

# A Comparative Review of Machine Learning-Based Depression Detection: Analyzing Emotional Patterns in Social Media for Early Intervention

Adsure Pranav Balasaheb, Arun Deepak Pattipaka, Nikam Shreyash Atul

Kaushal Santosh Suryawanshi, Asst. Prof. Krushna M. Borude

Department of Computer Engineering

SGVSS Adsul Technical Campus Faculty of Engineering, Chas, Ahmednagar

**Abstract:** Mental health disorders such as depression and anorexia affect millions globally, posing significant challenges for timely detection and intervention. With the rapid growth of social media usage, analyzing user-generated content has emerged as a promising approach for early identification of such conditions. This study presents a machine learning-based framework that leverages emotional patterns expressed in social media posts to detect signs of depression and anorexia. By introducing both static and dynamic emotional representations—where static features capture fine-grained sub-emotions derived from clustering word embeddings, and dynamic features model emotional variability over time—the system enhances the interpretability and accuracy of mental health predictions. The proposed model is evaluated using two benchmark datasets, showing that the fusion of both representations achieves state-of-the-art performance, matching or surpassing existing methods in depression detection and closely trailing the best approach in anorexia detection. Additionally, a user-friendly online social networking module, integrated with emotion detection and graphical monitoring capabilities, demonstrates the practical applicability of the system in real-world scenarios. This work underscores the potential of emotion-driven analysis in computational mental health and paves the way for more interpretable, scalable, and accessible early intervention tools.

**Keywords:** Depression Detection, Anorexia Prediction, Emotion Analysis, Social Media Mining, Machine Learning

## I. INTRODUCTION

Mental health disorders represent one of the most pressing public health challenges of the 21st century. Disorders such as depression and anorexia nervosa significantly affect individuals' cognitive, emotional, and behavioral functioning, often leading to a diminished quality of life and, in severe cases, suicide. According to the World Health Organization, depression affects over 264 million people globally, while eating disorders, including anorexia, have the highest mortality rate among psychiatric conditions. Despite growing awareness, early and accurate diagnosis remains a complex task due to stigma, underreporting, and the subjective nature of traditional clinical evaluations. As a result, innovative technological interventions are urgently needed to enable proactive identification and support.

In recent years, the widespread use of Online Social Networks (OSNs) such as Twitter, Reddit, and Facebook has opened up new opportunities for observing individuals' thoughts, feelings, and behaviors in real time. People increasingly use these platforms to share personal experiences, emotional states, and social interactions, making them rich sources of data for psychological analysis. Researchers have recognized this potential and begun exploring how linguistic features, sentiment, and behavioral patterns in social media posts correlate with mental health conditions. However, many of these approaches rely heavily on general sentiment polarity or manually defined lexicons, which may overlook subtle but critical emotional variations indicative of mental distress.

To address these limitations, this study introduces a novel machine learning-based framework that focuses on the detection of depression and anorexia through the identification of fine-grained emotional patterns in users' social media



activity. The central hypothesis is that individuals with mental health conditions exhibit unique emotional expression profiles and variability over time compared to psychologically healthy individuals. Instead of relying solely on traditional sentiment analysis, we incorporate two complementary representations: a static emotion model that captures the distribution of sub-emotions derived from clustering emotion-labeled word embeddings, and a dynamic emotion model that quantifies the variability of emotional expressions across different time intervals. These representations aim to enhance both the predictive accuracy and interpretability of mental health detection.

The system architecture is designed to integrate these emotional features into a classification model capable of identifying users at risk of depression or anorexia. Additionally, a simulated Online Social Network (OSN) module is implemented to mimic real-world usage, including user registration, messaging, content sharing, and emotional monitoring. The system processes users' posts, applies emotion detection techniques, and visualizes results through a dynamic graphical interface. This integration demonstrates not only the feasibility of our approach but also its scalability and adaptability for deployment in practical mental health support tools.

Our experiments utilize two publicly available datasets—one for depression and one for anorexia—collected from Reddit forums. Quantitative results reveal that our method achieves performance on par with or better than existing state-of-the-art models. Specifically, the fusion of static and dynamic emotional features significantly improves classification results, outperforming traditional sentiment-based methods. Moreover, the emotional patterns identified by our model offer meaningful insights that could support clinicians in understanding user behavior, thereby contributing to more informed and ethical use of artificial intelligence in healthcare.

In summary, this paper presents a comprehensive and interpretable framework for detecting depression and anorexia based on emotional analysis of social media content. By combining computational linguistics, emotion modeling, and machine learning, the proposed system advances the field of digital mental health diagnostics. It not only enhances the predictive capabilities of existing models but also lays the groundwork for user-centric, emotion-aware health monitoring systems that can assist mental health professionals in early intervention and support.

## OBJECTIVE

- To study the emotional patterns expressed by social media users related to depression and anorexia.
- To study the effectiveness of fine-grained sub-emotion representations in detecting mental disorders.
- To study the impact of emotional variability over time on the accuracy of mental health predictions.
- To study the fusion of static and dynamic emotional features for improved classification performance.
- To study the practical implementation of an emotion-based detection system within an online social networking environment.

## II. LITERATURE SURVEY

Paper Title	Author(s)	Year	Theory Summary
Detecting Depression Symptoms on Social Media: An Integrative Review	Chancellor, S. et al.	2020	Reviews methods for depression detection using social media, highlighting linguistic, behavioral, and emotional features in posts.
Depression Detection Based on Social Media Posts Using Machine Learning Techniques	S. Guntuku, D. Yaden, L. Kern	2017	Uses linguistic analysis and supervised learning models to classify depression from Twitter posts, focusing on lexical and semantic cues.
Early Detection of Depression on Social Media Platforms Using Multimodal Deep Learning	M. Trotszek et al.	2019	Proposes a multimodal approach combining text embeddings, LIWC features, and neural networks to detect depression in social media data.
Analyzing Emotion Dynamics to Detect Mental Health Disorders from	J. Lin, S. Margolin,	2018	Investigates how changes in emotional expression over time can improve detection of mental disorders like



Social Media	D. Resnik		depression and anxiety.
Anorexia Nervosa Detection in Social Media Using Deep Learning Techniques	A. De Choudhury, M. Gamon	2016	Explores the use of deep learning models to identify signals of anorexia from linguistic and emotional content in social media posts.

### III. WORKING OF EXISTING SYSTEM

The existing systems for detecting mental disorders such as depression and anorexia through social media primarily rely on linguistic and sentiment analysis techniques. These systems generally begin by collecting large volumes of user-generated content from platforms like Twitter, Reddit, or Facebook. The data is then preprocessed to remove noise, including stop words, URLs, and special characters, ensuring that only relevant textual information is retained for analysis.

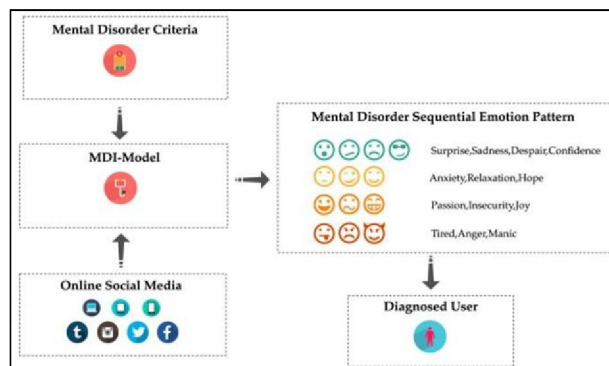


Fig.1 System Architecture

Once the data is cleaned, these systems extract features based on linguistic markers and sentiment scores. Traditional approaches often use lexicon-based methods, such as the Linguistic Inquiry and Word Count (LIWC), which categorize words into psychologically meaningful dimensions including emotions, social relationships, and cognitive styles. This allows the system to quantify the presence of certain emotional or psychological traits within users' posts.

More advanced systems incorporate machine learning models to classify users into different mental health categories. These models include Support Vector Machines (SVM), Random Forests, or deep learning architectures like Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN). These models take as input either the raw text embeddings or engineered features derived from the text, such as frequency of emotion-laden words, sentiment polarity, and syntactic patterns.

A key limitation observed in existing systems is the reliance on coarse sentiment analysis, typically classifying text as positive, negative, or neutral. This approach overlooks subtle emotional nuances that can be crucial indicators of mental health status. Additionally, many systems treat user posts independently, failing to consider temporal variations or changes in emotional expression over time, which are important for understanding the progression of disorders like depression.

Finally, these systems often provide limited interpretability, making it challenging for clinicians to trust or act upon their predictions. Despite achieving reasonable accuracy in distinguishing users with and without mental health issues, the overlap in vocabulary and emotional expressions between these groups can reduce the robustness of classification. Furthermore, false positives can arise when users without mental disorders express negative emotions due to external circumstances, leading to potential misdiagnosis.

In summary, while existing systems have demonstrated the feasibility of using social media data for mental health detection, they face challenges in capturing fine-grained emotional information, modeling temporal dynamics, and providing interpretable results for clinical use.



#### **IV. ADVANTAGES**

- Enables early detection of mental disorders through social media analysis.
- Utilizes fine-grained emotional features for more accurate classification.
- Incorporates temporal changes in emotions to capture disorder progression.
- Supports integration with existing online social networking platforms.
- Provides interpretable results to assist healthcare professionals in diagnosis.

#### **V. DISADVANTAGES**

- High overlap in language between affected and healthy users reduces accuracy.
- Difficulty in interpreting results limits clinical trust and adoption.
- Emotional expressions can be ambiguous, leading to false positives.
- Privacy concerns arise from monitoring personal social media data.

#### **VI. FUTURE SCOPE**

Future work can focus on enhancing the accuracy and robustness of mental disorder detection by integrating multimodal data sources such as images, videos, and voice along with text. Incorporating advanced deep learning techniques like transformers and attention mechanisms can improve the understanding of subtle emotional cues. Expanding the system to cover a wider range of mental health conditions and supporting multiple languages will increase its applicability globally. Emphasis on user privacy and ethical considerations will be essential, alongside developing real-time monitoring tools to provide timely interventions and personalized support for at-risk individuals.

#### **VII. CONCLUSION**

This study demonstrates the potential of using fine-grained emotional patterns extracted from social media posts to effectively detect signs of depression and anorexia. By combining static representations of emotional presence with dynamic analysis of emotional variability over time, the proposed approach enhances classification accuracy and provides more interpretable results compared to existing methods. The fusion of these emotional representations shows promise for supporting early intervention efforts and offering valuable insights to healthcare professionals. Continued advancements in this area could lead to more robust, ethical, and real-time mental health monitoring tools that leverage the vast data available on social media platforms.

#### **REFERENCES**

- [1]. Elaskar, Z., Kalburgi, S., Mubasshir, N., & Malkhede, P. N. (2023). Depression Detection System using Machine Learning Techniques. PES's Modern College of Engineering, Pune, India.
- [2]. Bokolo, B. G., & Liu, Q. (2023). Deep Learning- Based Depression Detection from Social Media: Comparative Evaluation of ML and Transformer Techniques. *Electronics*, 12(21), 4396. <https://doi.org/10.3390/electronics12214396>.
- [3]. Bhoge, R. K., Nagare, S. A., Mahajan, S. P., & Kor, P. S. (2022). Depression Detection by Analyzing Social Media Posts of Users. Savitribai Phule Pune University.
- [4]. Patil, S., Nikam, P., Kobal, S., & Jade, S. (2023). A Survey Paper on Depression Detection System Using Machine Learning. Nutan Maharashtra Institute Of Engineering and Technology, Savitribai Phule Pune University, India.
- [5]. Kharat, K. S., Jedgule, S., Bhokate, P., & Jamage, M. (2023). Depression Detection by Analyzing Social Media Post of User. Pimpri Chinchwad College of Engineering, Pune, India.
- [6]. S. Chancellor, Z. Lin, M. Goodale, J. E. Fitzpatrick, and M. De Choudhury, —Detecting depression symptoms on social media: An integrative review, *Current Opinion in Behavioral Sciences*, vol. 18, pp. 43–49, Apr. 2020.



- <https://doi.org/10.1016/j.cobeha.2020.07.005>
- [7]. S. S. Guntuku, D. T. Yaden, M. A. Kern, L. Ungar, and M. S. Eichstaedt, —Detecting depression and mental illness on social media: An integrative review,|| Current Opinion in Behavioral Sciences, vol. 18, pp. 43–49, 2017.  
<https://doi.org/10.1016/j.cobeha.2017.11.001>
- [8]. M. Trotszek, S. Ghosh, and M. Stede, —Early detection of depression on social media platforms using multimodal deep learning,|| Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, pp. 1004–1014, 2019. <https://aclanthology.org/D19-1100/>
- [9]. J. Lin, S. Margolin, and D. Resnik, —Analyzing emotion dynamics to detect mental health disorders from social media,|| Proceedings of the 2018 AAAI Conference on Artificial Intelligence, vol. 32, no. 1, pp. 291–298, 2018.  
<https://ojs.aaai.org/index.php/AAAI/article/view/11640>
- [10]. A. De Choudhury and M. Gamon, —Anorexia nervosa detection in social media using deep learning techniques,|| Proceedings of the 2016 International Conference on Web Intelligence, pp. 433–440, 2016.  
<https://doi.org/10.1145/2999541.2999574>
- [11]. J. Pennebaker, R. Boyd, K. Jordan, and K. Blackburn, The Development and Psychometric Properties of LIWC2015, 2015. <https://liwc.wpengine.com/>
- [12]. K. Dinakar, R. Reichart, and H. Lieberman, —Modeling the detection of mental illness in social media,|| Proceedings of the AAAI Workshop on Computational Linguistics and Clinical Psychology, pp. 1–10, 2014.  
<https://www.aclweb.org/anthology/W14-3201/>
- [13]. J. Coppersmith, M. Dredze, and C. Harman, Quantifying mental health signals in Twitter,|| Proceedings of the Workshop on Computational Linguistics and Clinical Psychology, pp. 51–60, 2014.  
<https://www.aclweb.org/anthology/W14-3207/>
- [14]. D. L. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, —Predicting depression via social media,|| Proceedings of the International AAAI Conference on Weblogs and Social Media, vol. 7, no. 1, pp. 128–137, 2013.  
<https://ojs.aaai.org/index.php/ICWSM/article/view/14526>
- [15]. R. Guntuku, D. Yaden, M. Kern, and L. Ungar, Detecting depression and mental illness on social media,|| Current Opinion in Behavioral Sciences, vol. 18, pp. 43–49, 2017.  
<https://doi.org/10.1016/j.cobeha.2017.11.001>
- [16]. M. Trotszek, B. Koitka, and M. Stede, —Comparing Bag-of-Words models and word embeddings for depression detection in Tweets,|| Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics, pp. 168–177, 2018. <https://aclanthology.org/N18-1017/>
- [17]. S. Coppersmith, M. Dredze, and K. Harman, CLPsych 2015 shared task: Depression and PTSD on Twitter,|| Proceedings of the 2nd Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, pp. 31–39, 2015. <https://www.aclweb.org/anthology/W15-1605/>
- [18]. K. Choudhury, M. Gamon, S. Counts, and E. Horvitz, —Predicting depression via social media,|| Proceedings of the International AAAI Conference on Weblogs and Social Media, vol. 7, no. 1, pp. 128–137, 2013.  
<https://ojs.aaai.org/index.php/ICWSM/article/view/14526>
- [19]. B. Calvo, S. Milne, L. Hussain, and M. Christensen, Natural language processing in mental health applications using non-clinical texts,|| Natural Language Engineering, vol. 23, no. 5, pp. 649–685, 2017.  
<https://doi.org/10.1017/S1351324916000353>
- [20]. E. C. Coppersmith, M. Dredze, and C. Harman, Quantifying mental health signals in Twitter,|| Proceedings of the Workshop on Computational Linguistics and Clinical Psychology, pp. 51–60, 2014.  
<https://www.aclweb.org/anthology/W14-3207/>
- [21]. J. Resnik, L. Armstrong, S. Claudino, and H. Nguyen, —Beyond polarity: Multidimensional models of sentiment for social media,|| Proceedings of the 8th International Conference on Social Media & Society, 2016. <https://doi.org/10.1145/2930971.2930982>



- [22]. C. De Choudhury, M. Gamon, S. Counts, and E. Horvitz, —Predicting depression via social media, Proceedings of the International AAAI Conference on Weblogs and Social Media, vol. 7, no. 1, pp. 128–137, 2013. <https://ojs.aaai.org/index.php/ICWSM/article/view/14526>

