Open Access Library Journal
2025, Volume 12, e13869
ISSN Online: 2333-9721
ISSN Print: 2333-9705

A Machine Learning Framework for Mood State Classification in Bipolar Disorder Using Clinical Features

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 Framework for Mood State Classification in Bipolar Disorder Using Clinical Features. *Open Access Library Journal*, **12**: e13869.
https://doi.org/10.4236/oalib.1113869

Received: June 28, 2025
Accepted: August 18, 2025
Published: August 21, 2025

Abstract

Bipolar disorder is a complex psychiatric condition characterized by alternating mood episodes, ranging from depression to mania. Accurate and timely detection of a patient's current mood state is critical for optimizing treatment strategies and preventing relapse. However, traditional clinical assessments are often subjective and prone to variability. This study proposes a data-driven, machine learning-based framework to classify mood states using synthetic but clinically informed patient data. We generated a dataset of 2000 virtual patients, incorporating key clinical variables such as lithium levels, sleep duration, stress levels, medication adherence, therapy attendance, and family history. Advanced feature engineering derived clinically relevant variables including therapeutic lithium ranges, adherence-stress ratios, and mood score formulations to reflect real-world variability and interaction effects. The pipeline integrates robust preprocessing techniques, SMOTE for class imbalance, and a stacked ensemble classifier combining Random Forest and XGBoost as base learners with logistic regression as the meta-classifier. Model evaluation across multiple metrics revealed an overall classification accuracy of 61%, with particularly strong performance in identifying manic states (F1 = 0.77). Visualization tools—including ROC curves, confusion matrices, feature importance plots, and PCA-reduced decision boundaries—were employed to enhance interpretability and clinical relevance. The model identified stress level, sleep deviation, and medication adherence as key predictors, aligning well with established psychiatric insights. While classification of euthymic and depressed states remains more challenging, this work demonstrates the feasibility and clinical utility of machine learning approaches in mood state prediction. It sets the stage for further research with real-world clinical data and emphasizes the importance of interpretable, feature-rich models in psychiatric decision support.

Subject Areas

Computational Psychiatry, Machine Learning

Keywords

Bipolar Disorder, Mood Classification, Machine Learning, Ensemble Learning, Explainable AI, Clinical Decision Support, Feature Engineering, Synthetic Data, Computational Psychiatry, Digital Health

1. Introduction

Bipolar disorder (BD) is a chronic and recurrent psychiatric condition marked by oscillating episodes of depression, euthymia, and mania [1]-[4]. These fluctuating mood states significantly impact a patient's cognition, functioning, and quality of life [5]-[8]. Accurate identification of a patient's current mood state is crucial for timely clinical interventions, medication adjustments, and the prevention of relapses [9]-[11]. Despite its clinical importance, mood state classification largely depends on patient self-reports, clinical interviews, and subjective rating scales, which are susceptible to bias, inconsistency, and underreporting [12]-[14]. Recent advances in machine learning (ML) and computational psychiatry offer promising avenues to address this diagnostic gap [15]-[17]. The integration of behavioural, biochemical, and therapy-related variables into predictive models has the potential to improve the objectivity and reproducibility of mood state classification [18]-[21]. However, many existing models suffer from limitations such as small sample sizes, lack of interpretability, and inadequate handling of class imbalance, particularly underrepresentation of transitional states like euthymia [22]-[27]. In this study, we present a comprehensive machine learning framework designed to classify mood states in patients with bipolar disorder using simulated clinical data enriched with domain knowledge. We employ realistic data generation for 2000 virtual patients, incorporating features including lithium levels, medication adherence, sleep patterns, stress levels, therapy attendance, and family psychiatric history. Advanced feature engineering introduces nonlinear clinical interactions such as therapeutic lithium ranges, adherence-stress ratios, and sleep deviation scores. To address the heterogeneity of the data and enhance generalization, we utilize a stacked ensemble classifier composed of Random Forest and XGBoost base models, with a logistic regression meta-learner. Our pipeline includes robust preprocessing, SMOTE-based balancing, and interpretable visualizations such as confusion matrices, ROC curves, and decision boundaries. This approach demonstrates the feasibility of mood state prediction with strong clinical interpretability, laying the groundwork for deployment in real-world psychiatric settings.

2. Methodology

2.1. Pipeline Overview

We developed a modular, end-to-end machine learning pipeline aimed at classifying mood states in patients with bipolar disorder. The pipeline is designed for interpretability, clinical realism, and scalability. It consists of these primary components: synthetic clinical data generation, clinically informed feature engineering, preprocessing and transformation, ensemble model training and tuning, and evaluation with interpretability outputs. The comprehensive structure is visually outlined in Figure 1, capturing each phase from raw data synthesis to model deployment.

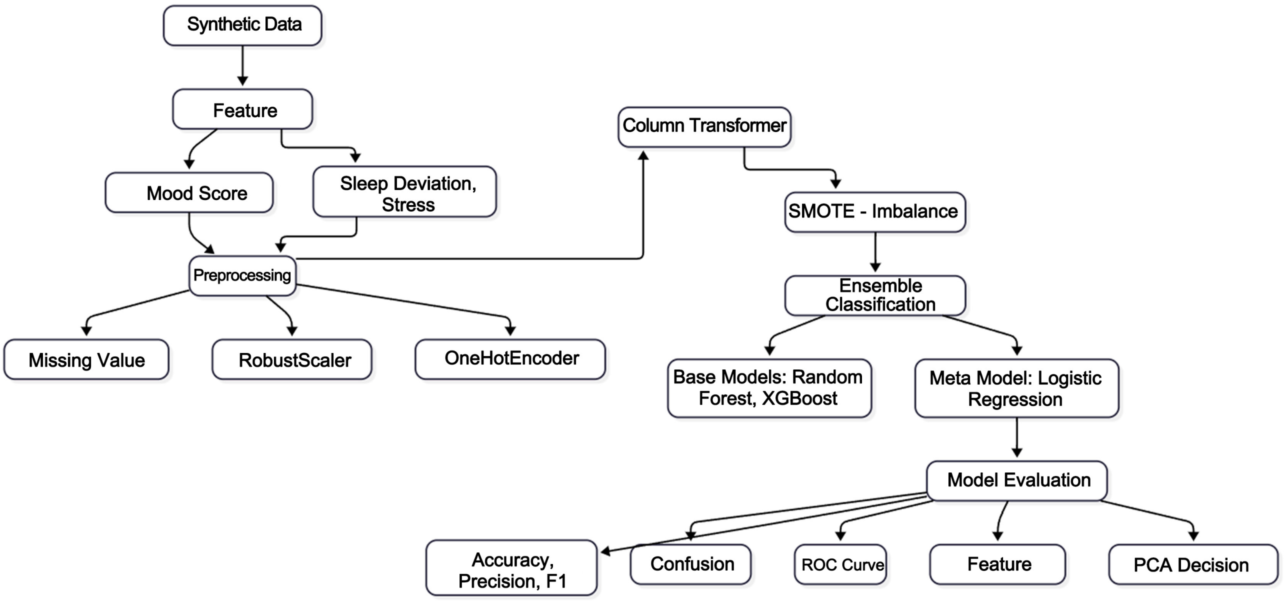

Figure 1. End-to-end AI-based machine learning pipeline.

2.2. Data Generation and Feature Engineering

We constructed a synthetic dataset representing 2000 patients diagnosed with bipolar disorder. The data was designed to mirror real-world clinical distributions and variance, incorporating variables such as:

- **Demographics**: Age (18 - 80 years), Gender (male/female)
- **Biological Measures**: Lithium serum levels
- **Behavioural Indicators**: Sleep hours, physical activity, social interaction
- **Clinical History**: Mood swing frequency, previous episodes, family history
- **Adherence Metrics**: Lithium adherence (%), therapy attendance (%)
- **Environmental Stress**: Perceived stress level (scale 0 - 10)

A continuous **mood score** was computed using a linear combination of the above features, weighted based on psychiatric relevance (e.g., higher weight to lithium level and stress). This continuous variable was then thresholder into three categories: **Depressed**, **Euthymic**, and **Manic**.

To increase model expressivity, we applied advanced feature engineering based on domain knowledge:

- **Therapeutic Lithium Range**: Binary indicator identifying if the lithium level lies between 0.6 and 1.0 mmol/L [28]-[30].
- **Sleep Deviation**: Absolute difference from 7 hours, considered optimal for mental health [31]-[33].
- **Adherence-Stress Ratio**: Captures patient stability by contrasting medication compliance against stress levels [34] [35].
- **High-Risk Flag**: Combines previous episodes >5 and positive family history to indicate recurrence risk [36]-[38].

The continuous mood score was discretized into three clinically relevant categories:

- Depressed: Mood Score ≤ -0.5
- Euthymic: $-0.5 <$ Mood Score < 0.5
- Manic: Mood Score ≥ 0.5

These thresholds were chosen to reflect a neutral clinical mood band around euthymia, where mood states within ±0.5 deviations are generally considered stable. Deviations beyond this band typically signify clinically observable manic or depressive episodes, consistent with diagnostic guidelines.

These engineered features introduced critical nonlinear dependencies often seen in psychiatric evaluations but rarely modelled explicitly in computational frameworks.

Synthetic Data Generation Details

Each clinical feature was generated using domain-informed distributions and parameter ranges to simulate realistic patient variability below in **Table 1**:

Table 1. Clinical feature distributions and parameter settings for synthetic data generation.

Feature	Distribution	Parameters	Range
Age	Uniform	18 to 80 years	[18, 80]
Lithium Level (mmol/L)	Normal	$\mu = 0.7$, $\sigma = 0.2$	[0.3, 1.2]
Sleep Hours	Normal	$\mu = 7$, $\sigma = 1.5$	[3, 12]
Stress Level	Uniform	0 to 10	[0, 10]
Lithium Adherence (%)	Beta (skewed high)	$\alpha = 5$, $\beta = 2$	[0, 100]
Therapy Attendance (%)	Beta	$\alpha = 4$, $\beta = 3$	[0, 100]
Mood Swing Frequency	Poisson	$\lambda = 2$	≥0
Family History	Bernoulli	p = 0.4	0 or 1

The continuous mood score was computed as follows:

$$\begin{aligned}\text{Mood Score} = {} & 0.4\times\text{Lithium Level} - 0.3\times\text{Stress Level} - 0.2\times\text{Sleep Deviation} \\ & + 0.15\times\text{Lithium Adherence} - 0.1\times\text{Therapy Attendance} \\ & + 0.05\times\text{Family History}\end{aligned}$$

Sleep Deviation = |Sleep Hours − 7|

Weights were selected based on psychiatric literature indicating that lithium levels and stress are the strongest mood determinants, while adherence and therapy contribute moderate effects.

2.3. Preprocessing and Imputation

Robust preprocessing was conducted to prepare the dataset for machine learning. Numerical features with missing values were imputed using the **median**, ensuring resistance to skewed distributions [39]-[41]. Categorical variables were imputed with their **mode** (most frequent value). To maintain numerical stability across features, we employed **RobustScaler**, which scales features based on interquartile ranges, minimizing the effect of outliers. A ColumnTransformer was used to separate numeric and categorical pipelines, each applying their own imputation and transformation sequence. For categorical variables, we applied OneHotEncoding, which preserves all unique values while enabling compatibility with tree-based models.

2.4. Model Architecture and Training

We employed a **stacked ensemble model** for classification, leveraging the strengths of diverse learners. The ensemble comprises:

- **Base Learners**:
    - **Random Forest Classifier**: An ensemble of decision trees that handles non-linearities and provides intrinsic feature importance.
    - **XGBoost Classifier**: A gradient boosting model offering fine control over regularization and tree structure, effective for tabular clinical data.
- **Meta-Learner**:
    - **Logistic Regression (Multinomial)**: Trained on predictions of the base models, providing probabilistic outputs and smooth decision boundaries.

To address the issue of class imbalance—particularly the underrepresentation of the Euthymic and Depressed states—we used **SMOTE (Synthetic Minority Over-sampling Technique)** after feature transformation. SMOTE generates synthetic examples in the feature space to balance the class distribution without mere duplication [42]-[45]. The simulated dataset resulted in a class imbalance favouring the manic category. This design partially reflects real-world clinical presentations, where manic episodes often trigger clinical attention and hospitalization more than euthymic or depressive states. Additionally, the imbalance was intentionally introduced to prioritize the detection of manic states due to their higher clinical risk and urgency. Class imbalance was partially addressed using SMOTE to synthetically balance underrepresented mood states. Hyperparameters were optimized using a grid search within stratified 5-fold cross-validation. The following ranges and selections were applied below in Table 2.

The logistic regression meta-learner used an L2 penalty with default solver settings for multinomial classification.

Table 2. Hyperparameter tuning grid and selected model configurations.

Model	Hyperparameter	Tuning Range	Selected Value
Random Forest	Number of Trees (n_estimators)	100, 300, 500	300
	Maximum Tree Depth (max_depth)	5, 10, None	None
	Minimum Samples Split (min_samples_split)	2, 5, 10	5
XGBoost	Learning Rate (learning_rate)	0.01, 0.1	0.1
	Maximum Tree Depth (max_depth)	3, 6, 10	6
	Number of Trees (n_estimators)	100, 200	200
	Subsample Ratio (subsample)	0.8, 1.0	1.0

The pipeline integrates all transformations, oversampling, and learning stages using Pipeline and make_pipeline constructs from sklearn and imblearn. Model validation was performed using **stratified 5-fold cross-validation** to preserve the proportion of each mood class across folds. Performance was assessed not only by accuracy but also via class-specific precision, recall, F1-score, and ROC-AUC.

The modularity of the pipeline ensures adaptability to real-world datasets and facilitates direct interpretability through integrated visual diagnostics and feature importance metrics [46]-[48].

3. Results

3.1. Clinical Visualization

To better understand the distribution of key clinical variables across mood states, we generated a series of visualizations that are collectively presented in Figure 2. These plots serve to highlight behavioural, biochemical, and psychological trends relevant to the classification task.

- Figure 2(A)**: Lithium Levels by Mood State** A boxplot illustrates the distribution of lithium serum levels across Depressed, Euthymic, and Manic patients. Notably, manic individuals tend to have higher lithium levels on average, although the variance is substantial [49] [50]. Depressed patients show a wider spread, potentially due to poor adherence or metabolic variation.
- Figure 2(B)**: Sleep Duration by Mood State** The violin plot reveals distinct sleep patterns. Manic patients exhibit reduced sleep durations, consistent with clinical profiles. In contrast, depressed individuals show extended sleep variability with a tail toward hypersomnia. Euthymic subjects are more cantered around the optimal range (~7 hours).
- Figure 2(C)**: Stress Levels by Mood State** Boxen plots visualize the stress distribution, showing that depressed patients consistently report higher stress, while euthymic individuals have the lowest median stress levels. This pattern supports the inclusion of stress as a significant predictor.
- Figure 2(D)**: Lithium Adherence Distribution** A histogram stratified by mood state demonstrates that manic patients generally show higher adherence

levels, while depressed patients often fall below the population mean. This feature may reflect therapy engagement or cognitive motivation.

- **Figure 2(E): Mood Score Distribution** A KDE plot highlights how the engineered mood_score effectively separates the three classes. Manic scores cluster on the positive side, depressed scores skew negative, and euthymic scores cluster near zero, affirming the variable's utility in labelling.
- **Figure 2(F): Correlation Heatmap** This matrix shows correlations among selected features including lithium level, adherence, sleep, stress, and current mood. Notably, stress and mood exhibit a moderate negative correlation (approximately −0.27), while adherence and mood are positively correlated.

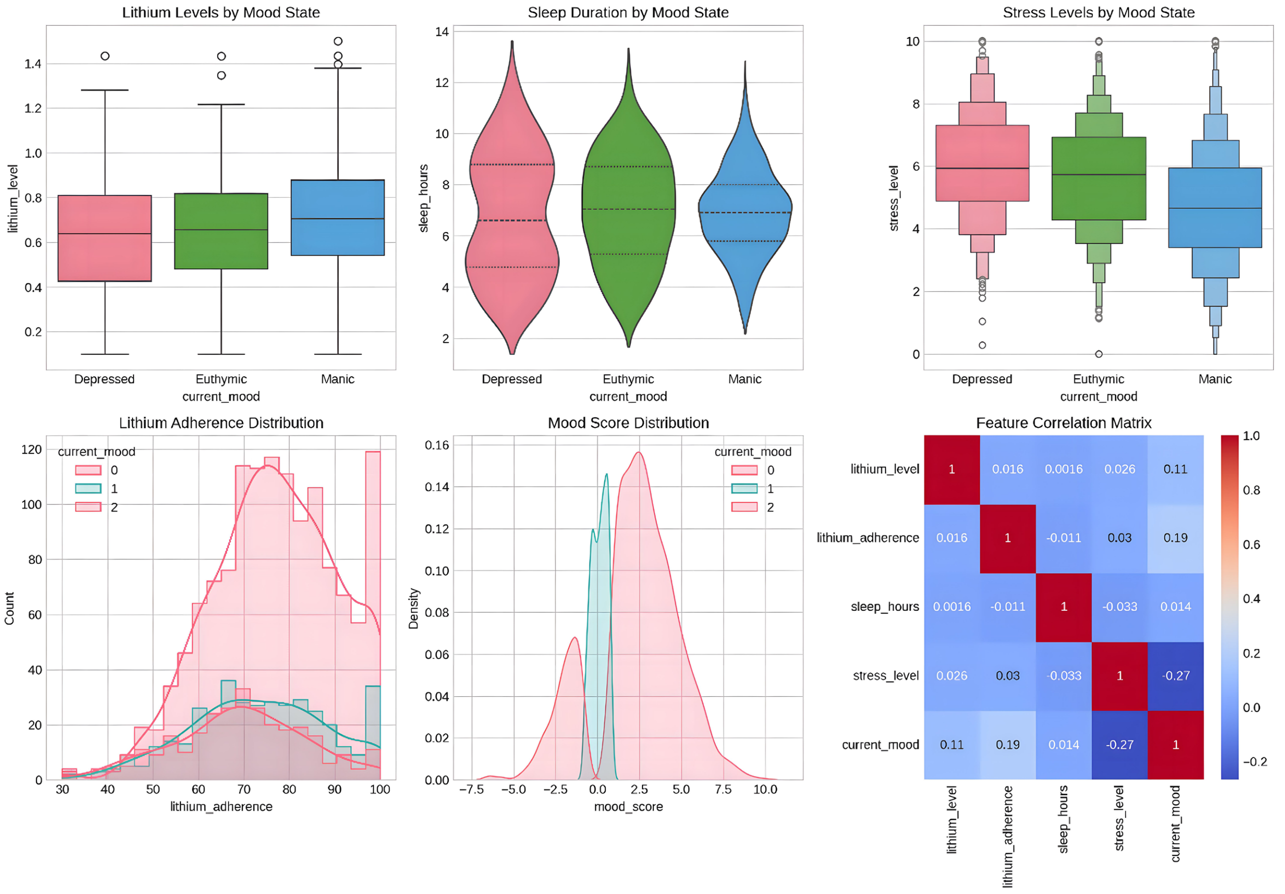

Figure 2. Clinical visualizations. (A) Lithium level by mood. (B) Sleep duration. (C) Stress levels. (D) Lithium adherence. (E) Mood score distribution. (F) Correlation heatmap.

These insights confirm the clinical validity of the features and justify their inclusion in the model pipeline. Together, the subplots in Figure 2 provide visual evidence that mood state classification is supported by underlying patterns in both behavioural and biochemical data.

3.2. Classification Metrics

The classification performance of the stacked ensemble model was evaluated using standard multi-class classification metrics: **accuracy**, **precision**, **recall**, and **F1-score**. These metrics were computed on a held-out test set comprising 25% of the dataset (500 patients).

The overall performance results are as follows:

- **Accuracy**: 61%—indicating that the model correctly predicted the mood state in 61 out of every 100 patients.
- **Precision**: 54.8%—representing the weighted average proportion of correct positive identifications among all predicted positives.
- **Recall**: 61%—indicating the model's ability to capture all relevant cases within each mood class.
- **F1-Score**: 56.9%—a harmonic mean balancing both precision and recall, reflecting the trade-off between false positives and false negatives.

Performance was further broken down by mood category in Table 3.

Table 3. Performance metrics by mood state showing precision, recall, F1-score, and support for each class (Depressed, Euthymic, Manic) in the test dataset.

Mood State	Precision	Recall	F1-Score	Support
Depressed	0.25	0.18	0.21	74
Euthymic	0.25	0.13	0.18	97
Manic	0.70	0.85	0.77	329
Overall	**0.55**	**0.61**	**0.57**	**500**

From the breakdown, it is evident that the classifier performs best on the **Manic** class, achieving a strong F1-score of 0.77, with high recall (0.85) and good precision (0.70). This suggests that the model is highly effective at identifying manic episodes—likely due to stronger signals in features like elevated mood score, reduced sleep, or stress-related behaviours that are more distinct in manic patients.

In contrast, the **Depressed** and **Euthymic** classes exhibit significantly lower precision and recall, with F1-scores below 0.2. This underperformance may stem from feature overlap or subtle clinical boundaries between depressive and euthymic states, which are often harder to distinguish even in real-world psychiatric assessments. The support column further reveals that class imbalance may be a contributing factor—manic cases (n = 329) dominate the dataset, while depressed (n = 74) and euthymic (n = 97) are underrepresented.

These findings highlight both the potential and limitations of the current model, emphasizing the need for either improved sampling or incorporation of additional discriminative features to better classify less prevalent mood states.

3.3. Confusion Matrix and ROC Analysis

To further investigate the model's behaviour on a per-class basis, we analysed the confusion matrix and ROC (Receiver Operating Characteristic) curves, both of which provide intuitive insights into model performance beyond scalar metrics.

Confusion Matrix

The confusion matrix quantifies correct and incorrect predictions across the three mood states [51]-[53]. As shown in Figure 3, many manic cases are correctly identified (true positives), which aligns with the high recall observed in Section 3.2. However, there is a clear tendency for the model to misclassify depressed and euthymic patients as manic. This pattern indicates that features distinguishing manic states are more dominant in the learned decision boundaries, while the depressed and euthymic signals are weaker or overlapping. Misclassifications between depressed and euthymic states are also evident, underscoring the clinical reality that these mood states often share overlapping behavioural and biochemical features. This reinforces the challenge of differentiating them purely through non-invasive clinical metrics.

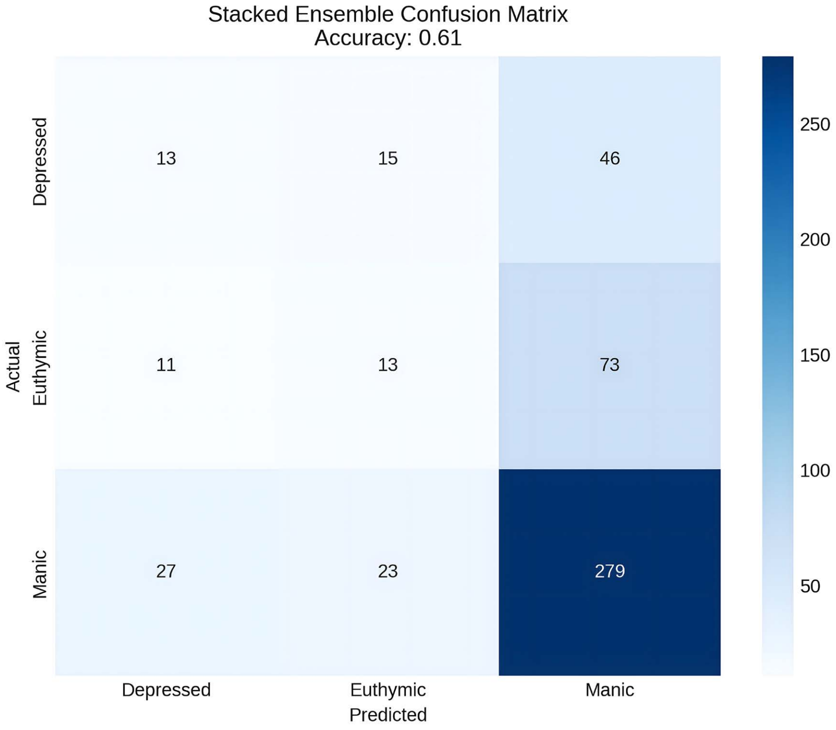

Figure 3. Confusion matrix illustrating classification performance across mood states. Strong diagonal for manic predictions; depressed and euthymic are frequently confused.

ROC Analysis

To evaluate the model's discriminative power, **one-vs-rest ROC curves** were generated for each mood state. These curves plot the trade-off between sensitivity (true positive rate) and specificity (1 – false positive rate).

As shown in Figure 4, the Area Under the Curve (AUC) values are as follows:

- **Manic**: 0.67
- **Depressed**: 0.68
- **Euthymic**: 0.59

These AUC scores confirm that the model is relatively better at distinguishing manic and depressed states, while euthymic classification remains more challenging.

The comparatively low AUC for euthymia reflects overlap with both depressed and manic profiles, likely due to its intermediate position in the mood continuum.

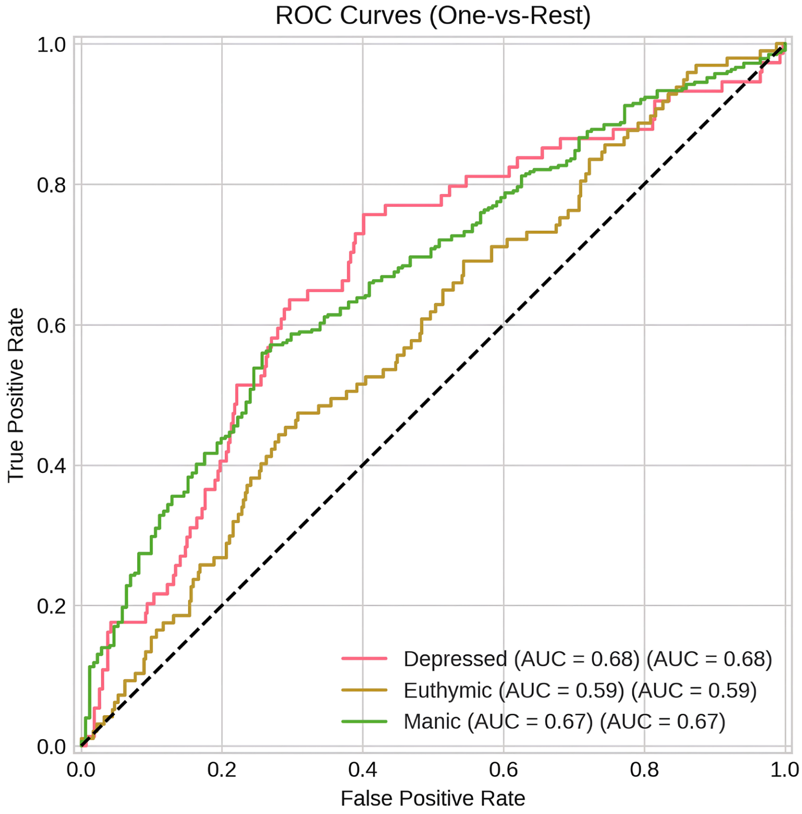

Figure 4. ROC curves for each class (one-vs-rest). Manic and depressed show moderate discriminative power; Euthymic remains difficult to separate.

3.4. Feature Importance

To interpret the decision logic of the ensemble model, we extracted feature importance values from the **Random Forest** component of the stacked classifier. Feature importance represents the contribution of each input variable to the model's predictions, measured by the average impurity reduction it provides across all decision trees.

As shown in Figure 5, the top 15 most influential features include:

- **Stress Level**: The most dominant feature, indicating its strong predictive relationship with mood state.
- **Sleep Deviation**: Suggests that deviations from optimal sleep duration (7 hours) are highly indicative of mood instability.
- **Lithium Adherence**: Reflects treatment compliance and correlates with mood stabilization, particularly for manic episodes.

Other notable contributors include therapy attendance, number of previous episodes, and family history, all of which are established psychiatric risk markers [54]-[56]. The visualization underscores the clinical plausibility of the model's behaviour and confirms the relevance of the engineered features.

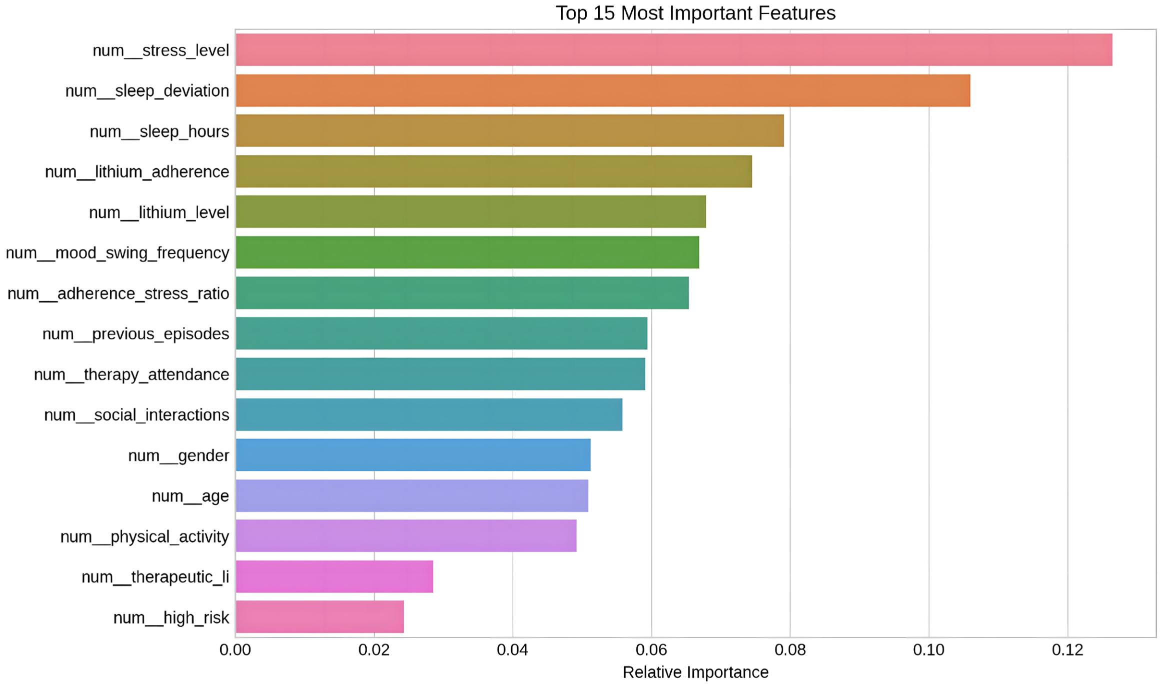

Figure 5. Top 15 most important features contributing to mood state classification. Stress, sleep deviation, and lithium adherence dominate.

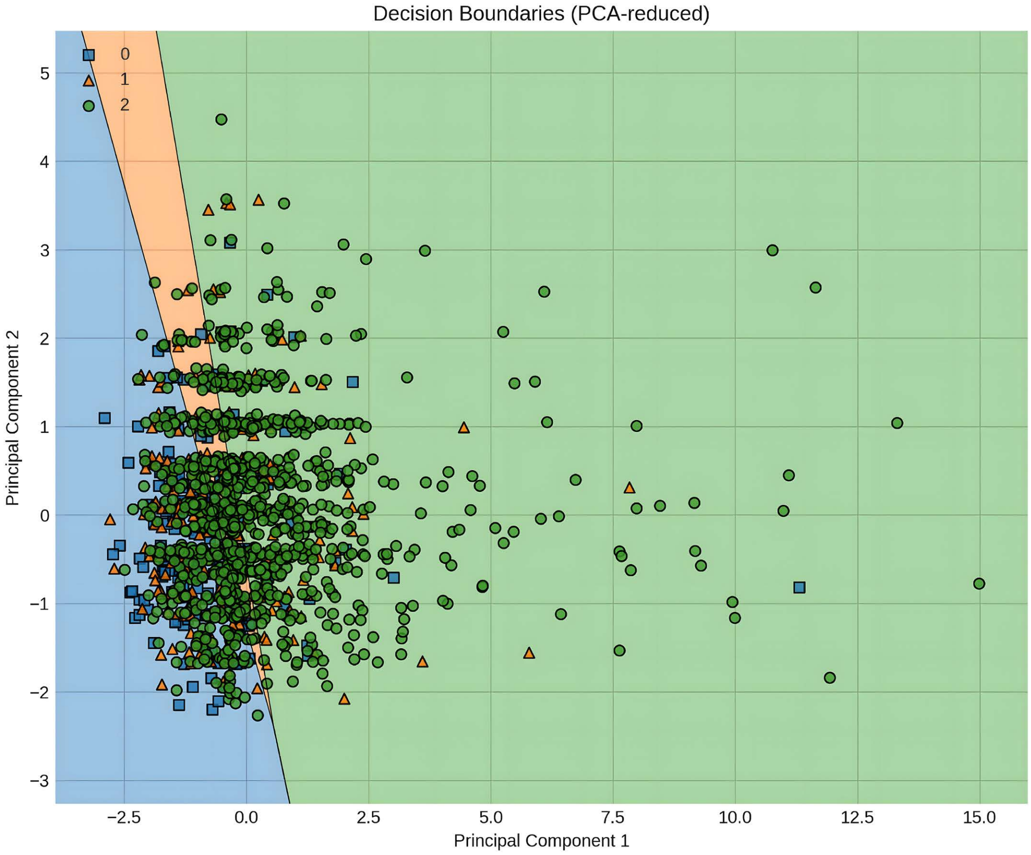

Figure 6. PCA-reduced 2D decision boundaries. Manic states are more separable, while depressed and euthymic states overlap.

3.5. Decision Boundaries

To visually assess how well the model separates mood classes in a reduced feature space, we applied Principal Component Analysis (PCA) to project the high-dimensional feature matrix into two dimensions [57]-[59]. A logistic regression classifier was then fitted to this PCA-reduced data to draw decision boundaries. As illustrated in **Figure 6**, there is some class overlap in the 2D space, particularly between depressed and euthymic samples. However, manic states form more distinct clusters, explaining why the model performs best in that category. The boundaries are non-linear and exhibit class-specific curvature, reflecting the complexity of the underlying relationships. This plot serves as a visual summary of the model's discriminative ability and supports prior conclusions regarding inter-class variability.

4. Discussion

This study presents a comprehensive and clinically grounded machine learning framework capable of classifying mood states in bipolar disorder using a combination of simulated patient data and domain-informed feature engineering. The performance of the stacked ensemble model, especially in detecting manic episodes, underscores the potential of computational methods to augment psychiatric decision-making.

One of the most significant outcomes is the model's high sensitivity and precision for **manic states**, which are often associated with risk-taking behaviour, impaired judgment, and hospitalization if left untreated. The model's robustness in identifying these cases suggests that behavioural and physiological indicators—such as decreased sleep, increased stress, and high adherence to lithium—offer sufficiently strong signals to support automated diagnosis. This reinforces the value of longitudinal, objective clinical data over self-reported mood scales, which may be limited by cognitive distortion or recall bias. However, the underperformance on euthymic and depressed states—as reflected by low F1-scores and overlapping decision boundaries—points to a critical limitation. The subtle transition between depression and euthymia is difficult to capture through commonly measured clinical variables. In real-world psychiatric contexts, these states often require nuanced clinical judgment based on tone of speech, affect, psychomotor activity, and patient history [60]-[62]. Future versions of this model could benefit from the integration of multimodal data such as voice analysis, wearable sleep metrics, and electronic health records (EHR). Another strength of this pipeline lies in its explainability. By prioritizing interpretable features like stress level, adherence ratios, and therapeutic lithium ranges, the model facilitates clinical trust and transparency—qualities often absent in black-box neural networks. The feature importance rankings and ROC analyses provide tangible evidence for how the model arrives at predictions, making it more suitable for integration into clinical workflows or electronic decision support tools.

The synthetic data strategy employed here also has methodological implications. By generating statistically controlled yet realistic patient profiles, we create a reproducible and ethical environment for testing diagnostic algorithms in the absence of public bipolar disorder datasets. This can be a stepping stone for validating the model on real-world cohorts from psychiatric hospitals or longitudinal mental health studies. This pipeline demonstrates that ensemble machine learning models, when combined with domain-driven engineering and explainability, can serve as powerful tools for mood state classification in bipolar disorder. The system's strengths in identifying manic states and providing interpretable outputs highlight its clinical relevance. However, performance gaps for euthymic and depressed states call for deeper data integration, emphasizing the need for future work to blend traditional psychiatric wisdom with modern AI capabilities.

5. Conclusion

In this study, we developed and validated a clinically interpretable machine learning framework for classifying mood states in bipolar disorder using structured synthetic data and domain-specific feature engineering. By integrating behavioural, biochemical, and therapy-related indicators, our ensemble-based model was able to capture clinically meaningful distinctions, particularly excelling in the identification of manic episodes. The stacked ensemble classifier, combining Random Forest and XGBoost with a logistic regression meta-learner, demonstrated promising performance with an overall accuracy of 61% and an F1-score of 0.77 for manic state prediction. The inclusion of explainable components such as feature importance rankings, decision boundaries, and ROC curves makes the model highly transparent—an essential quality for real-world deployment in psychiatric care. However, the reduced predictive power for euthymic and depressed states highlights the complexity of mood differentiation in bipolar disorder and suggests the need for richer data modalities in future iterations. Integrating multimodal sources such as wearable sensor data, longitudinal patient monitoring, and structured clinical notes could significantly enhance diagnostic resolution [63]-[66]. Importantly, the pipeline's modular and reproducible design—paired with its ethical use of simulated patient data—makes it a valuable prototype for clinical AI research. This work lays the groundwork for real-world validation, model refinement, and eventual integration into digital mental health platforms or clinical decision support systems. In conclusion, this research underscores the potential of interpretable, ensemble-based machine learning to support mood state classification in psychiatry. By prioritizing explainability, clinical alignment, and adaptability, our framework takes a critical step toward actionable, AI-assisted mental health diagnostics.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Miklowitz, D.J. and Johnson, S.L. (2017) Bipolar Disorder. In: Craighead, W.E.,

Miklowitz, D.J. and Craighead, L.W., Eds., *Psychopathology: History, Diagnosis, and Empirical Foundations, Third Edition*, Wiley, 317-363. https://doi.org/10.1002/9781394258949.ch8

[2] Vieta, E., Berk, M., Schulze, T.G., Carvalho, A.F., Suppes, T., Calabrese, J.R., *et al.* (2018) Bipolar Disorders. *Nature Reviews Disease Primers*, **4**, Article No. 18008. https://doi.org/10.1038/nrdp.2018.8

[3] Akiskal, H.S., Bourgeois, M.L., Angst, J., Post, R., Möller, H. and Hirschfeld, R. (2000) Re-Evaluating the Prevalence of and Diagnostic Composition within the Broad Clinical Spectrum of Bipolar Disorders. *Journal of Affective Disorders*, **59**, S5-S30. https://doi.org/10.1016/s0165-0327(00)00203-2

[4] Kim, Y., Santos, R., Gage, F.H. and Marchetto, M.C. (2017) Molecular Mechanisms of Bipolar Disorder: Progress Made and Future Challenges. *Frontiers in Cellular Neuroscience*, **11**, Article 30. https://doi.org/10.3389/fncel.2017.00030

[5] Cheng, K.K.F. and Lee, D.T.F. (2011) Effects of Pain, Fatigue, Insomnia, and Mood Disturbance on Functional Status and Quality of Life of Elderly Patients with Cancer. *Critical Reviews in Oncology/Hematology*, **78**, 127-137. https://doi.org/10.1016/j.critrevonc.2010.03.002

[6] Iosifescu, D.V. (2012) The Relation between Mood, Cognition and Psychosocial Functioning in Psychiatric Disorders. *European Neuropsychopharmacology*, **22**, S499-S504. https://doi.org/10.1016/j.euroneuro.2012.08.002

[7] Mikkelsen, M.E., Shull, W.H., Biester, R.C., Taichman, D.B., Lynch, S., Demissie, E., *et al.* (2008) Cognitive, Mood and Quality of Life Impairments in a Select Population of ARDS Survivors. *Respirology*, **14**, 76-82. https://doi.org/10.1111/j.1440-1843.2008.01419.x

[8] Molassiotis, A., Callaghan, P., Twinn, S.F., Lam, S.W., Chung, W.Y. and Li, C.K. (2002) A Pilot Study of the Effects of Cognitive-Behavioral Group Therapy and Peer Support/Counseling in Decreasing Psychologic Distress and Improving Quality of Life in Chinese Patients with Symptomatic HIV Disease. *AIDS Patient Care and STDs*, **16**, 83-96. https://doi.org/10.1089/10872910252806135

[9] Maj, M., Stein, D.J., Parker, G., Zimmerman, M., Fava, G.A., De Hert, M., *et al.* (2020) The Clinical Characterization of the Adult Patient with Depression Aimed at Personalization of Management. *World Psychiatry*, **19**, 269-293. https://doi.org/10.1002/wps.20771

[10] Ramasubbu, R., Beaulieu, S., Taylor, V.H., Schaffer, A. and McIntyre, R.S. (2012) The CANMAT Task Force Recommendations for the Management of Patients with Mood Disorders and Comorbid Medical Conditions: Diagnostic, Assessment, and Treatment Principles. *Annals of Clinical Psychiatry*, **24**, 82-90.

[11] Velligan, D.I., Weiden, P.J., Sajatovic, M., Scott, J., Carpenter, D., Ross, R., *et al.* (2010) Strategies for Addressing Adherence Problems in Patients with Serious and Persistent Mental Illness: Recommendations from the Expert Consensus Guidelines. *Journal of Psychiatric Practice*, **16**, 306-324. https://doi.org/10.1097/01.pra.0000388626.98662.a0

[12] Zentner, M. and Eerola, T. (2010) Self-Report Measures and Models. In: Juslin, P.N., Ed., *Handbook of Music and Emotion: Theory, Research, Applications*, Oxford University Press, 187-221. https://doi.org/10.1093/acprof:oso/9780199230143.003.0008

[13] Hopwood, C.J. and Bornstein, R.F. (2014) Multimethod Clinical Assessment. Guilford Publications.

[14] Lewis, A.J., Bertino, M.D., Bailey, C.M., Skewes, J., Lubman, D.I. and Toumbourou, J.W. (2014) Depression and Suicidal Behavior in Adolescents: A Multi-Informant and

Multi-Methods Approach to Diagnostic Classification. *Frontiers in Psychology*, **5**, Article 766. https://doi.org/10.3389/fpsyg.2014.00766

[15] Bzdok, D. and Meyer-Lindenberg, A. (2018) Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, **3**, 223-230. https://doi.org/10.1016/j.bpsc.2017.11.007

[16] 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

[17] Omiyefa, S. (2025) Artificial Intelligence and Machine Learning in Precision Mental Health Diagnostics and Predictive Treatment Models. *International Journal of Research Publication and Reviews*, **6**, 85-99. https://doi.org/10.55248/gengpi.6.0325.1107

[18] Tsapekos, D., Strawbridge, R., Cella, M., Goldsmith, K., Kalfas, M., Taylor, R.H., *et al.* (2023) Cognitive Remediation in Bipolar (CRiB2): Study Protocol for a Randomised Controlled Trial Assessing Efficacy and Mechanisms of Cognitive Remediation Therapy Compared to Treatment as Usual. *BMC Psychiatry*, **23**, Article No. 842. https://doi.org/10.1186/s12888-023-05327-1

[19] Owen, D.J. (2021) Individual Differences and Medication-Mediation in Chronic Illness Conditions: A Mixed Methods Approach to the Development of a Novel, Conceptual Framework. Master's Thesis, University of Derby (United Kingdom).

[20] McFadden, S. (2015) An Exploration of the Psychological Indicators of Aspirin Adherence, in Patients with Stable Coronary Artery Disease, Using a Direct Assay Measurement. Ph.D. Thesis, Royal College of Surgeons in Ireland.

[21] Pilarska, A., Pieczyńska, A. and Hojan, K. (2023) Neuropsychological Monitoring of Cognitive Function and ICF-Based Mental Components in Patients with Malignant Brain Tumours. *Frontiers in Psychology*, **14**, Article 1033185. https://doi.org/10.3389/fpsyg.2023.1033185

[22] Ostojic, D., Lalousis, P.A., Donohoe, G. and Morris, D.W. (2024) The Challenges of Using Machine Learning Models in Psychiatric Research and Clinical Practice. *European Neuropsychopharmacology*, **88**, 53-65. https://doi.org/10.1016/j.euroneuro.2024.08.005

[23] Ku, W.L. and Min, H. (2024) Evaluating Machine Learning Stability in Predicting Depression and Anxiety Amidst Subjective Response Errors. *Healthcare*, **12**, Article 625. https://doi.org/10.3390/healthcare12060625

[24] Corponi, F., Li, B.M., Anmella, G., Mas, A., Pacchiarotti, I., Valentí, M., *et al.* (2024) Automated Mood Disorder Symptoms Monitoring from Multivariate Time-Series Sensory Data: Getting the Full Picture Beyond a Single Number. *Translational Psychiatry*, **14**, Article No. 161. https://doi.org/10.1038/s41398-024-02876-1

[25] Mokheleli, T.D. (2023) A Comparison of Machine Learning Techniques for Predicting Mental Health Disorders. Master's Thesis, University of Johannesburg (South Africa).

[26] El-Mallakh, R.S. (1997) New Insights into the Course and Prognosis of Bipolar Illness. *Psychiatric Annals*, **27**, 478-481. https://doi.org/10.3928/0048-5713-19970701-08

[27] Kumar, S., Akhtar, Z., Satsangi, H., Sehrawat, S., Arora, N. and Bamal, K. (2024) Depression Prediction Using Machine Learning Techniques. In: Bathla, G., Kumar, S., Garg, H. and Saini, D., Eds., *Artificial Intelligence in Healthcare*, CRC Press, 241-265. https://doi.org/10.1201/9781003522096-12

[28] Grandjean, E.M. and Aubry, J. (2009) Lithium: Updated Human Knowledge Using

an Evidence-Based Approach: Part II: Clinical Pharmacology and Therapeutic Monitoring. *CNS Drugs*, **23**, 331-349. https://doi.org/10.2165/00023210-200923040-00005

[29] Ahmadi, F. (2022) Lithium Toxicity Following Co-Prescription of Lithium and ACEI/ARBs: A Population-Based Cohort Study. Master's Thesis, The University of Western Ontario.

[30] Sienaert, P., Geeraerts, I. and Wyckaert, S. (2013) How to Initiate Lithium Therapy: A Systematic Review of Dose Estimation and Level Prediction Methods. *Journal of Affective Disorders*, **146**, 15-33. https://doi.org/10.1016/j.jad.2012.08.013

[31] Castiglione-Fontanellaz, C.E.G., Schaufler, S., Wild, S., Hamann, C., Kaess, M. and Tarokh, L. (2023) Sleep Regularity in Healthy Adolescents: Associations with Sleep Duration, Sleep Quality, and Mental Health. *Journal of Sleep Research*, **32**, e13865. https://doi.org/10.1111/jsr.13865

[32] Wang, F., Sun, Z., Lin, F., Xu, Y., Wu, E., Sun, X., *et al.* (2024) Nonlinear Relationships between Sleep Duration, Mental Health, and Quality of Life: The Dangers of Less Sleep versus More Sleep. *Sleep Medicine*, **119**, 565-573. https://doi.org/10.1016/j.sleep.2024.05.043

[33] Baglioni, C., Nanovska, S., Regen, W., Spiegelhalder, K., Feige, B., Nissen, C., *et al.* (2016) Sleep and Mental Disorders: A Meta-Analysis of Polysomnographic Research. *Psychological Bulletin*, **142**, 969-990. https://doi.org/10.1037/bul0000053

[34] Sanjobo, N. (2007) Adherence to Antiretroviral Treatment in Zambia: A Qualitative Study of Patients and Health Professionals' Views. Master's Thesis, University of Oslo.

[35] VanWormer, A.M. (2010) Adherence in an Optimized Program of Mindfulness Based Stress Reduction. Ph.D. Thesis, University of Minnesota.

[36] Baillie, C.A., VanZandbergen, C., Tait, G., Hanish, A., Leas, B., French, B., *et al.* (2013) The Readmission Risk Flag: Using the Electronic Health Record to Automatically Identify Patients at Risk for 30-Day Readmission. *Journal of Hospital Medicine*, **8**, 689-695. https://doi.org/10.1002/jhm.2106

[37] Boydell, K.M., Volpe, T., Gladstone, B.M., Stasiulis, E. and Addington, J. (2012) Youth at Ultra High Risk for Psychosis: Using the Revised Network Episode Model to Examine Pathways to Mental Health Care. *Early Intervention in Psychiatry*, **7**, 170-186. https://doi.org/10.1111/j.1751-7893.2012.00350.x

[38] McEwan, T.E., Shea, D.E., Daffern, M., MacKenzie, R.D., Ogloff, J.R.P. and Mullen, P.E. (2016) The Reliability and Predictive Validity of the Stalking Risk Profile. *Assessment*, **25**, 259-276. https://doi.org/10.1177/1073191116653470

[39] Lee, K.J. and Carlin, J.B. (2016) Multiple Imputation in the Presence of Non-Normal Data. *Statistics in Medicine*, **36**, 606-617. https://doi.org/10.1002/sim.7173

[40] Liao, S.G., Lin, Y., Kang, D.D., Chandra, D., Bon, J., Kaminski, N., *et al.* (2014) Missing Value Imputation in High-Dimensional Phenomic Data: Imputable or Not, and How? *BMC Bioinformatics*, **15**, Article No. 346. https://doi.org/10.1186/s12859-014-0346-6

[41] Shadbahr, T., Roberts, M., Stanczuk, J., Gilbey, J., Teare, P., Dittmer, S., *et al.* (2023) The Impact of Imputation Quality on Machine Learning Classifiers for Datasets with Missing Values. *Communications Medicine*, **3**, Article No. 139. https://doi.org/10.1038/s43856-023-00356-z

[42] Mukherjee, M. and Khushi, M. (2021) SMOTE-ENC: A Novel Smote-Based Method to Generate Synthetic Data for Nominal and Continuous Features. *Applied System*

Innovation, **4**, Article 18. https://doi.org/10.3390/asi4010018

[43] Dablain, D., Krawczyk, B. and Chawla, N.V. (2023) DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data. *IEEE Transactions on Neural Networks and Learning Systems*, **34**, 6390-6404. https://doi.org/10.1109/tnnls.2021.3136503

[44] Fernandez, A., Garcia, S., Herrera, F. and Chawla, N.V. (2018) SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-Year Anniversary. *Journal of Artificial Intelligence Research*, **61**, 863-905. https://doi.org/10.1613/jair.1.11192

[45] Douzas, G., Bacao, F. and Last, F. (2018) Improving Imbalanced Learning through a Heuristic Oversampling Method Based on K-Means and Smote. *Information Sciences*, **465**, 1-20. https://doi.org/10.1016/j.ins.2018.06.056

[46] Pujatti, M., Calzà, D., Gobbi, A., Svaizer, P. and Cristoforetti, M. (2025) Mopidip: A Modular Real-Time Pipeline for Machinery Diagnosis and Prognosis Based on Deep Learning Algorithms. *Journal of Big Data*, **12**, Article No. 82. https://doi.org/10.1186/s40537-025-01123-8

[47] Zubair, M., Hussai, M., Al-Bashrawi, M.A., Ben-dechache, M. and Owais, M. (2025) A Comprehensive Review of Techniques, Algorithms, Advancements, Challenges, and Clinical Applications of Multi-Modal Medical Image Fusion for Improved Diagnosis. arXiv: 2505.14715.

[48] Tripathi, A., Waqas, A., Venkatesan, K., Yilmaz, Y. and Rasool, G. (2024) Building Flexible, Scalable, and Machine Learning-Ready Multimodal Oncology Datasets. *Sensors*, **24**, Article 1634. https://doi.org/10.3390/s24051634

[49] Murphy, D.L. (1974) Depression, Elation, and Lithium Carbonate Responses in Manic Patient Subgroups. *Archives of General Psychiatry*, **31**, 643-648. https://doi.org/10.1001/archpsyc.1974.01760170043007

[50] Rybakowski, J.K. (2014) Response to Lithium in Bipolar Disorder: Clinical and Genetic Findings. *ACS Chemical Neuroscience*, **5**, 413-421. https://doi.org/10.1021/cn5000277

[51] Picard, R.W., Vyzas, E. and Healey, J. (2001) Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **23**, 1175-1191. https://doi.org/10.1109/34.954607

[52] Sükei, E., Norbury, A., Perez-Rodriguez, M.M., Olmos, P.M. and Artés, A. (2021) Predicting Emotional States Using Behavioral Markers Derived from Passively Sensed Data: Data-Driven Machine Learning Approach. *JMIR mHealth and uHealth*, **9**, e24465. https://doi.org/10.2196/24465

[53] Ramirez Rios, F.E. and María Reyes Duke, A. (2023) Building of a Convolutional Neuronal Network for the Prediction of Mood States through Face Recognition Based on Object Detection with YOLOV8 and Python. 2023 *IEEE International Conference on Machine Learning and Applied Network Technologies* (*ICMLANT*), San Salvador, 14-15 December 2023, 1-6. https://doi.org/10.1109/icmlant59547.2023.10372862

[54] McGorry, P., Keshavan, M., Goldstone, S., Amminger, P., Allott, K., Berk, M., *et al.* (2014) Biomarkers and Clinical Staging in Psychiatry. *World Psychiatry*, **13**, 211-223. https://doi.org/10.1002/wps.20144

[55] Roberts, R.E., Roberts, C.R. and Chan, W. (2009) One-Year Incidence of Psychiatric Disorders and Associated Risk Factors among Adolescents in the Community. *Journal of Child Psychology and Psychiatry*, **50**, 405-415. https://doi.org/10.1111/j.1469-7610.2008.01969.x

[56] Gelenberg, A.J., Freeman, M.P., Markowitz, J.C., Rosenbaum, J.F., Thase, M.E.,

Trivedi, M.H. and Van Rhoads, R.S. (2010) American Psychiatric Association Practice Guidelines for the Treatment of Patients with Major Depressive Disorder. *American Journal of Psychiatry*, **167**, 9-118.

[57] Huang, D., Jiang, F., Li, K., Tong, G. and Zhou, G. (2022) Scaled PCA: A New Approach to Dimension Reduction. *Management Science*, **68**, 1678-1695. https://doi.org/10.1287/mnsc.2021.4020

[58] Nolazco-Flores, J.A., Faundez-Zanuy, M., Velázquez-Flores, O.A., Del-Valle-Soto, C., Cordasco, G. and Esposito, A. (2022) Mood State Detection in Handwritten Tasks Using PCA-mFCBF and Automated Machine Learning. *Sensors*, **22**, Article 1686. https://doi.org/10.3390/s22041686

[59] Nanga, S., Bawah, A.T., Acquaye, B.A., Billa, M., Baeta, F.D., Odai, N.A., *et al.* (2021) Review of Dimension Reduction Methods. *Journal of Data Analysis and Information Processing*, **9**, 189-231. https://doi.org/10.4236/jdaip.2021.93013

[60] Nelson, M.K. and Lawrence, M.P. (2025) Psychiatric Mental Health Nursing Case Studies, Volume 1: 100 Real-Life Scenarios for Clinical Practice in Mood, Anxiety, Psychosis, and Personality Disorders with Nursing Assessments, DSM-5 Diagnoses & Care Plans. Independently published.

[61] Dikaios, K., Rempel, S., Dumpala, S.H., Oore, S., Kiefte, M. and Uher, R. (2023) Applications of Speech Analysis in Psychiatry. *Harvard Review of Psychiatry*, **31**, 1-13. https://doi.org/10.1097/hrp.0000000000000356

[62] Goldberg, J.F. and Stahl, S.M. (2024) Clinical Reasoning and Decision-Making in Psychiatry. Cambridge University Press. https://doi.org/10.1017/9781009181549

[63] Mahato, K., Saha, T., Ding, S., Sandhu, S.S., Chang, A. and Wang, J. (2024) Hybrid Multimodal Wearable Sensors for Comprehensive Health Monitoring. *Nature Electronics*, **7**, 735-750. https://doi.org/10.1038/s41928-024-01247-4

[64] Singh, K., Piyush, P., Kumar, R., Chhabra, S., Goomer, N. and Kashyap, A. (2024) Multimodal Data Extraction & Fusion for Health Monitoring System and Early Diagnosis. 2024 *International Conference on Computational Intelligence and Computing Applications* (*ICCICA*), Samalkha, 23-24 May 2024, 216-220. https://doi.org/10.1109/iccica60014.2024.10585027

[65] Kline, A., Wang, H., Li, Y., Dennis, S., Hutch, M., Xu, Z., *et al.* (2022) Multimodal Machine Learning in Precision Health: A Scoping Review. *NPJ Digital Medicine*, **5**, Article No. 171. https://doi.org/10.1038/s41746-022-00712-8

[66] Chaabene, S., Boudaya, A., Bouaziz, B. and Chaari, L. (2025) An Overview of Methods and Techniques in Multimodal Data Fusion with Application to Healthcare. *International Journal of Data Science and Analytics*. https://doi.org/10.1007/s41060-025-00715-0