoro-D-glucose (FDG) and magnetic resonance spectroscopy (MRS) may identify biomarkers for predicting remission (absence of depression). Yet, no such image-based biomarkers have achieved

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

*Corresponding author. fali@buffalo.edu (F.Z. Ali).

¹Present address.

Author contributions

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship.

Human and animal rights

The authors declare that the work described has been carried out in accordance with the Declaration of Helsinki of the World Medical Association revised in 2013 for experiments involving humans as well as in accordance with the EU Directive 2010/63/EU for animal experiments.

Informed consent and patient details

Informed consent for participation in research was obtained from all participants prior to study.

Code availability

Standard software applications have been used for data analysis and are mentioned in the methodology.

Clinical trials registration

Advancing Personalized Antidepressant Treatment Using PET/MRI, [ClinicalTrials.gov: NCT02623205](https://clinicaltrials.gov/ct2/show/study/NCT02623205).

CRedit authorship contribution statement

Farzana Z. Ali: Conceptualization, Data curation, Formal analysis, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Kenneth Wengler:** Formal analysis, Visualization, Writing – review & editing. **Xiang He:** Formal analysis, Writing – review & editing. **Minh Hoai Nguyen:** Conceptualization, Investigation, Methodology, Software, Writing – review & editing. **Ramin V. Parsey:** Conceptualization, Investigation, Writing – review & editing. **Christine DeLorenzo:** Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.neuri.2022.100110>.

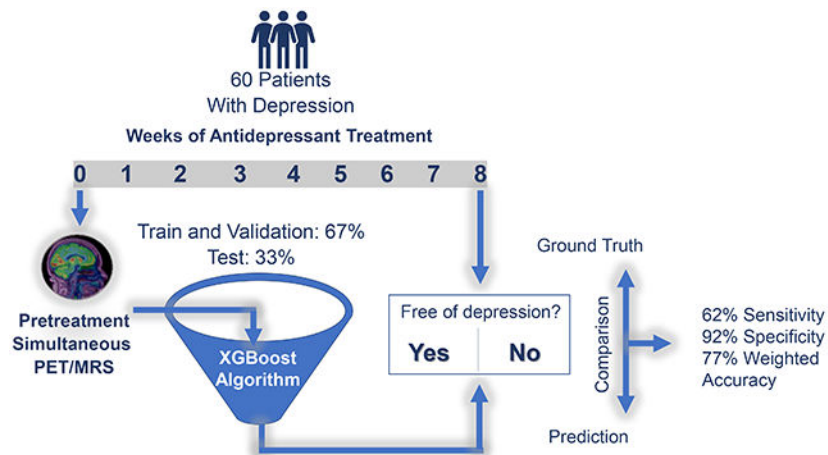
clinical validity. The purpose of this study was to identify biomarkers of remission using machine learning (ML) with pretreatment FDG-PET/MRS neuroimaging, to reduce patient suffering and economic burden from ineffective trials.

Methods: This study used simultaneous PET/MRS neuroimaging from a double-blind, placebo-controlled, randomized antidepressant trial on 60 participants with major depressive disorder (MDD) before initiating treatment. After eight weeks of treatment, those with ≤ 7 on 17-item Hamilton Depression Rating Scale were designated *a priori* as remitters (free of depression, 37%). Metabolic rate of glucose uptake (metabolism) from 22 brain regions were acquired from PET. Concentrations (mM) of glutamine and glutamate and gamma-aminobutyric acid (GABA) in anterior cingulate cortex were quantified from MRS. The data were randomly split into 67% train and cross-validation ($n = 40$), and 33% test ($n = 20$) sets. The imaging features, along with age, sex, handedness, and treatment assignment (selective serotonin reuptake inhibitor or SSRI vs. placebo) were entered into the eXtreme Gradient Boosting (XGBoost) classifier for training.

Results: In test data, the model showed 62% sensitivity, 92% specificity, and 77% weighted accuracy. Pretreatment metabolism of left hippocampus from PET was the most predictive of remission.

Conclusions: The pretreatment neuroimaging takes around 60 minutes but has potential to prevent weeks of failed treatment trials. This study effectively addresses common issues for neuroimaging analysis, such as small sample size, high dimensionality, and class imbalance.

Graphical Abstract



Keywords

Artificial intelligence; Imaging informatics; Medical imaging; FDG PET; Magnetic resonance spectroscopy; XGBoost

1. Introduction

There is an urgent need for reliable prediction of potential antidepressant failure in treatment of major depressive disorder (MDD). Pretreatment electroencephalogram (EEG) has predicted antidepressant efficacy with performance considerable for clinical utility [1].

EEG electrodes are placed on the surface of the brain to obtain functional measures, leading to lower spatial resolution compared to molecular neuroimaging, where the source of signal may not be apparent [2]. Molecular imaging modalities, such as positron emission tomography (PET) with 2-deoxy-2-[¹⁸F]fluoro-D-glucose (FDG) and proton magnetic resonance spectroscopy (¹H-MRS or MRS) have been useful for early assessment and risk stratification in patients with neurological, oncological and cardiovascular disorders [3]. Yet, they have not been proved clinically useful due to lack of accuracy [4].

FDG-PET is a sensitive indicator of cerebral function, leading to its extensive use for assessing response to the most widely used first-line MDD treatment, selective serotonin reuptake inhibitor (SSRI) [5]. However, these prior studies do not agree on the predictive region/measure and their prediction could not be replicated using conventional statistics [6].

Glutamine (Gln) is the precursor to the excitatory neurotransmitter, glutamate (Glu) and inhibitory neurotransmitter, γ -aminobutyric acid (GABA), and all three are essential amino acids for brain metabolism [7]. Prior magnetic resonance spectroscopy (MRS) studies showed that higher pretreatment Glu in the anterior cingulate cortex (ACC) [8] can predict better response, but other MRS studies showed no predictive potential for antidepressant response using Glu [9,10], GABA [9–11], or Glx (Gln/Glu) [12], adding to the lack of consensus on biomarkers of MDD treatment. These prior inconsistent findings using single imaging modalities with fewer functional features and conventional statistical models have motivated the use of machine learning (ML) with multimodal neuroimaging (both PET and MRS measures) for better prediction accuracy [13].

The purpose of this study was to identify biomarkers to predict remission (absence of depression) after eight weeks of treatment using pretreatment neuroimaging measures from i) PET and ii) MRS with a widely popular (due to superior performance) supervised, gradient tree boosting ensemble algorithm, eXtreme Gradient Boosting (XGBoost), which is a fast, scalable, and explainable artificial intelligence (AI) classifier with strong regularization [14]. XGBoost allows the adjustment of multiple hyperparameters to avoid overfitting and automatically ranks the most predictive features that can be used as biomarkers for remission. This article will have the following contribution regarding predictive biomarkers for antidepressant treatment.

- This study provides a novel instance of the development of a comprehensive machine learning model by integrating pretreatment brain functional measures from simultaneous PET/MRS in one framework.
- The pretreatment simultaneous PET/MRS used in the study takes around 60 minutes and has the potential to prevent weeks of failed treatment trials.
- Applying predictive measures from neuroimaging can reduce delay to effective treatment, patient suffering and economic burden, and enhance long-term functional outcomes.
- The findings may assist clinicians with treatment selection and shed light on the neurobiology of remission.

2. Material and methods

2.1. Study cohort

This study involved the analysis of simultaneously acquired PET/MRS neuroimaging data from a randomized, placebo-controlled, double-blind, single-site SSRI (Escitalopram) trial on 60 participants with MDD before initiating treatment. The study design and rationale for the data set have been previously described [6,10]. 43 regional measures of metabolic rate of glucose (MRGlu, mg/(min*100 mL)) from FDG-PET across multiple brain regions [6], and concentrations (mM) of Glx (Glu+Gln, a composite peak formed by Glu and Gln), GABA and the ratio of Glx to GABA (Glx/GABA) in the ACC from MRS [10] were quantified as described in prior publications. After eight weeks of treatment, those with ≤ 7 on the 17-item Hamilton Depression Rating Scale were designated *a priori* as remitters (free of depression).

2.2. Machine learning analysis

The ML model development with hyperparameters is illustrated in Fig. 1. The data partition was performed by random splitting into 67% train and cross-validation ($n = 40$), and 33% test ($n = 20$) sets, stratified by outcome (remitters vs. non-remitters) and treatment assignment (SSRI vs. placebo), to ensure comparable distribution in each data set. There is no established power calculation for XGBoost, so the conventional ML practice of evaluating the fitted model on validation set was used. The hyperparameter, *scale_pos_weight* [sum(negative instances)/sum(positive instances)] = (non-remitters/remitters) from the train set assigned greater weight to prediction of remitters. *OneClassSVM* identified and removed outliers from the train set. *Synthetic Minority Oversampling Technique (SMOTE)* oversampled the remitters class [15].

The hyperparameters for subsampling, number of trees, and depth of tree were optimized using *GridSearchCV* with 3 repetitions of stratified 10-fold cross-validation (preferred for depression research [16,17]). The 50 input features including the 46 imaging (43 PET and 3 MRS) measures, and information on age, sex, handedness, and treatment assignment were entered into XGBoost for training with optimized hyperparameters to predict remitters vs. non-remitters. The model performance was evaluated on the test data using confusion matrix (Fig. 2). Statistical and machine learning analyses were performed using STATA/SE 13.0 (StataCorp LLC, College Station, TX) and Python 3.9.0 (Python Software Foundation, Beaverton, OR).

3. Results

3.1. Study cohort

The study consisted of 60 participants with an age range of 18 to 64 years (mean \pm standard deviation: 30 ± 14 years). 37 (62%) were females, 30 (50%) were placed on SSRI and 51 (85%) were right-handed. After eight weeks of treatment, 22 participants remitted (37% remitters) and 38 participants did not (63% non-remitters). There was no significant difference between the non-remitters and remitters groups in the study sample in terms of age, sex, handedness, or treatment assignment.

3.2. Machine learning analysis

In the training set ($n = 40$), there were 14 remitters (35%) with 5 on SSRI and 9 on placebo, and 26 (65%) non-remitters with 15 on SSRI and 11 on placebo ($scale_pos_weight = 1.86$). The most predictive neuroimaging features based on “gain” (according to their contribution to the fitted model) are shown in Fig. 2. The cross-validated training and testing Receiver Operating Characteristic (ROC) Area under the Curves are shown in Fig. 3.

Supplemental Table 1 and Table 1 below show the performance of the fitted model on the unseen test data ($n = 20$) consisting of 8 remitters (40%) with 3 on SSRI and 5 on placebo, and 12 (60%) non-remitters with 7 on SSRI and 5 on placebo.

4. Discussion

4.1. Novelty

This study was the first effort to build a comprehensive predictive model using simultaneous PET/MRS data from randomized clinical trial for prediction of remission. This was also the first time a gradient boosting decision-tree-based algorithm was used for this purpose. In this novel architecture, the XGBoost hyperparameters were utilized following oversampling and outlier removal, which resulted in the current model’s weighted accuracy (77%) comparable to previous studies on the prediction of antidepressant treatment outcome with much larger sample size [17–20].

4.2. Generalizability

The splitting in the current study ensured an adequate test set (as opposed to the alternative practice of 80/20 or 90/10 train/test split) to protect against performance misestimation in MDD research [21]. This model’s generalizability is further strengthened through *Stratified Cross Validation* which is particularly useful for analyzing small data sets with unbalanced classes, as seen in our data set with the remitters and non-remitters class. This technique allows similar proportion of different classes in each fold to ensure all strata of the data is well represented.

4.3. Regularization

Instead of feature reduction, the current model optimized regularization hyperparameters that ensures higher accuracy and better uncertainty assessment [22]. These XGBoost regularization hyperparameters, a.k.a. penalty terms *alpha* (L1, *LASSO Regression*) and *lambda* (L2, *Ridge Regression*), shrink the coefficients of less relevant features toward 0 [23]. Using this technique, this model achieved accuracy higher than a previous model with 59% accuracy that used feature reduction to select 25 most predictive variables for remission after 12 weeks of SSRI treatment from 164 patient-reportable variables [18]. Including more features also protects the predictive model performance from being affected by influential data points [19].

4.4. Performance

The true positive rate and true negative rate for classification in the held-out test data are reasonably high compared to the negligible values of false positive rate and false negative

rate, attesting the acceptable performance of the model. At least 100 number of trees were used following convention when searching for the optimal number of trees, however the Receiver Operating Characteristic Area Under the Curve indicates that having a lower number of trees (~ 70) might have given slightly better performance.

4.5. Limitations

To address limitations related to small sample size in neuroimaging research, two most common data augmentation techniques for image classification including generative adversarial networks and unity game engine can be explored in future. However, this will require finding the optimal data augmentation strategy and developing evaluation systems to ensure quality of augmented data sets, while accounting for the computational cost for slower convergence.

4.6. Potential biomarkers

Even though there is a scarcity of predictive analytics using biomarkers with machine learning in neuropsychiatry, the insights from successful machine learning applications in neurological disorders such as stroke may be useful, considering the bidirectionality of stroke and depression [24]. The current model has been compared to predictive analytics using biomarkers from other modalities in Table 2. One such potential biomarker could be electromyography (EMG), since motor activity can be lower in depression [25–29], and increase with depression improvement [25,30,31]. The tree-based algorithm, random forest with real-time signals from thighs and calves of 287 participants has shown predictive accuracy > 90% for stroke [32]. Furthermore, electroencephalography (EEG) has been used for predicting stroke with Classification and Regression Trees (C&RT) algorithms with 89% accuracy [33]. With tree-based models such as C5.0 and random forest, EEG data has shown only around 70% accuracy for predicting stroke [34]. Pretreatment EEG measures have been useful for predicting antidepressant efficacy with > 87% accuracy using a mixture of factor analysis (MFA) classifier [1]. However, the biosignals received from EEG electrodes placed on the surface of the brain are less precise for locating the source of the signal as compared to PET/MRS used in the study [2]. Nonetheless, adding these cost-effective modalities to molecular neuroimaging may help develop a more comprehensive predictive model with improved sensitivity for predicting antidepressant response.

5. Conclusions

To our knowledge, this was the first effort to develop a gradient tree boosting classifier by integrating pretreatment multimodal molecular neuroimaging with easily interpretable brain functional measures in one framework, with accuracy comparable to previous predictive models. This study provides information on effectively addressing common issues related to neuroimaging analysis, such as small sample size, high dimensionality, and class imbalance. More importantly, the pretreatment neuroimaging takes around 60 minutes and has the potential to prevent weeks of failed treatment trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We acknowledge the support from Kathryn Hill and John D. Gardus, Renaissance School of Medicine at Stony Brook University.

Funding

This work was supported by the National Institute of Mental Health [R01MH104512, R36MH130050]; the Brain & Behavior Foundation; The Dana Foundation; The New York State Faculty Development Grant; and the Ruth L. Kirschstein National Research Service Award and National Health Service Corps Scholarship [T32GM127253] which support the corresponding author's training in the Scholars in Biomedical Sciences program.

Availability of data and materials

The data used in the study can be provided if required on request from the corresponding author.

References

- [1]. Khodayari-Rostamabad A, Reilly JP, Hasey GM, de Bruin H, MacCrimmon DJ, A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder, *Clin. Neurophysiol* 124 (2013) 1975–1985. [PubMed: 23684127]
- [2]. Burle B, Spieser L, Roger C, Casini L, Hasbroucq T, Vidal F, Spatial and temporal resolutions of EEG: is it really black and white? A scalp current density view, *Int. J. Psychophysiol* 97 (2015) 210–220. [PubMed: 25979156]
- [3]. Musafargani S, Ghosh KK, Mishra S, Mahalakshmi P, Padmanabhan P, Gulyás B, PET/MRI: a frontier in era of complementary hybrid imaging, *Eur. J. Hybrid Imag* 2 (2018) 12.
- [4]. Kang SG, Cho SE, Neuroimaging biomarkers for predicting treatment response and recurrence of major depressive disorder, *Int. J. Mol. Sci* (2020) 21.
- [5]. McGrath CL, Kelley ME, Holtzheimer PE, Dunlop BW, Craighead WE, Franco AR, et al. , Toward a neuroimaging treatment selection biomarker for major depressive disorder, *JAMA Psychiatr.* 70 (2013) 821–829.
- [6]. Hill KR, Gardus JD, Bartlett EA, Perlman G, Parsey RV, DeLorenzo C, Measuring brain glucose metabolism in order to predict response to antidepressant or placebo: a randomized clinical trial, *NeuroImage Clin.* 32 (2021) 102858. [PubMed: 34689056]
- [7]. Sanacora G, Gueorguieva R, Epperson C, Wu Y, Appel M, Rothman D, et al. , Subtype-specific alterations in gamma-aminobutyric acid and glutamate in patients with major depression, *Arch. Gen. Psychiatry* 61 (2004) 705–713. [PubMed: 15237082]
- [8]. Merkl A, Schubert F, Quante A, Luborzewski A, Brakemeier EL, Grimm S, et al. , Abnormal cingulate and prefrontal cortical neurochemistry in major depression after electroconvulsive therapy, *Biol. Psychiatry* 69 (2011) 772–779. [PubMed: 20951980]
- [9]. Salvadore G, Van der Veen J, Zhang Y, Marenco S, Machado-Vieira R, Baumann J, et al. , An investigation of amino acid neurotransmitters as potential predictors of clinical improvement to ketamine in depression, *Neuropsychopharmacology* 15 (2012) 1063–1072.
- [10]. Narayan GA, Hill KR, Wengler K, He X, Wang J, Yang J, et al. , Does the change in glutamate to GABA ratio correlate with change in depression severity? A randomized, double-blind clinical trial, *Mol. Psychiatry* (2022).
- [11]. Abdallah C, Niciu M, Fenton L, Fasula M, Jiang L, Black A, et al. , Decreased occipital cortical glutamate levels in response to successful cognitive behavioral therapy and pharmacotherapy for major depressive disorder, *Psychother. Psychosom* 83 (2014) 298–307. [PubMed: 25116726]

- [12]. Brennan BP, Admon R, Perriello C, LaFlamme EM, Athey AJ, Pizzagalli DA, et al. . Acute change in anterior cingulate cortex GABA, but not glutamine/glutamate, mediates antidepressant response to citalopram, *Psychiatry Res. Neuroimaging* 269 (2017) 9–16. [PubMed: 28892734]
- [13]. Calhoun VD, Sui J, Multimodal fusion of brain imaging data: a key to finding the missing link(s) in complex mental illness, *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 1 (2016) 230–244. [PubMed: 27347565]
- [14]. Chen T, Guestrin C, XGBoost: a scalable tree boosting system, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, San Francisco, California, USA, 2016, pp. 785–794.
- [15]. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP, SMOTE: synthetic minority over-sampling technique, *J. Artif. Intell. Res* 16 (2002) 321–357.
- [16]. Pearson R, Pisner D, Meyer B, Shumake J, Beevers CG, A machine learning ensemble to predict treatment outcomes following an Internet intervention for depression, *Psychol. Med* 49 (2019) 2330–2341. [PubMed: 30392475]
- [17]. Kautzky A, Baldinger-Melich P, Kranz GS, Vanicek T, Souery D, Montgomery S, et al. . A new prediction model for evaluating treatment-resistant depression, *J. Clin. Psychiatry* 78 (2017) 215–222. [PubMed: 28068461]
- [18]. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, et al. . Cross-trial prediction of treatment outcome in depression: a machine learning approach, *Lancet Psychiatry* 3 (2016) 243–250. [PubMed: 26803397]
- [19]. Browning M, Kingslake J, Dourish CT, Goodwin GM, Harmer CJ, Dawson GR, Predicting treatment response to antidepressant medication using early changes in emotional processing, *Eur. Neuropsychopharmacol* 29 (2019) 66–75. [PubMed: 30473402]
- [20]. Cattaneo A, Ferrari C, Uher R, Bocchio-Chiavetto L, Riva MA, et al. . Consortium tMI, Absolute measurements of macrophage migration inhibitory factor and interleukin-1- β mRNA levels accurately predict treatment response in depressed patients, *Int. J. Neuropsychopharmacol* (2016) 19.
- [21]. Flint C, Cearns M, Opel N, Redlich R, Mehler DMA, Emden D, et al. . Systematic misestimation of machine learning performance in neuroimaging studies of depression, *Neuropsychopharmacology* (2021).
- [22]. Greenland S, Invited commentary: variable selection versus shrinkage in the control of multiple confounders, *Am. J. Epidemiol* 167 (2008) 523–529. [PubMed: 18227100]
- [23]. Dankers F, Traverso A, Wee L, van Kuijk SMJ, Prediction modeling methodology, in: Kubben P, Dumontier M, Dekker A (Eds.), *Fundamentals of Clinical Data Science*, Springer, Cham (CH), 2019, pp. 101–120.
- [24]. Dong J-Y, Zhang Y-H, Tong J, Qin L-Q, Depression and risk of stroke, *Stroke* 43 (2012) 32–37. [PubMed: 22020036]
- [25]. Wielopolski J, Reich K, Clepce M, Fischer M, Sperling W, Kornhuber J, et al. . Physical activity and energy expenditure during depressive episodes of major depression, *J. Affect. Disord* 174 (2014) 310–316. [PubMed: 25532078]
- [26]. Buyukdura JS, McClintock SM, Croarkin PE, Psychomotor retardation in depression: biological underpinnings, measurement, and treatment, *Prog. Neuropsychopharmacol. Biol. Psychiatry* 35 (2011) 395–409. [PubMed: 21044654]
- [27]. Schuch F, Vancampfort D, Firth J, Rosenbaum S, Ward P, Reichert T, et al. . Physical activity and sedentary behavior in people with major depressive disorder: a systematic review and meta-analysis, *J. Affect. Disord* 210 (2017) 139–150. [PubMed: 28033521]
- [28]. Vancampfort D, Firth J, Schuch FB, Rosenbaum S, Mugisha J, Hallgren M, et al. . Sedentary behavior and physical activity levels in people with schizophrenia, bipolar disorder and major depressive disorder: a global systematic review and meta-analysis, *World Psychiatry* 16 (2017) 308–315. [PubMed: 28941119]
- [29]. Krane-Gartiser K, Henriksen TEG, Morken G, Vaaler A, Fasmer OB, Actigraphic assessment of motor activity in acutely admitted inpatients with bipolar disorder, *PLoS ONE* 9 (2014) e89574. [PubMed: 24586883]

- [30]. Bennabi D, Vandel P, Papaxanthis C, Pozzo T, Haffen E, Psychomotor retardation in depression: a systematic review of diagnostic, pathophysiologic, and therapeutic implications, *BioMed Res. Int* 2013 (2013) 158746. [PubMed: 24286073]
- [31]. Burton C, McKinsty B, T tar AS, Serrano-Blanco A, Pagliari C, Wolters M, Activity monitoring in patients with depression: a systematic review, *J. Affect. Disord* 145 (2013) 21–28. [PubMed: 22868056]
- [32]. Yu J, Park S, Kwon S-H, Ho CMB, Pyo C-S, Lee H, AI-based stroke disease prediction system using real-time electromyography signals, *Appl. Sci* 10 (2020) 6791.
- [33]. Hussain I, Park SJ, HealthSOS: real-time health monitoring system for stroke prognostics, *IEEE Access* 8 (2020) 213574–213586.
- [34]. Hussain I, Park SJ, Quantitative evaluation of task-induced neurological outcome after stroke, *Brain Sci.* (2021) 11.
- [35]. Hussain I, Park SJ, Prediction of myoelectric biomarkers in post-stroke gait, *Sensors (Basel)* 21 (2021).

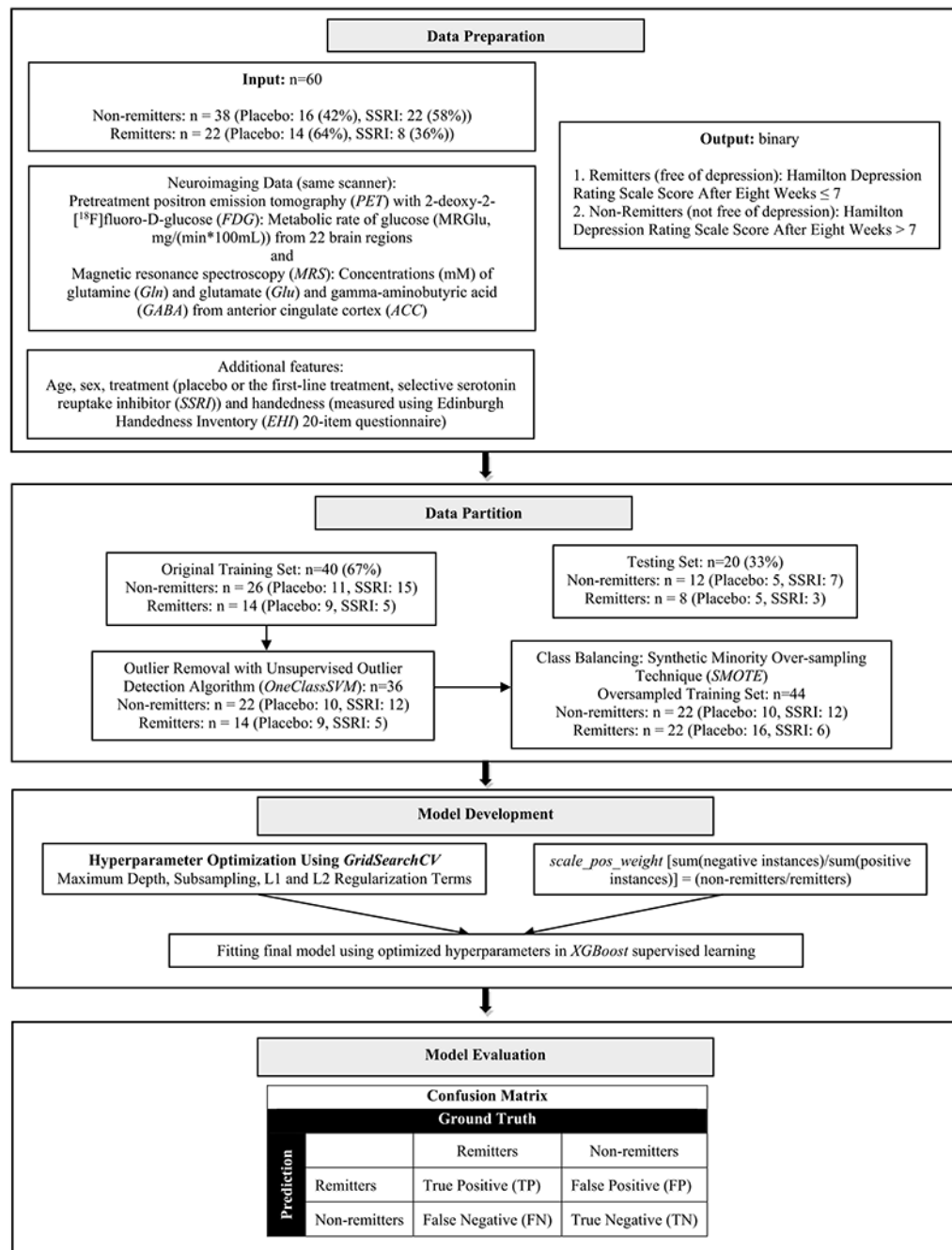


Fig. 1. XGBoost model development and evaluation.

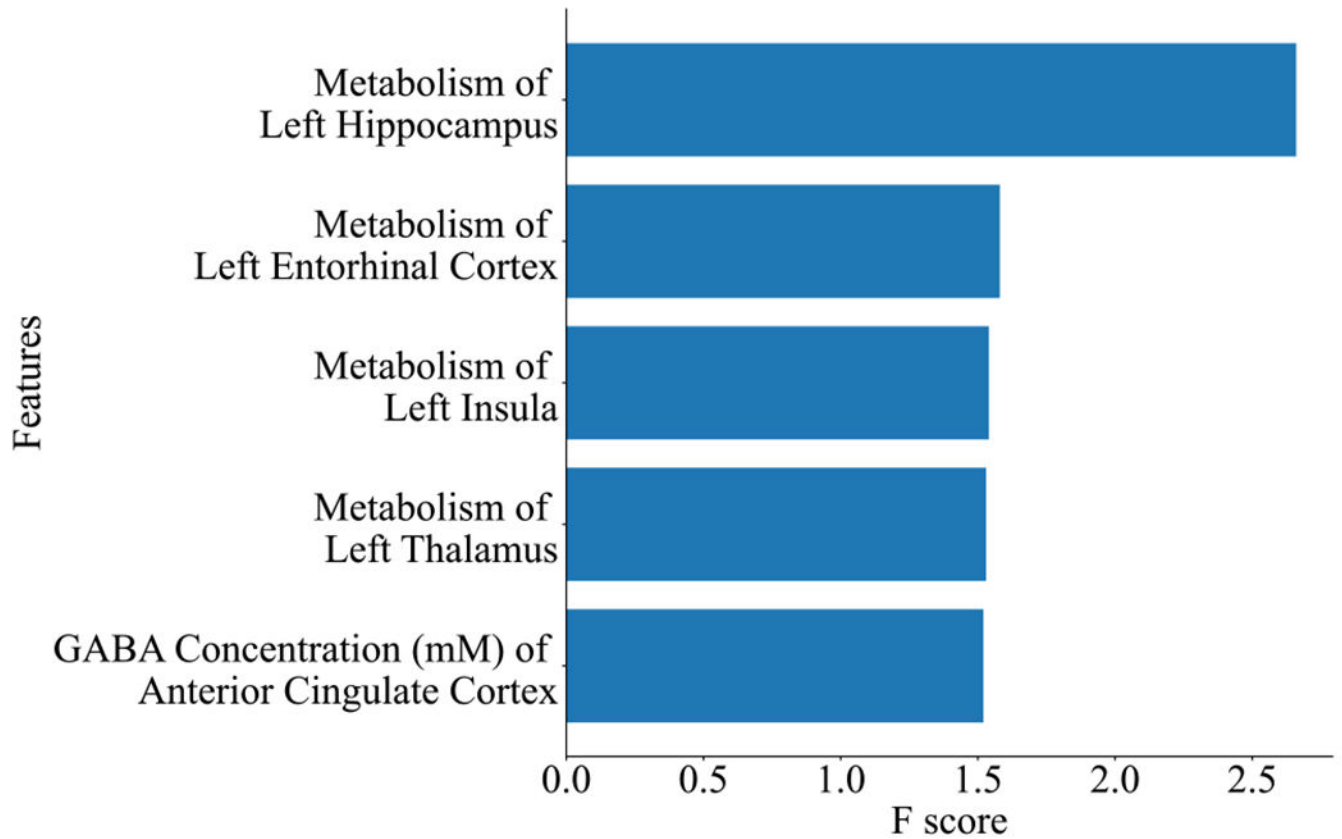


Fig. 2. The most predictive imaging features from the XGBoost model. The features are as follows: metabolism estimated by the metabolic rate of glucose (MRGlu, mg/(min*100 mL)) of left hippocampus, left entorhinal cortex, left insula, left thalamus and GABA (γ -aminobutyric acid) concentration of anterior cingulate cortex. F score: relative contribution of the feature to the prediction model.

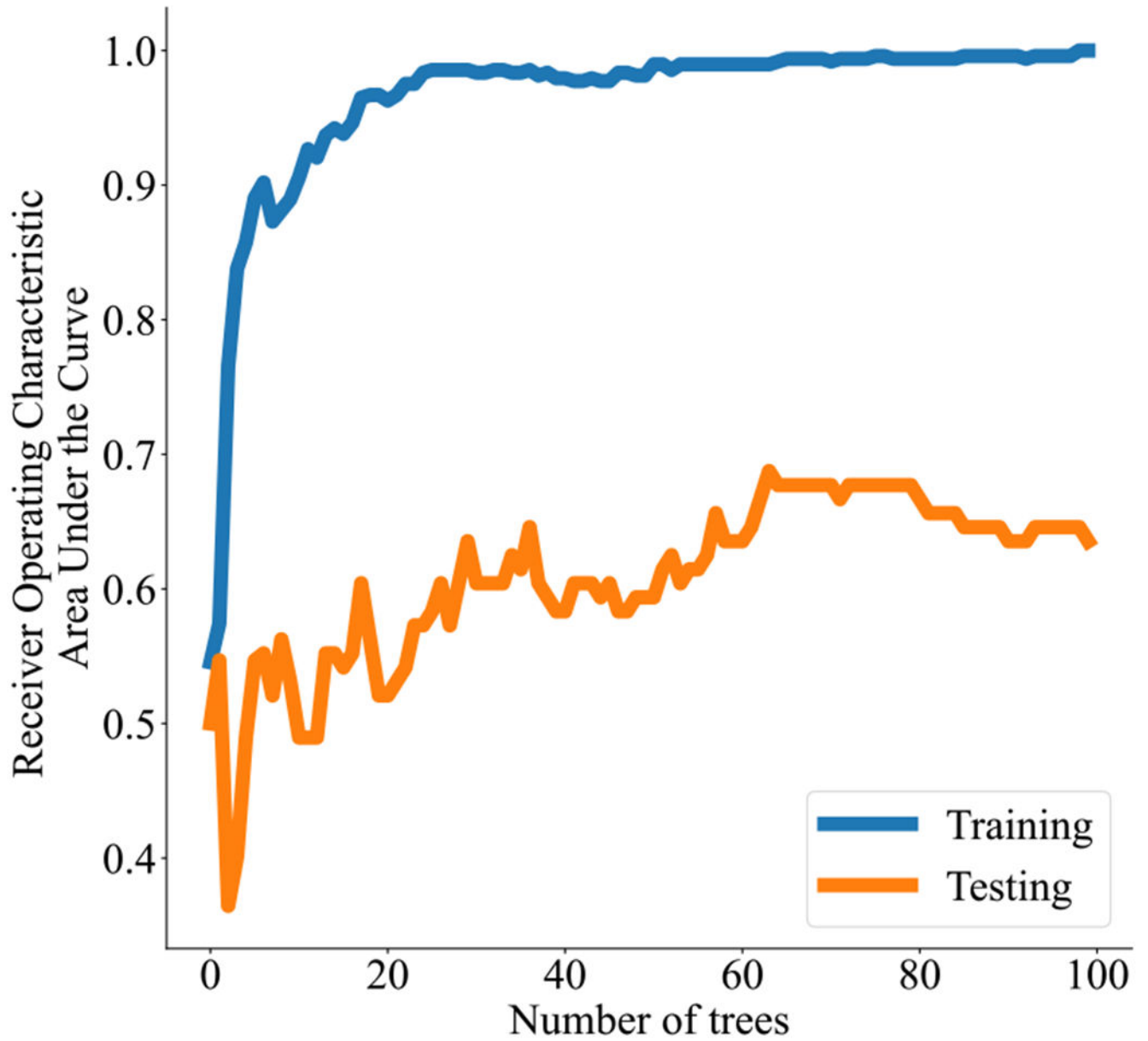


Fig. 3. XGBoost receiver operating characteristic area under the curve.

Table 1

The performance metrics of the fitted model on test data.

Performance Metrics	Recall/Sensitivity/True Positive Rate	False Negative Rate	Specificity/True Negative Rate	False Positive Rate	Weighted Accuracy	Precision/Positive Predictive Value	Negative Predictive Value	f1-Score	Sample Size
Overall	62%	38%	92%	8%	77%	83%	79%	0.71	20
Stratified by Treatment Assignment (SSRI vs. Placebo)									
SSRI	67%	33%	86%	14%	76%	67%	86%	0.67	10
Placebo	60%	40%	100%	0%	80%	100%	71%	0.75	10
Stratified by Sex (Female vs. Male)									
Female	50%	50%	100%	0%	75%	100%	75%	0.67	10
Male	75%	25%	83%	17%	79%	75%	83%	0.75	10

Table 2

A comparative table on the advantages and disadvantages of the proposed model.

Study	Advantages	Disadvantages
Hussain et al., 2020 [35]	Cost-effective for using EMG to address motor activity changes with antidepressant response	Accuracy lower than the current study using neural network model that requires relatively large sample
Yu et al., 2022 [32]	90% prediction accuracy with random forest	Lack of insight from neurobiology if only using EMG signals, but useful for multimodal analysis
Hussain et al., 2020 [33]	Portable EEG device for data collection, comparison of different algorithms	Removal of features before training the algorithm may lead to loss of information
Hussain et al., 2021 [34]	Wireless EEG devices used for data collection	Expertise will be required for EEG data interpretation and tree-based algorithms showed 70% accuracy