IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

# Depression Detection using Machine Learning Algorithms

**K. Jayanthi[1], K. Deepika[2], K. Rupanjali[3], K. Tharun Durga Chowdary[4], K. Sai Eswar[5]**
GMR Institute of Technology, Rajam, Andhra Pradesh, India
kanduladeepika403@gmail.com[2] and jayanthikandi5@gmail.com[1]

**Abstract:** *Depression is one of the most common terms we have experienced among children, adults, students, professionals, the elderly and especially the young. cause. It is necessary to consult a doctor. Before the depression can be cured, it is necessary to diagnose whether the person has depression. People think they're depressed when they're a little nervous, so this study is supposed to predict whether a person is depressed. Machine learning algorithms are used to detect depression. This study has six different machine learning classifiers that use a variety of sociodemographic and psychosocial information to determine whether a person is depressed. Machine learning classifiers such as Chatbots, Logistic Regression, Naive Bayes, Decision Trees, Random Forests, SVMs (Support Vector Machines), KNNs (k Nearest Neighbors) are applied, with sensitivity, specificity and accuracy as measures of performance parameters It will be added. After applying different methods, I found that the classes in the confusion matrix are unbalanced. The analysis showed that Random Forest achieved his highest accuracy of 84% in predicting depression.*

**Keywords:** Depression, classifiers, machine learning Algorithms, Accuracy

## I. INTRODUCTION

Depression, one of those dangerous and indigestible words that is common anywhere. Regardless of age, all types of people suffer from depression. 3.8% of the total population suffers from depression. In particular, 5% of adults and 5.7% of people over 60 years old suffer from depression. About 280 million people suffer from depression. It was a psychological problem that they had to face due to their personal and professional life problems. 700,000 people die each year from depression. From the survey, it is known that most of the students all over the world suffer from depression. In this case they have to be consulted by the best psychologists and they have to be treated because they are the future of modern technology the students are depressed and feel on the verge of suicide and end his life.

Because depression is part of healthcare. Artificial intelligence plays an important role in this health issue. AI has become a part of human life. From start to finish, AI is present in every moment and in every application. It is an important key in the medical field to diagnose any illness or disease. Not only diagnose but also cure disease. It eases human work and makes work fast. Machine learning as a subset of artificial intelligence is an excellent statistical method for predicting diseases or health problems. Because depression is part of the study of psychology. Machine learning has become a major part of psychometric assessment. The use of machine learning is increasing dramatically day by day. Machine learning related techniques play an important role in predicting future analytics.

In this machine learning approach, the methods involved are 1. Data collection, data preparation (data pre-processing), data evaluation or data analysis. Feature selection models are created using machine learning. In ML, they mainly fall into three categories 1. supervised learning 2. unsupervised learning and 3. reinforcement learning. In this study, supervised learning is considered. In supervised machine learning, there are two types of regression and classifier. Classification algorithms are used when the output variable is categorical, meaning there are two classes like Yes-No, Male-Female, True-False, etc. In the classification there are mainly algorithms (techniques) like K nearest neighbors, SVM (support vector machine), Naïve Bayes, Decision Trees, Random Forests, etc., In this study, depression is discovered performed using data collected from students and their problems collected through surveys and machine learning techniques applied to pre-processed data. The detection of depression can be achieved through supervised and unsupervised learning. For different data inputs, different types of machine techniques are applied. This may involve image processing, NLP (Natural Language Processing) emotion analysis.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

## II. LITERATURE SURVEY

**[1] Zulfiker, M. S., Kabir, N., Biswas, A. A., Nazneen, T., & Uddin, M. S. (2021). An in-depth analysis of machine learning approaches to predict depression.** ***Current research in behavioral sciences*****, *2*,100044.**

This paper presents an efficient integration model based on deep learning (DL) for efficient pre-processing, segmentation, detection and recognition.

This study investigated six machine learning classifiers that use different sociodemographic and psychosocial information to detect whether a person is depressed. Three different feature selection methods. And identify the most effective model for prediction using accuracy or score F1

The machine learning classifiers used in the study are the backbone. mRMR, selectionkBest and Boruta are feature selection models.

The applied machine learning classifiers are Adaboot, KNN, GB, Bagging.

The synthetic minority oversampling technique (SMOTE) is used to reduce the class imbalance of the training data.

AdaBoost classifier with selectkBest feature selection with 92.56%.

**[2] Moon, N. N., Mariam, A., Sharmin, S., Islam, M. M., Nur, F. N., & Debnath, N. (2021). Machine learning approach to predict the depression in job sectors in Bangladesh.** ***Current Research in Behavioral Sciences*****, *2*, 100058.**

This study predicts depression in Employment Sector Survey participants based on occupation, age, income, gender, children, expenses, assets, and their effect on personality sustainability in difficult times. Five machine learning algorithms are used to predict whether people will experience depression.

Statistical measures used in data collection. Apply machine learning algorithms and predict depression with ease.

Machine learning algorithms such as random forest, naive bayes, random forest regression, factor analysis.

Random Forest Classifier 99%

Random Forest Repressor 99%

Naïve Bayes 99%

K Neighbors Classifier 97%

**[3] Govindasamy, K. A., & Palanichamy, N. (2021, May). Depression detection using machine learning techniques on twitter data. In** ***2021 5th International conference on intelligent computing and control systems (ICICCS)*** **(pp. 960-966). IEEE**

This study detects the depression based on the twitter data by applying machine learning algorithms. Classifying the whole data into depressed and undepressed data. applying the classifiers and measuring the accuracies.

Using depression and non-depression factors for easy prediction of the depression.

The NBTree algorithm gives an accuracy of 97.31% to classify the depressive and non-depressive tweets on the 3000 tweets dataset and 92.34% on the 1000 tweets dataset. On the other hand, Naïve Bayes show 97.31% on the 3000 tweets dataset and 92.34 % on 1000 tweets datasets

**[4] AlSagri, H. S., &Ykhlef, M. (2020). Machine learning-based approach for depression detection in twitter using content and activity features.** ***IEICE Transactions on Information and Systems*****, *103*(8), 1825-1832**

This study predicts the depression among the social media users. Based on the tweeted data he/she can be detected whether the person depressed or not. By applying machine learning algorithms f1 scores and accuracies are measured

Social media has altered the world. Finding people depressed will be easy by their tweets and and their responses

Machine learning algorithms like Decision tree, naïve bayes, SVMSVM with highest accuracy 92.3%.

**[5] Tao, X., Chi, O., Delaney, P. J., Li, L., & Huang, J. (2021). Detecting depression using an ensemble classifier based on Quality of Life scales.** ***Brain Informatics*****, *8*(1), 1-15**

In this study, depression is detected based on the qualityof life Scales. MDD Major depressive detection is measured on the NHANES (national health and nutrition examination survey). Ensemble binary classifier and baseline algorithm are applied.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

Using ensemble classifier which takes the average of the machine learning classifiers ensemble classifier accuracy Fe = 0.228 · fsvm + 0.283 · fnb + 0.266 · fknn + 0.223 · fdt

**[6] Su, D., Zhang, X., He, K., & Chen, Y. (2021). Use of machine learning approach to predict depression in the elderly in China: A longitudinal study.** ***Journal of Affective Disorders*****, *282*, 289-298.**
In this study, depression is predicted using two techniques. LSTM (long short-term memory) and machine learning models.
The retrospective waves used in the LSTM model need to be further increased.
Long short-term memory (LSTM) and six machine learning (ML) models were used to predict different depression risk factors and the depression risks in the elderly population in the next two years.
AUROC of logistic regression was highest among the machine learning models.

**[7] Kumar, M. R., Pooja, K., Udathu, M., Prasanna, J. L., & Santhosh, C. (2022). Detection of Depression Using Machine Learning Algorithms.** ***International Journal of Online & Biomedical Engineering*****, *16*(4).**
In this study depression is detected from the social media posts by the users using three machine learning algorithms.
Algorithms like logistic regression, naive bayes, naive bayes multinomial model (natural language processing) are applied in this study.
Some features should be added for the data to get more accuracy for detection through sentiment analysis.
Several research problems were outlined at the beginning of this paper to provide a clear picture of this work. Sentiment analysis also taken as consideration for this research.
The accuracies score by Logistic regression 92.34, Naïve bayes 97.31 with highest accuracy.

**[8] Chiong, R., Budhi, G. S., Dhakal, S., &Chiong, F. (2021). A textual-based featuring approach for depression detection using machine learning classifiers and social media texts.** ***Computers in Biology and Medicine*****, *135*, 104499.**
In this study, depression is detected using social media text via text featuring by applying machine learning algorithms.
The algorithms used are ensemble classifiers, supervised machine learning algorithms like MLP (multilayer perception), logistic regression.
It should be noted that the approach presented in this paper uses supervised ML classifiers, and therefore, the approach is limited to using labelled datasets for training the classifiers.
In this study the text doesn't contain any words like depression or diagnosis.
RF model is better at depression detection using general text than other classifiers.

[9] **Monreale, A., Iavarone, B., Rossetto, E., & Beretta, A. (2022). Detecting Addiction, Anxiety, and Depression by Users Psychometric Profiles.**
In this paper, we study the problem of exploiting supervised learning approaches, based on users' psychometric profiles extracted from Reddit posts, to detect users dealing with Addiction, Anxiety, and Depression disorders.
These approaches are indeed valuable for extracting insights on the aspects of mental illnesses shared by different users but fail to understand the mental state of specific users.
Psychometrics profiles are easier to identify the depression through their activities.
Decision tree with 0.93 accuracy, Random forest with 0.95 accuracy, SVM with highest 0.96 accuracy.

[10] **Makhmutova, M., Kainkaryam, R., Ferreira, M., Min, J., Jaggi, M., & Clay, I. (2021, June). Prediction of self-reported depression scores using person-generated health data from a virtual 1-year mental health observational study. In** ***Workshop on Future of Digital Biomarkers*** **(pp. 4-11).**
In this paper, prediction of depression is based on the PHQ-9 Depression detection questionnaire score using machine learning.
Data for this study were cross-sectional (SHARE Wave 6), so we could not examine temporal associations between risk/protective factors and depression onset. The decision to analyze cross-sectional data was made on the basis that

social network variables of key concern were only collected in SHARE Waves 4 and 6, with additional risk/protective factors of focal interest only collected at Wave 6.
Questionnaire score increases the accuracy of the applied classifier.
Best predictive XGBoost models after extended hyperparameter tuning, with 95% confidence intervals.

[11] **Handing, E. P., Strobl, C., Jiao, Y., Feliciano, L., &Aichele, S. (2022). Predictors of depression among middle-aged and older men and women in Europe: A machine learning approach.** ***The Lancet Regional Health-Europe*****, *18*, 100391.**
In this study, prediction of depression included a broad array of sociodemographic, relational, health, lifestyle, and cognitive variables. Calculated for men and women separately with different parameters.
Machine learning algorithms applied for this model are random forest analysis, logistic regression.
In this study, they could not include SLE and related antecedents for affective disorders in our analyses, nor were we able to account for menopausal status or potentially important neurophysiological measures (e.g., cerebral white matter lesion prevalence) as these variables were not available from SHARE.
Predicting the depression in women and men separately due to various depression factors may change for men to women.
Finally, the model achieved 95% CI [1.90,2.08] in men, 95% CI [1.85,2.02] in women.

[12] **Solomonov, N., Lee, J., Banerjee, S., Flückiger, C., Kanellopoulos, D., Gunning, F. M., ... &Alexopoulos, G. S. (2021). Modifiable predictors of nonresponse to psychotherapies for late-life depression with executive dysfunction: a machine learning approach.** ***Molecular Psychiatry*****, *26*(9), 5190-5198.**
Predicting the depression using tree models and Latent Growth Mixture Models (LGMM) to detect subgroups with distinct trajectories of change in depression by mid-treatment.
A trajectory of depression nonresponse detected by week 6 predicts poor response at treatment end.
In this study, both detection and treatment of the depression is taken into consideration.
Random forests model with high prediction accuracy (80%).

**[13] Na, K. S., Cho, S. E., Geem, Z. W., & Kim, Y. K. (2020). Predicting future onset of depression among community dwelling adults in the Republic of Korea using a machine learning algorithm.** ***Neuroscience Letters*****, *721*, 134804.**
In this study, predicting future of Korean adults at depression by using machine learning techniques such as random forest and ensemble methods are used to predict the model.
The absence of biological factors could limit the predictive performance in this study.
This predictive model is expected to be used for early identification of individuals at risk for depression and secure time to intervention.
AUROC (Area Under the Receiver Operating Characteristics) score is calculated for the RF and ensemble models AUROC score 0.87, Accuracy 0.86%.

[14] **Islam, M., Kabir, M. A., Ahmed, A., Kamal, A. R. M., Wang, H., &Ulhaq, A. (2019). Depression detection from social network data using machine learning techniques.** ***Health information science and systems*****, *6*(1), 1-12.**
In this study, depression is detected using social network data through machine learning models and sentiment analysis.
Body of literature that suggests that more focused studies in depression analysis are needed.
Three types of factors (emotional process, temporal process, linguistic style) and trained a model to utilize each type of factor independently and jointly.
Decision Tree (DT) gives the highest accuracy than other ML approaches.

[15] **Priya, A., Garg, S., &Tigga, N. P. (2020). Predicting anxiety, depression and stress in modern life using machine learning algorithms.** ***Procedia Computer Science*****, *167*, 1258-1267.**
In this study, the data is collected from employed and unemployed through Depression, Anxiety and Stress Scale questionnaire (DASS 21) and five machine learning algorithms are applied and f1 scores are measured and added.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

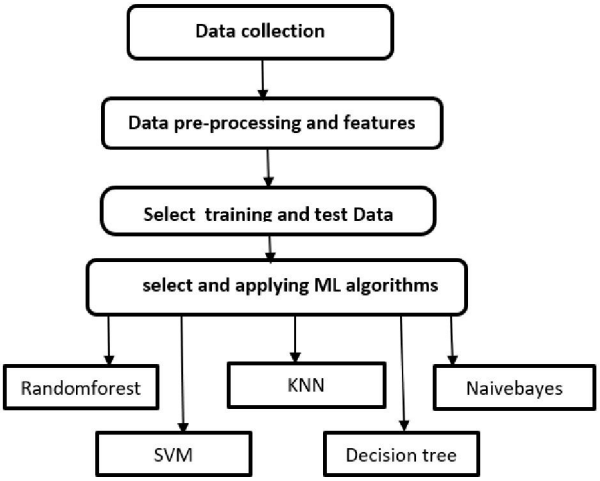
There would be a confusion that three metrics anxiety, depression, stress the performance will vary from one to another. Simple application of machine learning algorithms for detection like RF, SVM, KNN, Naïve bayes etc.
The accuracy of naïve Bayes 0.855 was found to be the highest, although Random Forest was identified as the best model.

## III. METHODOLOGY

### 3.1 Data Collection and Pre-Processing

In this study the data is collected from a survey from the people of survey in the covid time. The data is categorized into different variables based on their employment and unemployment, men and women etc., this data gets pre-processed. Data analysis is taking place Based on the correlation of the variables.

### 3.2 Flowchart




### 3.3 Algorithms

Machine learning algorithms like K-nearest neighbour, SVM, naivebayes, decision tree, random Forest are used for this study.

### A. k-nearest Neighbour

KNN is a supervised machine learning algorithm used for both regression and classification. It is non-parametric in nature. In this algorithm, the problem is to predict the target class label which is the most commonly represented class label among k most similar training examples for a given query point. Several distance measures can be used, such as Manhattan distance, Minkowski distance, Hamming distance, Euclidean distance etc. in KNN.
A common distance measure for determining the k nearest neighbors Dk is the Euclidean distance. In this study, Minkowski Distance is considered to determine the between these two points $E$ ($e1$, $e2$, …, $en$) and $F$ ($f1$, $f2$, …, $fn$) can be expressed as the following equation.

$$distance(E, F) = \left( \sum_{i=1}^{n} \left( |e_i - f_i| \right)^q \right)^{\frac{1}{q}}$$

### B. Random Forest

Random forest is a supervised learning technique for classification and regression algorithms in machine learning. It is based on the concept of group learning, which is a process of combining multiple classifiers to solve a complex problem and improve the performance of the model.
Random Forest is a classifier that holds several decision trees on different subsets of a given data set and takes the mean to improve the prediction accuracy of that data set.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

A higher number of trees in the forest leads to more accuracy and avoids the problem of over-fitting.

**C. Naïve Bayes Algorithm**

Naïve Bayes algorithm is a supervised learning algorithm, it is a statistical classifier derived from Bayes theorem and it considers the class feature to be independent of other features. The Naive Bayes classifier assumes that the effect of one object in one class is independent of another. Naive Bayes practices probabilistic approach that can predict test data class and excels at predicting multiple classes.

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

**D. Support Vector Machines**

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. The goal of the SVM algorithm is to generate the best decision line or boundary that can separate the n-dimensional space into classes so that we can easily place new data points in the right categories in the future. This best decision limit is a data point in the future matching category. This best decision boundary is called the hyperplane.

SVM has four widely used multiplication functions, namely linear function, polynomial function, sigmoid function and radial basis function (RBF).

In this study, SVM kernel functions are used in this study.

**E. Decision Tree**

In a decision tree, there are two nodes, the decision node and the leaf node. Decision nodes are used to make any decision and have many branches while leaf nodes are the output of those decisions and do not contain any other branches. Decisions or tests are made based on the characteristics of the given dataset.

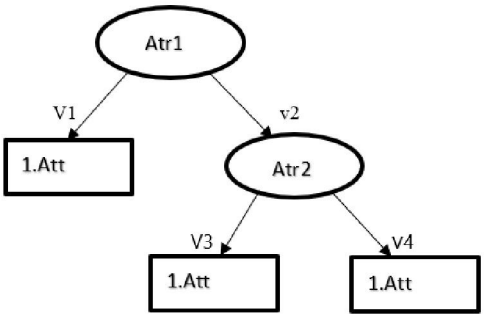



**Figure:** Decision tree example for NB Tree

## IV. EVALUATION MEASURES

**Accuracy**: It is the most commonly used measure to evaluate a classifier. It is defined as the degree of right predictions of a model. Based on the accuracy score the best model will be choose or selected.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

**Precision:** It is the ratio of the true positives and all the positives. It tells you that Out of all the positive classes we have predicted, how many are actually positive.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

**RECALL**: It tells you that out of all the positive classes, how many we predicted correctly. It is also known as sensitivity.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives + False Negatives}}$$

**F1-score:** The F1 score is defined as the harmonic mean of precision and recall. It is also used to compare the performance of the classifiers.

$$F1\ score = 2 * \frac{Precision\ * Recall}{Precision + Recall}$$

**ROC**: An ROCcurve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters:

- True Positive Rate
- False Positive Rate

### 4.1 Confusion Matrix

The confusion matrix is a matrix used to determine the performance of the classification models for a given set of test data. It can only be determined if the true values for test data are known. The matrix itself can be easily understood, but the related terminologies may be confusing. Since it shows the errors in the model performance in the form of a matrix, hence also known as an error matrix.

## V. RESULTS AND DISCUSSION

In this study, the test was performed on a data set consisting of 23 characteristics (columns) and a total of about 1429 records were collected based on their livelihood, financial status and age. use different classification algorithms (SVM, KNN, DT, Random Forest and NB) applied to the given data set.

Basically, the algorithm works well to detect the depression predicted from this study. Different tests are done to check if the person is depressed. Metrics such as accuracy score, precision, recall, and f1 score for each algorithm are measured. The ROC receiver operating curve is observed for the classifiers.

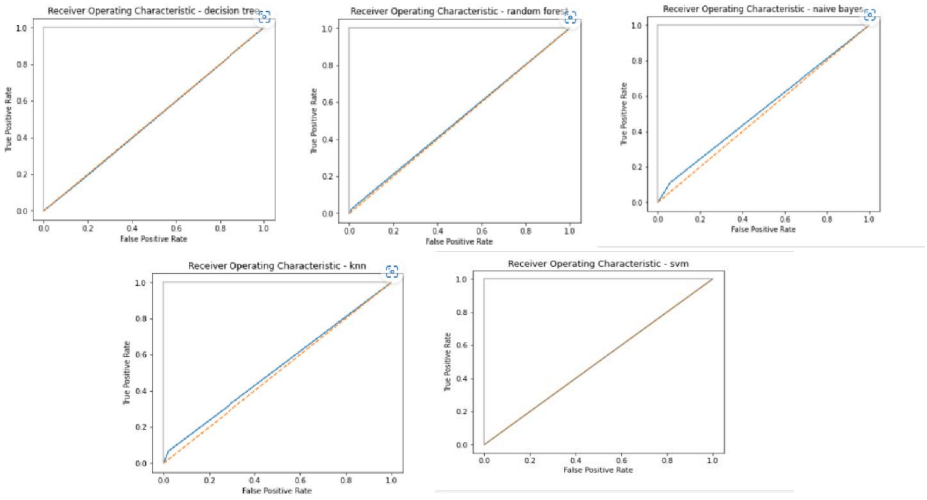
In this study, the accuracy of different SVM classifiers with 84%, naive with 81%, random forest with 84%, KNN with 84%, decision tree with 72%. Among all classifiers compared to other measures such as accuracy, recall, f1 score, support, and ROC random forest, the performance outperformed with high predictability of depression. Figure 1 below shows the ROC and AUC curves of the classifiers.

**Different Metric Measures for the Different Classifiers**

| classifier | accuracy | precision | recall | F1-score | support |
|---|---|---|---|---|---|
| SVM | 0.84 | 0.84 | 1.00 | 0.91 | 241 |
| Naïve-bayes | 0.81 | 0.85 | 0.94 | 0.89 | 241 |
| Random forest | 0.84 | 0.85 | 1.00 | 0.91 | 241 |
| KNN | 0.84 | 0.85 | 0.98 | 0.91 | 241 |
| Decision tree | 0.72 | 0.84 | 0.83 | 0.84 | 241 |

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

**Figure:** Graphs of ROC Curve




## VI. CONCLUSION

Many different factors can play a role in the development of depression in a person. This study tried to find out the most common factors that cause depression. First, a dataset was created that included 23 sociodemographic and psychosocial factors from 1429 participants to screen for depression. Various feature selection techniques have extracted the most important demographic and psychosocial factors responsible for the formation of depression. To predict depression in this study, five different machine learning classifiers were used. It can be asserted that the random forest classifier is almost the perfect model for predicting depression in the participants. He has an accuracy of 84%. Accuracy can be increased using advanced feature selection models. This study used a larger number of characteristics and factors to define depression. No biomarkers were included in the dataset to predict depression.

## REFERENCES

**[1].** Zulfiker, M. S., Kabir, N., Biswas, A. A., Nazneen, T., & Uddin, M. S. (2021). An in-depth analysis of machine learning approaches to predict depression. *Current research in behavioral sciences*, *2*,100044.

**[2].** Moon, N. N., Mariam, A., Sharmin, S., Islam, M. M., Nur, F. N., & Debnath, N. (2021). Machine learning approach to predict the depression in job sectors in Bangladesh. *Current Research in Behavioral Sciences*, *2*, 100058.

**[3].** Govindasamy, K. A., &Palanichamy, N. (2021, May). Depression detection using machine learning techniques on twitter data. In *2021 5th international conference on intelligent computing and control systems (ICICCS)* (pp. 960-966). IEEE

**[4].** AlSagri, H. S., &Ykhlef, M. (2020). Machine learning-based approach for depression detection in twitter using content and activity features. *IEICE Transactions on Information and Systems*, *103*(8), 1825-1832

**[5].** Tao, X., Chi, O., Delaney, P. J., Li, L., & Huang, J. (2021). Detecting depression using an ensemble classifier based on Quality of Life scales. *Brain Informatics*, *8*(1), 1-15

**[6].** Su, D., Zhang, X., He, K., & Chen, Y. (2021). Use of machine learning approach to predict depression in the elderly in China: A longitudinal study. *Journal of Affective Disorders*, *282*, 289

**[7].** Kumar, M. R., Pooja, K., Udathu, M., Prasanna, J. L., & Santhosh, C. (2022). Detection of Depression Using Machine Learning Algorithms. *International Journal of Online & Biomedical Engineering*, *16*(4).

**[8].** Chiong, R., Budhi, G. S., Dhakal, S., &Chiong, F. (2021). A textual-based featuring approach for depression detection using machine learning classifiers and social media texts. *Computers in Biology and Medicine*, *135*, 104499.

**[9].** Monreale, A., Iavarone, B., Rossetto, E., & Beretta, A. (2022). Detecting Addiction, Anxiety, and Depression by Users Psychometric Profiles

**[10].** Makhmutova, M., Kainkaryam, R., Ferreira, M., Min, J., Jaggi, M., & Clay, I. (2021, June). Prediction of

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 1, November 2022
Impact Factor: 6.252

self-reported depression scores using person-generated health data from a virtual 1-year mental health observational study. In *Workshop on Future of Digital Biomarkers* (pp. 4-11). Handing, E. P., Strobl, C., Jiao, Y., Feliciano, L., &Aichele, S. (2022). Predictors of depression among middle-aged and older men and women in Europe: A machine learning approach. *The Lancet Regional Health-Europe*, *18*,

**[11].** Handing, E. P., Strobl, C., Jiao, Y., Feliciano, L., &Aichele, S. (2022). Predictors of depression among middle-aged and older men and women in Europe: A machine learning approach. *The Lancet Regional Health-Europe*, *18*, 100391

**[12].** Solomonov, N., Lee, J., Banerjee, S., Flückiger, C., Kanellopoulos, D., Gunning, F. M., ... &Alexopoulos, G. S. (2021). Modifiable predictors of nonresponse to psychotherapies for late-life depression with executive dysfunction: a machine learning approach. *Molecular Psychiatry*, *26*(9), 5190-5198.

**[13].** Na, K. S., Cho, S. E., Geem, Z. W., & Kim, Y. K. (2020). Predicting future onset of depression among community dwelling adults in the Republic of Korea using a machine learning algorithm. *Neuroscience Letters*, *721*, 134804.

**[14].** Islam, M., Kabir, M. A., Ahmed, A., Kamal, A. R. M., Wang, H., &Ulhaq, A. (2019). Depression detection from social network data using machine learning techniques. *Health information science and systems*, *6*(1), 1-12.