IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 4, December 2024, pp. 4352~4363
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i4.pp4352-4363

DepXGBoot: Depression detection using a robust tuned extreme gradient boosting model generator

U. Ananthanagu, Pooja Agarwal
Department of Computer Science and Engineering, PES University, Bengaluru, India

Article Info

Article history:

Received Nov 23, 2023
Revised Mar 22, 2024
Accepted Apr 17, 2024

Keywords:

Classification
Depression
Electronic health records
Extreme gradient boosting
Machine learning

ABSTRACT

In terms of severity and prevalence, depression is the worst. Suicide rates have risen because of this and are on the rise universally. Consequently, effective diagnosis and therapy must reduce the impact of depression. There is often more than one factor at play when determining why someone has been diagnosed with depression. In addition to alcohol and substance abuse, other possible causes include problems with physical health, adverse reactions to medications, life-changing events, and social circumstances. In this paper, exploratory data analysis is conducted to understand the insights of the sensorimotor database depression comprising depressive experiences in individuals who are either unipolar or bipolar. This study proposes a robust tuned extreme gradient boosting model generator to automatically predict the state of depression. The performance is optimized by determining the best combination of hyperparameters for the extreme gradient boosting model. By harnessing the power of advanced machine learning methodologies, this study underscores comparative analysis and the importance of data-driven innovation in mental health and clinical practice. Future developments for the robust tuned extreme gradient boosting model's application and study to forecast depression in the sensorimotor database depression can be used to track changes in depressed states over time by integrating it with longitudinal and multimodal data.

Corresponding Author:

U. Ananthanagu
Department of Computer Science and Engineering, PES University
Bengaluru, India
Email: nagu.anantha@gmail.com

1. INTRODUCTION

One runs the risk of developing a mental disorder when subjected to persistent strain. These vulnerabilities often are examples of peer pressure, heart attacks, depression, and other adverse outcomes. Presently, one of the most common mental ailments is depression worldwide. It is estimated that five percent of adults worldwide suffer from this mental illness. It affects enormously the global health population and is among the leading causes of disability globally. It may even result in suicide. Depression may vary from moderate to severe, but all degrees are treatable. Early diagnosis and efficient patient care can reduce depression, which can also help in decreasing society's suffering [1], [2]. People with clinical depression suffer from severe impairment in their daily lives due to their stumpy mood or lack of interest in previously enjoyed activities [3]. The number of depressive symptoms, intensity, duration, and impact on a person's capacity to carry out their personal and professional lives determine the severity of a depressive episode [4]. Bipolar depression is a chronic psychological condition defined by intense mood fluctuations [5]–[7]. Bipolar disorder also manifests itself in the form of depression as a counterpoint to the manic state and is linked to higher activity, limited sleep, hyperactivity, target-oriented actions, and perceived self [8]. The lack of manic symptoms in someone suffering from unipolar depression makes the patient distinct from a bipolar person [9]. Both depression and bipolar illness are innate conditions that are best understood as intrinsic sensitivity to the natural state being disturbed by external factors [10].

More than half of those who take their own lives meet the definition of clinical depression. Major depressive disorder (MDD), often known as clinical depression, it's a serious mental disorder marked by persevering feelings of melancholy and pessimism as well as a decreased level of participation or enjoyment in activities. The symptoms of clinical depression can seriously obstruct an individual's general well-being by creating some issues with mental, physical, and actual working. It can affect individuals of various ages, personalities, and circumstances [11]. The traditional strategy for treating clinical depression typically involves a multimodal treatment plan that incorporates prescription, psychotherapy, and lifestyle changes. Many people with depression, for example, are unable or unable to acknowledge their emotional well-being issues. As a result, identifying adequate and effective methods to detect depression is a burgeoning area of research, and recent advances in instrumentation and sensor technologies opened up new vistas in the diagnosis of depression [12]. Advances in machine learning have positively affected representation learning, categorization, and forecasting models developed using data from electronic health records. These records contain information from various regimens, which are then cataloged in a specific order for every treatment session. These data may include statistical profiles, diagnostic tests, medications or bodily health problems, adverse medical consequences, significant life events, and environmental circumstances. The records are gathered from a variety of modalities, they are heterogeneous and computationally complicated [13], [14].

Depression is a predominant emotional well-being problem that significantly affects an individual's satisfaction and general prosperity. However, there might be inaccuracies and failures in the traditional methods for diagnosing depression, like self-reported surveys and clinical assessments. Subsequently, it is basic to make areas of strength for a dependable model for the distinguishing proof of gloom that can perceive discouraged side effects from various information sources, for example, social, physiological, and textual signals. The proposed study aims to foster a precise and powerful gloom discovery model by using fine-tuning techniques, a strong artificial intelligence algorithm, and extreme gradient boosting (XGBoost). Working on the early diagnosis, treatment, and medication for depression is the objective to improve the personal lifestyle and emotional wellness results of the individuals affected negatively by this illness.

These are the primary contributions of this work: i) create an innovative method to automatically forecast the stages of depression based on patterns found in the input data, such as physical health issues, unfavorable medical outcomes, important life events, and environmental conditions; ii) incorporated robust feature engineering techniques to increase the model's accuracy in identifying depressive symptoms; iii) validated the proposed model on the dataset to ensure its effectiveness across different populations and demographic groups; and iv) comparative analysis was carried out and validate the model's efficacy through cross-validation techniques and assess its performance indicators such as F1 score, recall, accuracy, and precision. The relevant works on this topic comprise section 2 of this study. Details of the suggested methodology and data pre-processing are provided in section 3. Results and discussion are shown in section 4, and the conclusion is emphasized in section 5.

2. RELATED WORK

Much research has been conducted in many domains of neuroscience, applied linguistics, healthcare, and psychology to truly comprehend the underlying causes of depression in individuals. This section provides a summary of various relevant studies, their benefits and shortcomings, and how our research is an improvement on that work in this sector. Bilal and Khan [15] investigated the automatic classification of users into sad and nondepressed groups using machine learning techniques on text data and emojis. The researchers suggest a technique for creating an automated system that can identify sad people by fuzing emojis with the emotional process. The benchmark datasets of social networking websites with text and emojis are used to classify users according to linguistic style, temporal process, emotional process, and other criteria. In this study, novel classifiers are presented, which are trained and assessed using various combinations of part-of-speech tags.

A method for detecting depression based on three modalities facial expressions, auditory, and gait was proposed by Dai *et al.* [16], every modality's performance is compared; facial expressions perform the best, followed by audio, while gaits perform the worst. This study shows that combining several modalities can enhance depression detection efficacy. This paper assesses the suggested model's efficacy using the open datasets emotion-gait, AVEC 2013, and AVEC 2014. Welch and Bishop [17] proposed a fusion technique using the Kalman filter with segmental and spectral features. They used classifiers to choose important video features and rich audio parameters to obtain a more accurate model. The authors identify emotional valence based on physiological variables such as pupil size, respiration signal, temperature, mouth length, and skin conductance as well as facial expressions and the NeuCube structure, that uses a developing spiking neural network (SNN) architecture. Applying feature-level fusion, the proposed method obtains a 73.15% classification accuracy for binary valence, it is equivalent to current deep learning methods. Interestingly, this precision is attained without the use of EEG, which is sometimes a need for other deep learning techniques [18]. The random forest regression technique with multi-modal fusion was used by Samareh *et al.* [19], based on 1,425 audio, 13 visual, and eight text features. This technique performed better than its predecessors. By creating a framework for those with mental health issues, Son *et al.* [20] addressed the poor use of mental health care. Using smartphone sensor data from both regular and depressed subjects, this study builds models for emergency and depression identification. this study uses several algorithms, including variational autoencoder (VAE), deep autoencoding gaussian mixture model (DAGMM), empirical cumulative distribution-based outlier detection (ECOD), copula-based outlier detection (COPOD), and light gradient boosting machine (LGBM), to identify emergencies and depression. With an even better 0.99 F1 score, the emergency detection model demonstrated its accuracy in identifying emergency circumstances.

Utilization of biomarker data from a large Dutch dataset in conjunction with machine learning model is applied to improve the identification of depression cases. Investigation of the identification of depression cases in the 11,081 Dutch citizens in the sample. using a variety of resampling techniques, including, over-sampling, over-undersampling, under-sampling, and random over-sampling examples (ROSE) sampling, to address the dataset's class imbalance issue. Using the XGBoost, machine learning technique, patients with mental illness are distinguished from healthy patients. Using the test dataset derived from the O-sample, the XGBoost model Xgb. O performed the best, achieving an F1 score of 0.9762, 0.9987 recall, 0.9729 high measures in accuracy, 0.9765 balanced accuracy, and 0.9548 precision [21].

Fitzpatrick *et al.* [22] presented Woebot, a completely automated conversational chatbot that treats cognitive behaviors in depressed young adults using demographic factors and patient health questionnaire (PHQ) scores to execute univariate exploration. This bot agent showed a significant reduction of depressive modes among users [23], [24]. Aldarwish and Ahmad [25] presented a method that categorizes people based on their psychological well-being. Artificial intelligence is at the heart of this system. Support vector machine (SVM) and naive Bayes (NB) are employed. Nandanwar and Nallamolu [26] offered a method for machine learning based on the AdaBoost classifier and synthetic minority oversampling technique (SMOTE) technique to predict depression using labeled information collected from Twitter.

In the wake of recent advancements in natural language processing (NLP) [27], [28], there are now many electronic health record data analysis methods. Digital health information, which is formed of time series sequences from many data modalities, can employ gradual learning on the sequences of words [29]. the bidirectional encoder representations from transformers for electronic health records (BEHRT) model proposed by Li *et al.* [30], is a sequence transduction model for use in electronic health record, which is particularly adept at predicting the possibility of 301 conditions in an individual's future visits. The system's adaptable design allows it to include diverse ideas to increase its precision. Jazaery and Guo [31] provided a novel method to estimate depression levels from visual data using a 3D convolutional neural network (CNN) to automatically learn spatiotemporal features from successive facial expressions.

According to the literature survey, the majority of depression detection research conducted nowadays makes use of homogeneous datasets, such as electronic health records or self-reported surveys. It is recommended to investigate a variety of data sources, such as social media posts, smartphone sensor data, and physiological signals, to improve robustness and generalizability. There is disagreement about which features are most informative, even though machine learning models like XGBoost show promise in the identification of depression. Enhancing model performance and interpretability through the identification and extraction of pertinent features from diverse sources should be the primary focus of future research. The poll also highlights the need for interpretable models because current ones are frequently viewed as "black boxes," which undermines confidence in depression detection criteria and understanding.

3. PROPOSED METHODOLOGY

This study aims to create a strong model using extreme gradient boosting XGBoost to detect depression early. In attempting to address the issue of depression detection, it is necessary to follow a systematic approach to obtain accurate and reliable results. Our proposed depression detection model, as shown in Figure 1, follows a well-structured methodology. It starts by collecting behavioral datasets relevant to depression. These datasets are pre-processed, normalized, converted to categorical data, and relevant features extracted. The core model generator is XGBoost, which provides scalability according to how complicated the data is. The algorithm is fine-tuned and optimized through hyperparameter fine-tuning to improve predictive accuracy. The algorithm is tested and validated across diverse demographics. The recall, F1 score, and accuracy of the depression symptom detection are evaluated. Ethical and privacy considerations are considered throughout the data collection, preparation, and model development. This evidence-based approach allows early intervention and customized treatments to improve the standard of living for people with depression and contributes significantly to psychiatric investigation. This is achieved through the use of the extreme gradient boost algorithm and the integration of strong preprocessing and validation techniques.

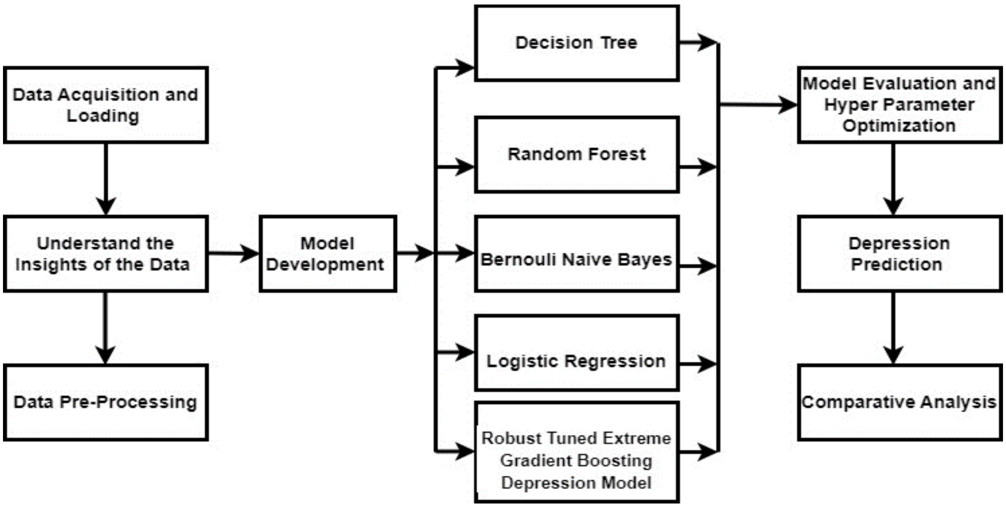
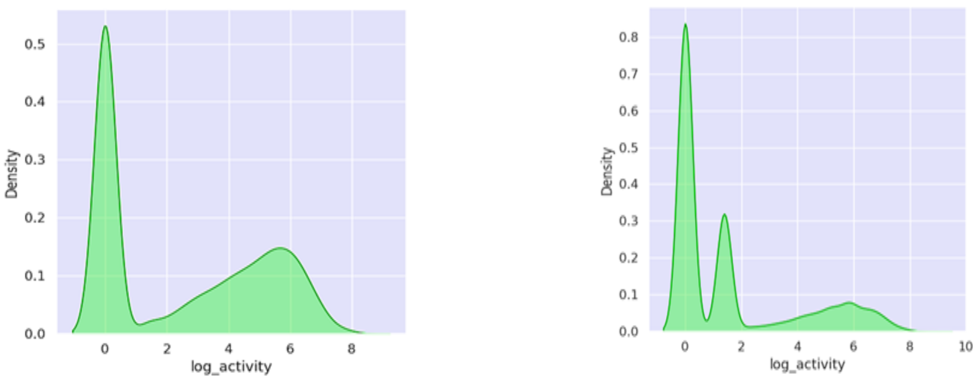


Figure 1. Proposed methodology

3.1. Data acquisition and loading

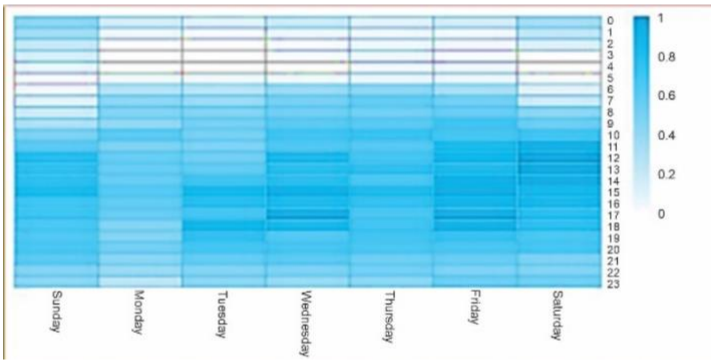
Nowadays, wearable sensors measure various types of activities of people wearing them. These gadgets also gather research information. This dataset offers sensor information from numerous days of continuous monitoring and specific statistical profiles of each patient [32], [33]. The right hand was utilized to wear an actigraph wristwatch, which tracked the individual's motor activity. Actigraph watches use piezoelectric accelerometers to integrate durability, volume, and movement duration in any direction. Motions greater than 0.05 g were captured at a 32 Hz sample rate. The Depresjon dataset contains data on patients without depression (controls) compared with those who do (cases). The dataset includes the sensorimotor recordings of depressed 23 patients and 32 healthy individuals. There were two categories of depressed patients: unipolar and bipolar. In the initial phase and the end of the motor activity recordings, a therapist used the montgomery-asberg depression rating scale (MADRS) grading system [34], [35] to determine the intensity of the continuous depression the patient was experiencing. Figure 2(a) shows the log activity of the condition, while Figure 2(b) shows the log activity of control patients and the density graph. Analyzing activities by weekday and hour could reveal how different times and days influence mood and behavior.



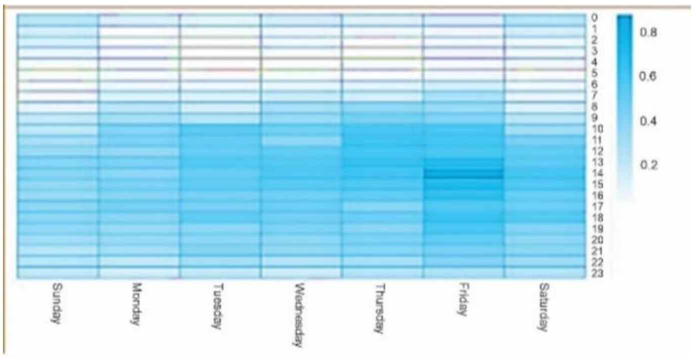
(a) (b)

Figure 2. Log activity of (a) density graph of the condition activity and (b) density graph of the control activity

Figure 3(a) shows that the control patient's physical activities are better as observed from the graph, which is highlighted in dark blue shades, whereas the Figure 3(b) shows condition patient's physical activities are less, which is observed from the graph with low blue color. This analysis could provide insights into potential triggers and patterns that contribute to the onset or exacerbation of depressive episodes. It highlights fluctuations in depressive symptoms based on routines, work-life balance, or other environmental factors.



(a)

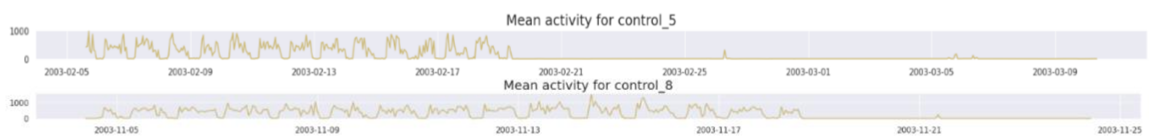


(b)

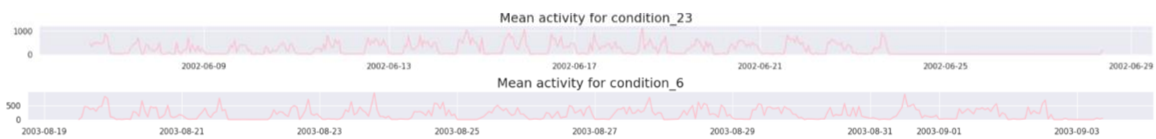
Figure 3. Group activities by weekday and hour: (a) control and (b) condition

3.2. Understanding the insights of the data

Multiplex data visualization presents the plot to understand the mean activity of control (healthy) and condition (with depression) subjects. The random mean activities of the subjects are generated in Figures 4(a) and 4(b). There are two datasets, which are called the 'activity,' and another is the 'scores'. Derived columns are generated for mean log activity, standard log activity, minimum log activity, maximum log activity, and zero proportion activity to make the data more efficient for the prediction model.



(a)



(b)

Figure 4. Mean activity of the random subjects: (a) control and (b) condition

In the ‘scores’ dataset, different levels of depression data are stored for condition subjects. During this stage, an unnecessary data column marking the subject number is dropped from the ‘scores’ data frame. Figure 5 shows the histogram plot and Figure 6 presents the density plot for the various data columns in this dataset. After exploratory data analysis (EDA) comes to the standardization process of data in order to get the data appropriate for the models of machine learning, the procedure begins with missing data handling in the ‘scores’ dataset. Some missing data portions are filled with ‘2.0’ to avoid empty spaces in the dataset. Some other missing data are replaced with ‘0’ or a range (example: ‘<6’ for the ‘edu’ column). Categorical data conversion is performed after that step for different data columns. In this case, the list of categorical columns include –'gender', 'age', 'afftype', 'melanch', 'inpatient', 'edu', 'marriage' and ‘work’. Then dummy, columns are introduced in various ranges for each categorical column. After that, the ‘activity’ and the ‘scores’ datasets are merged. The two redundant columns (‘source’, ‘state’) are dropped from this new merged dataset. After all these steps are completed, the whole dataset is divided into two sections: the training and testing datasets.

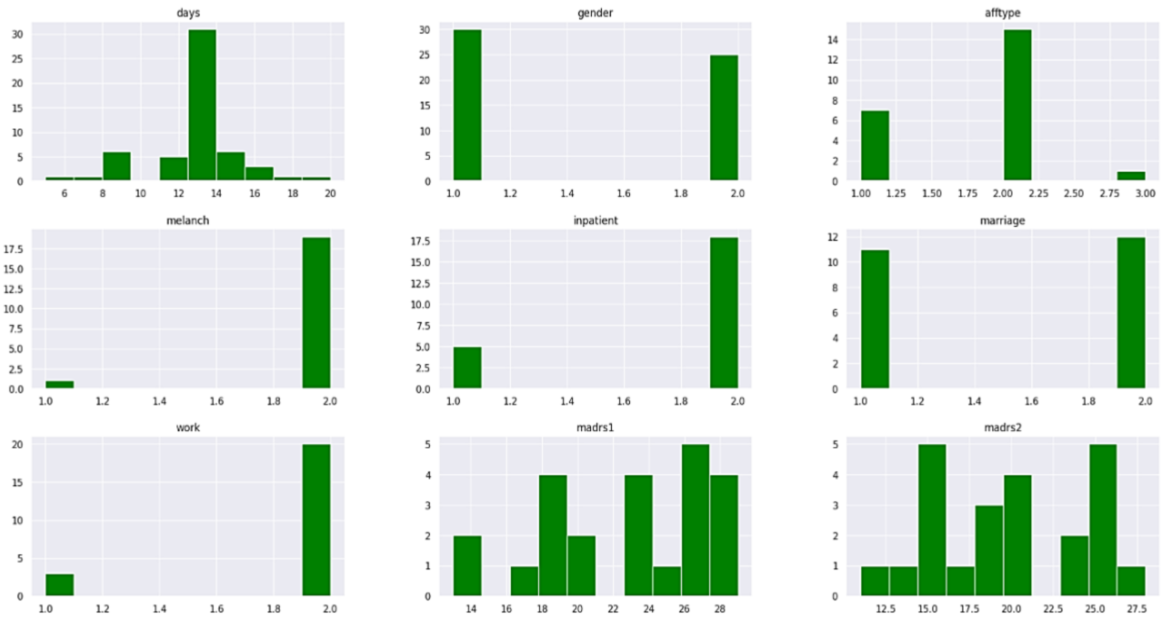


Figure 5. Histogram plot of various data present in scores

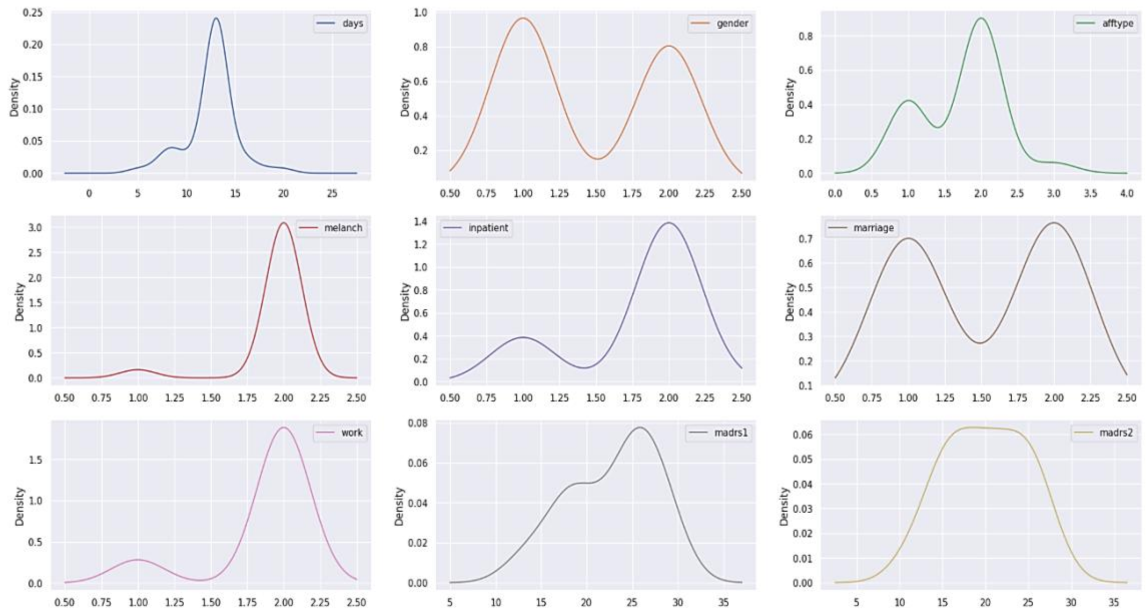


Figure 6. Density plot of various data columns in scores dataset

3.3. Model development

Preparing the dataset for model construction involves handling encoding categorical variables, scaling numerical features, missing values, and applying decision trees, logistic regression, NB, random forests, and the suggested model. Next, the pre-processed dataset is subjected to each method. To optimize information gain or Gini impurity for the decision tree, the model is trained to recursively partition the data depending on features. With a random forest, an ensemble of decision trees is built, using bootstrapped samples and random feature selection to reduce overfitting. NB leverages Bayes' theorem with the assumption of feature independence, and the model is trained to predict class probabilities based on the likelihood of features. Logistic regression, a linear classification model, is trained to predict the likelihood of the binary outcome using a logistic function. Finally, each model's performance parameters, including F1 score, precision, recall, and accuracy were assessed in a comparative study.

3.3.1. Decision tree classifier

A classifier in the form of a tree's arrangement is a supervised learning method called a decision tree. The inner nodes represent database properties, the branches indicate decision paths, and the leaf nodes represent cataloging. Finding all possible answers to a question or making decisions based on the available information can be done visually using this method. In this study, the maximum leaf node number was set to 10 for the decision tree classifier case. The finest nodes for the decision tree contain the fewest impurities. The random state determines the randomness of the estimator. Even when the separator is set to "best," each division adjusts the characteristics dynamically. When max has n attributes, the algorithm randomly selects max features before calculating the optimal partition. One must be chosen randomly if numerous splits have the same criterion improvement. The random state must be an integer for deterministic fitting. Here the set integer was 101.

3.3.2. Random forest classifier

Random forest strategy builds numerous "simple" decision trees during training and uses a clear majority for classification. This voting approach corrects the decision tree's tendency to overfit training data. Random forests bag individual trees during training. This bagging selects a random training set sample with replacement and fits trees. In this study, the forest had ten trees (n_estimators) and a maximum depth of 10. The number of concurrent jobs is specified by n_jobs. The trees are designed to run for fit, predict, decision path, and apply. Due to the parallelization of the backend implementation, the variablen-jobs were assigned 1. This model was then run with the training dataset to predict depression states.

3.3.3. Bernoulli naïve Bayes

Bayes' theorem-based NB classifiers create significant independence constraints among attributes. These Bayesian network models are simple. NBs relies on two assumptions. First, all features are conditionally independent. One feature does not influence another's performance. This is why "NB" is so-called. Second, all features are essential. To produce reliable forecasts, one must know all the features. In (1) depicts the Bayes theorem on which NB is based.

$$P(A|B) = \frac{P(B|A).P(A)}{P(B)} \tag{1}$$

Where A means event one and B means even two. P is the probability of an event happening. A supervised machine learning Bernoulli NB algorithm that comes in handy when the dataset has a binary output label distribution. It works with discrete data. This algorithm only accepts binary characteristics. In (2) and (3) describe Bernoulli's distribution.

$$p(x) = P[X = x] = \begin{cases} \mathrm{q = 1 - p,} & \mathrm{x = 0} \\ \mathrm{p} & \mathrm{x = 1} \end{cases} \tag{2}$$

$$X = \begin{cases} 1 & Bernoulli's\ trial = S \\ 0 & Bernoulli's\ trial = F \end{cases} \tag{3}$$

Where p means success and q means failure. From the (2) and (3), we can clearly see that x can work on only two instances (0 and 1). Based on this, Bernoulli's NB was applied to the training dataset to obtain an optimum prediction of depression state classes.

3.3.4. Logistic regression classifier

It is a popular supervised machine learning algorithm. A set of independent variables predicts the category-dependent variable. The result must be discrete or categorical. Logistic regression can categorize continuous and discrete data using probabilities. Logistic regression can be binomial (0 or 1, true or false), multinomial ('class A', 'class B', 'class C'.), or ordinal (types having quantifiable implications). This method works well with a large volume of data. In this study, C is given a value of 1e3 (+1.000E3) for the case of logistic regression model application on the training dataset. The value of C must be a positive float because it represents the inverse of regularization strength. Smaller values, as in SVMs, indicate greater regularization. Limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) is an optimization algorithm commonly used for logistic regression optimization. The 'LBFGS' solver is used here to optimize the model.

3.3.5. Proposed depression detection model: robust tuned XGBoost depression model generator

Boosting is a form of ensemble modeling that combines several less effective classifiers to construct a robust classifier. Every gradient boosting predictor corrects its predecessor's faults. XGBoost is such a boosting method that implements gradient-boosted decision trees. This algorithm builds judgment trees step-by-step and prioritizes weights. After weighting each independent variable, the operation is completed. Factors whose results were inaccurately predicted previously are given to their next decision tree with enhanced consequences. Then, the classifiers and predictors are integrated to create a more accurate model. XGBoost overcomes the disadvantages of random fittings (high bias and low variance) and provides several parameters to modify to produce a good, durable machine learning model. First, the training and testing dataset is further compiled using the regex for this model. The extreme gradient boosting classifier is then applied to this modified training dataset. In most cases, the parameters utilized by ordinary XGBoost are in the form of integers. However, in this innovative approach to machine learning, the dynamic value was used to tweak the characteristics of the XGBoost model. This was accomplished by feeding the dynamic value into the model's parameter. The algorithm for the robust tuned XGBoost depression model is as follows:

Algorithm: robust-tuned XGBoost depression model

Step 1: formally, let $\hat{y}_i^{(t)}$ be the prediction of the ith instance at the tth iteration, will need to add ft to minimize the following objective.

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{1}$$

Including ft enhances our model. Second-order approximation can optimize the aim quickly.

$$\mathcal{L}^{(t)} \simeq \sum_{i=1}^{n} [l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \tag{2}$$

where $\mathrm{g}_i = \partial_{\hat{y}_i^{(t-1)}} \mathrm{l}(y_i, \hat{y}_i^{(t-1)})$ and $\mathrm{h}_i = \partial^2_{\hat{y}_i^{(t-1)}} \mathrm{l}(y_i, \hat{y}_i^{(t-1)})$ loss function both first and second ordering gradient statistical analysis. We can simplify step 1 by removing the constant terms.

$$\tilde{\mathcal{L}}^{(t)} = \sum_{i=1}^{n} [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) \tag{3}$$

Step 2: define Ij={i|q(xi)=j} as instance set of leaf j. We can rewrite in (1) by expanding Ω as follows:

$$\begin{aligned} \tilde{\mathcal{L}}^{(t)} &= \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \\ &= \sum_{j=1}^{T} \left[\left(\sum_{i \epsilon I_j} g_i\right) w_j + \frac{1}{2}\left(\sum_{i \epsilon I_j} h_i + \lambda\right) w_j^2\right] + \gamma T \end{aligned} \tag{4}$$

For a fixed structure q(x), we compute the optimal weight w_j^* of leaf j as follows:

$$w_j^* = -\frac{\sum_{i \epsilon I_j} g_i}{\sum_{i \epsilon I_j} h_i + \lambda},$$

And calculate the corresponding value using:

$$\tilde{\mathcal{L}}^{(t)} = -\frac{1}{2} \sum_{i=1}^{T} \frac{\left(\sum_{i \epsilon I_j} g_i\right)^2}{\sum_{i \epsilon I_j} h_i + \lambda} + \gamma T \tag{5}$$

Step 3: utilized is a greedy algorithm that begins with a single leaf and incrementally adds branches to the tree. Accept IL and IR as examples of the left and right centers following the split. I=IL ∪ IR at that point, the misfortune decrease after the split is given by:

$$\mathcal{L}_{split} = \frac{1}{2}\left[\sum_{i=1}^{T} \frac{\left(\sum_{i \epsilon I_L} g_i\right)^2}{\sum_{i \epsilon I_L} h_i + \lambda} + \sum_{i=1}^{T} \frac{\left(\sum_{i \epsilon I_R} g_i\right)^2}{\sum_{i \epsilon I_R} h_i + \lambda} - \sum_{i=1}^{T} \frac{(\sum_{i \epsilon I} g_i)^2}{\sum_{i \epsilon I} h_i + \lambda}\right] - \gamma \tag{6}$$

This (6) is used to evaluate divided candidates in practice.
Step 4: create a parameter grid for hyperparameters such as n_estimators, learning rate, max depth, and scoring parameters. Consider the values n_estimators=[5, 10, 15, 20], learning rate=[0.01, 0.1], and max depth=[5, 15, 2].
Step 5: configure the stratified k fold to handle the imbalance dataset. The parameters number of splits=4, random state=101.
Step 6: perform grid search cross validation using the function $E(x_1, x_2) = \|d^{obs} - d^{pre}(x)\|$. Where error function E and x1 and x2 values and d^{obs} is the original value and $d^{pre}(x)$ predicted traverse through the set of extreme gradient-boosting model hyperparameter values with stratified K fold, we find the robust depression finding model with optimized hyperparameter. GridSearch cross validation accelerates XGBoost model development, improving model performance across a broad range of prediction tasks and use cases. Automating hyper-parameter tuning and cross-validation for robust model evaluation helps prevent problems like overfitting. Grid search with cross validation. The brute-force approach to find the best hyper-parameter for a specified set of data and model is by examining every possible parameter permutation and combination to see which parameters will allow the balanced xboost algorithm to work.

4. RESULTS AND DISCUSSION

The depression state in this kernel is grouped into various classes using XGBoost and implemented using Python with the use of scikit-learn. A case ‘condition’, having result value as 1, indicates that the patient has clinical depression while a ‘control’ condition which its outcome value is 0 means that the patient is normal. The XGBoost was used together with k-fold and grid-search cross-validation (CV) in this study.

An XGBoost model for depression detection had a 100% accuracy score. Hyperparameters such as n_estimators, learning_rate, and max_depth was tested on a control and condition patient dataset to optimize model performance. To determine the combination that maximized model performance, a grid search over a range of hyperparameter values was conducted during the experiment. The performances of each configuration were assessed by systematically varying hyperparameters were evaluated with recall, accuracy, F1-score, and precision for the prediction of both condition (depression) and categories of negative sentiment: condition (positive for depression) and control (not in depression). The results are tabulated in Table 1. Certain hyperparameter configurations led to better performance in terms of predicting depression. For example, higher values of parameters such as n_estimators tend to yield higher F1-score and recall for the condition class (higher number of instances of depression being correctly detected); similarly lower learning rates are tied to higher precision and accuracy for the control class (higher number of non-depressed cases correctly detected, leading to a lower probability of misclassifying non-depressed individuals). Overall, we can see that hyperparameter tuning is important in achieving the best possible performance for the XGBoost for depression detection model. By systematically exploring different hyperparameter combinations, we identified the most effective configuration for accurately identifying instances of depression in the dataset. For the learning rate value 0.01, max depth value of 5 and n estimators’ value of 50, we obtained highest accuracy of 100%.

Table 1. Experimental results obtained during robust tuned extreme gradient-boosting depression model implementation

Experiment	Class	n estimators	Learning rate	Max depth	Precision	Recall	F1-score	Accuracy
1	0	100	0.1	3	0.9	0.82	0.86	0.88
	1	100	0.1	3	0.85	0.92	0.88	0.87
2	0	200	0.1	5	0.91	0.84	0.88	0.9
	1	200	0.1	5	0.87	0.88	0.87	0.89
3	0	300	0.1	7	0.88	0.9	0.89	0.87
	1	300	0.1	7	0.89	0.86	0.88	0.86
4	0	100	0.01	3	0.79	0.81	0.8	0.82
	1	100	0.01	3	0.82	0.84	0.83	0.82
5	0	200	0.01	5	0.85	0.77	0.81	0.79
	1	200	0.01	5	0.84	0.8	0.82	0.8
6	0	50	0.01	5	1.00	1.00	1.00	1.00
	1	50	0.01	5	1.00	1.00	1.00	1.00
7	0	100	0.001	3	0.83	0.75	0.79	0.81
	1	100	0.001	3	0.79	0.81	0.8	0.78
8	0	200	0.001	5	0.77	0.82	0.79	0.8
	1	200	0.001	5	0.82	0.79	0.8	0.81
9	0	300	0.001	7	0.8	0.78	0.79	0.81
	1	300	0.001	7	0.8	0.85	0.82	0.84
10	0	150	0.05	4	0.86	0.81	0.83	0.85
	1	150	0.05	4	0.85	0.82	0.84	0.83

Table 2 presents the results for both (0 and 1) classes. To build a robust model, we first analyzed the existing models. The existing models such as random forest, NB, decision tree, and logistic regression models were implemented. These models were trained for the depression dataset, the results are tabulated in Table 1. The results obtained show that the proposed robust tuned extreme gradient boosting depression model gives the highest accuracy score of 100%. The proposed model is also compared with a linear SVM model using multi-modal active appearance model (AAM) and vocal prosody, which gives an accuracy of 79%, and manual FAACS (facial action coding system (FACS)) coding gives an accuracy of 89%. Another linear SVM model used region-based resting-state functional connectivity to gain an accuracy of 94.3%. When the linear kernel was used with sparse SVM, the model gave an accuracy of 78.95% [36].

Table 2. Comparison of recall, preceision, accuracy, and F1-score of the built models

Model developed	Class	Precision	Recall	F1-score	Accuracy
Decision tree	0	0.97	0.95	0.96	94.8
	1	0.92	0.94	0.94	
Random forest	0	0.98	0.97	0.97	96.8
	1	0.94	0.96	0.95	
Bernouli naive Bayes	0	0.97	0.96	0.96	95.5
	1	0.93	0.94	0.94	
Logistic regression	0	0.96	0.98	0.98	97.6
	1	0.96	0.97	0.96	
Robust-tuned XGBoost depression model	0	1.0	1.0	1.0	100
	1	1.0	1.0	1.0	

Pattern classification using functional magnetic resonance imaging (fMRI) and oxygen level in the blood led to 86% accuracy [37]. The Twitter post-based fastText model gave an accuracy of 92.52%, while the bag-of-words models based on the AdaBoost method improved the accuracy rate to 93.09%. Compared with these machine learning models; the proposed model increased the accuracy to 100%. In Table 3, a comparison of the accuracy for various models and our proposed model is presented.

Table 3. Comparison of different models of depression detection

Model	Accuracy (%)
SVM (linear) using multimodal AAM [36]	79
FAACS and vocal prosody [16]	89
Sparse L1-norm SVM (linear) [37]	78.95
SVM (linear) using fMRI [37]	94.3
SVM using blood O2 levels and fMRI [37]	86
fasttext [26]	92.52
Bag-of-words with Ada Boost [26]	93.09
Robust-tuned XGBoost depression model generator	100

5. CONCLUSION

A calibrated robust gradient-boosting depression model generator was developed using numerous criteria. To achieve this, data are incorporated from the Depresjon dataset to train the model. Depressive episodes in unipolar and bipolar people are included in this sensorimotor database. It also contains data from healthy people, based on input data patterns such as bodily health issues, adverse medical repercussions, key life events, and environmental situations. This study proposes unique ways to forecast depression states automatically. The model was tweaked robustly to diagnose depression. Grid search with CV uses brute force to determine the optimal hyperparameters for a specific dataset and model. Our research paired them with XGBoost to achieve the best results for detecting depression. The proposed model robust gradient boosting depression model generator gave a satisfactory output as the accuracy was high for the testing dataset. In the future, more aspects of a person’s medical and experience history can be included in the dataset to get better results. It can also help avoid unintentional outliers. This work will be extended to conduct longitudinal studies to track the progression and recurrence of depressive symptoms over time, enabling the development of dynamic models capable of predicting the likelihood of relapse and tailoring personalized intervention strategies. Explore the integration of multi-modal data sources, including genetic markers, neuroimaging data, and socioeconomic factors, to build comprehensive models that capture the complex interplay of biological, psychological, and environmental determinants contributing to depression.

REFERENCES

[1] M. Rahul, S. Deena, R. Shylesh, and B. Lanitha, "Detecting and analyzing depression: a comprehensive survey of assessment tools and techniques," in *6th International Conference on Inventive Computation Technologies, ICICT 2023*, 2023, pp. 749–753, doi: 10.1109/ICICT57646.2023.10134165.

[2] D. Shi, X. Lu, Y. Liu, J. Yuan, T. Pan, and Y. Li, "Research on depression recognition using machine learning from speech," in *2021 International Conference on Asian Language Processing, IALP 2021*, 2021, pp. 52–56, doi: 10.1109/IALP54817.2021.9675271.

[3] D. L. Segal, "Diagnostic and statistical manual of mental disorders (DSM-IV-TR)," in *The Corsini Encyclopedia of Psychology*, Wiley, 2010, pp. 1–3.

[4] C. Norman and L. Fotheringham, "Depression in adults," *InnovAiT: Education and inspiration for general practice*, vol. 14, no. 3, pp. 176–184, Mar. 2021, doi: 10.1177/1755738020978681.

[5] E. A. Carbone *et al.*, "Antisocial personality disorder in bipolar disorder: a systematic review," *Medicina*, vol. 57, no. 2, pp. 1–17, 2021, doi: 10.3390/medicina57020183.

[6] I. Sekulic, M. Gjurković, and J. Šnajder, "Not just depressed: bipolar disorder prediction on reddit," pp. 72–78, 2019, doi: 10.18653/v1/w18-6211.

[7] D. J. Martino and M. P. Valerio, "Bipolar depression: a historical perspective of the current concept, with a focus on future research," *Harvard Review of Psychiatry*, vol. 29, no. 5, pp. 351–360, 2021, doi: 10.1097/HRP.0000000000000309.

[8] R. S. McIntyre, M. Zimmerman, J. F. Goldberg, and M. B. First, "Differential diagnosis of major depressive disorder versus bipolar disorder," *Journal of Clinical Psychiatry*, vol. 80, no. 3, pp. 15–24, 2019, doi: 10.4088/JCP.OT18043AH2.

[9] L. Sirignano *et al.*, "Depression and bipolar disorder subtypes differ in their genetic correlations with biological rhythms," *Scientific Reports*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-19720-5.

[10] S. A. Galkin and N. A. Bokhan, "The differential diagnosis of unipolar and bipolar depression based on EEG signals," *Zhurnal nevrologii i psikhiatrii im. S.S. Korsakova*, vol. 122, no. 11, p. 51, 2022, doi: 10.17116/jnevro202212211151.

[11] R. Karrouri, Z. Hammani, Y. Otheman, and R. Benjelloun, "Major depressive disorder: validated treatments and future challenges," *World Journal of Clinical Cases*, vol. 9, no. 31, pp. 9350–9367, 2021, doi: 10.12998/wjcc.v9.i31.9350.

[12] A. Seal, R. Bajpai, J. Agnihotri, A. Yazidi, E. H. -Viedma, and O. Krejcar, "DeprNet: a deep convolution neural network framework for detecting depression using EEG," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–13, 2021, doi: 10.1109/TIM.2021.3053999.

[13] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1589–1604, 2018, doi: 10.1109/JBHI.2017.2767063.

[14] Y. Meng, W. Speier, M. Ong, and C. W. Arnold, "Multi-level embedding with topic modeling on electronic health records for predicting depression," *Studies in Computational Intelligence*, vol. 914, pp. 241–246, 2021, doi: 10.1007/978-3-030-53352-6_22.

[15] U. Bilal and F. H. Khan, "An analysis of depression detection techniques from online social networks," *Communications in Computer and Information Science*, vol. 1198, pp. 296–308, 2020, doi: 10.1007/978-981-15-5232-8_26.

[16] Z. Dai, Q. Li, Y. Shang, and X. Wang, "Depression detection based on facial expression, audio, and gaitt," *ITNEC 2023 - IEEE 6th Information Technology, Networking, Electronic and Automation Control Conference*, pp. 1568–1573, 2023, doi: 10.1109/ITNEC56291.2023.10082163.

[17] G. Welch and G. Bishop, "An introduction to the Kalman Filter," *University of North Carolina at Chapel Hill*, pp. 1–3, 2020.

[18] C. Tan, G. Ceballos, N. Kasabov, and N. P. Subramaniyam, "Fusionsense: emotion classification using feature fusion of multimodal data and deep learning in a brain-inspired spiking neural network," *Sensors*, vol. 20, no. 18, pp. 1–27, 2020, doi: 10.3390/s20185328.

[19] A. Samareh, Y. Jin, Z. Wang, X. Chang, and S. Huang, "Predicting depression severity by multi-modal feature engineering and fusion," in *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, 2018, vol. 32, no. 1, pp. 8147–8148, doi: 10.1609/aaai.v32i1.12152.

[20] M. Son, G. Lee, J. Y. Park, and M. Choi, "Research on depression and emergency detection model using smartphone sensors," *Korean Institute of Smart Media*, vol. 12, no. 3, pp. 9–18, Apr. 2023, doi: 10.30693/SMJ.2023.12.3.9.

[21] J. F. Dipnall *et al.*, "Fusing data mining, machine learning, and traditional statistics to detect biomarkers associated with depression," *PLOS ONE*, vol. 11, no. 2, Feb. 2016, doi: 10.1371/journal.pone.0148195.

[22] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial," *JMIR Mental Health*, vol. 4, no. 2, 2017, doi: 10.2196/mental.7785.

[23] J. J. Prochaska *et al.*, "A therapeutic relational agent for reducing problematic substance use (Woebot): development and usability study," *Journal of Medical Internet Research*, vol. 23, no. 3, 2021, doi: 10.2196/24850.

[24] H. M. Demirci, "User experience over time with conversational agents: case study of woebot on supporting subjective well-being," *M.Sc. Thesis*, School of Natural and Applied Sciences, Middle East Technical University, Çankaya, Turki, 2018.

[25] M. M. Aldarwish and H. F. Ahmad, "Predicting depression levels using social media posts," in *2017 IEEE 13th International Symposium on Autonomous Decentralized Systems, ISADS 2017*, 2017, pp. 277–280, doi: 10.1109/ISADS.2017.41.

[26] H. Nandanwar and S. Nallamolu, "Depression prediction on twitter using machine learning algorithms," *2021 2nd Global Conference for Advancement in Technology, GCAT 2021*, 2021, doi: 10.1109/GCAT52182.2021.9587695.

[27] Y. Juhn and H. Liu, "Artificial intelligence approaches using natural language processing to advance EHR-based clinical research," *Journal of Allergy and Clinical Immunology*, vol. 145, no. 2, pp. 463–469, 2020, doi: 10.1016/j.jaci.2019.12.897.

[28] Y. Luo *et al.*, "Natural language processing for EHR-based pharmacovigilance: a structured review," *Drug Safety*, vol. 40, no. 11, pp. 1075–1089, 2017, doi: 10.1007/s40264-017-0558-6.

[29] T. T. V. Vleck *et al.*, "Augmented intelligence with natural language processing applied to electronic health records for identifying patients with non-alcoholic fatty liver disease at risk for disease progression," *International Journal of Medical Informatics*, vol. 129, pp. 334–341, Sep. 2019, doi: 10.1016/j.ijmedinf.2019.06.028.

[30] Y. Li *et al.*, "BEHRT: transformer for electronic health records," *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020, doi: 10.1038/s41598-020-62922-y.

[31] M. Al Jazaery and G. Guo, "Video-based depression level analysis by encoding deep spatiotemporal features," *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 262–268, 2021, doi: 10.1109/TAFFC.2018.2870884.

[32] I. Peis *et al.*, "Actigraphic recording of motor activity in depressed inpatients: a novel computational approach to prediction of clinical course and hospital discharge," *Scientific Reports*, vol. 10, no. 1, 2020, doi: 10.1038/s41598-020-74425-x.

[33] Y. Tazawa *et al.*, "Actigraphy for evaluation of mood disorders: a systematic review and meta-analysis," *Journal of Affective Disorders*, vol. 253, pp. 257–269, 2019, doi: 10.1016/j.jad.2019.04.087.

[34] E. G. -Ceja *et al.*, "Depresjon: a motor activity database of depression episodes in unipolar and bipolar patients," in *Proceedings of the 9th ACM Multimedia Systems Conference, MMSys 2018*, 2018, pp. 472–477, doi: 10.1145/3204949.3208125.

[35] L. Z. -Calzada *et al.*, "Feature extraction in motor activity signal: towards a depression episodes detection in unipolar and bipolar patients," *Diagnostics*, vol. 9, no. 1, Jan. 2019, doi: 10.3390/diagnostics9010008.

[36] J. H. Majed, S. A. Nasser, A. Alkhayyat, and I. A. Hashim, "Artificial intelligent algorithms based depression detection system," in *IICETA 2022 - 5th International Conference on Engineering Technology and its Applications*, 2022, pp. 408–413, doi: 10.1109/IICETA54559.2022.9888485.

[37] X. Wang, Y. Ren, and W. Zhang, "Depression disorder classification of fMRI data using sparse low-rank functional brain network and graph-based features," *Computational and Mathematical Methods in Medicine*, vol. 2017, 2017, doi: 10.1155/2017/3609821.

BIOGRAPHIES OF AUTHORS

U. Ananthanagu currently pursuing a Ph.D. at PES University, India. He holds a master's degree in computer science and engineering from Jawaharlal Nehru Technological University, a bachelor's degree in Computer Science and Engineering from Visvesvaraya Technological University, and a Diploma in Computer Science and engineering from the Technical Board of Education. With 14 years of experience in academics, his primary areas of research interest include data science, machine learning, NLP, social media analysis, and transfer learning. He has published over 20 papers in international journals and conferences. He is a life member of the Computer Society of India (CSI) and the Indian Society for Technical Education (ISTE). He can be contacted at email: nagu.anantha@gmail.com or pes1pd20cs006@pesu.pes.edu.

Pooja Agarwal completed her bachelor's degree in 1998 from CCS University, Meerut, and her master's in computer science in 2001 from Banasthali Vidyapith, Rajasthan, and was awarded the Ph.D. degree in Computer Science primarily focusing in the area of machine learning and deep learning from Visvesvaraya Technological University in 2018. Since July 2008, she has been associated with PES institutions at PESIT Bangalore South Campus (now PES University-EC Campus). She has worked with over a hundred students on various projects at the undergraduate and master levels in the application areas of data science, machine learning, soft computing, natural language processing, data analytics, and deep architectures. She has around 35 publications in various reputed international conferences and journals. She is a member of IEEE and member IAENG. She can be contacted at email: poojaagarwal@pes.edu.