**Psychological wellbeing prediction and solutions using machine learning: A Systematic Literature Review**

*1 Yogita Hambir*

*Assitant Professor, Department of Computer Engineering*

*Army Institute of Technology, Pune*

*2Annu*

*Student, Department of Computer Engineering*

*Army Institute of Technology, Pune*

*3Priya*

*Student, Department of Computer Engineering*

*Army Institute of Technology, Pune*

*4Puja Nayak*

*Student, Department of Computer Engineering*

*Army Institute of Technology, Pune*

*5Mahima Chauhan*

*Student, Department of Computer Engineering*

*Army Institute of Technology, Pune*

**ABSTRACT:***These days, mental health is a big problem. It is crucial to find any problems and fix them before they have a significant effect. We strive to do this with the Mental Health Tracker App. Because users may be coping with mental illness, we will need to design the app to be extremely friendly and inviting.*

*The goal of our project is the development of a simple piece of machine learning software that tracks its users' progress while suggesting actions to help them improve their mental health. The user is posed with a number of questions, the application analyses their answers, proposes tasks, monitors their mental health, and displays the results on a dashboard. Machine learning was employed throughout this procedure.*

**KEYWORDS:** Random Forest, Naïve Bayes, Support Vector Machine, Recurrent Neural Network, Sentiment Analysis, Machine Learning Algorithms, Convolutional Neural Network, Logistic Regression, Linear Regression, Mental Health, Deep Learning, Binary Classification.

**INTRODUCTION**

Mental wellbeing of an individual is overall perspective of that individual and additionally gives a sign of a person’s general nature. Dysfunctional behaviour is a result of irregular characteristics in cerebrum science. The assessment of mental wellbeing is incredibly basic to comprehend and recommend treatments for patients with a digressed mental way of behaving. Most people are inclined to pressure while an are impacted by discouragement because of different reasons.

A managerial board of World Wellbeing Association (WHO)surveyed in 2011 that, by 2030, wretchedness will be the boss wellspring of overall infection trouble.

The mental health of a person may be used to gauge both that person's mental state and general character. Mental illness is brought on by chemical imbalances in the brain. Examining mental health is vitally necessary in order to understand and treat people with aberrant mental behaviour. The majority of individuals are vulnerable to stress, whereas other people suffer from depression for a number of causes.

According to a World Health Organization (WHO) administrative panel's 2011 forecast, depression would account for the majority of sickness burden globally by 2030. Early identification and treatment of mental health issues are essential. Those suffering from mental health issues can find relief via early identification, precise diagnosis, and efficient treatment. For the affected people, their families, and society at large, mental illness can have serious implications.

Interviews, observations, or questionnaires are typically used in conventional methods of mental health detection. Though labour- and time-intensive, traditional methods are frequently used. Mobile and wearable sensors have already been used to diagnose mental illness in the medical field. The majority of people who use these technologies, nevertheless, have been diagnosed with a mental condition and are under strict observation. The emotional, psychological, and social well-being of a person are all reflected in their mental health. One's feelings, ideas, and responses to circumstances are established. One can perform at their best when they have good mental health.

Childhood, adolescence, and adulthood are all essential life stages in terms of health.

Numerous variables, such as stress, social anxiety, depression, obsessive compulsive disorder, substance misuse, and others, can cause men substance abuse, personality issues, and obsessive compulsive disorder. To preserve a healthy life balance, it is becoming more and more crucial to recognise the early signs of mental illness. It is possible to completely utilise AI and machine learning capabilities to predict the start of mental disease. Real-time implementation of these apps will be advantageous.

**METHODOLOGY**

**1.NATURAL LANGUAGE PROCESSING**

Speech recognition, or speech-to-text as it is more often known, is the act of accurately converting voice input into text. Using speech recognition, our programme will reply to requests or commands made by voice. Speech recognition is particularly challenging since people often use poor grammar, speak fast, slur words together, use different accents and intonations, and speak in a variety of dialects. Natural language processing is being used to translate user audio input into texts, which will subsequently be fed into a machine learning model for prediction. Considering how ambiguous human language is, it is incredibly difficult to develop software that can accurately determine the intended meaning of text or audio data. Changes in phrase structure, sarcasm, idioms, metaphors, homonyms, homophones, and exceptions to grammar and usage norms

**2.LANGUAGE TRANSLATION API**

Accurate translation is crucial when creating a new website, mobile app, or software programme. It can be challenging to stay current with the most recent trends and advancements in the modern world because there are so many different languages and dialects. Because of this, developers require a simple method for including potent translation features in their works. Translation apps make communication easier by converting written material between two languages. The basic goal of a translation app is to make it possible for readers to read a translation as though it were the original. In our application we are using Libre language Translation API . Libre is an open source language translation software. we are using it API to take user input in different languages. We will convert different language inputs into English inputs and feed it into machine learning models.

**3.MACHINE LEARNING ALGORITHMS**

For the purpose of predicting mental health based on user input, we are applying machine learning algorithms. For model training, we are using an online data set. The data set will be made up of 30% of testing data and 70% of training data.

It needs less training time compared to other algorithms. Despite the vast amount of data, it still works well and makes accurate predictions about the outcome. Accuracy can still be maintained even if a sizable portion of the data is missing.

**RANDOM FOREST ALGORITHM**

The random forest algorithm is a well-liked machine learning method. It uses ensemble learning to blend different decision trees to predict the future.

This is how the random forest method functions:

Data preparation is necessary because the algorithm predicts the target variable using the input features (variables) from a labelled dataset. Usually, the training set and test set are separated into two sections of the dataset. Random forest chooses a subset of features at random for each tree in the input dataset. By reducing the correlation between the trees and adding unpredictability, this technique helps. A technique known as bootstrap aggregating, often known as "bagging," is used to build several decision trees. The decision trees are trained using the appropriate bootstrap sample for each tree. The tree divides the data during training based on the chosen features and their corresponding thresholds in an effort to produce nodes that best categorise the data into distinct groups or forecast numerical values.

Voting or averaging of predictions: After all decision trees have been trained, predictions are made on the test set. When it comes to categorization, the random forest algorithm uses majority vote to integrate all trees' predictions. The technique takes the predictions from all trees and averages them for regression.

**NAÏVE BAYES ALGORITHM**

A straightforward yet efficient classification technique, naive Bayes is based on the Bayes theorem and makes a strong assumption of feature independence. Although it is frequently employed in text categorization and spam filtering, it can also be utilised for many other classification tasks.

The algorithm requires a labelled dataset because it predicts the target variable using the input features (variables). Usually, a training set and a test set are created from the dataset. The features are presumed to be conditionally independent of one another given the class using naive Bayes. As a result of the class designation, it can be concluded that the existence or lack of a specific feature has no bearing on the existence or absence of other features.

**SUPPORT VECTOR ALGORITHM**

One of the most popular supervised learning techniques is the Support Vector Machine (SVM), which is used to solve Classification and Regression problems. But the majority of its applications are in Machine Learning Classification issues.

The objective of the SVM algorithm is to find the best decision boundary or line that can categorise the n-dimensional space, allowing us to swiftly classify fresh data points in the future. Known as a hyperplane, this optimal decision boundary.

**RECURRENT NEURAL NETWORK**

Recurrent neural networks (RNNs) are a subclass of neural networks in which the results of one phase are used as the input for the next. Although the inputs and outputs of traditional neural networks are independent of one another, in cases when it is important to anticipate the following word in a sentence, it is necessary to remember the preceding words. It was decided to employ RNN, which utilised a Hidden Layer, to address this problem. The Hidden state of RNNs, which holds onto some information about a sequence, is its most crucial and basic component.

All the data necessary for calculations is kept in an internal "memory" of RNNs. It uses the same process that produces the input on all outputs or both to produce the output.

**4.PYTHON LIBRARIES**

A collection of linked modules is referred to as a "Python library". It contains groups of code that can be used repeatedly in numerous programmes. It makes Python programming simpler and more usable for the programmer. because we don't need to create the same code for numerous programmes over and over again. Python libraries play a significant role in the fields of data science, machine learning, data visualisation, and other related fields. A comprehensive tool for creating static, animated, and interactive visualisation is offered by the Matplotlib package for Python. Matplotlib makes both basic and complicated tasks possible. Write stories that are publishable. Create interactive charts that can zoom, pan, and update. Layout and visual style can be modified. Export to several file types. synchronise with Jupyter Lab and graphical user interfaces. Use one of the many third-party programmes that are based on Matplotlib. To display users' weekly mental growth graphically, we utilise Matplotlib.

**5.REACT JS**

One of the most widely used JavaScript libraries for creating mobile and online applications is ReactJS. React, developed by Facebook, consists of components, which are reusable JavaScript code fragments used to design user interfaces (UI).Create self-contained, encapsulated components and combine them to create sophisticated user interfaces. Reactjs is being used for front end development for our app. ReactJS is a JavaScript toolkit used to create modular user interface elements.

Modular user interfaces are made using the React framework. It encourages the creation of reusable UI elements for dynamic data displays. In MVC, React is usually used as the V. You can programme more simply and quickly with React since it abstracts the DOM away from you. React is capable of driving native apps with React Native and server-side rendering using Node. React implements one-way reactive data flow, which is less boilerplate and

**6.NODE JS**

An open-source, cross-platform run time environment for JavaScript called Node.js allows JavaScript code to be executed outside of a web browser. Node.js is a well-liked, easy-to-learn web framework that is utilised by many well-known businesses, including Netflix and Uber.

When thinking of JavaScript, web development usually comes to mind. Actually, JavaScript could not be run outside of a browser before Node.js was created. Node.js is a well-liked choice for building a back end server and database since it enables us to run the code as an independent programme than something that can only be assessed in a browser context. Ruby's Event Machine and Python's Twisted both had an influence on the architecture of Node.js. An event is a concept that is expanded upon by Node.js. An event loop is provided as a runtime construct rather than as a library. In other systems, a blocking call always initiates the event-loop.

**RELATED WORK**

In [1] decision trees, RF, SVM, NB and K nearest neighbour algorithms have been used for predictions. All other algorithms were surpassed by the random forest classifier at predicting anxiety and depression.

In [2] binary classification, deep learning methods have been used. The paper was published in 2019. It has been concluded in this paper that human focused methods will be better for engagement and they determine good mode of representation for prediction of mental health.

In [3] SVM, KNN, Logistic Regression, Binary Classification, Decision tress are used. This paper has proposed deep learning methods to be more efficient in prediction of depression. The study was conducted in 2023. Deep learning methods make use of neural networks for training models.

In [4] attention deficient hyperactivity prediction has been done using various machine learning methods. Logistic Regression, Naïves Bayes, KNN, SVM, Decision trees are used. Decision trees found out to be most accurate one in prediction of disorder. The study was published in 2021.

[5] is a systemic review of mental health analysis using biometric signals. ML models are the most commonly used ones. Logistic Regression, SVM, RNN are the proposed models. SVM is the mostly used in various studies is found in this paper.

In [6] decision making responses in covid 19 are studied. Data driven policy, financial aid policy, counselling policy are the common ones used for covid 19. Data driven policy turned out to be the most efficient response policy in covid 19. The research paper is published in 2021.

A 2021 paper on deep learning approach for depression intensity estimation. In [7] supervised learning methods are being used. This method outperform the existing method of binary classification by 2%.

In [8] convolutional neural network is used for prediction. The subject wise accuracy turned out to be 0.9937 and the record wise accuracy turned out to be 0.914.

In [9] Logistic Regression, linear discriment analysis and Guassian naïve bayes was used. The study stated that people became more vulnerable to mental health disorders after the spread of covid 19 pandemic. The study took place in Australia in 2021.

In [10] this has been found that Future wellness-promoting technology will have the chance to use social media data to model users' emotional behaviour. The models used in the study are LSTM and CNN.

In [11], the collaborative effort of psychologists, psychiatric professionals, and computer scientists aims to more effectively mitigate stress-related diseases across Europe. The paper focused on machine learning, virtual reality, cognitive computing and multimodel signal processing.

In [12] By putting the suggested method into practise and lowering the false positive rate, an accuracy of 99% was achieved. The evaluation findings demonstrated that, in comparison to Naive Bayes, SVM, CNN, and Decision Trees, this framework delivers good accuracy, precision, recall, and F1-measures. The study was conducted in 2022 and methods used are RNN and LSTM.

Random Forest, SVM , neural network, and XGBoost and logistic regression are used in [13]. Random forest outperformed in accuracy. After that SVM was the most accurate one.

According to [14], ML is effective in treating mental health issues on its alone. The development of the DL technique, however, increases the likelihood of correctly diagnosing and forecasting a single illness.

Random sampling methodology is applied in [15]. The findings showed that the overall prediction accuracy for data on adolescent depression was 94.17%, outperforming the individual sample technique.

2022 saw the completion of the [16] study. When it came to accurately and precisely predicting depressed classes in just 315 milliseconds of training time, RF outperformed all other systems by 99%. The techniques utilised here are Gaussian Naive Bayes, Decision Tree (DT), XG Boost (XGB), and RF.

Support vector machines and multilayer perceptrons, which are both about 100% accurate, offer the best predictions of suicidal ideation as indicated by [18].

In [19], Two of the applications that have received the greatest research are physical exercise (29.6%;35/118) and mental health (27.9; 33/118). The two most common sensors in smart phones for gathering information that can be used to determine a user's health or well-being are accelerometers (57.7%; 67/118) and global positioning systems (GPS; 40.6%; 48/118.

[20] In all patients, MDD patients, BDI patients, and BD II patients, the accuracy of mood state prediction for the following three days was 65%, 65%, 64%, and 65%, with respective area under the curve (AUC) values of 0.7, 0.69, 0.67, and 0.67. the model used is Random Forest.

The use of active applications, exercise duration, and GPS features all changed before depression in the group with multiple comorbidities. In [21] Every three weeks, online questionnaires were distributed. For 16 weeks, study participants were enrolled. The North western University Institutional Review Board gave its blessing to all procedures, and prior to participation, informed consent was sought from every participant.

[22] review discovered that using machine learning algorithms in real-world studies to implement digital health interventions can be advantageous and successful. Gradient boosting, RF, Decision trees, Adaptive boosting and Naïve Bayes are the main methods used in the paper.

In [23], 370 college students' data were collected over the course of 16 months, and a thorough analysis revealed a 0.89 correlation between the real and expected call duration using generalised linear models, which is almost 2.65 times higher than the correlation value of population-level generalised linear mixed models. Generalised linear mixed model was used.

In [24], the survey examines existing SID techniques from a wide range of angles, including clinical approaches like patient-physician interaction and medical signal sensing; textual content analysis approaches like lexicon-based filtering and word cloud visualisation and deep learning-based representation learning approaches like CNN and LSTM-based text encoders.

In [25], There were found to be 300 studies devoted to utilising ML to improve mental health. The literature identified four key application fields, including detection and diagnosis; prognosis, therapy, and support.

The study's findings are summarised in [26], which states that while OSN offer great promise as data sources for mental health problem diagnosis, they cannot completely replace traditional mental health detection strategies based on interviews, self-reporting and questionnaires.

Distribution. In the paper, SVM, RF, KNN, Gaussian process, and Naive Bayes have all been applied.

Decision Tree was found to be the most effective algorithm due to its quick execution and high accuracy in [27] when using data mining to predict mental health. We employed RF, Naive Bayes, and decision trees for prediction.

Online social material is anticipated to be the main SID medium in the future, according to [28]. Therefore, it is imperative to develop new methods that may connect automatic machine detection with clinical mental health.

With 95% accuracy and 99% precision, the Random Forest algorithm outperforms all other algorithms in predicting depressed classes by 99%. RF required 315 milliseconds for training. [29] shows the efficiency of random forest algorithm in prediction of mental health disorders.

Using Random Forest, it was shown in [30] that accuracy was higher with 100 trees than with 10. Accuracy is often increased by including additional trees. This effect grows until it flattens out at a specific point.

With and without an impute missing values learner, the random forest classification accuracy values on the substance abuse dataset were 87.72% and 92.15%, respectively. Adding the learner to impute missing data increased classification efficiency by 4.43%. The sole algorithm utilised in [31] is random forest.

According to [32], the classifiers SVM, KNN, ensemble (bagging), and tree ensemble (random forest) provide the psychologically disturbed class label an overall score of 0.95. It was determined that the data sample was as mental distress genuinely fits into that category, the score, which is pretty close to 1, indicates such.

According to a 2016 [33] paper, Multilayer Perceptron, Multiclass Classifier, and LAD Tree have all been found to produce results for mental health issues in children that are more accurate than the others.

In order to discover sequences of a user's number of unfavourable sentiments by week over time as markers of mental illness, [34] studied Long Short-Term Memory (LSTM), a sort of gated recurrent neural network.

**CONCLUSION**

In today's environment, having a physically and psychologically sound individual has become crucial for society's seamless operation. Over the world, many people struggle with pathological issues. It is challenging to treat all those with mental illnesses under the country's present hospital management structure. We are developing an app that uses machine learning algorithms to identify mental illness at an early stage.

We are collecting user input in the form of a questionnaire for this purpose. We can accept input in any language and in audio format thanks to language translation API and natural language processing. Our application will address the issue of early mental disease identification. In the future, mental health can be predicted using the user's personal information, such as call duration, missed call count, screen time, walking and sleeping patterns, heart rate, and facial expressions. We can include a system that uses dynamic data rather than static data. We may expand this programme with new features like immediate psychiatrist consultation.

**REFERENCES**

1. Anupriya, Shruti Garg , Neha Prerna Tigga : Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms, IEEE,2019.
2. Georgia Tech, ERIC P.S Baumer : Human-Centered Machine Learning:The Case of Predicting Mental Health from Social Media,2019
3. MS. Purude Vaishali Narayanrao, Dr. P. Lalitha Surya Kumari : Analysis of Machine Learning Algorithms for Predicting Depression , IEEE ,2023.
4. Tianhua Chen, Grigoris Antoniou, Marios Adamou, Ilias Tachmazidis, Pan Su :Automatic Diagnosis of Attention Deficit Hyperactivity Disorder Using Machine Learning , IEEE 2021.
5. Jolly Ehiabhi and HaifengWang: Systematic Review of Machine Learning Models in Mental Health Analysis Based on Multi Channel Multi-Modal Biometric Signals, IEEE 2023
6. Shuo Yu , Member, Qing Qing, Chen Zhang , Ahsan Shehzad , Giles Oatley, and Feng Xia : Data-Driven Decision-Making in COVID-19 Response:, IEEE2021
7. Shreya Ghosh: A Deep Learning Approach Depression, Intensity Estimation via Social Media , IEEE 2021
8. Ayan Seal, Rishabh Bajpai , Jagriti Agnihotri , Anis Yazidi ,Enrique Herrera-Viedma : A Deep Convolution Neural Network Framework for Detecting Depression Using EEG , IEEE 2021
9. Jianlong Zhou , Senior Member, IEEE, Hamad Zogan, Shuiqiao Yang, Shoaib Jameel, Guandong Xu , and Fang Chen : Detecting Community Depression Dynamics Due to COVID-19 Pandemic in Australia ,IEEE 2021
10. Mario Ezra Arag , Adrian Pastor Lopez-Monroy, Luis Carlos Gonzalez-Gurrola , and Manuel Montes-y-Gomez : Detecting Mental Disorders in Social Media Through Emotional Patterns - The Case of Anorexia and Depression , IEEE 2023
11. Kresimir osi , Sinisa Popovi , Marko Sarlija & Ivan Kesedzi : Impact of human disasters and covid 19 pandemic on mental health :potential of digital psychiatry ,2020
12. Amna Amanat , Muhammad Rizwan , Abdul Rehman Javed , Maha Abdelhaq , Raed Alsaqour,Sharnil Pandya : Deep Learning for Depression Detection from Textual Data, 2022
13. Ashley E. TateID, Ryan C. McCabeID, Henrik Larsson1, Sebastian Lundstro , Paul Lichtenstein Predicting mental health problems in adolescence using machine learning Techniques , April 6, 2020
14. Ngumimi Karen Iyortsuun , Soo-Hyung Kim ,Min Jhon , Hyung-Jeong Yang and Sudarshan Pant : A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis , 17 January 2023
15. Siriporn Sawangarreerak , Putthiporn Thanathamathee : Random Forest with Sampling Techniques for Handling Imbalanced Prediction of University Student Depression , 5 November 2020
16. Umme Marzia HaqueID , Enamul Kabir, Rasheda Khanam : Detection of child depression using machine learning methods , December 16, 2021
17. Masanori Nagamine, Szu-Nian Yang, Yueh-Ming Tai, Chin Lin : Machine Learning Based Suicide Ideation7, JULY 2020
18. Safwan Wshah, Christian Skalka, Matthew Price : Prediction for Military Personnel ,2019
19. Predicting Posttraumatic Stress Disorder Risk - A Machine Learning Approach , 2019
20. Alina Trifan, Maryse Oliveira , José Luís Oliveira Passive Sensing of Health Outcomes Through Smartphones: Systematic Review of Current Solutions and Possible Limitation , 2019
21. Chul-Hyun Cho1,Taek Lee, Min-Gwan Kim, Hoh Peter In, Leen Kim ,Heon-Jeong LeeMood : Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational Cohort Study, 2019
22. Jonah Meyerhoff Tony Liu, Konrad P Kording Lyle H Unga Susan M Kaiser Chris J Karr : Evaluation of Changes in Depression, Anxiety, and Social Anxiety Using Smartphone Sensor Features: Longitudinal Cohort Study , 2021
23. Andreas K Triantafyllidis : Applications of Machine Learning in Real-Life Digital Health Interventions: Review of the Literature , 2019
24. Ahmad Alimadadi, Sachin Aryal, Ishan Manandhar, X Patricia B. Munroe,Bina Joe Predicting Next Call Duration : A Future Direction to Promote Mental Health in the Age of Lockdown, March 5, 2023.
25. Adrian B. R. Shatte, Delyse M. Hutchinson and Samantha J. Teague : Machine learning in mental health: a scoping review of methods and applications ,10 January 2019
26. R. A. Rahman :Application of Machine Learning Methods in Mental Health Detection: A Systematic Review, 8, 2020
27. Vidit Laijawala ,Aadesh Aachaliya :Classification Algorithms based Mental Health Prediction using Data Mining , 2019
28. Shaoxiong Ji , Shirui Pan , Member, Xue Li, Erik Cambria , Guodong Long , and Zi Huang : Suicidal Ideation Detection: A Review of Machine Learning Methods and Applications ,2021
29. Umme Marzia HaqueID , Enamul Kabir, Rasheda Khanam :Detection of child depression using machine learning methods , December 16, 2021
30. Predicting Generalized Anxiety Disorder Among Women Using Random Forest Approach
31. Horeya Abou-Warda1, Nahla A. Belal : A Random Forest Model for Mental Disorders Diagnostic Systems ,2017
32. M. Srividya1 & S. Mohanavalli1 & N. Bhalaj :Behavioral Modeling for Mental Health using Machine Learning Algorithms ,2018
33. Ms. Sumathi M.R : Prediction of Mental Health Problems Among Children Using Machine Learning Techniques , 2016
34. Shailesh Hinduja Mahbuba Afrin Sajib Mistry, Aneesh Krishna : Machine learning-based proactive social-sensor service for mental health monitoring using twitter data , 2022