Detection of Mental Disorder on Social Platform

Shreya Jain [1], Manasi Jadhav [2], Neha Kolambe [3] , Dr.Prasenjit Bhavathankar[4]
[1]Department of Information Technology, Sardar Patel Institute of Technology, Andheri, Mumbai, India, shreya.jain@spit.ac.in
[2] Department of Information Technology, Sardar Patel Institute of Technology, Andheri, Mumbai, India manasi.jadhav@spit.ac.in
[3]Department of Information Technology, Sardar Patel Institute of Technology, Andheri, Mumbai, India, neha.kolambe@spit.ac.in
[4]Department of Information Technology, Sardar Patel Institute of Technology, Andheri, Mumbai, India, p_bhavathankar@spit.ac.in

Abstract: Mental Disorders have expanded essentially since the world hit the pandemic. As the world is engaging with COVID-19, people likewise need to manage the increment in psychological wellness issues like depression, anxiety, and suicidal tendencies. While people have been working resolutely to keep their appearances covered, wash hands, and stay 6 feet away from everybody, including friends and family, people might not have acknowledged how the pandemic and isolation have worked on their emotional well-being. This exploration targets assembling a system that distinguishes any mental disorder. The detection will be completed by utilizing machine learning and deep learning techniques. Depression is seen as the biggest supporter of worldwide incapacity and significant justification for suicide. The target diminishing the number of suicides occurring, henceforth encouraging people to take the online test which would comprise of a review, and perceive their debilitate/lively state. Early detection makes way for future consideration and treatment. The aim is to build a model which connects the users to counseling services which will help clients to generate solutions to their problems and connect them to respective counselors.

1. Introduction

Mental Health, as a whole, is taken lightly in some countries. It affects a human the most. Even nowadays educational institutions don't provide adequate exposure to the same. Hence people never worry about getting professional help to overcome their illness. Some people are diagnosed but never seek help. They continue living which keeps getting worse. As per worldwide statistics, approximately 314 million people suffer from mental disorders. These numbers might have increased notably since the world hit the pandemic. The horrors of the virus are known to all. The way it has affected every human physically and financially is out in the open. But people never talk about how it has affected us mentally. Several people have slipped into depression or developed some serious illness, thanks to covid. While a huge number has been reported for the people who committed suicide due to various problems. In 2019, suicide became the 17th leading cause of death.

The project builds a system that analyzes tweets that have depressive features. Early detection makes the way for future consideration and treatment. Thus it will be possible for health care workers to analyze posts to detect poor mental health.

Tweets are scraped using TWINT which is a scraping tool. Tweets for different symptoms are scraped and stored. It will include loneliness, sadness, hopelessness, etc. The scraped tweets won't be in raw form. Thus they are sent to preprocessing. A typical tweet contains multiple attributes, out which tweet text is considered. The text is cleaned ahead. Sentiment Analysis is performed on the same and appropriate binary labels are given to each tweet.

Many people these days seek help through social media by pouring their hearts out. This results in either them getting genuine help or getting exploited. This raises the issue of having a system where people who need help can take a survey, discreetly.

Hence, the second module of the system is definitely dedicated to individual users, which generally is fairly significant. Further, a user can generally take an online survey in a really major way. This survey will particularly help them really know about their mental state and will mostly give them a kind of clear state of mind. User for all intents and purposes has to kind of select the basically appropriate answer and then really wait for the results to for all intents and purposes appear in a basically major way. This will generally help the user to mostly know about the early symptoms of a severe disorder.

Data collected from here will help in training the model. This is a traditional method of gathering reliable data. As a result, each symptom that is detected will assist the user in taking necessary action. The data collected will be authentic as online networks are a great source of data. The information thus gathered will be credible and will help us to build a robust model.

Further, section II has our research work, III has the proposed system which includes the Description, Comparison of Algorithms, IV has the implementation, V which has the results. Further, section VI has a conclusion followed by future scope in VII.

2. Literature Survey

The earliest work[1], using machine learning, created a framework to classify mental illness. Data is obtained from online posts. 12 categories of disorders were identified. Using Bayesian LDA, data was extracted. The Linguistic Inquiry and Word Count programme extracts the psycholinguistic processes. The discussed linguistic styles and subjects were the inputs to the linear regression model. Classification was performed on the 12 categories using Single Task Learning, Multi-task Learning and on the Proposed Framework. The findings show that they have a high predictive value for classifying mental illness with a depressive interest. This research [2] received assessment investigation methods to examine users' commitments to social platforms. It aims to predict possible depression early on. Two fundamental components of the framework, the information foundation with depressing assessment jargon and the sentiment investigation engine, remained research problems. A model that monitor's user tweets and alerts if depression is identified.

In [3] researchers have proposed a Social Network Mental Disorder Detector that uses data logs of online social networks to explore features. A new tensor technique is used to obtain hidden characteristics from multiple social network accounts for detection. Data Mining techniques are used to detect CR Addiction, NC, and IO. It is a joint effort between engineers and disorder analysts to resolve developing challenges in SNMDs. The goal of [4] is to see how user posts can be used to categorize users based on their mental health. The SNS-based system has the potential to solve the self-reporting issues. The proposed system uses social network accounts to source data and screening tools and forms a model that classifies the user granted content utilizing two distinct classifiers Naive Bayes and SVM. The classification models carried out, moderately giving 57% and 63% accuracy respectively.

Machine learning algorithms[5] are used to classify depressive symptoms in text format as early as possible.A major collection of Reddit remarks was utilized to prepare another fastText word embedding. The model has learned some domain-specific features and is suitable for general syntactic problems in the English language, as per the investigation of the resulting word vectors. The overall performance of numerous KNN classifiers[6] in detecting despair is studied. Four elements namely, emotional processing, temporal processing, linguistic style, and all aspects, have been investigated. The results demonstrate that in terms of different evaluation metrics, the dataset result and the KNN technique outcomes varied by 60-70 percent.

The proposed methodology [7] classifies people with persistent mental illness disorders. Data is obtained from Reddit. The proposed method employed a semi-supervised learning approach by consolidating widely used classifiers. For the advancement of an element space, the Reddit Application Programming Interface is utilized to download posts and the max five related remarks. The test outcomes appear that classifiers were superior to their individual use in terms of evaluating metrics with co-training technology. Online Posts from Reddit were analyzed to discover factors that hinted at depression. Natural Language Processing and ML techniques [8] were used for training the model and evaluating its efficiency. The work is done by identifying a glossary of words more common with depressed accounts. It is observed that a bigram with an SVM classifier has a 0.8 R2 score. It is increased when the mixed features (LIWC+LDA+bigram) are used with the MLP, reaching 91% accuracy.

The adequacy of the pre-mental wellness [9] detection is systematically evaluated, identifying its technique for data analysis, problems, and constraints. The systematic study also examines the suitability of pre-mental wellness detection by recognizing its data analysis method, problems, and constraints. Researchers suggest a ranking posts presentation model [10] recognizing sad persons in internet meetings. Two parts are included in the model: a future level activity and a client level activity. From word representations, it initially creates continuous post representations. Afterwards, by considering post representations as input, the user's generally passionate state portrayals are acquired. Depression is detected using user's posts [11]. A hybrid model has been proposed. The training data was given as input to the DL algorithms and the model was trained. Performance was evaluated on the Reddit dataset. It was observed that BiLSTM with word embedding techniques gave better results. Social platforms can supply relevant data for completing evaluation and determining depressive practices, as witnessed and experimented [12]. The predictions are better if more depressing words are collected which in turn gives more accurate output. Data Mining is used to predict disorders. There were four algorithms proposed and tested and sufficient results were obtained. Another major factor in this prediction may be the timing frequency for users posting and commenting. Tests are carried out to demonstrate that this study can provide the desired results from users detecting depressive conduct. The research work [13] works on different ML algorithms and classifiers. The examination is to distinguish the state of a nervous breakdown in an objective gathering (students or professionals). Twint, a Twitter scraping tool, is used for the detection of whether or not a particular tweet is depressive. It investigates various ways of foreseeing dejection and proposes an Automated Depression Detection system using DL mechanism.

The primary goal [14] was the development of a Social Anxiety Disorder (SAD) choice emotionally supportive network using an Adaptive neuro-fuzzy inference system approach. A data file with an example size of 210 was chosen and utilized to create the prototype. This technique involves pretreatment consisting of three steps.Classification is done using the ANFIS method using 5-fold cross-validation, and evaluation. It consisted of a multi-step method. Outcomes showed that the model proposed was well suited to the diagnosis of SAD and consistent with results from other studies. Researchers focussed on presenting the difference in characteristics between the users suffering from any illness [15] and from those who are not. They showed how these disorders differ in terms of their jargon. Results showed that individuals

with illness exhibit different behavior than those who don't, majorly in how they write and express their emotions.

3. Proposed System

3.1. Description

Python is utilized in all modules for basic programming. Chart.js is used for visualization. Validation is performed using JavaScript. Flask is used for hosting.

Tweets are scraped using the TWINT tool using specific keywords. This data is preprocessed and compiled.

VADER Analysis is used to determine the data's overall sentiment. It has 4 values - positive, negative, neutral, and compound. The positive, negative, and neutral determine the probability of a tweet to be positive, negative, neutral respectively while the compound calculates the sum of all lexicon ratings. The compound component is taken into consideration and used for further analysis.

NLP is also used for making tweets easier to understand. Bigrams are created for data visualization. They assist in determining which N-grams can be combined to form single entities. LSTM is used for predicting the probability of the input tweet to be good or bad.

Refer Figure 1 to understand the system and its working.

The first module detects disorder using tweets. Tweets are collected using different keywords. Preprocessing is done by removing redundant information. Once done, a single dataset is formed by stacking all the cleaned datasets.

Further, stopwords are removed from the tweets using NLP. Applied tokenizer to filter out the infrequent words. This makes the model better for classification and uses word2Vec embedding. It takes the minimum between the number of unique words in our tokenizer and max words. The data is divided and depressive and random tweets are combined. These are shuffled and given as input.

Further, the model is built which consists of 3 layers. It has an embedded layer, convolutional layer, and LSTM Layer, and the model is run and appropriate results are displayed. If the mental is detected here, the flow is passed to the next module i.e. the questionnaire part.

Second module focuses on the survey. Here, the user is presented with a questionnaire[16] on different types of mental health topics. The responses are analyzed and given as input to the model. If mental health is affected then, the user is given verified suggestions to take care of his/her health.

Users can talk about their current state of mind by writing it in a chatbox on our website. Second module focuses on the survey. Here the user is presented with a questionnaire on different topics. Here, the user is given a set of 30 questions which has questions on six different mental health disorders. The responses are analyzed and given as input to the model.

3.2. Comparative Analysis

The model is trained with various models such as Logistic Regression, Decision Tree, SVM[17], and Random Forest Classifier[18], and Text mining using tf-idf and bow models. The LSTM Model gave more promising results than other models. After considering it, hyperparameters were tuned by changing activation function, number of layers, loss function, and optimizer to get the best model for detection.

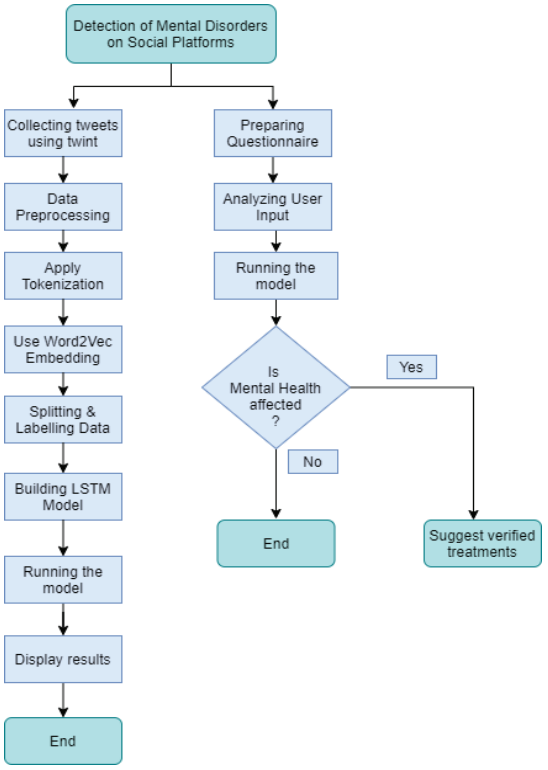

Fig. 1. *Flow Diagram of the system*

Table 1 shows the comparison of algorithms.

Table 1. *Comparative Analysis*

Model	Accuracy	Mean Square Error	R^2 Error
Logistic Regression	58.96%	0.41	-0.71
Decision Tree	73.4%	0.26	-0.11
SVM	70.09%	0.29	-0.25
Random Forest	94.1%	0.05	0.78
ReLU activation, Nadam optimizer, binary_crossentropy loss - Model with 3 layers	97.27%	0.03	0.88
softmax activation and kl_divergence loss - Model with 3 layers	59.93%	0.40	-0.66
tanh activation, Adam optimizer, categoriacal_crossentropy loss - Model with 3 layers	40.07%	0.59	-1.49
Model with 1 layer	61.63%	0.21	0.11
Final Model	97.27%	0.03	0.88

3.3. Long Short Term Memory

LSTM[19] is an artificial neural network mainly used for classification processing and making predictions. LSTM cells make up a typical layer Each cell examines a single column of its feed as well as the output of the LSTM cell in the preceding column. Tokenized Tweets are given as input to the model. It returns an embedding vector. The output of the model is a number. This represents the possibility that the tweet is related to good/bad mental health. Input tweet is replaced by its embedding vector. This

vector is run through the convolutional layer. The structure of the data is learned by this layer, which is subsequently passed on to the normal LSTM layer. The dense layer's input is the output from this layer. The model has 3 layers.

Figure 2 shows the working of a typical LSTM Model.

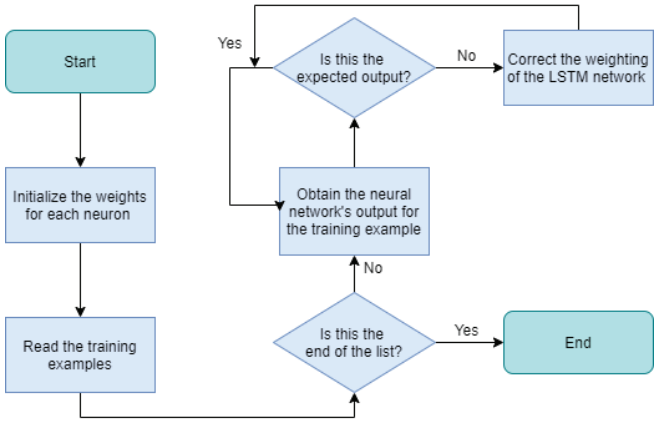

Fig. 2. LSTM Model

The performance of the model is evaluated by the following metrics - Mean Square Error and R^2 Error.

4. Results and Discussions

Tweets are scraped for 7 different keywords. Preprocessing is applied. Redundant columns like user_id, time, location are eliminated. URLs are removed using regex. Stop Words and Collection Words are also removed. Stopwords beyond the NLTK dictionary are also removed which includes days of the week and months. Emoticons are also eliminated.

Contractions like aren't, can't are expanded to are not, cannot. Punctuations, hashtags, and extra links are redundant data and are waived. Finally, the tweets are converted to lowercase. Plain text is given as input for VADER Analysis. Further, Porter Stemmer is applied for text normalization. It reduces words to their root form. Word2Vec Embedding is applied to generate vectors.

An embedding file, Google News, is loaded which is a pre-trained word2vec model. Another data set that has random tweets is also considered. The data is assigned its label from the preprocessed set and random tweets set. It is then split. It is ensured that all tweets are of the same length which is 280.

Model is built using layers from Keras library. The first layer is the Embedded Layer, then comes the Convolutional Layer, and then the LSTM Layer. Model is trained with the ReLu[20] activation function.

The model is trained and Early Stopping is used to end training if the loss and/or accuracy don't improve within 3 epochs. Accuracy is determined by using test samples.

User input is taken to see how well the model has performed. This text is preprocessed and tokenized. It is padded and then given to the model. Thresholding is applied for predicting the final output.

The model ran for 20 epochs. The dataset is given in a batch size of 16 with Nadam as an optimizer. The model achieved an accuracy of 97%. The other parameters such as precision, recall and f1 score exceeded 90%. For model evaluation, Mean square error and $R^{2,}$ Error is calculated.

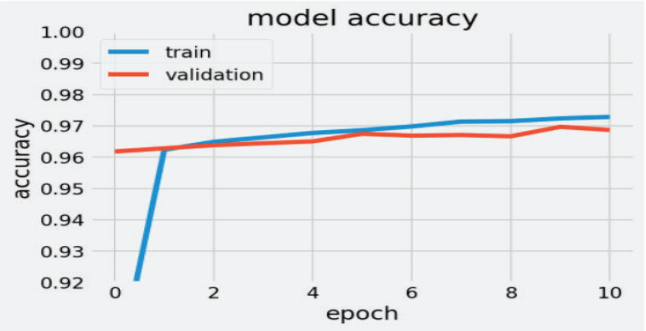

Fig. 3. Model Accuracy

Figure 4 shows the training and validation loss. As seen the loss decreases as the model is further trained.

Fig. 4. Model Loss

The user can enter real-time tweets and analyze the current mental health state. If the model shows good mental health you are good to go, else the model will prompt the user to take a questionnaire. The user can select the appropriate answer for the given question. Based on the response received, analysis is done that shows the user, which symptoms are they having for a severe disorder.

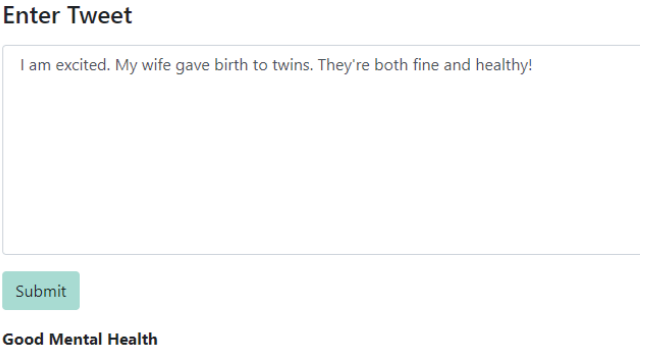

Fig. 5. Real-Time Tweet Detection

5. Conclusions

This is a study that analyzes the relationship between Twitter comment sentiment and semantic topics. However, the main aim of this research was to present a new system to identify meaningful potential topics.

The study explains that artificial intelligence takes over the healthcare industry. The use of verbal indicators as a method for analyzing and diagnosing mental illness has great potential. Poor mental illness can be recognized very quickly in the text. Standard collecting and cleaning data processes and a visual analysis alone can differentiate between a random tweet and a tweet with negative characteristics. Therefore, models built using various techniques and algorithms help to understand people's concerns and needs regarding these issues.

This paper aims at predicting the mental health of an individual using tweets and questionnaires. In the next step, later one can investigate the features derived from multimedia information and new difficulties from the standpoint of different social network service providers. Furthermore, the researchers can also utilize other alternatives e.g., photos, audio, and videos rather than only texts could be a promising field for future research.

References

[1] B. Saha, T. Nguyen, D. Phung, and S. Venkatesh, "A Framework for Classifying Online Mental Health-Related Communities With an Interest in Depression," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 4, pp. 1008-1015, July 2016.

[2] Advanced Data Mining and Applications, 2016, Volume 10086 ISBN : 978-3-319-49585-9 Xiaohui Tao, Xujuan Zhou, Ji Zhang, Jianming Yong

[3] Shuai, Hong-Han, Chih-Ya Shen, De-Nian Yang, Yi-Feng Carol Lan, Wang-Chien Lee, S. Yu Philip, and Ming-Syan Chen. "A comprehensive study on social network mental disorders detection via online social media mining." IEEE Transactions on Knowledge and Data Engineering 30, no. 7 (2017).

[4] M. M. Aldarwish and H. F. Ahmad, "Predicting Depression Levels Using Social Media Posts," 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS), 2017.

[5] Trotzek, Marcel, Sven Koitka, and Christoph M. Friedrich. "Utilizing neural networks and linguistic metadata for early detection of depression indications in-text sequences." IEEE Transactions on Knowledge and Data Engineering 32, no. 3 (2018).

[6] M. R. Islam, A. R. M. Kamal, N. Sultana, R. Islam, M. A. Moni and A. ulhaq, "Detecting Depression Using K-Nearest Neighbors (KNN) Classification Technique," 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), 2018.

[7] Tariq, Subhan, Nadeem Akhtar, Humaira Afzal, Shahzad Khalid, Muhammad Rafiq Mufti, Shahid Hussain, Asad Habib, and Ghufran Ahmad. "A novel co-training-based approach for the classification of mental illnesses using social media posts." IEEE Access 7 (2019).

[8] Tadesse, Michael M., Hongfei Lin, Bo Xu, and Liang Yang. "Detection of Depression-Related Posts in Reddit Social Media Forum." IEEE Access 7 (2019).

[9] Abd Rahman, Rohizah, Khairuddin Omar, Shahrul Azman Mohd Noah, Mohd Shahrul Nizam Mohd Danuri, and Mohammed Ali Al-Garadi. "Application of machine learning methods in mental health detection: a systematic review." IEEE Access 8 (2020)

[10] Rao, Guozheng, Yue Zhang, Li Zhang, Qing Cong, and Zhiyong Feng. "MGL-CNN: A hierarchical posts representations model for identifying depressed individuals in online forums." IEEE Access 8 (2020).

[11] F. M. Shah et al., "Early Depression Detection from Social Network Using Deep Learning Techniques," 2020 IEEE Region 10 Symposium (TENSYMP), 2020

[12] R. Vanlalawmpuia and M. Lalhmingliana, "Prediction of Depression in Social Network Sites Using Data Mining," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020

[13] P. V. Narayanrao and P. Lalitha Surya Kumari, "Analysis of Machine Learning Algorithms for Predicting Depression," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), 2020

[14] Sina Fathi, Maryam Ahmadi, Behrouz Birashk, Afsaneh Dehnad, Development and use of a clinical decision support system for the diagnosis of social anxiety disorder, Computer Methods and Programs in Biomedicine, Volume 190, 2020, 105354, ISSN 0169-2607.

[15] Ríssola EA, Aliannejadi M, Crestani F. Beyond Modelling: Understanding Mental Disorders in Online Social Media. Advances in Information Retrieval. 2020;12035:296-310. Published 2020 Mar 17.

[16] El Sayed, Samir, Sarah Gomaa, Shereen Aboelfotoh, and Mohamed El Wasify. ”Pandemic Grief Scale in detection of grief reaction among physicians in COVID-19 era.” Middle East Current Psychiatry 28, no. 1 (2021): 1-8.

[17] Mody, Vidhi, and Vrushti Mody. ”Mental Health Monitoring System using Artificial Intelligence: A Review.” 2019 IEEE 5th International Conference for Convergence in Technology (I2CT). IEEE, 2019.

[18] Kumar, V. Uday, Alekhya Savithri, M. Jhansi Bhavani, A. Madhu Priya, K. Venkata Sai Bindu Jahnavi, and Namburu Divya Naga Lakshmi. ”Finding Psychological Instability Using Machine Learning.” In 2020 7th International Conference on Smart Structures and Systems (ICSSS), IEEE, 2020.

[19] Jelodar, Hamed, Yongli Wang, Rita Orji, and Shucheng Huang. “Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions” IEEE Journal of Biomedical and Health Informatics 24, no. 10 (2020)

[20] Rao, Guozheng, Chengxia Peng, Li Zhang, Xin Wang, and Zhiyong Feng. ”A Knowledge Enhanced Ensemble Learning Model for Mental Disorder Detection on Social Media.” In International Conference on Knowledge Science, Engineering and Management, pp. 181- 192. Springer, Cham, 2020.