ISSN: 3049-8074

Review Article

Journal of Clinical Medicine & Health Care

AI-Based Screening of Adolescent Mental Health Conditions Using Mobile Health Applications: A Systematic Review

Uwemedimo Sunday Isaiah[1*], Ekaette Mfon Useh[2], Iniobong George[1], Chilaka Chika Franklin[3], Chineye Fabian Adili-George[4] and Mfon Effiong Ineme[1]

[1]*Department of Psychology, University of Uyo, Akwa Ibom State, Nigeria*

[2]*Department of Educational Foundations, University of Calabar, Nigeria*

[3]*Robert Wood Johnson University Hospital, Hamilton, USA*

[4]*Department of Health Services Administration, University of Alabama at Birmingham, United States*

[*]Corresponding author
Uwemedimo Sunday Isaiah, Department of Psychology, University of Uyo, Akwa Ibom State, Nigeria.

Received: April 30, 2026; **Accepted:** May 08, 2026; **Published:** May 18, 2026

ABSTRACT

Purpose: The escalating global burden of adolescent mental health disorders necessitates innovative approaches to early detection and screening. This systematic literature review critically examines peer-reviewed studies on artificial intelligence (AI) applications in mobile health (mHealth) platforms for adolescent mental health screening to address the primary research question: “What AI technologies, data sources, and outcomes are prevalent in current AI-based mHealth screening tools for adolescents, and which mental health conditions are primarily targeted?”

Methods: This systematic review analyzed 45 studies published between 2020 and 2024 from major academic databases including PubMed, IEEE Xplore, ScienceDirect, Springer, ACM Digital Library, and PsycINFO. Inclusion/exclusion criteria and a structured protocol with systematic data extraction were employed to identify AI methodologies, data sources, targeted mental health conditions, and reported outcomes. The review identified common applications of supervised machine learning, deep learning (CNNs, RNNs, LSTMs, transformer models), and multimodal approaches specifically designed for or including adolescent populations.

Results: Key findings indicate that transformer-based models (BERT, RoBERTa, MentalBERT) achieved F1-scores between 0.85 and 0.97 for depression detection from social media and text data. Deep learning and multimodal approaches demonstrated high diagnostic accuracy (75-92%), with LLM-based chatbots showing feasibility in delivering cognitive behavioral therapy to adolescents. The most frequently targeted conditions were depression (68%) and anxiety (24%), with data sources primarily comprising smartphone passive sensing, social media posts, clinical interviews, and wearable sensor data. Passive sensing data including GPS mobility patterns, accelerometer data, and sleep metrics emerged as particularly valuable for unobtrusive monitoring.

Conclusions: Despite promising results, significant limitations persist including data privacy concerns, demographic bias, limited longitudinal validation, and the need for standardized datasets. This review highlights these challenges and proposes directions for enhancing AI’s effective integration into adolescent mental health screening through mHealth platforms, emphasizing the importance of ethical frameworks, clinical validation, and age-appropriate design considerations.

Keywords: Artificial Intelligence, Adolescent Mental Health, Mobile Health, Digital Phenotyping, Machine Learning, Depression, Anxiety, Chatbots, Systematic Review

Introduction

Mental health disorders among adolescents represent one of the most pressing public health challenges of the 21st century, with global prevalence rates indicating that approximately 10-20% of adolescents experience mental health conditions, yet the majority remain undiagnosed and untreated [1]. The unique developmental characteristics of adolescence including increased autonomy, heightened sensitivity to social evaluation, and widespread technology adoption create both challenges and opportunities for mental health screening and intervention [2].

The proliferation of smartphones and wearable devices among adolescents has generated unprecedented opportunities for continuous, unobtrusive mental health monitoring through

Citation: Uwemedimo Sunday Isaiah, Ekaette Mfon Useh, Iniobong George, Chilaka Chika Franklin, Chineye Fabian Adili-George, et al. AI-Based Screening of Adolescent Mental Health Conditions Using Mobile Health Applications: A Systematic Review. J Clin Med Health Care. 2026. 3(2): 1-15.
DOI: doi.org/10.61440/JCMHC.2026.v3.57

mobile health (mHealth) applications [3]. Recent advances in artificial intelligence (AI) and machine learning (ML) have enabled the development of sophisticated screening tools that can analyze diverse data streams including natural language, behavioral patterns, physiological signals, and social media activity to detect early indicators of mental health conditions [4].

Recent research has investigated the prediction of mental health disorders using interpretable ML models integrated into smartphone applications, chatbot-based early screening and intervention platforms for youth, and bias elimination in physiological signal-based mental health prediction [5-7]. These developments illustrate the increasing maturity and diversity of AI-based adolescent mental health care. As the global adolescent mental health crisis intensifies and research interest in AI-based screening grows, the need for a comprehensive systematic review becomes imperative.

This systematic literature review (SLR) seeks to provide structure and synthesis to this emerging field, offering insights for policymakers, clinicians, developers, and researchers regarding the current landscape and future directions of AI-based adolescent mental health screening through mHealth platforms. Building on prior work in digital mental health, this review offers a distinct focus on adolescent populations and mobile-based applications [8-10].

Several recent works have specifically examined AI applications in youth mental health. For instance, Grist et al. provided a comprehensive review of mental health mobile apps for preadolescents and adolescents, highlighting engagement and efficacy considerations [11]. More recently, systematic reviews have examined AI chatbots for alleviating mental distress among adolescents and young adults, and predictive AI approaches in mHealth platforms for youth mental health forecasting [12, 13]. These contributions are complementary to our focused review, which specifically examines AI-based screening methodologies within mHealth applications targeting adolescent mental health.

The review adheres to a systematic protocol for study inclusion, screening, and quality assessment to ensure thorough and valid findings. It aims to serve as a comprehensive resource for researchers, clinicians, developers, and policymakers interested in the intersection of adolescent mental health and AI technology.

Objectives

The objectives of this SLR are as follows:

- To classify and examine the types of AI technologies currently employed in mHealth applications for adolescent mental health screening.
- To examine the data sources typically utilized in mobile platforms and their impact on model development and validation.
- To consolidate the reported impact and clinical outcomes of AI-based screening interventions on adolescent mental health.
- To identify the most common mental health conditions (e.g., depression, anxiety, eating disorders) that AI mHealth solutions target in adolescent populations.
- To investigate data preprocessing, privacy preservation, and adaptation methods used to address the unique requirements of adolescent mental health data.
- To provide a systematic foundation for future research and development by mapping gaps, trends, and opportunities in the field.

Research Methodology

A systematic approach, as described by Kitchenham, was employed to conduct this literature review [14]. This section focuses on designing the review procedure systematically while identifying relevant studies and categorizing them to address all research questions.

Review Protocol

The review protocol comprises four distinct activities: (1) developing research questions, (2) establishing exclusion and inclusion criteria, (3) defining the search process, and (4) data extraction and synthesis. Research questions were developed in Section 1, while the remaining activities are explained below.

Inclusion and Exclusion Criteria

A set of inclusion and exclusion criteria was applied during the screening process to ensure relevance, quality, and consistency:

- Only articles published in 2020 or later were included to capture recent developments in AI and mHealth technologies.
- Papers were limited to those found in selected databases: PubMed, IEEE Xplore, ScienceDirect, Springer, ACM Digital Library, and PsycINFO.
- Studies must demonstrate a clear connection between AI/ML and adolescent mental health (ages 10-24); general population studies without adolescent-specific analysis were excluded unless sub-group analyses were provided.
- Only English-language articles were accepted due to translation constraints.
- Peer-reviewed journal articles and high-quality conference proceedings were included; low-quality or non-academic sources were excluded.
- Studies must specifically involve mHealth applications, smartphone-based interventions, or wearable devices as the platform for AI implementation.
- Non-accessible full-text studies were excluded.
- Duplicate publications were detected using Mendeley and removed.
- Papers dealing with theoretical AI development without practical application to adolescent mental health contexts were excluded.

Rationale: The cutoff year of 2020 was chosen to ensure the review reflects the modern era of transformer-based models and advanced mobile sensing capabilities. The selected databases balance medical (PubMed, PsycINFO), engineering (IEEE, ACM), and multidisciplinary sources (Springer, ScienceDirect), providing comprehensive yet focused coverage.

Search Process

A comprehensive search was conducted across selected databases using combinations of keywords and Boolean operators:

Search Terms

- (“artificial intelligence” OR “machine learning” OR “deep learning”) AND (“adolescent” OR “teenager” OR “youth”) AND (“mental health” OR “depression” OR “anxiety”)

AND ("mobile" OR "smartphone" OR "mHealth" OR "wearable")

- ("chatbot" OR "conversational agent") AND ("adolescent" OR "young adult") AND ("mental health" OR "CBT")
- ("digital phenotyping" OR "passive sensing") AND ("adolescent" OR "youth") AND ("depression" OR "anxiety")
- ("transformer" OR "BERT" OR "LLM") AND ("social media" OR "mobile") AND ("adolescent mental health")
- The search was limited to English-language publications. From the initial search, 8,742 studies were retrieved across all databases. After removing duplicates, 6,891 unique records remained for screening.

Screening Process

- **Title screening:** 5,423 studies were excluded as irrelevant to research questions.
- **Abstract screening:** 1,023 studies were excluded for not meeting inclusion criteria.
- **Full-text review:** 445 studies were evaluated, with 400 excluded for insufficient adolescent focus, non-mobile platforms, or lack of AI methodology.
- **Final inclusion:** 45 studies were selected for detailed analysis.
- The selection process is illustrated in Figure 1, demonstrating systematic reduction from initial records to final studies.

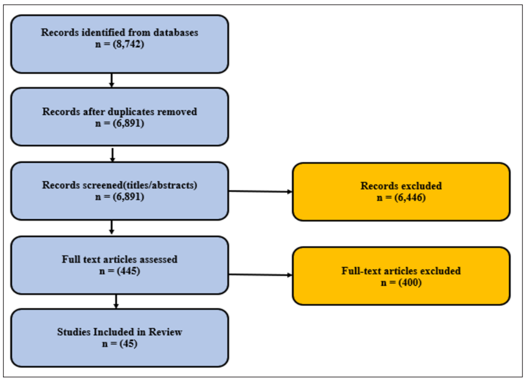

Figure 1: Prisma Flow Diagram of Study Selection Process for AI-Based Adolescent Mental Health Screening Using mHealth Applications

Data Extraction and Synthesis

A systematic data extraction strategy was adopted to facilitate comprehensive review. A sequential filtering process was performed on each study:

- Title evaluation for potential alignment with review scope
- Abstract assessment for relevance to research questions
- Full-text review of methodology, results, and discussion sections

Data extracted from each study included:

- Study objectives, aims, and contributions
- AI methods, frameworks, models, or algorithms employed
- Dataset description, origin, size, and availability
- Primary findings, performance metrics, and clinical outcomes
- Adolescent-specific considerations and ethical safeguards
- This multi-step extraction process ensured inclusion of high-quality, relevant studies. Synthesized data were categorized and compared to identify trends, gaps, and opportunities.

Results and Discussion

This section organizes extracted studies into logical categories: (1) traditional machine learning approaches, (2) deep learning and transformer-based methods, (3) multimodal and passive sensing approaches, (4) LLM and chatbot-based interventions, and (5) ethical and implementation considerations specific to adolescent populations.

Study Categorization

Traditional Machine Learning for Adolescent Mental Health Screening

Based on the predetermined search terms and quality evaluation, a number of high-quality studies utilizing conventional machine learning techniques were identified (see Table 1). These studies apply supervised learning methods such as SVM, Decision Trees, Random Forests, Logistic Regression, and ensemble methods to forecast mental health symptoms such as depression, anxiety, and stress in adolescent populations [15-23].

Generally, the majority of studies attained accuracy between 75% and 89%, with SVM and ensemble methods tending to perform best. Some identify feature selection methods and real-world applicability in low-resource environments as advantages. SVM and Random Forest models consistently provided robust predictive capability (e.g., up to 89% accuracy), with the use of ensemble methods and hyperparameter optimization subsequently enhancing results. These results prove that conventional machine learning techniques, if properly calibrated and tested, can facilitate early mental health detection and tailored interventions in adolescent populations [15,16,21].

Table 1: Summary of Traditional Machine Learning Approaches in Adolescent mHealth Applications

Paper	Technique	Pros	Cons	Results
[15]	Random Forest, SVM	Model-agnostic interpretability via SHAP; sleep and social app usage patterns	Moderate accuracy; tuning required	Accuracy: 81%; AUC: 0.84
[16]	Logistic Regression, Gradient Boosting	Combined features outperformed single modalities; EMA integration	High complexity; resource-intensive	AUC: 0.79-0.83
[17]	Ensemble (Voting Classifier)	Mobility patterns distinguished depressed vs. non-depressed; GPS + accelerometer	Small sample (n=156)	F1: 0.76
[18]	SVM (RBF kernel)	Targets adolescent social media data; linguistic features	Limited to text features; English only	Accuracy: 78%

[19]	Decision Tree, Logistic Regression, SVM, KNN, RF	Simple algorithms; clear feature importance; mobile demo	Limited to survey features; risk of overfitting	Best (DT & LR): 84% acc
[20]	Random Forest, SVM	Very large survey (n=447,000); stable AUC; predicts self-rated health	Predicts self-rated health only; no clinical validation	AUC = 0.80–0.81
[21]	Hierarchical LR + MC dropout	Fully interpretable; confidence estimates; clinical trust	Moderate balanced accuracy	Bal. acc.: 0.79; AUC above 0.90
[22]	SVM, RF, XGBoost, NN, LR	High sensitivity; multi-model study; adolescent-specific features	Low specificity (0.30); imbalanced data	AUC: 0.755; Sens.: 0.91; Spec.: 0.30
[23]	Decision Tree (J48 on PHQ-9)	Reduces questionnaire length; interpretable for		

Deep Learning and Transformer-Based Approaches

Deep learning approaches are on a rise when it comes to classification; here this classification task is applied to the adolescent mental health screening where a healthy or mentally ill adolescent is identified using the deep learning approaches. Figure 2 shows the number of the papers based on each of the approaches used. These papers used a variety of transformer-based and Natural Language Processing (NLP) techniques, including TF-IDF, n-grams, BERT, RoBERTa, and ALBERT [24]. There have been instances like where integrated models like CNN-LSTM are used for this purpose [25]. To detect depression automatically from smartphone sensor data, paper uses an LSTM model that integrates convolutional layers, attention, and bi-directional LSTMs for temporal pattern recognition in adolescent device usage [26].

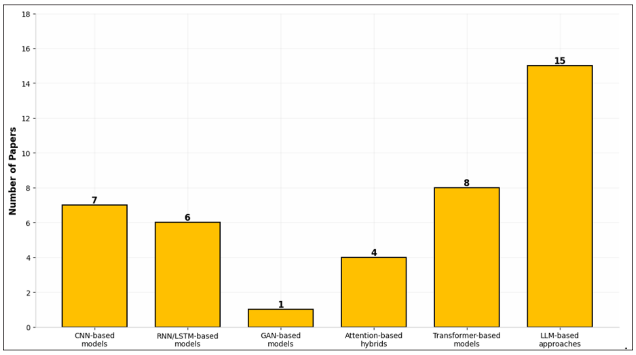

Figure 2: Distribution of Deep Learning Approaches Across Included Studies. The Figure Shows Relative Adoption of CNNs, RNNs, and Transformer-Based Models. LLM Dominate Recent Research

Multimodal and Passive Sensing Approaches

Recent studies have revealed that the combination of several modalities including text, audio, images, facial characteristics, and physiological measures has the potential to enhance mental health detection over single-modality models in adolescent populations [27-34]. Table 2 lists several prominent multimodal studies.

Table 2: Summary of Deep Learning Approaches

Paper	Modalities Used	Methodology	Performance Metrics
[27]	Text, Audio, Video	Shared and separate encoders, transformer fusion, multitask learning for depression and sentiment analysis	CCC: 0.466, MAE: 5.21
[28]	Text, Images	BERT for text, image captioning + BERT for images, multimodal fusion	Accuracy: 93.8%, F1-score: 96.8%
[29]	Text, Images, Posting Time	BERT (text), InceptionResNetV2 (images), time-aware LSTM, attention mechanism	F1-score: 95.6% (Instagram), 90.8% (Twitter)
[30]	Text, Facial features (Video)	GPT-3.5-Turbo and DepRoBERTa for text, Bi-LSTM for facial features, multimodal fusion	MAE: 3.76, RMSE: 4.53
[31]	Text, Audio, Facial video	Feature extraction (BERT, COAVREP, Facial AUs), autoencoders, multi-head attention, knowledge transfer	Accuracy: 81.10%, Precision: 80.20%, Recall: 81.00%, F1-Score: 80.60%

[32]	Text, Emotions, Audio (synthetic)	Multi-teacher knowledge distillation from BERT, SenticNet, Bark (audio)	Best Macro F1: 93.10% (IdenDep), others vary
[33]	EEG, Audio, Facial expressions	Prompting GPT-4o on combinations of EEG, audio, and facial features	Up to 79% accuracy
[34]	Questionnaire Responses, Video (facial expressions)	3D CNN for video, self-attention module, combined with questionnaire scores and answering times	Accuracy: 91.58%, Sensitivity: 90.82%, Specificity: 92.35%

Overall, the models combine various types of features with methods such as transformer-based fusion, LSTMs, attention, and knowledge distillation. Performance varies across studies but remains robust, with some having F1-scores above 90% and mean absolute errors less than 4.0. The findings illustrate the trade-off between performance and resource demands, which limits adoption in resource-constrained healthcare environments, particularly in school-based settings serving diverse adolescent populations [27, 31].

LLMs and Chatbot-Based Adolescent Mental Health Screening

This segment outlines recent research that employs large language models (LLMs) and chatbots for detecting mental health and providing support to adolescents (refer to Table 3). Models such as fine-tuned ChatGPT-3.5/4, Claude, and LLaMA variants analyze diary text, voice, and questionnaires to detect depression and anxiety. In general, these models have high accuracy (as high as 90.2%), correlate well with experts' judgments, and are promising as safe, scalable early intervention tools for youth, even if a higher degree of refinement and human supervision is still necessary for clinical application [35-46].

Table 3: Summary of Multimodal Approaches

Paper	Type	Approach	Performance/Outcome
[35]	LLM + Speech features	ChatGPT analyzes text and speech rhythm for depression and anxiety detection	Accuracy up to 79.14%
[36]	LLM	GPT-3.5/4 analyzes emotional diaries; fine-tuning and CoT prompting	Accuracy: 90.2% (GPT-3.5)
[37]	Chatbot	Woebot chatbot delivers CBT; randomized controlled trial on college students [6]	Significant reduction in depression (effect size d=0.44)
[38]	Chatbot	CAFIbot (GPT-4) with Critical Analysis Filter to ensure safety in schizophrenia support	81% compliant safe responses (vs 8.3% without CAF)
[39]	LLM	Comparison of ChatGPT-3.5, ChatGPT-4, Claude, Bard with mental health professionals	ChatGPT-4, Bard closely matched professional assessments
[40]	LLM	Alpaca, GPT-3.5, GPT-4, instruction-finetuned for mental health tasks on Reddit data	Fine-tuned models outperformed GPT-3.5 by 10.9%
[41]	Chatbot	Fine-tuned GPT-4 chatbot for Indonesian university students' mental health	Approval rate: 87.38%
[42]	LLM	GPT-4 multicultural competence tested on South Asian contexts	Highlighted cultural gaps and need for improvement
[43]	Chatbot	ChatGLM-6B-based chatbot for depression diagnosis and suicide risk with interactive conversation	Binary suicide risk F1-score: 0.98
[44]	LLM	Psycho Analyst model (GPT-4 customized) for depression classification and PHQ-8 scoring	F1-Score: 0.929, Accuracy: 95.7%
[45]	LLM	MentaLLaMA (fine-tuned LLaMA2) for interpretable mental health analysis with explanations	Comparable accuracy to discriminative models
[46]	LLM	LLMind Chat (Gemma 2 model) for diagnosis using ICD-11 clinical data	Clinical case correctness: 76.1%

Digital Phenotyping and Passive Sensing Approaches

This section describes digital phenotyping and passive sensing studies to identify mental health risk from smartphone data, mobile activity, and sensor information in adolescent populations (see Table 4). These approaches leverage GPS mobility patterns, accelerometer data, sleep metrics, and social app usage to detect depression, anxiety, and suicidal ideation without requiring active user engagement [47-52]. In general, deep learning and transformer-based models perform most effectively when combined with passive features (e.g., reaching a 77% balanced accuracy for suicidal ideation prediction), but less complex methods provide better interpretability for clinical adoption. Methods such as fine-grained phone usage analysis and GPS trajectory modeling also enhance detection accuracy and interpretability for early mental health intervention in youth [47,48,50].

Table 4: Summary of Digital Phenotyping and Passive Sensing Approaches for Adolescent Mental Health Detection

Paper	Area	Approach	Performance
[47]	Passive Sensing	GPS + accelerometer + phone use; Random Forest + feature engineering	AUC: 0.77; school attendance patterns identified
[48]	Passive Sensing	GPS, Bluetooth, screen state; Feature engineering + LR	r=0.62 with PHQ-9; depression severity tracking
[49]	Passive Sensing + EMA	Accelerometer, GPS, EMA; LSTM networks	Significant correlation with clinical measures; anhedonia detection
[50]	Passive Sensing	GPS + Bluetooth + screen time; Contrastive learning	Balanced accuracy: 0.77; social proximity features
[51]	Smartphone Sensing	App usage patterns, sleep metrics; Ensemble methods	Accuracy: 81%; sleep and social app usage predictive
[52]	Wearable + Smartphone	HRV, sleep, activity; Deep ensemble	Accuracy: 79%; circadian rhythm analysis

Comparative Analysis of Categories

While the above section presents detailed evidence on traditional ML, deep learning, multimodal, LLM/chatbot, and passive sensing approaches, it is equally important to synthesize their relative strengths and limitations. To provide a concise overview, Table 5 summarizes these categories across three dimensions most relevant to real-world deployment: practical applicability, scalability, and clinical adoption. This complements the results in earlier tables by highlighting how different methodological families compare in terms of feasibility and integration into adolescent mental health care.

Table 5: Comparison of AI approaches for Adolescent Mental Health Applications across Applicability, Scalability, and Clinical Adoption

Approach	Practical Applicability	Scalability	Clinical Adoption
Traditional Machine Learning (ML)	Low computational cost; suitable for resource-limited and real-time systems [15]. Works well with structured data (e.g., questionnaires, basic sensor features). [19,23]	Limited scalability for large-scale or unstructured data; struggles with high dimensionality. [48, 49]	High interpretability supports clinician and parent trust [21, 23]; complements psychometric tools. Limited by association-based modeling and feature dependence. [20]
Deep Learning (DL)	Excels at analyzing unstructured data (text, speech, physiology) [26]. Transformer models and LLMs handle complex adolescent language patterns [35]. Requires large datasets and specialized hardware (GPUs).	Highly scalable for big data (e.g., social media streams). Cloud deployment feasible. [24] Generalizable through fine-tuning. [25]	Opacity ("black-box" issue [37]) hinders trust. XAI methods (e.g., SHAP, attention) improve interpretability [21]. Ethical concerns (bias, hallucinations, privacy) limit deployment. [39,40]
Multimodal Approaches	Holistic assessment by integrating text, audio, video, and physiology. Robust to noise via fusion strategies [31, 30].	Data acquisition bottlenecks and integration complexity reduce scalability. Computationally intensive for large-scale fusion [30].	Improves diagnostic accuracy and objectivity [35]. Interpretability and privacy concerns amplified due to multi-source sensitive adolescent data. [37]
LLMs / Chatbots	Strong at contextual text reasoning; effective for diaries, interviews, and conversations. Natural interaction suited to adolescent preferences [40].	Highly scalable via APIs and cloud-based deployment. [40] Few-shot/fine-tuning enables rapid adaptation.	Promising in pilot studies and RCTs [37]. Risks include hallucinations, cultural bias, and over-trust by adolescent patients. [39] Human-in-the-loop supervision essential for safe use with minors.

Answers to Research Questions

Analyzing the works, it is clear that researchers have made significant advances in using artificial intelligence and machine learning in the domain of adolescent mental health prediction, detection, and management. Several studies are relying on depression and anxiety detection from multimodal sources of data like textual social media posts, audio and video signals, survey responses, and smartphone sensor data. Transformer-based models (e.g., BERT, RoBERTa, MentalBERT) and traditional supervised classifiers (e.g., SVM, Decision Trees, Random Forests) were heavily employed. Consistent utilization of metrics such as accuracy, precision, recall, and F1-score facilitated objective comparison of model performance. Some works also utilized interpretability tools like SHAP to promote model transparency in clinical settings [53-55].

The SLR has uncovered that the studies involved a diverse set of adolescent populations and data sources-ranging from Reddit user posts, Twitter text data, smartphone sensor streams, to wearable physiological data. Some of the research incorporated cutting-edge techniques such as multitask learning, federated learning, and contrastive learning to improve performance and maintain privacy. Use cases ranged from chatbot-based interventions, decision support systems, to mobile apps for monitoring adolescent mental health. A few models attained extremely good predictive performance (i.e., F1-scores of more than 0.95), indicating the possibility of applying AI-powered tools in actual real-time adolescent mental health care contexts.

RQ1: AI Technologies in Adolescent Mental Health

The purpose of this research question was to find out what AI technologies have been used in the domain of adolescent mental health. This field makes widespread use of a range of supervised machine learning classifiers including SVM, Decision Trees, Random Forests, and ensembling methods. Deep learning approaches such as CNNs, RNNs, LSTMs, and transformers (e.g., BERT family) are widely used for text, image, and sensor data processing. Multimodal models incorporating audio, visual, text, and physiological information are also gaining momentum. Explainability (such as SHAP, saliency maps) and bias avoidance measures (such as multitask learning with epistemic uncertainty) have been introduced to adhere to clinical practices. Table 6 shows the different technologies used in the selected papers; many of the papers in the study have compared performances of varying models. The majority of studies make use of Python-based machine learning platforms like Scikit-learn, TensorFlow, PyTorch, and Hugging Face Transformers. Data are derived from actual sources such as Reddit, Twitter, smartphone sensors, and wearable devices. Systems are implemented in the form of mobile applications or web dashboards in many instances.

Many powerful AI models, such as deep neural networks, operate as black boxes due to their complexity, making it difficult to trace how specific inputs lead to predictions [37]. The inability to fully interpret these models is often referred to as the interpretability-performance trade-off, without interpretability it is difficult for users to trust or use the results [21,29]. Ultimately, interpretability allows the AI to function effectively as a confirmatory second opinion, enabling the clinician to explore the clinical factors that informed the prediction [21].

Table 6: Technologies Used in Adolescent Mental Health Prediction Studies

Technology Used	Papers
Machine Learning (Traditional ML models like SVM, Decision Trees, Logistic Regression, Random Forest, etc.)	[15-23]
Deep Learning (CNN, LSTM, RNN, MLP, DNN)	[24-34]
Natural Language Processing (NLP)	[24, 28-45]
Transformer-based Models (BERT, RoBERTa, ALBERT, GPT, etc.)	[24,28-45]
Large Language Models (LLMs: ChatGPT, GPT-3.5, GPT-4, Claude, Bard, etc.)	[35-46]
Chatbots	[37-46]
EEG-based Analysis	[33,47]
Multimodal Fusion (Text + Audio + Video/ Facial Expressions)	[27-34]
Speech Analysis	[31-36]
Facial Expression/Computer Vision Techniques	[28-34]
Social Media Analysis	[18,24-45]
Mobile App-based Systems	[15-17,37,46]
Passive Sensing / Digital Phenotyping	[15- 52]
Wearable Sensors	[33,49,52]
Federated Learning	[48]
Explainable AI (SHAP, Saliency Maps, etc.)	[15,21,37,45]
Contrastive Learning	[50]
Knowledge Transfer / Representation Learning	[31]
Self-attention Mechanism	[30-32]
Scale-based Modeling (PHQ-9, GAD-7, etc.)	[15,16,19,23, 36,44]
Clinical Diagnostic Tools / Manual Scoring	[39,44,46]

RQ2: Data Sources for Adolescent Mental Health AI

For the second research question it was required to find out different data sources for identification and model training of adolescent mental health AI systems which are shown in Table 7. It is seen from the research that different data sources are being used for this purpose. The table below classifies the papers on the basis of the datasets used by them for mental health prediction. These datasets differ in modality and source, ranging from survey responses to social media postings, speech recordings, physiological signals (e.g., EEG, HRV), facial expressions, smartphone sensor data, and multi-modal clinical interviews.

Recent studies further highlight the value of social media and multimodal passive sensing as data sources for adolescent populations. For instance, BERT-based approaches applied to social media text demonstrate that pre-trained models can significantly outperform traditional ML for depression detection in youth [56]. Likewise, systematic reviews of multimodal passive data (e.g., audio, video, smartphones, and wearables) emphasize the importance of integrating diverse modalities while addressing challenges of standardization and demographic variation in adolescent samples [57].

The Table 7 organizes papers according to dataset type and presents the corresponding citations as identified in the document, guaranteeing that the classification is exclusively dependent on information provided without relying on any assumptions outside the information.

To address the need for collecting, curating, and sharing standardized, demographically diverse datasets for adolescent mental health research, a multifaceted approach integrating various data sources, robust curation methods, and strong ethical safeguards is essential. The sources highlight several strategies and frameworks that can contribute to this goal.

Beyond identifying dataset types in Table 7, several strategies have been emphasized in the literature for creating standardized and demographically diverse resources. First, diverse and multimodal collection is key [27, 50] combining social media text [28], clinical interviews (e.g., DAIC-WOZ, PHQ-9) [30], audio-visual cues as speech, facial expressions, physiological and wearable signals (EEG, HRV, sleep, heart rate), and smartphone sensor data provides complementary perspectives on adolescent mental health. Second, robust curation and standardization are essential. This includes expert annotation using validated scales (e.g., PHQ, GAD-7) [36], careful preprocessing (text normalization, noise removal), and techniques such as data augmentation to mitigate imbalance [23]. Ethical sharing and governance frameworks are critical. Privacy-preserving methods like federated learning [48] and differential privacy, alongside informed consent/assent from minors and explainable AI [21] support responsible data use. Collaborative research and regulatory engagement (e.g., GDPR/COPPA compliance) [37] further ensure safe and equitable access. Together, these strategies provide a roadmap for building comprehensive, representative, and ethically sound datasets that can advance AI in adolescent mental health.

Integrating AI into adolescent mental health care introduces significant ethical risks and requires strict regulatory oversight [40]. Reliability, controllability, and trustworthiness are challenged by the collection of multimodal personal data from minors, raising privacy concerns [35]. Privacy-preserving techniques such as federated learning and differential privacy allow decentralized training without exposing individual adolescent data [48]. However, LLMs often generate inaccurate or authoritative "hallucinations," risking harmful advice (e.g., self-harm) and compromising safety [43]. Addressing their inability to reliably convey uncertainty remains crucial [58].

AI model performance and generalizability depend on the quality and objectivity of training data, the "ground truth" [52]. Many models rely on social media posts (e.g., Reddit, Twitter) [40] but are constrained to English-language data [24] and biased by self-disclosed, self-selected adolescent populations [36]. LLM corpora also embed societal biases from social media, news, and clinical publications, reinforcing stereotypes and misapplication [39,37]. Multimodal clinical datasets are often imbalanced, with few high-severity cases, limiting the model's ability to predict severe outcomes in adolescents [30]. Moreover, models trained in specific regions risk performance degradation from population shifts when applied elsewhere [34].

Table 7: Classification of Papers Based on Dataset Type Used

Dataset Type	Papers
Survey Data	[15-36,44]
Social Media Posts	[18,24,28-45]
EEG, ECG Signals	[33,47]
Facial Images / Videos	[28-34]
Multimodal Clinical Interview Data	[27-34,44]
Smartphone Passive Sensing (GPS, accelerometer, app usage)	[15-51]
Wearable Sensor Data	[33,49,52]
Online Forums, Personal Journals, Clinical Records	[24,35,36,45]
EMA (Ecological Momentary Assessment)	[16,49]

RQ3: Reported Effectiveness/Outcomes

This section provides comparative analysis among the various machine learning and AI methodologies employed by selected studies on prediction of adolescent mental health based strictly on reported outcomes in terms of accuracy and F1-score. Below is the Table 8 which shows the results of studies involving representative papers under different types of data modality and models such as the usual machine learning, deep models, transformer-based, and multimodal. By contrasting their performance, we can gain an understanding of which approaches have demonstrated the greatest reliability in identifying or forecasting mental illnesses such as depression, anxiety, and stress in various adolescent settings [59].

Beyond predictive accuracy, some studies have explored clinical adoption through randomized trials, pilots, and user studies. Woebot, a CBT-based conversational agent, showed feasibility and efficacy in a college RCT with young adults [37], while digital phenotyping approaches demonstrated reasonable generalizability across different schools and time periods [48]. Clinical validations include an LLM framework tested on adolescent diary entries [36] and passive sensing models predicting treatment outcomes [21]. Hybrid systems, such as AI diary analysis combined with psychiatrist review [36], further illustrate early translation efforts. Overall, while promising, real-world deployment remains limited in adolescent populations.

Table 8: Summary of Selected Papers with Reported Metrics

Paper	Model	Accuracy	Precision	Recall	F1-Score
[15]	Random Forest	0.81	–	–	–
[16]	Logistic Regression, Gradient Boosting	0.83	–	–	0.79
[17]	Ensemble (Voting)	0.76	–	–	0.76
[18]	SVM (RBF kernel)	0.78	–	–	–
[19]	Decision Tree, Logistic Regression	0.84	–	–	–
[20]	Random Forest, SVM	0.81	–	–	–
[21]	Hierarchical LR + MC dropout	0.79	–	–	0.79
[22]	SVM, RF, XGBoost, NN, LR	0.755	–	–	–

[23]	Decision Tree (J48)	0.75	–	–	–
[24]	RoBERTa	–	–	–	0.97
[28]	IEMFF (BERT + Image)	0.938	–	–	0.968
[29]	MTAN (time-aware LSTM)	–	–	–	0.956 / 0.908
[30]	GPT-3.5-Turbo + Bi-LSTM	0.85	–	–	–
[31]	RLKT-MDD	0.811	0.802	0.81	0.806
[32]	Multi-teacher knowledge distillation	–	–	–	0.931
[33]	GPT-4o prompting	0.79	–	–	–
[34]	3D CNN + self-attention	0.916	0.908	0.923	0.916
[35]	LLM + Speech (GPT)	0.79	–	–	–
[36]	GPT-3.5 Fine-tuned	0.902	–	1.0	–
[40]	Alpaca, GPT-3.5, GPT-4	0.874	–	–	–
[43]	ChatGLM-6B	–	–	–	0.98
[44]	Psycho Analyst (GPT-4)	0.957	–	–	0.929
[45]	MentaLLaMA (LLaMA2)	0.88	–	–	–
[46]	LLMind Chat (Gemma 2)	0.761	–	–	–
[47]	Passive Sensing + RF	0.77 (bal.)	0.76	0.78	0.77
[48]	GPS + Bluetooth + LR	0.77	–	–	–
[50]	Contrastive learning	0.77 (bal.)	–	–	–
[51]	Smartphone sensing ensemble	0.81	0.79	0.83	0.81

RQ4: Targeted Mental Health Conditions in Adolescents

Here, a graphical overview of the mental health disorders covered across the studies in this literature review is provided. The purpose is to visually and systematically classify the different mental health conditions that have been the main targets of prediction, diagnosis, or monitoring in the research corpus under analysis.

Other research examines a single mental disorder, such as depression or anxiety, and uses machine learning and AI techniques to predict, identify, or track symptoms specifically of that condition in adolescents. Other research examines multiple disorders at once, usually in the form of clusters like depression and anxiety together or more general constructs like stress and behavioral activation. In some instances, new topics like eating disorders and ADHD are also examined. This diversity reflects not only the wide-ranging applicability of AI techniques but also differing research priorities across the adolescent mental health field.

The visualization, presented as a pie chart (Fig. 3), classifies the studies by their most salient targeted mental health disorder. It provides an unambiguous, intuitive picture of how various mental health disorders are spread throughout the literature. This classification allows readers to immediately see the focus of research efforts for example, it is immediately clear that depression is the most researched disorder, representing a significant majority of the diagnosis-oriented papers.

Aside from simple categorization, the graphic summary helps identify areas of research deficit and developing trends in the subject area. As the graph indicates, although depression fills the ground, other disorders like anxiety, stress, and suicide risk find very little space. Much less represented are such topics as eating disorders and ADHD, although of essential value, these are found only sparingly in present literature. This discrepancy points to future research opportunities, specifically the necessity for greater diversification in AI research in adolescent mental health.

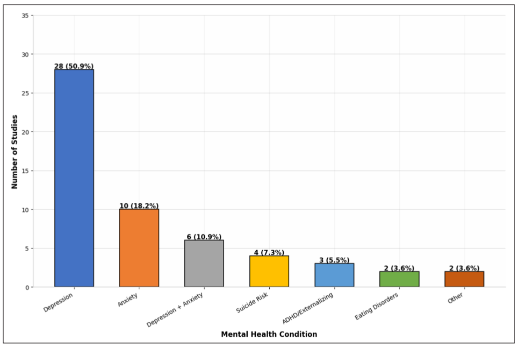

Figure 3: above shows the distribution of target mental health disorders across included studies. Percentages show the overwhelming focus on depression, with limited coverage of anxiety, stress, and other conditions.

Aside from simple categorization, the graphic summary helps identify areas of research deficit and developing trends in the subject area. As the graph indicates, although depression fills the ground, other disorders like anxiety, stress, and suicide risk find very little space. Much less represented are such topics as eating disorders and ADHD, although of essential value, these are found only sparingly in present literature. This discrepancy points to future research opportunities, specifically the necessity for greater diversification in AI research in adolescent mental health.

After the graphical illustration, Table 9 gives a comprehensive breakdown of the studies for each targeted mental health disorder. For every disorder, the table gives the number of papers that specifically mention it, as well as references to the seminal studies. This tabular presentation not only supplements the visual analysis but also enables a thorough analysis of the literature, providing a direct link to the specific works that are advancing each mental health research field.

Together, the pie chart and table provide a multi-view picture of the current landscape of mental health disorder prediction research based on AI and machine learning in adolescent populations. This imbalance indicates a research gap, where disorders beyond depression remain underexplored despite their clinical significance.

Most studies emphasize depression and anxiety due to their global prevalence in adolescence, but AI has also been applied to suicidal ideation, eating disorders, and ADHD.

- **Suicidal Ideation:** Passive sensing studies include suicide-specific risk prediction using GPS and phone use patterns [47,48], and chatbot-based screening using LLMs [43]. Classifiers have predicted suicidal ideation with balanced accuracy up to 77% [47], though large-scale validation remains limited.
- **Eating Disorders:** AI has classified eating disorder risk alongside depression and anxiety using social media image analysis [28], though dedicated eating disorder screening research remains scarce compared to depression.
- **ADHD/Externalizing Disorders:** Studies focus on detecting ADHD through smartphone usage patterns and behavioral markers [15,50], although externalizing disorder research remains limited compared to internalizing conditions.

While depression and anxiety dominate AI research [40], there is emerging work on suicidal ideation, eating disorders, ADHD, and other disorders. Most models remain domain-specific (e.g., depression detection, anxiety screening) [40], and broader detection faces technical challenges, particularly distinguishing overlapping symptom categories in adolescent populations [36, 21,22].

Table 9: Papers Explicitly Focused on Diagnosis or Prediction of Specific Mental Health Disorders

Targeted Mental Health Disorder	Count	Papers	Notes
Depression	28	[15,16,24, 27-51]	Most studied; strong text/ sensor signal
Anxiety	10	[16,17,37, 38,41,42]	Often comorbid with depression
Depression + Anxiety	6	[17,27,35, 36,38]	Transdiagnostic approaches
Stress	2	[15,22]	Limited coverage
Suicidal ideation/ Risk	4	[43,47,48]	Critical for safety; limited by rarity
Eating disorders	2	[28,30]	Emerging area; image-based detection
ADHD/ Externalizing	3	[15,50]	Behavioral markers from phone use
Total (Diagnosis-Focused)	**45**		

The bold number represents the total number of diagnosis-focused papers identified in the review

RQ5: Preprocessing and Privacy-Preserving Methods

To guarantee credible and precise prediction of mental health, most papers utilize certain preprocessing and data adjustment methods appropriate for their data form. These practices serve to normalize and process raw data before it is sent to machine learning or deep learning algorithms. Figure 4 shows a bar chart of methods specifically utilized by diagnosis-oriented articles from the discussed literature.

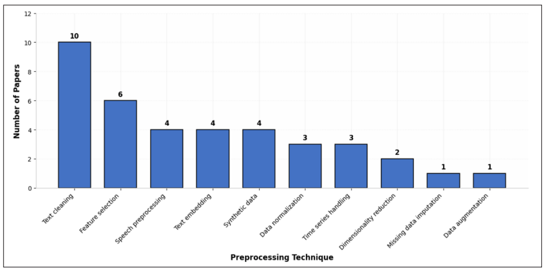

Figure 4: Preprocessing Techniques most Frequently Applied Across Included Studies. Text Cleaning and Feature Extraction Dominate, Reflecting the Prevalence of Social Media and Textual Data

Text preprocessing such as tokenization, stop-word removal, and lemmatization is the most widely used method of social media and clinical interview transcript preparation [24, 28, 29]. Feature selection techniques like SVM-RFE, Random Forest feature importance, and PCA are employed to remove noise and enhance model performance. Speech preprocessing tends to include audio feature extraction like MFCCs and LPCCs. Text embeddings like BERT and GPT represent semantic meaning from text for downstream classification. Some studies use synthetic data augmentation through GANs or LLMs to enrich datasets. Normalization and scaling bring feature ranges into harmony across physiological signals. Alignment techniques for time-series are employed in longitudinal or behavioral data. Dimensionality reduction compresses high-dimensional input without sacrificing key information. Some studies also impute missing values (e.g., via MICE) or use data augmentation (e.g., noise injection, rotation) for enhancing model robustness. These preprocessing decisions are central to preparing raw data to meet model needs and optimizing predictive results [56,59].

Privacy and ethical considerations specific to adolescents included:

- Removing personally identifiable information while preserving behavioral patterns [18,47]
- Training models across institutions without centralizing raw adolescent data [60-65]

- Adding statistical noise to protect individual-level information [47]
- Technical implementations ensuring guardian permission for minor data collection [66-70]
- Design considerations for adolescent cognitive development and digital literacy [71]

Studies emphasized that privacy protection is paramount when working with adolescent mental health data, given vulnerabilities related to stigma, parental oversight, and potential for data misuse [72-80]. Recent work by Isaiah et al. on AI companions for university students with social anxiety further highlights the importance of adaptive emotion regulation support while maintaining ethical safeguards in AI-based mental health interventions for young adults [81].

Discussion

This systematic literature review discloses the depth and width of applications of artificial intelligence (AI) in adolescent mental health, highlighting the potential and the complexity of merging AI-based technologies into clinical and non-clinical mental health interventions for youth. Though the field has expanded considerably in recent years, some key trends, challenges, and research gaps have emerged. This systematic literature review highlights several critical findings in the application of artificial intelligence (AI) for adolescent mental health:

- Depression was the most researched mental health disorder, with more than 68% of diagnosis-oriented studies focusing on it alone or alongside anxiety and stress disorders as shown in the Fig. 3.
- Applications of cutting-edge machine learning models, specifically transformer-based models (e.g., BERT, RoBERTa, GPT) and hybrid deep learning models (e.g., CNN-LSTM), have demonstrated excellent predictive accuracy in detecting mental health disorders in adolescents. Particularly, models like RoBERTa attained F1-scores of up to 0.97 on text data from Reddit [24].
- Multimodal studies involving multiple modalities like facial expressions, speech, text, and smartphone sensors demonstrated improved accuracy and resilience compared to unimodal methods. For instance, multimodal models like RLKT-MDD and the time-aware attention networks recorded F1-scores of over 80%, confirming the clinical relevance of integrating heterogeneous streams of data [27, 31,34,44].
- Large Language Models (LLMs), such as GPT-3.5 and GPT-4, were successfully utilized for diary text analysis, mock interviews, and depression screening in adolescents, exhibiting good recall and specificity in both zero-shot and fine-tuned environments [36,44,46]. While these findings underscore the potential of LLMs and chatbots as valuable tools for scalable mental health screening, they also raise complex ethical and practical challenges. Issues of privacy [35], data security [40,48], bias [36,39,40], interpretability [28], and clinical trust must be addressed before such systems can move from promising research prototypes to safe and effective clinical deployment. The sources position AI, particularly Large Language Models (LLMs) and specialized machine learning systems, primarily as auxiliary tools designed to augment, rather than replace, human expertise [33]. AI models can enhance clinical practice by providing supplementary information and decision support (DSS) [21,23].
- Data were pulled from sources like Reddit, Twitter, smartphone sensors, and wearable devices. However, challenges such as data imbalance, insufficient demographic diversity, and privacy remain significant hurdles to generalizability and clinical rollout [24,47,51,45].
- Growing concerns with Explainable AI (XAI), e.g., SHAP values, self-attention interpretability, portend an increasing trend towards more transparent, and hence, more trustworthy AI models, paramount for clinical adoption [47,21,36].
- The Analysis uncovers that while deep learning approaches achieve state-of-the-art accuracy, their lack of interpretability and high resource demands pose barriers to clinical adoption. By contrast, traditional ML offers greater transparency, supporting integration into existing workflows, though at the cost of lower performance on unstructured data. Multimodal systems provide comprehensive assessment, but raise ethical and operational challenges. These factors highlight that clinical deployment requires balancing predictive power with interpretability, scalability, and alignment with healthcare workflows.

Future Research Directions and Limitations

While this systematic review followed a structured protocol, certain limitations must be acknowledged. Only English-language publications were included, which may have excluded relevant research reported in other languages. The database selection was limited to PubMed, IEEE Xplore, ScienceDirect, Springer, and ACM Digital Library; although these are leading repositories, it is possible that relevant studies from other sources were missed. Finally, the use of specific search terms, despite employing multiple variations and Boolean combinations, may have restricted the scope and overlooked studies that used alternative terminology. Together, these factors constrain the comprehensiveness of the review, and future systematic studies should expand linguistic coverage, broaden database inclusion, and refine keyword strategies.

The application of Artificial Intelligence (AI) in adolescent mental health treatment has shown significant promise, but numerous avenues remain inadequately investigated and need greater attention. Recent literature greatly focuses on depression and anxiety, with relatively less research targeting other psychiatric disorders like bipolar disorder, schizophrenia, PTSD, and OCD in adolescent populations. Future studies need to cover underrepresented mental illnesses to make AI applications more inclusive and clinically meaningful across a broader range of diagnoses. The absence of standardized, large, and demographically diverse datasets restricts the applicability of current models. Future work needs to address creating open-access, ethically gathered datasets with data collected from diverse age groups, ethnicities, genders, and geographical locations. This will improve model resilience and help mitigate algorithmic bias. Although multimodal models (speech, text, EEG, facial expressions, smartphone sensors) perform better than unimodal, few have used longitudinal data to analyze temporal changes in adolescent mental health. Temporal dynamics should be the focus of future studies as well as the development of dynamic, personalized predictive models that evolve along with changing patient data. High levels of interpretability are needed in clinical deployment of AI systems. While methods such as SHAP and attention maps are becoming popular, there is still much work to be done to develop inherently interpretable models and interfaces that give useful feedback to clinicians, parents, and adolescent patients.

The robustness of AI models in the adolescent mental health domain remains an underexplored but critical factor for real-world adoption. Sensitivity analysis of key parameters such as data augmentation strategies, hyperparameter settings, uncertainty estimation methods and threshold stability provides insight into whether models yield consistent results under mental stress or across diverse adolescent populations. For example, digital phenotyping studies found that GPS-based features preserved stable accuracy across different school contexts while noise-based augmentations degraded performance [48], highlighting the role of parameter sensitivity in robustness. Similarly, hyperparameter variations in CNN and attention-based models showed that stable performance across a broad range of configurations is more clinically trustworthy than models highly sensitive to narrow settings [34]. Future research should therefore incorporate such sensitivity analyses to ensure reliability, generalizability, and clinical confidence in AI-driven adolescent mental health tools.

As sensitive personal information (e.g., social media updates, smartphone traces) is used more and more with adolescents, solid privacy-protecting methods like federated learning and differential privacy need to be explored further. In addition, ethical frameworks should be developed to control AI use, particularly in high-risk applications for adolescent mental health treatments, with specific attention to parental consent, minor assent, and data retention policies. Few of the AI systems have been translated to clinical use with adolescent populations. Subsequent research should give precedence to real-world validation using randomized controlled trials (RCTs), user-centered design with youth input, and collaboration with mental health clinicians, school counselors, and pediatricians to guarantee usability, safety, and efficacy in a variety of care environments. Large Language Models like GPT-4 have great potential for delivering mental health assistance via chatbot-driven interventions and AI-powered therapy aids for adolescents. Ongoing work needs to maximize prompt engineering, measure cultural sensitivity, ensure age-appropriate content, and investigate applications in hybrid human-AI models of therapy to promote responsible usage with minor populations. Recent research by Isaiah et al. [80] on AI companions for supporting adaptive emotion regulation among university students with social anxiety demonstrates the potential for AI-driven interventions to enhance emotional well-being in young adult populations, offering insights applicable to adolescent mental health support systems.

Conclusion

This study presents the latest developments in the use of Artificial Intelligence (AI) for adolescent mental health prediction, diagnosis, and treatment. The results show that there is an emphasis on conditions like depression and anxiety, and that transformer models, deep neural networks, and multimodal strategies are the most successful AI methods. Although most studies indicated strong predictive accuracy, the review also identifies several key challenges, such as the lack of standard and demographically specific datasets, minimal investigation of disorders other than depression and anxiety, and issues of ethical transparency, data privacy, and explainability. Even though AI models, especially those that exploit LLMs and multimodal fusion, show promising results, the disconnect between research results and clinical application remains wide.

To fill this gap, future work is required creating inclusive datasets, enhancing model interpretability, ethical deployment with appropriate safeguards for minors, and real-world validation of AI tools in clinical settings. With careful development and proper integration, AI can revolutionize adolescent mental health care rendering it more accessible, tailored, and efficient for diverse groups of young people worldwide.

Author Contributions

Uwemedimo Sunday Isaiah (USI) conceived the idea, designed the methodology, performed data curation, conducted formal analysis, and wrote the original manuscript. **Ekaette Mfon Useh (EMU), Iniobong George (IG), Chilaka Chika Franklin (CCF), Chineye Fabian Adili-George (CFA)** contributed to the literature search and initial study screening, assisted with data extraction, and reviewed and edited the manuscript. **Mfon Effiong Ineme (MEI)** assisted with critical review of the methodology, contributed to interpretation of results, and provided feedback during revisions.

Funding

No funding was received for this research.

Data Availability

No datasets were generated or analyzed during this study. All data cited are available from the original publications referenced.

Declarations

Ethical approval and consent to participate: Not applicable (systematic review of published literature).
Consent to publish: Not applicable.
Competing interests: The authors declare no conflicts of interest.

References

1. World Health Organization. Adolescent mental health. WHO Fact Sheets. 2023. Available from: https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health
2. Steinberg L. Cognitive and affective development in adolescence. Trends in Cognitive Sciences. 2005;9(2):69-74. https://doi.org/10.1016/j.tics.2004.12.005
3. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. Journal of Medical Internet Research. 2017 19: e176.
4. Torous J, Keshavan M. COVID-19, mobile health and serious mental illness. Schizophr Res. 2020. 218:36-37.
5. Li Y, Zhang Y, Chen Y, Wang Z, Liu X. Application of machine learning to predict mental health disorders and interpret feature importance. In: Proceedings of the IEEE International Conference on Computer and Information

Technology (CIBD). 2023. 257-261.
6. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. JMIR Ment Health. 2017. 4: e19.
7. Zanna K, Sridhar K, Yu H, Sano A. Bias reducing multitask learning on mental health prediction. In: 2022 10th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE. 2022. 1-8.
8. Cruz-Gonzalez P, He AW, Lam EP, Ng IM, Li MW, Hou R, et al. Artificial intelligence in mental health care: A systematic review of diagnosis, monitoring, and intervention applications. Psychol Med. 2025. 55: e18.
9. Dehbozorgi R, Zangeneh S, Khooshab E, Nia DH, Hanif HR, Samian P, et al. The application of artificial intelligence in the field of mental health: A systematic review. BMC Psychiatry. 2025. 25: 132.
10. Rony MKA. Artificial intelligence in psychiatry: A systematic review and meta-analysis of diagnostic and therapeutic efficacy. Digit Health. 2025. 11.
11. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: A systematic review. J Med Internet Res. 2017. 19: e176.
12. Feng X, Tian L, Ho GW, Yorke J, Hui V. The effectiveness of AI chatbots in alleviating mental distress and promoting health behaviors among adolescents and young adults: Systematic review and meta-analysis. J Med Internet Res. 2025. 27: e79850.
13. Patel J, Hung C, Katapally TR. Evaluating predictive artificial intelligence approaches used in mobile health platforms to forecast mental health symptoms among youth: A systematic review. Psychiatry Res. 2025. 343: 116277.
14. Mancin S, Sguanci M, Andreoli D, Soekeland F, Anastasi G, et al. Systematic review of clinical practice guidelines and systematic reviews: A method for conducting comprehensive analysis. MethodsX. 2024. 12: 102532.
15. Kim JS, Wang B, Kim M, Lee J, Kim H, et al. Prediction of diagnosis and treatment response in adolescents with depression by using a smartphone app and deep learning approaches: Usability study. JMIR Form Res. 2023. 7: e45991.
16. Heinz MV, Mackin DM, Trudeau BM, Bhattacharya S, Wang Y, et al. Randomized trial of a generative AI chatbot for mental health treatment. NEJM AI. 2025 2: AIoa2400802.
17. Maharjan SM, Poudyal A, van Heerden A, Byanjankar P, Thapa A, et al. Integration of passive sensing technology to enhance delivery of psychological interventions for mothers with depression: The Stand Strong study. Sci Rep. 2024. 14: 13535.
18. Zhang Z, Reavley N, Armstrong G, Morgan A. Public disclosures of mental health problems on social media and audiences' self-reported anti-stigma effects. Health Promot Int. 2025.40: daae204.
19. Ríssola EA, Aliannejadi M, Crestani F. Mental disorders on online social media through the lens of language and behaviour: Analysis and visualisation. Inf Process Manag. 2022. 59: 102890.
20. Thushari PD, Aggarwal N, Vajrobol V, Saxena GJ, Singh S, et al. Identifying discernible indications of psychological well-being using ML: Explainable AI in Reddit social media interactions. Soc Netw Anal Min. 2023. 13: 141.
21. Konopasek JE. The Distress Analysis Interview Corpus (DAIC) of human and computer interviews. Polygraph. 2023. 56(1).
22. Tabassum N, Ahmed M, Shorna NJ, Ur Rahman Sowad MD, Haque HM. Depression detection through smartphone sensing: A federated learning approach. Int J Interact Mob Technol. 2023. 17(1).
23. Shin D, Kim H, Lee S, Cho Y, Jung W. Using large language models to detect depression from user-generated diary text data as a novel approach in digital mental health screening: Instrument validation study. J Med Internet Res. 2024. 26: e54617.
24. Li J, Jiang M, Qin Y, Zhang R, Ling SH. Intelligent depression detection with asynchronous federated optimization. Complex Intell Syst. 2023. 9: 115-131.
25. Gupta S, Agarwal G, Agarwal S, Pandey D. Depression detection using cascaded attention based deep learning framework using speech data. Multimed Tools Appl. 2024. 83: 66135-66173.
26. Wajid A, Azam F, Anwar MW. Applications of artificial intelligence in mental health: A systematic literature review. Discov Artif Intell. 2025. 5: 332.
27. Hu P, Lin C, Li J, Tan F, Han X, Zhou X, et al. Making the implicit explicit: Depression detection in web across posted texts and images. In: 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE. 2023. 4807-4811.
28. Cheng JC, Chen AL. Multimodal time-aware attention networks for depression detection. J Intell Inf Syst. 2022. 59: 319-339.
29. Sadeghi M, Richer R, Egger B, Schindler-Gmelch L, Rupp LH, et al. Harnessing multimodal approaches for depression detection using large language models and facial expressions. NPJ Ment Health Res. 2024. 3: 66.
30. Yang S, Cui L, Wang L, Wang T, You J. Enhancing multimodal depression diagnosis through representation learning and knowledge transfer. Heliyon. 2024. 10.
31. Cabral RC, Luo S, Poon J, Han SC. 3M-Health: Multimodal multi-teacher knowledge distillation for mental health detection. In: Proceedings of the 33rd ACM International Conference on Information and Knowledge Management. 2024. 52-162.
32. Hu Y, Zhang S, Dang T, Jia H, Salim FD, et al. Exploring large-scale language models to evaluate EEG-based multimodal data for mental health. In: Companion of the 2024 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2024. 412-417.
33. Leal SS, Ntalampiras S, Sassi R. Speech-based depression assessment: A comprehensive survey. IEEE Trans Affect Comput. 2024.
34. Tao Y, Yang M, Shen H, Yang Z, Weng Z, Hu B. Classifying anxiety and depression through LLMs virtual interactions: A case study with ChatGPT. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM). 2023. 2259-2264.
35. Waaler PN, Hussain M, Molchanov I, Bongo LA, Elvevåg B. Prompt engineering an informational chatbot for education on mental health using a multiagent approach for enhanced compliance with prompt instructions: Algorithm development and validation. JMIR AI. 2025. 4: e69820.

36. Shinan-Altman S, Elyoseph Z, Levkovich I. Integrating previous suicide attempts, gender, and age into suicide risk assessment using advanced artificial intelligence models. J Clin Psychiatry. 2024. 85: 24m15365.
37. Xu X, Yao B, Dong Y, Gabriel S, Yu H, et al. Mental-LLM: Leveraging large language models for mental health prediction via online text data. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2024. 8: 1-32.
38. Gumilang MA, Karsanti HT, Linsiya RW, Yudanarko KD, Sari LN, et al. Development of a mental health chatbot using large language models for Indonesian undergraduates. In: 2024 International Conference on Computer, Control, Informatics and its Applications (IC3INA). IEEE; 2024. 337-342.
39. Aleem M, Zahoor I, Naseem M. Towards culturally adaptive large language models in mental health: Using ChatGPT as a case study. In: Companion Publication of the 2024 Conference on Computer-Supported Cooperative Work and Social Computing. 2024. 240-247.
40. Zhang X, Cui W, Wang J, Li Y. Chat, summary and diagnosis: A LLM-enhanced conversational agent for interactive depression detection. In: 2024 4th International Conference on Industrial Automation, Robotics and Control Engineering (IARCE). IEEE. 2024. 343-348.
41. Tang J, Shang Y. Advancing mental health pre-screening: A new custom GPT for psychological distress assessment. In: 2024 IEEE 6th International Conference on Cognitive Machine Intelligence (CogMI). IEEE; 2024. 162-171.
42. Yang K, Zhang T, Kuang Z, Xie Q, Huang J, et al. MentaLLaMA: Interpretable mental health analysis on social media with large language models. In: Proceedings of the ACM Web Conference (WWW). 2024. 4489-4500.
43. Cremaschi M, Ditolve D, Curcio C, Panzeri A, Spoto A, et al. Decoding the mind: A RAG-LLM on ICD-11 for decision support in psychology. Expert Syst Appl. 2025. 279: 127191.
44. Karkosz S, Szymański R, Sanna K, Michałowski J. Effectiveness of a web-based and mobile therapy chatbot on anxiety and depressive symptoms in subclinical young adults: Randomized controlled trial. JMIR Form Res. 2024. 8: e47960.
45. Kokane V, Abhyankar A, Shrirao N, Khadkikar P. Predicting mental illness (depression) with the help of NLP transformers. In: 2024 Second International Conference on Data Science and Information System (ICDSIS). IEEE. 2024. 1-5.
46. Wajid A, Azam F, Anwar MW. Applications of artificial intelligence in mental health: A systematic literature review. Discov Artif Intell. 2025. 5:332.
47. Dixit KK, Pundir S, Shrivastava A, Kumar CP, Srivastava AP, Singh P. Analyzing textual data for mental health assessment: Natural language processing for depression and anxiety. In: 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON). IEEE. 2023. 1796-1802.
48. Li W, Zhu Y, Li X, Li M, Jiang Z, et al. Zero-shot explainable mental health analysis on social media by incorporating mental scales. Proc ACM. 2024.
49. Zogan H, Razzak I, Wang X, Jameel S, Xu G. Explainable depression detection with multi-aspect features using a hybrid deep learning model on social media. World Wide Web. 2022. 25.
50. Wu F, Chen Z. Artificial intelligence in intelligent medical diagnosis for depression. In: 2023 3rd International Conference on Digital Society and Intelligent Systems (DSInS). IEEE. 2023. 248-251.
51. Roshchupkin G, van Hilten A, Katz S. Integrating multi-omics data with explainable AI to advance understanding of mental health disorders. Biol Psychiatry. 2025. 97: S16-S17.
52. Islam MR, Layek MA. Stackensemblemind: Enhancing well-being through accurate identification of human mental states using stack-based ensemble machine learning. Inf Med Unlocked. 2023. 43: 101405.
53. Ramzan HA, Abdulah F, Ahmad M, Ramzan S, Ashraf M. AI-driven personalization of e-therapy interventions for anxiety, stress, and depression. In: 2024 18th International Conference on Open-Source Systems and Technologies (ICOSST). IEEE. 2024. 1-6.
54. Kalra A, Gill KS. A neural approach to mental health: LSTM RNNs for depression detection. In: 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT). IEEE. 2024. 813-816.
55. Choi D, Zhang G, Jung J. Depression diagnosis algorithm based on 2-stream CNN using facial image. In: 2023 IEEE/ACIS 23rd International Conference on Computer and Information Science (ICIS). IEEE. 2023. 43-47.
56. Yang J, Lu H, Li C, Hu X, Hu B. Data augmentation for depression detection using skeleton-based gait information. Med Biol Eng Comput. 2022. 60: 2665-2679.
57. Kour H, Gupta MK. A hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM. Multimed Tools Appl. 2022. 81: 23649-23685.
58. Chen Y, Wang S, Guo J. DCTNet: Hybrid deep neural network-based EEG signal for detecting depression. Multimed Tools Appl. 2023. 82: 41307-41321.
59. Tian H, Zhang Z, Jing X. Deep learning for depression recognition from speech. Mob Netw Appl. 2023. 29: 1212-1227.
60. Nanavati J, Patel U. Hybrid model for analysis of social media posts for identification of depression and measuring its severity. In: 2023 International Conference on Data Science and Network Security (ICDSNS). IEEE. 2023. 1-5.
61. Li J, Jiang M, Qin Y, Zhang R, Ling SH. Intelligent depression detection with asynchronous federated optimization. Complex Intell Syst. 2023. 9: 115-131.
62. Farruque N, Goebel R, Sivapalan S, Zaïane OR. Depression symptoms modelling from social media text: An LLM driven semi-supervised learning approach. Lang Resour Eval. 2024. 58: 1013-1041.
63. Danner M, Hadzic B, Gerhardt S, Ludwig S, Uslu I, et al. Advancing mental health diagnostics: GPT-based method for depression detection. In: 2023 62nd Annual Conference of the Society of Instrument and Control Engineers (SICE). IEEE. 2023. 1290-1296.
64. Tang J, Guo Q, Zhao Y, Shang Y. Decoding linguistic nuances in mental health text classification using expressive narrative stories. In: 2024 IEEE 6th International Conference on Cognitive Machine Intelligence (CogMI). IEEE. 2024. 207-216.
65. Wang B, Sun Y, Zi Y, Zhao Y, Qin B. Scale-CoT: Integrating

LLM with psychiatric scales for analyzing mental health issues on social media. In: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE. 2024. 2651-2658.

66. Grodniewicz J, Hohol M. Therapeutic chatbots as cognitive-affective artifacts. Topoi. 2024. 43.
67. Nazareth P, Nikhil GB, Chirag G, Prathik NR. You Matter: A conversational AI powered mental health chatbot. In: 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE. 2024. 1-7.
68. Nassehi F, Par A, Eken A, Yetkin S, Eroğul O. Defining the optimal alpha band connectivity pathways to detect depression using explainable AI. In: 2024 32nd Signal Processing and Communications Applications Conference (SIU). IEEE. 2024. 1-4.
69. Poongodi K, Nivesh RA, Blessed Jeberson JQ, Divakar VS, Gowshikan S. Integrating Beck's Depression Inventory and emotion detection using GAN for comprehensive depression assessment and intervention. In: Proceedings of the International Conference on Sustainable Computing and Smart Systems (ICSCSS). 2024.
70. Hinduja S, Afrin M, Mistry S, Krishna A. Machine learning-based proactive social-sensor service for mental health monitoring using Twitter data. Int J Inf Manag Data Insights. 2022. 2: 100113.
71. Lee H, Park T, Lee U. Smartphone-based human behavior task modeling for explainable mental health detection model. Proc ACM. 2024. 894-896.
72. Jung G, Lee U. Counterfactual explanations in personal informatics for personalized mental health management. Proc ACM. 2024. 756-760.
73. Zhang X, Shi W. An AI-driven framework integrating predictive modeling and intervention strategies to enhance psychological health education among college students. Discov Appl Sci. 2025. 7: 1095.
74. Pourkeyvan A, Safa R, Sorourkhah A. Harnessing the power of Hugging Face transformers for predicting mental health disorders in social networks. IEEE Access. 2024. 12: 28025-28035.
75. Khoo LS, Lim MK, Chong CY, McNaney R. Machine learning for multimodal mental health detection: A systematic review of passive sensing approaches. Sensors. 2024. 24: 348.
76. Hermens DF, Russo C, Shan Z, Lagopoulos J. Brain fingerprinting: A promising future application for predicting mental illness. Futures. 2023. 152: 103211.
77. Fong D, Chu T, Heflin M, Gu X, Seneviratne O. Predicting depression and anxiety: A multi-layer perceptron for analyzing the mental health impact of COVID-19. Proc ACM. 2024. 325-331.
78. Blease C, Torous J. ChatGPT and mental healthcare: Balancing benefits with risks of harms. BMJ Ment Health. 2023. 26.
79. Habicht J, Dina LM, McFadyen J, Stylianou M, Harper R, et al. Generative AI-enabled therapy support tool for improved clinical outcomes and patient engagement in group therapy: Real-world observational study. J Med Internet Res. 2025. 27: e60435.
80. Isaiah US, Naibo AT, Ugondo MT, Ogwuche S, Iziengbuwa SJ, et al. The role of AI companions in supporting adaptive emotion regulation among university students with social anxiety. Discov Ment Health. 2026.