UNVEILING DEPRESSION: A REVIEW OF NATURAL LANGUAGE PROCESSING AND MACHINE LEARNING APPROACHES FOR DETECTING SYMPTOMS IN SOCIAL MEDIA TEXTS

[1]Chinmay Kshirsagar

[1]Student
[1]Prof Dhananjay Bhagat,
[1]Assitant Professor, Nagpur, India

Abstract: The increasing number of people suffering from mental health issues, especially depression, is a growing problem on a worldwide scale. The abundance of textual data made available by people's frequent online expressions of emotion and ideas on social media platforms is a promising resource for the early diagnosis and treatment of depression. In this review article, we take a look at the latest methods for identifying depressive symptoms in social media posts using machine learning and natural language processing. We review the pros and cons of many techniques for reliably detecting depressive symptoms, including topic modelling, sentiment analysis, deep learning, linguistic feature extraction, and sentiment analysis. Data privacy issues, sample bias, and language subtleties are just a few of the obstacles we investigate when it comes to using social media data for mental health evaluation. Our hope is that this review will shed light on the present state of depression diagnosis from social media texts and point the way toward future studies that will improve these approaches while also considering their ethical implications.

KeyWords-. Machine Learning, NLP, Depression,Textual Data.

I. INTRODUCTION

An increasing number of people across all socioeconomic groups are suffering from mental health issues, with depression being one of the most common. This has become a major worldwide problem in recent years. By 2030, depression is expected to surpass all other causes of disability globally, as reported by the World Health Organization (WHO). Amidst this burgeoning crisis, the advent of social media platforms has transformed the landscape of mental health discourse. Individuals increasingly turn to online platforms to express their thoughts, feelings, and experiences, presenting a unique opportunity for leveraging textual data for early detection and intervention of depression.

The vast expanse of social media data offers unprecedented insights into individuals' emotional and psychological states, providing researchers and practitioners with a rich source of information for understanding and addressing mental health concerns. Natural Language Processing (NLP) and Machine Learning (ML) techniques have emerged as powerful tools for mining and analyzing this textual data to detect signs and symptoms of depression. By harnessing the computational power of NLP and ML, researchers can sift through large volumes of social media texts to identify linguistic patterns, sentiment cues, and semantic structures indicative of depressive symptoms.

The purpose of this review article is to provide a thorough analysis of the most recent developments in natural language processing and machine learning as they pertain to the detection of depressive symptoms in online social media posts. We explore many techniques that may be used to detect depressive symptoms, such as sentiment analysis, topic modelling, deep learning, and linguistic feature extraction. We explain how each of these approaches works and what limits they have. Moreover, we go into the several complex issues related to using social media data for mental health evaluation, including ethical concerns and technological barriers. Key challenges include ensuring data privacy and confidentiality, mitigating sample bias inherent in social media datasets, and navigating the intricacies of language and culture in diverse online communities. By critically examining these challenges, we aim to foster a nuanced understanding of the complexities involved in leveraging social media data for depression detection and pave the way for ethically responsible and methodologically rigorous research practices.

We hope that scholars and practitioners will find this review useful in understanding the state of the art of depression identification using social media writings. Additionally, we aim to outline promising directions for future research aimed at enhancing the effectiveness and ethical implications of NLP and ML methodologies in this domain. Ultimately, our overarching goal is to contribute to the advancement of innovative approaches for unveiling depression and facilitating timely interventions to support individuals' mental well-being in the digital age.

II. Literature Review:

Depression is a prevalent mental health issue globally, and its detection through social media textual data analysis has gained significant attention in recent years. Various studies have explored the application of natural language processing (NLP) and machine learning (ML) techniques for this purpose.

Kabir et al. (2023) introduced a methodology for depression detection from social media textual data using NLP and ML techniques. They emphasized the importance of linguistic features and sentiment analysis in identifying depressive patterns. Similarly, K. S et al. (2024) focused on fake news detection using ML and NLP techniques, highlighting the significance of identifying misinformation to prevent its spread.

In a different context, Illahi (2022) proposed an ensemble machine learning approach for stress detection in social media texts. Their study emphasized the effectiveness of combining multiple ML algorithms to improve detection accuracy. Nadeem et al. (2022) presented a hybrid deep learning SSCL framework for depression detection, showcasing the power of deep learning architectures in handling complex social networking data.

Sentiment analysis has also been widely employed in this domain. Patwardhan et al. (2023) conducted sentiment analysis on Twitter data related to the National Eligibility-Cum Entrance Test, demonstrating the utility of sentiment analysis in understanding public opinion. Similarly, Inamdar et al. (2023) utilized sentiment analysis for mental stress detection on Reddit posts, illustrating the potential of NLP in gauging emotional states.

Furthermore, Kashid et al. (2023) developed a depression monitoring system via social media data, integrating ML techniques into a comprehensive framework for continuous monitoring. Dhaker et al. (2023) conducted a comprehensive review of data analysis techniques for depression detection on social media, providing insights into various approaches and their effectiveness.

In addition to depression, other mental health issues such as cyberbullying have also been addressed using ML and NLP techniques. Sahana et al. (2023) conducted a comparative analysis of ML techniques for cyberbullying detection, emphasizing the importance of context and linguistic patterns. Similarly, Nijhawan et al. (2022) focused on stress detection using NLP and ML over social interactions, highlighting the need for proactive measures to address online harassment.

Moreover, the detection of fake news has emerged as a critical area of research. Prachi et al. (2022) developed algorithms for fake news detection using ML and NLP, aiming to curb the spread of misinformation. Agravat et al. (2024) focused on fake social media profile detection, underscoring the importance of identifying and reporting fraudulent accounts.

In summary, recent studies have demonstrated the potential of NLP and ML techniques in detecting various mental health issues and combating misinformation on social media platforms. These advancements pave the way for more effective interventions and support systems to promote mental well-being and ensure the integrity of online information.

III. Conclusion:

In conclusion, this review has shed light on the burgeoning field of natural language processing (NLP) and machine learning (ML) in detecting symptoms of depression within social media texts. Through an exploration of diverse methodologies and studies, it becomes evident that the fusion of NLP and ML offers promising avenues for early detection and intervention in mental health disorders.

The reviewed literature showcases the efficacy of various linguistic features and machine learning algorithms in identifying markers of depression, ranging from sentiment analysis to linguistic style and syntactic patterns. Ensemble learning techniques, deep learning architectures, and hybrid frameworks have emerged as powerful tools for capturing the nuanced linguistic cues embedded within social media data.

Despite the progress made, challenges persist, particularly concerning data privacy, model generalizability, and interpretability. Ethical considerations surrounding user consent and data anonymization underscore the importance of responsible data handling practices in mental health research.

Moving forward, interdisciplinary collaborations and advancements in explainable AI (XAI) will be instrumental in addressing these challenges and fostering trust among users and healthcare practitioners. Moreover, future research directions may involve exploring multimodal approaches that integrate textual, visual, and audio-visual data sources to provide a more comprehensive understanding of mental health states. In essence, the review underscores the transformative potential of NLP and ML in unveiling depression symptoms in social media texts. By navigating ethical considerations, embracing interdisciplinary collaborations, and advancing XAI techniques, researchers can pave the way for more robust, ethical, and culturally sensitive approaches to mental health detection in the digital age.

IV. References:

1. F. Kabir, M. A. Hossain, A. F. M. Rahman, and S. Z. Mishu, "Depression Detection From Social Media Textual Data Using Natural Language Processing and Machine Learning Techniques," in 26th International Conference on Computer and Information Technology (ICCIT), 2023. DOI: 10.1109/iccit60459.2023.10441612.

2. K. S, A. J. V, M. S, D. R, A. G, and T. R. V, "Unveiling Fake News with Machine Learning and Natural Language Processing Techniques," in Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 2024. DOI: 10.1109/ic-etite58242.2024.10493692.

3. M. Illahi, "Ensemble Machine Learning Approach for Stress Detection in Social Media Texts," in Quaid-e-Awam University Research Journal of Engineering, Science & Technology, 2022. DOI: 10.52584/qrj.2002.15.

4. Nadeem et al., "Depression Detection Based on Hybrid Deep Learning SSCL Framework Using Self-Attention Mechanism: An Application to Social Networking Data," in Italian National Conference on Sensors, 2022. DOI: 10.3390/s22249775.

5. V. Patwardhan et al., "Analysing The Sentiments Of Marathi-English Code-Mixed Social Media Data Using Machine Learning Techniques," in International Conference on Emerging Smart Computing and Informatics (ESCI), 2023. DOI: 10.1109/esci56872.2023.10100304.

6. T. Sravanthi et al., "Detection of Mentally Distressed Social Media Profiles Using Machine Learning Techniques," in IOP Conference Series: Materials Science and Engineering, 2020. DOI: 10.1088/1757-899x/981/2/022056.

7. E. Chandralekha et al., "Sentiment Analysis of National Eligibility-Cum Entrance Test on Twitter Data Using Machine Learning Techniques," in IoT, Cloud and Data Science, 2023. DOI: 10.4028/p-8fy5ca.

8. L. D'Cruz et al., "Depression Prediction from Combined Reddit and Twitter Data using Machine Learning," in 2nd International Conference for Innovation in Technology (INOCON), 2023. DOI: 10.1109/inocon57975.2023.10101174.

9. V. Sahana et al., "A Comparative Analysis of Machine Learning Techniques for Cyberbullying Detection on FormSpring in Textual Modality," in International Journal of Computer Network and Information Security, 2023. DOI: 10.5815/ijcnis.2023.04.04.

10. M. K. Kabir et al., "Detection of Depression Severity from Bengali Social Media Posts for Mental Health: A Study Using Natural Language Processing Techniques (Preprint)," in JMIR formative research, 2022. DOI: 10.2196/36118.

11. N. Hudda et al., "Detection of Cyberbullying on Social Media," in International Journal of Computer Applications, 2023. DOI: 10.5120/ijca2023922702.

12. S. Inamdar et al., "Machine Learning Driven Mental Stress Detection on Reddit Posts Using Natural Language Processing," in Human-Centric Intelligent Systems, 2023. DOI: 10.1007/s44230-023-00020-8.

13. O. Kashid et al., "Depression Monitoring System via Social Media Data using Machine Learning frameworkk," in International Journal for Research in Applied Science and Engineering Technology, 2023. DOI: 10.22214/ijraset.2023.51811.

14. T. Dhaker et al., "Detecting Depression on Social Media: A Comprehensive Review of Data Analysis, Deep Learning, NLP, and Machine Learning Approaches," in International Journal of Scientific Research in Computer Science Engineering and Information Technology, 2023. DOI: 10.32628/cseit2390510.

15. T. Nijhawan et al., "Stress detection using natural language processing and machine learning over social interactions," in Journal of Big Data, 2022. DOI: 10.1186/s40537-022-00575-6.

16. N. N. Prachi et al., "Detection of Fake News Using Machine Learning and Natural Language Processing Algorithms," in Journal of Advances in Information Technology, 2022. DOI: 10.12720/jait.13.6.652-661.

17. G et al., "Detection of Stress Over Social Interactions Using Machine Learning and Natural Language Processing," in International Conferences on Contemporary Computing and Informatics, 2023. DOI: 10.1109/ic3i59117.2023.10397916.

18. M. S. Islam et al., "Natural Language Processing and Machine Learning Approaches to Detect Bangla Hate Speech on Social Media," in 26th International Conference on Computer and Information Technology (ICCIT), 2023. DOI: 10.1109/iccit60459.2023.10441452.

19. Agravat et al., "Fake Social Media Profile Detection and Reporting Using Machine Learning," in International Journal of Advanced Research in Science, Communication and Technology, 2024. DOI: 10.48175/ijarsct-16695.

20. K. K. Dixit et al., "Analyzing Textual Data for Mental Health Assessment: Natural Language Processing for Depression and Anxiety," in IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics Engineering, 2023. DOI: 10.1109/upcon59197.2023.10434291.

21. M. A. Wani et al., "Depression Screening in Humans With AI and Deep Learning Techniques," in IEEE Transactions on Computational Social Systems, 2023. DOI: 10.1109/tcss.2022.3200213.

22. J. Sharma and V. Tomer, "Depression detection using sentiment analysis of social media data," in Technologies Applied to Electronics Teaching, 2022. DOI: 10.1063/5.0104192.

23. P. Mann et al., "Detecting Depression from Social Media Data as a Multiple-Instance Learning Task," in Affective Computing and Intelligent Interaction, 2022. DOI: 10.1109/acii55700.2022.9953811.

24. W. Zhang and S. Ram, "A Comprehensive Methodology for Extracting Signal from Social Media Text Using Natural Language Processing and Machine Learning," SSRN, 2015. DOI: 10.2139/ssrn.2880834.