*Journal of Harbin Engineering University* Vol 44 No. 8
*ISSN: 1006-7043* August 2023

# Heterogeneous Bootstrapped Ensemble Model for an Early Assessment of MDD

**Devesh Kumar Upadhyay, Subrajeet Mohapatra, and Niraj Kumar Singh**
Department of Computer Science and Engineering, Birla Institute of Technology, Mesra, Ranchi, India
Corresponding author: Devesh Kumar Upadhyay

**Abstract**

**Background:** Major depressive disorder (MDD) is a psychiatric illness that has far-reaching consequences for mental and physical health, eating and sleeping habits, professional employment activities, and lifestyle. Individualized diagnosis based on the DSM-5 has long been a strategic imperative and is entirely dependent on the clinical psychologist's and psychiatrist's subjective behavioural interpretation.

**Method:** A Machine Learning (ML)-based Heterogeneous bootstrapped ensemble method for an automatic detection approach for detecting MDD based on an individual's behavioural characteristics. In this study, the behavioural data of 503 undergraduates were collected.

**Result:** In this study, we found that the probability of occurrence of MDD in male participants pursuing technical education is higher than in females of the same group. This study found that the Heterogeneous bootstrapped ensemble method performed well with an accuracy of 87.38%, followed by SVM and Bayes Net with accuracy levels of 83.34 and 81.72, respectively.

**Conclusion:** This research revealed that students of technical streams are more likely to develop MDD. Another fact that came to light was that undergraduates from urban areas are more prone to MDD than rural areas. The early detection of MDD among undergraduates is essential for mental health practitioners, and the Heterogeneous bootstrapped ensemble method approach has demonstrated exemplary performance in this regard.

**Index Terms:** Psychological Disorder, Depression, Major depressive disorder (MDD), Cognitive psychology, Support Vector Machine, Confusion matrix, Heterogeneous bootstrapped ensemble.

## 1. Introduction

Major Depressive Disorder (Mdd) Is A Severe Mood Disorder That Affects A Person's Mood And Psychosocial And Occupational Functioning [1]. Depressed Emotions Are A Normal Part Of Life. They Can Arise Due To A Significant Loss Or A Hectic Lifestyle, Leading To Negative Emotions That Last For A Short Period Of Time. However, If These Feelings Last For Weeks Or Even Months, They Are Cause For Concern And Can Cause Distress To The Individual And His Loved Ones. Depression Affects More Than 300 Million People Globally And Is The Leading Cause Of Suicide [2]. Every Year, Over 800,000 People Commit Suicide. Depression Is The Second Largest Cause Of Death Among Those Aged 15 To 29 Years Old, According To The Global Burden Of Diseases Research [3]. Depression Is One Of The Most Common Mental Illnesses In India. It Has Engulfed Practically Every Age Group, Accounting For 12 Percent Of Global Cases, Which Is Expected To Rise To 20% By 2030[4]. Because The Symptoms Of Depression Are Frequently Mistaken With Those Of Other Conditions, Making It Difficult To Diagnose [5], The Concussion Caused By Depression Might Worsen Over Time. If Depression Is Not Treated Promptly, Its Side Effects, Such As Drug And Alcohol Addiction, Sleep Problems Or Insomnia, And So Forth, Can Reduce Hippocampal Volume [6]. In Contrast To Recent Studies, It Has Been Discovered That The Deterioration Of Immunological Aspects Of The Human Body, Such As The Cardiovascular System, Endocrine System, And Central Nervous System (Cns), As A Result Of Prolonged Depression, Can Expose The Body To Metabiological Dysfunction And The Risk Of Infection. Depression Is A Severe Mental Illness, And Researchers Have Used Artificial Intelligence And Mathematical Modeling To Study It. These Studies Uncover The Characteristics Of A Wide Range Of Physiological Diseases. Eeg Research Has Been Done, Which Entails Detecting The Brain's Neural Activities Using Numerous Electrodes That Measure The Low And Peak Value Of Electrical Activity In The Brain.

The Cathodes Detect The Electrical Pulses, And The Instrument Displays Them As Signals, Which May Be Further Examined Using Various Filters And Neural Network Training Approaches. One Of These Technologies' Shortcomings Is The Introduction Of Noise Into The Signal, Which Can Degrade The Original Signal And Lead To Erroneous Results [7]. For Numerous Years, Speech Analysis Has Been Used To Treat Depression. This Method Focuses On Speech's Phonetics, Semantics, And Syntax [8]. When People Are Depressed, Their Speech Qualities Change Dramatically [9]. Compared To Normal People, The Range Of Frequency Of Many Components Of Speech, Such As Loudness, Speaking Pace, Pitch, And Articulation, Is Reduced [10]. Depressed People Exhibit The Aspect That They Converse Without Any Emotions And With Interruptions In Between The Speech [11-14]. This Is Demonstrated By The Fact That Depression Increases The Subject's Unhappiness, And As A Result, The Subject's Capacity To Speak And Communicate Appropriately Deteriorates.

A Combination Of Facial Characteristics And Body Part Movement Analysis Is Another Fascinating Study In Depression Analysis. The Face Analysis System Analyzes Each Expression Using The Facial Action Coding System (Facs), Which Contains A Combination Of Facial Muscles. Aside From Facs, Local And Integrated Feature Analyzers Strive To Identify Expressions By Treating The Face As A Whole. When A Person Is Indifferent To Mood Swings, In Some Type Of Pain, Or Showing Indications Of Agitation, Researchers Have Focused On Head And Hand Parts Movement As Well As The Total Body Movement Cues [15-18]. Our Approach To Depression Analysis Is Considerably Different From The Methodologies Discussed Before. In This Study, The Notion Of Behavioural Characteristics Is Introduced, And A Set Of Questions Encompassing The Various Features Of Depression Were Framed For This Aim. These Questions Were Created And Prepared Under The Supervision Of A Medical Practitioner, And The Participants Were Also Questioned In The Presence Of Certified Staff. This Paper's Ambition Encompasses A Variety Of Elements. This Research Presents A Reliable Strategy For Distinguishing Healthy People From Depressive People. Prolonged Exposure To Depressive Symptoms Can Lead To Severe Mental Illnesses And Can Also Cause Temporary Or Permanent Damage To Brain Cells; Thus, Early Detection Is Critical.

Machine Learning Is One Of The Most Reliable Approaches For Analyzing Such Random Features Because It Is Based On Iterative Algorithm Learning. Following Pattern Recognition Principles, A Typical Machine Learning Algorithm Collects The Unique Features Displayed By The Data Set And Applies Them To Various Categorization Algorithms [19]. The Above-Mentioned Eeg Method Has Drawbacks; It Is A Complex And Sensitive Piece Of Equipment That The Average Individual Cannot Utilize. Eeg Requires A Significant Amount Of Data To Be Processed, Which Is Generated At The Electrodes, And So Takes Time To Complete. Because Eeg Is Founded On The Principles Of Signals And Waveforms, It Is Susceptible To Noise In All Forms. These Sounds Can Affect The Quality Of Captured Data As Well As The Signal Quality. The Features In The Voice Data Recorded By The Device Can Be Affected By The Introduction Of Noise, Which Poses A Problem For Speech Analysis. In The Case Of Signals, It Is A Well-Known Truth That There Will Always Be Some Noise, And So These Errors Must Be Corrected. The Major Disadvantage Of Voice Recognition Is That It Fails When The Patient Refuses To Speak. Motion Sensing-Based Technology Is More Popular, But It Has Its Own Set Of Drawbacks. For Example, Facial Recognition Requires A Camera With A High Megapixel Count And A Relatively Sluggish Recording Rate Per Second. These Devices Are Expensive And Require A Specific Setup. Body Movement Tracking Is Prone To Error Because What Appears To Be Erratic Behavior Can Represent A Person's Normal Habitual Movement.

Based On Our Literature Review, We Found That Various Researchers Used Different Methods To Make Predictions. Even Though The Approaches Mentioned Earlier And Algorithms Perform Well, Each Has Its Own Set Of Limitations. Different Algorithms And Techniques Have Been Discovered To Perform Differently Depending On The Circumstances; No Single Approach Has Been Found To Be Suitable In All Cases. Furthermore, New Algorithms And Methods Are Required For Efficient Prediction Because Of The Importance Of

The Extent Of Improvement In The Prediction Process.

The Significant Contributions Of This Work Are As Follows:

- To Review The Application Of Heterogeneous Bootstrapped Stacked Ensemble Techniques, Especially In Major Depressive Disorder Characterization.
- To Establish A Premise For The Need For Heterogeneous Bootstrapped Stacked Ensemble Techniques In Psychiatry And Psychology Streams Of Medical Science.
- The Main Idea Is To Build The Heterogeneous Stacked Ensemble Classifier To Make The System More Robust And Reliable For Diagnosing Mdd.
- To Determine Whether The Ensemble Technique's Performance Is Better Than Or Equal To The Best Of Its Individual Base Models.
- The Goal Of This Paper Is The Early Assessment Of Mdd Among People And To Distinguish Between Healthy And Depressed People.

The Outline Of This Paper Is As Follows: -Section 2 Contains An Explanation Of The Methodology Used And Various Characteristics Of Mdd. Section 3 Consists Of A Brief Description Of Validation. The Performance Analysis With The Confusion Matrix And Formulas Used Are Explained In Section 4. The Result Analysis Is Described In Section 5. Section 6 Describes The Conclusion Of This Work.

## 2. Methodology

The Details Of The Behavioural Characteristics Used For The Study And The Layout Of The Proposed System Are Described In This Section. Figure 1 Depicts The Flow Chart Of A Proposed Approach For Developing A Machine Learning-Based Mdd Detection System.

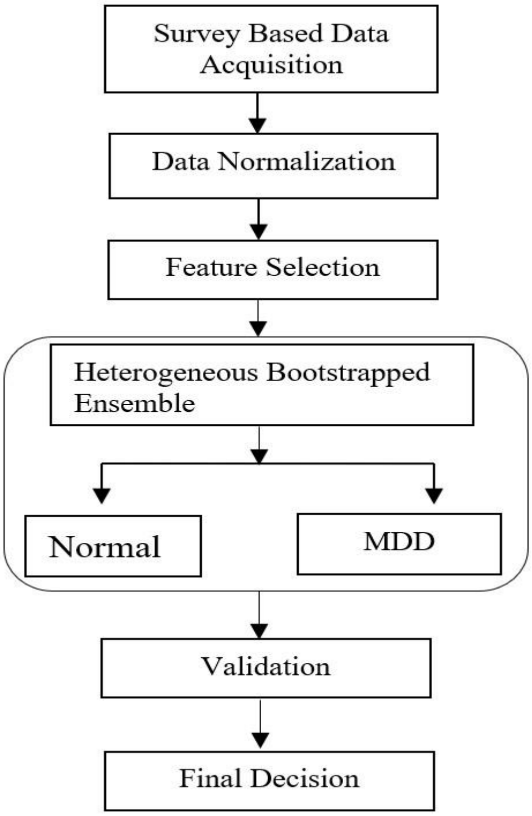


**Figure 1: Flowchart showing the steps involved in the proposed study for the assessment of PDD**

The relevant information was gathered using questionnaires based on the behavioural characteristics of MDD. Questionnaires are used to collect the samples, which are gathered in the presence of medical personnel. The data was collected and fed into an algorithm that involved data shuffling. The data shuffle was performed as a preprocessing step to avoid data monitoring. The data was separated into training and testing after it was preprocessed. Cross-validation is used to evaluate the prediction model at the end of the method.

### *2.1. QUESTIONNAIRES AND CHARACTERISTICS DISCUSSION*

Mood disorder, anxiety, rage, loss of interest in a regular job, hobbies, and other symptoms are all indicators of a brain problem. When present for an extended period of time, such symptoms might lead to depression. The questions in a questionnaire were created under the supervision of a medical practitioner. They were based on the numerous symptoms of depression, such as depressed mood, rage, anxiety, etc. The survey was created to capture every symptom of depression and any odd behavior. Before moving on, the characteristics of depression are described below:

a) Depressed mood: This is a transient state of mind brought on by a significant loss or stress in one's life. In their daily lives, people experience feelings of emptiness, hopelessness, worthlessness, and hatred. If these feelings persist over an extended period of time, they should be taken seriously.

b) Anxiety psychic: When people are anxious, they pay more attention to little details that aren't particularly important.

c) Anxiety somatic: The person experiences bodily symptoms such as perspiration, dry mouth, indigestion, cramps, as well as palpitations, and headaches.

d) Agitation: - In this state, the individual is nervously aroused and occasionally feels excited about no apparent cause. Nail biting, lip biting, hair pulling, twitching fingers, aimless wandering, and other aberrant activities are examples of these experiences.

e) Early Insomnia: In this scenario, the patients cannot fall asleep or may fall asleep after waiting for a long time. Insomnia refers to a lack of or insufficient sleep.

f) Insomnia Middle: In this scenario, people can sleep earlier or at regular times, but their sleep is disrupted in the middle of the night, waking up every 1 to 2 hours or even earlier.

g) Late Insomnia: In this scenario, the person's sleep is not disrupted in the morning of the sleeping cycle. The person usually wakes up early in the morning but then goes back to sleep. It is also possible that if a person gets out of bed, they will be unable to fall asleep again.

h) Insight: Insight tells how much they are aware of their current medical condition or how much they are aware of the sickness but blame others for it or deny being unwell at all.

i) Weight loss: This is a reasonably primary characteristic that can be easily observed by comparing the patient's weight before and after the onset of MDD symptoms.

j) Ideas or acts of self-injury or suicide: - The person suffering from MDD can begin hating themselves or do not want to live, injuring themselves by cutting their hand or even trying to commit suicide.

k) Feelings of guilt and unworthiness: - The persons experience feelings of unworthiness or feelings of guilt, they start blaming themselves for an act they did in the past, and they start thinking that the present disease is a punishment for that.

l) Work and activity: - Individuals suffering from MDD lose interest in everything; their daily routine work, career, and even hobbies are no longer appealing to them. In reality, even general activities no longer deliver enjoyment or pleasure.

m) Somatic symptoms (Gastrointestinal): - The individuals encounter loss in appetite, and sometimes they may not eat as well. On encouragement by others, they may eat, but they feel heavy in the abdomen. The patients may also suffer from constipation.

n) Somatic symptoms: - The sufferers may complain of heaviness in limb, chest, or head. They may also encounter back pain, lethargy, tiredness, and loss of vitality.

o) Genital symptoms: - The genital symptoms are often related to loss of libido or lack of sexual drive and menstruation abnormalities.

p) Psychomotor retardation: - The person's capacity to think, respond, communicate, and grasp is affected. The chief symptoms are slowness in speech and may exhibit absolute stupor.

q) Hypochondriasis: - The patient feels hallucinated about his health and always complains about the health-related difficulties that may not be present in reality. If the patient's demands are not met, the patient may feel irritated.

This study's data is entirely based on the Hamilton Depression Rating Scale (HDRS17) [20]. The responses of the participants ranged from 0 to 4. By accumulating the values, the final scores were calculated. As shown in table 1, the final ratings were categorized according to their severity range.

**Table 1: Showing severity level and scores of MDD**

| Severity Level | Scores |
|---|---|
| Normal | 0-7 |
| Mild | 8-13 |
| Moderate | 14-18 |
| Severe | 19-22 |
| Very Severe/Extreme | Above 23 |

### *2.2. DATA ACQUISITION*

The data for this research work was collected from the leading universities of eastern India. The students are admitted from pan India and abroad countries like Nepal, Bangladesh, Sri-Lanka, and the Middle East countries. For this study, we included only those undergraduates who have had some symptoms of depression in the past six months. Another inclusion criterion was age; students aged between 18 to 25 years were included, and the others were excluded. With the approval of dignitaries from the two leading educational organizations, data was collected using questionnaires based on MDD's behavioral characteristics under the guidance of dignitaries and research scholars from the medical college. The ethics committee approval cum certificate was obtained from the leading universities, and the details of the certificates are *(Protocol/Certificate Reference number: CSE/PHD/2021/12 & date: 14th Dec. 2021)* and *(Certificate Reference number: UMU/R/12-21/279 & date: 30th Dec. 2021)* respectively. Initially, we explained the study's objective and took consent from each individual. For this study, the data of 503 participants were gathered; out of those, 393 were normal persons and 110 people suffering from MDD. The patients were between the ages of 18 and 25. The medical professionals conducted one-on-one interviews with each person, recording their responses to the questions.

In the first stage of the interview, the individuals were asked some straightforward and polite questions to put them at ease.

**Table 2: Showing sociodemographic characteristics of the participants in total, normal and MDD groups.**

| Variables | Total (*n=503*) | Normal (*n=393*) | MDD (*n=110*) |
|---|---|---|---|
| Age | | | |
| Age group (18-20) | 124(24.65%) | 94(23.92%) | 30(27.27%) |
| Age group (21-22) | 141(28.03%) | 108(27.48%) | 33(30%) |
| Age group (23-25) | 238(47.32%) | 191(48.60%) | 47(42.73%) |
| Gender | | | |
| Male | 358(71.17%) | 294(74.80%) | 64(58.18%) |
| Female | 145(28.83%) | 99(25.20%) | 46(41.82%) |
| Place | | | |
| Rural | 143(28.43%) | 109(27.73%) | 34(30.90%) |
| Urban | 360(71.57%) | 284(72.27%) | 76(69.10%) |
| Family Type | | | |
| Nuclear Family | 318(63.22%) | 245(62.34%) | 73(66.36%) |
| Joint Family | 185(36.78%) | 148(37.66%) | 37(33.64%) |
| Financial Status | | | |
| Low (Economic weaker section) | 186(36.98%) | 140(35.62%) | 46(41.82%) |
| Middle Class | 264(52.48%) | 216(54.96%) | 48(43.64%) |

| | | | |
|---|---|---|---|
| Upper Class | 53(10.54%) | 37(7.32%) | 16(14.54%) |
| Education | | | |
| Technical | 211(41.95%) | 168(42.75%) | 43(39.10%) |
| Non-Technical | 298(59.25%) | 225(57.25%) | 67(60.90%) |

### *2.3 DATA PREPROCESSING*

All the data collected were passed into the pre-processing phase, where data shuffling was done. In machine learning, the dataset is divided into training and testing data. Data Shuffling is performed to remove the bias patterns from the dataset before training is committed. Another advantage of shuffling data is that our neural network becomes more generalized, i.e., for every set of input values, there is a set of output values depending upon them. In our study, the shuffling of data is done in python by using a *random.shuffle()* function. This random function algorithm takes the entire data set and randomizes it by exchanging the rows and columns.

### *2.4 FEATURE SELECTION*

We minimize the features since higher dimensionality is never a good option because it is extremely complex and it increases over fitting. Feature selection is the process of choosing the subset of most acceptable and relevant features to be used in model construction can be done automatically or manually. The dataset is used to choose features by either including the crucial features or removing the superfluous features, without altering the dataset's original features.

The SelectKBest method uses a score function that applies parameters to the *(X, Y).* An array of scores is returned by this function. For each feature *X[:,i]* of *X*, scores are calculated. It occasionally additionally returns the p value, which is neither required nor needed. The first *k* features of *X* with the greatest scores are essentially what this feature selection technique returns. Consider passing the ANOVA as a score function. The ANOVA statistic between each characteristic of *X* and *Y*, which are supposed to be class labels, will be computed using this approach. If the computed value is lower, it means that the feature is unrelated to the class label, while a higher value means that the feature is associated to the class label non-randomly. In this manner, we only pick the top k feature

### *2.5 CLASSIFICATION*

Classification algorithms are used to accurately detect the occurrence of MDD with the help of behavioural characteristics, which is advantageous to both physicians and patients. KNN, MLP, BN, RBFN, and SVM are the classification techniques commonly used here. The performances of each algorithm were calculated in terms of accuracy, specificity, sensitivity, precision, and f-score. Classification is the process of placing an object into a specified group or class [21]. A classifier must have some memory to learn the items. The training and testing components of a dataset are separated. The training data sets are utilised to identify the classes or groups that will be worked on and the testing data sets.

#### *2.5.1) K-NEAREST NEIGHBOR (KNN)*

KNN is a non-parametric classifier that divides classes into groups based on related patterns. KNN does not make any assumptions about the data sets and relies only on neighbourly majority voting. This approach uses the Euclidean distance to compute the distance between the two locations. Our research chose five neighbours *(k=5)* and applied Euclidean distance *(p = 2)* to calculate distance. Because of its simplicity and adequacy, KNN [22] is frequently used to categorise future data.

$$P(y = j|X = x) = \frac{1}{K}\sum_{i} = {}_{A} I(y^i = j) \quad (1)$$

#### *2.5.2) SUPPORT VECTOR MACHINE (SVM)*

The SVM classifier is based on the separating hyperplane idea [23]. Each class is on the opposite side of the hyperplane, which splits the space into two sections. The hyperplane mapping is done in such a way that the distance between the two classes is maximum. The SVM's primary goal is to create an ideal hyperplane that can lower the error function.

##### *2.5.2.1 LINEAR SVC*

A Linear SVC (Support Vector Classifier) is designed to fit to the data you provide and provide a "best fit" hyperplane that divides or categorizes your data. Following that, you may input some features to your classifier to check what the "predicted" class is after you've obtained the hyperplane. This makes this algorithm particularly ideal for our purposes, however it can be used in a variety of circumstances. The general equation of linear SVC can be written as:

$$\min_{\omega,b} \frac{1}{2}\omega^T\omega + C\sum_{i=1} max\left(0,1 - y_i(\omega^T \emptyset(x_i) + b)\right), \quad (2)$$

***2.5.2.2 SVM-GAUSSIAN RADIAL BASIS FUNCTION (RBF)***

It is one of the most commonly used kernel functions in SVM. It is generally utilized with non-linear data. It assists in proper separation when there is no prior knowledge of data.

*F($x_i$, xj) = exp(-gamma * ||$x_i$ - xj||^2)* (3)

The gamma value might be anywhere between 0 and 1. The gamma value must be entered manually into the code. The gamma value of 0.1 is the best option.

***2.5.3) MULTILAYER PERCEPTRON (MLP)***

The MLP classifier contains three layers: the input layer, the second, which is the hidden layer, and the third, which is the output layer [24]. The non-linear activation function, which may be implemented except for polynomial functions, is the MLP classifier's major strength. MLP learning is an iterative process in which the weights are modified after each iteration based on the difference in error between the input and output values. The process is repeated until an optimal result is achieved.

***2.5.4) RADIAL BASIS NEURAL NETWORK (RBFN)***

The RBFN is a two-layered network with a radial basis function as the activation function in each hidden layer. In RBFN, the presented input values are non-linear, while the output values are linear. The output layer is a weighted sum of the outputs of the hidden layers [25]. This classifier's primary goal is to predict the desired value using various combinations of radial kernels, also known as Gaussian functions, as the activation function.

The key benefit of this classifier is that it can be used with any input data set by using data points as centers and approximation algorithms.

The kernel is given by:

$$F(x_i, x_j) = exp(- d(x_i,x_j)^2 /2l^2) \quad (4)$$

where l is the length scale of the kernel and $d(x_i,x_j)$ is the Euclidean distance.

***2.5.5) BAYESIAN NETWORK (BN)***

The Bayes network classifier is a probabilistic graphical model that uses directed acyclic networks to describe a set of variables and their conditional relationships. The Bayes network classifier is used to depict probability correlations between illness and symptoms. It is most commonly used to forecast various diseases using symptoms [26].

***2.5.6) HETEROGENEOUS BOOTSTRAPPED ENSEMBLE (HBE) CLASSIFIER***

The HBE classification approach enhances classification performance by increasing the variety and complementarity of base predictors by mixing a group of different classifiers. The basis classifiers are usually combined using a weighted technique, particularly the simple average method, with stacking-based base classifiers combination methods proven to enhance significantly heterogeneous classification ensemble models' performances [27]. The two steps of the stacking-based base classifiers combination approach are building a collection of different base classifiers and feeding the first stage predictions into a meta-classifier to generate the final predictions. In our research, we had considered KNN, MLP, SVM, RBFN, and Bayesian Network as base classifiers. and SVM with RBF kernel as a meta-classifier in this experiment. The purpose of a stacked ensemble classifier is to generate an ensemble model, which is a collection of SVM kernels with a decision function that looks like this:

$$f(x) = sgn \{ \sum_{i=1}^{n} \alpha_i^*. y_i \sum_d \beta_d \, k_d(x_i, x_j) + b^*\} \quad (5)$$

Ensemble classifiers are used to solve classification problems where a fusion of classifiers or a mixture of kernels from a single classifier leads to a final result.

Figure 2 shows the architecture of a HBE method with an input dataset of n records. A collection of new sub-databases is created by randomly collecting m samples with a replacement where *m<n*. Bootstrapping involves choosing m numbers of data at random and replacing them with data from the dataset. Bootstrapped samples refer to the new input dataset provided to the basic classifiers. In the first phase, the each base learner classifier is trained with the bootstrapped sample database, and in the second phase, the test data from the dataset *D* is given as an input to each base classifier.

Each base classifier's output is stacked and used as an input feature for the meta-classifier, which generates the final prediction result. The heterogeneous stacked model makes the ensemble classifier more efficient and durable, which delivers a substantially higher result.

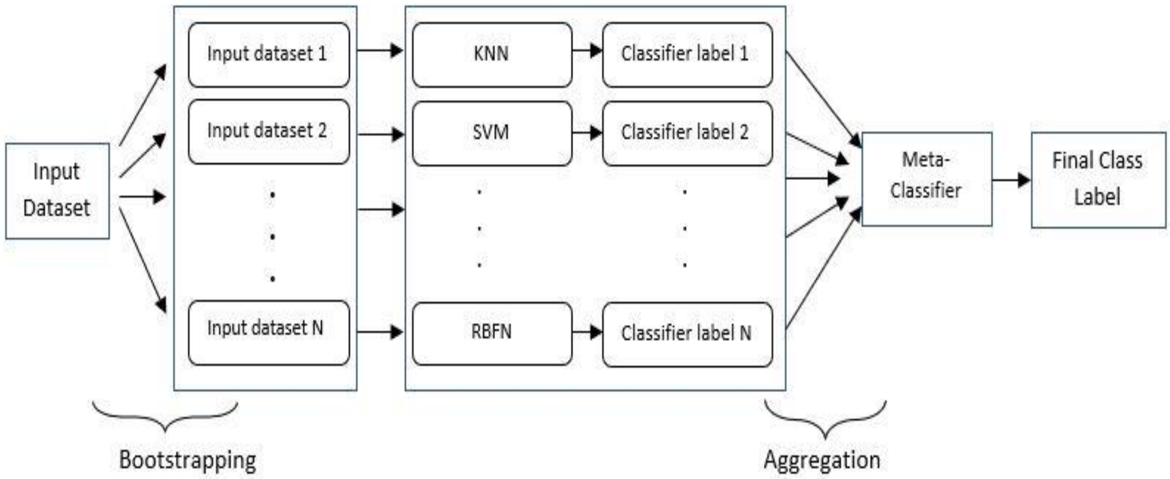


**Figure 2: Showing architecture of the Heterogeneous bootstrapped ensemble model**

The proposed algorithm 1 describes the heterogeneous bootstrapped ensemble method in which Dataset *D = {(xi, yi), i=1,…,m},* split into *j* samples *(D1$_j$),* is provided as an input to the method. These samples are then distributed to n base classifiers with replacement. The process begins by creating a random number between *1* and *j*, represented by *k,* called *genrateRandom (1, j).* The classifier *Ci* will be trained with each record of the data sample *dk,* and *hi* will hold the prediction result for the same if the *k$^{th}$* data sample is present in *D'*. For creating an input dataset for the meta-classifier, a set of the predicted and initial input labels is created, designated z, and is unionid to *D''*. The classification outcome from the meta-classifier is represented by *h'* at the end. In our dataset *D*, where I has a value between *1* and *m, xi* is a vector that encodes the attribute values of the *i*$^{th}$ input instance, and yi is its corresponding feature label.

*Proposed algorithm:*

**Input:** - Initialize:

Training dataset D = {($x_i$, $y_i$), i=1,…,n}

% $x_i$ is a vector representing the attribute values of the $i^{th}$ input instance and $y_i$ is the associated feature label. %

Base classifier $C_1$, $C_2$, …, $C_k$;

Meta classifier C;

**Process:**

for t = 1 to *k* do

{

$h_t$ = $C_t$(D) //Train each base level learner $h_t$ by applying the base-level classifier $C_t$ to the original dataset D

}

D$'$= Ø; // Generate a new dataset

for t = 1, 2, …. *n* do

{

for t=1,2,… k do

{

$Z_{it}$ = $h_t(x_i)$ // use $h_t$ to classify the training example $x_i$

}

D$'$=D$'$U{(($z_{i1}$, $z_{i2}$,….$z_{iT}$),$y_i$)}

}

h$'$=C(D$'$) // Train a meta level learner h$'$ by applying the second level learning classifier C to the new dataset D$'$

**Output:**

*Finally, from this we can conclude that the output will be*

H(x)= h$'$($h_1$(x), $h_2$(x),…..$h_k$(x))

### 3. Validation

When a classifier is trained on a particular dataset, it will not attain accuracy when applied to different datasets; hence cross-validation is introduced to solve this problem. Cross-validation does not improve classifier accuracy, but it does improve their stability and generalisation. The *K-fold* cross-validation approach is chosen because it is basic, easy, and uses the entire dataset for training and validation [28]. For each dataset, this method entailed building a K-fold partition and executing it for K times to determine accuracy for each consecutive run.

## 4. Performance Analysis

The performance analysis is carried out using machine learning methods to accurately examine the classifiers' ability to predict the outcome. The confusion matrix is used to calculate accuracy, sensitivity, and specificity in the performance analysis, as shown in table 3.

**Table 3: Showing Confusion matrix for performance assessment**

| Classifier Output | Psychiatrist Opinion | |
|---|---|---|
| | Positive (MDD) | Negative (Normal) |
| Positive (MDD) | TP (True Positive): Correctly identified | FP (False Positive): Correctly rejected |
| Negative (Normal) | FN (False Negative): Incorrectly rejected | TN (True Negative): Incorrectly identified |

The formulas that govern the confusion matrices for performance evaluation are discussed below:

$$\textit{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \text{ x } 100 \quad (6)$$

$$\textit{Specificity} = \frac{TN}{TN+FP} \text{ x } 100 \quad (7)$$

$$\textit{Sensitivity} = \frac{TP}{TP+FN} \text{ x } 100 \quad (8)$$

$$\textit{Precision} = \frac{TP}{TP+FP}$$

(9)

$$\textit{F Measure} = 2\,x\,\frac{(Precision\ x\ Sensitivity)}{(Precision + Sensitivity)}$$

(10)

## 5. Experiments and Result Analysis

### *5.1 STATISTICAL ANALYSIS*

In this paper, SPSS software (Producer: IBM, Version: SPSS25, Origin: Chicago, IL USA) was used to examine the data. All the continuous variables are converted into categorical variables like *Age*, divided into three groups for bringing uniformity in data such that one-way ANOVA can be applied. The descriptive table (Table 1) in section 2 shows that categorical variables were reported as numbers and percentages. The significance level considered is 5% ($\alpha$ = 0.05) as per the standard practice. Initially, the data was assumed to be distributed normally which then further verified through normality test as shown in table 4. Although the *skewness* and *kurtosis* values in a normal distribution are both zero, *skewness* and *kurtosis* values between *-2* and *+2* are acceptable for psychometric applications [29]. In the result analysis, we found that almost all the characteristics variables lie within the limit except some critical behavioral characteristics variables like *suicidal tendencies*, *retardation*, and *Hypochondriasis*.

**Table 4: Descriptive Statistics demonstrating Normality test of the study subjects**

| **Descriptive Statistics** | | | | | |
|---|---|---|---|---|---|
| | **Mean ± Std. Deviation** | **Skewness** | | **Kurtosis** | |
| **Variables** | **Statistic** | **Statistic** | **Std. Error** | **Statistic** | **Std. Error** |
| Depressed Mood | 0.939 ± 0.89 | 1.117 | 0.109 | 1.258 | 0.217 |
| Feelings Of Guilt | 1.123 ± 0.71 | 1.585 | 0.109 | 1.567 | 0.217 |
| Suicidal Thoughts | 0.649 ± 0.2 | 3.773 | 0.109 | 4.879 | 0.217 |
| Insomnia - Initial | 0.696 ± 0.57 | 0.814 | 0.109 | -0.567 | 0.217 |
| Insomnia – Middle | 0.619 ± 0.4 | 1.313 | 0.109 | 0.607 | 0.217 |
| Insomnia - Delayed | 0.632 ± 0.42 | 1.224 | 0.109 | 0.358 | 0.217 |
| Work And Interests | 1.223 ± 1.18 | 0.386 | 0.109 | -1.265 | 0.217 |
| Retardation | 0.95 ± 0.57 | 1.864 | 0.109 | 3.111 | 0.217 |
| Agitation | 0.618 ± 0.5 | 0.827 | 0.109 | -0.32 | 0.217 |
| Anxiety - Psychic | 0.919 ± 0.62 | 1.493 | 0.109 | 1.834 | 0.217 |

| | | | | | |
|---|---|---|---|---|---|
| Anxiety Somatic | 0.82 ± 0.53 | 1.371 | 0.109 | 0.962 | 0.217 |
| Somatic Symptoms | 0.512 ± 0.28 | 1.632 | 0.109 | 1.781 | 0.217 |
| General Symptoms | 0.528 ± 0.35 | 1.171 | 0.109 | 0.343 | 0.217 |
| General Somatic Symptoms | 0.448 ± 0.19 | 2.248 | 0.109 | 4.434 | 0.217 |
| Hypochondriasis | 0.757 ± 0.28 | 3.485 | 0.109 | 3.116 | 0.217 |
| Weight Loss | 0.56 ± 0.34 | 1.389 | 0.109 | 0.961 | 0.217 |
| Insight | 0.573 ± 0.46 | 0.797 | 0.109 | -0.365 | 0.217 |

For analyzing the impacts of the characteristics variables among different groups, we applied one-way ANOVA based on various *age groups*, as shown in Table 5. From the results of table 5, we found that only eight characteristics variables have their levels of significance below the considered threshold. The characteristics variables like *anxiety psychic* and *Insight* have *p-values* of nearly zero. When we considered *gender* for the detailed statistical analysis, we found that the probability of MDD occurring among females is 31.72%, much higher than that of males (17.88%). Females from the age group *(23-25)* years are more prone to MDD. Around 50% of females suffering from MDD have somatic symptoms of anxiety; 52.17% of female subjects are from the age group *(23-25)* years.

**Table 5: One-way ANOVA analysis of characteristics variables**

| ANOVA | | | | | |
|---|---|---|---|---|---|
| Characteristic Variables | Sum of Squares | df | Mean Square | F | p-values |
| DEPRESSED MOOD | 0.884 | 2 | 0.442 | 0.500 | 0.607 |
| FEELINGS OF GUILT | 0.880 | 2 | 0.440 | 0.348 | 0.706 |
| SUICIDAL THOUGHTS | 0.087 | 2 | 0.043 | 0.102 | 0.903 |
| INSOMNIA - Initial | 1.991 | 2 | 0.995 | 2.063 | 0.128 |
| INSOMNIA - Middle | 0.847 | 2 | 0.424 | 1.106 | 0.332 |
| INSOMNIA - Delayed | 3.211 | 2 | 1.606 | 4.066 | 0.018 |
| WORK AND INTERESTS | 5.205 | 2 | 2.602 | 1.744 | 0.176 |
| RETARDATION | 5.249 | 2 | 2.625 | 2.931 | 0.054 |
| AGITATION | 1.048 | 2 | 0.524 | 1.373 | 0.254 |
| ANXIETY - PSYCHIC | 13.558 | 2 | 6.779 | 8.258 | 0.001 |
| ANXIETY SOMATIC | 6.088 | 2 | 3.044 | 4.597 | 0.011 |
| SOMATIC SYMPTOMS | 1.889 | 2 | 0.945 | 3.645 | 0.027 |
| GENERAL SYMPTOMS | 0.198 | 2 | 0.099 | 0.354 | 0.702 |
| GENERAL SOMATIC SYMPTOMS | 1.238 | 2 | 0.619 | 3.106 | 0.046 |
| HYPOCHONDRIASIS | 5.547 | 2 | 2.774 | 4.911 | 0.008 |
| WEIGHT LOSS | 0.242 | 2 | 0.121 | 0.385 | 0.680 |
| INSIGHT | 8.229 | 2 | 4.115 | 13.130 | 0.001 |

We also considered *gender* for the statistical analysis in which we found eight significant variables. Variables like *depressed mood*, *feelings of guilt*, *work and interest*, and *agitation* have much more impact as their *p-value* are close to zero in comparison to other variables. Interestingly, 63.04% of undergraduate girls have *feelings of guilt* about their past activities, which is one of the main reasons for the occurrence of MDD among females. At the same time, suicidal tendencies are found prominently among males compared to females of the same group. But we can't consider it for further detailed analysis due to the presence of only two groups. When *financial status* is considered, we found only five variables with p-value less than the threshold set. Those five variables also show their significance between the groups when considering *Age* as a factor.

**Table 6: Results of Games-Howell Post-Hoc analysis showing multiple comparisons among the age groups**

| Multiple Comparisons | | | | | | | |
|---|---|---|---|---|---|---|---|
| Games-Howell Post-hoc test | | | | | | | |
| Characteristic Variables | Age groups | | Mean Difference (I-J) | Std. Error | p-values | 95% Confidence Interval | |
| | | | | | | Lower Bound | Upper Bound |
| INSOMNIA - Delayed | 21-22 | 23-25 | .190* | 0.068 | 0.015 | 0.03 | 0.35 |
| RETARDATION | 21-22 | 23-25 | 0.243 | 0.106 | 0.048 | -0.01 | 0.49 |
| ANXIETY - PSYCHIC | 18-20 | 21-22 | .428* | 0.12 | 0.001 | 0.15 | 0.71 |
| | | 23-25 | .336* | 0.117 | 0.012 | 0.06 | 0.61 |
| ANXIETY SOMATIC | 18-20 | 23-25 | .273* | 0.094 | 0.012 | 0.05 | 0.5 |
| SOMATIC SYMPTOMS | 21-22 | 23-25 | .145* | 0.059 | 0.039 | 0.01 | 0.28 |
| HYPOCHONDRIASIS | 21-22 | 23-25 | .225* | 0.082 | 0.019 | 0.03 | 0.42 |
| INSIGHT | 18-20 | 21-22 | .186* | 0.072 | 0.028 | 0.02 | 0.36 |
| | | 23-25 | .317* | 0.061 | 0.001 | 0.17 | 0.46 |
| **. The mean difference is significant at the 0.05 level.* | | | | | | | |

For further detailed statistical analysis, we chose to perform the *Games Howell post-hoc test*, shown in the table [6]. For the *post-hoc test*, we had only considered the variables with the calculated *p*-values less than the desired threshold. In multiple comparisons between the age groups, we found that there are significant differences (*p=0.015*) between the age group (*21-22*) and the age group (*23-25*) for the case of *Insomnia delayed*. When we considered the *anxiety psychic* characteristic variable, we observed that there are significant differences between the age group (*18-20*) and age group (*21-22*) (*p=0.001*) & for age group *(18-20)* and age group *(23-25)(p=0.012)* and similar differences are noticed for characteristic variable *anxiety somatic*. From the table [6], we can notice that the variable *Insight* has much more impact on the age groups (18-20) and (23-25), where the *p*-values are approaching zero. However, it has a tremendous effect on all the groups. Another remarkable impact can be observed for the variable *hypochondriasis*, where the age groups *(21-22)* and *(23-25),* with *p=0.019*, exhibit significant difference.

The consistency of a measurement method's or test results over time is referred to as reliability [30]. Reliability is the propensity for consistency in repeated measurements of the same phenomenon. The reliability coefficient *Cronbach's alpha* of the *HDRS* scale range between *0.65* and *0.98*; for this study, *Cronbach's alpha* calculated value is *0.855* [31]. This finding suggests that the rating scale employed in the study is trustworthy.

### *5.2. MACHINE LEARNING ANALYSIS*

In this research paper, python is used for carrying out processing and computational tasks. It is installed on Windows™ 10 Pro, Intel® Core™ i7-8550U CPU with RAM storage of 16 GB. All of the rigorous experiments, from data preparation through classifier training and validation, are carried out on the same hardware and software platform. The proposed classifier's performance is evaluated in terms of accuracy, sensitivity, specificity, precision, and f-score. We had used k-fold cross-validation with *k = 5*, and the data is partitioned into training and testing datasets. The stacked ensemble classifier is trained on the training subset before being evaluated on the testing dataset. All of the experiments are run on the set-up mentioned earlier, with classifiers at two layers of the HBE model.

The desired result was achieved by following the methods stated in the flow chart, as shown in figure 1. The first step was to design surveys based on the behavioural characteristics of MDD patients.

Individual responses were recorded under the supervision of medical specialists; see section II for more information (A). After feature selection, the discriminant features are determined using the statistical feature selection technique [32]. For feature categorization, 17 significant characteristics were determined using ANOVA. The relationship between various features and their accompanying *p-values*, which reflect the significance of the features, is shown in Figure 3.

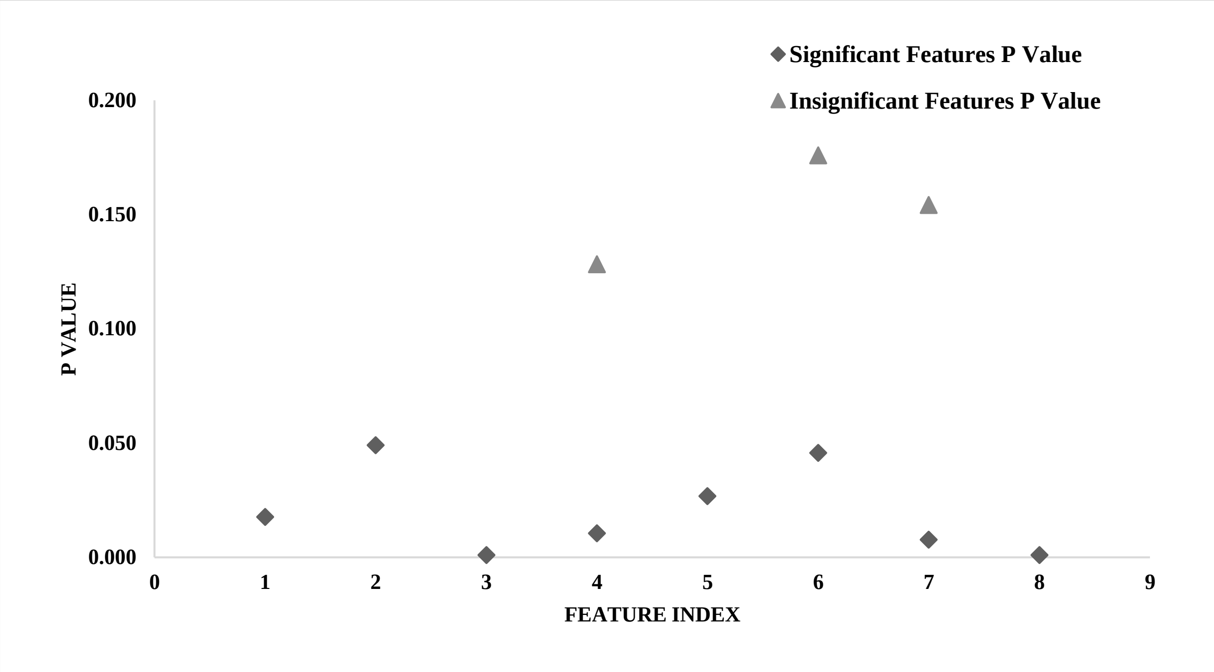


**Figure 3: Scatter chart showing plot between feature index and p-value for showing significant features ($p<0.05$) and non-significant features($p>0.05$).**

In Section II, we'll go over data preprocessing (C). K-Fold validation was performed for all of the classifiers KNN, SVM RBFN, MLP, Bayes Net and HBE Model.

**Table 7: Comparing the accuracy of HBE classifier with KNN, MLP, SVM, RBFN and Bayes Net classifier**

| Classifiers | Fold | | | | | Average Accuracy |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| KNN | 72.8 | 72.5 | 72.8 | 69.8 | 73.4 | 72.26 |
| MLP | 76.6 | 73.4 | 74.5 | 76.3 | 75.8 | 75.32 |
| SVM | 83.4 | 82.1 | 81.9 | 84.5 | 84.8 | 83.34 |
| RBFN | 78.4 | 78.2 | 78.8 | 80.1 | 79.2 | 78.94 |
| Bayes Net | 81.7 | 82.3 | 81.8 | 81.2 | 81.6 | 81.72 |
| HBE | 86.7 | 86.9 | 88.2 | 87.8 | 87.3 | 87.38 |

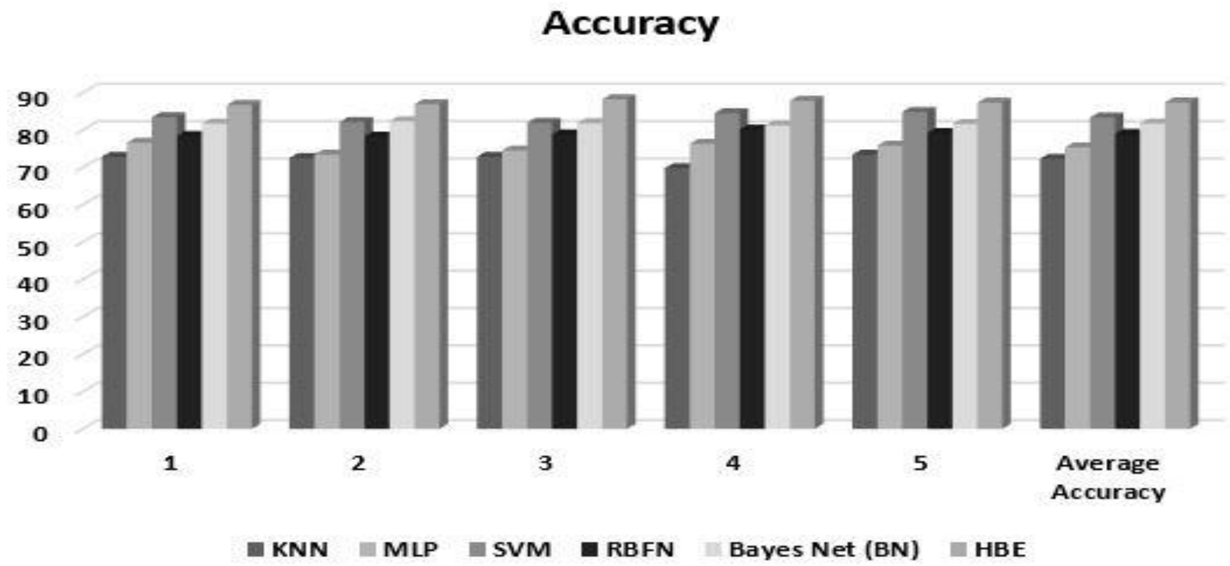


**Figure 4: Showing accuracy of Stacked HBE classifier along with standard classifiers using 5–fold.**

The accuracy prediction of several classifiers is shown in Table 4 and Figure 4, with the HBE classifier leading with an accuracy of 87.38 percent and the SVM classifier behind with an accuracy of 83.34 percent.

The specificity rate of all the classifiers is shown in Figure 5. The HBE classifier has the highest specificity level of 86.24 percent, followed by SVM and Bayes Net with a specificity level of 82.88 percent and 80.64 percent. Figure 5 compares all the classifiers and shows that the HBE algorithm surpasses them all in every aspect.

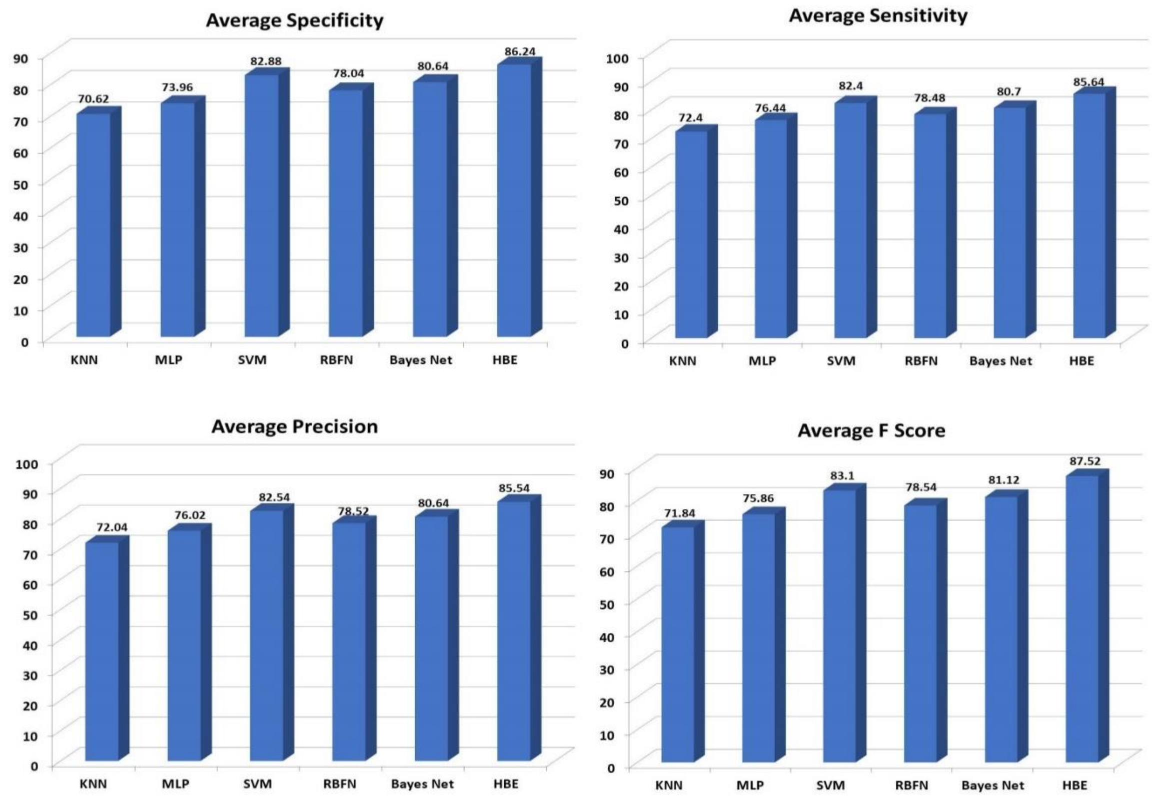


**Figure 5: Graphical representation demonstrating the performance of HBE classifier in terms of Specificity, Sensitivity, Precision and F-score along with KNN, MLP, SVM, RBFN and Bayes Net classifier**

Sensitivity is the percentage of negatives accurately identified by the classifier. Figure 5 demonstrate that the average sensitivity of the HBE classifier is 85.64 percent, which is much greater than the sensitivity of the RBFN classifier, which is 78.48 percent after a 5-fold operation. The rate of precision is a measure of how often a classifier predicts correctly when it is actually correct, or simply a percentage of all positives correctly detected by the classifiers. Figure 6 displays the precision rates of the classifiers, with HBE having the best performance with 85.54 percent, whereas KNN have least performance with 72.04 percent, as seen visually in figure above. The F1-Measure or F1-Score is a popular way to calculate test performance in statistical analysis. The precision and recall values are used to calculate the F1-score, with precision equaling the number of correctly identified true positive (TP) values divided by the total number of positive values and recall equaling the number of correctly identified true positive (TP) values divided by the actual number of positive values. From Figure 5, we can easily conclude that the f-score of HBE is performing best with 87.52 percent.

In addition, we used the AUC (Area under the Receiver Operating Characteristic Curve) curve to assess the model further. The AUC-ROC method is the most commonly used method for evaluating classification models. The AUC represents the separability measure, and the ROC represents the model's probability curve. When the AUC is high, the model has a high level of classification accuracy. The False Positive Rate (FPR) on the x-axis and True Positive Rate (TPR) on the y-axis are plotted on the AUC-ROC curve.

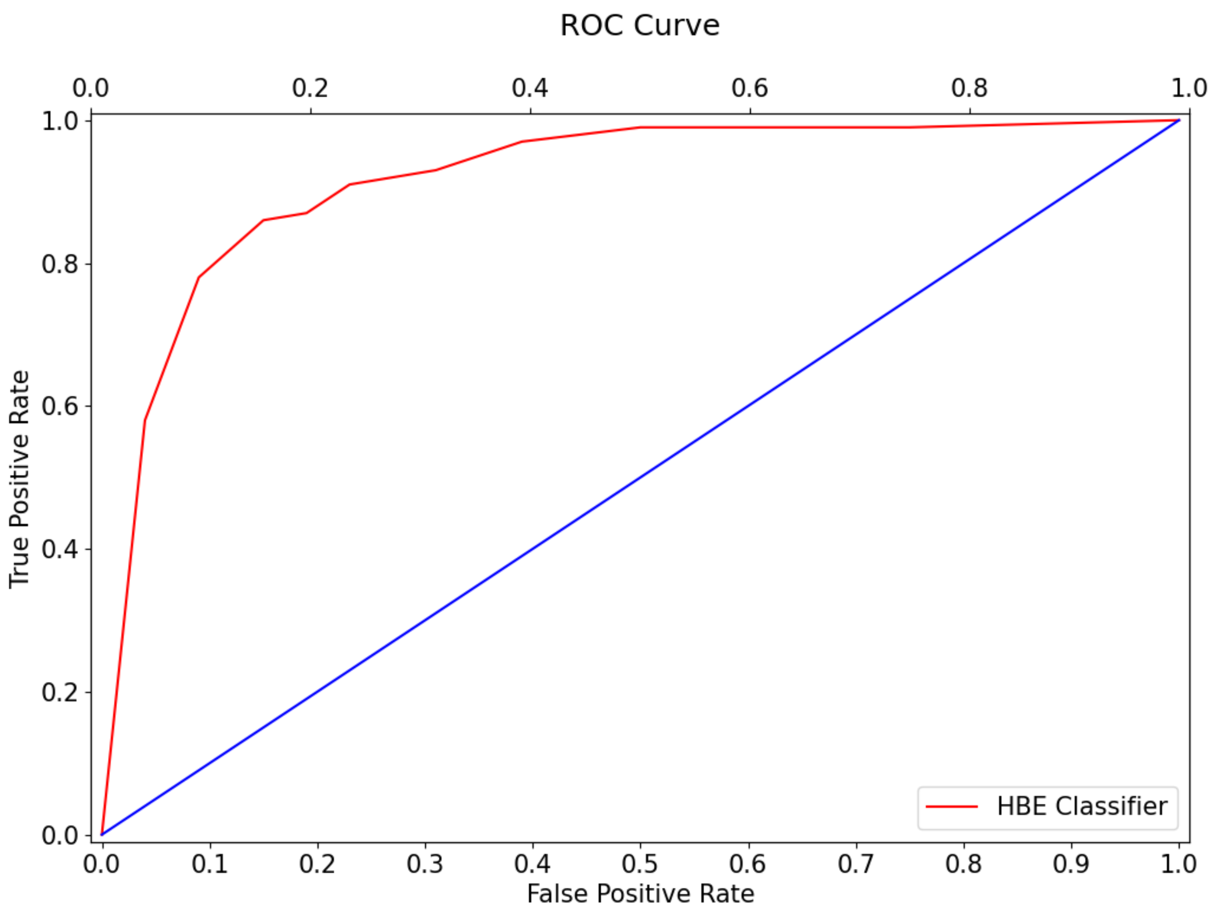


**Figure 6: ROC-AUC CURVE for HBE**

The computation time of any classifier is the time taken by the classifier to complete the classification process. Figure 7 represents the computation time for all the classifiers. From the below data, we can conclude that the SVM is taking minimum time among all these classifiers, followed by the Bayes net classifier. Although the RBFN performance is good in terms of accuracy, it consumes more time.

The computation time depends on your computer's processor and may vary from system to system.

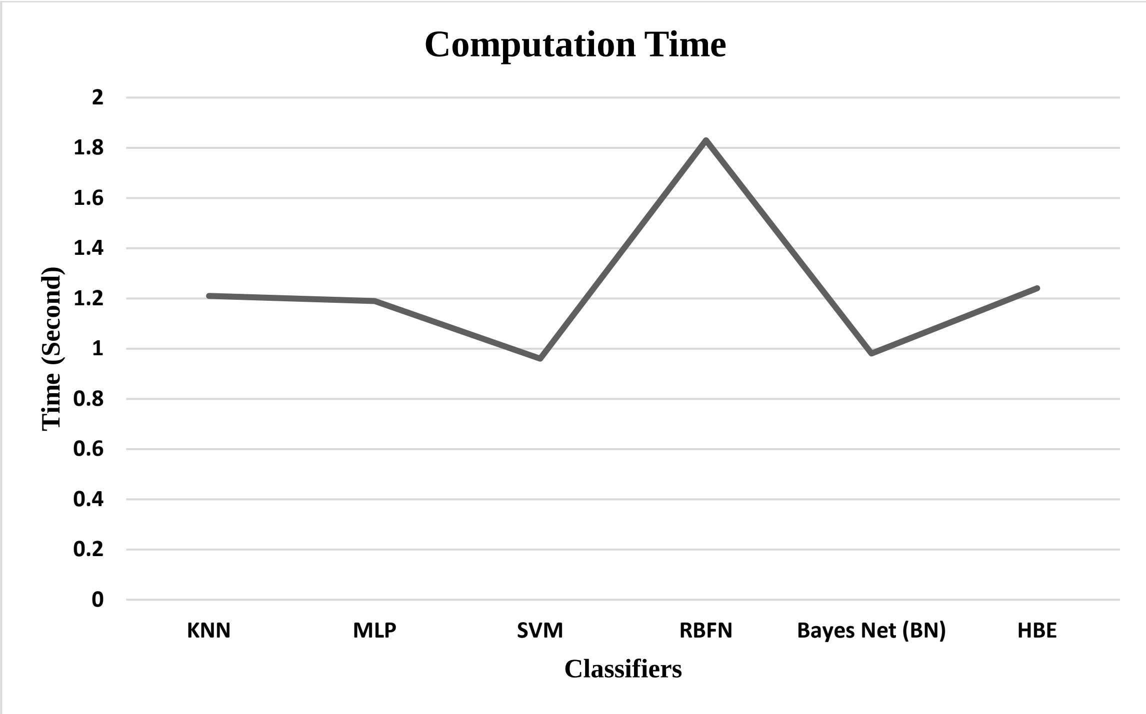


**Figure 7: Showing average computational time of KNN, MLP, SVM, RBFN and Bayes Net classifier**

In any event, the above result supports the idea that HBE classifier systems outperform SVM classifier, Bayes Net and other examined classifiers like as RBFN, KNN and MLP.

***5.3. COMPARISON WITH EXISTING APPROACHES***

In this section, we compared our proposed methodology to several existing approaches. And for comparing with various techniques, we used the depression, anxiety and stress dataset, which have 21 features. The depression dataset includes Depression related behavioural features. When applied to the depression dataset, our proposed technique has an accuracy of 87.9%. The suggested model is compared to current models in Table 8.

**Table 8: Comparison with existing similar approaches**

| **Models used** | **Accuracy** |
|---|---|
| Choudhury et al. (2019) [33] | 75% |
| Jalandra et al. (2022) [34] | 80 % |
| Proposed work [Heterogeneous Bootstrapped Ensemble] | 87.9% |

In an analysis of previous methods, we can easily conclude that our new proposed model outperforms the previously proposed ones in terms of accuracy, sensitivity, specificity, precision, and f-score due to its diversity and richness.

## 6. Conclusion

MDD is one of the most widespread psychological disorders in the world presently. The incarnation of modern life has had numerous detrimental consequences for people of all ages. Due to the rise in responsibilities, work pressure, and stressed work culture, the human brain has been subjected to pressure even at a young age. All of these factors have a negative impact on people and can lead to serious mental illnesses such as MDD, which has resulted in an exponential growth in MDD cases. Early identification of MDD is critical since long-term exposure to MDD can result in serious mental illness.

In data analysis, we found that females from higher-income families are comparatively more prone to MDD than other groups. The probability of MDD is much higher among females who come from a nuclear family and have low incomes. An in-depth analysis, we found that the undergraduates from the age group *(23-25)* have an intense feeling of loss of interest in work and activity, including their day-to-day tasks and college activities, and even lost practicing their hobbies. Our research also found that anxiety and insight affect males and females of all age groups and significantly impact their lifestyles. Interestingly, the behavioral factor of suicidal tendency is observed prominently among male students.

Various machine learning models were utilized in this study to create a computer-assisted approach for MDD prediction. For this purpose, multiple classifiers were clubbed together in such a way that they should predict more efficiently. Due to its accurate values and diversity, HBE classifier systems are becoming more prominent in medical diagnosis and healthcare. The proposed method is considerably affordable and simple as compared to previous detection methods. The goal of this study was to develop an algorithm that could be utilized to assess MDD more accurately, specifically, and sensitivity while also requiring less time. According to the above data, the Heterogeneous Stacked ensemble performed better than the other classifiers analyzed in this research article, with an accuracy level of 87.38 percent.

### *6.1 Limitations and future work:*

This study's emphasis on the HDRS questionnaire as the gold standard for identifying depression is most likely a research limitation. The study's other limitation is that the respondent is its only data source; thus, if they choose not to respond, the study would lag. This study lacks follow-up data, which could be added to future work. As a survey-based study, participant bias could not be removed entirely. Studies with higher sample sizes and longitudinal designs can be planned to get around these restrictions.

The findings of this study have opened up many new areas for future research. Future research will focus on developing more efficient ensemble approaches to overcome the limitations of this work. The Heterogeneous ensemble model will also be used to detect various types of mental illnesses. Furthermore, the system can be improved by introducing new adjustments. This method could be standardised, and it would undoubtedly aid doctors and others in diagnosing MDD.

**Acknowledgment**

All of the authors would like to thank everyone who took the time to complete the questionnaire and participate in this study. The study would not have been successful if the subjects had not given their consent. Dr. Ajay Kumar Bakhla, Rajendra Institute of Medical Sciences Ranchi, India, Dr. Varun S. Mehta, associate professor at Central Institute of Psychiatry, and Madhu Gupta, research scholar at Ranchi Institute of Neuro-Psychiatry & Allied Sciences, Ranchi, are thanked for their help in completing this work.

**Funding Source**

For this research we have not received any fund from any agency.

**Ethics Committee Approval:**

The ethics committee approval cum certificate was obtained from the leading universities, and the details of the certificates are (University name: Birla Institute of Technology, Mesra, Ranchi, India and Protocol/Certificate Reference number: BIT/CSE/PHD/2021/12 & date: 14th Dec. 2021) and (University name: Usha Martin University, Ranchi India and Certificate Reference number: UMU/R/12-21/279 & date: 30th Dec. 2021) respectively.

**Informed Consent:**

Written informed consent was obtained from all participants who participated in this study.

**Declaration Of Interests**

The authors have no conflicts of interest to declare.

**References**

[1] Angst F, Stassen HH, Clayton PJ, et al. Mortality of patients with mood disorders: follow-up over 34–38 years. J Affect Disord 2002; 68:167–181

[2] Depression in India let's talk, World Health organization (WHO): World Health Day (April 2017).

[3] Kapil Joshi, Himanshu Gupta, Prashant Chaudhary and Punit Sharma. "Survey on different enhanced k-means clustering algorithm", in (IJETT), vol.27, no. 4, pp.178–182, Sept 2015.

[4] World Population Ageing 2019 [online]. Available: https://www.un.org/en/development/desa/population/ publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf [Accessed: 1-Sep-2020]

[5] Sharifa Alghowinem. "Multimodal Analysis of Verbal and Nonverbal Behaviour on the Example of Clinical Depression". PhD thesis, The Australian National University, May 2015.

[6] Y. I. Sheline, M. H. Gado and H. C. Kraemer. "Untreated depression and hippocampal volume loss", American Journal of Psychiatry, vol. 160(8), pp.1516–1518, Aug 2003.

[7] Behshad Hosseinifard, Mohammad Hassan Moradi, and Reza Rostami. "Classifying depression patients and normal subjects using machine learning techniques", In 19th Iranian Conference on Electrical Engineering, 2011.

[8] Michelle Renee Morales and Rivka Levitan. "Speech vs. text: A comparative analysis of features for depression detection systems", in SLT, pp. 136–143, Dec 2016.

[9] Jyoti Joshi. "An automated framework for depression analysis" In Humaine Association Conference on Affective Computing and Intelligent Interaction, pp. 630–635, Sept 2013.

[10] J. C. Mundt, P. J. Snyder, M. S. Cannizzaro, K. Chappie and D. S. Geralts. "Voice acoustic measures of depression severity and treatment response collected via interactive voice response technology" Journal of Neurolinguistics, vol.20 (1),pp.50–64, 2007.

[11] K. R. Scherer."Vocal assessment of affective disorders" Lawrence Erlbaum Associates, 1987.

[12] S. Alghowinem, R. Goecke, M.Wagner, J. Epps, M. Breakspear, and G. Parker."From joyous to clinically depressed: Mood detection using spontaneous speech" In Proceedings of the 25th International Florida Artificial Intelligence Research Society Conference, pp.141–146, 2012.

[13] A. Ozdas, R. Shiavi, S. Silverman, M. Silverman and D. Wilkes. "Analysis of fundamental frequency for near term suicidal risk assessment", In IEEE International Conference on Systems, Man, and Cybernetics, Aug, 2000.

[14] E. Moore, M. Clements, J. Peifer and L. Weisser."Critical analysis of the impact of glottal features in the classification of clinical depression in speech", IEEE Transactions on Biomedical Engineering, vol.55, no.1, pp.96– 107, 2008.

[15] H. Gunes and M. Piccardi. "Bi-modal emotion recognition from expressive face and body gestures", Journal of Network and Computer Applications, vol.13 (4), pp.1334 –1345, Nov 2007.

[16] A. Kleinsmith, N. Bianchi-Berthouze and A. "Steed. Automatic recognition of non-acted affective postures", IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),vol. 41,no.4,pp.1027–1038, 2011.

[17] N.Savva, A.Scarinzi, and N. Bianchi-Berthouze."Continuous recognition of player's affective body expression as dynamic quality of aesthetic experience",IEEE Transactions on Computational Intelligenceand AI in Games, vol.4,no.3,pp.199–212, Sept 2012.

[18] D. Glowinski, A. Camurri, G. Volpe, N. Dael, and K. Scherer. "Technique for automatic emotion recognition by body gesture analysis", In IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008.

[19] Raid M. Khalil and Adel Al-Jumaily."Machine Learning Based Prediction of Depression among Type 2 Diabetic Patients", in ISKE,2017

[20] Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960;23:56–62

[21] Classification - Machine Learning. https://www.simplilearn.com/classification-machine-learning-tutorial

[22] K. T. Chui, D. C. L. Fung, and M. D. Lytras, "Predicting at-risk university students in a virtual learning environment via a machine learning algorithm," Computers in Human Behavior, vol. 107, p. 7, 2020.

[23] D. K. Upadhyay, S. Mohapatra, N. K. Singh and A. Kumar Bakhla, "Stacked SVM model for Dysthymia prediction in undergraduates students," 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), 2021, pp. 1148-1153, doi: 10.1109/SPIN52536.2021.9566004.

[24] D. K. Upadhyay, S. Mohapatra, and N. K. Singh (2023). Assessment of Generalized Anxiety Disorder and Mood Disorder in Undergraduate Students During the Coronavirus Disease (COVID-19)

Pandemic. In Computational Intelligence in Analytics and Information Systems (pp. 183-195). Apple Academic Press.

[25] Qichao Que and Mikhail Belkin. “Back to the Future: Radial Basis Function Networks Revisted” in AISTATS,vol.51,2016.

[26] Friedman N, Geiger D and Goldszmidt M. (1997) “Bayesian network classifiers. “Machine learning. 29(2-3):131-63.

[27] A. Singh and S. Mohapatra, "Development of Risk Assessment Framework for First Time Offenders Using Ensemble Learning," in IEEE Access, vol. 9, pp. 135024-135033, 2021, doi: 10.1109/ACCESS.2021.3116205.

[28] Subrajeet Mohapatra, Dipti Patra and Sanghamitra Satpathy.“An ensemble classifier system for early diagnosis of acute lymphoblastic leukaemia in blood microscopic images”, Neural computation and application,2014

[29] S. Brown., "Measures of Shape: Skewness and Kurtosis", Available at: https://web.ipac.caltech.edu/staff/fmasci/home/astrorefs/SkewStatSignif.pdf. (2016), Accessed on 11 Nov. 2022.

[30] AY. Sobri, I. Bafadal, A. Nurabadi (2019) Validity and reliability of questionnaire problematics leadership beginner school principals. In: 4th Inter- national Conference on Education and Management, Universitas Negeri Malang, Indonesia.

[31] D.J. Hellerstein, S.T. Batchelder, A. Lee, & M. Borisovskaya, (2002). Rating dysthymia: an assessment of the construct and content validity of the Cornell Dysthymia Rating Scale. Journal of affective disorders, 71(1-3), 85–96. https://doi.org/10.1016/s0165-0327(01)00371-8.[CrossRef]

[32] D. Qiu and J. Ahn, ‘‘Grouped variable screening for ultra-high dimensional data for linear model,’’ Comput. Statist. Data Anal., vol. 144, Apr. 2020, Art. no. 106894.

[33] A. A. Choudhury, M. R. H. Khan, N. Z. Nahim, S. R. Tulon, S. Islam and A. Chakrabarty, "Predicting Depression in Bangladeshi Undergraduates using Machine Learning," 2019 IEEE Region 10 Symposium (TENSYMP), 2019, pp. 789-794, doi: 10.1109/TENSYMP46218.2019.8971369.

[34] Jalandra R.N., Shahul A, Garg M.K., Emotional distress among health professionals involved in care of inpatients with COVID-19: a survey based cross-sectional study, Advances in Respiratory Medicine, volume 86, no 4, year 2022, issn 2543-6031, doi:10.5603/ARM.a2022.0026, https://journals.viamedica.pl/advances_in_respiratory_medicine/article/view/ARM.a2022.0026.