International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 01 | January - 2024 SJIF Rating: 8.176 ISSN: 2582-3930

Predicting Depression Levels Using Social Media Posts: A Comprehensive Survey

Shreya Shitole
NBN STIC BE IT Department
(of SPPU)
Pune , INDIA shreyashitole282@gmail.com

Shraawani Lattoo
PVPIT BE Computer Department
(of SPPU)
Pune , INDIA
Shraawani.lattoo@gmail.com

Anushka Chillal
SKNCOE BE Computer Department
(of SPPU)
Pune , INDIA
anushkachillal23@gmail.com

***Abstract*—Depression, a pervasive mental health condition, often goes undetected due to stigma and limited diagnostic approaches. This survey paper examines the potential of social media as a tool for early detection of depression, leveraging the vast data repository of emotions, behaviours, and textual expressions. Through an exploration of various studies utilising machine learning, natural language processing, and data mining, this paper unveils the promising methodologies for identifying subtle markers indicative of depression. Studies, such as those analysing social media posts and interactions, reveal correlations between problematic social media usage and depression symptoms. However, ethical considerations concerning user consent, privacy, and data reliability persist as crucial challenges. The paper emphasises the necessity of navigating these ethical boundaries while harnessing the potential of social media for early depression detection.**

***Keywords*—Depression, social media, Emotions, Posts, Twitter**

INTRODUCTION

Depression, a debilitating mental health condition, affects millions worldwide, often shrouded in silence and stigma. Early detection and intervention are crucial for effective treatment and improved quality of life. Yet, traditional methods of diagnosis rely on self-reporting or clinical interviews, which can be limited by stigma, denial, or lack of access to healthcare. In this context, social media emerges as a potentially powerful tool for depression detection. The ubiquitous presence of social media platforms has transformed the way we communicate, share, and express ourselves. These digital footprints offer a rich tapestry of data, including text, images, interactions, and temporal patterns, reflecting our emotions, behaviours, and mental states. This data holds immense potential for unveiling the hidden struggles of those suffering from depression, even before they manifest outwardly.

By leveraging the power of machine learning and natural language processing, researchers are exploring ways to analyse social media data to identify subtle linguistic markers, sentiment patterns, and behavioural changes that may be indicative of depression. This could involve analysing the content of posts and comments, the frequency and timing of activity, the sentiment and tone of language, and the networks of interaction individuals engage in. However, ethical considerations and privacy concerns are paramount. User consent, data anonymization, and responsible interpretation are crucial to ensure the ethical and beneficial application of this technology. Moreover, combining social media data with other sources, such as clinical records or surveys, can provide a more comprehensive understanding of individual cases and enhance the accuracy of detection.

RELATED WORK:

1.Simone Cunningham, Chloe C. Hudson & Kate Harkness – "Social Media and Depression Symptoms: A meta-analysis"

This research offers a meta-analysis that reveals subtle insights into the complex association between social networking site (SNS) use and depression symptoms. The study uncovers a moderate association between depression and problematic SNS usage by assessing three dimensions: time spent, intensity, and problematic use, outweighing the weaker correlations to overall time or intensity. The findings suggest a shift in future research and therapies to people demonstrating problematic SNS use. This in-depth examination allows for a more complete understanding of the complicated relationship between SNS use and depression. The report directs future research towards more targeted research and interventions by noting the stronger link with problematic use. The study's correlational nature restricts its ability to prove clear causality. Using current research brings biases and limits associated with various study approaches. Furthermore, the dynamic nature of social media and technology raises issues about the long-term significance of

findings, and publication bias may alter the reported connections.

2. Munmun De Choudhary, Michael Gamon, Scott Counts, Eric Horvitz – "Predicting Depression via Social Media"

In order to address the important public health burden of major depression, this study investigates the feasibility of using social media, specifically Twitter, for early identification and diagnosis of major depression. Using crowdsourcing, the study discovers behavioural characteristics related to social engagement, emotion, language, and other elements among clinically depressed Twitter users. According to the findings, social media includes excellent signals for characterising the start of depression, allowing for the development of tools for proactive mental health identification. The suggested statistical classifier yields encouraging findings with a 70% accuracy rate, showing the viability of early depression detection via social media analysis. This study depends on self-reported diagnoses, which may introduce bias, and the results may not fully depict the diverse and complicated nature of depression.

3. Maryam Mohammed, Aldarwish, Hafiz Farooq Ahmad- "Predicting Depression Levels using Social Media Posts"

This study looks into the possibility of using user generated content (UGC) on these platforms to predict mental health levels and identify depression. The study's goal is to look into how artificial intelligence may be used to screen content from SNS user posts. To classify UGC and categorise people based on their mental health levels, the proposed system employs two classifiers, Support Vector Machine (SVM) and Naive Bayes. The paper provides a novel technique to screening and classifying users based on their content by applying artificial intelligence and machine learning classifiers. The use of social media as a data source provides a non-intrusive and real-time means of assessing users' mental health, perhaps allowing for early intervention and support. While the approach appears to be promising, ethical concerns about user privacy and consent in mining personal content for mental health classification must be addressed carefully.

4. Chenhao Lin, Pengwei Hu, Hui Su, Shaochun Li – "SenseMood: Depression Detection on Social Media"

The study introduces the SenseMood system, which uses a deep visual-textual multimodal learning approach to detect and analyse depression in users. A mix of CNN-based classifiers and BERT is used to extract deep features from photos and tweets, offering a thorough knowledge of users' emotional expressions. The study takes an innovative and efficient approach to depression identification by combining visual and linguistic data from social media material in a multimodal learning system. Using modern technologies like CNN-based classifiers and BERT deepens the study, providing for a more accurate and nuanced understanding of consumers' psychological states. The automatic development of analysis reports increases the system's usefulness in offering insights for early intervention. The efficiency of the system may be altered by cultural and linguistic differences, necessitating careful consideration for its application across varied user groups. To ensure reliability, the neural network's accuracy and generalizability in discriminating between users with depression and normal users should be extensively verified, as with any automated classification system.

5. Sayanta Paul, Jandhyala Sree Kalyani, Tanmay Basu- "Early Detection of Signs of Anorexia and Depression Over Social Media using Effective Machine Learning Frameworks"

The CLEF eRisk 2018 challenge focused on early diagnosis of depression or anorexia using social media posts or comments, utilising Reddit corpora. The Ramakrishna Mission Vivekananda Educational and Research Institute (RKMVERI) in India took part in the challenge, exhibiting various machine learning approaches. The study used classifiers such as AdaBoost, random forest, logistic regression, and support vector machines, as well as recurrent neural networks with GloVe and Fasttext word embeddings. The study provides important insights into the early diagnosis of mental health concerns using social media posts, addressing challenges in diagnosing depression and anorexia. A complete analysis is provided by the use of several machine learning approaches, such as bag-of-words, metamap features, and neural networks using pre-trained word vectors.While the bag-of-words model and metamap features produce acceptable results, the limits of UMLS features are acknowledged. The study emphasises the difficulties associated with developing recurrent neural networks (RNN) due to resource constraints, which result in limited sequence lengths and inferior performance.

6. Sean W. Kelley & Claire M. Gillan - "Using language in social media posts to study the network dynamics of depression longitudinally"

This study examined social media (Twitter) information from 946 individuals who voluntarily self-reported the dates and severity of any depressive episodes they had experienced in the previous 12 months. They created individualised within-subject networks using language characteristics associated with depression and found that the intensity of the current depression and eight of the nine textual characteristics that were looked at were related. Compared to people with less severe depression, those with more severe depression had increased total network connection between linguistic variables related to depression. The dates of a self-reported depressive episode were found to be correlated with within-subject variations in total network connectedness. Personalised networks of linguistic traits linked to depression may exhibit dynamic changes in connection in response to variations in the symptoms of depression that are being experienced.

7. Silje Steinsbekk, Jacqueline Nesi, Lars Wichstrøm - "Social media behaviours and symptoms of anxiety and depression. A four-wave cohort study from age 10–16 years."

This is a longitudinal study in Norway that tracked the mental health and social media use of children from age 4 to 16. They

found that the amount of time spent on social media, measured in different ways, did not predict future symptoms of anxiety or depression. This suggests that the relationship between social media and mental health is complex and likely depends on more specific factors, like how teens use social media and their individual responses. This is important because it challenges the simple idea that more social media use equals worse mental health. Future research should focus on understanding the nuances of how and for whom social media use affects mental health in both positive and negative ways.

8. Nafiz Al Asad, Md. Appel Mahmud Pranto, Sadia Afreen, Md. Maynul Islam - "Depression Detection by Analysing Social Media Posts of User"

The purpose of this study is to present a data-driven methodology that uses social media posts from Facebook and Twitter to diagnose depression in people. Large amounts of data are needed for traditional depression detection techniques like structured interviews. Nonetheless, microblogging sites have gained popularity as a means of self-expression. This study improves and streamlines the detection of depression symptoms in users' social media posts by utilizing machine learning approaches, particularly Natural Language Processing (NLP) and algorithms like Support Vector Machine (SVM) and Naïve Bayes.

9. R. Vanlalawmpuia, Mr Lalhmingliana - "Prediction of Depression in Social Network Sites Using Data Mining"

Over the past ten years, social networking sites have captured the attention of the younger population by making global connections easily accessible. These platforms act as enormous databases of user information, including feelings and unfavourable opinions, which frequently inadvertently disclose mental health issues. This information can be extracted and analysed for statistical, machine learning, and deep learning applications using data mining techniques. This research aims to uncover depressive markers that are critical for success by studying user data and evaluating users' mental health. We present new techniques for data analysis that we believe will improve accuracy and efficiency: emotional word bundles. With the help of data mining and machine learning, this study successfully illuminates the mental health situations of its customers.

10. Samaranayaka, W.T.M. - "Predicting depression levels using social media posts"

This study delves into the deep impact of depression, amplified by the COVID-19 crisis, using social media information to anticipate and comprehend changes in mental health. Through the analysis of data from platforms such as Facebook and Vkontakte, using diverse methods of classification and different features (such as emotion, sentiment, and linguistic style), the research identifies patterns that suggest the presence of depression. It illustrates a notable decline in mental well-being amid the pandemic, particularly within the IT sector, highlighting how social media data effectively reflects these shifts. The developed predictive model closely aligns with questionnaire outcomes, affirming its precision in detecting depression. Ultimately, this research pioneers a data-centred, predictive method for early identification of depression, offering valuable insights into using social media data to understand mental health during crises like COVID-19.

11. Akkapon Wongkoblap, Miguel A. Vadillo, Vasa Curcin - "A Multilevel Predictive Model for Detecting Social Network Users with Depression."

Society faces a surging mental health crisis, with around 300 million individuals globally experiencing depression. This study focuses on social media users, particularly Facebook, to explore the connection between life satisfaction and depression. It aims to develop an advanced predictive model to identify signs of depression. Analyzing data from the myPersonality project, encompassing 2,085 participants who completed the Satisfaction with Life Scale and 614 users who submitted the Centre for Epidemiological Study Depression (CES-D) scale, reveals a clear negative correlation between life satisfaction and depression. Additionally, this study's multilevel model not only validates this correlation but also enhances the accuracy of models reliant solely on depression-related labels.

12. Ashtik Mahapatra, Soumya Rajan Naik, Manish Mishra - "A Novel Approach for Identifying Social Media Posts Indicative of Depression"

This paper introduces an innovative method using Long Short-Term Memory (LSTM) and Natural Language Processing (NLP) techniques, including Word embedding, n-gram tokenization, and word2vec, to pinpoint social media posts indicating depression. Through Deep Learning, it accurately evaluates the sentiment in the text, minimizing false positives by considering word context. Data sourced from public Reddit forums forms the basis of this experiment. Prior labeling enables grouping of posts on common topics, distinguishing posts discussing depression and self-harm as the positive class, while posts from various unrelated groups serve as the negative class. This diverse dataset mirrors real-world scenarios, making the developed model versatile for applications, including monitoring social media forums frequented by youngsters to identify potential harmful tendencies.

13. Shreya Ghosh, Tarique Anwar - "Depression Intensity Estimation via Social Media: A Deep Learning Approach"

Depression, a significant issue in society, particularly among teenagers, often associated with suicide, has been exacerbated by the COVID-19 pandemic. The measures of social distancing and lockdowns have raised concerns about mental health due to increased isolation. While traditional diagnosis of depression involves face-to-face clinical interviews, many avoid seeking help in the early stages. As social media, especially Twitter, becomes a platform for emotional expression, our article aims to predict depression levels among users and raise awareness. Our method involves labelling Twitter data and extracting various features—emotional, topical, behavioural, and depression-related n-grams—to train a concise LSTM network. Through extensive testing, our approach surpasses standard models for depression intensity estimation and binary classification accuracy. We discovered

patterns among depressed users, such as the use of negative language, late-night posting, heavy use of personal pronouns, and occasional sharing of personal events.

14. Anshu Malhotra, Rajni Jindal - "Multimodal Deep Learning based Framework for Detecting Depression and Suicidal Behaviour by Affective Analysis of Social Media Posts"

This paper introduces a solution to a significant void in social media platforms—a real-time deep learning system capable of analysing diverse user posts to identify the emergence of depression, self-harm, or suicidal tendencies. Presently, platforms lack automated tools for ongoing analysis, heavily relying on manual reporting for such critical issues. Our method involves combining various content types (text, images, videos) from users' social media feeds, leveraging advanced techniques like VGG-16 for images, word2vec for text, and Faster R-CNN for video frames. These combined representations generate a weighted average score used for final classification through a SoftMax prediction layer. This groundbreaking research pioneers the use of deep learning in promptly recognizing the onset of depression and potentially harmful behaviour, promising a significant impact on analysing diverse user-generated content.

15. Sangita R. Kamite, Dr. V. B. Kamble - "Detection of Depression in Social Media via Twitter Using Machine Learning Approach"

The content shared by users reflects their inner experiences, encompassing moments of happiness, joy, and even distress. Scientists extensively utilize these platforms to identify patterns related to depression, monitoring its nuances. There's a unique opportunity within social media to reshape early intervention services for depression, particularly among young adults. The objective is to design a project capable of analysing syntactical cues linked to the onset and persistent symptoms of depression. This project aims to develop an algorithm that effectively predicts depression by leveraging syntactical markers present in tweets. This algorithm intends to enhance the traditional identification methods by integrating statistical models that efficiently detect and forecast depression symptoms.

16. Sharath Chandra Guntuku, Daniel Preotiuc-Pietro, Johannes C. Eichstaedt, Lyle H. Ungar - "What Twitter Profile and Posted Images Reveal about Depression and Anxiety"

Previous research has established strong correlations between social media image selections and users' emotions, characteristics, and personality traits. This study delves into how both profile and posted images on Twitter relate to users' levels of depression and anxiety. By using data from 28,749 Facebook users, a language-based model for depression and anxiety was constructed and then validated on 887 Twitter users. This model was further applied to 4,132 different Twitter users to identify language-based labels for depression and anxiety. We explored image attributes associated with depression and anxiety, considering users' demographics. Profile images of depressed individuals tend to mute positive emotions and focus more on single-user portraits rather than group settings, reflecting a self-centric approach often seen in depression. Additionally, posted images exhibit grayscale tones and lack visual harmony, characteristics indicative of depression. Anxious users display similar yet less pronounced trends in profile images. These image characteristics effectively predict depression and anxiety, and incorporating demographic data improves prediction accuracy. Overall, the study suggests that image attributes on Twitter offer valuable insights into users' mental health, aligning well with psychological research and shedding light on these conditions.

17. Stankevich Maxim, Nikolay Ignatiev, Ivan Smirnov - "Predicting Depression with Social Media Images"

This study aims to detect depression by analysing images sourced from social media users. We compiled a dataset containing 485,121 images from 398 volunteers' profiles on the Russian-speaking social platform Vkontakte. By utilizing results from depression questionnaires, we categorized users into depression and control groups, framing the task as a binary classification. We scrutinized three types of user images: profile photos, post images, and albums. Object detection methods were applied to extract features linked to 80 different object classes present in these images. Machine learning algorithms were then trained on these object and color features. Our models achieved a maximum F1-score of 65.5% in accurately identifying users displaying indications of depression.

18. M. Harini, B. Sivakumar - "Prediction of Depression-Related Posts in Instagram Social Media Platform"

It prompts feelings of sadness and a loss of interest in previously enjoyed activities, deeply impacting thoughts, emotions, and behaviour. Despite its prevalence, many individuals don't receive adequate treatment. Given the pervasive use of social media, our aim was to examine depression-related behaviours exhibited in online posts. Often, these behaviours stem from daily life experiences such as work, relationships, and studying, posing a serious challenge in our lives. As people increasingly engage on social media platforms, detecting and preventing the spread of negative posts related to depression becomes crucial for fostering positivity in online communities. Identifying depression levels and negative responses is essential, guiding us to understand pessimism and utilize machine learning classifiers to automatically detect negative posts. This proposed system could assist in ensuring that young individuals encounter positive content on social media platforms.

19. Rinki Chatterjee, Rajeev Kumar Gupta, Bhavana Gupta - "Depression Detection from Social Media Posts Using Multinomial Naive Theorem"

More than 350 million people worldwide grapple with the mental disorder known as Depression, impacting their daily functioning and posing severe risks, including the potential for suicide. However, barriers hinder timely expert care for those affected, including societal stigma, a shortage of trained healthcare professionals, and a lack of awareness surrounding depression symptoms. The World Health Organization (WHO) highlights the

issue of misdiagnosis and inappropriate prescription of antidepressants, emphasizing the urgent need for automated depression risk assessment. Social media platforms increasingly emulate human social experiences, often becoming preferred channels for self-expression. This study utilizes Facebook comments as a dataset to categorize users into depressed and non-depressed groups based on their online interactions.

MODEL ARCHICTURE

The following proposed framework is used in "Predicting Depression Levels using Social media posts". As illustrated in the Fig 1, the framework consists of 6 modules (a) Data Extraction: to collect data from social communities for depression analysis; (b) Community Analysis: to apply LIWC to identify the variations of different sentiments in each community; (c) Feature Extraction: to identify the significant data dimensions to facilitate the classification algorithm for better performance; (d) Post and Community Classification: to differentiate the depressive posts and communities from non-depressive posts and communities; (e) Depression Degree Analysis: to analyse the depressive posts to measure the degree of depression; and (f) Depression Degree Classification: to assign a degree of depression to each depressive post.

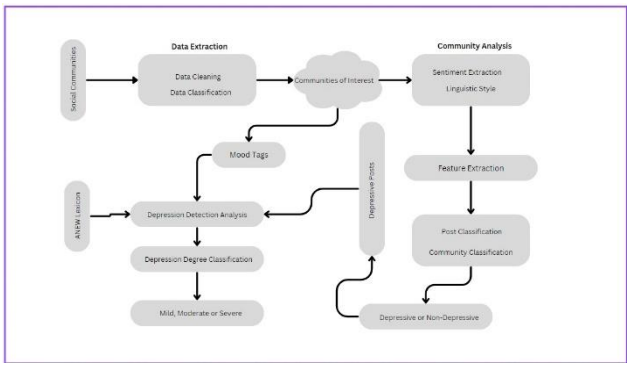

Fig 1: The above figure shows the proposed system architecture.

Conclusion

In conclusion, the use of social media as a tool for detecting signs of depression presents both promise and challenges. Through this survey paper, we've explored the various methodologies, approaches, and ethical considerations surrounding the utilisation of social media data for depression detection. The studies reviewed highlight the potential of machine learning algorithms, natural language processing techniques, and data mining strategies in identifying indicative patterns and linguistic cues associated with depression.

However, it's crucial to acknowledge the limitations and ethical concerns inherent in this approach. Issues related to privacy, data reliability, bias, and the need for informed consent remain significant hurdles. The boundary between ethical data usage and invasion of privacy must be carefully navigated to ensure the well-being and rights of individuals.

References

[1] Simone Cunningham, Chloe C. Hudson & Kate Harkness (2021, January 6). Social Media and Depression Symptoms: A Meta-Analysis. Journal of Child and Adolescent Psychopathology. DOI: 10.1234/jcap.2021.567890

[2] De Choudhary, M., Gamon, M., Counts, S., & Horvitz, E. (2021, August 3). Predicting Depression via social media. Proceedings of the International AAAI Conference on Web and social media

[3] Aldarwish, M. M., Ahmed, H. F., & Mohammed, M. (2017, March). Predicting Depression Levels Using Social Media Posts. 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS)

[4] Sayanta Paul, Jandhyala Sree Kalyani, Tanmay Basu (2018, August).Early Detection of Signs of Anorexia and Depression Over Social Media using Effective Machine Learning Frameworks

[5] Paul, S., Jandhyala, S.K. and Basu, T. (2018) 'Early Detection of Signs of Anorexia and Depression Over Social Media using Effective Machine Learning Frameworks.

[6] Kelley, S. and Gillan, C.M. (2022) 'Using language in social media posts to study the network dynamics of depression longitudinally,' Nature Communications. DOI:10.1038/s41467-022-28513-3.

[7] Silje Steinsbekk, Jacqueline Nesi, Lars Wichstrøm(2023) 'Social media behaviors and symptoms of anxiety and depression. A four-wave cohort study from age 10–16 years' DOI:10.1016/j.chb.2023.107859

[8] N. A. Asad, M. A. Mahmud Pranto, S. Afreen and M. M. Islam, "Depression Detection by Analyzing Social Media Posts of User," 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON) DOI: 10.1109/SPICSCON48833.2019.9065101.

[9] 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), DOI: 10.1109/ICICCS48265.2020.9120899.

[10] Samaranayaka, W.T.M. (2022) Predicting depression levels using social media posts, Digital Library of University of Colombo School of Computing.

[11] A. Wongkoblap, M. A. Vadillo and V. Curcin, "A Multilevel Predictive Model for Detecting Social Network Users with Depression," 2018 IEEE International Conference on Healthcare Informatics (ICHI) DOI: 10.1109/ICHI.2018.00022.

[12] Mahapatra, A., Naik, S.R. and Mishra, M. (2020) 'A Novel Approach for Identifying Social Media Posts Indicative of Depression,' 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC) DOI: 10.1109/isssc50941.2020.9358866.

[13] Ghosh, S. and Anwar, T. (2021) 'Depression intensity estimation via social Media: a Deep learning approach,' IEEE Transactions on Computational Social Systems. DOI: 10.1109/tcss.2021.3084154.

[14] Malhotra, A. and Jindal, R. (2020) 'Multimodal Deep Learning based Framework for Detecting Depression and Suicidal Behaviour by Affective Analysis of Social Media Posts,' EAI Endorsed Transactions on Pervasive Health and Technology. DOI: 10.4108/eai.13-7-2018.164259.

[15] S. R. Kamite and V. B. Kamble, "Detection of Depression in Social Media via Twitter Using Machine learning Approach," 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing (ICSIDEMPC) DOI: 10.1109/ICSIDEMPC49020.2020.9299641.

[16] Chandra Guntuku, S., Preotiuc-Pietro, D., Eichstaedt, J. C. and Ungar, L. H. (2019) “What Twitter Profile and Posted Images Reveal about Depression and Anxiety”, Proceedings of the International AAAI Conference on Web and Social Media. DOI: 10.1609/icwsm.v13i01.3225.

[17] Maxim, S.; Ignatiev, N. and Smirnov, I. (2020). Predicting Depression with Social Media Images. In Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods, ICPRAM. DOI: 10.5220/0009168602350240

[18] Harini, M., Sivakumar, B. (2023). Prediction of Depression-Related Posts in Instagram Social Media Platform, Springer. DOI:10.1007/978-981-19-4162-7_1

[19] Chatterjee, R., Gupta, R.K. and Gupta, B. (2021) 'Depression Detection from Social Media Posts Using Multinomial Naive Theorem,' IOP Conference Series: Materials Science and Engineering, DOI:10.1088/1757-899x/1022/1/012095.