International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

# A Systematic Review on Strategies of Depression Detection from Social Media

**Mohammad Kaosain Akbar [1], Sk. Musfique Ahmed [2]**
[1]*Department of Computer Science and Engineering, United International University*
[2]*Department of Computer Science and Engineering, United International University*

***Abstract***
Depression is considered as one of the severe global challenges of our current society. Millions of people suffer from depression each day. Beside feeling sad and lose of interests, depressed person may face various heath complexities or may commit suicide. Therefore, depression is said to be a "Global Burden". Traditionally, psychiatrists conduct face to face interviews with a patient referring to depression criteria. Unfortunately, many patients feel shy to consult doctor which deteriorates their mental health even more. Additionally, everyday large number of people are sharing their thoughts and feelings on various social media platforms. Since people are spending huge amount of time online, researchers are actively linking online system with depression detection to understand the mental health of the individuals or even the entire population. So far, various efforts have been made to detect depression in users through their posts, tweets and messages on different social media platforms which generated successful outcomes. This paper reviews recent theories which were developed to detect depression from different social media. This survey aims in providing information about current depression detection methods, highlighting their core features, data source, findings and certain drawbacks, promoting further research in the field of depression detection and prevention.

***Index Terms- Depression, Dataset, Neural Network, Machine Learning, Sentiment Analysis, Facebook, Twitter, Reddit***

## I.INTRODUCTION

Depression is a common mental disorder. More than 350 million people suffers from depression [1]. According to World Health Organization, reportedly 8 million people commit suicide each year, resulting from chronic depression. Moreover, depression can affect health causing cardiovascular disease, cancer, diabetes and obesity. Depressed person also suffers from negative impacts on family, relationships, education, work, sleeping and even eating habits. It is estimated that within next 20 years, depression will be a leading disability among many first world countries such as United States. Advancement in Computational Science increases the chances of automatically detecting depressed users based on how they communicate, and social media is one of the vital mediums of communication.

Social media is an important part of everyone's life. Attractive content, photos and posts are shared between users in those social media platforms. These posts and contents being shared often predicts user's happiness, reactions to different events or unemployment. Therefore, social media can be characterized as an assessment tool for mental health. Often depressed people tend to feel shy in taking appointment at the psychiatrist or even expressing their thoughts and feelings to their family members. Social media allows depressed people to share their thoughts without coming into contact with other humans. In the recent years, numerous research was conducted on social media to understand the behaviors of depressed users. Social media gave researchers the scope to capture user behavior and activity in a natural setup so that the data obtained from the user do not have any experimenter demand effects or memory biasness.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

Over the years, researchers developed different depression detection techniques, majority of which showed an outstanding result with very high accuracy in depression detection. Various concepts Machine Learning and Neural Network theories were integrated in different models to identify and detect depression by analyzing text or images posted or shared on social media. This paper primarily focuses on discussing recent methods and strategies which were adopted, tested and used by different researchers around the globe to detect depression among users across multiple number of social media platforms. The paper presented the fundamental insights behind the implementation of each of these methods, source of the user-generated data, approach in training these individual models to detect depression from text or images successfully.

This study aims to deliver the answers to the following questions:
- What are the most common sources of depression data?
- What types of data are used in identifying depression?
- How data from the online platform is prepared for depression detection?
- What core techniques are used in extracting out information about depression various types of data?
- How data is further processed to provide better accuracy in detection
- How depressed data are distinguished from the normal data?

The reviewed literatures will provide a clear understanding to these research questions. In this paper, a systematic review is given to various online depression identifying strategies, a recommendation from author is presented and finally the study ends with a brief discussion and conclusion.

## II. RELATED WORKS

The literature review section is divided into two main parts. First part discusses about different principles and symptoms of Depression and the later part explains role social media as tremendous source of user-generated data.

### *A. Different Principles and Symptoms of Depression*

Before the internet, patients were diagnosed with depression based on numerous globally accepted survey questionnaires. CES-D Scale [2] was a well-known set of 20 questionnaires that focused on patient's feeling of guilt or sleeping information. Beck's Depression Inventory [3] comprised of 21 questions on psychological state of a patient. Majority of the questions were multiple choice and the total score obtained by an individual patient represented his/her depression level. The Diagnosis and Statistical Manual of Mental Disorders (DSM) [4] is another principal method of identifying depression which defined a total of nine different kinds of depressive indicators. The state of a patient's mental health is diagnosed by the symptoms using DSM.

### *B. Social Media as a source of user generated data.*

Introduction of social media changed the way people exchanged feelings, interests, opinions and emotions. Researchers and Psychiatrists became more interested in analyzing the user behavior in detecting depression through users' online activities. Various social media platforms namely Facebook, Twitter and Reddit acts as a source of data of different users' emotions, feelings and opinions and online behavior. Facebook was created in 2004 as website to just connect Harvard students with one another [5]. By the middle of 2019, Facebook had 2.5 billion users [6]. 500 million of stores are posted every day. On average, a user is connected to at least 155 people. Additionally, pictures, videos and different posts shared by individual Facebook users acts as a great source of identifying and understanding users' mental health and can be a vital medium of detecting depression online.

### *C. Machine Learning*

The field of Artificial Intelligence which occupies the analyzing of data that provides the capability to the computer-based device to learn without the need for someone to program the system is termed as Machine Learning. Data analysis helps computer systems to learn from both current and future

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

data. Machine Learning process seeks for hidden pattern and predicts how the data would behave in the upcoming days. Machine Learning is often considered like as Data Mining.

*D. Data Mining*

The process of analyzing and extracting data in different perspective so that we can get useful information out of the data. Even protecting future data is often done through the help of data mining. Therefore, Data Mining plays a vital role in aiding revenue of various data-driven sections

## III. METHODOLOGY

In this paper, we have collected, reviewed and assessed various research papers on depression detection strategies which have published in different journals and conference proceedings. We have defined our research questions and tried to reach an appropriate conclusion after reviewing those articles. Here we have explored the various studied techniques that are used to identify depression in online platforms over years of 2013 to 2020, with their success in identifying the depressed person. We were interested in identifying various algorithms that were used, how user data was processed to identify depression and what were the findings of the experiments.

The inclusion criteria of this study are:

- The publication year of the literatures on techniques for identifying depression must be within the year 2013 to 2020
- The literatures must focus on techniques used in identifying depressed person from online platforms.
- The studies must have findings and accuracy of the depression detection techniques through appropriate experiments.

In addition to the above inclusion criteria, this paper also has certain exclusion criteria which are:

- Depression identifying techniques must be computer-based and any literature which involved physical techniques to identify user depression were not reviewed in this paper.
- The reviewed literatures were written and published only in English Language.
- Although various books chapters were focused on detecting depression, most of them did not cover up procedures to detect depression completely online. Therefore, any chapters from books were not included in this Study

| **Name of the online library** | **Search Result: *"Online Depression Detection Technique"*** | **Search Result: *"Online Depression Detection"*** | **Number of Articles selected from the source** |
|---|---|---|---|
| ACM Digital Library | 219 | 305 | 8 |
| IEEE Xplore | 12 | 39 | 7 |
| Research Gate | 183 | 261 | 5 |
| | Total: 414 | Total: 605 | Total: 20 |

Table 1: Source of the Searched and Selected Literature

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

For this study, it was observed that 1019 research articles were available in different online digital libraries when searched with the search item "Online Depression Detection" and "Online Depression Detection Technique". After considering all the inclusion and exclusion criteria, 20 literatures among 1019, were selected for this systematic review. Table 1 shows the details about the sources of literature selection. The selected literatures showed direct contribution in depression detection from online platform. The literatures were collected from three databases – ACM Digital Library, Research Gate and IEEE Xplore. The articles were critically reviewed by the author to understand the different various ways of identifying depression among people in online platforms and how much successful and accurate each technique is in identifying depression among users

| **Authors** | **Datasets Sources** | **General Techniques** | **Output Accuracy** |
|---|---|---|---|
| M. Tadesse et al [7] | Reddit Posts | Neural Network and LDA | 91% Accuracy of Depression detection |
| Choudhury et al [8] | Twitter Tweets | SVM | 73% Accuracy of Depression detection |
| Ziwe et al [9] | Twitter Tweets | Cronbach Alpha formula | 70% Accuracy of Depression detection |
| Park et al [11] | Facebook Posts | CES-D* self-report | Used Standard CES-D questionnaires (100%) |
| Lin et al [12] | Twitter Tweets | CNN* and BERT | Accuracy is 88.4% and test accuracy is 93% |
| Wang et al [13] | Reddit Posts | CNN* and Scikit-Learn | Test accuracy is 65% |
| Maupome et al [16] | Reddit Posts | NLP and ANN* | 87.8% Accuracy of Depression detection |
| Stankevich et al [17] | Reddit Posts | SVM* and Random Forest | Test accuracy is 63.36% |
| Shen et al [18] | Reddit Posts | Neural Network | Test accuracy is 85% |
| Shen et al [19] | Reddit and Weibo Posts | DNN-FATC* | Test accuracy is 80% |
| Zuorba et al [20] | Twitter and Facebook | LDA with NLP and Logistic Regression | 99.9162% based on the outcome of the confusion matrix |
| Yazdavar et al [21] | Twitter Posts | Lexicon based technique with semi supervised topic modeling techniques | 71% Accuracy with SVM based model |
| Wang et al [22] | Sina Micro Blog | Node and Linkage features | 95% Output Accuracy |
| Katchapakirin et al [23] | Facebook Posts | NLP with SVM, Random Forest and Deep Learning | Random Forest (84.6%) and Deep Learning (85%) |
| Asad et al [24] | Twitter and Facebook Posts | SVM and Naïve Bayes for training Machine Learning Model and NLP for uncluttering the data | Accuracy of Naïve Bayes (74%) |

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

| | | | |
|---|---|---|---|
| Seah et al [25] | Reddit Posts | LDA and Python Gensim Package | Successfully identifies depression and suicidal posts into 14 topics (100%) |
| Uddin et al [26] | Bengali Twitter Posts | LSTM* Deep Recurrent Model | 86.3% after tuning layers of LSTM |
| Deshpande et al [27] | Twitter post | SVM and Naïve Bayes Classifiers | Naïve Bayes 83.3% and SVM 79.7% |
| Chen et al [28] | Twitter Posts | EMOTIVE, Random Forest and SVM | Emotive + Random Forest 92.82% (accuracy) |
| Ray et al [29] | DAIC-Wizard-of-Oz dataset | Neural Network, Deep LSTM, BLSTM Network | 17.52% better than the baseline performance |

Table 2: General Techniques and Data sources of the reviewed systems

*CES-D = Center for Epidemiologic Studies Depression Scale; *CNN = Convolutional Neural Network; *BERT= Bidirectional Encoder Representations from Transformers; *ANN = Artificial Neural Network; *NLP= Natural Language Processing; *DNN-FATC= Deep Neural Network model with Feature Adaptive Transformation & Combination; *SVM=Support Vector Machine; LSTM=Long Short-Term Memory

## IV. SYSTEMATIC REVIEW ON DETECTING DEPRESSION

This section highlights recent techniques which were developed to detect depression from social media.

M. Tadesse et al. [7] examined posts from Reddit users for detecting factors which reveal depression attitudes of those users. The data composed of 1293 depressive posts and 548 standard posts. Natural Language Processing and Machine Learning techniques were used to train the data. LIWC dictionary, LDA topics and N-gram features were employed so that data could be processed by different classifiers. Depression of the posts were by a framework which was developed using Logistic Regression, Support Vector Machine, Random Forest, Adaptive Boosting and Multilayer Perceptron (MLP). Using the combination of LIWC, LDA and bigram text encoding methods along with MLP, 91% accuracy is achieved in detecting depression from the post. Choudhury et al. [8] used a crowdsourcing data of 637 twitter users posts and these posts were characterized based on time, emotion, language styles, n-gram, user engagement and user ego-network.

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

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F_1 = 2.\frac{precision.recall}{precision + recall}$$

Figure 1. Output Calculation

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

The researchers developed a SVM classifier with an RBF kernel which predicted depression with an accuracy of 73% and a precision of 0.82. The researchers also introduced a Social Media Depression Index (SMDI) that help in characterizing levels of depression in population across geography, demography and seasonal pattern. When SMDI is compared to various other reported depression rates, SMDI showed a very high correlation. Major drawbacks include that user do not get any information about his/her mental state; data and depression detection result are used by researchers only.

Ziwe and Chua [9] developed a web application which performs sentiment analysis from text and written communication. The application displays users their depression state in form of charts through a web interface. Dictionaries of Github.com and SSD lexicon [10] were used in understanding the depression lexicon from texts. Cronbach Alpha formula was used as a classification function for depression. Output comes either classifying a post as a depressed or not a depressed. One of the major drawbacks of the system is that only publicly posted tweets are detected for depression.

$$\text{SMDI} = \frac{\text{depressed posts at a given time} - \text{mean of the depressed posts}}{\text{standard deviation of the depressed posts}} - \frac{\text{regular posts at a given time} - \text{mean of the regular posts}}{\text{standard deviation of the regular posts}}$$

Figure2: SMDI calculation formula

Park et al. [11] implemented a Web application within Facebook which has an easy access to Center of Epidemiological Studies – Depression (CES-D) scale. The application give feedback on depression tests also provide tips and games. 234 participants aged between 20 to 30 from a university tested the application. 20% of the participants were found out to be depressed. It was observed that depressed user gets fewer likes, comments, commented less on other posts. They also have more wall posts and remained active for long hours at night. The drawbacks of the method include tested users being from a specific demography, no linguistic features were examined and since the system will predict depression only from Facebook posts.

Lin et al. [12], analyzed the social media data composed of images from profile pictures and texts from posts of 1202 depressed twitter users and 300 million non-depressed users. Visual features from images were extracted using CNN classifier and text signals were extracted using BERT. Both CNN and BERT were trained using testing data set which comprised of 280 depressed user and 280 regular users. Afterwards, image and textual features were combined using Neural Network formula with basic linear function that stated whether a user is depressed or not. The system generated an accuracy of 88.4% with a high F1-score of around 93%. Wang et al. [13] analyzed the CLEF eRisk 2018 [14] dataset for early detection of depression as well as Anorexia from posts of Reddit users. Testing and training datasets were both obtained from Reddit and TF-IDF [15] scores were computed from the dataset using the Scikit-Learn. CNN model is trained using the training data in order to predict a Reddit post being a depressive post or not.

Diego Maupome and Marie-Jean Meurs [16] developed a system which translates user writing to emotion analysis with the help of on unsupervised topic extraction and a neural network model. Topic extraction helped the system to summarize a lengthy text to an interpretable manner and is extracted using an unsupervised statistical model in NLP called LDA from eRisk 2018 dataset composing posts of 1707 Reddit users. The extracted topic is then provided to a class of Artificial Neural Network (ANN) called Multilayer Perceptron (MLP) which then predicts a label on each user being depressed or not. Stankevich et al. [17] processed the text messages of 887 reddit users obtained from CLEF

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

eRisk 2017 dataset for detecting depression. Features such as tf-idf, word embeddings and bi-grams along with complex features like morphology and stylometric were evaluated from the text of user posts. SVM and Random Forest machine learning algorithms were used for classification purpose. In different experiments with various combination among features and ml algorithm, the model with SVM as classifier along with tf-idf, stylometric and morphology as the features generated the best F1-score which is 63.36%. It was observed that depressed users used less unique words in their posts Often depressed users posted their messages at night in compared to non-depressed users.

| | |
|---|---|
| Total number of users | 489 |
| (1) #users detected with depression | 117 |
| (2) #users with no signs of depression | 157 |
| Total number of Twitter posts | 69,514 |
| (1) Positive class (depression-indicative) | 23,984 |
| (2) Negative class (standard posts) | 45,530 |
| Mean number of posts per user | 253.7 |
| Variance of number of posts per user | 24.88 |
| Mean number of posts per day per user | 2.79 |
| Variance of number of posts per day per user | 7.04 |

Table 3: Twitter Post statistics from [8]

Shen at el [18] harvested social media data to detect depression of the user using a proposed Modal Depressive Dictionary Learning Model (MDL). Data set is composed of tweets three subsets of more than 300 million twitter users – Depressed users, regular users and unlabeled users. After preprocessing the data, several depressive feature groups are declared namely Social Network, User Profile, Visual, Emotional, Topic-Level and Domain-specific feature. The proposed model presents users in these six groups. The depression prediction results are then compared with results obtained from Naïve Bayesian, Multiple Social Networking Learning (MSNL) and Wasserstein Dictionary Learning (WDL). MDL gave the best performance in comparison to the three ML algorithms, resulting 85% in F1-score, clearly indicating that MDL is extremely efficient in depression detection. Another similar extended work of the earlier work done by Shen et al. [19] using twitter as a source domain of 2788 user data, depression is detected from a target domain such as Weibo of 1160 user data. After analyzing the textural, visual, user profile, posting behavior and social interaction features from both the datasets, two challenges were emerged while predicting target domain data using source data. In order to tackle these challenges, a cross domain Deep Neural Network model with Feature Adaptive Transformation and Combination strategy (DNN-FATC). When compared, the effectiveness of DNN-FATC transfer method had a higher F1 score of around 80%, in respect to other transfer methods. This indicates a better chance of depression detection in cross domains where one is the source domain and the other one is a target domain.

Zuorba et al [20] initially used Data Mining techniques to identify keywords related to depression from Facebook and Twitter posts. Two lexicons were used to identify user sadness from the post. For classifying user as either “sad” or “not sad”, Logistic Regression was used. Logistic Regression is used to find the value of the β parameter to find the best decision – either being “sad” or “not sad”.

$$y = \begin{cases} 1 & \beta_0 + \beta_1 x + \varepsilon > 0 \\ 0 & else \end{cases}$$

Figure 3: β parameter calculation for Logistic Regression

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

The depressed users were further clustered using the Standard Topic Modeling Toolbox (STMT) into certain topics. LDA was used on those topics to find the reason behind certain user being depressed. The logistic regression model had an accuracy of around 99.92% and generated confusion matrix in support to its accuracy.

| | Depressed (Predicted) | Not Depressed (Predicted) |
|---|---|---|
| Depressed (Actual) | 62811 | 68 |
| Not Depressed (Actual) | 0 | 18270 |

Figure 4: Confusion Matrix of the Logistic Regression

Yazdavar et al [21] proposed a system which initially identifies depressed words from twitter posts with the help of word clusters. The lexicon is used to extract symptoms from those words. The semi supervised model named CC-SVM along with lexicon generated a desirable output accuracy of around 71% in identifying depression from twitter posts. The authors also generated different symptoms of depression over a period which was learned by ssToT model, shown in Figure 5.

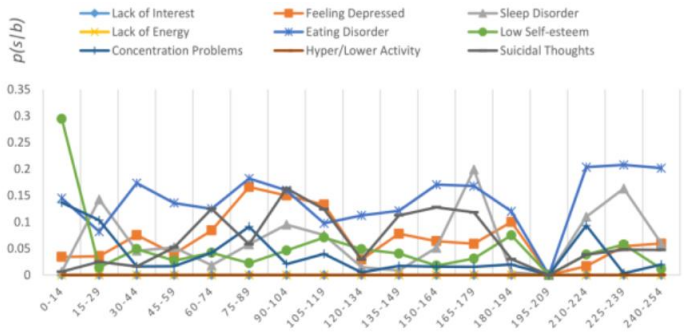


Figure 5. Trends of depression generated by the ssToT model in [21]

Wang et al [22] proposed a model which finds a user being depressed based on node and linkage. Tie strength and interaction content analysis is used for finding linkage feature. The model involving both node and linkage have an accuracy of 95%. Katchapakirin et al [23] developed a model which first translates Thai Facebook posts in English language. Then NLP techniques are used to analyze the sentiments. Then two machine learning techniques (Random Forest and SVM) and one deep learning technique is used to find whether a particular post is a depressed post or not. The SVM generated an accuracy score of 68.57% where-else accuracy of Random Forest and Deep Learning technique was around 85%.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

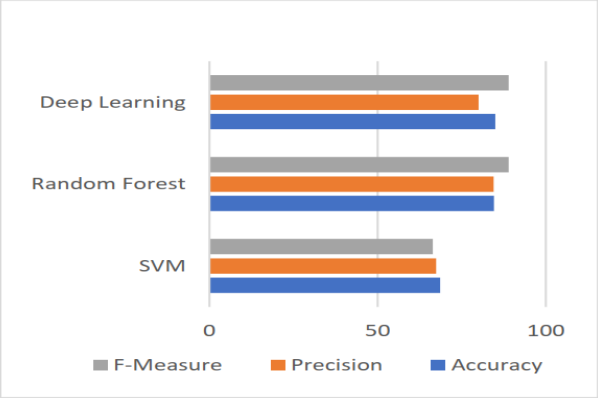


Figure 6: Bar chart comparing the output generated by [23].

Asad et al [24] proposed a method where first user social media data is processed and provided to a machine learning model where SVM is used to vectorize the post and Naïve Bayes is used to determine whether a post is depressed, normal or neutral. NLP is used next to unclutter the data. Naïve Bayes had an accuracy of 75% in predicting whether a Facebook or twitter post is depressed.

| | Depressed (Predicted) | Not Depressed (Predicted) |
|---|---|---|
| Depressed (Actual) | 17 | 13 |
| Not Depressed (Actual) | 0 | 20 |

Figure 7: Confusing Matrix showing the Naïve Bayes accuracy in predicting depression in [24]

Seah et al [25] used LDA for discovering topics. Python Gensim package was then used to conduct topic modeling on dataset from Reddit. The authors of this literature categorized depression and suicide into 14 topics which elaborates how suicidal and depressive thoughts begin in the first place. In Uddin et al [26], LSTM Deep Recurrent Neural network was used in order to analyze Bengali twitter posts for depression. Tuning different paraments of the LSTM lead to different output accuracy. The best LSTM tuning configuration can be seen in the Figure 8. Accuracy is high 78.9% when size of LSTM is 128, 81.1% when LSTM batch size is 25 with 10 epoch and 86.3% when LSTM has 5 layers.

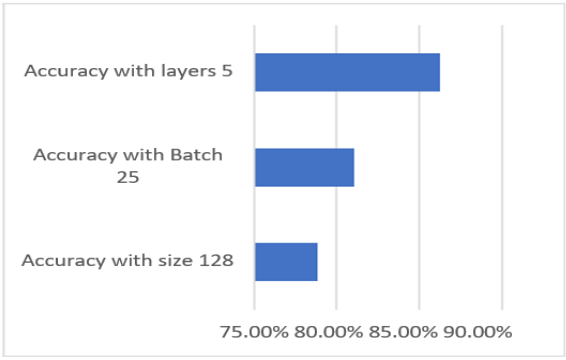


Figure 8: Accuracy comparison in [26]

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.6 Aug-Sept – 2022
DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

Deshpande et al [27] applied Naïve Bayes and SVM on a twitter dataset to find out whether a post indicates depression or not. The authors found that Naïve Bayes showed an accuracy of 83% in predicting depression where-else SVM showed a slight less accuracy of 79%.

| | Precision | Recall | Accuracy |
|---|---|---|---|
| **Naïve Bayes** | 83.6% | 83% | 83% |
| **SVM** | 80.4% | 79% | 79% |

Table 4: Comparison between depression prediction output between Naïve Bayes and SVM.

In a work by Chen et al [28], basic emotions from twitter posts were measured by employing emotive system. Next, various feature sets were constructed. Then Emotive along with other Machine Learning algorithms were used compare the best result and it was found that emotive with random forest generated the best. Ray et al [29] introduced a method of fusing video, text and audio information through a network using a combination of Neural Network, Deep LSTM, BLSTM Network to detect whether a particular data is indicating depression or not.

## V. RECOMMENDATION.

Based on the reviewed literatures, we would like to recommend that most of the researchers and scientists studied social media data to detect depression, but depression as whole also depends on the time when various posts are made. In future, researchers should also keep time parameters before implementing any depression detection models. Moreover, various machine learning and data mining techniques were used to detect depression, but all these techniques were solely focused on either Facebook or twitter dataset, data from different blogs were analyzed at a very small scale. Only a handful of techniques used other methods rather than the conventional textual analysis to identify depression in any user. In future researchers in this field should focus more into other features rather than only textual features. Furthermore, most of the output's accuracy was shown through F1 Measurements and other accuracy measures but almost all of the literatures did not validate their output by any psychologists or medical officers. It would be highly recommended to researchers ensure proper validation of their findings so that the derived model can be used even for basic medical purpose.

| Depressed-User Action Features | Indication About the Depressed User |
|---|---|
| Posting tweets between 23:00 to 06:00 | Majority depressed user might suffer from Insomnia |
| 52% of the words from the tweets are negative | Depressed user expresses more negative emotions |
| 26% of the words from tweets are often personal pronouns | Depressed users have suppressed monologues and strong self-awareness |
| Posting depression symptom words 165% more than regular users | Willingness to share what they faced in their daily lives |

Table 5: General Findings of multiple paper regarding attitudes of depressed users from twitter data sets

DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

## VI. CONCLUSION

In this paper we have discussed several depression detection strategies from social media. It is observed that majority of the methods used Machine Learning Algorithms and Neural Network approaches. The datasets for most of the methods were from Twitter and Reddit. Classifiers in combinations with different text encoding techniques generated standard accuracy for almost all the models. All the methods had negligible limitations. Even-though, we see each method predict with very high accuracy, yet none of the discussed procedures are used for certified clinical purpose. It is expected that researchers will device more accurate algorithm which might be accepted clinically.

## VII. REFERENCES

[1]**World Health Organization**. https://www.who.int/news-room/factsheets/detail/depression
[2]Lenore Sawyer Radloff. The ces-d scale: **A self-report depression scale for research in the general population. Applied Psychological Measurement, 1(3):385–401, 1977.**
[3]Aaron Beck, Martin Mendelson, J Mock,and J Erbaugh. **An inventory for measuring depression. Archives of General Psychiatry, 4(6):561–571, 1961.**
[4]Whooley and Owen. **Diagnostic and Statistical Manual of Mental Disorders (DSM). John Wiley & Sons, Ltd, 2014.**
[5] **History** website: https://www.history.com/this-day-in history/facebook-launches-mark-zuckerberg
[6]**Omnicore Facebook Statistics.** https://www.omnicoreagency.com/facebookstatistics/
[7]Detection of Depression-Related Posts in Reddit Social Media Forum (IEEE)Link: M. M. Tadesse, H. Lin, B. Xu and L. Yang, "Detection of Depression-Related Posts in Reddit Social Media Forum," in IEEE Access, vol. 7, pp. 44883-44893, 2019, doi: 10.1109/ACCESS.2019.2909180.
[8]Social Media as a Measurement Tool of Depression in Populations (ACM) Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Social media as a measurement tool of depression in populations. In Proceedings of the 5th Annual ACM Web Science Conference (WebSci '13). Association for Computing Machinery, New York, NY, USA, 47–56. DOI:https://doi.org/10.1145/2464464.2464480
[9]An Application for Classifying Depression in Tweets (ACM) Bernice Yeow Ziwei and Hui Na Chua. 2019. An Application for Classifying Depression in Tweets. In Proceedings of the 2nd International Conference on Computing and Big Data (ICCBD 2019). Association for Computing Machinery, New York, NY, USA, 37–41. DOI:https://doi.org/10.1145/3366650.3366653
[10] Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media (ACM) Amir Hossein Yazdavar, Hussein S. Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (ASONAM '17). Association for Computing Machinery, New York, NY, USA, 1191–1198. DOI:https://doi.org/10.1145/3110025.3123028
[11] Manifestation of Depression and Loneliness on Social Networks: A Case Study of Young Adults on Facebook (ACM) Sungkyu Park, Inyeop Kim, Sang Won Lee, Jaehyun Yoo, Bumseok Jeong, and Meeyoung Cha. 2015. Manifestation of Depression and Loneliness on Social Networks: A Case Study of Young Adults on Facebook. In Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15). Association for Computing Machinery, New York, NY, USA, 557–570. DOI:https://doi.org/10.1145/2675133.2675139
[12] SenseMood: Depression Detection on Social Media (ACM) Chenhao Lin, Pengwei Hu, Hui Su, Shaochun Li, Jing Mei, Jie Zhou, and Henry Leung. 2020. SenseMood: Depression Detection on Social Media. In Proceedings of the 2020 International Conference on Multimedia Retrieval (ICMR '20). Association for Computing Machinery, New York, NY, USA, 407–411. DOI:https://doi.org/10.1145/3372278.3391932

[13] A Neural Network Approach to Early Risk Detection of Depression and Anorexia on Social Media Text (RG) Wang, YT., Huang, HH., Chen, HH.: A Neural Network Approach to Early RiskDetection of Depression and Anorexia on Social Media Text. Proceedings of the 9thInternational Conference of the CLEF Association, CLEF 2018, Avignon, France (2018)
[14]**Erisk 2018: Early Risk Prediction On The Internet** https://early.irlab.org/2018/index.html
[15]**TF-IDF**: http://www.tfidf.com/
[16]Diego Maupome and Marie-Jean Meurs., **Using Topic Extraction on Social Media Content for the Early Detection of Depression**
[17] Maxim Stankevich, Vadim Isakov, Dmitry Devyatkin and Ivan Smirnov, **Feature Engineering for Depression Detection in Social Media**
[18] Guangyao Shen, Jia Jia, Liqiang Nie, Fuli Feng, Cunjun Zhang, Tianrui Hu, Tat-Seng Chua and Wenwu Zhu, **Depression Detection via Harvesting Social Media: A Multimodal Dictionary Learning Solution**
[19] Tiancheng Shen, Jia Jia, Guangyao Shen, Fuli Feng, Xiangnan He, Huanbo Luan, Jie Tang, Thanassis Tiropanis, Tat-Seng Chua and Wendy Hall., **Cross-Domain Depression Detection via Harvesting Social Media**
[20] A Framework for Identifying Excessive Sadness in Students through Twitter and Facebook in the Philippines (ACM) Hussain D. Zuorba, Celine Louise O. Olan, and Anthonette D. Cantara. 2017. A Framework for Identifying Excessive Sadness in Students through Twitter and Facebook in the Philippines. In Proceedings of the International Conference on Bioinformatics Research and Applications 2017 (ICBRA 2017). Association for Computing Machinery, New York, NY, USA, 52–56. DOI:https://doi.org/10.1145/3175587.3175600ilippines
[21] Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media (ACM) Amir Hossein Yazdavar, Hussein S. Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (ASONAM '17). Association for Computing Machinery, New York, NY, USA, 1191–1198. DOI:https://doi.org/10.1145/3110025.312302
[22] A Neural Network Approach to Early Risk Detection of Depression and Anorexia on Social Media Text (RG) Wang, YT., Huang, HH., Chen, HH.: A Neural Network Approach to Early RiskDetection of Depression and Anorexia on Social Media Text. Proceedings of the 9thInternational Conference of the CLEF Association, CLEF 2018, Avignon, France (2018)
[23] Facebook Social Media for Depression Detection in the Thai Community (IEEE) K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot and Y. Kaewpitakkun, "Facebook Social Media for Depression Detection in the Thai Community," 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), Nakhonpathom, 2018, pp. 1-6, doi: 10.1109/JCSSE.2018.8457362.
[24] Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media (ACM) Amir Hossein Yazdavar, Hussein S. Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (ASONAM '17). Association for Computing Machinery, NewYork, NY, USA, 1191–1198. DOI:https://doi.org/10.1145/3110025.3123028
[25] Data Mining Approach to the Detection of Suicide in Social Media: A Case Study of Singapore (IEEE) J. H. K. Seah and K. Jin Shim, "Data Mining Approach to the Detection of Suicide in Social Media: A Case Study of Singapore," 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 2018, pp. 5442-5444, doi: 10.1109/BigData.2018.8622528.

DOI:10.46647/ijetms.2022.v06i05.136 ISSN: 2581-4621

[26] Depression Analysis from Social Media Data in Bangla Language using Long Short-Term Memory (LSTM) Recurrent Neural Network Technique (IEEE) A. H. Uddin, D. Bapery and A. S. M. Arif, "Depression Analysis from Social Media Data in Bangla Language using Long Short Term Memory (LSTM) Recurrent Neural Network Technique," 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2019, pp. 1-4, doi: 10.1109/IC4ME247184.2019.9036528

[27] Depression detection using emotion artificial intelligence (IEEE) M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 858-862, doi: 10.1109/ISS1.2017.8389299.

[28] What about Mood Swings: Identifying Depression on Twitter with Temporal Measures of Emotions (ACM) Xuetong Chen, Martin D. Sykora, Thomas W. Jackson, and Suzanne Elayan. 2018. What about Mood Swings: Identifying Depression on Twitter with Temporal Measures of Emotions. In Companion Proceedings of the The Web Conference 2018 (WWW '18). International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 1653–1660. DOI:https://doi.org/10.1145/3184558.3191624

[29] Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media (ACM) Amir Hossein Yazdavar, Hussein S. Al-Olimat, Monireh Ebrahimi, Goonmeet Bajaj, Tanvi Banerjee, Krishnaprasad Thirunarayan, Jyotishman Pathak, and Amit Sheth. 2017. Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 (ASONAM '17). Association for Computing Machinery, New York, NY, USA, 1191–1198. DOI:https://doi.org/10.1145/3110025.3123028