




ORIGINAL

## Detection of bipolar disorder by means of ensemble machine learning classifier

### Detección del trastorno bipolar mediante un clasificador de aprendizaje automático conjunto

Lingeswari Sivagnanam<sup>1</sup>  , N. Karthikeyani Visalakshi<sup>1</sup>  

<sup>1</sup>Government Arts and Science College, Computer Science. Kangeyam, India.

Cite as: Sivagnanam L, Visalakshi NK. Detection of bipolar disorder by means of ensemble machine learning classifier. Data and Metadata. 2023; 2:134. <https://doi.org/10.56294/dm2023134>

Submitted: 02-08-2023

Revised: 07-10-2023

Accepted: 03-12-2023

Published: 04-12-2023

Editor: Prof. Dr. Javier González Argote 

#### ABSTRACT

The accurate diagnosis of bipolar disorder is extremely challenging, due to unpredictable mood swings, behaviors, sleep, judgment, and inability to think, which makes it difficult to make a proper diagnosis. This paper aims to investigate the application of ensemble classifiers in classifying bipolar disorder and to compare their performance with existing methods. Herein, the work involves a thorough analysis of diagnostic precision and performance metrics. According to a study, an existing classifier achieved an accuracy rate of 87 % in bipolar disorder classification. In addition, the two most widely used classifiers, which are Random Forest and Decision Tree, achieved accuracy rates of 90 % and 86 %, respectively. These results highlight the performance baseline against which the proposed ensemble classifier is evaluated. Notably, the proposed methodology ensemble classifier shows excellent results in bipolar disorder classification thereby, achieving an impressive accuracy rate of 98 %. This considerable improvement in accuracy marks a significant stride in diagnostic precision, showcasing the potential of ensemble classifiers in enhancing bipolar disorder detection. The results of this study have given substantial implications for the field of mental health diagnosis, offering a promising avenue for a more accurate and reliable classification of bipolar disorder. This research reinforces the significance of advanced machine learning techniques and their potential to revolutionize the approach to diagnose and to manage mental health conditions.

**Keywords:** Bipolar Disorder; Machine Learning; Ensemble Classifier.

#### RESUMEN

El diagnóstico preciso del trastorno bipolar es extremadamente difícil, debido a la imprevisibilidad de los cambios de humor, el comportamiento, el sueño, el juicio y la incapacidad para pensar, lo que dificulta la realización de un diagnóstico adecuado. Este trabajo tiene como objetivo investigar la aplicación de clasificadores ensemble en la clasificación del trastorno bipolar y comparar su rendimiento con los métodos existentes. Para ello, el trabajo incluye un análisis exhaustivo de la precisión diagnóstica y las métricas de rendimiento. Según un estudio, un clasificador existente alcanzó una tasa de precisión del 87 % en la clasificación del trastorno bipolar. Además, los dos clasificadores más utilizados, que son Random Forest y Decision Tree, alcanzaron tasas de precisión del 90 % y el 86 %, respectivamente. Estos resultados ponen de relieve la línea de base de rendimiento con la que se evalúa el clasificador conjunto propuesto. En particular, el clasificador conjunto propuesto muestra excelentes resultados en la clasificación del trastorno bipolar, alcanzando una impresionante tasa de precisión del 98 %. Esta considerable mejora en la precisión marca un avance significativo en la precisión diagnóstica, mostrando el potencial de los clasificadores ensemble para mejorar la detección del trastorno bipolar. The results of this study have given substantial implications for the field of mental health diagnosis, offering a promising avenue for a more accurate and reliable classification of bipolar disorder. This research reinforces the significance of advanced machine learning techniques

and their potential to revolutionize the approach to diagnose and to manage mental health conditions.

**Keywords:** Bipolar Disorder; Machine Learning; Ensemble Classifier.

## INTRODUCTION

Bipolar Disorder is one of the serious mental health disorders typically don't occur immediately; rather, they evolve gradually and exhibits its symptoms. This research endeavors to address this challenge by focusing on the development of accurate, accessible, and dependable diagnostic methods with Feature Engineering (FE) techniques and classifier models that have been employed by previous researchers for diagnosing mental disorders. The primary objective is to boost the reliability of bipolar disorder diagnosis, which eventually results in improved patient care and better outcomes by applying machine learning ensemble classifier.

### Related Works

Machine learning is progressively finding applications in the field of psychiatry, aiming to enhance genetic analysis,<sup>(1,2)</sup> identifying the most distinctive differences among patient groups and control groups, and contributing to the diagnostic classification process. The data comprises the results of SNP analysis from five studies and transcriptomic analysis from two studies.<sup>(3)</sup> Exploring data from the relatively underexplored part of the cerebellum can introduce novel and valuable bio signatures to our understanding of BD's pathogenesis, progression, and even treatment response and resistance.<sup>(4)</sup> Machine learning feature selection algorithms opted for biomarkers based on their collective predictive value for the model rather than relying merely on individual (univariate) p-values or correlations with the outcome.<sup>(5)</sup> Consequently, feature selection algorithms may discard the statistically significant markers that do not add substantial value to the final model and may uncover individually insignificant markers, and become meaningful when considered in combination with others.<sup>(6)</sup> In two studies<sup>7,8)</sup> the novel neural network model is intended to provide a precise diagnosis of bipolar disorder in patients. In an article<sup>(9)</sup> promising results have been found in predicting tasks while researching machine learning algorithms including Support Vector Machines, Random Forest, Naive Bayes, and Multilayer Perceptron. Some articles<sup>(10,11)</sup> presents the proposed approach, which is a blend of varoius deep learning architectures: bidirectional long-term memory (biLSTMs) and convolutional neural networks (CNNs).

## METHODS

The methodology focuses on bipolar disorder within the realm of machine learning research is to enhance the lives of those affected by intricate mental health condition. Likewise, it helps in early detection, expediting diagnosis, and intervention, which is paramount for effective management. Additionally, machine learning models can assess the risk of severe episodes and self-harm, enabling timely intervention by healthcare providers. Extensive research, stigma reduction, and advancements in support systems collectively contribute to a holistic approach to addressing the challenges of bipolar disorder. This ultimately nurtures a more empathetic and evidence-based comprehension of the condition and promotes advanced public health efforts directed at enhancing the overall mental well-being of the community. The following figure is the proposed structure diagram.

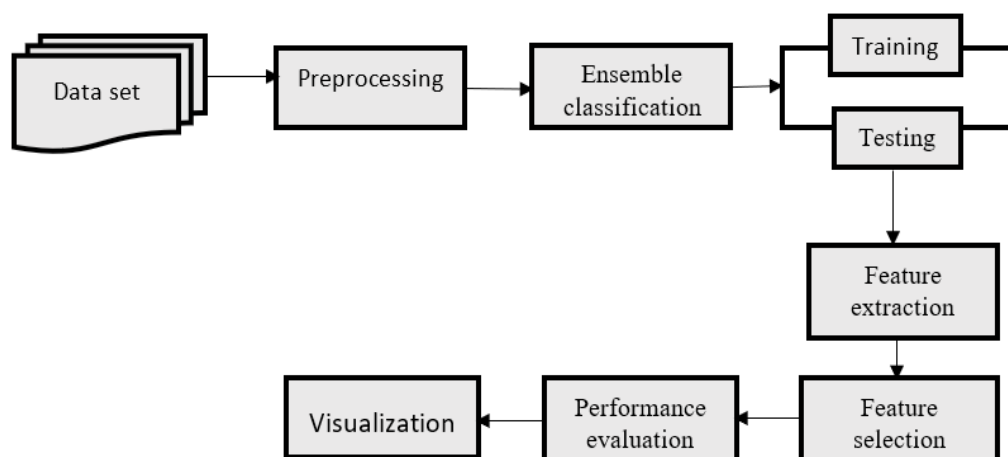


Figure 1. Proposed structure diagram

### *Dataset*

The dataset collection deals with the perusal records of patient's data. This contains patients' medical history as well as treatment details. Data relevant to the research has been analysed with the help of Clinical Specialists of the mental health research centre. The retrieved data has been stored in excel sheet in .CSV format. 183 samples of data were collected. Around twenty five features such as age, gender, previous treatment, family history, symptoms duration, psychiatric illness, ECT, etc.

### *Data Preprocessing*

Preprocessing the data is to guarantee its quality. This involves identifying and managing missing values, addressing data outliers, and rectifying any inconsistencies or inaccuracies. Ensuring Proper data cleaning is essential for precise model training.

### *Feature Engineering*

Create or choose suitable features (variables) that are relevant to the problem. These Features might include demographic information, symptom assessments, medication records, and genetic markers. The process of Feature engineering has a substantial impact on model performance.

### *Data Splitting*

Splitting the dataset into three distinct subsets: the training set, the validation set, and the test set. Training sets are used to train a machine learning model, validation sets are used to tune hyperparameters and select models, and test sets are used to evaluate the final model.

### *Model Selection*

Select the machine learning algorithms and models that most pertain to the problem. This selection may depend on the nature of the data and the specific task, while common options include decision trees, random forests, support vector machines, Relevant vector machine, Multilayer perceptron or neural networks.

### *Model Training*

Instruct the chosen machine learning models on the training dataset using the selected features. During this phase, the models will learn patterns and relationships within the data.

### *Hyperparameter Tuning:*

Optimize the model's hyperparameters utilizing the validation dataset. Make adjustments to parameters like learning rates or tree depths to succeed in optimal model performance.

### *Model Evaluation and Implementation*

Assess the model's effectiveness by utilizing the test dataset. Standard assessment metrics for bipolar disorder-related tasks include accuracy, F1 score, ROC, or clinical metrics relevant to the specific problem. The implementation tool used is WEKA for experimental results.

### *Model Deployment*

Implementing the trained machine learning model in a clinical or research environment was to predict or offer recommendations. This could entail integrating the model into previous healthcare systems or applications.

### *Monitoring and Feedback*

Consistently monitor the model's performance in real-world applications and collect feedback from healthcare professionals and users. Used this feedback to improve the model and refine its recommendations as time progresses.

### *Ethical Considerations*

Ensure that the machine learning model adheres to ethical guidelines and respects patient privacy and data security. Comply with relevant regulations and acquire the required consent, data privacy concerns rigorously.

### *Collaboration*

Making a strong collaboration with mental health professionals, clinicians, and domain experts in the entire implementation process. Their expertise and recommendations are important in ensuring the model's align with the highest clinical standards.

### *Iterative Improvement*

Iteratively enhance to refine both the model and the data preprocessing procedures, based on feedback and fresh insights. The mental health field is dynamic, and ongoing improvement is imperative to ensure the model's accuracy and applicability. Implementing machine learning for bipolar disorder involves an intricate and multifaceted process. It necessitates careful planning, and a strong focus on ethical and privacy considerations to create effective and safe solutions in this sensitive domain.

### Ensemble classifiers

The implementation of ensemble classifiers in the domain of bipolar disorder involves the application of ensemble learning techniques to enhance the accuracy and dependability of machine learning models employed for tasks like diagnosis, prediction, or outcome assessment. Commencing the process requires careful preparation of the dataset by cleaning, engineering features, and splitting the data. Subsequently, a diverse selection of base machine learning models such as decision trees, support vector machines, and logistic regression, among others was made. Ensemble methods like Bagging, Boosting, and stacking are then implemented to combine the predictions from these base models. Bagging works to diminish variance and enhances model stability by training on bootstrapped samples, while boosting focuses on sequential training to rectify errors and increase accuracy. Stacking combines the strengths of distinct models using a meta-model trained on their predictions. The process involves hyperparameter tuning and cross-validation to ensure optimal performance. Finally, the Adaboost ensemble classifiers are assessed using appropriate metrics and deployed in clinical or research settings, adhering to ethical guidelines and monitoring their performance over time. Closely with mental health professionals is crucial to align the models with the best clinical practices in managing bipolar disorder.

Algorithm: Ensemble Classifier for Bipolar Disorder using AdaBoost

1. Input
2. Initialize the Features of (X) and target label (Y) from the preprocessed data set
3. Output
4. parameters are evaluated Accuracy, sensitivity, specificity-measure
5. Step 1: Data Splitting
6.  $D_s = D_1, D_2, D_3, \dots, D_n$  data splitting
7. set  $(X_{train}, Y_{train})$ , and test data set  $(X_{test}, Y_{test})$  split (80 % training, 20 % testing)
8. Step 2:
9. Initialize the ensemble classifier
10. Create the ensemble classifier with hyperparameters
11. Base estimator (DT, RF or any other suitable classifier)
12. Number of boosting stages (estimators)
13. Learning rate (Accuracy, sensitivity, specificity)
14. Random state for reproducibility
15. Step 3: Train the Classifier
16. Training classifier initialized
17. Train the ensemble classifier on training data set  $(X_{train}, Y_{train})$  using the fit method.
18. Step 4: Make Predictions
19. Test the ensemble classifier on test data set  $(X_{test}, Y_{test})$  using the fit method.
20. Step 5: Model Evaluation
  - Evaluate the model's accuracy by comparing predicted labels ( $y_{pred}$ ) to true labels ( $y_{test}$ ) using the `accuracy_score` function.
21. Step 6: Display Accuracy
  - Display or return the accuracy of the AdaBoost ensemble classifier on the test data.

## RESULTS

The results involves in the accuracy rates of the different machine learning classifiers with its performance metrics as follows

### Confusion matrix

A confusion matrix provides a detailed account of the classification of the problem's prediction results. These Predictions fall into two types: correct and incorrect. The determination of correct or incorrect predictions depends on broken down and counted values. The finest thing that makes it particularly valuable is its ability to reveal the error type made by that specific classifier and how the classifier made the error.

### Description of the Terms

Positive (P):                      The Observation is positive

Negative (N):	The Observation is not positive
True Positive (TP):	The Observation is positive and is predicted as positive.
False Negative (FN):	The Observation is positive but is predicted as negative.
True Negative (TN):	The Observation is negative and is predicted as negative.
False Positive (FP):	The Observation is negative but is predicted as positive.

**Recall**

The recall is solved by dividing the correctly classified positive examples by the total positive samples. When there are minimal FN and robust recall values, it indicates that examples were accurately recognized. It is calculated as follows;

$$\text{Recall} = (TP) / (TP)+FN$$

**Precision**

Precision is determined by dividing the total accurately classified positive samples by the total predicted positive samples. Elevated precision values signify positive outcomes. It is calculated as follows;

$$\text{Precision} = (TP) / (TP)+FP$$

**F-measure**

The F-measure is derived from calculating Recall and Precision. Unlike, the Arithmetic mean, it employs the Harmonic mean measurement, as it is particularly effective with higher values. Meanwhile, the F-measure consistently yields a value lower than either Recall or Precision. The calculation is as follows:

$$F - \text{measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

**ROC**

The Receiver Operating Characteristic (ROC) is a fundamental tool in machine learning for assessing the performance of binary classification models. It offers valuable insights into a model's ability to discriminate between two classes, such as identifying disease presence or absence. The table 1 above, presents an assessment of outcomes in the training set, which includes accurately classified and misclassified instances. The parameters mentioned above outline the respective findings.

Correctly Classified Instances	162	98,5246 %
Incorrectly Classified Instances	21	2,4754 %
Kappa statistic		0,6988
Mean absolute error		0,0853
Root mean squared error		0,1918
Relative absolute error		38,0724 %
Root relative squared error		57,7147 %
Total Number of Instances		183

Table 2 illustrates a range of various performance metrics, such as accuracy, rates of true positive (TP) and false positive (FP) rates, precision, recall, F-measure, and ROC curve.

TP Rate	0,879
FP Rate	0,013
Precision	0,935
Recall	0,879
F-measure	0,906
ROC	0,997

Table 3 shows the maximum and minimum values, along with the standard deviation, and mean values.

Statistics	Value
------------	-------

Minimum	1
Maximum	183
Standard deviation	52,972
Mean	92

Table 4, represents the outcomes of the current methodology against those of the proposed ensemble classifier. The previous classifiers achieved an accuracy of 87 %, while the Random Forest and Decision Tree classifiers achieved an accuracy of 90 % and 86 %, respectively. Notably, the proposed ensemble classifier surpasses the other earlier methods, by achieving an impressive accuracy of 98 %. The graph results show as figure 2.

Classification models	Accuracy rates
KNN	87 %
Random forest	90 %
Decision tree	86 %
Ensemble classifier	98 %

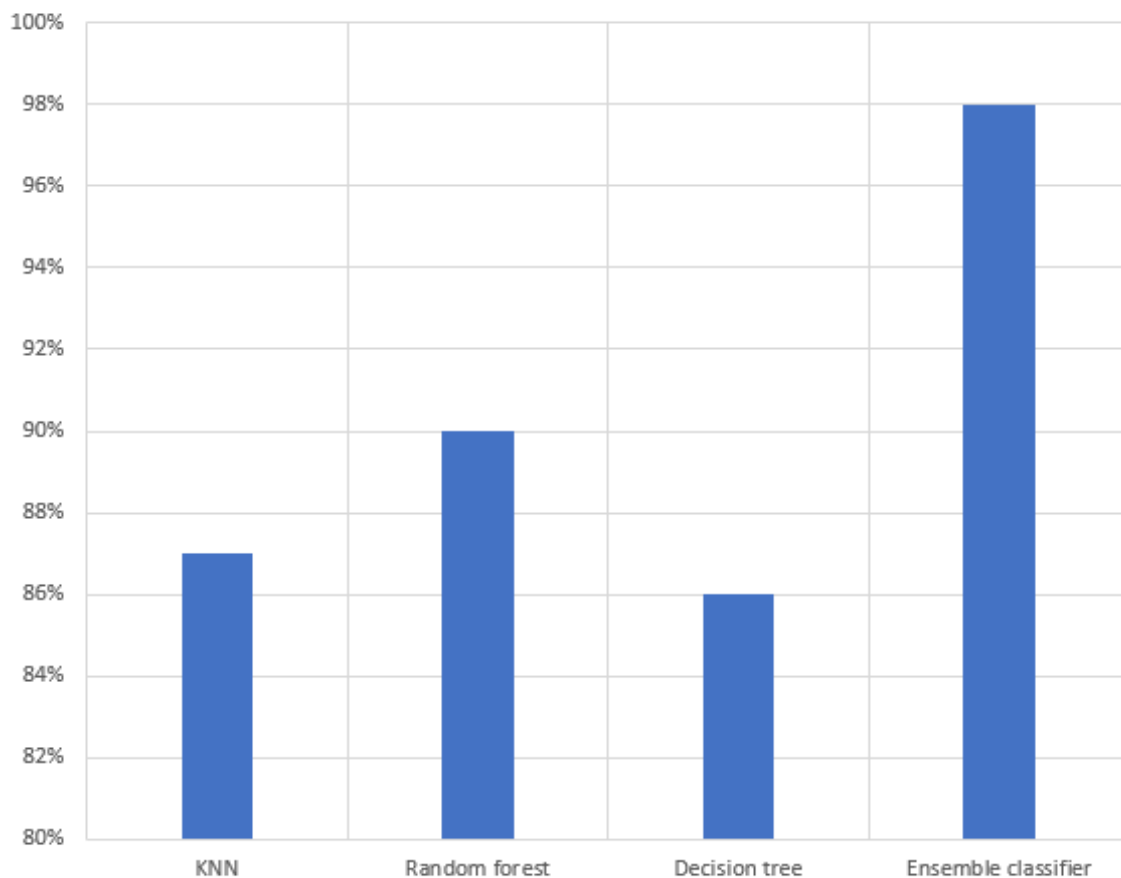


Figure 2. Comparison between existing algorithms

## DISCUSSION

The considerable enhancement in accuracy has the potential to revolutionize the precision and dependability of bipolar disorder diagnosis. Additionally, this study advances the domain of mental health diagnosis by demonstrating the practical applicability of advanced machine learning techniques. The impetus of this research is to unlock the difficulties of bipolar disorder diagnosis. This includes its intricate nature, characterized by abnormal mood swings, meanwhile diagnosing this mental health condition proves to be a formidable challenge. Previous diagnostic approaches often struggle to grasp the nuances of this condition accurately.<sup>(13)</sup>

The impetus behind this study is the conviction that enhanced diagnostic accuracy is important for the well-being of individuals who are struggling with bipolar disorder. Moreover, the rapid advancements in machine learning provides a distinctive opportunity to establish in the field of mental health diagnostics, offering more

precise and reliable approach. The novelty of this research offers insights into the application of ensemble classifiers for bipolar disorder classification, further understanding the critical aspects of mental health treatment.

## CONCLUSIONS

This research has concluded a substantial advancement in the field of bipolar disorder classification. Adaboost ensemble classifier outperforms existing methods, achieving an impressive accuracy rate of 98 %. This notable performance significantly exceeds the existing accuracy rate of both existing classifiers 87 % and the Random Forest (90 %) and Decision Tree (86 %) classifiers. Thus, these outcomes emphasize the potential of ensemble classifiers as a beneficial tool for enhancing the precision and reliability of bipolar disorder. The implications of these discoveries are more intense. This study can be further enhanced by combining various types of machine learning models into a single ensemble classifier, proposed in the detection of bipolar disorder. This approach aims to achieve even more accurate and reliable in bipolar disorder classification, benefiting early interventions and personalized treatments.

## BIBLIOGRAPHIC REFERENCES

1. Bracher-Smith M, Crawford K, Escott-Price V. Machine learning for genetic prediction of psychiatric disorders: a systematic review. *Mol Psychiatry*. 2021; 26:70-79.
2. Karthik S, Sudha M. Predicting bipolar disorder and schizophrenia based on non-overlapping genetic phenotypes using deep neural network. *Evol Intell*. 2021; 14:619-634.
3. Canova-Barríos C, Machuca-Contreras F. Interoperability standards in Health Information Systems: systematic review. *Seminars in Medical Writing and Education*. 2022;1:7. <https://doi.org/10.56294/mw20227>
4. Wang Y, Wang J, Jia Y, Zhong S, Niu M, Sun Y, Qi Z, Zhao L, Huang L, Huang R. Shared and specific intrinsic functional connectivity patterns in unmedicated bipolar disorder and major depressive disorder. *Sci Rep*. 2017; 7:3570.
5. Hernandez LM, Kim M, Hoftman GD, Haney JR, de la Torre-Ubieta L, Pasaniuc B, Gandal MJ. Transcriptomic insight into the polygenic mechanisms underlying psychiatric disorders. *Biol Psychiatry*. 2021; 89:54-64.
6. Aghaizu ND, Jolly S, Samra SK, Kalmar B, Craessaerts K, Greensmith L, Salinas PC, De Strooper B, Whiting PJ. Microglial expression of the Wnt signaling modulator DKK2 differs between human Alzheimer's disease brains and mouse neurodegeneration models. *ENEURO*.0306-22.2022 *eneuro* 10.
7. Alnafisah RS, Reigle J, Eladawi MA, O'Donovan SM, Funk AJ, Meller J, Mccullumsmith RE, Shukla R. Assessing the effects of antipsychotic medications on schizophrenia functional analysis: a postmortem proteome study. *Neuropsychopharmacology*. 2022; 47:2033-2041.
8. Yashaswini KA, Aditya Kishore Saxena. A Novel Predictive Scheme for Confirming State of Bipolar Disorder using Recurrent Decision Tree. *Int J Adv Comput Sci Appl*. 2022;13(2).
9. Li Z, Li W, Wei Y, Gui G, Zhang R, Liu H, Chen Y, Jiang Y. Deep learning based automatic diagnosis of first-episode psychosis, bipolar disorder and healthy controls. *Comput Med Imaging Graph*. 2021. doi:10.1016/j.compmedimag.2021.101882.
10. Rotenberg LS, Borges-Júnior RG, Lafer B, Salvini R, Dias RDS. Exploring machine learning to predict depressive relapses of bipolar disorder patients. *J Affect Disord*. 2021. doi:10.1016/j.jad.2021.08.127.
11. Ganasigamony WJ, Muthuraj AA Selvaraj. Computer assisted diagnosis of bipolar disorder using invariant features. *Concurrency and Computation: Practice and Experience*. 2022;e6984.
12. Araujo Inastrilla CR. Big Data in Health Information Systems. *Seminars in Medical Writing and Education*. 2022; 1:6. <https://doi.org/10.56294/mw20226>
13. Kour H, Gupta MK. An hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM. *Multimed Tools Appl*. 2022. doi:10.1007/s11042-022-12648-y.

## **FINANCING**

The authors did not receive financing for the development of this research.

## **CONFLICT OF INTEREST**

The authors declare that there is no conflict of interest.

## **AUTHORSHIP CONTRIBUTION**

*Conceptualization:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Data curation:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Formal analysis:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Acquisition of funds:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Research:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Methodology:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Project management:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Resources:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Software:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Supervision:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Validation:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Display:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Drafting - original draft:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.

*Writing - proofreading and editing:* Lingeswari Sivagnanam and N. Karthikeyani Visalakshi.