TuijinJishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 46 No. 2 (2025)

A Comprehensive Review on Mental Health Prediction Using Social Media Data

Subhashree Mohapatra[1], Aisha Abeid Abdulrahman[2], Dr. Kusum Lata[3]

[1, 2, 3]Department of Engineering and Technology, Sharda University, Greater Noida, Uttar Pradesh, India

Abstract:-Mental health affects how we think, feel, act, handle stress, relate to others, and make choices. Machine learning is increasingly used for early mental illness detection, particularly through social media data. This paper reviews machine learning models, algorithms, and applications for mental health prediction. Our methodology integrates sentiment analysis, multimodal data fusion, and AI-driven treatment. We emphasize multilingual models and cultural adaptation in mental health applications. Additionally, we explore AI-driven chatbots for real-time support and intervention. This review systematically evaluates existing approaches, identifies research gaps, and highlights the need for inclusive, robust, and multimodal frameworks. We discuss the advantages and challenges of using machine learning in mental health care and introduce novel concepts like real-time AI interventions, digital twins for mental health, and multilingual sentiment analysis. Finally, we outline future directions, focusing on multimodal analysis and chatbot-based mental health support.

Keywords*: AI-driven Mental Health Support, Mental Health Prediction, Multimodal Machine Learning, Social Media Analytics.*

1. Introduction

How people cope with stress, relatedness to others, and informed decision-making on a daily basis all fall under the condition of mental health. Mental health disorders affect millions around the globe, but many remain undiagnosed by stigma, access to care, or even self-awareness. As people share their feelings more often than ever on social media, researchers around the world have recognized the potential of this unique and underexplored data source for analyzing mental health.

Recent developments in machine learning (ML) and natural language processing (NLP) have allowed researchers to use these digital footprints to detect early indicators of mental illness[1]. It presents current research on machine learning models and techniques for predicting mental health, particularly those that utilize social media data, providing the basis for this review.

Approaches here can be traditional machine learning, deep learning, transformers-based models (e.g., BERT, XLM-R, mBERT etc.) or cross-modal approaches that combine text, image and/or video. Multilingual adaptation and cultural sensitivity, as well as new applications, such as real-time AI-driven chatbots, are some of emerging trends that our review highlights.

Here, we present a systematic assessment of the existing literature, summarising critical gaps and areas of opportunity moving forward in the pursuit of inclusive and ethically aligned AI for mental health care. This paper makes thefollowing contributions:

- A systematic review of traditional and modern machine learning techniques for mental health prediction using social media.
- Analysis of deep learning and transformer-based models for sentiment and emotion classification.
- Survey of real-time AI interventions, including chatbot-based mental health support.

- Discussion of ethical considerations, data privacy, and the development of standardized benchmarking frameworks.

This review also seeks to summarizes the existing studies and provide a structured survey of the methods, challenges, and future directions in this domain. The rest of the paper is organized as follows: In Section 2, we summarize previous work on mental health and sentiment analysis in social media; In Section 3, we provide a review of the methodologies for mental health prediction; Section 4 summarizes common datasets and benchmarking practices; in Section 5, we present ethical considerations and challenges; in Section 6, we discuss real-world applications and case studies; and finally, in Section 7, we outline future directions, including recent advances in AI-driven interventions and digital twins that could be applied to mental health monitoring.

2. Mental Health and Sentiment Analysis inSocial Media

The presence of mental health and social media have received significant interest in recent years due to theintegration of online platforms into our daily lives and the growing demand for scalable mental health solutions. Mental health is our emotional, psychological, and social well-being that affects how we think, feel, and act. Diseases such us depression, anxiety, stress, [2] affect over millions of people on the earth, mostly undiagnosed because of social stigma or lack of mental health issues service. Historically, most diagnostics occur via self-reported surveys and clinical interviews, which, lacks the ability to reflect real-time psychological flows, and can often miss at-risk individuals in the spaces they frequent spend time digitally.

Sentiment Analysis — the computational study of opinions, sentiments, attitudes, and emotions expressed in text — has proved to be an effective method for identifying mental health markers in user-generated text that occurred on social media platforms. It enables researchers to extract users' affective states from the posts based on the polarity (positive, negative, or neutral) and subjectivity of posts. While early methods relied on rule-based systems and lexicons, recent developments use machine learning and deep learning to allow readers to express complex affective emotions [3].

Unstructured data from social media like Twitter, Reddit, Facebook, and Instagram contain rich information about what users are thinking, feeling, and doing on a daily basis. For many users, these platforms have become digital diaries, in which they share their battles and emotional states, sometimes more openly than in interpersonal exchanges [3]. Reddit, for example, is popular in research due to its anonymous and topic-specific structure and is frequently referred to as a social media site where mental health discussions take place. Twitter is crisp and now; Facebook is chew and context. Instagram is an image-centric social media platform but largely ignores the value of emotional expression considered in caption, hashtags, and user engagement behavior to the emotional content of the posts [4].

Previous works found that the linguistic and behavioral aspects in social media posts are predictive of a number of mental health conditions. As an example, De Choudhury et al. (2013) employed Reddit data to model depressive behavior using a frequency of posts, sentiment shifts and linguistic markers [5]. Similarly, Coppersmith et al. (2014) developed annotated datasets through self-reporting of mental health diagnoses from Twitter users and conducted natural language processing (NLP) techniques for the classification [6]. More recently, Yazdavar et al. (2020), applied deep learning models to the detection of PTSD and suicidal ideation on Reddit and Twitter [7]. In other words, sentiment analysis on social media data can yield advanced warning signs of problems around mental health in these studies, providing an opportunity for preemptive care.

Nonetheless, challenges remain including data privacy, generalizability, platform bias, and ethical considerations regarding the use of sensitive information. Even though Twitter and Reddit prevail in the research context, platforms such as Instagram have not been mapped extensively even though they are frequently utilized by younger segments ofthe population—a cohort that is especially sensitive to mental health problems. Instagram's multimodal data represents a rich yet under-explored source of information due to the highly visual nature of the platform, where emotionally loaded captions and hashtags complement images of food, clothes, places, reminiscence, and basically any information one could share with the world. This review aims to bridge such gaps by highlighting methodological advancements considering different data types, and by calling for further research to be conducted across less studied data and platforms.

Numerous studies have explored the relationship between social media activity and mental health. Some have focused on linguistic markers and sentiment patterns, while others have incorporated multimodal approaches, including images and network analysis. The studies discussed in this section demonstrate the potential of mental health research, particularly via studies creating analyses from Instagram, Twitter, Reddit and Facebook data. Methodologies have progressed from conventional lexicon-based approaches to deep learning and transformer-based models, resulting in exponentially increased prediction performance [9]. However, some challenges still need to be addressed, including data imbalance, model interpretability, ethical issues, and particularly the lack of more multilingual and multimodal support [8]. In particular, Instagram, despite being one of the most popular platforms, especially among young people has attracted relatively little interest in the area of mental health prediction [7]. Since Instagram places a premium on visual content and aesthetic-oriented self-expression, the platform is a rich of underutilized source of data for detecting emotional and psychological signals. Consequently, our review not only connects the methodological development but also highlights this important research gap that we aim to fill in the subsequent sections by exploring multimodal techniques and cross-platform studies. Therefore, the following section will present an analysis of the computational approaches adopted to predict mental health based on its frameworks, together with their relative advantages and disadvantages.

3. Methodologies for Mental Health Prediction

The recent field of mental prediction based on social media has undergone a significant evolution over the years, from the initial use of machine learning technologies to the application of deep learning and transformer-based architectures. We systematically review these methods, comparing their evolution and utility in improving predictive performance.

A. Traditional Machine Learning Methods

Early prediction methods in the field of mental health were based on traditional machine learning (ML) models like Support Vector Machines (SVM), Random Forests, and Naïve Bayes classifiers. These approaches were based on feature engineering with linguistic, psychological, andbehavioral cues. For example, Tsakalidis et al. (2018) utilized SVMs to classify depression symptoms in twitter data using linguistic and behavioral features [16]. Similarly, Ernala et al. (2019) analyzed mood disorders with Random Forest models of Facebook activity logs [17]. While interpretable, these methods were constrained by their reliance on feature extraction, which does not always afford sufficient vocabulary and generalization capabilities.

B. Deep Learning-Based Methods

During the deep learning genesis, neural network architectures (e.g., Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN)) have become popular because they are capable of automatically learning feature representations, Orabi et al. (2018) used Long Short-Term Memory (LSTM) networks for depression detection on social media sites, achieving better prediction results than conventional classifiers [21]. Similarly, Tadesse et al. (2019) used CNNs and LSTMs to predict mental health disorders from Reddit conversations [18]. More recently, Rissola et al. for sentiment-based mental health scoring [18], Vascellaro et al. (2021) incorporated BiLSTMs with an attention mechanism for an improved context through sentiment-focused BiLSTMs [19]. Although these models made it easier to extract essential features, they could not address context, especially in the form of producing answersas a response to queries, and were not designed to handle multilingual data.

C. Transformer-Based Methods

The recent advances of transformer-based models, particularly BERT, XLM-RoBERTa and mBERT, have improved the ability to make predictions for mental health. Such language models use self-attention mechanisms to capture long-range dependency within text. Del Esposte et al. (2021) showed significant improvements in

---

applying BERT-based models compared to traditional ML and deep learning methodologies to detect depression from Reddit posts [21]. Similarly, Cohan et al. [20] Yang et al. (2021) utilized RoBERTa to monitor psychological distress within social media discussions and reported state-of-the-art performance. For multilingual data, some researchers like Sawhney et al. have experimented with XLM-R and mBERT. (2022), focused on challenges with cross-lingual mental health detection [22]. More recently, Zhang et al. (2023) introduced an approach that integrates BERT and GNN to boost mental health inference from multimodal evidence in social media [23].

The table below provides a comprehensive summary of recent research papers categorized by their methodological approach, including performance metrics and identified research gaps.

Table I. Summary of Methodological Approaches in Mental Health Prediction Using Social Media Data

Year	Study	Approach	Key Findings	Performance Metrics	Identified Gaps
2015	Resnik et al.	Lexicon-based Sentiment Analysis	Showed potential for mental healthmonitoring	Accuracy ~65%	Limited contextual understanding
2016	De Choudhury et al.	Topic Modeling	Linked social media discussions to mentalhealth signals	Precision ~0.70	Lacked temporal modeling and personalization
2017	Benton et al.	SVM, Feature Engineering	Improved early detection viahandcrafted features	F1-score ~0.68	Dependency on manual feature design
2018	Tsakalidis et al.	SVM, Linguistic Features	Behavioral and linguistic featureshelped detect depression	Accuracy ~72%	Poor performance on complex expressions
2018	Orabi et al.	LSTM	Enhanced depression detection using LSTM	F1-score ~0.75	Limited handling of multilingual content
2019	Ernala et al.	Random Forest	Facebook logs provided mood disorder insights	Accuracy ~70%	Dataset limited to a single platform
2019	Tadesse et al.	CNN, LSTM	DL models outperformed ML onReddit data	F1-score ~0.78	Temporal and multimodal aspects not incorporated
2020	Coppersmith et al.	BERT-based Classification	Improved diagnostic classification throughcontextual embedding	Accuracy ~82%	Lacked interpretability
2021	Rissola et al.	BiLSTM +Attention	Attention mechanisms improved contextualunderstanding	F1-score ~0.80	Resource intensive
2021	Del Esposte et al.	BERT	Achieved high accuracy in depression detection	Accuracy ~85%	Limited cross-lingual support
2021	Cohan et al.	RoBERTa	State-of-the-art resultsin psychological distress detection	F1-score ~0.83	Need for domain-specific tuning
2022	Sawhney et al.	XLM-R, mBERT	Improved cross-lingual prediction performance	Accuracy ~86%	Underutilization of multimodal data

Year	Study	Approach	Key Findings	Performance Metrics	Identified Gaps
2023	Zhang et al.	BERT + GNNFusion	Multimodal integration boosted predictionaccuracy	F1-score ~0.87	Complexity in model architecture
2024	Patel et al.	Multimodal Learning	Fusion of text and images enhancedemotion prediction	Accuracy ~88%	Limited real-time application
2025	Lee et al.	Self-Supervised Learning	Utilized unlabeled data to extract meaningfulmental health signals	F1-score ~0.89	Explainability and ethical concerns remain

The shift to deep learning and transformer-based models has highlighted the importance of advanced feature extraction and context-aware systems. Early models involved laborious feature engineering and often failed to capture fine-grained signals in user-generated content. But transformers, which provide contextual embeddings and attention, can be trained to make more context-sensitive predictions.

Moreover, advancements in multilingual modeling with models like XLM-R and mBERT have allowed non-English speakers to be better represented; which is particularly important for global mental health monitoring. In fact, recent works started to leverage multimodal learning, using text, images, and even user behavior graphs, to maximize general predictive performance.

Thus, another important need is to release standard datasets and benchmarks to enable fair methodology comparisons. The receive section examines these datasets and their contributions to reproducible and equitable research on mental health prediction.

4. Datasets and Benchmarking

Datasets are critical for the development and assessment of mental health prediction models. Researchers have utilized a diverse set of datasets, from hand-annotated corpora to freely accessed social media content. They are usually categorized according to the platform used (Twitter, Reddit, Facebook), the mental health affliction (depression, anxiety, PTSD), and the type of annotation (self-disclosure, clinical).

Twitter datasets were extensively used early on, benefitting from easy access and their public nature. Such datasets include some of the CLPsych 2015 and 2016 datasets, that label posts for mental health conditions from self-reported user disclosures [25]. Reddit has also become a go-to source, specifically subreddits like r/depression and r/anxiety where users often share their personal struggles with mental health. For instance, spaces such as the Dreaddit dataset [26] and Self-report Mental Health Diagnoses (SMHD) [27] corpora have been popular.

The quality and complexity of these datasets are heterogeneous. Although it can be small but clinically verified, and some datasets are large but noisy relying on user-generated tags or indirect labels. Due to the ethical and practical consequences of annotating data for mental health, researchers have investigated weak supervision and distant labeling methods.

To overcome these drawbacks, recent works have used pretrained models trained on massive corpora. Models like BERT, RoBERTa, XLM-RoBERTa, and mBERT are pre-trained on large general-purpose datasets and fine-tuned on the task-specific mental health data. This transfer learning paradigm allows the lean of few annotated images while showing excellent performance. Recent datasets such as GoEmotions [28], DAIC-WOZ [29], and Distress Analysis Interview Corpus offer rich, structured benchmarks to supplement pretrained models for a range of emotion and distress recognition tasks.

Benchmarking continues to be essential for assessing models on diverse tasks and domains. These metrics includeaccuracy, precision, recall, F1-score and AUC. Still, these feature vectors fail to offer a direct comparison due to differences in data preprocessing, class imbalance, and quality of annotations. Standardized benchmarking has been facilitated in recent efforts such as the eRisk shared tasks and the CLPsych workshops by releasing shared datasets and evaluation protocols.

While vast amounts of data are available from Twitter and Reddit, platforms such as Instagram lag due to difficulties in data access and annotation. Future work can study multimodal signals (e.g., image + caption) from rich visual platforms.

The following part describes the challenges and ethical implications of using social media data for mental health prediction, including privacy, bias, and potential misuse.

5. Challenges and Ethical Considerations

Predicting mental health from social media data shows great promise, but also comes with significant challenges and ethical implications. For the sake of specificity, these range from challenges around access and quality of data to challenges associated to the social context of the data such as privacy, algorithmic bias or misuse of predictive models.

A. Privacy, Consent, and Ethical Data Uset

Although social media offers a rich data source for mental health analysis, users seldom give permission for their post to be used for research purposes, creating ethical issues—particularly when it concerns sensitive mental health information. Anonymization is very challenging, especially regarding multimodal content, like images and metadata 33. There is also the risk of misuse of the data by third parties such as employers or insurers which might lead to surveillance or stigmatization.

B. Annotation Validity and Algorithmic Bias

Models often utilize self-reported diagnoses or platform activity (e.g., subreddit membership) that may not meet clinical standards, with implications for model validity and generalizability [35]. In addition, machine learning models may retain and enhance biases associated with age, gender, or cultural context—particularly when pretrained transformers such as BERT or mBERT are used.

C. Platform Limitations and Reproducibility Issues

Most studies focus on text-heavy platforms like Twitter and Reddit, leaving out visually rich platforms like Instagram or TikTok due to data access restrictions and multimodal complexity [29]. Additionally, reproducibility remains a challenge as many studies use proprietary datasets and do not share code or models, limiting transparency and progress in the field [30].

6. Applications and Real-World Case Studies

With the increasing sophistication of mental health prediction models, their integration into real-world applications such as early diagnosis tools and therapeutic support systems has been enabled.IOS strip. In this section, we will discuss essential areas where these technologies have been adopted, along with case studies in the real world, showcasing their opportunities and challenges.

A. chatbots AI-Powered Mental Health Support and Early Intervention

AI-based chatbots such as Woebot and Wysa employ natural language processing and CBT frameworks to provide therapeutic interaction and have shown significant efficacy in clinical trials. On a larger level, social media monitoring is leveraged by organizations such as Crisis Text Line and Facebook to catch suicide risk and distress through machine learning models — though these raise significant ethical discussions.

B. Integration into Clinical, Educational, and Workplace Settings

Healthcare organizations like Mount Sinai Hospital have started combining social media–based prediction models with electronic health record (EHR) electronic systems for the screening of mood disorders [31]. Other research [32] used social media analytics to monitor population mental health in times of crisis, such as COVID-

19. At the same time, schools and companies are adopting predictive technology to identify signals of burnout or anxiety from digital communications, but using digital communications raises privacy issues.

C. Cross-Lingual and Multimodal Applications

Recent advancements highlight the global applicability and growing sophistication of mental health prediction tools, particularly in cross-lingual and multimodal settings. These approaches are especially important for expanding access to mental health technologies across diverse populations and linguistic groups. For instance, Sawhney et al. (2022) employed cross-lingual transformer models such as XLM-R to detect depressive symptoms from multilingual Twitter data, demonstrating that mental health signals can be effectively captured across languages without relying solely on English corpora. This is a significant step toward bridging the gap in mental health support for low-resource language communities.

Additionally, researchers have explored multimodal analysis on platforms like Instagram, which offer both visual and textual content. In these studies, not only were captions and hashtags analyzed, but also visual features such as color saturation, brightness, and facial expressions within images, which can subtly signal emotional or psychological states. For example, Reece and Danforth (2017) showed that markers such as lower brightness, less saturation, and more faces in Instagram photos were correlated with depressive symptoms, offering compelling evidence for the utility of image-based features in mental health detection.

These findings emphasize the potential of combining multiple data modalities and supporting diverse languages, both of which are essential for creating inclusive, culturally aware, and globally scalable mental health monitoring systems.

Looking ahead, Part VII and VIII turns to broader conclusions and future directions in research and development, which is needed to ensure that these technologies remain safe, inclusive, and impactful.

7. Future Directions

The field of digital mental health prediction is still nascent but growing quickly with the introduction of advanced machine learning, deep learning, and multimodal AI systems. With the rising availability of data pertaining to data sets that are spread across a variety of digital platforms, there does arise significant scope to develop models that are more accurate, but also inclusive, culturally aware and context-aware. But the development of these models raises issues, especially dealing with ethics, fairness, and interpretability. It requires researchers to traverse a minefield of data privacy issues, algorithmic bias and the clinical relevance of AI outputs. Consequently, upcoming applied research and development needs to be directed toward systems that are robust, generalizable, and user-centered in design and can be embedded in real-world mental health settings. These pathways prioritize cross-cutting work across disciplines and the promotion of methodological transparency, and will necessarily involve proactive support strategies that are preventive and personalized rather than reliant on diagnosis, to enable transformative mental health action.

A. Enhancing Model Capability Through Multimodal, Multilingual, and Personalized Approaches

Future work is taking in the direction of multimodal models which can analyze text, images, audio and video from Instagram ready platforms like TikTok and YouTube. Another priority is multilingual prediction, especially in low-resource languages, using cross-lingual transformers and few-shot learning. An additional trend is the need for personalized, context-aware models, which adapt to user-specific baselines and events, protecting privacy using methods such as federated learning.

B. Improving Real-World Robustness, Ethics, and Explainability

Mental health is not an objective and static measure. Future studies should investigate more personalized models incorporating either individual-level baseline or between subjects' variability in their linguistic features. Adding context-sensitive signals, such as event timelines, life changes or geolocation (with user permission), can enhance accuracy and minimize false positives. Methods such as federated learning could also enable personalization without compromising privacy.

C. Toward Scalable, Preventive, and Interdisciplinary Solutions

Mental health prediction is by nature interdisciplinary, demanding communication between AI experts, clinicians, ethicists, and public health professionals. Scale will depend on partners in health care and technology. The long-range plan is to transition from reactive care to preventive assistance, embedding predictions in wearables, wellness apps, and digital phenotyping systems that can aid early interventions before symptoms appear.

8. Conclusion

This survey explored the evolution of mental health prediction using social media data, tracing the shift from traditional machine learning approaches to state-of-the-art transformer-based models. A key takeaway is the growing emphasis on context-aware, multilingual, and multimodal techniques that enhance predictive accuracy. However, the foundation of these advancements lies in the availability of high-quality, diverse datasets.

Our review highlights the critical need for well-curated, multilingual datasets and standardized benchmarks to drive meaningful progress. Despite the wealth of user-generated content on platforms like Instagram, it remains underutilized in mental health research. The lack of large-scale, annotated datasets across multiple languages and cultural contexts presents a significant barrier to model generalization and fairness. Addressing this gap will require concerted efforts in data collection, annotation, and the ethical handling of sensitive information.

Furthermore, while technical advancements continue, ethical concerns—such as user privacy, algorithmic bias, and responsible AI deployment—must be addressed through interdisciplinary collaboration. The future of mental health prediction depends on a synergy between machine learning researchers, mental health professionals, and policymakers to develop robust, ethical, and impactful solutions. By prioritizing dataset diversity and ethical AI practices, this field can move closer to realizing its potential for early intervention and accessible mental health support.

References

[1] Smith et al., "Machine Learning for Mental Health Prediction: A Review," Journal of AI & Health, 2022. 1
[2] Resnik et al. (2015). Lexicon-based Approaches for Mental Health Monitoring. 3
[3] De Choudhury et al. (2016). Topic Modeling for Mental Health Insights. 4
[4] Benton et al. (2017). Feature-Based SVM for Depression Detection. 5
[5] Tsakalidis et al. (2018). Depression Detection via SVMs and Behavioral Features. 6
[6] Orabi et al. (2018). Deep Learning for Depression Detection. 7
[7] Ernala et al. (2019). Random Forest Analysis of Mood Disorders from Social Media. 8
[8] Tadesse et al. (2019). CNN and LSTM Models for Mental Health Prediction. 9
[9] Del Esposte et al. (2021). BERT for Depression Detection on Reddit.
[10] Cohan et al. (2021). RoBERTa-based Psychological Distress Analysis. 13
[11] Sawhney et al. (2022). Cross-Lingual Mental Health Prediction Using Transformers.14
[12] Zhang et al. (2023). BERT-GNN Fusion for Multimodal Mental Health Prediction. 15
[13] Patel et al. (2024). Multimodal Learning for Emotion Prediction. 16
[14] Lee et al. (2025). Self-Supervised Learning for Social Media Mental Health Insights. 17
[15] Resnik et al. (2015). Lexicon-based Approaches for Mental Health Monitoring.18
[16] De Choudhury et al. (2016). Topic Modeling for Mental Health Insights. 19
[17] Benton et al. (2017). Feature-Based SVM for Depression Detection. 20
[18] Orabi et al. (2018). Deep Learning for Depression Detection. 22
[19] Ernala et al. (2019). Random Forest Analysis of Mood Disorders from Social Media.23
[20] Tadesse et al. (2019). CNN and LSTM Models for Mental Health Prediction. 24
[21] Coppersmith et al. (2020). BERT-based Diagnostic Classifier. 25
[22] Rissola et al. (2021). Attention-Based BiLSTMs for Mental Health Analysis. 26
[23] Del Esposte et al. (2021). BERT for Depression Detection on Reddit.27
[24] Learning for Social Media Mental Health Insights.32

[25] Coppersmith, G., Leary, R., Whyne, E., & Wood, T. (2015). Quantifying mental health signals in Twitter. Proceedings of the Workshop on Computational Linguistics and Clinical Psychology.33
[26] Turcan, E., & McKeown, K. (2019). Dreaddit: A Reddit dataset for stress analysis in social media. Proceedings of the Tenth Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis.34
[27] Cohan, A., Desmet, B., Yates, A., Soldaini, L., MacAvaney, S., &Goharian, N. (2018). SMHD: A large-scale resource for exploring online language usage for multiple mental health conditions. Proceedings of the 27th International Conference on Computational Linguistics (COLING).35
[28] Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., &Jurafsky, D. (2020). GoEmotions: A dataset of fine-grained emotions. Proceedings of ACL 2020. 37
[29] Gratch, J., Artstein, R., Lucas, G. M., Stratou, G., Scherer, S., &Morency, L. P. (2014). The Distress Analysis Interview Corpus of human and computer interviews. LREC 2014. 38
[30] Chancellor, S., & De Choudhury, M. (2020). Methods in predictive techniques for mental health status on social media: a critical review. NPJ Digital Medicine, 3(1), 1-11. 40
[31] Gillespie, T. (2018). Facebook's AI suicide prevention: Who watches the watchers? Big Data & Society. 51
[32] Birnbaum, M. L., et al. (2017). Using machine learning on social media to identify individuals with psychotic disorders. Schizophrenia Research. 52
[33] Morris, R. R. (2015). Social networking site interventions in mental health: The promise and challenge of peer support. Journal of Medical Internet Research. 55
[34] Sawhney, R., et al. (2022). Multilingual mental health detection using XLM-R. Proceedings of the 60th ACL.