University of Mohaghegh Ardabili 1978
RTBS Journal of A Review of Theorizing of Behavioral Sciences
License number from the Ministry of Culture and Islamic Guidance

Research Paper

Artificial Intelligence for Early Detection of Anxiety and Depression: A Rapid Review and Theoretical Perspective on Speech, Text, and Online Behavioral Indicators

Farzaneh Abbasi[1*], Mohammad Reza Abedi[2], Zahra Yousefi[3]

1. Payam-e Noor University, Isfahan (Mobarakeh Center) Branch, Isfahan, Iran.
Corresponding author: farzaneh-abbasi@outlook.com
2. Department of Counselling, Faculty of Educational Sciences and Psychology, University of Isfahan, Isfahan, Iran.
3. Department of Clinical Psychology, Islamic Azad University, Isfahan (Khorasgan) Branch, Isfahan, Iran.

Received: 2025/01/24
Revised: 2025/04/19
Accepted: 2025/06/20

Use your device to scan and read the article online

Abstract
Anxiety and depression are among the most prevalent mental health disorders worldwide, and early detection remains a clinical priority. This rapid review synthesizes evidence on the application of artificial intelligence (AI) for detecting anxiety and depression using speech, text, and online behavioral data. A systematic search was conducted in Scopus, PubMed, IEEE Xplore, ACM Digital Library, Web of Science, and Google Scholar for English-language studies published between 2013 and 2025. Studies applying machine learning, deep learning, natural language processing (NLP), or speech analysis for mental health detection were included. Fifteen studies (including empirical, review, and model-based studies) met the inclusion criteria. AI models have been shown to identify depression-related indicators, such as negative language, self-referential expressions, reduced social interaction, and acoustic speech changes. Anxiety-related indicators included worry-related language, emotional instability, and irregular behavioral patterns. Multimodal and transformer-based models demonstrated improved contextual understanding. AI-based systems show potential for early detection of anxiety and depression; however, they should not be considered diagnostic tools due to limitations in data quality, bias, and clinical validation.

Keywords:
Artificial Intelligence, Depression, Anxiety, Natural Language Processing, Digital Mental Health

Citation: Abbasi, F., Abedi, M., & Yousefi, Z. (2025). Artificial Intelligence for Early Detection of Anxiety and Depression: A Rapid Review and Theoretical Perspective on Speech, Text, and Online Behavioral Indicators. *A Review of Theorizing of Behavioral Sciences, 2*(4), 1-14. doi: 10.22098/j9032.2026.20011.1064

Introduction

Mental disorders such as anxiety and depression represent a major global health burden and are among the most prevalent psychiatric conditions worldwide (*WHO, 2023*). These disorders are associated with impaired emotional functioning, reduced quality of life, occupational and academic difficulties, and increased healthcare demands. Early identification is therefore essential because delayed diagnosis may contribute to symptom progression, functional impairment, and more complex treatment needs.

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR), depressive disorders are characterized by persistent low mood, loss of interest or pleasure, cognitive and somatic symptoms, and impaired daily functioning (*APA, 2013*). Anxiety disorders involve excessive fear, worry, physiological arousal, and avoidance behaviors that interfere with normal activities. Because many early symptoms may appear subtly in language, communication patterns, and behavioral routines, digital and speech-based indicators may provide useful complementary information for early psychological screening.

Recent advances in artificial intelligence (AI) have created new opportunities for the early identification of mental health conditions through computational analysis of speech, text, and online behavior, including machine learning (ML), deep learning, natural language processing (NLP), and large language models (LLMs), have enabled computational identification of psychological patterns associated with anxiety and depression (*Owen et al., 2024*; *Yang et al., 2024*). These technologies can process large volumes of unstructured data and detect subtle linguistic, emotional, and behavioral signals that may not be easily observable in conventional clinical settings.

Social media platforms, voice recordings, and other digital interaction environments have become important sources of behavioral and psychological data for mental health research. Recent studies have highlighted the growing use of NLP and speech-based AI systems for identifying depression- and anxiety-related indicators from online communication and vocal characteristics (*Maran et al., 2025*; *Scherbakov et al., 2025*). In addition, emerging research suggests that transformer-based models and multimodal AI systems may improve contextual interpretation of emotional and psychological states across different data modalities (*Sharma et al., 2025*; *Yang et al., 2024*).

Despite the growing number of studies on AI-based mental health detection, existing evidence remains fragmented across different modalities, including text, speech, and online behavioral data. Many studies focus on a single disorder or a single data source, limiting cross-modal comparison and reducing the practical interpretability of findings for mental health applications (*Le Glaz et al., 2021*; *Owen et al., 2024*). Therefore, the aim of the present rapid review was to synthesize recent evidence on the application of AI in identifying anxiety- and depression-related indicators through speech, text, and online behavioral data.

Method

- Study Design

This study is a rapid review conducted in accordance with the PRISMA 2020 reporting guidelines to ensure transparent reporting of study identification, screening, eligibility assessment, and inclusion processes. The PRISMA framework was adapted for rapid evidence synthesis, with the recognition that full systematic review protocol requirements were not applied.

Given the rapidly evolving nature of AI applications in mental health, particularly in natural language processing (NLP), speech analysis, and multimodal machine learning, a rapid review approach was considered appropriate to enable timely synthesis of the most recent evidence while maintaining methodological transparency.

- Search Strategy

A structured literature search was performed in Scopus, PubMed, IEEE Xplore, ACM Digital Library, Web of Science, and Google Scholar. The search covered English-language peer-reviewed journal articles and full conference papers published between 2013 and 2025.

Search queries were constructed using combinations of controlled keywords and free-text terms related to artificial intelligence and mental health. The following core search strings were applied and adapted across databases:

- “Artificial intelligence” AND depression detection
- “Machine learning” AND anxiety
- “Natural language processing” AND mental health
- “Deep learning” AND depression
- “Speech analysis” AND depression
- “Online behavior” AND anxiety
- “Social media” AND mental health
- “Large language model” AND mental health

Boolean operators (AND, OR) were used to combine and refine search terms across databases. Reference lists of relevant studies were also screened to identify additional eligible publications.

- Eligibility Criteria

Studies were included if they met all of the following criteria: (1) published in English between 2013 and 2025; (2) investigated anxiety, depression, or psychological distress; (3) applied AI-based methods such as machine learning, deep learning, natural language processing, or speech analysis; (4) utilized speech, text, social media, or online behavioral data; and (5) reported empirical findings related to mental health indicators.

Studies were excluded if they were non-peer-reviewed publications (e.g., editorials, letters, opinion pieces, or abstracts without full text), did not employ AI-based methods, focused solely on general social media behavior without mental health analysis, lacked sufficient methodological detail, or were unrelated to anxiety, depression, or psychological distress detection.

The inclusion of empirical, review, and model-based studies was intended to capture the breadth of rapidly evolving AI methodologies in mental health research and to enable a comprehensive mapping of emerging evidence across heterogeneous study designs.

- Study Selection

A total of 84 records were identified through database searching. After removing 18 duplicates, 66 records were screened based on title and abstract.

During screening, 38 records were excluded due to irrelevance to AI-based mental health detection or failure to meet inclusion criteria.

The remaining 28 full-text articles were assessed for eligibility. Of these, 13 studies were excluded due to lack of relevance to AI-based mental health detection or insufficient methodological reporting.

Finally, 15 studies were included in the qualitative synthesis.

The study selection process followed the PRISMA 2020 four-stage framework: Identification, Screening, Eligibility, and Included studies. Some studies contributed to multiple thematic categories; therefore, studies were not mutually exclusive across depression and anxiety sections. The study selection process is summarized in Figure 1.

- Data Extraction and Synthesis

Data extraction was conducted using a structured data-charting form. Extracted variables included authors, publication year, target mental health condition, data modality, AI methodology, and key findings.

The included studies were analyzed using thematic narrative synthesis. Findings were organized into two primary categories:

(1) depression-related indicators and

(2) anxiety-related indicators.

Within each category, studies were further grouped according to data modality, including text-based analysis, speech-based features, and online behavioral patterns.

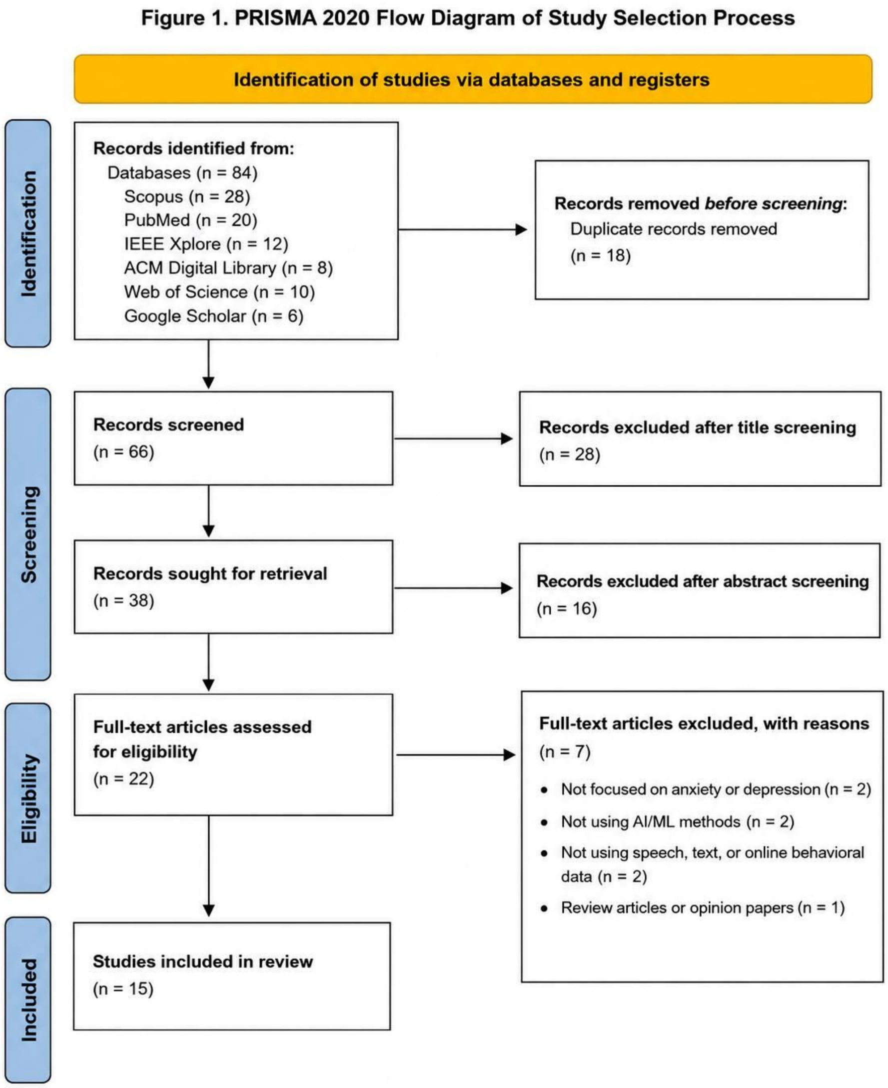

Figure 1. PRISMA 2020 flow diagram illustrating the study selection process

- Quality Assessment

Methodological quality of the included studies was assessed using a structured critical appraisal approach adapted from the Joanna Briggs Institute (JBI) principles. Given the heterogeneity of study designs, including machine learning experiments, natural language processing studies, speech-based analyses, and review articles, a uniform quantitative risk-of-bias tool was not considered appropriate. Instead, a domain-based scoring framework was applied to ensure consistency and transparency.

Each study was evaluated across five domains:

(1) clarity of research objectives,

(2) appropriateness of AI methodology,

(3) transparency of data sources,

(4) reproducibility of analytical procedures, and

(5) relevance and validity of reported mental health outcomes.

Each domain was rated as 0 (not addressed), 1 (partially addressed), or 2 (fully addressed), resulting in a total quality score ranging from 0 to 10.

Based on total scores, studies were categorized as:

- High quality (8–10)
- Moderate quality (5–7)
- Low quality (0–4)

Two independent reviewers were intended for assessment; however, due to the rapid review design, appraisal was conducted by a single reviewer with cross-checking of ambiguous cases through consensus discussion.

No studies were excluded solely on the basis of methodological quality, in line with the exploratory nature of rapid evidence synthesis. Instead, quality ratings were used to inform interpretation of findings and to contextualize the strength of evidence across studies.

Results

- Characteristics of Included Studies

The final synthesis included studies using various artificial intelligence approaches, including machine learning, deep learning, natural language processing (NLP), transformer-based models, and speech analysis systems. The reviewed studies analyzed multiple forms of digital data, including social media text, online behavioral activity, speech recordings, and multimodal communication patterns.

The findings were organized into two primary categories:

- AI-based studies on depression-related indicators
- AI-based studies on anxiety-related indicators
- Depression-Related Indicators Identified by AI Systems

The reviewed depression-related studies suggest that AI-based systems can identify several recurring linguistic, behavioral, and speech-related indicators associated with depressive symptoms. Text-based studies commonly reported negative emotional language, hopelessness-related expressions, increased self-referential language, and reduced linguistic diversity as important depression-related markers (*Eichstaedt et al., 2018*; *Tadesse et al., 2019*). Online behavioral analyses additionally identified reduced social interaction, altered engagement patterns, and visual behavioral indicators associated with depression (*Reece & Danforth, 2017*).

Speech-based studies demonstrated that depressive symptoms may also be reflected through acoustic and vocal characteristics, including reduced vocal energy, slower speech rate, monotonic tone, and increased pauses (*Gratch et al., 2014*; *Maran et al.,*

10.22098/J9032.2026.20011.1064

2025). More recent studies further suggest that transformer-based models and large language models may improve contextual understanding of psychological language patterns and enhance interpretability of depression-related signals (*Yang et al., 2024*).

Overall, the findings indicate that AI systems can detect indirect behavioral and linguistic markers associated with depressive symptoms across multiple digital modalities.

Table 1. Summary of AI-based studies on depression-related indicators (empirical and model/review studies)

Article	Authors/Year	Data Source	AI Method	Main Findings
Facebook language predicts depression in medical records	*Eichstaedt et al. (2018)*	Facebook posts and medical records	NLP and predictive modeling	Linguistic markers (negative affect, self-reference) were associated with depression
Instagram photos reveal predictive markers of depression	*Reece and Danforth (2017)*	Instagram images and behavioral data	Machine learning, image analysis	Visual and behavioral patterns were predictive of depressive symptoms
Detection of depression-related posts in Reddit forum	*Tadesse et al. (2019)*	Reddit posts	Deep learning, NLP	Depression-related posts exhibited distinct linguistic/emotional patterns
Automatic speech analysis for depression detection	*Gratch et al. (2014)*	Clinical interview speech	Speech signal processing	Acoustic features reflected depressive states (reduced energy, slower speech)
Speech-based depression detection: systematic review	*Maran et al. (2025)*	Speech datasets	Meta-analysis / speech AI	Consistent association between acoustic features and depression
AI for mental health analysis using social media	*Owen et al. (2024)*	Social media datasets	NLP / AI synthesis	AI systems effectively extract depression-related signals from online text
MentaLLaMA: interpretable mental health LLM	*Yang et al. (2024)*	Mental health text datasets	Large language models	Improved contextual understanding of depressive language patterns

- Anxiety-Related Indicators Identified by AI Systems

The reviewed anxiety-related studies suggest that AI-based systems have been reported to identify several linguistic, behavioral, and speech-related indicators associated with anxiety across digital environments. Text-based studies commonly reported that anxiety is reflected through worry-related expressions, fear and uncertainty language, emotionally charged wording, and repetitive concern-oriented communication patterns (*Benton et al., 2017*; *Coppersmith et al., 2014*). These linguistic features were consistently detected using NLP and machine learning approaches applied to social media datasets.

Behavioral studies further indicate that anxiety may be associated with irregular online activity patterns, fluctuating engagement levels, and increased emotional reactivity in digital interactions (*De Choudhury et al., 2013*). These findings suggest that anxiety is not only reflected in language use but also in temporal and behavioral dynamics of online communication.

Review-based studies additionally highlight that NLP methods are effective in identifying anxiety-related semantic and emotional cues in large-scale datasets, although anxiety detection remains methodologically challenging due to its overlap with general psychological distress and other mental health conditions (*Guntuku et al., 2017*; *Le Glaz et al., 2021*).

More recent studies emphasize the growing role of advanced AI architectures, including transformer-based models and large language models (LLMs), in improving contextual understanding of anxiety-related language patterns and enhancing detection performance (*Owen et al., 2024*; *Yang et al., 2024*). These models are particularly effective in capturing subtle emotional and contextual variations that are difficult to detect using traditional NLP approaches. Transformer-based NLP models, such as BERT (*Devlin et al., 2019*), have improved performance in mental health–related text classification tasks by capturing contextual semantic relationships and subtle emotional cues (*Yang et al., 2024*).

Speech-based studies also demonstrate that anxiety-related states may be reflected in acoustic features such as increased hesitation, altered speech rhythm, and vocal tension (*Gratch et al., 2014*). Speech-based studies suggest that acoustic and prosodic features have been explored as potential indicators of psychiatric conditions, including anxiety and depression, though findings remain inconsistent across datasets and methodologies. These findings suggest that multimodal AI systems combining text, behavioral, and speech data may provide more comprehensive and reliable detection of anxiety-related indicators.

Overall, the reviewed studies suggest that AI systems can identify indirect linguistic, behavioral, and speech-related markers associated with anxiety across multiple digital modalities. However, compared with depression-related evidence, anxiety-focused findings remain more heterogeneous due to variability in datasets, methodologies, and the lack of consistent disorder-specific definitions.

Table 2. Summary of AI-based studies on anxiety-related indicators (empirical, review, and methodological AI studies)

Article	Authors/Year	Data Source	AI Method	Main Findings
Quantifying mental health signals in Twitter	*Coppersmith et al. (2014)*	Twitter posts	NLP, machine learning	Linguistic patterns associated with anxiety and psychological distress identified
Multitask learning for mental health classification	*Benton et al. (2017)*	Social media text	Multitask deep learning	Simultaneous detection of anxiety and related mental health conditions
Predicting postpartum emotional changes	*De Choudhury et al. (2013)*	Social media activity	Behavioral analytics	Anxiety-related behavioral changes detected in online activity

Detecting mental illness in social media	*Guntuku et al. (2017)*	Social media datasets	NLP review synthesis	Anxiety associated with stress and emotional instability patterns
NLP and social determinants of mental health	*Scherbakov et al. (2025)*	Psychosocial datasets	NLP analysis	Anxiety-related linguistic and social indicators identified
AI in adolescent mental health	*Sharma et al. (2025)*	Adolescent datasets	Machine learning	Emotional distress and anxiety-related behavioral signals detected
NLP in mental health applications	*Le Glaz et al. (2021)*	Text datasets	NLP systematic review	Anxiety-related semantic and emotional patterns identified
Speech-based psychiatric biomarkers	*Low et al. (2020)*	Speech datasets	Acoustic modeling	Anxiety associated with prosodic and speech variability
Distress analysis interview corpus	*Gratch et al. (2014)*	Clinical interviews	Speech analysis	Vocal features associated with anxiety and emotional distress
Mental Lama (LLM-based analysis)	*Yang et al. (2024)*	Mental health text datasets	LLMs	Improved contextual detection of anxiety-related language
BERT contextual language model	*Devlin et al. (2019)*	NLP benchmarks	Transformer model	Contextual embeddings improved downstream classification performance

Discussion

- Interpretation of Depression-Related Findings

The reviewed depression-related studies suggest that depression is associated with observable behavioral and linguistic changes that can be captured in digital communication data. Across text-based studies, depressive states were commonly associated with negative emotional vocabulary, increased self-referential language, hopelessness-related expressions, and reduced linguistic diversity (*Eichstaedt et al., 2018*; *Tadesse et al., 2019*). These findings are consistent with previous research indicating that depression may alter both emotional expression and communication behavior in digital environments.

Natural language processing (NLP) and transformer-based models appear particularly useful in this context because they can identify semantic and contextual relationships beyond simple keyword frequency. Recent large language model (LLM)-based systems have demonstrated improved capability in interpreting complex psychological language patterns and contextual emotional signals (*Yang et al., 2024*). Such models may improve the sensitivity of AI-based mental health screening by identifying subtle linguistic variations associated with depressive symptoms.

Online behavioral studies further suggest that depression may be reflected through reduced social interaction, altered engagement patterns, and changes in visual or communication behavior (*Reece & Danforth, 2017*). Similarly, speech-based studies indicate that depressive symptoms may influence acoustic and vocal characteristics, including reduced vocal energy, slower speech rate, and increased pauses (*Gratch et al., 2014*; *Maran et al., 2025*). These speech-related features may reflect psychomotor slowing and reduced emotional expressiveness, which are commonly associated with depressive states.

Overall, the reviewed evidence suggests that AI-based approaches can detect indirect linguistic, behavioral, and speech-related markers associated with depression across multiple digital modalities. However, these systems primarily identify probabilistic behavioral patterns rather than clinically definitive diagnostic features.

- Interpretation of Anxiety-Related Findings

The reviewed anxiety-related studies indicate that AI systems can identify anxiety-related psychological patterns through language, behavioral activity, and digital interaction signals. Anxiety-related indicators were commonly associated with expressions of worry, uncertainty, fear, heightened emotional tension, and unstable behavioral engagement patterns (*Benton et al., 2017*; *Coppersmith et al., 2014*).

NLP-based systems may be particularly effective in detecting anxiety-related signals because anxiety is often reflected linguistically through repeated concern, future-oriented threat language, emotional instability, and uncertainty-related expressions. Recent studies have emphasized the growing role of AI-driven language analysis in identifying emotional distress and anxiety-related psychological markers in online communication environments (*Owen et al., 2024*; *Scherbakov et al., 2025*).

Behavioral analyses also suggest that anxiety may influence patterns of online activity, including fluctuating engagement, irregular interaction frequency, and heightened emotional responsiveness in digital environments. AI-based systems can process these behavioral signals at large scale and identify recurring patterns that may not be easily observable through traditional assessment approaches.

However, compared with depression-related research, anxiety detection appears less consistent across studies and datasets. Several reviewed studies examined anxiety together with broader psychological distress or multiple mental health conditions rather than as an isolated disorder. This heterogeneity may reduce the comparability of findings and complicate interpretation of anxiety-specific indicators.

Recent transformer-based and speech-analysis studies further suggest that multimodal AI systems may improve detection sensitivity for anxiety-related psychological states by integrating contextual linguistic and acoustic features.

- Why AI-Based Systems Can Detect Mental Health Patterns

AI-based approaches are increasingly used in mental health research because they can process large volumes of complex and unstructured digital data, particularly in large-scale digital phenotyping and social media-based mental health inference. Machine learning and deep learning systems can identify recurring behavioral and linguistic patterns across datasets, while NLP models can analyze semantic, contextual, and emotional features of language (*Le Glaz et al., 2021*).

Speech-based AI systems can additionally extract acoustic features that may not be easily recognized through conventional observation alone. Similarly, transformer-based architectures and large language models may improve contextual interpretation by capturing long-range semantic relationships and nuanced emotional expressions in text (*Yang et al., 2024*). These capabilities explain why AI systems have demonstrated growing potential as supportive tools for early mental health screening. Nevertheless, most reviewed studies relied on indirect behavioral proxies rather than clinically verified psychiatric assessment. Consequently, AI-generated outputs should be interpreted cautiously and should not be considered equivalent to formal clinical diagnosis.

- Limitations

This rapid review has several limitations. First, only English-language studies between January 2013 and March 2025 were included, which may have introduced language and publication bias. Second, the reviewed studies used heterogeneous datasets, AI methodologies, outcome measures, and definitions of psychological distress, limiting direct comparison across studies.

Third, many studies relied on publicly available social media or online behavioral datasets rather than clinically validated psychiatric populations. As a result, some identified indicators may not fully correspond to clinically diagnosed anxiety or depression. Fourth, although a structured narrative quality appraisal was conducted, no formal meta-analysis was performed; therefore, the findings should be interpreted qualitatively rather than as pooled estimates of diagnostic performance.

Finally, cultural, linguistic, and demographic differences across datasets may affect the generalizability of AI-based mental health indicators across populations and contexts.

The increasing use of AI in mental health research also raises important ethical concerns related to privacy, informed consent, algorithmic bias, and responsible use of sensitive psychological data. Since many AI systems rely on social media and digital behavioral data, ensuring confidentiality and transparent data governance remains essential for ethical implementation in mental health contexts.

The included studies comprised heterogeneous methodological designs, including empirical machine learning studies, review articles, and large language model–based investigations. Therefore, the synthesis is narrative rather than statistically aggregated.

The findings of this review have several implications for behavioral science research. Across the reviewed studies, artificial intelligence systems identified patterns in language use, speech characteristics, and online behavioral activities that were associated with anxiety- and depression-related indicators. These findings suggest that observable digital behaviors may provide measurable representations of psychological processes that have traditionally been examined through self-report measures and clinical assessment.

The reviewed evidence also highlights the growing role of computational approaches in connecting behavioral data with psychological constructs. Natural language processing, speech analysis, and machine learning methods enable the identification of recurring patterns in communication and behavior that may reflect emotional distress, negative affect, worry-related cognition, and changes in social engagement. In this way, AI-based approaches may contribute to the operationalization and large-scale observation of behavioral indicators relevant to mental health research.

In addition, the findings support the emerging concept of digital phenotyping, which refers to the use of data generated through digital interactions to characterize behavioral and psychological functioning. The reviewed studies demonstrate how speech, text, and online behavioral data can be analyzed to identify indicators associated with anxiety and depression. From a behavioral science perspective, these developments provide new opportunities to study psychological phenomena in naturalistic environments while complementing traditional research and assessment methods. However, the interpretation of such indicators should remain grounded in established psychological theory and supported by appropriate clinical validation.

Conclusion

This rapid review suggests that artificial intelligence can identify recurring linguistic, behavioral, and speech-related indicators associated with anxiety and depression across heterogeneous digital data sources, including text, speech, and online behavioral signals. The findings indicate that AI-based systems may support early psychological screening by capturing subtle patterns in communication and behavior that are not easily detectable through traditional assessment methods.

The findings of this review have important implications for counselors, psychologists, and other mental health professionals. AI-assisted approaches may serve as supportive tools for the early identification of individuals at risk of anxiety and depression and may complement conventional assessment and monitoring practices in digital and community-based mental health settings. However, current evidence remains largely exploratory, and AI-based systems should be considered supportive technologies rather than standalone diagnostic tools. Their implementation requires rigorous clinical validation, transparency, and careful attention to ethical issues, including privacy, informed consent, and algorithmic bias.

Future research should prioritize clinically validated and multimodal datasets, interpretable AI architectures, longitudinal investigations, and cross-cultural validation studies to improve the reliability, fairness, and practical applicability of AI-based mental health detection systems.

Conflict of Interest Statement

The authors declare no conflict of interest.

Funding Statement

This research received no external funding.

References

APA. (2013). Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc, 21*(21), 591-643. doi:https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787

Benton, A., Mitchell, M., & Hovy, D. (2017). Multi-task learning for mental health using social media text. *arXiv preprint arXiv:1712.03538*. doi:https://doi.org/10.48550/arXiv.1712.03538

Coppersmith, G., Dredze, M., & Harman, C. (2014). *Quantifying mental health signals in Twitter.* Paper presented at the Proceedings of the workshop on computational linguistics and clinical psychology: From linguistic signal to clinical reality. https://aclanthology.org/W14-3207.pdf

De Choudhury, M., Counts, S., & Horvitz, E. (2013). *Predicting postpartum changes in emotion and behavior via social media.* Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems. https://doi.org/10.1145/2470654.2466447

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *Bert: Pre-training of deep bidirectional transformers for language understanding.* Paper presented at the Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers). https://doi.org/10.18653/v1/N19-1423

Eichstaedt, J. C., Smith, R. J., Merchant, R. M., Ungar, L. H., Crutchley, P., Preoţiuc-Pietro, D., . . . Schwartz, H. A. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences, 115*(44), 11203-11208. doi:https://doi.org/10.1073/pnas.1802331115

Gratch, J., Artstein, R., Lucas, G. M., Stratou, G., Scherer, S., Nazarian, A., . . . Marsella, S. (2014). *The distress analysis interview corpus of human and computer interviews.* Paper presented at the Lrec. https://d1wqtxts1xzle7.cloudfront.net/98764174

Guntuku, S. C., Yaden, D. B., Kern, M. L., Ungar, L. H., & Eichstaedt, J. C. (2017). Detecting depression and mental illness on social media: an integrative review. *Current Opinion in Behavioral Sciences, 18*, 43-49. doi:https://doi.org/10.1016/j.cobeha.2017.07.005

Le Glaz, A., Haralambous, Y., Kim-Dufor, D.-H., Lenca, P., Billot, R., Ryan, T. C., . . . Berrouiguet, S. (2021). Machine learning and natural language processing in mental health: systematic review. *Journal of medical Internet research, 23*(5), e15708. doi:https://doi.org/10.2196/15708

Low, D. M., Bentley, K. H., & Ghosh, S. S. (2020). Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope investigative otolaryngology, 5*(1), 96-116. doi:https://doi.org/10.1002/lio2.354

Maran, P. L., Braquehais, M. D., Vlaic, A., Alonzo-Castillo, M. T., Vendrell-Serres, J., Ramos-Quiroga, J. A., & Rodríguez-Urrutia, A. (2025). Performance of automatic speech analysis in detecting depression: systematic review and meta-analysis. *JMIR Mental Health, 12*(1), e67802. doi:https://doi.org/10.2196/67802

Owen, D., Lynham, A. J., Smart, S. E., Pardiñas, A. F., & Camacho Collados, J. (2024). Ai for analyzing mental health disorders among social media users: Quarter-century narrative review of progress and challenges. *Journal of medical Internet research, 26*, e59225. doi:https://doi.org/10.2196/59225

Reece, A. G., & Danforth, C. M. (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science, 6*(1), 15. doi:https://doi.org/10.1140/epjds/s13688-017-0110-z

Scherbakov, D. A., Hubig, N. C., Lenert, L. A., Alekseyenko, A. V., & Obeid, J. S. (2025). Natural language processing and social determinants of health in mental health research: AI-assisted scoping review. *JMIR Mental Health, 12*(1), e67192. doi:https://doi.org/10.2196/67192

Sharma, G., Yaffe, M. J., Ghadiri, P., Gandhi, R., Pinkham, L., Gore, G., & Abbasgholizadeh-Rahimi, S. (2025). Use of artificial intelligence in adolescents' mental health care: systematic scoping review of current applications and future directions. *JMIR Mental Health, 12*, e70438. doi:https://doi.org/10.2196/70438

Tadesse, M. M., Lin, H., Xu, B., & Yang, L. (2019). Detection of depression-related posts in reddit social media forum. *Ieee Access, 7*, 44883-44893. doi:https://doi.org/10.1109/ACCESS.2019.2909180

WHO. (2023). *Depressive disorder (depression)*. Retrieved from https://www.who.int/news-room/fact-sheets/detail/depression

Yang, K., Zhang, T., Kuang, Z., Xie, Q., Huang, J., & Ananiadou, S. (2024). *MentaLLaMA: interpretable mental health analysis on social media with large language models.* Paper presented at the Proceedings of the ACM Web Conference 2024. https://doi.org/10.1145/3589334.3648137