e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

A COMPREHENSIVE REVIEW ON DEPRESSION DETECTION AND MENTAL HEALTH TRACKER USING MACHINE LEARNING

Sunil D. Kale*1, Gaurav Rasal*2, Pratiksha Ganjave*3, Rushikesh Markad*4, Yash Kalekar*5

*1Affiliation, Professor, Department Of Computer Engineering, Pimpri Chinchwad College Of Engineering, Pune, India.

*2,3,4,5Affiliation, Student, Department Of Computer Engineering, Pimpri Chinchwad College Of Engineering, Pune, India.

DOI : https://www.doi.org/10.56726/IRJMETS37194

ABSTRACT

Depression is very commonly found and a quite serious type of mental disturbance. There is one thing about depression is it's one of the curable medical illnesses but the worst thing about depression is that the people who are suffering from depression won't get to know that they are suffering from some sort of mental illness. So, a system to predict depression in its early stage can be a savior for people and they can consult a psychiatrist and find a suitable cure for it. Here, we presented a comprehensive review on depression detection and mental health tracker with conclusion and future scope to explorer this topic further. Moreover, we are proposing a model which will take more than one input parameter for depression prediction to achieve better result of accuracy.

Keywords: Mental Health, Depression, Mental Balance, Increased Productivity, Psychiatrist, Psychology, Machine Learning, Natural Language Processing (NLP), Deep Learning, Text Processing.

I. INTRODUCTION

Depression, also known as Major Depressive Disorder (MDD), is a mental illness that affects people all around the world is a major mental illness found in one of every 15 grown-ups every year which is about 6.7% and around 16.6% of people will come across some sort of depression at a certain stage in their life. There are various reasons behind depression like overthinking, loss of money, unemployment, isolation from social activities and many more. One of the major reasons is that things are not going or happening in the way that people expect. Depression is very common amongst teenagers and people in their mid-twenties.

Depression leads to lower productivity, lack of enthusiasm for doing daily tasks, feeling of worthlessness, feeling of sadness, inability to think and concentrate and change in appetite. The surrounding environment where these people work or live also gets affected by their behaviour. People suffering from depression harm themselves and in the worst cases, depression can lead them to commit suicide.

So, to address this problem we're proposing a model which can detect depression in its early stage which is the major objective of this paper and the depression can be cured and people can live their life happily. This proposed model can be integrated with smart watches, fitness bands and also with mobile phones which are common amenities for most people and can be made easily available.

II. LITERATURE REVIEW

According to research done by authors in [1], The paralinguistic features of depressed people are different from the non-depressed. So, they did an analysis of these features. Data collection and preprocessing was done firstly. The feature extraction from the audio file data was done by applying the MFCC. The MFCC file was then fed to the classification model which was implemented by using CNN.

In [2], authors fabricated an online model that permits the investigation of multiple features of users concerning two particular mental illnesses. That model also provides minimal outputs which further can be utilized to develop more composite models for understanding a user's mental health virtually. In advance, the model can be utilized to gather additional data of users by a return feedback.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

The authors of the paper [3] investigated and compared various algorithms for analyzing social media texts and posts of users in order to predict depression among them. Peter Norvig's code for spelling correction and Natural Language Toolkit components were used. The experiments were carried out using a 10-fold CV method.

The conclusion was that of the two datasets used, Shen et al's dataset and Eye's dataset; Shen et al's predicted depression more accurately and is a better choice. They also discovered that some changes to the dataset were required to reduce overfitting in the model.

In [4], authors have proposed an automated Major Depressive Disorder(MDD) detection system based on deep learning. The model was set up on spectrogram images Convolutional Neural Networks(CNN). By using hold-out validation strategy they were able to obtain 99.58% of classification accuracy The limitation of this model was it was trained on a small set of data of only 64 persons which were further classified as 34 MDD patients and 30 healthy persons.

According to a study in [5,] The MHDeep framework was suggested by the authors. The system enables continuous and universal diagnosis of schizoid personality, major depressive illness, and bipolar disorder by combining data retrieved from commercial warehouse management software with the knowledge purifying power of DNNs. The concept makes use of a synthetic data generation module to alert users when huge datasets are unavailable. Iterative internet backbone growth and pruning was employed to decide on the model's architecture and strength training during training. They used data collected from 74 persons to evaluate the model. The experiment's findings showed that the models were effective in terms of computation. The MHDeep may be used for routine diagnostics and everyday monitoring due to its computational complexity efficiency and accuracy.

In the paper [6], The researchers have studied online depression communities and considered their differential factors to other online communities. Affect, psycholinguistic processes and topics within content were the contexts under the examination. Machine learning and statistical methods were employed to differentiate online messages between depressed and control communities. All the studied aspects were observed to be significantly different between these two groups.

The researchers in [7] have suggested a Novel stratified DL network that combines various multiple associated layers to consolidate people's discernible depiction and end user posting. In their system Depression detection was done by using online social media data as an input which was divided by summarizing relevant user's social media post history to axiomatically select the markable user data. The axiomatic summarization provided one advantage of being able to focus only on the most markable data in the model training which crucially helped to reduce the Hughes phenomenon problem. Which helps to provide better contingent knowledge or information to the model that helps to focus on the task where the model provides inputs instead of detecting depression only.

In [8], the authors have processed the information from social media for sentiment analysis for foreboding and sadness detection by applying Artificial Intelligence techniques. They observed that the Deep Learning with Multi Class Classification showed the higher precision value for exploration. The amassed data was used to understand the feelings, discernment and replications which are being conveyed by them. These analyzed sentiments were used to predict human deportment. The data was bifurcated into three classes positive, negative and neutral.

During the study, they concluded that the emoticons and emojis present in the data are the principal parameters for predicting the sentiments of users.

In the paper [9] they have offered a wide view of mental health computing through garnering social media. In this, they systematically studied the issues with two classic mental health problems, like stress and depression. They developed a set of benchmark physical world datasets, defined several distinct feature groups and proposed a series of effective detection models. They prepare in-depth data analysis to reveal the underlying on-line behaviors below the mental health problems.

In [10], the sociological habits and views of twitter users were used by the authors for the prediction of depression. The user was within 2 months of completing the Patient Health Questionnaire-9. This was used as the final measurement. They implemented the SVM-RFE and ANOVA feature selection methods for improving

performance of their model. Also the Random Forests technique among the other machine learning techniques given the highest accuracy for the prediction. The limitation of their model is it processed the information of the twitter account holders who had agreed with the disclosure, for the model.

In [11], they have reviewed diverse machine learning techniques and algorithms for depression prediction. This study considered the verbal and non-verbal actions of depressed and non-depressed persons.. They reviewed data assembly and data contrivances for developing a prediction model. The model was not standalone diagnostics, but it was profitable for distant evaluation, support systems and awareness. It provides many intuitions and also acknowledges many questions for the future investigation.

The reference [12] presents a new technology for personalized behavior classification. The technology gives the plan from memory based conjunct filtering and utilizes the behavior interrelation square to gather the behavior applicability among individuals. It mixes relevancy metrics and association rule mining to get unique behavior rules. They incorporated their technology on a passive sensing dataset collected from up to 97 undergraduate students over a semester. Whose depression symptoms were calculated by their BDI-II score at the end of semester. They further assessed the method by duplicating the pipeline on other datasets gathered from different institutions to obtain similar results. And, the technology also generated thoroughly illustratable rules that get the homogeneity and the heterogeneity in students' behavior related to depression.

Text classifiers were trained for detection of depression in [13]. Comparison of hybrid and ensemble methods was done to improve the performance. The conclusion of the research was ensemble models provided better results than the hybrid and deep learning based classification models. The Reddit social media dataset was used for the study. Also they observed that the models utilizing more than one lexical features were performing better than the models which rely on a single lexicon feature.

A full-fledged IIS based on pattern detection is suggested by the study [14] to produce useful features for video processing across the entire movie. STFAM, SPP, a programmable module, and a 3-dimensional block were used to "encode" the multi-scale patterns prevalent throughout the facial areas into an attractive representation.

The approach was validated using two independent depression datasets, and the researchers found encouraging estimation results. The primary goal of the study was to develop multimodal fusion techniques based on pattern recognition for the evaluation and diagnosis of depression in the long term.

In paper [15] the authors have explored and researched various sets of features for the sake of depression prediction in social media. In order to differentiate reedit users, they worked with the text of their messages and investigated how bag-of-words, word embedding and bigrams works with 2017 data. The applicability of some extra features was evaluated for their applicability. The additional features improve the classification results in most of the cases. They achieved decent results in comparison with the CELF/eRisk 2017 task reports with 2017 data. The future scope is to apply semantic role labelling to explore some other features and test different classification models to get better output.

The paper [16] describes correspondence of spatial cubes to generate brain networks as starting features for detecting depression in healthy individuals. The authors proposed a Machine Learning workflow to extricate features and to differentiate MDD vs control on sMRI and FMRI. The major focus of the workflow was to utilise the cube resemblance as starting feature. While most studies pass over balancing the dataset, the author intercepted a more challenging problem of the exceedingly imbalanced datasets.

The purpose of the article [17]'s author was to examine the use of mobile phones in observing depression in a clinically depressed adolescent population. The author presents the findings of the mobile phone and virtual utilisation assessment for Depression (SOLVD)-Teen experiment. There were two main hypotheses. First, they postulated that a combination of mobile phone sensor use and patient self-reports of daily behaviour could predict depressive states when compared to established clinical psychometric tools like the Patient Health Questionnaire-9 (PHQ-9), Hamilton Rating Scale for Depression (HAM-D), and Hamilton Anxiety Rating Scale (HAM-A). Second, as parents usually assist in their children's upbringing, parental reports on the teen's mood might be a crucial factor in monitoring and foretelling the teen's depressive status.

The [18] explores the availability of an Android application. This app was constructed to track features in the domains namely behavior, coping, social history, cognitive majors, and lifestyle vitals related with possibility

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

for depression, as well as variables that might reduce the risk. The association between the domain variables and base depression, standard daily behavior of the 14-day study time, and mood changes were tested to assess the value of the domain variables in the application. The usability and effective recording and capturing of data of the health and wellness UDTracker app was tremendously successful. The opinions of users about the app were easy to use, great user experience and straightforwardness of the functionality.

The results of paper [19] indicate that the primary indicators of children's mental health are behavioural and emotional problems, which are measured as both dimensional and categorical events, a known need for professional assistance with emotional or behavioral problems, and level of functioning, which is represented by academic progress and social ability. The severity of the mental condition is directly impacted by the inherent persistence of behavior problems associated with it (number, intensity, frequency).

The paper [20] focuses on benefits and limitations of AI technologies as a medium for recognition and interference, applicable for mental health problems. Chabot's are a systematic way of bringing mental health services via a module. For providing idiomatic solutions to mental health problems Precision Therapy and Diagnostic modules are proven effective. Although it would be early to say that these platforms will 'replace' the 5 need for one-on-one, unique services, they certainly display a viable pathway to personalized mental care and are a main part of the mental health outlook.

Table 1. Overview of different studies

Ref	Year	Area focused	Algorithms reviewed	Limitations	Accuracy	Dataset used
[1]	2022	Use of speech-language recognition to improve the detection of depression in Arabic speech.	Short-Time Fourier Transform (STFT) Constant-Q Transform (CQT)	Only Audio file is acceptable	98%	Audio dataset Of depressed and non-depressed persons
[2]	2016	For the detection of mental disorder in a user, especially twitter users, predictive models were used which rely on speech and behavioural patterns.	Term frequency-inverse document frequency (TF-IDF) Pattern of Life features (PLF)	Applicable for limited mental disorders.	95%	BPD and BD patient datasets.
[3]	2021	The goal of this study was to determine whether machine learning can be used to analyse social media posts to identify	Support Vector Machine (SVM), Multilayer Perceptron (MLP), Logistic Regression (LR)	The scenarios where the algorithms won't perform well were identified	99.80%	Eye's dataset and Shen et al. Dataset.

		depressive symptoms in users.				
[4]	2022	A deep learning system is learned using spectrogram pictures derived from the signals of EEG.	Short-Time Fourier Transform (STFT), F1-score, Artificial Neural Network (ANN), Support Vector Machine (SVM).	Limited CPU memory. 2D-CNN-based systems are computationally challenging and need long learning time.	98.11%	A dataset containing 15 MDD people 15 healthy people.
[5]	2014	Study the characteristics of online depression communities in comparison with those joining other online communities.	Lasso Regression	Psycholinguistic features were chosen instead of latent topics which have a better predictive power.	N/A	The dataset was built using social media posts extracted from LiveJournal.
[6]	2022	combining information obtained from economically obtain WMSs with DNN knowledge refining power For endless and widespread diagnosis of three important mental health disorders, schizoaffective disorder, major depression, and bipolar disorder are included.	Grow-and-prune synthesis Decision tree Naive Bayes	Applicable to small datasets	70%-80%	Instagram photos Phone call logs, microphone
[7]	2021	Learning the result of social media to speculate depression	Deep neural network (DNN),Recurrent Neural Network	-	90%	Pedregosa et al. Shen et al. Chung et al.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

			(RNN), Natural language processing (NLP)			
[10]	2022	Depression speculates using sociological characters and feedback of twitter end users.	ANOVA, SVM-RFE,Decision Tree, Support Vector Machine, Random Forests, Naive Bayes	Cannot get information from end users who do not want to reveal their data.	ANOVA-76.57% SVM RFE - 76.46%	Twitter users dataset
[13]	2022	Comparison of hybrid and ensemble models for depression network.	LSTM Networks, Logistic Regression	Scope of improvement for the values of evaluation metrics.	Ensemble Model - 75%	The Reddit social media dataset
[14]	2021	A 3-d CNN model equipped with STFAM on audio-visual emotion challenge to model closely depression related discriminative patterns.	STFAM (Spatiotemporal feature aggregation module)	Multimodal fusion (audio, video, text, etc.) scheme based on pattern recognition for depression recognition and Analysis can be used.	8.42 RMSE and 6.78 MAE on the test set.	AVEC2013 and AVEC2014 datasets

III. RELATED WORK

A number of researches are done and algorithms are developed for detecting depression and other mental state related problems differing in their use cases and also the form of input they accept for predicting the results. The prior work completed in depression detection is discussed in this section. In Machine learning, wherever the system is capable of learning by itself while not like any programming thus it improves machine potency by sensing and describing drive information consistency and pattern. Within the existing system, depression detection was achieved by taking various inputs like audio files, questionnaires and through face expression. Past work on detecting depression has largely relied on messages, posts and tweets from social media, using either a approach based on textual features or a descriptive-based features approach. Linguistic features of the social media text, like words, n-grams, POS, and other linguistic characteristics, are the focused features in textual-based features. Some algorithms which were used for the prediction of the depression are a collection of some manifested methods of preprocessing, highlighting an ML algorithm from the preceding studies like DT, RF, Naïve Bayes, SVM, and KNN were used to forecast stress, anxiety, and depression. The Great Depression, Anxiety, and Stress Scale questionnaire (DASS 21) examined 348 individuals's data. The analysis designated

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:04/April-2023 Impact Factor- 7.868 www.irjmets.com

that Naïve Bayes achieved the best accuracy of 85.50% for forecasting depression. The approach followed by the researchers was mainly based on the stages shown in the following **Figure 1.**

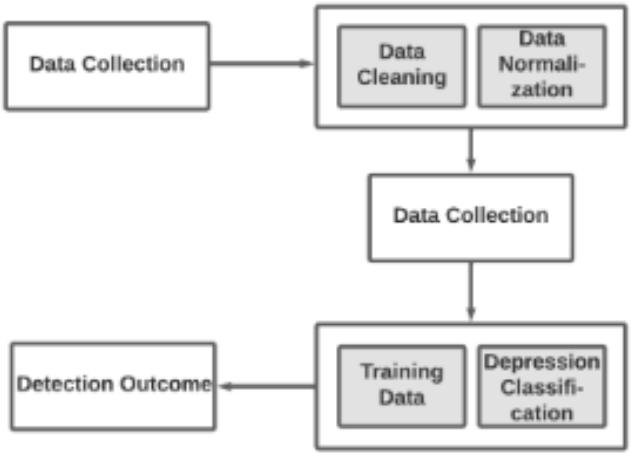

Figure 1: Depression Model

IV. RESULT

By studying and analyzing all the prior work done in the field of depression and other mental issues detection using Machine Learning, Deep Learning and CNN we came to know that the researchers have focused on only one type of input for the prediction. To achieve better accuracy in our model we have planned to use more than one input parameter for the processing and predictions will be based on the combined results from the model.

V. CONCLUSION

Mental health is a difficulty that's both delicate and vital at the instant. It's necessary for a healthy and balanced lifestyle. A mental state has an influence on one's thoughts, actions, and feelings. It can have an impression on a person's productivity and effectiveness. This paper elaborates our approach to detecting signs of depression using ML technology and algorithms that may be productively effective. Our ML systems can be proven productive for the early detection of a user's mental state and suggest some tasks to recover the user from that state.

VI. REFERENCES

[1] Depression Detection in Arabic Using Speech Language Recognition Zainab Alsharif, Salma Elhag, Sulhi Alfakeh

[2] MIDAS: Mental Illness Detection and Analysis via Social Media Elvis Saravia, Chun-Hao Chang, Renaud Jollet De Lorenzo, Yi-Shin Chen

[3] A textual-based featuring approach for depression detection using machine learning classifiers and social media texts, Raymond Chiong a,*, Gregorius Satia Budhi a,b,**, Sandeep Dhakal a, Fabian Chiong c.

[4] Decision Support System for major depression detection using spectrogram and convolution neural network with EEG signals,Hui Wen Loh1, Chui Ping Ooi1, Emrah Aydemir4, Turker Tuncer5, Sengul Dogan5, U Rajendra Acharya1,2,3,6,7,*

[5] Affective and Content Analysis of Online Depression Communities Thin Nguyeny, Dinh Phungy, IEEE Member, Bo Daoy, Svetha Venkateshy, IEEE Fellow, Michael Berkzy

[6] MHDeep: Mental Health Disorder Detection System based on Wearable Sensors and Artificial Neural Networks Shayan Hassantabar, Joe Zhang, Hongxu Yin, And Niraj K. Jha, Princeton University, USA

[7] DepressionNet: A Novel Summarization Boosted Deep Framework for Depression Detection on Social Media Hamad Zogan, Imran Razzak, Shoaib Jameel, Guandong Xu

[8] Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review Nirmal Varghese Babu1 · E. Grace Mary Kanaga1

[9] Mental Health Computing via Harvesting Social Media Data. Jia jia.

[10] Big data analytics on social networks for real-time depression detection Jitimon Angskun1,2, Suda Tipprasert1 and Thara Angskun1,2*

[11] On the review of image and video-based depression detection using machine learning Arselan Ashraf1 , Teddy Surya Gunawan2, Bob Subhan Riza3, Edy Victor Haryanto4, Zuriati Janin5

[12] Leveraging Collaborative-Filtering for Personalised Behaviour Modelling: A Case Study of Depression Detection among College Students, Xuhai Xu,Prerna Chikersal And Janine M. Dutcher, Yasaman S. Sefidgar, Woosuk Seo, Michael J. Tumminia, Daniella K. Villalba, Sheldon Cohen, Kasey G. Creswell, And J. David Creswell,Afsaneh Doryab,Paula S. Nurius, Eve Riskin, Anind K. Dey, And Jennifer Mankoff

[13] Ensemble Hybrid Learning Methods for Automated Depression Detection Luna Ansari , Shaoxiong Ji , Qian Chen, and Erik Cambria , Fellow, IEEE

[14] Intelligent system for depression scale estimation with facial expressions and case study in industrial intelligence Lang He1,2 | Chenguang Guo3 | Prayag Tiwari4 | Hari Mohan Pandey5 | Wei Dang

[15] Feature Engineering for Depression Detection in Social Media Maxim Stankevich, Vadim Isakov, Dmitry Devyatkin and Ivan Smirnov

[16] Depression detection from sMRI and rs-fMRI images using machine learning Marzieh Mousavian 1· Jianhua Chen 1 · Zachary Traylor 2 · Steven Greening3

[17] Tracking and Predicting Depressive Symptoms of Adolescents Using Smartphone-Based Self-Reports, Parental Evaluations, and Passive Phone Sensor Data: Development and Usability Study Jian Cao1, PhD; Anh Lan Truong2 , MD; Sophia Banu2 , MD; Asim A Shah2 , MD; Ashutosh Sabharwal1 , PhD; Nidal Moukaddam2 , MD, PhD

[18] Developing A Smartphone Application For Depression: Tracking Risk And Wellness Factors Garret Richard Sacco

[19] Tracking Children's Mental Health in the 21st Century: Lessons from the 2014 OCHS Michael H. Boyle, PhD1 , Laura Duncan, MA1,2 , Katholiki Georgiades, PhD1 , Jinette Comeau, PhD3,4, Graham J. Reid, PhD4,5, Warren O'Briain, MA6 , Robert Lampard, PhD7 , and Charlotte Waddell, MSc, MD, FRCP(C)8 ; 2014 Ontario Child Health Study Team9

[20] Mental Health Monitoring System using Artificial Intelligence: A Review Vidhi Mody1, Vrushti Mody 2.