

An efficient model to detect social network mental disorders using machine learning techniques

¹Er. Astha Shrivastava, ²Dr. Rohit kumar Miri

¹Research Scholar, Department of Computer Science and Engineering, Dr. C. V. Raman University, Kota, Bilaspur , C.G., India

²Department of Computer Science and Engineering, Dr. C. V. Raman University, Kota, Bilaspur , C.G. ,India

Email : ¹astha2989@gmail.com, ²rohitmiri@gmail.com

Abstract—A person is said to be suffering from a mental disorder if their cognition, emotion control, or behavior are all affected. Despite the prevalence of mental illness, it is often under diagnosed. There is a lot of curiosity about the link between depression and social media. Individuals' well-being is negatively impacted by social media use, according to several experts. Many attempts have been made to analyze individual postings using Machine Learning (ML) methods to identify mental patients on social media. Datasets taken from different social media platforms and the sorts of characteristics used in the identification of mental disorders are discussed in this study. Research methods included decision trees, random forests, support vector machines, recurrent neural networks, convolutional neural networks as well as logistic regression. In comparison to the other methods, the Convolutional Neural Network, also known as (CNN), achieved the best accuracy, which was 91.08 percent, for the diagnosis of mental conditions. Accuracy rates of 85.87 percent, 81.22 percent, and 85 percent were achieved by using other methods such as random forests, Support Vector Machines (SVM), and recurrent neural networks (RNN), respectively. It has been shown that ML techniques, when useful to text data collected from users of social media platforms, can be an effective method for detecting depression and can one day serve as supplementary tools in the field of public mental health.

Keywords—machine learning; depression; social media; SVM; CNN.

I. INTRODUCTION

Anxiety and depression are the ninth and tenth most prevalent causes of disease burden worldwide [1]. They have the potential to harm one's physical, emotional, and mental well-being [2]. Anxiety and suicidal thoughts are common symptoms of depression. Sleep disturbances, fatigue, poor energy, and a sense of worthlessness are all common [3]. Depression detection helps reduce these hazards. Depression is traditionally diagnosed using subjective measures or clinical diagnoses from attending professionals[4]. Results from conventional scales given in the typical way are influenced by context, the patient's mental state at the time, the clinician-patient relationship at that time, and the patient's current mood and memory bias [5]. Traditional diagnostic methods are time-consuming [6]. In the early stages of depression, individuals may be uninformed or embarrassed by their depressive symptoms and not seek professional help. The world health organization (WHO) estimates that 350 million individuals worldwide suffer from depression [7]. ML uses complex statistical and probabilistic approaches to build systems that improve with experience [8]. It is a valuable technique for forecasting mental health. It allows researchers to get critical data, deliver individualized experiences, and build clever automated systems [9]. Machine learning techniques such as SVM, RF, and Artificial neural network (ANN) are used to anticipate and classify future occurrences [10]. The WHO publishes reports on the state of various hurdles in diagnosing mental health disorders on a regional basis and urges researchers to arm themselves with the scientific information necessary to tackle the problem of mental health [11]. As a result of advances in technology, there are a variety of methods that may be used to forecast an individual's mental

health [12]. The amount of research conducted in mental health has lately expanded, which has led to the knowledge and publications on many aspects of mental health [13]. This information and these publications may be used on a diverse assortment of issues [14]. A person's ability to think, feel, and behave normally could all be negatively impacted by mental illness. The most prominent sign of depression is an alteration in mood, which most often takes the form of an overwhelming sense of melancholy [15]. Impatience, anger, and lack of interest may be the predominant symptoms of depression at times [16].

Social media mental disorder

The effect of social media on society are complex and debated. Social media might contribute to depression since it harms people's well-being. An excessive amount of time spent on social media may also contribute to sleep problems in adolescents[17]. Those who are eager to communicate their thoughts are more likely to look for support on social media than those who are reluctant to do so [18]. According to the authors, the use of social media is associated with poor levels of self-esteem among young people, which may be a precursor to mental illnesses. Many people believe that social media is an effective tool for strengthening social relationships and gaining social support via the exchange of information because of its widespread use. Self-disclosure may provide insight into the likelihood that an individual would suffer from depression, as revealed by the individual. Studies link social media self-disclosure to depression [19]. In social media, user profile information discloses user traits, thus academics may use it to predict sadness. Figure 1 shows how depression is addressed on social media, which has grown rapidly. Informants spend 1 to 8 hours every day using the internet, and they feel frustrated, furious, bored, ostracized, separated from the world, or that something is missing when they can't. The informants have their dedication to working [20].

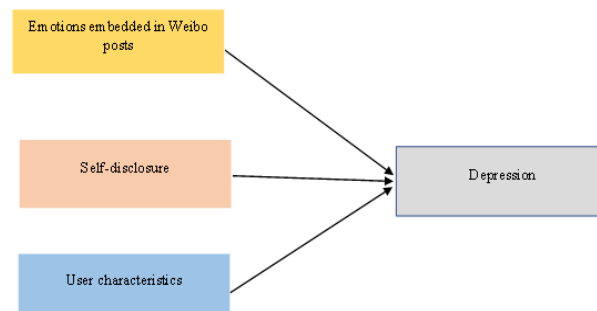


Fig.1. A conceptual model of depression is indicated by social media content [21].

Machine learning (ML) in Mental Health

ML allows machines to "learn" and anticipate outcomes automatically. ML has helped patients and clinicians in healthcare. Complex datasets in cancer, radiology, cardiology, and pathology enable clinicians to make better judgments by evaluating images. EHR software has begun digitizing patient data, invoices, and prescriptions in various nations. ML approaches have strengths in managing big data since scalable

ML algorithms to handle large-scale data were investigated and designed before the relevance of big data appeared [22]. World Health Organization (WHO) says depression causes mental disease. They need an intelligent system to recognize symptoms and learn from datasets to accurately forecast a person's sentiment [23]. In mental health, these typical ML algorithms are used: SVM, RF, KNN, GBM, and Naive Bayesian (NB). In the field of mental health, popular machine learning techniques include the Gradient Boosting Machine (GBM), Random Forest, the SVM, and Naive Bayesian [24]. The datasets are analyzed to obtain essential information, which is the objective of the Machine Learning approaches [25]. Most of both supervised and unsupervised learning takes place in the medical field. One other kind of learning that is used in the process of data analysis is known as reinforcement learning or RL.

II. LITERATURE REVIEW

In this section, there are various author has given their outcomes which has described given below.

Zogan, et al., (2022) [26] introduced that MDHAN finds people who are depressed on social media and provides an explanation of the model prediction. They've looked at how to improve user-submitted tweets by utilizing Twitter as a research tool. Semantic sequence characteristics taken from user timelines are encoded using algorithms for the tweet and word attention, and tweet and word significance is appraised (posts). Patterns that explain findings may be found using hierarchical attention models. Deep learning and multi-aspect features may be combined in a way that surpasses several well-known and robust baseline techniques, as shown by findings. Increased accuracy in predicting the amount of depression in social media users was achieved using that method. It's safe to say MDHAN does an excellent job and gives enough facts to back up its predictions.

Lavanya, et al., (2021) [27] presented that the text is categorized using Deep Learning (DL) models on social media health networks. This study's primary contribution is that it helps to provide insight for training the data and categorizing the text by analysing and extracting the raw input, as well as creating output using Natural Language Processing (NLP). As a result of a growth in the amount of complexity included in healthcare data, Recurrent Neural Networks (RNNs) and CNNs are gaining more and more traction in the field of healthcare application development. The end objective is to enhance text classifier performance based on efficacy to boost text processing speed and accuracy by utilizing an appropriate approach to achieve promising results for the future.

Chancellor, et al., (2019) [28] studied the depiction of human players in human-centered machine learning research in academic journals (HCML). These representations are investigated using a thematic discourse analysis of 55 pieces, with social media status prediction of mental health as a case study. Person, Scientific, Data/Machine Learning, and Disorder/Patient/social media are the five distinct discourses that contribute to the total picture of the human subject. Human dehumanization may occur because of these five discourses' divergent perspectives on humanity. Multidisciplinary research poses a challenge to interdisciplinarity, scientific methods, and online communities. It also offers ideas for better HCML research in this new sector.

Burdisso, et al., (2019) [29] explained that a new approach for text classification, SS3, that naturally supports these properties has been introduced. A general framework for dealing with all forms of early risk detection (ERD) difficulties was built in SS3. Models such as support vector machines, maximum likelihood, and neural networks don't operate very well in this scenario. This is because they either do not offer incremental categorization or learning or function as black boxes. It was based on the results of the CLEF's eRisk2017 pilot project on the early detection of depression. The vast majority of the thirty entries submitted for consideration in this competition used cutting-edge methods and technology. These models and standard classifiers were defeated by this classifier, despite its lower processing cost and the ability to explain its reasoning. This was shown by the results of the studies.

Chacheda, et al., (2019) [30] studied that depressive illness, commonly referred to as major depressive disorder, affects around 300 million individuals worldwide (MDD). As a result, depression is one of the most common mental health conditions. An early diagnosis is critical for swift treatment, which may prevent the sickness from becoming more serious. Suggested machine learning methods included single and dual. A single RF classifier with two threshold functions is used in the first case, whereas two separate RF classifiers are used in the second. Early detections are rewarded, whereas late detections are penalized, in this assessment. To combat depression in its earliest stages using social media, they believe that this research might be useful.

Shatte, et al., (2019) [31] introduced that the use of machine learning in the area of mental health has shown to have a variety of positive effects, especially in terms of diagnosis, treatment, and support, as well as research and clinical administration. The research that was done on this subject led to the discovery of four key application sectors, which are as follows: (i) detection and diagnosis; (ii) prognosis, treatment, and support; (iii) public health; and (iv) research and clinical administration. One of the most common mental health problems people faced was depression, followed by schizophrenia and then Alzheimer's disease. In addition to the decision trees, support vector machines, neural networks, and clustering, latent Dirichlet allocation was also used. Most research concentrate on the identification and diagnosis of mental health issues; thus, ML might help other aspects of psychological performance. The obstacles and potential of employing ML methods are highlighted.

Guntuku, et al., (2019) [32] presented that it is possible to modify and scale up the language-based models that have been created to assess county-level changes. Twitter now has access to county-level language because of the Streaming API, and a variety of tactics are studied for adapting user-level Facebook models to Twitter. Compared to traditional survey-based measurements of stress, domain-adapted and scaled social media stress indicators outperform. Higher stress levels are connected not just with worse health but also with a lack of access to medical services and a worse socioeconomic standing. Their conversation would end with an examination of the consequences of utilizing social media as a new instrument for measuring the levels of stress experienced by people as well as by counties.

Islam, et al., (2018) [33] described that use of social networks may share their thoughts and contact their interested friends, feelings, videos, and images that express their moods,

feelings, and emotions. This opens the possibility of analyzing the thoughts and sentiments of users on social networks to learn more about their current state of mind and outlook on life. An efficient and scalable implementation of the suggested strategy using machine learning techniques. The findings indicate that in a variety of studies, the Decision Tree (DT) technique provides the greatest accuracy in comparison to other ML methods for locating depression. Using approaches based on machine learning, Facebook users with mental health issues may find high-quality answers to their difficulties.

Du, et al., (2018) [34] computed that in this study, researchers employed a combination of deep learning and a transfer learning mechanism to detect suicide-related mental pressures in tweets. With the help of a CNN, researchers were able to construct a binary classifier. The topic of stressor identification is approached using techniques that are based on RNN since it is conceived of as an ordinarily Named Entity Recognition (NER) challenge. The findings point to the possibility of using techniques based on deep learning for the task of automating the detection of sources of stress in social media.

Paul, et al., (2018) [35] examined the effectiveness of several machine learning algorithms in the early detection of anorexia and depression. Ad a boost, random forest, support vector machine, and logistic regression classifiers have all been used in the basic bag of words model to find documents in the various corpora that pertain to anorexia or depression. Meta map, a method for extracting biological ideas, was also used to extract anorexia and depression-related phrases. GloVe and Fast text word embeddings are also used to evaluate the performance of the recurrent neural network. To train Glove and Fast text, specialized corpora, like Wikipedia, were used to create pre-trained word vectors. This classifier surpasses all other techniques in task 1 in terms of precision on the training set, and it also does well in task 2 in terms of accuracy throughout all runs in the challenge. The bag of words model-based support vector machine classifier outperforms the other techniques in terms of task2 measurement. There is evidence to indicate that these frameworks perform well on the test data provided for the contest. Table I shows the summary of reviewed literature which has described given below.

TABLE I. Depicts the summary of the reviewed literature

Sr. no.	Author	Techniques	Literature of review	Outcomes
1.	Zogan, et al., (2022) [26]		MDHAN	It's safe to say MDHAN does an excellent job and gives enough facts to back up its predictions.
2.	Lavanya, et al., (2021) [27]		Deep learning (DL) models	The end objective is to enhance text classifier performance based on efficacy to boost text processing speed and accuracy by utilizing an appropriate approach to achieve promising results for the future.
3.	Chancellor, et al., (2019) [28]		Human-centered machine learning (HCML)	It also offers ideas for better HCML research in this new sector.
4.	Burdisso, et al., (2019) [29]		SS3's design	Despite being less computationally costly and being able to explain its logic
5.	Chacheda, et al., (2019) [30]		Singleton and dual machine learning	To combat depression in its earliest stages using social media, they believe that this research might be useful.
6.	Shatte, et al., (2019) [31]		Decision trees, convolutional neural networks, Support vector machines, clustering., and latent Dirichlet allocation.	Most research concentrate on the identification and diagnosis of mental health issues; thus, ML might help other aspects of psychological performance.
7.	Guntuku, et al., (2019) [32]		Language-based models	The consequences of utilizing social media as a new instrument for measuring the levels of stress experienced by people as well as by counties.
8.	Islam, et al., (2018) [33]		Decision Tree (DT) technique	Facebook users with mental health issues may find high-quality answers to their difficulties.
9.	Du, et al., (2018) [34]		Recurrent neural networks (RNN)	The findings point to the possibility of using techniques based on deep learning for the task of automating the detection of sources of stress in social media.
10.	Paul, et al., (2018) [35]		Ad a boost, random forest, support vector machine, and logistic regression	The bag of words model-based support vector machine classifier outperforms the other techniques in terms of task2 measurement.

III. COMPARISON

It is the primary premise of this comparative study to compare the accuracy of research on employing machine learning methods and machine learning classifiers to identify social network mental problems, which offers a prediction model based on machine learning to determine the correct therapies for social media mental disorder. The highest approach of the CNN technique has an accuracy of 91.08% for mental disorders as compared to other techniques. Other techniques such as Random Forest, Recurrent neural network, SVM, etc had these accuracies of 85.87%, 81.22%, and 85%. Table II describes the comparative study of the models that are used by various authors for a mental disorder with the highest accuracy of CNN. It has been shown that machine learning

techniques, when useful to text data collected from users of social media platforms, can be an effective method for detecting depression and can one day serve as supplementary tools in the field of public mental health. In social media, user profile information discloses user traits, thus academics may use it to predict sadness. Machine learning techniques such as SVM, RF, and ANN are used to anticipate and classify future occurrences. The amount of research conducted in mental health has lately expanded, which has led to the knowledge and publications on many aspects of mental health.

TABLE II. Depicts the comparative study of the models that are used by various authors for mental disorders with the highest accuracy of CNN.

SR.n o	Comparison table		
	Author	Technique	Accuracy
1	Shatte, et al., (2019) [31]	Support vector machines, Convolutional Neural Networks	85.87% 91.08%
2	Islam, et al., (2018) [33]	Decision trees	80.69%
3	Du, et al., (2018) [34]	Recurrent neural networks (RNN)	85%
4	Paul, et al., (2018) [35]	Logistic regression, Random Forest	79.63% 81.22%

Figure 2 depicts the comparison graph for the classifier's accuracy in the existing literature which has illustrated given below.

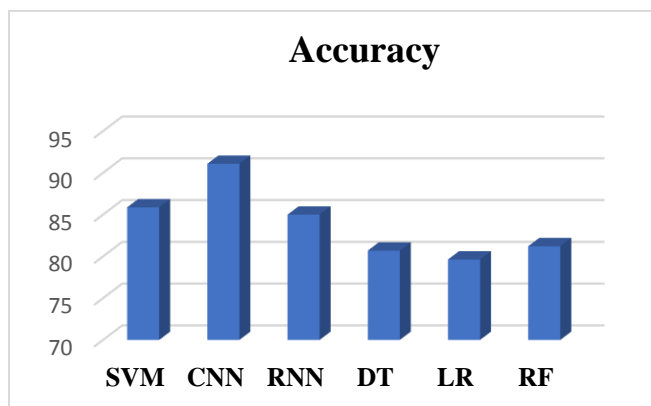


Fig.2: Graphical representation of the accuracy of various classifiers.

IV. CONCLUSION

In this study, a machine learning-based approach for detecting mental diseases in social networks is explored. In this analysis, the definition of depression and the most frequent risk factors for depression are explored. A study comparing the models used by different authors for health monitoring systems found that CNN had an accuracy rate of 99.08 percent after evaluating the proposed model, which is encouraging for our needs. The researchers used ML models to predict depression based on the characteristics. Seventeen different research looked at how words in online texts could be used to extract various kinds of data about the writer's feelings or linguistic style or chronological process. As a result of the experiments discussed in this review, it can be concluded that ML techniques can be useful in diagnosing depression using text data from social media. Additionally, it appears proper and applicable for these approaches to operate as a supplemental tool to the more conventional ways that have been created for diagnosing depression. This would be the case since it seems reasonable and relevant. However, further research is still necessary for the areas of sample size, generalizability, optimization of prediction techniques and characteristics, privacy concerns, and general research ethics. In the work that we will be doing in the future, they want to employ a different way to extract paraphrases from a greater range of emotional characteristics. In addition to this, they want to leverage data from other datasets to see whether our approaches are effective.

REFERENCES

- [1] Brundtland G. Mental health: a call for action by world health ministers. In: Geneva: World Health Organization. Geneva: World Health Organization; 2014.
- [2] Liu, Danxia, Xing Lin Feng, Farooq Ahmed, Muhammad Shahid, and Jing Guo. "Detecting and measuring depression on social media using a machine learning approach: a systematic review." JMIR Mental Health 9, no. 3 (2022): e27244
- [3] World Health Organization. Depression and other common mental disorders: global health estimates. No. WHO/MSD/MER/2017.2. World Health Organization, 2017.
- [4] Alonso, Jordi, Zhaorui Liu, Sara Evans – Lacko, Ekaterina Sadikova, Nancy Sampson, Somnath Chatterji, Jibril Abdulmalik, et al. "Treatment gap for anxiety disorders is global: Results of the World Mental Health Surveys in 21 countries." Depression and anxiety 35, no. 3 (2018): 195-208.
- [5] American Psychiatric Association, D. S., and American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. Vol. 5. Washington, DC: American psychiatric association, 2013.
- [6] Kumar, Mrinal, Mark Dredze, Glen Coppersmith, and Munmun De Choudhury. "Detecting changes in suicide content manifested in social media following celebrity suicides." In Proceedings of the 26th ACM conference on Hypertext & Social Media, pp. 85-94. 2015.
- [7] Coppersmith, Glen, Ryan Leary, Patrick Crutchley, and Alex Fine. "Natural language processing of social media as screening for suicide risk." Biomedical informatics insights 10 (2018): 1178222618792860.
- [8] Aung, Aye Phyu Phyu, Senthilnath Jayavel, Xiaoli Li, and Bo An. "Planning Sequential Interventions to tackle Depression in Large Uncertain Social Networks using Deep Reinforcement Learning." Neurocomputing (2022)
- [9] Ghosh, Shreya, and Tarique Anwar. "Depression Intensity Estimation via Social Media: A Deep Learning Approach." IEEE Transactions on Computational Social Systems 8, no. 6 (2021): 1465-1474.
- [10] Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." Science 349, no. 6245 (2015): 255-260
- [11] Dabek, Filip, and Jesus J. Caban. "A neural network-based model for predicting psychological conditions." In International conference on brain informatics and health, pp. 252-261. Springer, Cham, 2015
- [12] Jung, Hyesil, Hyeoun-Ae Park, and Tae-Min Song. "Ontology-based approach to social data sentiment analysis: detection of adolescent depression signals." Journal of medical internet research 19, no. 7 (2017): e7452.
- [13] Funk, M. "Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level." Retrieved on 30 (2016)
- [14] Punithavathi, R., M. Sharmila, T. Avudaiappan, I. Raj, S. Kanchana, and Samson Alemayehu Mamo. "Empirical Investigation for Predicting Depression from Different Machine Learning-Based Voice Recognition Techniques." Evidence-Based Complementary and Alternative Medicine 2022 (2022).
- [15] Hadzic, Maja, Meifania Chen, and Tharam S. Dillon. "Towards the mental health ontology." In 2008 IEEE International Conference on Bioinformatics and Biomedicine, pp. 284-288. IEEE, 2008
- [16] Sharma, Eva, and Munmun De Choudhury. "Mental health support and its relationship to linguistic accommodation in online communities." In Proceedings of the 2018 CHI conference on human factors in computing systems, pp. 1-13. 2018.
- [17] Chung, Jetli, and Jason Teo. "Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges." Applied Computational Intelligence and Soft Computing 2022 (2022)
- [18] Andalibi, Nazanin, Pinar Ozturk, and Andrea Forte. "Sensitive Self-disclosures, Responses, and Social Support on Instagram: the case of# depression." In Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing, pp. 1485-1500. 2017.
- [19] Jeske, Debora, Sonia Lippke, and Kenneth S. Shultz. "Predicting self-disclosure in recruitment in the context of social media screening." Employee Responsibilities and Rights Journal 31, no. 2 (2019): 99-112
- [20] Kurniasih, Nuning. "Internet addiction, lifestyle or mental disorder? a phenomenological study on social media addiction in Indonesia." KnE Social Sciences (2017): 135-144.

- [21] Lyu, Danyang, Zixuan Wang, Yutong Du, Rob K. Marjerison, and Rongjuan Chen. "Using social media content to identify mental health problems: The case of depression in Sina Weibo." *Review of Integrative Business and Economics Research* (2020).
- [22] Zhou, Zhirou, Tsung-Chin Wu, Bokai Wang, Hongyue Wang, Xin M. Tu, and Changyong Feng. "Machine learning methods in psychiatry: a brief introduction." *General psychiatry* 33, no. 1 (2020).
- [23] Zhou, Lina, Shimei Pan, Jianwu Wang, and Athanasios V. Vasilakos. "Machine learning on big data: Opportunities and challenges." *Neurocomputing* 237 (2017): 350-361
- [24] Qiao, Jiang. "A systematic review of machine learning approaches for mental disorder prediction on social media." In *2020 International Conference on Computing and Data Science (CDS)*, pp. 433-438. IEEE, 2020.
- [25] Cho, Gyeongcheol, Jinyeong Yim, Younyoung Choi, Jungmin Ko, and Seoung-Hwan Lee. "Review of machine learning algorithms for diagnosing mental illness." *Psychiatry investigation* 16, no. 4 (2019): 262.
- [26] Zogan, Hamad, Imran Razzak, Xianzhi Wang, Shoaib Jameel, and Guandong Xu. "Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media." *World Wide Web* (2022): 1-24
- [27] Lavanya, P. M., and E. Sasikala. "Deep learning techniques on text classification using Natural language processing (NLP) in social healthcare network: A comprehensive survey." In *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*, pp. 603-609. IEEE, 2021
- [28] Chancellor, Stevie, Eric PS Baumer, and Munmun De Choudhury. "Who is the "human" in human-centered machine learning: The case of predicting mental health from social media." *Proceedings of the ACM on Human-Computer Interaction* 3, no. CSCW (2019): 1-32
- [29] Burdisso, Sergio G., Marcelo Errecalde, and Manuel Montes-y-Gómez. "A text classification framework for simple and effective early depression detection over social media streams." *Expert Systems with Applications* 133 (2019): 182-197
- [30] Cacheda, Fidel, Diego Fernandez, Francisco J. Novoa, and Victor Carneiro. "Early detection of depression: social network analysis and random forest techniques." *Journal of medical Internet research* 21, no. 6 (2019): e12554.
- [31] Shatte, Adrian BR, Delyse M. Hutchinson, and Samantha J. Teague. "Machine learning in mental health: a scoping review of methods and applications." *Psychological medicine* 49, no. 9 (2019): 1426-1448
- [32] Guntuku, Sharath Chandra, Anneke Buffone, Kokil Jaidka, Johannes C. Eichstaedt, and Lyle H. Ungar. "Understanding and measuring psychological stress using social media." In *Proceedings of the International AAAI Conference on Web and social media*, vol. 13, pp. 214-225. 2019
- [33] Islam, Md, Muhammad Ashad Kabir, Ashir Ahmed, Abu Raihan M. Kamal, Hua Wang, and Anwaar Ulhaq. "Depression detection from social network data using machine learning techniques." *Health information science and systems* 6, no. 1 (2018): 1-12
- [34] Du, Jingcheng, Yaoyun Zhang, Jianhong Luo, Yuxi Jia, Qiang Wei, Cui Tao, and Hua Xu. "Extracting psychiatric stressors for suicide from social media using deep learning." *BMC medical informatics and decision making* 18, no. 2 (2018): 77-87
- [35] Paul, Sayanta, Sree Kalyani Jandhyala, and Tanmay Basu. "Early Detection of Signs of Anorexia and Depression Over Social Media using Effective Machine Learning Frameworks." In *CLEF (Working notes)*. 2018