Asia-Pacific J.Mgmt. Tech. Volume 3(2) 39-55
https://doi.org/10.46977/apjmt.2022v03i02.004

AJMT

Asia-Pacific Journal of Management and Technology

Online ISSN: 2652-6840

journal homepage www.ajmt.org

Original Article

A Survey of Machine Learning Based Methods for the Diagnosis of Mental Health

Angshuman Sarkar*, Shambo Saurav Mallik

Narasinha Dutt College (C.U) Tikiapara, Howrah, West Bengal 711101, India

*Correspondence E-mail: anshuman_sarkar2000@yahoo.co.in

Abstract

Introduction: Mental illnesses like depression, schizophrenia, bipolar disorder, etc. have become widespread in today's society. More than 7.5% of Indians suffer from some kind of mental disorder. Early detection of mental illness is important for treatment as well as to prevent self-harm. Traditionally, the diagnosis involved answering a specifically designed questionnaire at the doctor's clinic. **Methods:** Mental health data, however, can also be collected from other sources like social media posts, wearable smart-devices, etc. Manual analysis of the patient data may not reveal all the information. Hence, diagnostic errors are common. In recent times Machine Learning (ML) algorithms have been successfully employed for the identification of critical symptoms, the development of diagnostic modules, and the personalization of therapy. **Result** - Authors have systematically reviewed 40 papers that used Machine Learning based techniques for the diagnosis of mental disorders and chose 25 of them for the survey. This paper provides an overview of the application of ML in mental healthcare. It has been focused here on the state-of-the-art Machine Learning based work on mental health. **Discussion**: The majority of the papers consulted so far concentrated on finding whether a subject belongs to a diagnostic category. Classifying into one broad category, however, does not take into account the variability of the symptoms. **Conclusion**: The authors identified some major research prospects in the early detection of schizophrenia.

Keywords: *Mental Health; Machine Learning; Bipolar Disorder; Schizophrenia; Depression*

Introduction

Mental disorders affect a person's behavior, mind, emotions, and physical health as well. Some of the common afflictions are schizophrenia, autism, attention deficit disorder, bipolar disorder, etc. 450 million adults and children are suffering from one or the other mental health disorder (World Health Organization, 2001). A study conducted in 2017 found that in the USA alone, 46.6 million adults are suffering from mental health problems (Abuse, 2020), while 7.5% of Indians suffer from some mental illness. The treatment gap in India is around 83% [Gururaj]. It is also recognised as one of the prevalent public health issues as many patients with mental health tend to opt for self-harm such as suicide. Thus, the diagnosis and detection of psychiatric problems at the earliest is very important in order to impart accurate and personalize treatment.

Unlike the other forms of illness that require pathological and other laboratory tests for diagnosis, the diagnosis of psychiatric problems requires a patient's own report and a specifically designed

Received 5 Sep 2022; Received in revised form 16 Sep 2022; Accepted 25 Sep 2022

https://doi.org/10.46977/apjmt.2022v03i02.004

questionnaire. Social media, smartphones, and neuroimaging devices also allow researchers and psychiatrists to collect personal health and behavioural data from patients (Chen, Mao & Liu, 2014). A robust technique is required to analyse these vast amounts of raw data and extract the necessary features to detect the psychiatric disorder. Different soft-computing-based techniques such as Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) are being constantly used nowadays in the analysis of data and making a final diagnosis. Hence, these are useful decision-making tools in the hands of clinicians.

The main strength of AI lies in its ability to evaluate a large dataset rapidly. It has thus been used in various fields of medical diagnosis, such as fetal heart rate analysis, cancer detection, CT scan analysis, etc. The AI-based methods were found to overcome the limitations of clinicians. ML builds systems using statistical and probabilistic models that learn from the data using either supervised or unsupervised algorithms (Lovejoy, 2019). The four major areas of application of ML in mental health are diagnosis, prognosis, research, and public health (Shatte, Hutchinson & Teague, 2019). Knowledge can be gained from complex data using DL, which transmutes the data through layers of nonlinear computational processing units. DL has recently been used in the study of brain dynamics and the behavior of patients. It is found to be useful for covering complex structures in high-dimensional data. A block diagram of machine learning-based methods for the diagnosis of mental health problems is given in Figure 1.

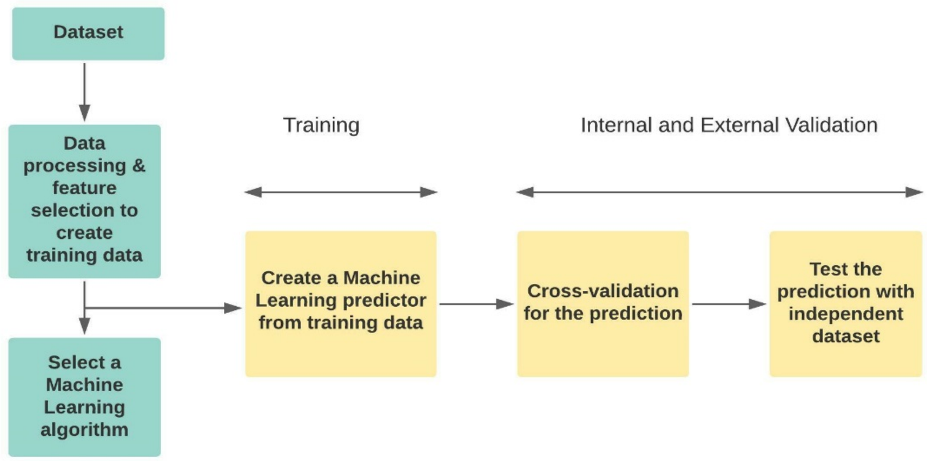

Figure 1: Block diagram of machine learning based methods for the diagnosis of mental health problems

1. Standard Guidelines for the Detection of Mental Health Condition

WHO considers the American Psychiatric Association's (APA) Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Statistical Classification of Diseases and Related Health Problems (ICD) the standard guidelines for the diagnosis of mental disorders. The current versions are DSM-IV and ICD-10, respectively. Questionnaires prepared following the DSM-IV guidelines are commonly used as the screening tool by clinicians. Subtle changes in the brain due to mental illness cannot be detected by MRI, CT scan, or EEG. In spite of the guidelines, mental illnesses are usually over or under diagnosed. Observation of the clinician and the symptoms are subjective in nature. Despite the guidelines, over or under diagnosis often occurs.

Various research works that investigated the application of various soft-computing based techniques in the field of mental health have been briefly discussed in this paper.

https://doi.org/10.46977/apjmt.2022v03i02.004

2. Some Common Types of Mental Health Disorders

Depression

Depression is considered as the feeling of sadness or anger that interferes with a person's everyday activities. People suffering from this disorder usually suffer from mood swings, irritability, sleep disorders, weight loss or gain, difficulty concentrating etc. To diagnose depression, doctors initially prescribe some tests to rule out other neurological and endocrine malfunctions that sometimes cause depression. Once these causes are eliminated, the doctors ask specific questions to find out the root cause. Drugs and psychotherapy may be prescribed after the diagnosis.

Bipolar Disorder

Patients suffer from extreme mood swings that go from a period of elation to a period of extreme depression. Other symptoms include talking very fast, restlessness, garbled thoughts, change in appetite, severe fatigue, lack of concentration, suicidal thoughts, etc. (Healthline, 2021). Detecting bipolar disorder is not simple. It requires at least one depressive and one manic or hypomanic episode. Questionnaires are given to patients during these periods to assess their mental state. For proper diagnosis, doctors follow the Diagnostic and Statistical Manual of Mental Disorders (DSM). Lab tests are also done to eliminate other causes (Healthline, 2021).

Diagnosis of bipolar disorder sometimes takes a long time because many of the symptoms are in common with other mental illnesses such as schizophrenia, depression, etc.

Schizophrenia

It is considered one of the most serious mental disorders. The person suffering from it loses touch with reality. The symptoms comprise hallucinations, delusions, disordered thinking, abnormal motor behavior, neglect of personal hygiene etc. (Mayo Clinic, 2020). Screening tests to rule out other disorders, CT scan or MRI, and psychiatric evaluation such as cognitive tests, personality tests, inkblot or Rorschach tests may be used in the diagnosis.

Personality Disorder

The patients suffer from unnatural pattern of thinking, functioning and behaving, leading to problems in relationships, social activities, work, and study. They are classed as follows (Mayo Clinic, 2020):

a) Schizoid
b) Schizotypal
c) Antisocial
d) Borderline.
e) Histrionic disorder.
f) Narcissistic personality.
g) Obsessive-compulsive disorder (OCD).

Diagnosis is sometimes difficult because some of the traits are common with other disorders. Some of the areas that should be noted are the way one identifies oneself, emotional response, dealing with people, and control over impulses. It, however, sometimes become difficult to determine the type of personality disorder as they share similar symptoms (Mayo Clinic, 2020).

3. Challenges in the Diagnosis of Mental Health Disorders

Institute of Medicine (IOM), in September 2015, released a report about diagnostic errors in medicine. This method is extensive in both scope and method. According to the report diagnostic errors can arise due to various reasons such as poor communication by the patients and their relatives, fatigue and stress of the physicians, lack of experience of the physician, imperfect reasoning. According to IOM there are five major diagnostic challenges specific to mental health where improvements are required.

i. Since there is no diagnostic test; the diagnosis is made based on the patient's reports and the clinicians observations. Diagnostic system thus has inherent limitations.

ii. Differentiating a physical symptom from mental health symptom can be a major challenge.

https://doi.org/10.46977/apjmt.2022v03i02.004

iii. Underlying medical condition may be responsible for the mental health symptoms.
iv. It has been observed that psychiatric diagnoses are often missed among elderly patients.
v. Some physical conditions may be the manifestation of some underlying psychiatric condition such as anxiety or mood swing.
vi. Error in clinical reasoning may have cognitive bias that leads to inappropriate action by the physician.

Methodology

Mental health issues are an affliction that affects people of various ages, ethnicity, gender, and socio-economic status. To build a robust system to detect mental health conditions using ML, a large and diverse dataset is required to alleviate bias. However, acquiring data is complicated because many patients do not easily come out with their problems as there is a stigma attached to them. Error, uncertainty, and bias also exist in the deployment of an intelligent system. All these factors combined together may introduce a lack of generalization. Solution, thus, lies in inter-disciplinary approach involving ML, human-computer interface, and the physicians.

The authors went through 40 different papers on mental health from different sources. The authors chose papers that dealt with the application of ML in mental health. The obtained research work was divided into four categories depending upon the type of data that was used:

i. Smart-device based.
ii. Text/ Social Media based.
iii. Structured-data based
iv. Multi-model system based.

Main motivations for applying ML to mental health assessment are:

i. Availability of natural data generated by the patients themselves through their interactions in social-media or through wearable devices. Such data are usually are less biased.
ii. ML-based approaches reduce the time needed to study the participants because the proxy data of mental health is available from public platform.
iii. Detection and diagnosis of the mental health related issues are possible at the earliest stage, thus, leading to the commencement of treatment online.
iv. More accurate diagnosis.

A block diagram for the diagnosis and classification of mental health issues is given in Figure 2.

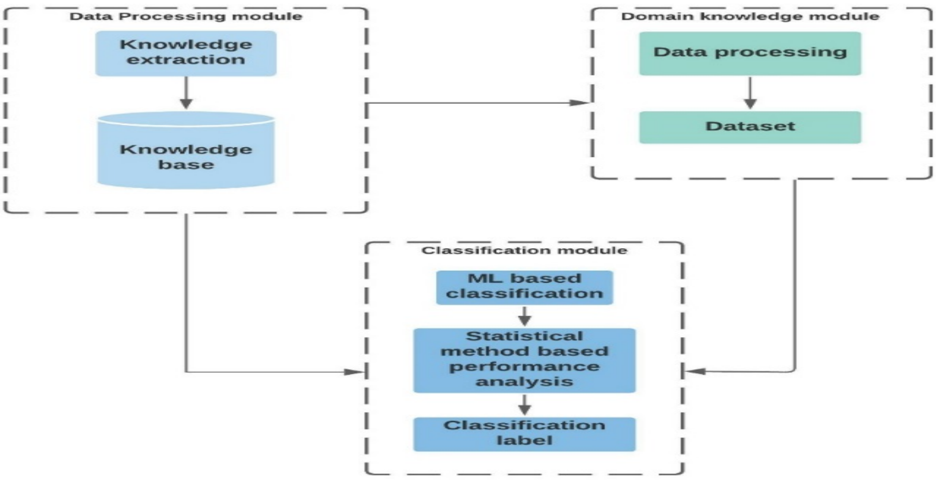

Figure 2: Block diagram representing the diagnosis and classification of mental health issues using machine learning

https://doi.org/10.46977/apjmt.2022v03i02.004

Existing Works on Mental Health Using Machine Learning Based Approach

Smart-device/ Sensors Based

Though most of the data are collected through questionnaire for the evaluation of mental health, many important data can be collected from smart devices. There are wide range of sensors integrated into smart watches to detect the signs of anxiety. High anxiety affects the heart rate pattern which then triggers a warning to the wearer or the clinician. Many apps using soft-computing based techniques and IOT, have been developed in recent times to monitor mental health. Some of the apps and their functions are mentioned in Table 1 (Truschel & Tzeses, 2021; Moturu *et al.*, 2011; Walker, 2015).

Table 1: Metal health monitoring and diagnosis apps

App Name	Mental Health	Functions
Ginger.io (Truschel & Tzeses, 2021	Depression	Sensor data, information from phone. Monitors depression by observing communication pattern to conclude if the patient is segregating from others, mood swings etc. Generates an alarm to the physician.
nOCD (Truschel & Tzeses, 2021	Obsessive Compulsive Disorder	When OCD is detected it assesses the severity and provides a guideline.
Schizophrenia HealthStorylines (Moturu *et al*., 2011)	Schizophrenia	Helps Schizophrenia patients with self-monitoring. Keeps note of mood swing, symptoms etc. and gives reminder about medication.
eMoods (Walker, 2015)	Bipolar Disorder	It keeps track of psychotic symptoms and mood swings. It provides a monthly report of mood swing.
PTSD Coach (Moturu *et al*., 2011)	Post-Traumatic Stress Disorder (PTSD)	Helps the patient with self-assessment by observing their mood and behavior patterns. One can customise the app to one's needs.

*1)*Chang, Chan and Canny (2011) used audio data of the patients for the detection of mental state of the patients. Five features of the speech and their patterns are given in the Table 2 below. Using ML based algorithms they categorized the human voice and developed automated monitoring system for the detection of depression (Chang, Chan & Canny, 2011).

Table 2: Features of voice used for detection of mental state

Feature	Pattern
Rate of speech	slow, rapid
Flow of speech	hesitant, long pauses, stuttering
Intensity of speech	loud, soft
Clarity	clear, slurred
Liveliness	pressured, monotonous, explosive
Quantity	verbose, scant

Source: Chang, Chan & Canny, 2011

https://doi.org/10.46977/apjmt.2022v03i02.004

2) van den Broek, van der Sluis and Dijkstra (2013) used audio data from 25 female patients of PTSD. They had specifically chosen female patients because a study by Blain, Galovski and Robinson (2010) reported that greater number of females than males suffer from PTSD. The data collection consisted of story-telling (ST) and reliving (RL) sessions. During ST all the participants read the same story that induced some kind of emotion. In the RL study participants recalled their last panic attack and the last happy moment of their life. Result of the speech analysis were compared to the analysis made by the physicians based on set of standard questionnaires. Five speech signal features which are sensitive to stress, and emotion were considered:

i. Intensity
ii. Fundamental frequency
iii. Zero crossing rate
iv. Amplitude
v. High-frequency power.

The signal processing phase consisted of removal of outliers, normalization of data, and derivation of parameter from the feature set. Choosing the set of features is crucial for the accuracy of the classification of stress. Since the number of input parameters was large, in the range of 2^{78}, it was essential to optimize the search space. They used a linear regression model (LRM) based heuristic search to optimize the parameters. It is an optimal relation between one independent variable and many dependent variables. LRM is represented as

$$y = l_0 + l_1 x_1 + \cdots . + l_i x_i + \varepsilon$$

l_i, with $i = 1, 2, \ldots, i$ are the regression coefficients, dependent variables x_i, and ε is the unobserved random noise. Parameter set was reduced to 28 using 32 iterations.

28 parameters were fed to Principal Component Analysis (PCA) was used to further reduce the dimensionality of the speech signal. k-Nearest Neighbour (kNN), Support Vector Machine (SVM), and Multilayer Perceptron (MLP) were used for classification. Percentage accuracy obtained with kNN, SVM, and MLP were 89.74, 89.74, and 82.37 respectively. The accuracy of the previous ML based models to recognize emotional speech ranged from 25% - 87%. However, these works handled 4 – 7 emotion levels (Wu, Falk, & Chan, 2011). The work was a significant step towards demonstrating stress through an acoustic model. Statistical models were built on the basis of a selection from 78 parameters of five features of speech, which showed reliable and robust stress classification.

3) To identify depression in patients Mitra *et al*. (2014) used AVEC-2014 challenge dataset that consists of 150 audio-visual records of 84 subjects. Since depression affects a person's production system, the tell-tale signs can be identified from the articulatory motions. The organs and their features are given in Table 3

Table 3: Features extracted from articulatory motion

Constriction Organ	**Vocal Tract Variables**	**Dynamic Range**	
		Max	**Min**
Lip	Lip Aperture (LA)	27.00 mm	-4.00 mm
	Lip Protrusion (LP)	12.00 mm	8.08 mm
Tongue Tip	Tongue tip constriction degree (TTCD)	31.07 mm	-4.00 mm
	Tongue tip constriction location (TTCL)	80.00 mm	0 mm
Tongue Body	Tongue body constriction degree (TBCD)	12.5 mm	-2.00 mm
	Tongue body constriction location (TBCL)	180°	87°
Velum	Velum (VEL)	0.20	-0.20
Glottis	Glottis (GLO)	0.74	0

Source: Mitra et al., 2014

Received 5 Sep 2022; Received in revised form 16 Sep 2022; Accepted 25 Sep 2022

https://doi.org/10.46977/apjmt.2022v03i02.004

Deep Neural Network (DNN) with 150, 200, 100, 80, 60, and 40 neurons were used. The number of neurons in each layer was optimized and the depth of the network and the depth of the network was augmented incrementally (Mitra *et al*., 2014). Before training and classification frame-level features were converted to waveform level features using i-vectors. Support vector regression was used to model the features. Root mean square error (RMSE) was found to be between 9.18 – 11.87 and mean absolute error (MAE) was between 7.68 – 9.99.

4) Frogner *et al*. (2019) use ML to detect depression in patients. Experiment group was divided into two – 23 participants with unipolar and bipolar disorder (condition group) whose motor activities were recorded and the other was the control group consisting of 32 people without depression. Convolution neural network (CNN) was used to categorize between depressed and non-depressed people and different levels of depression as well. The system could also predict Montgomery Åsberg Depression Rating Scale scores. Three experiments were performed:

i. Experiment 1 classified the participants into belonging to condition group and control group. 80% data were used for training using 3-fold cross validation out of which 40% were used for validation and 20% were for testing. Mean loss was found to be 0.06 and mean accuracy was 0.98.

ii. Experiment 2 determined the degree of depression of the participants using MADRS score which is a rating system to conclude how depressed a person is. Score of 0 – 10 means not depressed while a score of 11 – 19 is mild depression and above 20 means moderate depression. In this phase the shortest time segments were omitted and 96 hours long segment was used for classification. The highest accuracy was 72% for moderate depression.

iii. Experiment 3 predicted the MADRS score of the participants. 60% data were used for training with 2700 epochs and 40% were used for testing. 3-fold cross-validation was used for the training purpose. Mean square error (MSE) was found to be 4.

They compared the work with the works of Garcia-Ceja *et al*. (2018). The F1 score of Garcia was 0.66 and 0.67 with DNN and Random Forest respectively, while Frogner's method yielded a score of 0.7 without oversampling.

5) Mallol-Ragolta, Dhamija and Boult (2018) built a system for the diagnosis of the severity of PTSD symptoms. PTSD is a mental disorder that affects people who suffered some traumatic event in life. 6.8% adults in the USA alone suffer from this disorder. They had used The Engagement Arousal Self-Efficacy (EASE) multimodal dataset of 110 patients, consisting of facial and audio data, physiological signals and self-reports by the patients. The data were collected when the patients were undergoing treatment. The data were collected from the first two sessions which consisted of triggers (TR) and relaxation (RX) modules. Participants answered two sets of questionnaires – PCL-5 questionnaires were answered in the beginning of the treatment while coping self-efficacy (CSE-T) were answered before and after modules allocated during each session. The severity scores estimated from PCL-5 are called ΔPCL.

The aim of their work was to quantify the augmentation or regression severity of the PTSD symptom from the skin-conductance features extracted from TR, RX, and TR+RX datasets. The steps consist of

1. **Signal Processing:** EASE dataset, which consists of skin-conductance signal with a sampling rate of 256 Hz with 16 bit quantizer.
2. **Feature Extraction:** Features are extracted from the processed signals.
    i. The set of original signals S_{org} is the set that consists of mean and standard deviation of the signal, number of peaks per second, mean of the peak's magnitude, and the mean time to reach the peak amplitude.
    ii. The set of decomposed signal S_{dec} is the set that consists of mean and standard deviation of the tonic components of the signal, number of peaks per second, mean of the peak's magnitude, and the mean time to reach the peak amplitude.

https://doi.org/10.46977/apjmt.2022v03i02.004

Before proceeding to the next stage the features are z-normalized.

3. **Machine Learning:** ΔPCL was predicted using Support Vector Regressor (SVR) with the Radial Basis Function (RBF) kernel with both monomodal (CSE-T or S) and multimodal (CSE- T) data. Validation of TR+RX, TR and RX were done using cross validation. MSE between actual and the expected ΔPCL was used to estimate the performance of the system.

They compared the predictive power of the model trained with S_{org} and S_{dec} features of the S signal S signals corresponding to TR+RX, TR and RX modules. Comparison with CSE-T are given in Table 4. It was established that S signal model trained with S_{dec} feature from RX treatment module outperforms all the other modules.

Table 4: MSE comparison of CSE-IT and S signal models

Statistical Parameters	CSE-T	S Signal					
		TR + RX		TR		RX	
		S_{org}	S_{dec}	S_{org}	S_{dec}	S_{org}	S_{dec}
Mean	294.9	138.3	135.1	178.8	179.9	207.9	116.2
Median	137.3	34.8	46.4	47.7	50.1	48.9	48.1
Std. Deviation	466.1	254.3	231.7	320.1	294.4	333.2	199.4

Source: Mallol-Ragolta, Dhamija & Boult, 2018

6) Efforts have been made to evaluate mental health using mobile application and other wearable devices. Problems had arisen because these apps posed security risk because they accessed scores of personal data such as location, sleep pattern, conversations etc. These systems mainly emphasise on individual factors and overlook personal interaction aspects.

Zakaria, Balan and Lee (2019) proposed a scalable system called StressMon that automatically and non-intrusively identified persons displaying signs of extreme stress or depression in a work environment. It uses location information from the Wi-Fi to surmise both the individual and their group interaction patterns. The system was evaluated by the university students who worked in groups in class projects. Components of StressMon are shown in Figure 3.

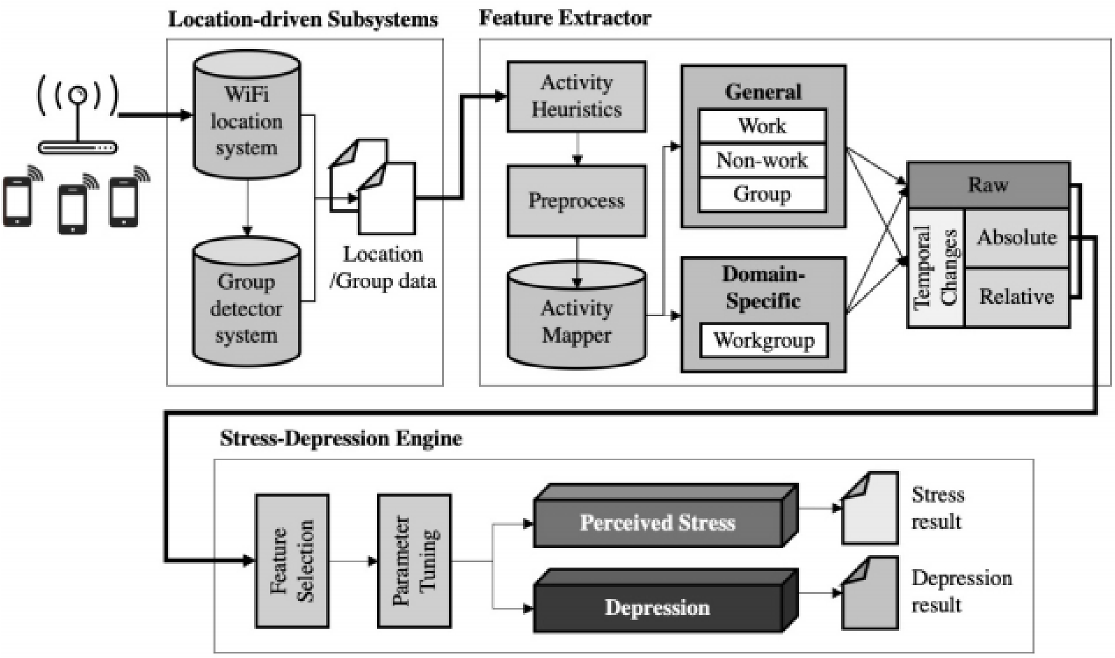

Figure 3: Components of StressMon (Zakaria, Balan & Lee, 2019)

https://doi.org/10.46977/apjmt.2022v03i02.004

Logistic Regression, SVM, and Random Forest were used for feature selection and classification, and their performances were compared. Input data for the stress-depression analysis consisted of location and group features averaged over three days. Stress level output were classified as *severe* or *normal* whereas, depression level was classified as *depressed* and *non-depressed*. The system was found to perform best with Random Forest with AUC, True Positive, and True Negative respectively as 0.97, 99.6%, and 72%.

The strength of StressMon was that it did not need to train a new stress model to detect severe stress among the participants. High accuracy for the detection of critical cases ensured early detection of stress and depression symptoms.

7) A smartphone can be used as a self-reporting tool to assess mental health. There have been a number of studies related to mood detection. However, those were conducted in controlled experiments with a limited number of participants. For the widespread acceptance of smartphones as decision support devices, there was a lack of a robust system to forestall one's mood for scant data. Spathis *et al*. (2019) worked with 735, 778 self-monitoring data collected from smartphones and other monitoring devices of 33,000 participants over a span of three years. They proposed a ML based model to predict the future mood sequences. Recurrent Neural Network (RNN) was used as it allows diagnosis for participants with partial self-reports. Traditional machine learning algorithms such as random forest, linear regression, etc., are known to forecast only one scalar value. The result usually presents a compound error that can skew the input distribution for upcoming forecasts. RNN, on the other hand, is increasingly used to model high-dimensional, non-sequential data.

The forecast model consists of an encoder and a decoder, both of which are RNN. It allows regression into multiple steps in the future. Since RNNs are difficult to train, Long Short Term Memory (LSTM) units help to overcome this problem [Sepp]. The sliding window of step 1 was used over the mood sequence of each participant, giving four weeks of past and one week of future sequential mood sequence 80% of the data was used for training and 20% was used for testing to avoid overfitting.

The implementation was based on Keras with Tensorflow in the backend. Mood variation, personality, and day of the week were found to play a major role in the performance of the model. This multi-task performing model was found to have a higher accuracy than the traditional ML models.

Text/ Social Media Based Approach

Millions of people use social media for interaction with people, and they end up sharing their emotions and moods. Analysis of the patterns of Facebook and Twitter use has revealed that important information related to one's mental health can be obtained. McClellan *et al*. (2017) analyzed Twitter data using terms related to depression and suicide from 2011–2014 and developed an experimental model to foresee trends in communication. They used an autoregressive integrated moving average (ARIMA) model to perform the analysis and concluded that the appropriate ARIMA model for the monitoring of depression and suicide is of the order 7. The trend is shown in Figure 4. They concluded that this model can be used successfully to identify periods of heightened mental health-related activity on Twitter.

Yazdavar *et al*. (2020) performed an exhaustive analysis of visual and contextual data of 400 million tweets from 45,000 Twitter users to identify depression in individuals. They used statistical methods to fuse diverse sets of features from these data and developed a multimodal network. The feature sets were input to a SVM classifier with a linear kernel.

De Choudhury *et al*. (2016) built a classifier to predict depression using features such as language, style of expression, emotion, etc. from written words in Twitter posts. Yates, Cohan and Goharian (2017) developed an ANN-based model to detect depression in users by merging tweets into a representation of the user's activities.

https://doi.org/10.46977/apjmt.2022v03i02.004

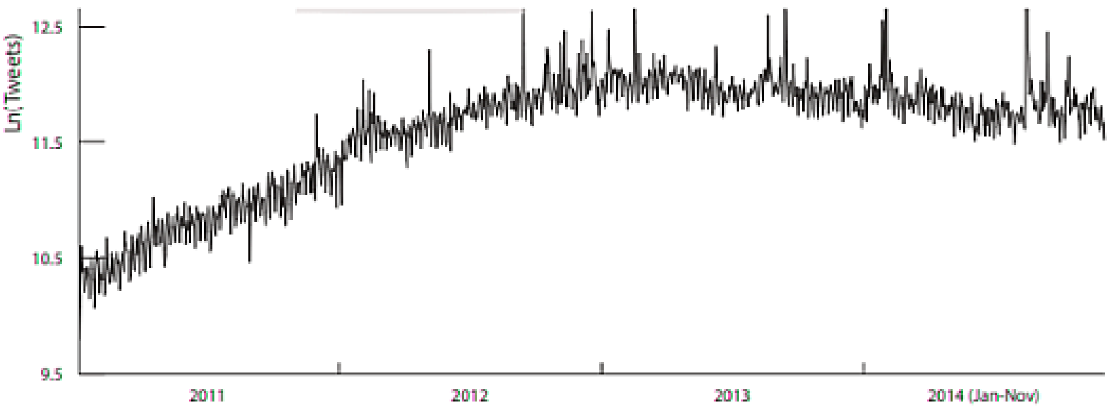

Figure 4: Natural log transformation of tweets to monitor suicide and depression from 01/01/2011 – 28/12/2014 (McClellan *et al*., 2017)

1) Social media posts can be used as a clue to detect mental disorder at an early stage. Fatima *et al*. (2018) applied machine learning and statistical analysis techniques to distinguish between depressive and non-depressive posts in social media. They studied different social networking sites such as Twitter, Facebook etc. Data were also extracted from LiveJournal which is a platform where people join different groups of interest and discuss various issues. Difference between the communities was identified using Linguistic Inquiry Word-Count (LIWC).

They trained the system with 3626 social media posts and tested with 200 depressive and 200 non- depressive posts using 10-fold cross validation. 10-fold cross validation was also done on community posts from which 200 data each from depressed and non-depressed patients were used. Random Forest was used because of its ability to do multiclass classification with high accuracy. The results were compared with the performance of SVM. Levels of depression were classified as *mild*, *moderate*, and *severe*. The comparison is shown in Table 5.

Table 5: Statistical measure of the performance comparison of the two classifiers

Classifier	Data Source	Precision	Recall	F-Measure	Accuracy
Random Forest	Social media post	0.892	0.905	0.897	0.898
	Community post	0.9	1.0	0.947	0.950
SVM	Social media post	0.818	0.783	0.799	0.820
	Community post	0.875	0.885	0.879	0.895

Random Forest was found to perform better than SVM. The accuracy was 90% and 95% respectively for the classification of depressive post and depressive community. For the classification of the depression they had used only the social media post. Accuracy of the classification of the degree of depression.

1) Stress is a part and parcel of campus life. Violent incidents on college campuses have become common nowadays. They focused on the other 12 such incidents on different campuses in the US. Saha and De Choudhury (2017) pulled the Reddit post of the students and used ML-based classification to analyse the level of stress after gun violence on the campus. The forum structure of Reddit is used both for content sharing and getting information from numerous communities.

https://doi.org/10.46977/apjmt.2022v03i02.004

They quantified the temporal dynamics of stress using both time domain analysis and frequency domain analysis. In order to assess the linguistic expression in the post, two forms of language analysis were used – i) psycholinguistic characterization and ii) incident-specific lexical analysis.

A stress classifier was built using binary SVM with a 5-fold cross-validation technique. Sentiments were classified as positive, negative, and neutral. Accuracy was found to be 82%. It was observed that psychological stress may be automatically detected from social media content by employing supervised learning approaches.

2) Joshi *et al*. (2018) used Deep learning feature extraction algorithm to evaluate the mental health of persons from their social media posting. They collected 1.2 million Twitter data from 166 users and analysed phrases to detect suicidal intention of the users. Stop words were removed from the texts and the features were categorized into structural features and behavioural features. Behavioural features were identified manually, whereas, structural features were identified using deep learning.
Automatic analysis of the language in the tweet helped identify the structural features that are indicative of emotions and sentiments. The feature set consisted of 14 features such as number of followers, retweets, hashtags, time of tweet etc. were used to train the classifier and its accuracy is measured. Percentage of abnormal tweets was the feature that was obtained from deep learning. Thus, there were 13 behavioural and one structured feature. All these features were used for the final classification of the user. The average accuracy was 89% with Deep Learning.

3) Yazdavar *et al*. (2020) used a semi-supervised model to analyze the data from 2000 Twitter profiles to detect the signs of clinical depression. The dataset was constructed from 4500 tweets where the users confessed of depression and also from 2000 tweets where the users did not declare to be depressed. Since, in social media people use varied terms to convey certain emotions, their primary challenge was to generate personalised set of seed terms of each user. A term was recognized as seed if it conveyed negative sentiment. They first identified the words that generally appear in the conversation of a depressed individual to create a lexicon of depression symptoms. A semi-supervised statistical model then extracts, classifies, and monitors the symptoms of depression on a continuous basis.
Three people were asked to judge and annotate each tweet using PHQ-9 categories as label. Cohen's inter-rater agreement was $\kappa = 0.74$, indicating good agreement. Two approaches were used for the detection of symptoms – processing the datasets using a bottom-up approach to discover specific word clusters conveying depression and the second approach combines the first one with the top-down method using the lexicon terms to identify the symptoms. Sentiment analysis was performed using Python TextBlob. The proposed model predicted the presence of each of the nine symptoms with 68% accuracy with an average precision of 72%.

4) Suicide is one of the serious health hazards all over the world. In India alone the rate of suicide is 10.5 per 100,000, while the worldwide figure is 11.6 per 100,000 [ndtv]. More than 90% of the people who commit suicide were found to suffer from mental disorder. Assessments by the psychologists is usually the prime method of predicting future attempts, however, these predictions usually lack accuracy. Suicide risks of an individual can be detected from different text-based contents such as chat, email, social media contents etc. Nobles *et al.* (2018) collected SMS from individuals to build a database and recognize the distinctive patterns in communication. A total of more than one million messages were obtained from 26 individuals. They built a deep neural network (DNN) model to differentiate between languages used during the periods of suicidality and the periods of depression. The sensitivity was 81% and the false negative (FN) was 44%.
Adamou *et al*. (2018) used the data of 130 mental health patients who died by suicide between 2013 – 2016. This data consisted of progress notes, medical assessments, and events that demarcated the clinical trajectory of the patients. Various machine learning algorithms such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree, and Ridge Logistic Regression were used to train the models

https://doi.org/10.46977/apjmt.2022v03i02.004

with 10-fold cross-validation. The final model would be able to predict from patient's medical records the future tendency to attempt suicide.

5) Chen *et al*. (2018) used 8 basic features of emotions from Twitter post over a period of time. These features were anger, disgust, fear, happiness, sadness, surprise, shame, and confusion. Both temporal and non-temporal analysis of the features revealed that the expressions in the twits can give valuable insight into the level of depression of a person. Prediction was a binary classification that was performed using SVM and Random Forest classifiers. They achieved an accuracy of 89.83% and 92.17% respectively with SVM and Random Forest.

Structured Data Based Approach

*1)*It is difficult to distinguish between Type-I and Type-II bipolar disorder because of the similarity between the symptoms. Mood disorder questionnaire (MDQ) [Hirschfield] and hypomania checklist (HCL-32) [Angst] are some of the self-reporting tools that were used by the psychiatrists, however, they had recall bias. Based on the works of Robins [Sachs] affective disorder evaluation (ADE), which has 145 detailed questions, was developed to help the process of making diagnosis. Feng *et al*. (2018) applied minimum redundancy maximum relevance (mRMR) feature selection approach to downsize the ADE to 112 questions and proposed a new model to effectively differentiate between Type-I and Type-II bipolar disorder.

They collected the data of 281 Type-I and 79 Type-II patients. Data were analyzed using five machine learning algorithms – Random Forest, Support Vector Machine, Lasso, LDA, Logistic Regression, with 10-fold cross-validation. Using their optimal feature set, Random Forest performed better than the other four algorithms with AUC ≥ 0.9.

*2)*According to WHO, by 2030 mental health disorder, worldwide, will be a major issue to reckon with. In India around 50 million people suffer from depression that requires clinical intervention [WHO]. Nevertheless, the treatment facility in the country is poor with only one clinical psychologist for every 1.3 million people and one psychiatrist for every 330,000. It is essential to diagnose mental health related problems at the onset, for the patient to lead a normal life. Existence of a machine learning based decision-making model, thus, would be helpful at this juncture for both the patients and the clinicians.

Srividya, Mohanavalli and Bhalaji (2018) proposed a system to profile the behavior of an individual. Their target population was divided into two groups – 300 high school/ college students and 356 young professionals with less than five years' experience. They prepared a set of 20 questions to identify the levels of engagement, happiness, perseverance, optimism, and connectedness. Based on the scores obtained, an individual could be labelled as optimistic, barely satisfied with life, or mentally distressed. Various machine learning algorithms such as SVM, RF, KNN, Naïve Bayes, Decision Tree, and Bagging were used for the classification. 80% data were used for training and the rest were used for testing. The performance of the classifiers was analyzed using the metrics like accuracy, precision, specificity, and F-score. They concluded that the best performance was obtained with SVM, KNN, and RF because all the metrics had value greater than 0.8 for all the three classes.

3) Diagnosis of depression may not always be easy as it may have similarities with some of physical ailments. A person is said to have depression comorbidity if he/ she suffers from both depressive disorder and other form of disease. Though there had been lot of research in the field of application of machine learning in the diagnosis of depression, there had been very little research how the same techniques can be applied when comorbid illness exists. One of the limitations in such research was the lack of depression dataset in comorbid population. A multi-dimensional classification (MDC) is required to evaluate the correlation between the depression and the comorbid illness, recognizing and evaluating both simultaneously.

Ojeme and Mbogho (2016) proposed a multi-dimensional prediction model based on Multi-dimensional Bayesian Network Classifier (MBC) that is capable of providing a consistent diagnostic result. 1090 data

https://doi.org/10.46977/apjmt.2022v03i02.004

instances were collected through semi-structured interviews and previously diagnosed cases. 22 main features such as sad mood, suicidal tendency, psychomotor agitation, weight loss/ gain, insomnia etc. were identified. Two class labels – depression diagnosis and depression comorbid, were assigned. Performance of the classifiers were evaluated using Hamming score, Hamming loss, and Exact-match. Best result was obtained with Bayesian Classifier Chain (BCC), and Probabilistic Classifier Chain (PCC) with Hamming score = 0.91, Hamming loss = 0.09, and Exact-match = 0.826.

Multi-modal System Based Approach

*1)*To ensure treatment and recovery early diagnosis of depression is crucial. Researches had revealed that machine learning algorithms can identify initial signs which might otherwise be difficult to detect. This can be achieved using wearable intelligent devices or other smart devices that can interact with the patient. Rastogi, Keshtkar and Miah (2018) proposed a fully-automated, reinforced learning Human-Robot Interaction (HRI) system based on the cognitive behavioural therapy (CBT) model for the recognition of early signs of depression. The robot-interaction framework is comprised of five modules:

- Multi-modal data – Patient data can be collected either by direct interaction with the robot or data from different sources can be uploaded.
- Social robot and dialogue system – Patients with early symptoms can communicate with the robot.
- Psycho-linguistic features – Consists of prosodic features, facial expressions as obtained from video, and sentiment expression from the transcript.
- Machine learning and natural language processing (NLP) methods – Facial feature analysis and NLP using various machine learning algorithms.
- Psychology feedback – Needed to review a patient's behavior.

This system is still in proposal state and all the modules are yet to be fully implemented.

*2)*Ray *et al*. (2019) used behavioural clues to automate the diagnosis of depression and identify the stage of illness. Features were extracted from analysis of emotion from facial expression, audio clips, and text data containing depressive words mined from social media platforms. They had used Extended Distress Analysis Interview Corpus dataset (E-DAIC) which contain audio-visual recording of patient's interviews where they described various psychological distress conditions. Different regression models were created for audio, video, and text modalities to foretell the severity of depression.

*3)*Tsiakas *et al*. (2015) proposed a system called Adoption Multimodal Dialogue System for Depression and Anxiety (DADS) which interacts with an individual to extract the requisite information. The system perceives multimodal input – text, audio, and facial expression from the video clips. They performed a multi-level screening to keep the number of questions to a minimum. Screening takes place incrementally depending upon the score obtained at each level. For the system implementation Robot Operating System (ROS) was used for Human Computer Interaction (HCI). Audio emotion recognition was done using Hidden Markov Model (HMM), and facial recognition was classified using Python Wrapper for Indico. The algorithm was trained using Reinforcement Learning.

Discussion

The survey done reveals that in recent time there is a growth in ML application in the domain of mental health. Majority of the past researches described methods of detection and diagnosis, but few used ML based approach to understand mental health behaviours. One of the reasons might be the lack of availability of sufficient and diverse data. Though a large number of data can be obtained from health records, the access is often restricted. Number of people who take part in the study are low. As a result of these restrictions the study design is hampered, leading the researchers to access the public database or social media. Such data may be suboptimal and also lacks the clinical validation. A better access to data is needed for automating the diagnostic process.

https://doi.org/10.46977/apjmt.2022v03i02.004

It was found that patients sometimes feel uncomfortable to share their social-media data as there might be sensitive or personal information. Though such personal data are required for developing a robust ML model there are issues of how to safeguard that the participants' sensitive information. Due to the difficulties in accessing data, proposals have been made to create benchmark repositories, the access to which can be reviewed by an ethical committee.

The majority of the papers consulted so far concentrated on finding whether a subject belongs to a diagnostic category. Classifying into one broad category, however, does not take into account the variability of the symptoms. Psychiatric practitioners need differential diagnostic tools that can discriminate between several illness categories that have similar symptoms and can identify optimal treatments. ML approaches can be used for targeted adjustment of treatments. Jeong and Breazeal (2016) proposed Just-in-time Adaptive Intervention (JITAI), which aims to provide the right amount of intervention at the right time by adapting to the patient's changing needs. On the other hand, some of the existing ML-based models exhibited a high misclassification rate that made the interventions risky, e.g., StressMon (Zakaria, Balan & Lee, 2019), which was built to detect stress and depression at the earliest phase, had a high misclassification rate of 18.2%.

One of the main reasons for such high misclassification is the lack of normalisation in clinical screening. E.g., in the case of social media data analysis, self-confessed or sentence-based labelling by the participant does not conform to standardized clinical assessment. Also, a specific period of distress in one's life may not necessarily imply the presence of a mental health issue. Control groups included in the study are chosen randomly. More often than not their past mental health symptoms are not taken into account [Chancellor]. The output of the ML-based model that used publicly available datasets should also be interpreted cautiously. It was found that approximately half of these studies overestimated the accuracy of the algorithms because they used record-wise cross-validation [Saeb].

One of the major challenges in the design of an ML-based system is the cohort of clinically interpretable output. To attain this goal, model output and an explanation of the predictions are included in the model. Interactive visualization and user interface design play an important role in producing all-inclusive mapping.

Conclusion and Future Work:

The review given in this paper is not complete as the research works in this field are still in their infancy and more and more works are added each day. The authors have highlighted recent trends in the application of ML in the field of mental health. In the course of this study it was found that there are scopes of research in the following key domains:

1. Early detection and diagnosis of critical mental health conditions such as suicidal tendency arising out of depression, schizophrenia etc.
2. Diagnosis of less explored mental health conditions such as anxiety, personality disorders, neurodevelopmental problems etc.
3. Most of the research areas explored so far applied supervised classification techniques that made use of labelled datasets. There are scopes for research using less standard data for providing personalised, real-time ML-based evaluation and intervention.
4. ML-based models suffer from challenges due to quality and quantity of data. Creation of comprehensive dataset needs to be built by the collaboration between the researchers and clinical practitioners.

Acknowledgment:

The authors are sincerely grateful to Dr. Subhabrata Bhakta, M.D. in Psychiatry & Clinical Psychology, a renowned psychiatrist in Kolkata, for his inspiration and support.

https://doi.org/10.46977/apjmt.2022v03i02.004

Conflict of Interest:

The authors have no conflict of interest.

References:

Abuse, S. (2020). Key substance use and mental health indicators in the United States: results from the 2019 National Survey on Drug Use and Health.

Adamou, M., Antoniou, G., Greasidou, E., Lagani, V., Charonyktakis, P., Tsamardinos, I., & Doyle, M. (2018). Toward automatic risk assessment to support suicide prevention. *Crisis*.

Blain, L. M., Galovski, T. E., & Robinson, T. (2010). Gender differences in recovery from posttraumatic stress disorder: A critical review. *Aggression and violent behavior*, *15*(6), 463-474.

Chang, K. H., Chan, M. K., & Canny, J. (2011). Analyzethis: Unobtrusive mental health monitoring by voice. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1951-1956).

Chang, K. H., Chan, M. K., & Canny, J. (2011). Analyzethis: Unobtrusive mental health monitoring by voice. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (pp. 1951-1956).

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171-209.

Chen, X., Sykora, M. D., Jackson, T. W., & Elayan, S. (2018, April). What about mood swings: Identifying depression on twitter with temporal measures of emotions. In *Companion Proceedings of the The Web Conference 2018* (pp. 1653-1660).

De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. (2016, May). Discovering shifts to suicidal ideation from mental health content in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2098-2110).

Fatima, I., Mukhtar, H., Ahmad, H. F., & Rajpoot, K. (2018). Analysis of user-generated content from online social communities to characterise and predict depression degree. *Journal of Information Science*, *44*(5), 683-695.

Feng, C., Gao, H., Ling, X. B., Ji, J., & Ma, Y. (2018, March). Shorten bipolarity checklist for the differentiation of subtypes of bipolar disorder using machine learning. In *Proceedings of the 2018 6th International Conference on Bioinformatics and Computational Biology* (pp. 162-166).

Frogner, J. I., Noori, F. M., Halvorsen, P., Hicks, S. A., Garcia-Ceja, E., Torresen, J., & Riegler, M. A. (2019, October). One-dimensional convolutional neural networks on motor activity measurements in detection of depression. In *Proceedings of the 4th International Workshop on Multimedia for Personal Health & Health Care* (pp. 9-15).

Garcia-Ceja, E., Riegler, M., Jakobsen, P., Torresen, J., Nordgreen, T., Oedegaard, K. J., & Fasmer, O. B. (2018, June). Motor activity based classification of depression in unipolar and bipolar patients. In *2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS)* (pp. 316-321). IEEE.

Healthline (2nd November 2021). Could It Be Bipolar Disorder? Signs to Look For. https://www.healthline.com/health/could-it-be-bipolar-signs-to-look-for

Jeong, S., & Breazeal, C. L. (2016, October). Improving smartphone users' affect and wellbeing with personalized positive psychology interventions. In *Proceedings of the Fourth International Conference on Human Agent Interaction* (pp. 131-137).

Joshi, D. J., Makhija, M., Nabar, Y., Nehete, N., & Patwardhan, M. S. (2018, January). Mental health analysis using deep learning for feature extraction. In *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data* (pp. 356-359).

Lovejoy, C. A. (2019). Technology and mental health: the role of artificial intelligence. *European Psychiatry*, *55*, 1-3.

Mallol-Ragolta, A., Dhamija, S., & Boult, T. E. (2018, October). A multimodal approach for predicting changes in PTSD symptom severity. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (pp. 324-333).

https://doi.org/10.46977/apjmt.2022v03i02.004

Mayo Clinic. (2020). Schizophrenia. https://www.mayoclinic.org/diseases-conditions/schizophrenia/symptoms-causes/syc-20354443

McClellan, C., Ali, M. M., Mutter, R., Kroutil, L., & Landwehr, J. (2017). Using social media to monitor mental health discussions− evidence from Twitter. *Journal of the American Medical Informatics Association*, *24*(3), 496-502.

Mitra, V., Shriberg, E., McLaren, M., Kathol, A., Richey, C., Vergyri, D., & Graciarena, M. (2014, November). The SRI AVEC-2014 evaluation system. In *Proceedings of the 4th International Workshop on Audio/Visual Emotion Challenge* (pp. 93-101).

Moturu, S. T., Khayal, I., Aharony, N., Pan, W., & Pentland, A. (2011, August). Sleep, mood and sociability in a healthy population. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5267-5270). IEEE.

Nobles, A. L., Glenn, J. J., Kowsari, K., Teachman, B. A., & Barnes, L. E. (2018, April). Identification of imminent suicide risk among young adults using text messages. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 1-11).

Ojeme, B., & Mbogho, A. (2016). Selecting learning algorithms for simultaneous identification of depression and comorbid disorders. *Procedia Computer Science*, *96*, 1294-1303.

Rastogi, N., Keshtkar, F., & Miah, M. S. (2018, October). A multi-modal human robot interaction framework based on cognitive behavioral therapy model. In *Proceedings of the Workshop on Human-Habitat for Health (H3): Human-Habitat Multimodal Interaction for Promoting Health and Well-Being in the Internet of Things Era* (pp. 1-6).

Ray, A., Kumar, S., Reddy, R., Mukherjee, P., & Garg, R. (2019, October). Multi-level attention network using text, audio and video for depression prediction. In *Proceedings of the 9th International on Audio/Visual Emotion Challenge and Workshop* (pp. 81-88).

Saha, K., & De Choudhury, M. (2017). Modeling stress with social media around incidents of gun violence on college campuses. *Proceedings of the ACM on Human-Computer Interaction*, *1*(CSCW), 1-27.

Shatte, A. B., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychological Medicine*, *49*(9), 1426-1448.

Spathis, D., Servia-Rodriguez, S., Farrahi, K., Mascolo, C., & Rentfrow, J. (2019, July). Sequence multi-task learning to forecast mental wellbeing from sparse self-reported data. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2886-2894).

Srividya, M., Mohanavalli, S., & Bhalaji, N. (2018). Behavioral modeling for mental health using machine learning algorithms. *Journal of medical systems*, *42*(5), 1-12.

Truschel, J. & Tzeses, J. (22nd December 2021). Top Mental Health Apps: An Effective Alternative for When You Can't Afford Therapy?. https://www.psycom.net/25-best-mental-health-apps

Tsiakas, K., Watts, L., Lutterodt, C., Giannakopoulos, T., Papangelis, A., Gatchel, R., ... & Makedon, F. (2015, July). A multimodal adaptive dialogue manager for depressive and anxiety disorder screening: a Wizard-of-Oz experiment. In *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (pp. 1-4).

van den Broek, E. L., van der Sluis, F., & Dijkstra, T. (2013). Cross-validation of bimodal health-related stress assessment. *Personal and Ubiquitous Computing*, *17*(2), 215-227.

Walker, J. (2015). Can a Smartphone Tell If You're Depressed?. *The Wall Street Journal*.

World Health Organization. (2001). The World Health Report 2001: Mental health: new understanding, new hope.

Wu, S., Falk, T. H., & Chan, W. Y. (2011). Automatic speech emotion recognition using modulation spectral features. *Speech communication*, *53*(5), 768-785.

Yates, A., Cohan, A., & Goharian, N. (2017). Depression and self-harm risk assessment in online forums. *arXiv preprint arXiv:1709.01848*.

https://doi.org/10.46977/apjmt.2022v03i02.004

Yazdavar, A. H., Mahdavinejad, M. S., Bajaj, G., Romine, W., Sheth, A., Monadjemi, A. H., ... & Hitzler, P. (2020). Multimodal mental health analysis in social media. *Plos one*, *15*(4), e0226248.

Zakaria, C., Balan, R., & Lee, Y. (2019). StressMon: scalable detection of perceived stress and depression using passive sensing of changes in work Routines and group interactions. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 1-29.