Open Access Library Journal
2025, Volume 12, e12894
ISSN Online: 2333-9721
ISSN Print: 2333-9705

A Machine Learning Approach to Predicting Treatment Outcomes in Bipolar Depression with OCD Comorbidity

Rocco de Filippis[1*], Abdullah Al Foysal[2]

[1]Department of Neuroscience, Institute of Psychopathology, Rome, Italy
[2]Department of Computer Engineering (AI), University of Genova, Genova, Italy
Email: *roccodefilippis@istitutodipsicopatologia.it, niloyhasanfoysal440@gmail.com

How to cite this paper: de Filippis, R. and Al Foysal, A. (2025) A Machine Learning Approach to Predicting Treatment Outcomes in Bipolar Depression with OCD Comorbidity. *Open Access Library Journal*, **12**: e12894.
https://doi.org/10.4236/oalib.1112894

Received: January 1, 2025
Accepted: February 25, 2025
Published: February 28, 2025

Abstract

Bipolar depression with comorbid obsessive-compulsive disorder (OCD) presents a significant clinical challenge due to its complex symptomatology, unpredictable treatment responses, and high relapse rates. Traditional approaches to treatment planning lack reliable tools for predicting patient-specific outcomes, leaving clinicians with limited options for personalizing care. This study leverages advanced machine learning (ML), specifically XGBoost, to develop a predictive framework capable of classifying treatment responses while identifying key predictors such as age, clinical scores (HDRS, YBOCS), and treatment characteristics (quetiapine dose). By incorporating interpretability techniques such as SHAP (SHapley Additive exPlanations), the model provides transparent insights into how individual features influence predictions, making the outputs actionable for clinical decision-making. Furthermore, probabilistic predictions are evaluated and calibrated using isotonic regression to ensure reliability, particularly for high-stakes applications in psychiatry. Through detailed visual analyses, including confusion matrices, ROC-AUC curves, SHAP plots, and calibration curves, this research bridges the gap between data-driven methodologies and clinical practice, offering a robust framework for advancing personalized treatment strategies in bipolar depression with OCD comorbidity.

Subject Areas

Psychiatry, Machine Learning, Healthcare

Keywords

Bipolar Depression, Obsessive-Compulsive Disorder (OCD), Machine Learning (ML), XGBoost, Treatment Response Prediction, SHAP (SHapley

Additive exPlanations)

1. Introduction

Bipolar depression, when coupled with obsessive-compulsive disorder (OCD), creates a unique and challenging clinical scenario characterized by increased symptom severity, greater treatment resistance, and heightened relapse rates [1]. The coexistence of these two psychiatric conditions often leads to overlapping symptomatology that complicates diagnosis and treatment [2]. For instance, depressive episodes in bipolar disorder may exacerbate obsessive-compulsive tendencies, while OCD symptoms, such as compulsive behaviors and intrusive thoughts, may interfere with mood stabilization efforts [3] [4]. As a result, these patients are at a higher risk of poor clinical outcomes, including suboptimal treatment responses and reduced quality of life. Despite the availability of pharmacological and psychotherapeutic interventions, including mood stabilizers, atypical antipsychotics, and selective serotonin reuptake inhibitors (SSRIs), the variability in treatment outcomes across individuals remains a critical challenge for clinicians [5] [6]. The lack of predictive tools in psychiatry further complicates treatment planning for this vulnerable population. Clinicians often rely on trial-and-error approaches, which are time-consuming and may expose patients to unnecessary side effects or delays in achieving therapeutic benefits [7] [8]. Additionally, while certain medications, such as quetiapine, have shown promise in managing bipolar depression with OCD, there is no standardized method to predict which patients will respond positively to a particular intervention [9]. This creates a critical gap in the ability to deliver personalized care, where treatments are tailored to the specific needs and characteristics of individual patients. In recent years, advancements in machine learning (ML) have revolutionized various fields, including healthcare, by enabling data-driven predictions and uncovering patterns in complex datasets [10]. ML techniques, such as XGBoost, offer the potential to analyze multidimensional clinical data and predict treatment outcomes with greater accuracy than traditional statistical models [11]. Moreover, explainability tools like SHAP (SHapley Additive exPlanations) provide transparency by highlighting the influence of individual features on model predictions, thereby enhancing clinician trust and facilitating actionable insights [12]. When combined with calibration techniques, ML models can further improve the reliability of probabilistic predictions, ensuring that predictions align with observed clinical outcomes [13].

This study aims to harness the power of ML to address the unmet need for predictive tools in managing bipolar depression with OCD comorbidity. Specifically, it seeks to develop a robust and interpretable predictive framework that can classify treatment responses, identify key predictors, and provide reliable probability estimates. By focusing on both accuracy and interpretability, this research bridges the gap between ML-driven methodologies and real-world clinical applications,

offering a pathway to personalized psychiatric care.

2. Research Objectives

This research is guided by the following objectives:

1) Predicting Treatment Responses: To leverage ML models, particularly XGBoost, for accurately classifying treatment responses into well-defined categories, such as "Very Much Improved", "Much Improved", "Minimally Improved", and "No Change" [14].

2) Identifying Influential Predictors: To determine the most critical factors influencing treatment outcomes, such as demographic variables (age, sex), clinical scores (HDRS, YBOCS, YMRS), and treatment characteristics [15] [16] (e.g. quetiapine dose).

3) Evaluating Model Reliability: To assess the reliability of the ML model's predictions through calibration techniques, ensuring that probabilistic outputs are clinically meaningful and aligned with real-world observations [17] [18].

4) Providing Actionable Insights: To use interpretability tools like SHAP to translate complex ML outputs into comprehensible insights, empowering clinicians to make informed treatment decisions and tailor interventions to individual patient profiles [19].

By achieving these objectives, this study aims to advance the field of psychiatry by integrating ML-based predictions with clinical decision-making, ultimately improving treatment outcomes and quality of life for patients with bipolar depression and OCD comorbidity.

3. Literature Review

3.1. Bipolar Depression with OCD Comorbidity

Bipolar depression and obsessive-compulsive disorder (OCD) are both severe psychiatric conditions, and their comorbidity presents unique challenges for diagnosis and treatment [20] [21]. Research indicates that up to 20% of individuals with bipolar disorder may also meet the criteria for OCD, leading to heightened symptom complexity and increased clinical burden [22] [23]. This comorbidity is associated with overlapping symptoms, such as intrusive thoughts and mood dysregulation, which can exacerbate depressive episodes and complicate treatment planning [24]. Moreover, comorbid OCD in bipolar patients is linked to higher rates of treatment resistance, reduced quality of life, and a greater risk of relapse compared to patients with bipolar disorder alone [25].

Pharmacological treatments, including atypical antipsychotics like quetiapine and mood stabilizers such as lithium and valproate, have demonstrated varying degrees of efficacy in managing this dual diagnosis [26]. However, while quetiapine shows promise in alleviating depressive and OCD symptoms, clinicians often lack a reliable framework for predicting individual patient responses [27]. Current treatment strategies rely heavily on trial-and-error approaches, which are both time-consuming and potentially harmful due to the risk of adverse effects, such as

treatment-induced manic episodes [28]. This underscores the urgent need for predictive tools that can guide clinicians in selecting the most effective treatment strategies for this high-risk population.

3.2. Machine Learning in Psychiatry

Machine learning (ML) has emerged as a transformative technology in psychiatry, with applications ranging from diagnostic classification to predicting treatment outcomes in conditions like depression, schizophrenia, and anxiety disorders [29]. ML algorithms, such as decision trees, random forests, and gradient boosting models like XGBoost, are particularly suited for analyzing complex, multidimensional datasets [30]. These techniques enable the identification of patterns and relationships within clinical data that might be overlooked by traditional statistical methods.

In the context of psychiatric comorbidities, ML offers the potential to uncover insights into how combinations of features—such as demographic factors, symptom severity scores, and medication dosages—affect treatment outcomes [31]. However, despite the growing body of research on ML applications in mental health, there is a noticeable gap in its use for bipolar depression with OCD comorbidity [32]. Most studies focus on single disorders, and limited work has addressed the interpretability and clinical utility of ML models in predicting treatment responses for complex, dual-diagnosis populations. This highlights the need for research that not only applies ML to this underserved area but also ensures that the models are interpretable and actionable for clinical decision-making [33].

3.3. Calibration and Interpretability

One of the key challenges in implementing machine learning models in psychiatry is ensuring that predictions are both accurate and clinically actionable [34]. Interpretability tools, such as SHAP (SHapley Additive exPlanations), play a vital role in bridging the gap between complex ML models and clinical applications [35]. SHAP values provide detailed insights into how specific features—such as age, baseline severity scores, or medication doses—influence model predictions, making the outputs comprehensible and trustworthy for clinicians [36]. This is particularly important in psychiatry, where decisions often have significant consequences for patient outcomes. In addition to interpretability, the reliability of probabilistic predictions is crucial in high-stakes domains like psychiatry. Calibration tools ensure that the predicted probabilities generated by ML models correspond closely to observed outcomes. For example, a well-calibrated model predicting a 70% likelihood of treatment improvement should match the actual observed improvement rate of 70% in similar cases [37] [38]. Calibration curves are widely used to assess and improve the reliability of probabilistic outputs, ensuring that models provide clinically meaningful predictions [39]. Without calibration, even highly accurate models can lead to misleading probabilities, undermining their utility in practice [40].

Despite the availability of tools like SHAP and calibration techniques such as isotonic regression, limited studies have applied them to dual-diagnosis conditions like bipolar depression with OCD [41]. Addressing this gap is critical to developing ML models that are not only predictive but also interpretable and reliable, ultimately fostering greater adoption in clinical practice [42].

4. Methodology

4.1. Dataset

This study utilized a simulated clinical dataset designed to reflect real-world characteristics of patients with bipolar depression and comorbid obsessive-compulsive disorder (OCD) [43]. The dataset included the following features:

1) Demographics:

- Age: A continuous variable reflecting the age of the patient, a known factor influencing treatment outcomes in psychiatry [44].
- Sex: A categorical variable (male or female) to account for potential gender-based differences in treatment response.

2) Clinical Scores:

- Hamilton Depression Rating Scale (HDRS): A widely used measure of depression severity [45].
- Yale-Brown Obsessive-Compulsive Scale (YBOCS): A clinical metric assessing OCD symptom severity [46].
- Young Mania Rating Scale (YMRS): A scale evaluating manic symptoms, relevant for bipolar disorder [47].

3) Treatment Variables:

- Quetiapine Dose: A continuous variable capturing the dosage of quetiapine administered, an atypical antipsychotic commonly used to treat bipolar depression and OCD symptoms [48].

The simulated dataset was designed to closely replicate real-world clinical conditions by incorporating key demographic, clinical, and treatment-related variables observed in previous studies of bipolar depression and OCD comorbidity. For example, the distributions for age, clinical scores (HDRS, YBOCS), and quetiapine dosages were modelled based on published statistical data to ensure realistic variability. Furthermore, comorbidity-related complexities, such as overlapping symptom severity, were integrated to mimic clinical heterogeneity. However, it is acknowledged that simulated data cannot fully capture the noise, missing values, and inconsistencies inherent in real-world clinical datasets, which may impact the model's generalizability. This limitation underscores the need for future validation using diverse, real-world datasets from multiple clinical settings.

The target variable was the treatment response, categorized into four distinct classes based on clinical outcomes:

- "Very Much Improved";
- "Much Improved";
- "Minimally Improved";

- "No Change".

This multi-class target variable allowed the model to capture the spectrum of treatment efficacy, ranging from significant improvement to no observable change.

4.2. Machine Learning Framework

To predict treatment responses and derive clinically meaningful insights, a robust machine learning framework was implemented, as described below:

1) Model Selection:

- The XGBoost (Extreme Gradient Boosting) algorithm was chosen for its ability to handle complex, high-dimensional datasets [49]. XGBoost is a gradient boosting framework that builds an ensemble of decision trees, optimizing for predictive performance while preventing overfitting through regularization [50].
- Its ability to handle multi-class classification problems made it a suitable choice for this study, given the four treatment response categories.

2) Evaluation Metrics: To comprehensively evaluate model performance, the following metrics and techniques were used:

- Confusion Matrix: Provided a detailed breakdown of model predictions versus actual outcomes, enabling identification of specific misclassification patterns.
- ROC-AUC (Receiver Operating Characteristic-Area Under the Curve): Measured the model's ability to discriminate between classes, with higher AUC scores indicating better separability.
- SHAP (SHapley Additive exPlanations): Used to interpret the model's predictions by quantifying the contribution of each feature to the outcome, ensuring transparency and clinical applicability.
- Calibration Curves: Assessed the reliability of probabilistic predictions by comparing predicted probabilities to observed frequencies, addressing the critical need for confidence in clinical predictions.

3) Baseline Comparison:

- A logistic regression model was implemented as the baseline for performance benchmarking. Logistic regression, while interpretable and widely used in clinical research, is limited in handling complex, non-linear relationships [51]. Comparing XGBoost to logistic regression helped demonstrate the added value of advanced ML techniques in capturing intricate patterns within the dataset [52].

XGBoost was chosen as the primary model for its ability to handle complex, high-dimensional data and model non-linear feature interactions effectively, outperforming alternatives like Random Forest and SVM in terms of computational efficiency and regularization. Additionally, its compatibility with SHAP values ensures interpretability, critical for clinical applications. Logistic regression was used as a baseline due to its simplicity and widespread use in psychiatry, though its limitations in modelling non-linear relationships highlight the advantages of advanced methods like XGBoost in addressing the complexities of dual-diagnosis conditions.

4.3. Workflow Overview

The methodology followed a systematic workflow to ensure robustness and reliability:

1) Data Preprocessing:

- Data normalization and encoding of categorical variables (e.g. sex).
- Splitting the dataset into training and testing subsets to prevent overfitting and evaluate generalizability.

2) Model Training:

- The XGBoost model was trained using the training data, with hyperparameter tuning to optimize performance. Parameters such as tree depth, learning rate, and number of estimators were fine-tuned using cross-validation.

3) Model Evaluation:

- The trained model was tested on unseen testing data.
- Predictions were analyzed using confusion matrices to identify correct classifications and misclassifications across the four response categories.
- ROC-AUC scores were computed for each class to evaluate discriminatory power.
- SHAP values were analyzed to interpret feature contributions, providing insights into which factors most strongly influenced predictions.

4) Calibration:

- A calibration curve was generated to evaluate the reliability of probabilistic predictions.
- Calibration techniques such as isotonic regression was applied to adjust probability estimates, ensuring that predicted probabilities aligned with observed outcomes.

5) Baseline Comparison:

- The performance of the XGBoost model was compared to logistic regression using the same evaluation metrics. This comparison highlighted the advantages of using advanced ML techniques for complex, multi-class classification problems.

5. Results

5.1. Confusion Matrix

The confusion matrix (Figure 1) visualizes classification accuracy across all response classes. Diagonal values represent correct predictions, while off-diagonal values show misclassifications. The model performed best for "Very Much Improved" (Class 0), but there was significant confusion between "Much Improved" (Class 1) and "Minimally Improved" (Class 2).

5.2. SHAP Summary Plot

SHAP values (Figure 2) highlight the most influential features. Age, HDRS_Baseline, and YBOCS_Baseline emerged as the top predictors. This provides interpretability by showing how specific features contribute to predictions.

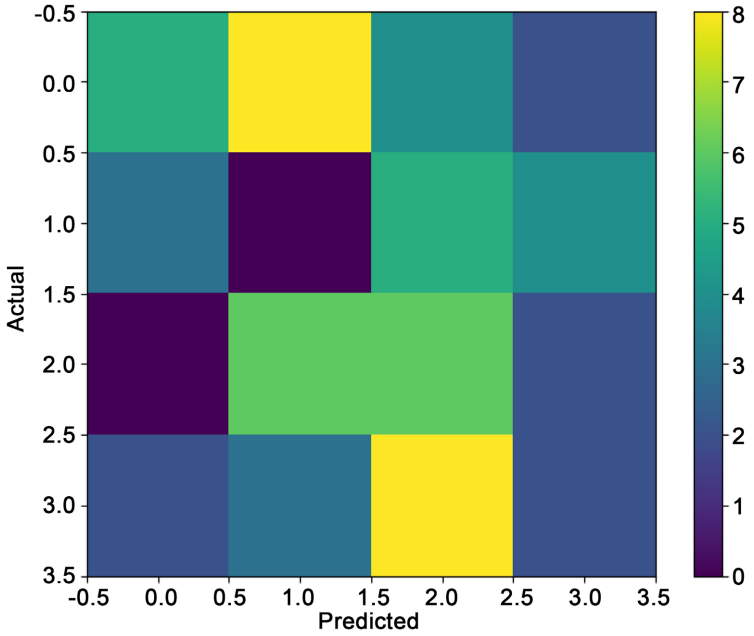

Figure 1. Confusion matrix—classification performance overview.

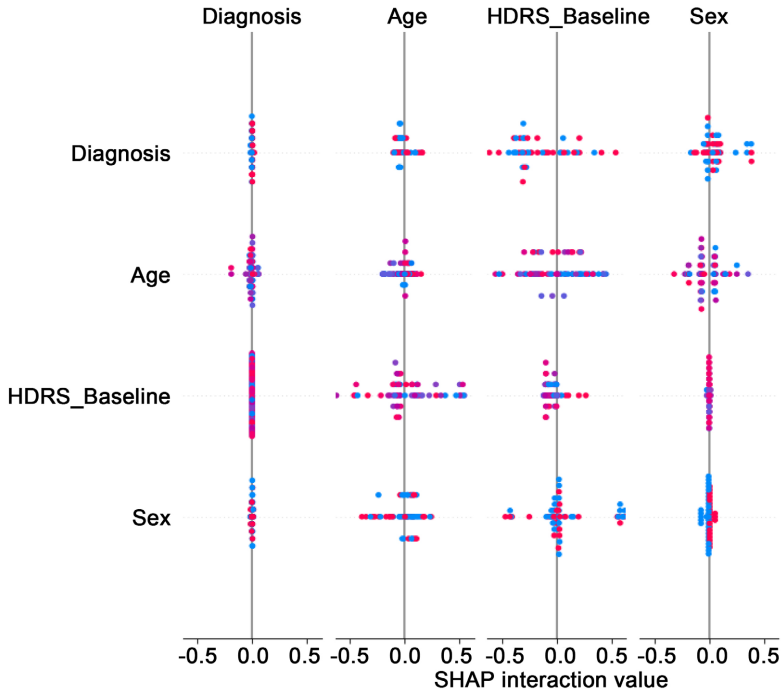

Figure 2. SHAP summary plot—key features driving predictions.

5.3. Annotated Confusion Matrix

Figure 3 adds labels to the confusion matrix, improving clarity about class-specific errors. The highest misclassification rates occurred between Class 1 ("Much Improved") and Class 2 ("Minimally Improved"). The confusion between Classes 1 ("Much Improved") and 2 ("Minimally Improved") can be attributed to overlapping feature distributions and insufficient granularity in the dataset. For example,

similar baseline HDRS or YBOCS scores may not fully capture subtle distinctions in treatment response levels. This overlap highlights the need for additional features that better differentiate these classes.

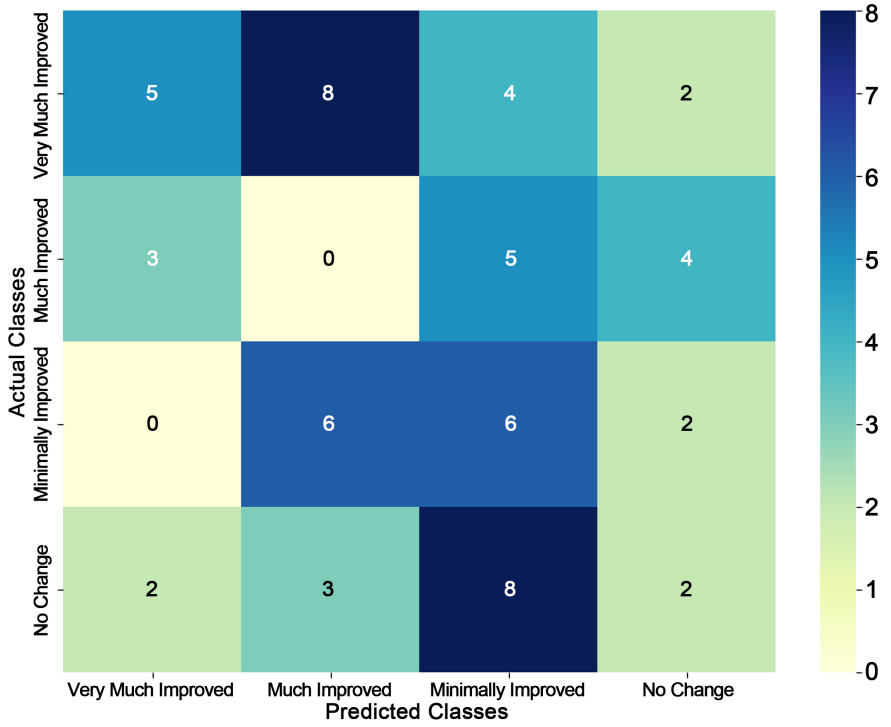

Figure 3. Annotated confusion matrix—error distribution in predictions.

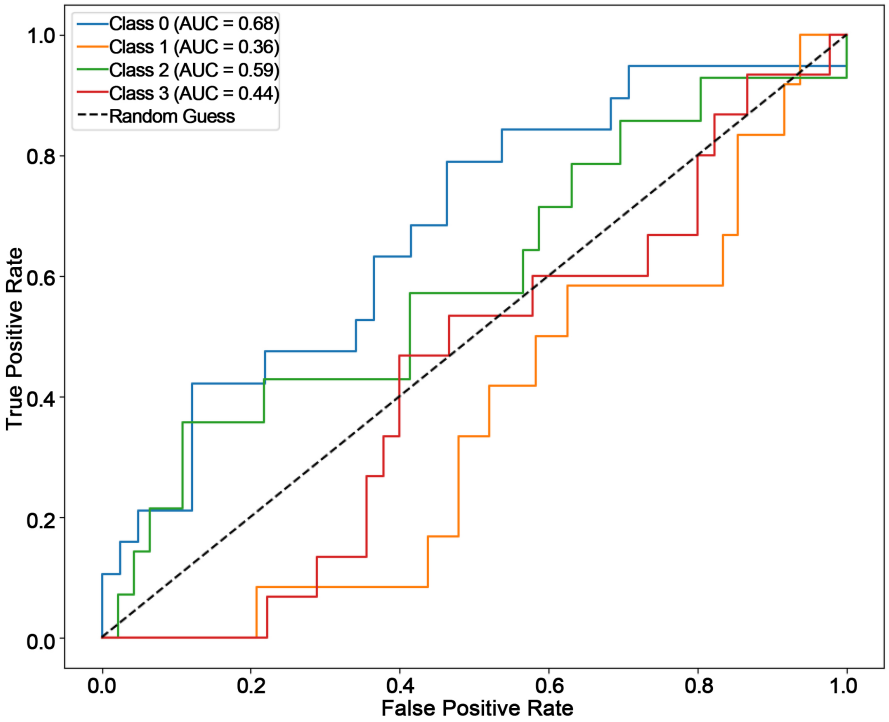

Figure 4. Multi-class ROC-AUC curve—class-specific discrimination ability.

5.4. Multi-Class ROC-AUC Curve

The ROC-AUC curve (Figure 4) evaluates the model's ability to discriminate between classes. Class 0 ("Very Much Improved") achieved the highest AUC (0.68), while the other classes showed weaker separations, indicating areas for improvement.

5.5. Feature Importance

Feature importance (Figure 5) ranks predictors based on their influence in the model. Age and HDRS_Baseline are the most significant predictors, underscoring their relevance in determining outcomes [53] [54].

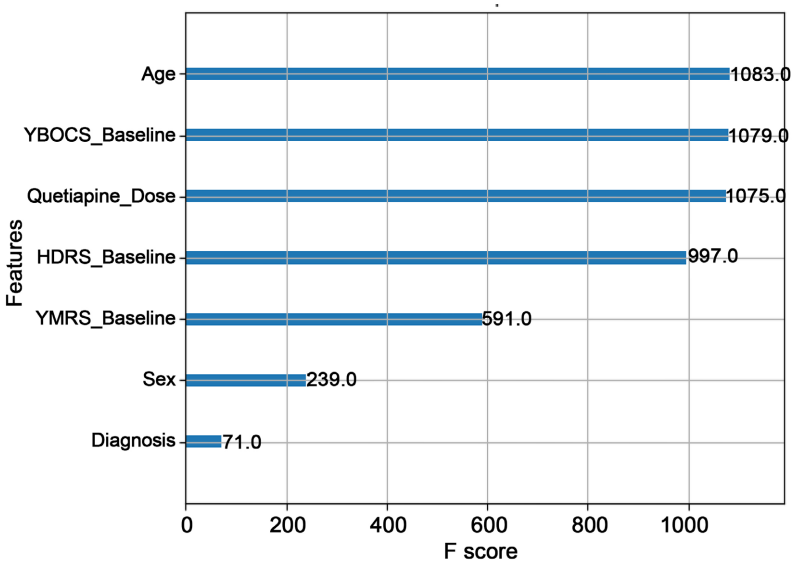

Figure 5. Feature importance—ranked predictors of outcomes.

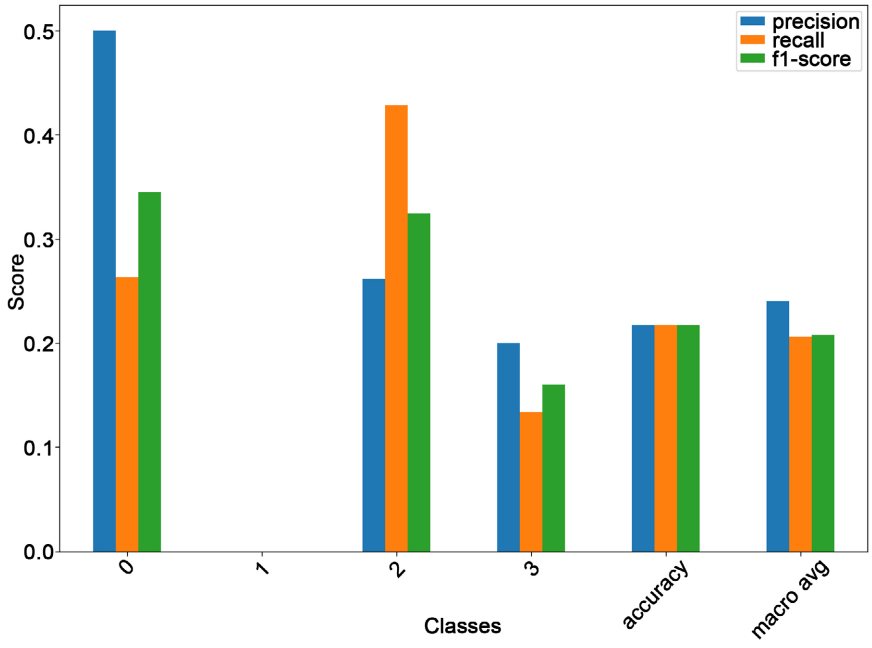

Figure 6. Precision, recall, and F1-score—model performance metrics.

5.6. Precision, Recall, and F1-Score

Figure 6 compares precision, recall, and F1-scores across classes. Class 0 performed best, while metrics for Class 1 and Class 2 revealed challenges in model generalization.

5.7. Accuracy Comparison

Figure 7 compares the accuracy of XGBoost with logistic regression. Interestingly, the baseline model slightly outperformed XGBoost, highlighting the importance of further model refinement.

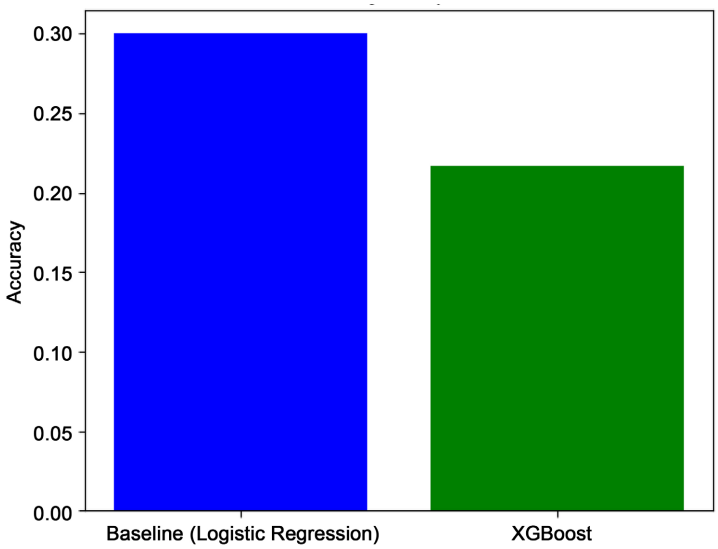

Figure 7. Accuracy comparison—baseline vs. XGBoost.

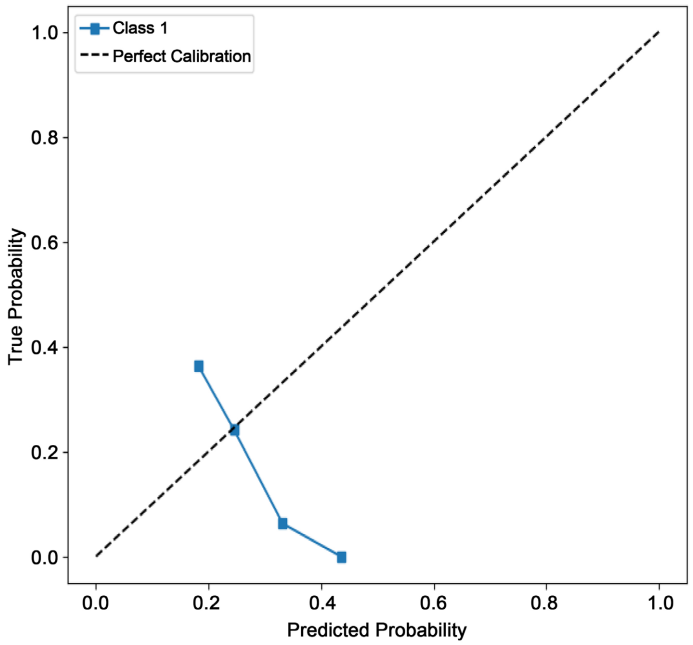

Figure 8. Calibration curve—probability calibration for Class 1.

5.8. Calibration Curve

Figure 8 illustrates the calibration curve for Class 1. The model underestimates probabilities, as evidenced by the curve deviating below the ideal line, suggesting the need for recalibration.

6. Discussion

Insights

The findings from this study highlight several critical insights into the use of machine learning (ML) for predicting treatment responses in bipolar depression with OCD comorbidity [55] [56]. The use of SHAP (SHapley Additive exPlanations) provided interpretable and actionable insights into the model's predictions [57]. Notably, features such as Age and baseline clinical scores, including the Hamilton Depression Rating Scale (HDRS) and the Yale-Brown Obsessive-Compulsive Scale (YBOCS), emerged as the most influential predictors [58]. This aligns with clinical evidence that age and symptom severity are key factors influencing treatment outcomes. By quantifying the contributions of these predictors, SHAP facilitates transparency, making the model's outputs more interpretable for clinicians. Calibration curves further revealed important findings about the reliability of the model's probabilistic predictions. While the XGBoost model demonstrated strong predictive capabilities in terms of classification accuracy and discrimination (as evidenced by the ROC-AUC scores), the calibration analysis indicated a gap in probability alignment, particularly for Class 1 ("Much Improved"). This underlines the need for post-hoc calibration methods, such as isotonic regression, to ensure that predicted probabilities are both reliable and clinically meaningful. Without calibration, even highly accurate models risk misleading clinicians in high-stakes decision-making scenarios, emphasizing the importance of evaluating both accuracy and reliability. Addressing the misclassification between Classes 1 and 2 requires incorporating additional features that capture nuances in symptom progression and treatment dynamics. Temporal data, such as changes in HDRS and YBOCS scores over the treatment period, could provide critical insights into response trajectories. Additionally, incorporating psychosocial variables, medication adherence patterns, or genetic markers could enhance class separability. Future work should explore ensemble methods or hybrid approaches to better capture complex interactions between features and improve classification accuracy.

7. Limitations

Despite its strengths, the study has several limitations that must be addressed:

1) Simulated Data:

- The use of simulated clinical data, while reflective of real-world characteristics, limits the generalizability of the findings. Real-world datasets often contain complexities, such as missing values, noise, and variability in clinical practices, which may impact the model's performance. Thus, validation using real-world clinical datasets is essential to establish the model's practical applicability.

2) Misclassifications in Classes 1 and 2:

- The confusion matrices and performance metrics revealed significant misclassifications between Classes 1 ("Much Improved") and 2 ("Minimally Improved"). This suggests that overlapping feature distributions in these categories may have reduced the model's ability to differentiate between them. Such overlaps are common in clinical datasets where subtle differences in symptom severity or response levels may not be adequately captured by the available features. Addressing this limitation requires feature engineering or the inclusion of additional clinical variables to improve class separability.

3) Probabilistic Predictions:

- While XGBoost provided strong classification performance, its probabilistic outputs were less reliable, particularly for intermediate classes. The calibration curve for Class 1 demonstrated significant deviation from the ideal calibration line, indicating a systematic underestimation of probabilities. This limitation underscores the importance of incorporating calibration techniques to improve the reliability of predictions for clinical use.

8. Future Directions

Building on the insights and limitations of this study, several key directions for future research are proposed:

1) Integration of Real-World Datasets:

- To enhance the robustness and generalizability of the findings, future studies should incorporate real-world clinical datasets. These datasets should include diverse patient populations, varying treatment protocols, and longitudinal follow-up data to capture the complexity of bipolar depression with OCD comorbidity [59]. Real-world validation will also help identify potential biases in the model and refine its performance in practical settings [60] [61].

2) Advanced Ensemble Techniques:

- To address misclassifications and improve model performance for underrepresented classes, ensemble techniques such as stacked models or hybrid approaches can be explored. Combining models with complementary strengths may improve class separability, particularly for Classes 1 and 2, where feature overlap poses challenges. Additionally, techniques like class rebalancing or oversampling methods (e.g. SMOTE) can address imbalances in class representation.

3) Improved Calibration Methods:

- Calibration of probabilistic predictions is critical for clinical reliability [62]. Future research should apply advanced calibration techniques, such as isotonic regression or Platt scaling, to ensure that the model's probabilities align with observed outcomes. Furthermore, calibration methods should be evaluated on a class-specific basis to address variability in prediction reliability across different response categories.

4) Incorporating Temporal Data:

- Future models could benefit from incorporating longitudinal or temporal data

to capture changes in symptom severity over time. This approach could provide dynamic predictions, enabling clinicians to monitor and adjust treatment plans more effectively.

5) Expanding Feature Space:

- Including additional features, such as genetic markers, comorbid medical conditions, and psychosocial factors, may enhance the model's ability to predict treatment outcomes. Multimodal data integration, combining clinical, genetic, and imaging data, could further improve the model's predictive accuracy.

9. Conclusion

This research highlights the transformative potential of machine learning (ML), specifically XGBoost, in predicting treatment outcomes for patients with bipolar depression and OCD comorbidity—a clinically complex and underserved population. By leveraging advanced analytical techniques, the study not only achieves accurate classification of treatment responses, but also provides critical insights into the factors driving these outcomes. The integration of interpretability tools, such as SHAP (SHapley Additive exPlanations), ensures that the model's predictions are transparent and clinically actionable, empowering clinicians to make informed decisions grounded in data. Furthermore, the application of calibration methods addresses the critical need for reliable probabilistic predictions, bridging the gap between theoretical ML models and their practical utility in real-world clinical settings. While the study underscores the efficacy of XGBoost in handling complex, multidimensional clinical data, it also acknowledges limitations, including the use of simulated datasets and challenges in distinguishing closely related response categories. These findings point to opportunities for further refinement through real-world validation, advanced ensemble techniques, and expanded feature sets that capture the nuances of patient characteristics and treatment dynamics. In conclusion, this research establishes a robust framework for integrating ML into personalized psychiatric care, paving the way for more precise, reliable, and interpretable treatment planning. By addressing the critical challenges of prediction accuracy, interpretability, and calibration, this study offers a significant step toward improving outcomes for individuals with bipolar depression and OCD comorbidity, ultimately enhancing the quality and effectiveness of mental health care.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Fineberg, N.A., Hollander, E., Pallanti, S., Walitza, S., Grünblatt, E., Dell'Osso, B.M., *et al.* (2020) Clinical Advances in Obsessive-Compulsive Disorder: A Position Statement by the International College of Obsessive-Compulsive Spectrum Disorders. *International Clinical Psychopharmacology*, **35**, 173-193. https://doi.org/10.1097/yic.0000000000000314

[2] Brady, K.T., Killeen, T.K., Brewerton, T. and Lucerini, S. (2000) Comorbidity of

Psychiatric Disorders and Posttraumatic Stress Disorder. *Journal of Clinical Psychiatry*, **61**, 22-32.

[3] McIntyre, R.S., Alda, M., Baldessarini, R.J., Bauer, M., Berk, M., Correll, C.U., *et al.* (2022) The Clinical Characterization of the Adult Patient with Bipolar Disorder Aimed at Personalization of Management. *World Psychiatry*, **21**, 364-387. https://doi.org/10.1002/wps.20997

[4] Attiullah, N., Eisen, J.L. and Rasmussen, S.A. (2000) Clinical Features of Obsessive-Compulsive Disorder. *Psychiatric Clinics of North America*, **23**, 469-491. https://doi.org/10.1016/s0193-953x(05)70175-1

[5] Bauer, M., Severus, E., Möller, H. and Young, A.H. (2017) Pharmacological Treatment of Unipolar Depressive Disorders: Summary of WFSBP Guidelines. *International Journal of Psychiatry in Clinical Practice*, **21**, 166-176. https://doi.org/10.1080/13651501.2017.1306082

[6] Nemeroff, C.B. (2007) The Burden of Severe Depression: A Review of Diagnostic Challenges and Treatment Alternatives. *Journal of Psychiatric Research*, **41**, 189-206. https://doi.org/10.1016/j.jpsychires.2006.05.008

[7] Dawoodbhoy, F.M., Delaney, J., Cecula, P., Yu, J., Peacock, I., Tan, J., *et al.* (2021) AI in Patient Flow: Applications of Artificial Intelligence to Improve Patient Flow in NHS Acute Mental Health Inpatient Units. *Heliyon*, **7**, e06993. https://doi.org/10.1016/j.heliyon.2021.e06993

[8] Squires, M., Tao, X., Elangovan, S., Gururajan, R., Zhou, X., Acharya, U.R., *et al.* (2023) Deep Learning and Machine Learning in Psychiatry: A Survey of Current Progress in Depression Detection, Diagnosis and Treatment. *Brain Informatics*, **10**, Article No. 10. https://doi.org/10.1186/s40708-023-00188-6

[9] Fineberg, N.A., Reghunandanan, S., Simpson, H.B., Phillips, K.A., Richter, M.A., Matthews, K., *et al.* (2015) Obsessive-Compulsive Disorder (OCD): Practical Strategies for Pharmacological and Somatic Treatment in Adults. *Psychiatry Research*, **227**, 114-125. https://doi.org/10.1016/j.psychres.2014.12.003

[10] Sunny, M.N.M., Saki, M.B.H., Nahian, A.A., Ahmed, S.W., Shorif, M.N., Atayeva, J., *et al.* (2024) Optimizing Healthcare Outcomes through Data-Driven Predictive Modeling. *Journal of Intelligent Learning Systems and Applications*, **16**, 384-402. https://doi.org/10.4236/jilsa.2024.164019

[11] Inoue, T., Ichikawa, D., Ueno, T., Cheong, M., Inoue, T., Whetstone, W.D., *et al.* (2020) XGBoost, a Machine Learning Method, Predicts Neurological Recovery in Patients with Cervical Spinal Cord Injury. *Neurotrauma Reports*, **1**, 8-16. https://doi.org/10.1089/neur.2020.0009

[12] Kalusivalingam, A.K., Sharma, A., Patel, N. and Singh, V. (2021) Leveraging SHAP and LIME for Enhanced Explainability in AI-Driven Diagnostic Systems. *International Journal of AI and ML*, **2**, 1-23.

[13] Staffa, S.J. and Zurakowski, D. (2021) Statistical Development and Validation of Clinical Prediction Models. *Anesthesiology*, **135**, 396-405. https://doi.org/10.1097/aln.0000000000003871

[14] Shende, C., Sahoo, S., Sam, S., Patel, P., Morillo, R., Wang, X., *et al.* (2023) Predicting Symptom Improvement during Depression Treatment Using Sleep Sensory Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **7**, 1-21. https://doi.org/10.1145/3610932

[15] McLaughlin, N.C.R., Dougherty, D.D., Eskandar, E., Ward, H., Foote, K.D., Malone, D.A., *et al.* (2021) Double Blind Randomized Controlled Trial of Deep Brain Stimulation for Obsessive-Compulsive Disorder: Clinical Trial Design. *Contemporary Clinical*

Trials Communications, **22**, Article ID: 100785.
https://doi.org/10.1016/j.conctc.2021.100785

[16] Rodriguez, C.I., Kegeles, L.S., Levinson, A., Feng, T., Marcus, S.M., Vermes, D., *et al.* (2013) Randomized Controlled Crossover Trial of Ketamine in Obsessive-Compulsive Disorder: Proof-of-Concept. *Neuropsychopharmacology*, **38**, 2475-2483. https://doi.org/10.1038/npp.2013.150

[17] D'Amour, A., Heller, K., Moldovan, D., Adlam, B., Alipanahi, B., Beutel, A., Chen, C., *et al.* (2022) Under Specification Presents Challenges for Credibility in Modern Machine Learning. *Journal of Machine Learning Research*, **23**, 1-61.

[18] Thiagarajan, J.J., Venkatesh, B., Rajan, D. and Sattigeri, P. (2020) Improving Reliability of Clinical Models Using Prediction Calibration. In: Sudre, C.H., *et al.*, Eds., *Uncertainty for Safe Utilization of Machine Learning in Medical Imaging, and Graphs in Biomedical Image Analysis*, Springer, 71-80. https://doi.org/10.1007/978-3-030-60365-6_8

[19] Wani, N.A., Kumar, R., Mamta, Bedi, J. and Rida, I. (2024) Explainable AI-Driven IoMT Fusion: Unravelling Techniques, Opportunities, and Challenges with Explainable AI in Healthcare. *Information Fusion*, **110**, Article ID: 102472. https://doi.org/10.1016/j.inffus.2024.102472

[20] Amerio, A., Odone, A., Liapis, C.C. and Ghaemi, S.N. (2014) Diagnostic Validity of Comorbid Bipolar Disorder and Obsessive-Compulsive Disorder: A Systematic Review. *Acta Psychiatrica Scandinavica*, **129**, 343-358. https://doi.org/10.1111/acps.12250

[21] Mucci, F., Toni, C., Favaretto, E., Vannucchi, G., Marazziti, D. and Perugi, G. (2019) Obsessive-Compulsive Disorder with Comorbid Bipolar Disorders: Clinical Features and Treatment Implications. *Current Medicinal Chemistry*, **25**, 5722-5730. https://doi.org/10.2174/0929867324666171108145127

[22] de Filippis, R., Aguglia, A., Costanza, A., Benatti, B., Placenti, V., Vai, E., *et al.* (2024) Obsessive-Compulsive Disorder as an Epiphenomenon of Comorbid Bipolar Disorder? An Updated Systematic Review. *Journal of Clinical Medicine*, **13**, Article 1230. https://doi.org/10.3390/jcm13051230

[23] Sagman, D. and Tohen, M. (2012) Comorbidity in Bipolar Disorder: The Complexity of Diagnosis and Treatment. *Psychiatric Times*, **29**, 30.

[24] Belzer, K. and Schneier, F.R. (2004) Comorbidity of Anxiety and Depressive Disorders: Issues in Conceptualization, Assessment, and Treatment. *Journal of Psychiatric Practice*, **10**, 296-306. https://doi.org/10.1097/00131746-200409000-00003

[25] McIntyre, R.S., Konarski, J.Z. and Yatham, L.N. (2004) Comorbidity in Bipolar Disorder: A Framework for Rational Treatment Selection. *Human Psychopharmacology: Clinical and Experimental*, **19**, 369-386. https://doi.org/10.1002/hup.612

[26] Martinotti, G., Chiappini, S., Mosca, A., Miuli, A., Santovito, M.C., Pettorruso, M., *et al.* (2022) Atypical Antipsychotic Drugs in Dual Disorders: Current Evidence for Clinical Practice. *Current Pharmaceutical Design*, **28**, 2241-2259. https://doi.org/10.2174/1381612828666220623092853

[27] Chekroud, A.M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., *et al.* (2021) The Promise of Machine Learning in Predicting Treatment Outcomes in Psychiatry. *World Psychiatry*, **20**, 154-170. https://doi.org/10.1002/wps.20882

[28] Ochs, L. (2006) The Low Energy Neurofeedback System (LENS): Theory, Background, and Introduction. *Journal of Neurotherapy*, **10**, 5-39. https://doi.org/10.1300/j184v10n02_02

[29] Shatte, A.B.R., Hutchinson, D.M. and Teague, S.J. (2019) Machine Learning in Mental Health: A Scoping Review of Methods and Applications. *Psychological Medicine*, **49**, 1426-1448. https://doi.org/10.1017/s0033291719000151

[30] Kavzoglu, T. and Teke, A. (2022) Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). *Arabian Journal for Science and Engineering*, **47**, 7367-7385. https://doi.org/10.1007/s13369-022-06560-8

[31] Perna, G., Alciati, A., Daccò, S., Grassi, M. and Caldirola, D. (2020) Personalized Psychiatry and Depression: The Role of Sociodemographic and Clinical Variables. *Psychiatry Investigation*, **17**, 193-206. https://doi.org/10.30773/pi.2019.0289

[32] Fisher, A.L., Arora, K., Maehashi, S., Schweitzer, D. and Akefe, I.O. (2024) Unveiling the Neurolipidome of Obsessive-Compulsive Disorder: A Scoping Review Navigating Future Diagnostic and Therapeutic Applications. *Neuroscience & Biobehavioral Reviews*, **166**, Article ID: 105885. https://doi.org/10.1016/j.neubiorev.2024.105885

[33] Pyles, D.A.M., Muniz, K., Cade, A. and Silva, R. (1997) A Behavioral Diagnostic Paradigm for Integrating Behavior-Analytic and Psychopharmacological Interventions for People with a Dual Diagnosis. *Research in Developmental Disabilities*, **18**, 185-214. https://doi.org/10.1016/s0891-4222(97)00003-6

[34] Tai, A.M.Y., Albuquerque, A., Carmona, N.E., Subramanieapillai, M., Cha, D.S., Sheko, M., *et al.* (2019) Machine Learning and Big Data: Implications for Disease Modeling and Therapeutic Discovery in Psychiatry. *Artificial Intelligence in Medicine*, **99**, Article ID: 101704. https://doi.org/10.1016/j.artmed.2019.101704

[35] Rane, N., Choudhary, S. and Rane, J. (2023) Explainable Artificial Intelligence (XAI) in Healthcare: Interpretable Models for Clinical Decision Support. *SSRN Electronic Journal*.

[36] Kanyongo, W. and Ezugwu, A.E. (2023) Feature Selection and Importance of Predictors of Non-Communicable Diseases Medication Adherence from Machine Learning Research Perspectives. *Informatics in Medicine Unlocked*, **38**, Article ID: 101232. https://doi.org/10.1016/j.imu.2023.101232

[37] Van Calster, B., McLernon, D.J., van Smeden, M., Wynants, L. and Steyerberg, E.W. (2019) Calibration: The Achilles Heel of Predictive Analytics. *BMC Medicine*, **17**, Article No. 230. https://doi.org/10.1186/s12916-019-1466-7

[38] Peetluk, L.S., Rebeiro, P.F., Ridolfi, F.M., Andrade, B.B., Cordeiro-Santos, M., Kritski, A., Durovni, B., *et al.* (2022) A Clinical Prediction Model for Unsuccessful Pulmonary Tuberculosis Treatment Outcomes. *Clinical Infectious Diseases*, **74**, 973-982.

[39] Huang, Y., Li, W., Macheret, F., Gabriel, R.A. and Ohno-Machado, L. (2020) A Tutorial on Calibration Measurements and Calibration Models for Clinical Prediction Models. *Journal of the American Medical Informatics Association*, **27**, 621-633. https://doi.org/10.1093/jamia/ocz228

[40] Keren, G. (1991) Calibration and Probability Judgements: Conceptual and Methodological Issues. *Acta Psychologica*, **77**, 217-273. https://doi.org/10.1016/0001-6918(91)90036-y

[41] Tiet, Q.Q. and Mausbach, B. (2007) Treatments for Patients with Dual Diagnosis: A Review. *Alcoholism*: *Clinical and Experimental Research*, **31**, 513-536. https://doi.org/10.1111/j.1530-0277.2007.00336.x

[42] Anuyah, S., Singh, M.K. and Nyavor, H. (2024) Advancing Clinical Trial Outcomes Using Deep Learning and Predictive Modelling: Bridging Precision Medicine and Patient-Centered Care. arXiv: 2412.07050.

[43] Meehan, A.J., Lewis, S.J., Fazel, S., Fusar-Poli, P., Steyerberg, E.W., Stahl, D., *et al.* (2022) Clinical Prediction Models in Psychiatry: A Systematic Review of Two Decades of Progress and Challenges. *Molecular Psychiatry*, **27**, 2700-2708.

https://doi.org/10.1038/s41380-022-01528-4

[44] Souery, D., Oswald, P., Massat, I., Bailer, U., Bollen, J., Demyttenaere, K., *et al.* (2007) Clinical Factors Associated with Treatment Resistance in Major Depressive Disorder: Results from a European Multicenter Study. *The Journal of Clinical Psychiatry*, **68**, 1062-1070. https://doi.org/10.4088/jcp.v68n0713

[45] Gibbons, R.D., Clark, D.C. and Kupfer, D.J. (1993) Exactly What Does the Hamilton Depression Rating Scale Measure? *Journal of Psychiatric Research*, **27**, 259-273. https://doi.org/10.1016/0022-3956(93)90037-3

[46] Rosario-Campos, M.C., Miguel, E.C., Quatrano, S., Chacon, P., Ferrao, Y., Findley, D., *et al.* (2006) The Dimensional Yale-Brown Obsessive-Compulsive Scale (DY-BOCS): An Instrument for Assessing Obsessive-Compulsive Symptom Dimensions. *Molecular Psychiatry*, **11**, 495-504. https://doi.org/10.1038/sj.mp.4001798

[47] Rabinowitz, J., Young, R.C., Yavorsky, C., Williams, J.B.W., Sedway, J., Marino, P., *et al.* (2024) Consistency Checks to Improve Measurement with the Young Mania Rating Scale (YMRS). *Journal of Affective Disorders*, **345**, 24-31. https://doi.org/10.1016/j.jad.2023.10.098

[48] Bandelow, B., Chouinard, G., Bobes, J., Ahokas, A., Eggens, I., Liu, S., *et al.* (2009) Extended-Release Quetiapine Fumarate (Quetiapine XR): A Once-Daily Monotherapy Effective in Generalized Anxiety Disorder. Data from a Randomized, Double-Blind, Placebo- and Active-Controlled Study. *The International Journal of Neuropsychopharmacology*, **13**, 305-320. https://doi.org/10.1017/s1461145709990423

[49] Kavzoglu, T. and Teke, A. (2022) Advanced Hyperparameter Optimization for Improved Spatial Prediction of Shallow Landslides Using Extreme Gradient Boosting (XGBoost). *Bulletin of Engineering Geology and the Environment*, **81**, Article No. 201. https://doi.org/10.1007/s10064-022-02708-w

[50] Trizoglou, P., Liu, X. and Lin, Z. (2021) Fault Detection by an Ensemble Framework of Extreme Gradient Boosting (XGBoost) in the Operation of Offshore Wind Turbines. *Renewable Energy*, **179**, 945-962. https://doi.org/10.1016/j.renene.2021.07.085

[51] Levy, J.J. and O'Malley, A.J. (2020) Don't Dismiss Logistic Regression: The Case for Sensible Extraction of Interactions in the Era of Machine Learning. *BMC Medical Research Methodology*, **20**, Article No. 171. https://doi.org/10.1186/s12874-020-01046-3

[52] Fatima, S., Hussain, A., Amir, S.B., Ahmed, S.H. and Aslam, S.M.H. (2023) XGBoost and Random Forest Algorithms: An in Depth Analysis. *Pakistan Journal of Scientific Research*, **3**, 26-31. https://doi.org/10.57041/pjosr.v3i1.946

[53] de Filippis, R. and Foysal, A.A. (2024) Case Report: The Role of Monoamine Oxidase Inhibitors in Treating Resistant Depression. *Open Access Library Journal*, **11**, 1-12. https://doi.org/10.4236/oalib.1112369

[54] de Filippis, R. and Foysal, A.A. (2024) Case Report: The Role of Tricyclic Antidepressants in Treating Resistant Depression. *Open Access Library Journal*, **11**, 1-12. https://doi.org/10.4236/oalib.1112368

[55] Adler, D.A., Yang, Y., Viranda, T., Xu, X., Mohr, D.C., Van Meter, A.R., *et al.* (2024) Beyond Detection: Towards Actionable Sensing Research in Clinical Mental Healthcare. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, **8**, 1-33. https://doi.org/10.1145/3699755

[56] de Filippis, R. and Foysal, A.A. (2024) Evaluating Pharmacological and Rehabilitation Strategies for Effective Management of Bipolar Disorder: A Comprehensive Clinical Study. *Advances in Bioscience and Biotechnology*, **15**, 406-431. https://doi.org/10.4236/abb.2024.157025

[57] Feretzakis, G., Sakagianni, A., Anastasiou, A., Kapogianni, I., Bazakidou, E., Koufopoulos,

P., *et al*. (2024) Integrating Shapley Values into Machine Learning Techniques for Enhanced Predictions of Hospital Admissions. *Applied Sciences*, **14**, Article 5925. https://doi.org/10.3390/app14135925

[58] Hazari, N., Narayanaswamy, J.C. and Arumugham, S.S. (2016) Predictors of Response to Serotonin Reuptake Inhibitors in Obsessive-Compulsive Disorder. *Expert Review of Neurotherapeutics*, **16**, 1175-1191. https://doi.org/10.1080/14737175.2016.1199960

[59] Cederlöf, M., Lichtenstein, P., Larsson, H., Boman, M., Rück, C., Landén, M., *et al*. (2014) Obsessive-Compulsive Disorder, Psychosis, and Bipolarity: A Longitudinal Cohort and Multigenerational Family Study. *Schizophrenia Bulletin*, **41**, 1076-1083. https://doi.org/10.1093/schbul/sbu169

[60] de Filippis, R. and Foysal, A.A. (2024) Securing Predictive Psychological Assessments: The Synergy of Blockchain Technology and Artificial Intelligence. *Open Access Library Journal*, **11**, 1-23. https://doi.org/10.4236/oalib.1112378

[61] de Filippis, R. and Foysal, A.A. (2024) Decoding Emotions: How AI and Machine Learning Unravel the Human Psyche. *Voice of the Publisher*, **10**, 382-399. https://doi.org/10.4236/vp.2024.104030

[62] de Filippis, R. and Foysal, A.A. (2024) Harnessing the Power of Artificial Intelligence in Neuromuscular Disease Rehabilitation: A Comprehensive Review and Algorithmic Approach. *Advances in Bioscience and Biotechnology*, **15**, 289-309. https://doi.org/10.4236/abb.2024.155018