BIPOLAR DISORDER DIAGNOSIS

Lorenzo Maria Mazzotti

ABSTRACT

Abstract

This study aims to explore the predictive capabilities of various machine learning models for the diagnosis of bipolar disorder using a structured dataset containing mental health variables. Regression and multiple classification techniques were employed, including logistic regression, classification trees, random forests, XGBoost, and k-nearest neighbors (KNN). Advanced techniques such as PCA and t-SNE were used for dimensionality reduction, followed by agglomerative and K-Means clustering to investigate potential intrinsic groupings within the data.

Contents

1 Data

1.1 Dataset Source

Dataset Source: Karbalaeipour, Hengameh; Damari, Siavash; Zolfagharnasab, Mohammad Hossein; Haghdadi, Amin, 2023, "A Collection of 120 Psychology Patients with 17 Essential Symptoms to Diagnose Mania Bipolar Disorder, Depressive Bipolar Disorder, Major Depressive Disorder, and Normal Individuals", `https://doi.org/10.7910/DVN/0FNET5`, Harvard Dataverse, V1

1.2 Dataset Description

The dataset, assembled by Karbalaeipour et al. (2023), comprises 120 patients divided equally among four categories: Normal, Mania Bipolar Disorder (Bipolar type 1), Depressive Bipolar Disorder (Bipolar type 2), and Major Depressive Disorder. It features 17 essential symptoms used by psychiatrists to diagnose these disorders, recorded in CSV format.

1.3 Variables Description

Below is Table 1 describing each variable:

Table 1: Variable Description

Variable	**Description**
Patient Number	A unique identifier for each patient.
Sadness	Categorical variable indicating the frequency of feeling sad, with levels like "Usually", "Seldom", "Sometimes", and "Most-Often".
Euphoric	Categorical variable describing the frequency of euphoric feelings, similarly categorized.
Exhausted	Measures how often the patient feels exhausted.
Sleep Disorder	Indicates the frequency of sleep disorders.
Mood Swing	Binary variable ("YES" or "NO") indicating the presence of mood swings.
Suicidal Thoughts	Binary variable indicating whether the patient has suicidal thoughts.
Anorexia	Binary variable showing whether the patient experiences anorexia (likely indicating loss of appetite rather than the eating disorder).
Authority Respect	Binary variable capturing the patient's respect for authority.
Try-Explanation	Indicates whether the patient often tries to explain their actions or feelings.
Aggressive Response	Binary variable indicating if the patient often responds aggressively.

Table 1 – *Variable Description*

Variable	**Description**
Ignore & Move-On	Binary variable describing whether the patient tends to ignore problems and move on.
Nervous Break-down	Binary variable indicating if the patient has had a nervous breakdown.
Admit Mistakes	Binary variable showing whether the patient admits to making mistakes.
Overthinking	Indicates whether the patient has a tendency to overthink.
Sexual Activity	Numeric variable (originally formatted as "X From 10") measuring the level of sexual activity.
Concentration	Similarly formatted numeric variable assessing the level of concentration.
Optimism	Numeric variable (formatted as "X From 10") assessing the level of optimism.
Expert Diagnose	The expert's diagnosis for the patient, with categories such as "Bipolar Type-1", "Bipolar Type-2", "Depression", and "Normal".

2 Data Preparation

This section outlines the preprocessing steps applied to the dataset to ensure compatibility with machine learning algorithms, improve model interpretability, and enhance overall effectiveness.

2.1 Numeric Conversion of Ordinal and Ratio Data

A dedicated function was implemented to extract numeric values by stripping textual annotations from fields that originally combined numeric scores with descriptive text. For instance, fields such as "Sexual Activity", "Concentration", and "Optimism" were converted into integer values. This conversion is critical for enabling quantitative analysis.

2.2 Binary Encoding

Binary outcomes, which were initially recorded as "YES" or "NO", were transformed into numeric values (1 for "YES" and 0 for "NO"). This transformation applies to variables such as mood swings, suicidal thoughts, and other binary diagnostic criteria, thereby simplifying their integration into various machine learning models.

2.3 Mapping Ordinal Variables

Ordinal variables, including "Sadness" and "Euphoric", were encoded as integers to preserve their inherent ordering. For example, categories such as "Seldom" and "Sometimes" were mapped to a progressive scale that quantifies the frequency or severity of the symptoms, making these variables more amenable to analysis.

2.4 Target Variable Preparation

A new binary variable, "Bipolar Diagnose", was created to identify cases specifically of Bipolar Type-1 and Type-2, based on the expert diagnoses provided. This focused target variable aids in differentiating bipolar disorder from other conditions in the dataset.

2.5 Dropping and Normalizing Columns

In the interest of standardizing the data, selected features were normalized to a range between 0 and 1. This step standardizes the scale across variables, mitigating bias induced by disparate scales. Additionally, irrelevant columns, such as "Patient Number", were removed to prevent non-predictive content from influencing analytical outcomes. A specific subset of the dataset was also retained for clustering analyses, ensuring that variables like "Expert Diagnose" and "Mood Swing" were excluded from certain analyses to avoid data leakage.

2.6 Final Dataset Adjustments for Machine Learning

For the purpose of supervised learning, columns that could lead to data leakage or that were perfect predictors of the outcome were excluded. This final adjustment ensures that the model training is based solely on features that truly represent the underlying patterns in the data.

3 Descriptive Statistics

In the context of the present study, the descriptive goal is to elucidate the relationships among various psychological symptoms and their correlation with the diagnosis of bipolar disorder. To achieve this, several analytical techniques were employed.

3.1 Correlation Analysis

A correlation heatmap was generated to visually assess the strength and direction of relationships between all variables. This method provides a comprehensive overview, highlighting how different symptoms relate to each other and their collective influence on bipolar disorder diagnosis.

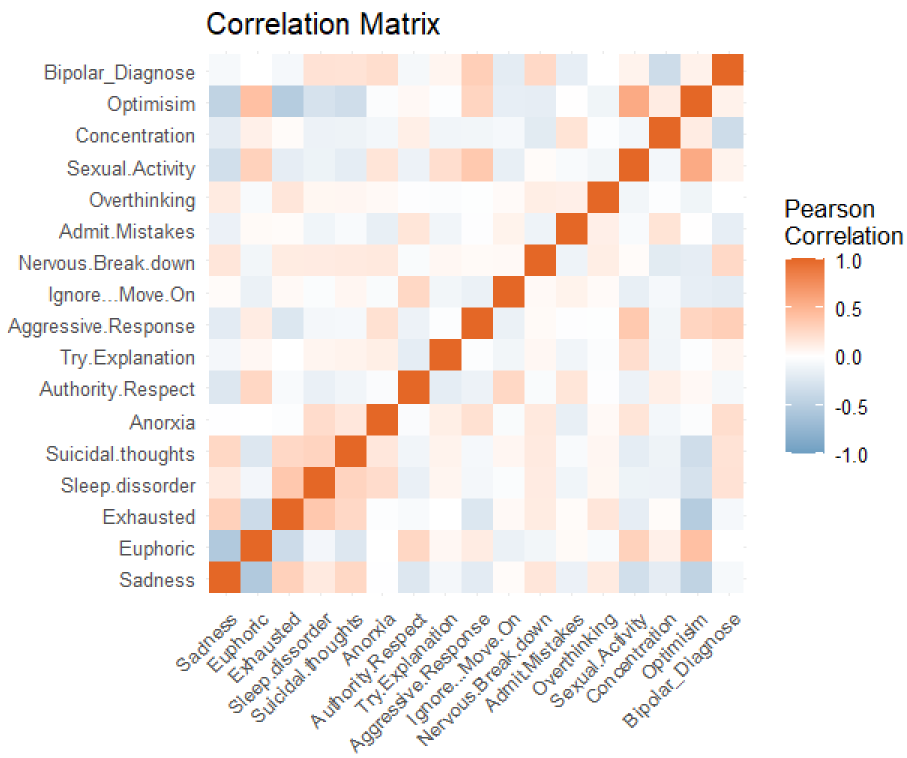

Figure 1: Correlation Heatmap

Positive correlations were observed between bipolar diagnosis and symptoms such as Aggressive Response, Nervous Breakdown, Anorexia, Suicidal Thoughts, and Sleep Disorder. These associations suggest that these symptoms are more prevalent or intense in individuals diagnosed with bipolar disorder. Negative correlations were identified for Concentration, Admit Mistakes, and Ignore and Move On with the bipolar diagnosis, indicating that these features are less pronounced in diagnosed individuals.

3.2 Correlation Network Graph

To further explore the interdependencies among variables, a correlation network graph was constructed, allowing for a dynamic visualization of how each symptom interrelates within the broader context of mental health assessments.

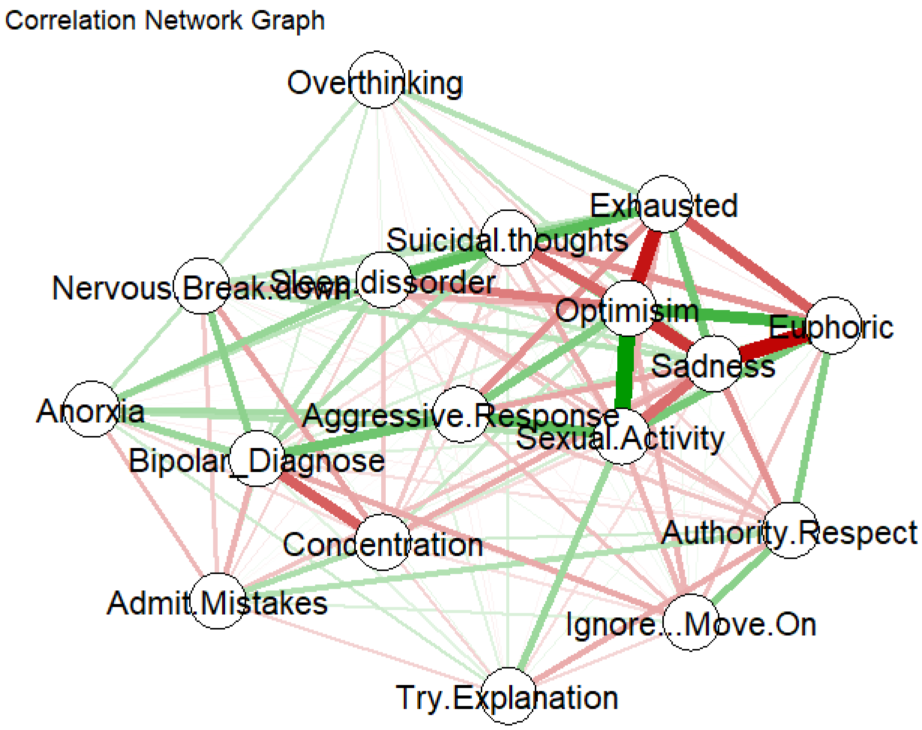

Figure 2: Correlation Network Graph

The correlation network graph corroborates the findings observed in the correlation heatmap, offering a visual representation that enhances our understanding of the relationships among psychological symptoms and their link to the diagnosis of bipolar disorder. This graphically intuitive presentation allows for the identification of not only confirmed correlations but also additional interactions between variables.

4 Supervised Learning Analysis

4.1 Introduction

In the supervised learning phase of the study, the aim was to ascertain the diagnostic accuracy for bipolar disorder using various predictive modeling techniques. These models were selected based on their distinct approaches to classification, providing a robust framework for the comparison and validation of results.

4.2 Model Training and Validation

Initially, the dataset was divided into training and testing sets, maintaining 80% of the data for training to ensure adequate learning while reserving 20% for testing to evaluate model performance. This split was randomized to prevent any bias.

4.3 Models Tested

The following classification models were tested:

Logistic Regression: Constructed to estimate the probabilities of bipolar disorder, evaluated using a confusion matrix with accuracy, precision, recall, and F1-score.

K-Nearest Neighbors (KNN): The optimal number of neighbors was determined through cross-validation to establish an effective benchmark.

Decision Tree: Developed to visualize decision rules and understand the hierarchical importance of symptoms.

Random Forest: Applied with extensive hyperparameter tuning to enhance prediction accuracy.

XGBoost: Tuned with cross-validation to prevent overfitting, optimizing hyperparameters such as learning rate and tree depth.

Each model's output was evaluated against the test set, and performance was documented to establish the most effective approach for diagnosing bipolar disorder.

4.4 Logistic Regression

Logistic regression was used as the foundational model for the classification task. Given its suitability for binary classification, it provided a strong baseline for comparison with other methods.

Table 2: GLM Coefficients for `Bipolar_Diagnose`

	Estimate	Std. Error	z value	Pr(> $\|z\|$)
(Intercept)	1.4812	1.7898	0.828	0.40791
Sadness	-3.2330	1.4479	-2.233	0.02556*
Euphoric	-1.8947	1.6069	-1.179	0.23836
Exhausted	-0.4392	1.0722	-0.410	0.68209
Sleep.dissorder	1.4459	1.0796	1.339	0.18045
Suicidal.thoughts	2.2453	0.8172	2.748	0.00601**
Anorxia	0.6556	0.6617	0.991	0.32181
Authority.Respect	0.1587	0.7909	0.201	0.84094
Try.Explanation	0.7686	0.6654	1.155	0.24807
Aggressive.Response	1.5941	0.6776	2.353	0.01864*
Ignore...Move.On	-1.9728	0.6563	-3.006	0.00265**
Nervous.Break.down	1.3633	0.6585	2.070	0.03843*
Admit.Mistakes	-0.9282	0.6444	-1.440	0.14974
Overthinking	0.4333	0.6464	0.670	0.50270
Sexual.Activity	-3.2493	1.6663	-1.950	0.05118·
Concentration	-4.2555	1.4778	-2.880	0.00398**
Optimisim	4.4607	1.9932	2.238	0.02523*

4.4.1 Key Findings from the GLM Regression

The regression results in Table 2 indicate several notable relationships. Higher levels of sadness are significantly associated with lower odds of a bipolar diagnosis ($\beta = -3.2330$, $p = 0.0256$), while increased suicidal thoughts correspond to a significantly higher likelihood ($\beta = 2.2453$, $p = 0.0060$). Similarly, more frequent aggressive responses and nervous breakdowns are linked with greater odds of diagnosis, with coefficients of 1.5941 ($p = 0.0186$) and 1.3633 ($p = 0.0384$), respectively. In contrast, the tendency to "ignore and move on" significantly reduces the likelihood of diagnosis ($\beta = -1.9728$, $p = 0.0027$). Furthermore, difficulties with concentration are associated with lower odds ($\beta = -4.2555$, $p = 0.0040$), whereas higher levels of optimism appear unexpectedly related to increased odds ($\beta = 4.4607$, $p = 0.0252$). These findings emphasize the complex interplay between emotional states, behavioral responses, and cognitive factors in bipolar diagnosis.

Table 3: Logistic Regression Performance Metrics

Model	**Balanced Accuracy**	**Precision**	**Recall**	**F1**
Logistic Regression	0.638	0.846	0.611	0.709

4.5 K-Nearest Neighbors (KNN)

Following logistic regression, the KNN method was explored to capture nonlinear relationships in the dataset.

Table 4: K-Nearest Neighbors Performance Metrics

Model	Balanced Accuracy	Precision	Recall	F1
KNN	0.722	0.916	0.611	0.733

4.6 Decision Tree

The Decision Tree model was employed to enhance interpretability and provide clear decision rules.

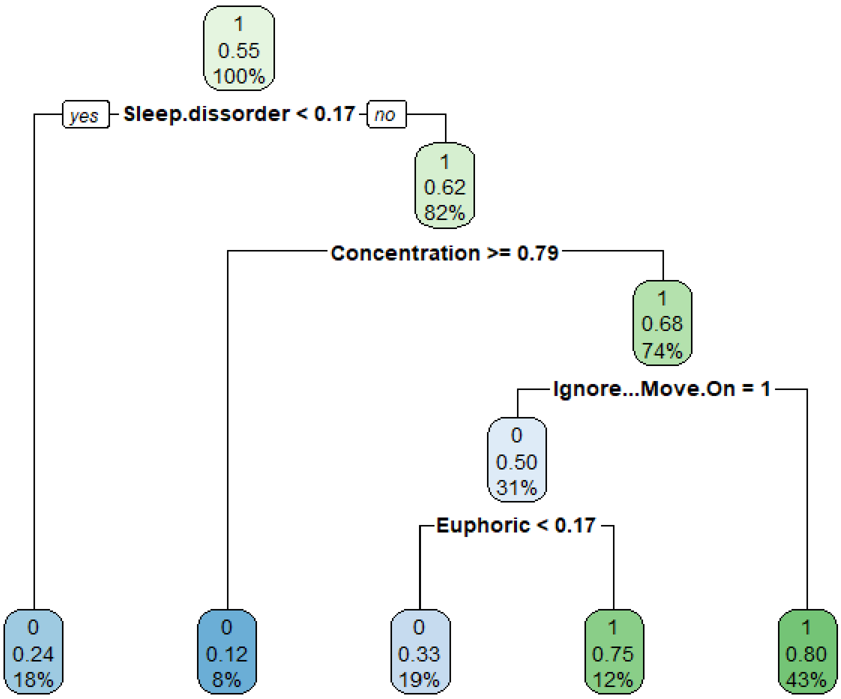

Figure 3: Decision Tree Visualization

Table 5: Decision Tree Performance Metrics

Model	Balanced Accuracy	Precision	Recall	F1
Decision Tree	0.694	0.909	0.555	0.689

4.7 Random Forest

A Random Forest classifier was implemented with extensive hyperparameter tuning. The best-performing parameters were:

Table 6: Optimal Random Forest Hyperparameters

Hyperparameter	Value
mtry	3
nodesize	1
minsplit	2
maxdepth	10

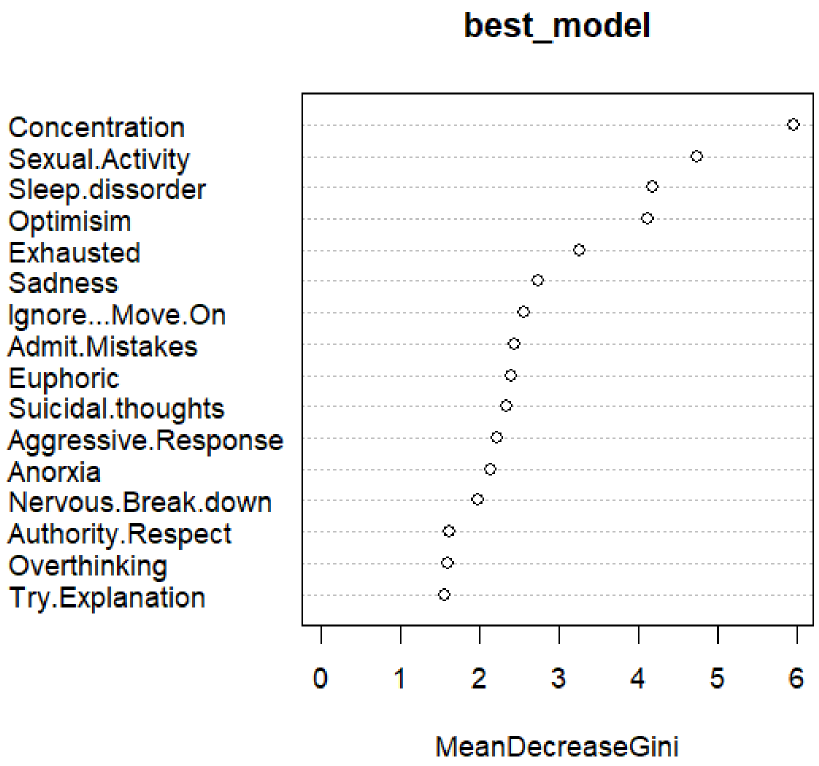

Figure 4: Random Forest Variable Importance

Table 7: Random Forest Performance Metrics

Model	Balanced Accuracy	Precision	Recall	F1
Random Forest	0.694	0.866	0.722	0.787

4.8 XGBoost

XGBoost was employed to further enhance predictive accuracy, leveraging parallel processing and extensive hyperparameter tuning.

Table 8: Optimal XGBoost Hyperparameters

Hyperparameter	**Value**
Number of Boosting Rounds (nrounds)	50
Maximum Depth (max_depth)	12
Learning Rate (eta)	0.3
Gamma (gamma)	0.5
Column Subsample Ratio (colsample_bytree)	0.5
Minimum Child Weight (min_child_weight)	5
Training Instance Subsample Ratio (subsample)	0.5

Table 9: XGBoost Performance Metrics

Model	**Balanced Accuracy**	**Precision**	**Recall**	**F1**
XGBoost	0.777	0.928	0.722	0.812

4.9 Results Summary

A comparison of all models tested in this study is provided in the table below:

Table 10: Summary of Supervised Learning Model Performance

Model	**Balanced Accuracy**	**Precision**	**Recall**	**F1**
Logistic Regression	0.638	0.846	0.611	0.709
K-Nearest Neighbors	0.722	0.916	0.611	0.733
Decision Tree	0.694	0.909	0.555	0.689
Random Forest	0.694	0.866	0.722	0.787
XGBoost	0.777	0.928	0.722	0.812

4.10 Supervised Learning Conclusion

Among the models tested, **XGBoost demonstrated the highest performance**, achieving an F1 score of 0.812. This suggests that XGBoost is highly effective at correctly identifying true cases of bipolar disorder while minimizing false positives. The balanced accuracy of 0.777 further supports its capability as a robust diagnostic tool.

While Random Forest also performed well, particularly in terms of recall, the overall balance achieved by XGBoost makes it the preferred model for this dataset. Given its strong results across multiple evaluation metrics, XGBoost could serve as a reliable clinical decision-making tool for diagnosing bipolar disorder.

5 Unsupervised Learning Analysis

5.1 Introduction

Delving into the unsupervised learning segment of the study, the aim was to explore various clustering techniques to potentially illuminate the differences among diagnostic categories, including normal, depressed, bipolar type 1, and bipolar type 2. To achieve this, a comprehensive suite of advanced methods was implemented, encompassing Principal Component Analysis (PCA), Uniform Manifold Approximation and Projection (UMAP), and t-Distributed Stochastic Neighbor Embedding (t-SNE), coupled with K-Means and hierarchical clustering approaches.

5.2 Methodological Framework

5.2.1 PCA with Clustering

PCA: Dimensionality was reduced through PCA, which simplifies the complexity while retaining the essence of the data, crucial for effective clustering.

K-Means and Hierarchical Clustering: Post-PCA, both K-Means and hierarchical clustering techniques were applied to the dataset. K-Means is utilized to capture central tendencies, while hierarchical clustering is employed to understand nested data groupings. This combination offers detailed structural insights into the data clusters.

5.2.2 t-SNE with Clustering

t-SNE: Known for its effectiveness in managing complex poly-dimensional data, t-SNE was employed as it provides a basis for deeply nuanced data separations.

Clustering on t-SNE: K-Means clustering was used to scrutinize the t-SNE results, aiming to identify clear patterns and clusters that correlate with clinical diagnoses.

The integration of these advanced techniques allows for rigorous analysis and visualization of the dataset from multiple angles. By clustering the data post-application of PCA, UMAP, and t-SNE, the aim was to graphically highlight differences between diagnostic categories based on expert opinions and to determine if unsupervised learning can offer new insights or confirm existing knowledge. This multifaceted approach enhances the ability to discern and visualize the underlying patterns that distinguish between different mental health states.

5.3 Principal Component Analysis (PCA)

PCA was performed on the standardized features extracted from the dataset, which included various psychological and behavioral variables. The PCA results indicated that the first principal component (PC1) accounted for approximately 18.87% of the variance, with a cumulative proportion of variance gradually increasing across the components. The first eight principal components were retained based on their cumulative explanatory power, covering about 70% of the total variance. This provided a substantial reduction in dimensionality while retaining the core information.

5.4 Elbow Method for Optimal Cluster Determination

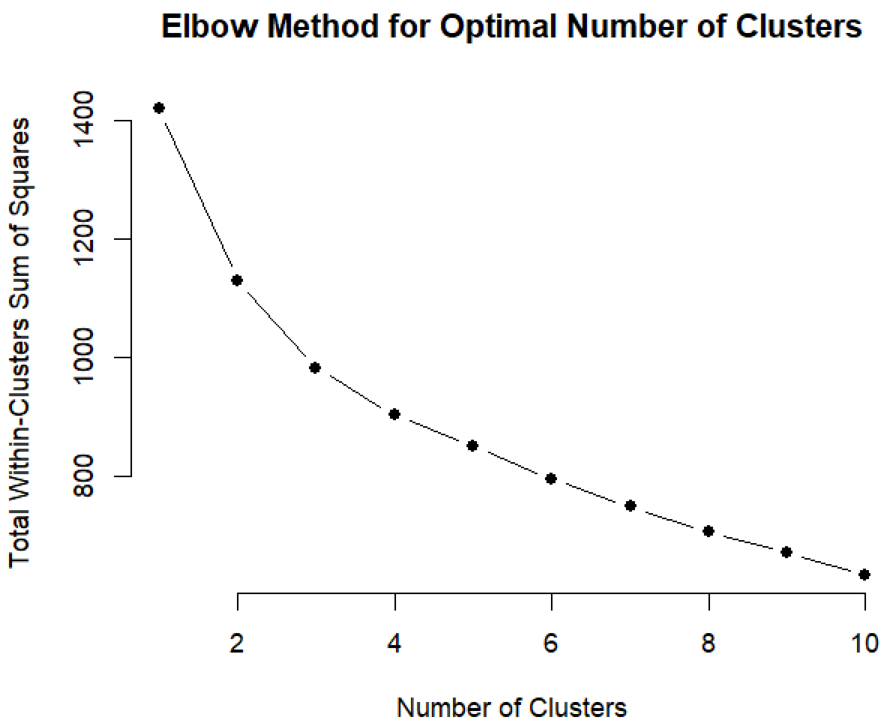

Figure 5: Elbow Method Analysis

From the Elbow Method analysis, it was determined that four clusters are the ideal number for the dataset. This optimal clustering not only helps in structuring the data effectively but also sets the stage for investigating if distinct expert diagnoses can be visually differentiated within each cluster. This approach is particularly useful for identifying whether specific mental health conditions cluster together, potentially enhancing the understanding of underlying patterns.

5.5 PCA with K-Means Clustering

With the optimal number of clusters determined, K-Means clustering was conducted to group the data into four distinct clusters. This clustering was visualized using the first two principal components, and the clusters were further differentiated using shapes based on expert diagnoses. This visualization aimed to explore if distinct psychological profiles corresponding to different mental health conditions (such as normal, depressed, bipolar type 1, and bipolar type 2) could be identified graphically.

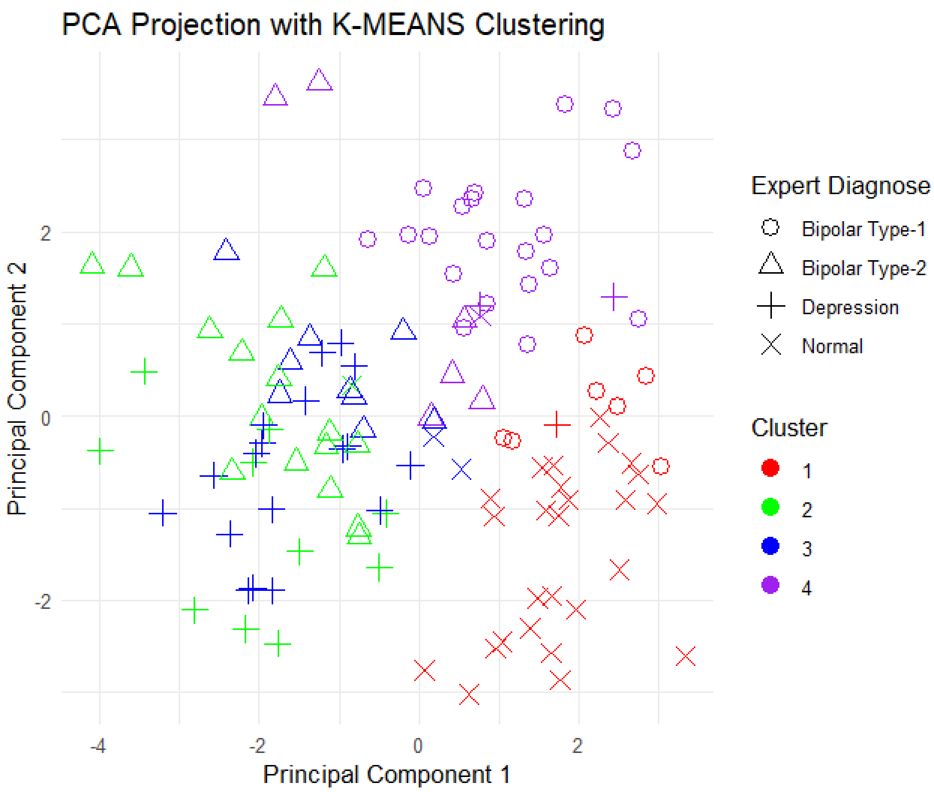

Figure 6: PCA with K-Means Clustering Visualization

Average Silhouette Width: 0.153

5.5.1 PCA with KMeans: Evaluation of Clustering Performance

The results obtained using PCA with KMeans are not satisfactory. Although some separation is visible in the visualization, the average Silhouette Width of 0.153 indicates that the clusters are not well-defined. This suggests that exploring alternative techniques could be beneficial to improve the clustering outcomes.

5.6 PCA with Hierarchical Clustering

After implementing PCA followed by K-Means clustering to initially explore data grouping, hierarchical clustering on the PCA-reduced data was utilized to refine the clustering approach. A dendrogram using the complete linkage method was constructed to visually assess the hierarchical structure of the data. Cutting the dendrogram at four clusters was determined to be optimal, consistent with earlier findings from the K-Means approach.

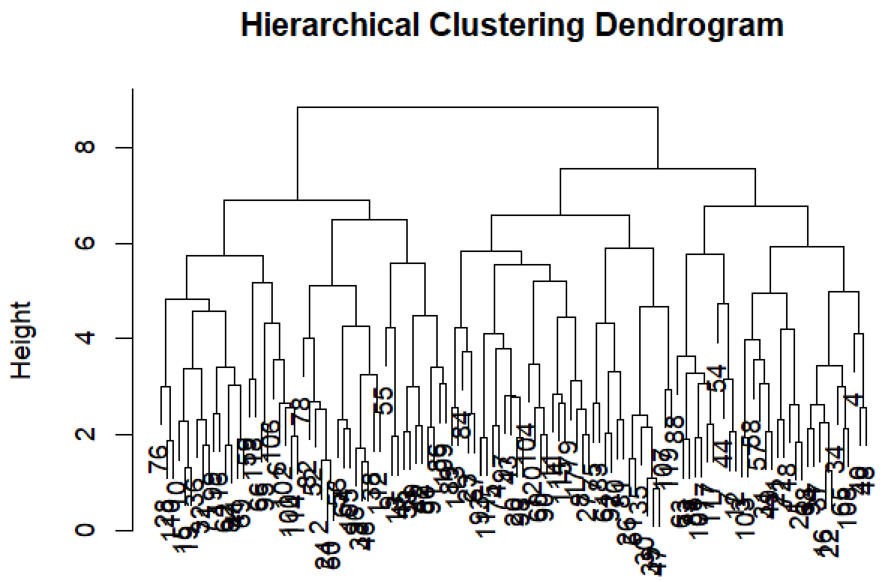

Figure 7: Hierarchical Clustering Dendrogram

Average Silhouette Width: 0.099

5.7 PCA Clustering Final Considerations

The relatively low silhouette scores obtained from both PCA-KMeans and PCA-hierarchical clustering indicated only moderate separation between the clusters. This result suggests that the PCA-based methods were limited in their ability to clearly distinguish between the underlying groupings in the data.

5.8 Introduction to t-SNE for Improved Clustering

t-SNE is particularly well-suited for capturing the complex non-linear relationships inherent in high-dimensional data. By implementing t-SNE, the objective is to determine whether better clustering results can be achieved, potentially revealing more distinct groupings or uncovering subtle patterns not captured by previous methods such as PCA and UMAP. This approach further investigates the intricacies of mental health diagnoses and enhances the clarity of the findings.

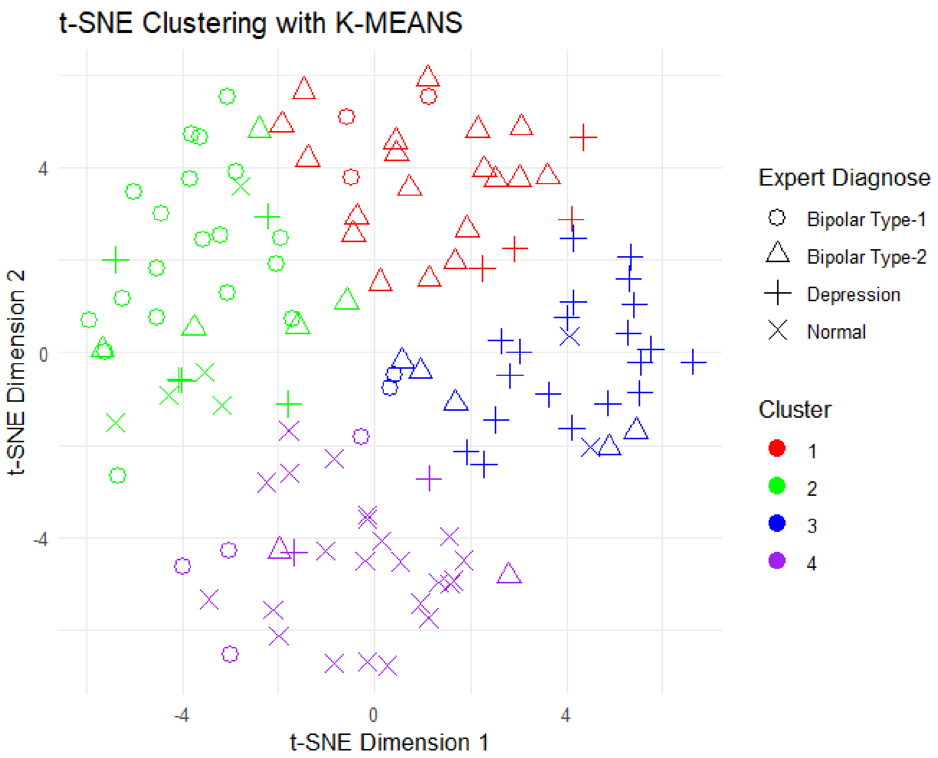

Figure 8: t-SNE with K-Means Clustering

Average Silhouette Width: 0.42

5.8.1 t-SNE with KMeans: Enhanced Cluster Definition

Since the t-SNE with KMeans clustering exhibits visible separation in the visualization and achieves an Average Silhouette Width of 0.42—which is significantly higher than that obtained with PCA clustering—it was decided to conduct a detailed analysis of each cluster.

5.9 t-SNE Clustering Discussion

The clustering results derived from the t-SNE with K-Means technique, which achieved a silhouette score of 0.42, effectively categorize individuals into distinct groups that align closely with their mental health diagnoses. Each cluster exhibits unique characteristics corresponding to specific psychological patterns and behaviors, thereby providing a deeper understanding of the nuances among different mental health conditions.

Table 11: Cluster Summary Statistics

Cluster	Sadness	Euphoric	Exhausted	Sleep.dissorder	Mood.Swing
1	0.6538462	0.1666667	0.6923077	0.5641026	0.8461538
2	0.4411765	0.3235294	0.3921569	0.4705882	0.6764706
3	0.7126437	0.1609195	0.6896552	0.5862069	0.2068966
4	0.3010753	0.5591398	0.4516129	0.3440860	0.1935484

Cluster	Suicidal.thoughts	Anorxia	Authority.Respect	Try.Explanation
1	0.96153846	0.5384615	0.53846154	0.5384615
2	0.29411765	0.5588235	0.05882353	0.5588235
3	0.65517241	0.1034483	0.10344828	0.4137931
4	0.06451613	0.3225806	0.90322581	0.3870968

Cluster	Aggr.Response	Ignore.Move.On	Nervous.Break.down	Admit.Mistakes
1	0.3846154	0.5384615	0.8461538	0.3076923
2	0.9705882	0.2058824	0.4705882	0.2941176
3	0.2068966	0.4137931	0.5172414	0.6551724
4	0.2903226	0.5483871	0.2903226	0.7096774

Cluster	Overthinking	Sexual.Activity	Concentration	Optimisim
1	0.4230769	0.3605769	0.3131868	0.2980769
2	0.4117647	0.6323529	0.4243697	0.5551471
3	0.8965517	0.3275862	0.5073892	0.2672414
4	0.4516129	0.5080645	0.5944700	0.5685484

- **Cluster 1:** Primarily comprising individuals diagnosed with bipolar type 2, this cluster is marked by significant emotional distress, high levels of sadness, and the highest incidence of suicidal thoughts. It also shows pronounced mood swings and symptoms of anorexia, reflecting the volatile emotional states associated with bipolar type 2 disorder.

- **Cluster 2:** Largely associated with bipolar type 1, this cluster is distinguished by high levels of aggressiveness and a notable disregard for authority. It exhibits considerable mood swings, although not as severe as in Cluster 1, indicating a commonality of emotional instability across bipolar disorders.

- **Cluster 3:** Mostly includes individuals diagnosed with depression, featuring significant levels of sadness and a propensity for overthinking—typical symptoms of depression. While suicidal thoughts are present, they are less pronounced than in the bipolar clusters, and aggressive behaviors or issues with authority are absent.

- **Cluster 4:** Predominantly consists of individuals categorized as normal, showing the lowest levels of sadness, exhaustion, and suicidal thoughts. This cluster stands out for its high scores in optimism and concentration, indicating a generally stable and healthy emotional state.

6 Conclusion

This study has demonstrated the potential of machine learning to enhance the diagnosis of bipolar disorder by harnessing both supervised and unsupervised methods. Our analysis revealed that models like XGBoost can effectively capture the intricate relationships among psychological symptoms, with key predictors—such as sadness, suicidal thoughts, aggressive responses, and concentration difficulties—playing a critical role in the diagnostic process.

The clustering analyses further enriched our understanding by uncovering natural groupings within the patient data. These groupings closely mirrored expert diagnostic categories, suggesting that individuals with similar symptom profiles tend to cluster together in meaningful ways. Such insights not only validate the clinical relevance of the selected features but also highlight the promise of advanced dimensionality reduction techniques in revealing the subtle structure of mental health disorders.

Overall, the findings from this work provide a strong foundation for future research aimed at refining diagnostic tools. By integrating more diverse data sources and exploring additional modeling strategies, there is significant potential to further improve diagnostic accuracy and ultimately support more personalized and effective treatment approaches in mental health care.