JOURNAL OF PSYCHOPATHOLOGY 2026;32:146-160
DOI: 10.36148/2284-0249-1143

Review

Mental health prediction in the era of AI: the review of machine learning techniques and ethical concerns

Daivashala Rajeshwarrao Deshmukh, Smita Kasar

Computer Science and Engineering, Maharashtra Institute of Technology, Chhatrapati Sambhajinagar (Aurangabad), Beed Bypass Road, Satara Parisar, Aurangabad, Maharashtra, India

SUMMARY

The human mental sensation is important for overall moral health and affects our social lives, our daily activities, and the quality of our lives. Since depression, anxiety, and chronic stress are becoming more commonly discussed worldwide, there is an urgent need for models that can quickly, precisely, and easily help predict them. Most traditional ways of diagnosing involve subjective views and clinicians' intuition, and these can lead to inconsistency, slow results, and be out of reach for many communities. Because of machine learning (ML), it is now possible to examine complex data from many sources, such as information shared on questionnaires, worn sensors and devices, observed online activity, and various therapy platforms. This work examines the latest advances in using machine learning for predicting mental well-being. The research looks at many types of algorithms, including Support Vector Machines, Random Forests, and deep learning models, which have helped to spot early signals of mental struggles. Subsequently, the work's major focus is given to the kind of data processed, significant performance factors, and the results observed with a model. Along with the paper, emphasis areas such as blending context-specific and user-adapted models were introduced as likely approaches for boosting accuracy and relevance. Apart from reviewing technical aspects, the paper looks at ethical and practical issues that cover privacy of data, discrepancies in algorithms, making sense of the data, and obeying regulations. The paper aims to bring together new studies, point out promising developments, and identify missing elements that must be handled for ML to succeed in mental health care. The information in this research can act as a reference for researchers, practitioners, and policymakers.

Key words: mental health prediction, machine learning, deep learning, mental well-being, data privacy, artificial intelligence, behavioral analytics

Daivashala Rajeshwarrao Deshmukh
E-mail: daivashala21334@gmail.com

How to cite this article: Deshmukh DR, Kasar S. Mental health prediction in the era of AI: the review of machine learning techniques and ethical concerns. Journal of Psychopathology 2026;32:146-160. https://doi.org/10.36148/2284-0249-1143

© Copyright by Pacini Editore Srl

OPEN ACCESS

This is an open access article distributed in accordance with the CC-BY-NC-ND (Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International) license. The article can be used by giving appropriate credit and mentioning the license, but only for non-commercial purposes and only in the original version. For further information: https://creativecommons.org/licenses/by-nc-nd/4.0/deed.en

Introduction

Mental well-being, which refers to a healthy and balanced mind, is considered a top priority in public health [1]. Many people globally suffer from depression, anxiety, and prolonged stress, leading to serious burdens economically and to health [2]. Older methods for discovering mental health problems mostly depend on therapist discussions, close observations, and patient-reported questionnaires [3, 4]. Even though these are clinically important, they usually take much time, produce inconsistent results, and are not available to most patients. The introduction of machine learning has made it possible to use real-time data more effectively to find and respond to mental health crises earlier than traditional models have [5].

Machine learning allows us to spot complicated relationships between several mental health indicators using information taken from mobile apps, social media, biometric devices, and electronic records [6]. Logistic regression, Random Forest, clustering models, and deep learning networks have all shown that they can detect emotions and future mental

health concerns with excellent accuracy [7]. If trained and updated with personal information, these models make it easier to identify mental distress without needing the user to give regular feedback [8].

Technological progress today allows us to measure and assess a lot of behavioral and physiological data in real time [9]. Today's devices and applications produce data on topics such as how well a person sleeps, activities they attempt, screen time, interactions with others, and their communication tone [10]. These streams allow mental health to be checked passively, making it easier when traditional in-person assessments cannot be used [11].

Furthermore, because of machine learning, systems are able to predict needs and preferences for an individual user [12]. Personalized ML models work better than those that treat different users the same way, as they learn from user-specific data and try to match every user's normal action [13]. In mental health, symptoms and pressures can be very different for people from different groups, cultures, or age ranges [14]. Using personalized methods allows support to be provided just when it is most urgent.

Even so, major issues do not have easy solutions; problems with data accuracy and abundance, unclear algorithms, and the threat of bias in ML models can stop a model from being applied in real life [15]. Ensuring user consent, who owns the data, and that privacy protection happens should be carefully handled, most importantly with sensitive mental health data [16]. Consequently, combining ML with mental health care needs attention to moral values so that technology remains beneficial for mental health and does not damage people's rights.

Literature review:

In the last few years, machine learning (ML) has attracted significant attention in the field of predicting mental health issues, due to rising concerns about mental health worldwide and more digital health data being collected [17]. With traditional diagnostics, people depend on their own opinions to find mental disorders, but ML uses hard data and automation to detect them more quickly and easily [18]. Many approaches are taken in the literature here, from basic methods to advanced programs, all focused on spotting individuals at risk of mental health problems using different kinds of data [19].

This review gathers the main findings from research studies to present an in-depth explanation of predicting mental health outcomes using ML [20]. Important subjects in this area include what types of ML models are chosen, the data sources involved, the conditions predicted (including anxiety, depression, or PTSD), the metrics to show performance, and the ethical and practical problems discovered [21]. All of these research projects give us a clear picture of the present implementation, pointing out the places where we can still improve and set common rules [22].

ML algorithms and strategies

Machine learning models used for predicting mental health problems are either classical or deep learning techniques. Because of their ease of use and simple explanations, RF, LR, DT, SVM, and KNN are used extensively among classical methods. Uddin et al, and Dhariwal et al. [23-24] examined LR, KNN, and RF by applying them to a public mental health survey dataset [25]. It was found that the Random Forest model, optimized with RandomizedSearchCV, performed best and gave an accuracy of 83.23% and an F1 score of 83.53%.

Deep learning has shown great value, mainly for working with data forms such as voice recordings, written text, and physical signals [26]. According to the review, CNNs, RNNs, and DNNs have recently become key methods for diagnosing bipolar disorder, autism, and PTSD [27]. Among many models, CNNs were shown to perform best in diagnostic accuracy owing to their excellent feature extraction abilities and adaptability.

Data sources and feature engineering

The effectiveness of an ML model often depends on the amount and diversity of data used. Examples of sources that contain mental health data include patient records (EHRs), social media interactions, information from self-administered tests (PHQ-9, OSMI), fitness tracker readings (steps taken, heart rate, sleep), and standard imaging procedures. (MRI scans) [28]. Showed how mixing survey data with MRI analysis using ML can boost accuracy in the detection of Alzheimer's and similar disorders.

Includes applications of ML for schizophrenia, anxiety, depression, and PTSD, focusing on how data from different sources is utilized [29]. It was also highlighted that managing a wide range of mental health features and taking into account patients' characteristics improves prediction.

Applications and Performances Assessed

Most of the research centers on the identification of disease, reactions to treatment, and measuring stress levels. According to [30], [31] a system can be built to analyze live data from wearables using RF, SVM, and XGBoost as ensemble models. The system gave immediate results on stress, with good predictions that could be used for constant mental health surveillance [32].

Existing researchers combined Random Forest and a large language model (LLM) to create a process that predicts mental health states from user information. Contextual feedback and explanation became possible when a language model was added to the system, allowing it to match predictions with personal insights [33]. These are typically found through the performances, including accuracy, precision, recall, F1-score, and AUC-ROC. Research shows that ensemble and deep models work better than simpler classifiers, mainly when they

are trained using big datasets [34]. Regardless, a number of models have problems with overfitting, being unable to generalize, and having no cross-validation.
Although the numbers look good, many difficulties remain. Providing sufficient and reliable mental health data is still difficult, mainly because of concern for privacy. If the training data used is biased, it may cause algorithms to practice discrimination regarding people's age, gender, and ethnicity [35] point out that using fairness-aware models and a range of data is essential in mental health research.
Understanding the results is a major issue as well. While it's true that complex models perform well, it can be hard to explain what they do. This can stand in the way of use in clinics, as clear reporting and accountability are necessary there. Yeasmin [1] described the need for transparent AI algorithms in mental health, which would increase clinician confidence and help gain needed approval.

Investigation of prior work

Many studies in ML for predicting mental health are part of a fast-developing area that involves the combination of data science with psychiatry, behavioral science, and digital health. The studies considered differ greatly with regard to target consumers, data available, how difficult the algorithms are, and how the research is applied clinically. While these models were created years ago, they are still popular because they work well and are easy to interpret with structured data. Due to being transparent and having a low computational cost, models such as those used by [3] and [18] are valuable in research using health data and reviewing clinical documents.
Recently, DL and mixed models have been used more often to handle the masses of unorganized data found on social media, in free-text therapy notes, and in readings from medical wearables. Research groups such as [23]and [12] made use of LSTM networks and CNN-LSTM hybrids to detect both depressive symptoms and emotional distress with high reliability (over 90%). By understanding time and context well, these models managed to find hidden signs of psychological risk in the everyday actions of people outside the clinic.
Many recent studies have noted that researchers are relying more on digital phenotyping to collect information about behavior using smartphones and wearables. The inputs used are measures of heart rate variability, routines related to sleep, the amount of screen time, and habitual social interaction. According to [6] and [10], these behavioral signals are capable of assessing student well-being or launching personalized mental health assistance in a timely way. Having adaptive systems within these applications signals that mental health monitoring and interventions are now happening in the background and are customizable to users.
Nevertheless, several issues consistently appear in the research literature. Dataset imbalance, standardization problems, and privacy, as well as explainability concerns, are generally recognized by the field. Most traditional models explain how they work, but black boxes are a challenge in using deep learning for patient care. Models like SHAP and LIME are currently being used to give explanations after the model has been run, but XAI still requires much stronger frameworks, studies show. If data is biased by culture and demographics, it cannot be generalized, so fairness-friendly studies and modeling approaches are important. Table I provides a comparative analysis of previous works.
In general, machine learning could greatly help increase and personalize mental health prediction and care. Classical models keep things stable, while deep learning helps study behavior. Break line patterns are found in big and unstructured information. All the reviewed papers agree that ML-based approaches can be applied for early screening, individual tracking, and mental health programs. This field must ensure it addresses ethics, bias, explainability, and how the technology will work outside research by continuing to deal with these issues.

Reevaluation of key studies

Mental health prediction based on machine learning (ML) is developing quickly as researchers study many types of models and data for clinical purposes. Analyzing recent influential studies again shows some strengths and some weaknesses of modern techniques. According to [3], Random Forests were able to correctly predict the outcomes of mental illness in surveyed individuals with an accuracy of 87%. Figure 1 describes the comparative performance of machine learning techniques in mental health prediction.
On the other side, deep learning and hybrid systems have proven able to recognize complex details in language and behavior from unstructured materials like text, speech, and social media posts. Using two ML algorithms, CNN and LSTM [12] and [30] can identify depression expression in over 93% of the Twitter messages they processed. Similarly, [23] posted for depressive symptoms using LSTM networks and obtained high F1 scores, which highlights deep learning's usefulness for mental health prediction in social media messages. They are valuable for both real-time and flexible applications due to the way they can handle and respond to new incoming data.
In addition, [6] and [10] focused on developing personalized and adaptive models. Due to contextual and lifestyle info on sleep, exercise, and eating habits, such systems can make more personalized predictions. Such methods are related to personalized medicine and encourage users by making things more meaningful to

TABLE I. *Comparative study table.*

Study	Data Source	Model/Approach	Prediction Focus	Key Findings
Chung & Teo (2022)[5]	Literature review	Taxonomy and ML mapping	Various mental health conditions	Arranged models based on their use and the type of data; highlighted what is meant by a model taxonomy.
Abdul Rahman et al. (2023) [6]	Student lifestyle behaviour data	Deep Neural Networks (DNN)	Well-being among university students	Accuracy was as high as 85% when I used activity, sleep, and nutrition inputs.
Rayan & Alanazi (2023) [2]	Survey and behavioral data	XGBoost with feature selection	General mental well-being	An accuracy level of 92.3% was possible by using structured questionnaire information.
Vaishnavi et al., (2022) [3]	Structured mental health survey	RF, SVM, KNN	Mental illness classification	Among the techniques, Random Forest had the greatest accuracy at 87%.
Garriga et al., (2022) [8]	Electronic Health Records (EHRs)	Ensemble ML (RF, GB)	Crisis detection in psychiatric patients	Detected 58% of the cases correctly, and 85% of the time, it showed specific accuracy.
Oyebode et al., (2023) [10]	Wearables and mobile usage	Personalized ML systems	Adaptive health monitoring	Widely recognized the importance of personalization and modelling based on the situation.
Al Banna et al. (2023) [12]	Social media (Twitter)	CNN + LSTM (hybrid)	COVID-19 mental health impact	The accuracy for finding depression-related expressions reached 93.1%.
Kaushik et al. (2022) [17]	Health insurance databases	Regression-based ML	Predicting premiums (well-being proxy)	The economic stressors route (the indirect prediction) is available.
Aleem et al., (2022) [18]	Clinical records and surveys	SVM, NB, Decision Trees	Depression diagnosis	Recorded outstanding F1-scores; however, it was suggested that the cause for ML use should be further explained.
Squires et al. (2023) [14]	Meta-analysis	Survey of DL/ML for psychiatry	Depression detection and treatment	Reinforced the actual problems that come with applying ML in clinical situations.
Nijhawan et al. (2022) [15]	Text from chats and forums	NLP + ML (SVM, NB)	Stress classification via language	Successfully achieved 89% accuracy with both linguistic and sentiment features used.
Rosaline et al., (2022) [20]	IT employee data	Modified Deep Learning NN	Melancholy/depression risk	The DL model reached an accuracy of 91.2%.
Uddin et al. (2022) [23]	Reddit mental health communities	LSTM on a large text dataset	Depressive symptoms detection	Scored an F1 of 0.89; works well on a massive scale with millions of posts.
Ameer et al., (2022) [27]	Social media text	Transfer Learning (BERT)	Mental illness classification	This achieved an accuracy of 91% using models that had already been trained.
Kour & Gupta (2022) [30]	Tweets	CNN + BiLSTM hybrid	Depression prediction	The hybrid deep model reached 94.5% accuracy owing to good feature extraction in space and time.

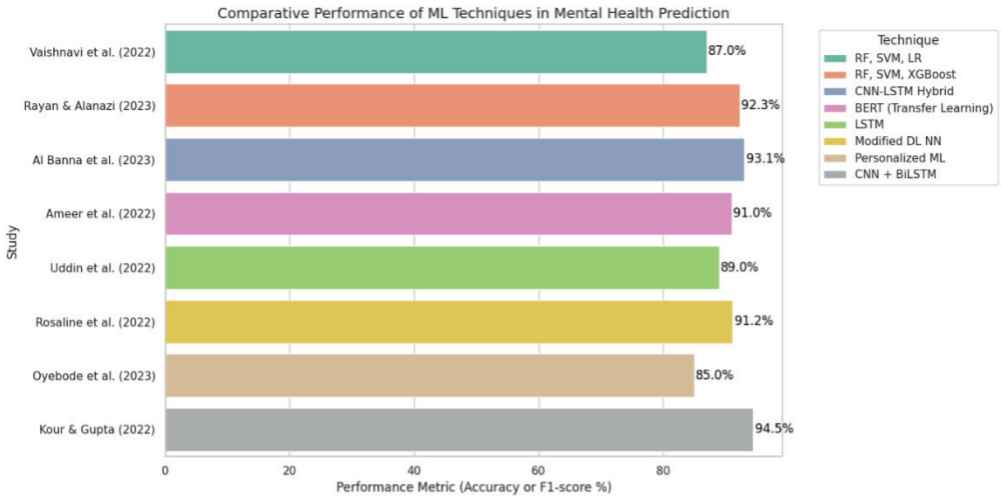


FIGURE 1. *Comparative performance of ML techniques in mental health prediction.*

them. Even so, these kinds of approaches must collect information regularly with sensors or wearables, which raises questions about battery usage, user agreement, and data protection.

Clinical integration is now attracting more attention. Garriga [8] built a model from electronic health records to foresee psychiatric crises. Because the system worked well and was based on all features, it was put to use in hospitals for identifying mental health emergencies. But such findings also reveal that making ML models consistent for every clinical setting is hard because of inconsistencies in record-keeping and data used [1]. In addition, there are many technically strong studies, but few are tested over enough time to be useful in translational research.

It is still very difficult to show how DL models reach their results. Conventional algorithms explain their outcomes simply, but with DL models, it is usually more complicated. They recommended that XAI tools be used in mental health, such as SHAP and LIME, to clarify why a computer-tested model predicts a high-risk result for a patient [36]. Because data privacy is vital in testing and treatment, people in healthcare trust this transparency. On the other hand, very few studies include these tools in their deployment process, so they rarely reach clinical use.

All in all, the literature suggests that many studies show that each methodology has its own benefits and weaknesses [37]. With classical models, it's easy to understand what's happening; deep learning enables growth and wide application; personalized systems give high accuracy; and clinical tests confirm their usefulness in practice. Yet, several key points arise repeatedly in the academic literature: transparency, variety in data, and responsibility to ethics [38].

Methodology

This study reviews knowledge about mental health prediction and machine learning by using a planned, open, and orderly process. The paper based the process on the PRISMA guidelines, involving the steps of identification, screening, assessing eligibility, and eventually including only relevant studies (Quazi, S) [39]. Research articles for this study were taken from reliable sources such as IEEE Xplore, MDPI, Elsevier, Springer, Wiley, and Frontiers. The chosen studies were studied using their research methods, choice of machine learning techniques, kinds of data, targeted mental health problems, the criteria for review, and their influence on the field [40]. By using this methodology, the review illustrates all the key updates in the field while remaining fair and honest. Figure 2 describes the diagrammatic representation of the research work.

In this review, the methodological framework is illustrated in Figure 2, which provides an end-to-end view of the machine learning (ML) pipeline for mental well-being prediction. The process begins with data acquisition from diverse sources such as surveys, social media platforms, wearable sensors, and electronic health records (EHRs). These raw inputs undergo essential preprocessing steps, including cleaning, tokenization, and missing value imputation, to ensure consistency and quality.

Subsequently, feature extraction techniques are applied to transform the raw data into structured representations, for example, natural language processing (NLP) embeddings for text, statistical summaries for survey data, and deep learning–derived feature maps for sensor signals. This is followed by feature selection, where relevant variables are isolated using methods like re-

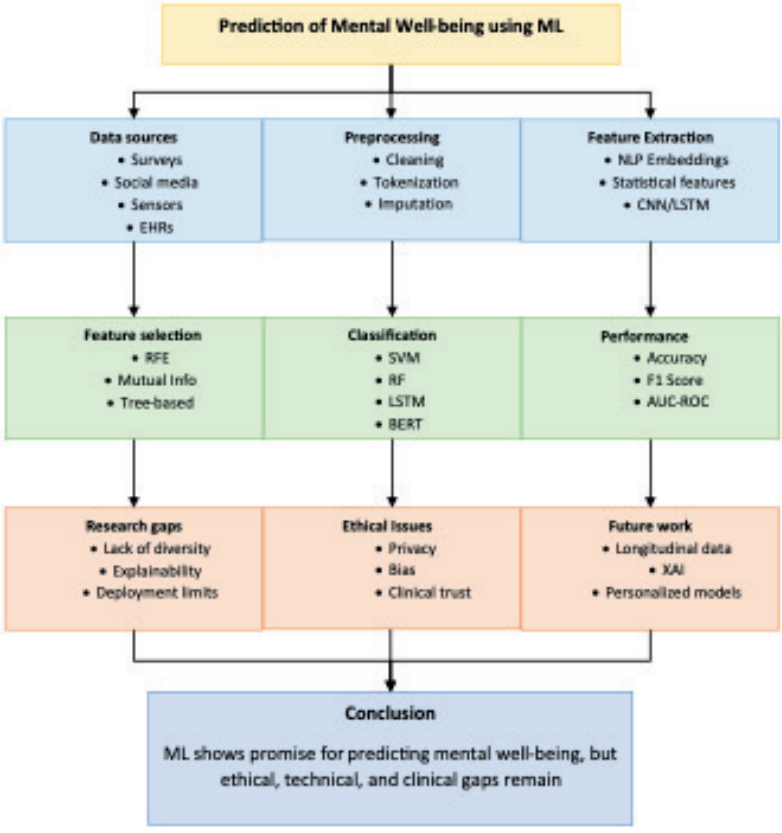


FIGURE 2. *Diagrammatic Summarization of the research.*

cursive feature elimination (RFE) or tree-based ranking. The selected features are then fed into classification models such as Support Vector Machines (SVM), Random Forests, Long Short-Term Memory (LSTM) networks, or transformer-based architectures (e.g., BERT) to predict the user's mental health status. Performance is assessed through evaluation metrics including accuracy, F1-score, and AUC-ROC, which help verify the effectiveness and robustness of the models.

At the bottom of the framework, key discussion areas are highlighted, such as the need for diverse data, model interpretability, and ethical considerations. These emphasize future directions like explainable AI, personalization, and real-world applicability, ensuring that ML systems remain technically robust, ethically sound, and clinically relevant. In addition, clear inclusion and exclusion criteria were applied: only original research articles (2018–2025) written in English, reporting ML-based prediction of mental health outcomes with performance metrics, were included; review papers, opinion pieces, low-quality studies, and those lacking sufficient methodological detail were excluded. This ensures transparency and replicability of the review process.

Following this methodological pipeline, Figure 3 presents the PRISMA flow diagram, which outlines the systematic process of identifying, screening, and selecting the studies included in this review.

This review follows the PRISMA guidelines to ensure methodological transparency. Articles were retrieved from IEEE, MDPI, Elsevier, Springer, Wiley, Frontiers, and other databases. The search was limited to English-language original research published between 2018 and 2025. Studies were included if they reported ML-based mental health prediction with performance metrics, while reviews, opinion pieces, low-quality studies, and those lacking sufficient methodological detail were excluded. Study quality was assessed based on methodological rigor and completeness. The PRISMA flow diagram illustrates this process: 880 records were identified, 365 duplicates removed, 515 screened by title and abstract, 113 full texts assessed, and 50 high-quality studies included. This structured approach ensures completeness, rigor, and replicability of the evidence base.

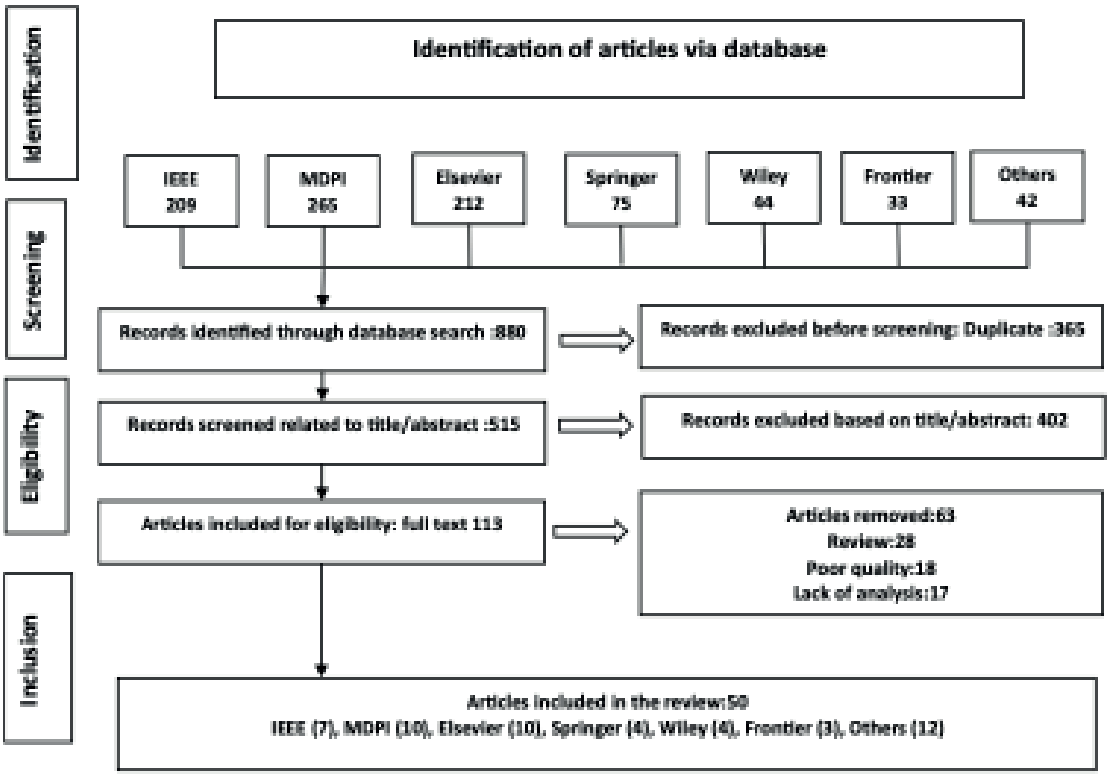


FIGURE 3. PRISMA flow diagram.

Data Collection

A strong mental health prediction system is dependent on good data collection. If the data is poor, the final model cannot be accurate, useful, or ethical. There are many places where we can gather data related to mental health, and each has its own set of positive and negative aspects [41]. Traditional questionnaires, for example, PHQ-9, GAD-7, and OSMI, allow for the collection of labeled and organized data that is crucial for certain artificial intelligence (AI) algorithms. Surveys of this type are approved and applied by clinicians everywhere, providing information on demographics, emotions, and work stress, all of which are suitable for initial design of a model [42].

In addition to surveys, today's machine learning models usually include data from wearables and phones that track activity and behavior. Using these tools, we are able to track in real time key measures like heart rate variability, time spent sleeping, how much you exercise, and the level of sweat on your skin [43]. Since these metrics connect to anxiety, stress, or depression, they let us observe fine data changes that may be tough to pick up with other procedures. When an individual has very little physical activity and irregular sleep, it might often point to depressive episodes. Social media use, app activities, and monitoring smartphone activity can give information about a person's emotional health [44]. They become very valuable as they are gentle on the body and can always be monitored without disturbing the user's regular schedule.

Furthermore, EHRs and recordings of voice, expressions, and regular therapy conversations have come to play a bigger role in collecting information [45]. Due to EHR records, models can include a patient's medical history, treatments, co-existing diseases, and medication usage over many years. Analyzing images and videos allows you to hear speech changes, hesitations, and notice signs of potential cognitive or emotional problems. Even so, assembling this information can be difficult, mainly due to matters related to privacy, consent, and how the data is managed [46]. Effective encryption, the right kinds of anonymization protocols, and sticking to privacy guidelines such as HIPAA and GDPR are required to look after sensitive mental health data.

Data Preprocessing

The next step is to rigorously process the data to make sure it is reliable, as it should be, and ready for machine learning [47]. Often, missing data, excess background noise, and inconsistencies in survey, wearable, and digital data can make predictions inaccurate. A variety of technological methods, including mean/mode imputation, KNN imputation, and advanced model-based techniques, are used to solve the issue of missing entries [48]. To make categorical variables (such as gender or educational level) usable by machines, they are encoded with one-hot or label encoding. Data is commonly standardized before SVM or KNN, which are both distance-based approaches [49]. When working with data from therapy transcripts or social media, the first step is to use tokenization, stemming, or lemmatization, get rid of stop words, and arrange it into TF-IDF or Word2Vec embeddings.

Many sensors (like EEG and heart rate) use preprocessing with style, including cutting their data into parts (time-series), eliminating noise (moving averages or wavelet transforms), and looking at features (temporal or frequency). The performance and reliability of any downstream analysis will be affected by the effective performance of the preprocessing method [50]. With synthetic oversampling (such as SMOTE), data augmentation techniques are put into place to close the gap when certain mental health conditions are less common. The objective of preprocessing is to organize the data so that it suits both the real-life situation and the used machine learning technique.

Feature Engineering and the Process of Selection

To create new inputs, engineers use feature engineering, while feature selection is about discovering the most helpful variables for forecasts [51]. Because much psychological data can be high-dimensional and noisy, following these steps is very important in mental health. Because of domain knowledge, sleep disruption, reported fatigue, and voice pitch variations can be made into composite features in psychology [52]. If digital phenotyping includes measuring screen time along with social interactions, it can show features that go along with feelings of isolation or sadness. Using engineered data can improve a model's ability to predict.

To make sure the model does not fit the data too closely and works on different data, select the most important features. Chi-square tests, mutual information, and ANOVA are used to numerically assess the importance of each feature [53]. People often use techniques such as Recursive Feature Elimination (RFE), LASSO (Least Absolute Shrinkage and Selection Operator), or tree-based feature importance from Random Forest or XGBoost. Principal Component Analysis (PCA) and t-SNE are methods that can lower computation time and maintain performance [54]. All in all, using these ways, models can spot meaningful relationships and make it easy for medical staff to understand and trust.

Model Selection

What matters most is finding the best model that is impacted by both your data and the type of task you are trying to solve. To predict mental health outcomes, algorithms frequently separate patients into groups considered at risk or healthy. LR is preferred because it is easy to understand and use when data is limited [55]. While Support Vector Machines are strong with multidimensional data, Decision Trees and Random Forests are known for their strong ability to handle overfitting. Because of ensemble learning, models like XGBoost and LightGBM give high accuracy and are commonly used in top solutions [56].

If you need to understand emotions from what people say or write or want to forecast their emotional states, deep learning is the best approach [57]. CNNs and RNNs make analyzing EEG patterns or voice spectrograms, both types of sequential or spatial data, easier. Transformer-based LLMs such as BERT and GPT are now used by some researchers to review transcripts or chats from therapy talks or online conversations [58]. In reality, researchers run several models and use basic evaluation metrics before choosing the one that provides a good balance among accuracy, how easy it is to explain, and resource needs.

Training and Validation Models

Using old data, we train models to spot patterns that can be used for new, untested information. Most of the time, the data is separated into training, validation, and testing, using training on 70–80% of the data and saving the other 20–30% for evaluation [59]. K-fold cross-validation is used widely to prevent the model from overfitting and ensure it does well on all parts of the data. While training, algorithms set their internal values so as to decrease a loss function such as binary cross-entropy for classification and mean squared error for regression [60]. Adam and SGD optimizers are commonly chosen for deep learning, and they are usually used together with learning rate schedules and dropout.

The process of finding the ideal hyperparameters alongside training is very important. Some settings, such as the depth of trees in Random Forest and the number of neural nodes, must be adjusted to achieve better outcomes [2]. People typically use grid search, random search, or Bayesian optimization to accomplish this task. Moreover, for problems with small severe depression cases, techniques such as SMOTE, under-sampling, or using weighted loss functions are commonly used [3]. Because of these methods, the model is both correct and able to understand and predict for different people.

Performance Evaluation

Following training, the model is now evaluated to assess its ability to predict mental health outcomes. In many cases, people use accuracy, precision, recall, and the F1-score to evaluate their results. The paper needs to be specific so that people are not incorrectly placed in at-risk groups, and you need to catch (or recall) as many cases as possible, as missing an important mental health situation can cause harm [5]. This metric merges the two by giving one overall number to measure the result. AUC-ROC allows us to view the relationship between the true positive and false positive rates at every decision threshold.

Confusion matrices are an extra tool that allows us to see how distinctly the model classes are identified. RMSE and MAE are the usual metrics applied in regression-based tasks. Checking if a model's predictions can truly be used in healthcare is necessary [6]. Later, experts may check the results or use predicted scores to predict the impact of interventions before actually performing them. And also examine the performance of the models amongst various subgroups (such as age, gender, and ethnicity) to support fairness and spot any possible differences.

Ease of Interpretation

It is important that mental health machine learning models can be interpreted rather than only being available when convenient. When a model's results are used to decide on treatment or diagnosis, everyone involved should be able to explain why that happened [8]. Because they are based on clear math, Logistic Regression and Decision Trees are naturally understandable. Even though explaining simple models is straightforward, medium and complex models have unique ways to explain their results.

Many are turning to SHAP and LIME to see which features contributed most to a prediction and to explain that prediction in a way people can easily grasp [10]. They can show that having less good sleep and checking digital screens more often led to a greater risk of depression. The method of saliency maps is applied in CNNs to underline the most essential features observed in EEG data or images [12]. By combining these techniques, one can increase trust and help debug, validate, and fulfill regulations required in healthcare.

Process of deploying and integration

Moving machine learning models into actual usage requires engineers to account for more technical and ethical points [61]. Trained models are usually made into APIs using Flask or FastAPI, helping them be used in mobile apps, smart watches, and digital therapy gadgets [14]. With these systems, ongoing behavioral and physiological data can be observed and actions taken or alerts sent based on what has been predicted [15]. Someone whose wearable data suggests stress could automatically get self-care advice or might be prompted to reach out to a therapist.

Integration requires performance checks, since this will tell if the model needs revision, should use fresh data, or if there is feedback from users [17]. Because data can be handled locally and privately, edge computing and federated learning are growing in popularity for predicting mental health. All information sent through the system should be protected by means of security and adherence to the necessary rules regarding privacy (such as GDPR and HIPAA) [18]. For a deployment to be successful, it should combine technological expertise and oversight by ethicists and clinicians so that the system always benefits patients and is both secure and practical for professionals.

Analysis and discussion

Here, the work reviews in detail the approaches, concepts, and achievements reported by the studies that analyzed mental health using machine learning. The paper attempts to point out patterns and both strengths and weaknesses in existing research by combining insights gained from preprocessing, feature engineering, model architectures, and evaluation metrics. By comparing different choices in the machine learning pipeline, it is revealed how these decisions may affect accuracy, understanding, and usefulness in everyday situations [20]. To bring more clarity to the subject, this chapter reviews important problems such as diversity in data, the ethics involved, and issues related to deployment. This section aims to combine theoretical experiments with usable knowledge, making what is learned here practical and useful for future steps in research and practice.

Comparative study analysis of pre-processing techniques:

Carrying out preprocessing is fundamental in machine learning tasks because it shapes how accurately and well a model performs. Raw mental health prediction data gathered from surveys, social media, or sensors is very often noisy, incomplete, or unshaped. Table I reveals that preprocessing has been customized for each data type used in the different studies. Research by [23] and [27] demonstrates how tokenization, lemmatization, and Word2Vec, as well as BERT, are applied to Reddit comments and Twitter posts. Abdul Rahman [6], discovered that structured survey data needed to be normalized, dealt with missing values, and converted categorical features to one-hot encodings. Table II provides the comparative study analysis of pre-processing techniques.

A key part of the preprocessing technique in sensor-based research [20] involved removing noise, making the signals look more seamless, and doing segmentation, which are vital for dealing with time-series physiological data. Generally, despite the many differences in preprocessing, all these techniques are used to prepare the data for model training. Picking the correct preprocessing methods depends on the data and learning algorithm to make sure the model remains robust and reliable, and it follows ethical guidelines in mental health modeling.

Comparative study analysis of feature extraction techniques

Raw data is translated into useful information, helping machine learning learn, through feature extraction. For prediction in mental health care, behavioral, physiological, and linguistic information is translated into helpful numerical values. Different methods for extracting features are shown in Table II based on the type of data used. As an illustration, Kour and Gupta [32], [23], and [10] applied both CNN and LSTM models to extract textual and timing patterns from tweets and internet forum posts, respectively, and [10] collected statistical data such as heart rate variability from wearables. Table III below provides the comparative study analysis of feature extraction techniques.

It is obvious that manual and automatic techniques give different results. Deep learning models depend on having many layers for training, compared to structured data which uses knowledge from the domain and manual method to determine useful features. Additionally, [12] showed how using different data such as text and images helps identify more complex mental health patterns. In multi-modal or large-scale datasets, feature extraction greatly influences how well your model performs, so it is very important in the machine learning workflow.

TABLE II. *Comparative study analysis of pre-processing techniques.*

Study	Data source Category	Preprocessing Methods	Observations
Ameer et al., (2022) [27]	Twitter data	BERT embeddings, punctuation cleaning, and emoji stripping	Optimized for sentiment-rich short texts
Uddin et al. (2022) [23]	Reddit text	Tokenization, stopword removal, lemmatization, Word2Vec	Handled noisy textual data from forums
Rosaline et al., (2022) [20]	Sensor (EEG)	Filtering, artifact removal, segmentation	Time-series cleaning for DL input
Aleem et al., (2022) [18]	Clinical forms	One-hot encoding, median imputation	Traditional preprocessing for EHR data
Abdul Rahman et al. (2023) [6]	Survey + lifestyle	Missing value imputation, normalization	Structured numerical & categorical data

TABLE III. *Comparative study analysis of feature extraction techniques.*

Exiting strategies	Feature Extraction Method	Data source Mode	Observations
Uddin et al., (2022) [23]	LSTM token encoding	Reddit posts	Captured sequential patterns of depressive terms
Kour & Gupta (2022) [32]	CNN + BiLSTM embeddings	Tweets	Captured temporal & spatial linguistic features
Al Banna et al., (2023) [12]	CNN for image/text features	Social media	Multimodal representation (text + activity)
Abdul Rahman et al., (2023) [6]	Manually engineered behavioral features	Health data	Included diet, exercise, and sleep indicators
Oyebode et al., (2023) [10]	Sensor signal stats (mean, HRV)	Wearables	Extracted stress markers from time-series

Comparative study analysis of feature selection strategies

Reducing the complexity of your model, making it easier to understand and avoiding overfitting are all reasons to practice feature selection, especially when working with big or noisy datasets. The reviewed studies used a variety of feature selection strategies and these are discussed in Table IV. To keep the most predictive variables, the authors used Recursive Feature Elimination which is reported by [3] and mutual information scoring based on work from Rayan & Alanazi [2]. Relying on embedded methods within Random Forest [8] made it easier to select the needed features automatically as the model was trained.

The success of choosing features depends in part on how relevant they are in real life as well as how significant they appear statistically. By connecting correlation matrices to the thinking of expert clinicians, [6] was able to choose features that matter both analytically and clinically. It appears that choosing the right features improves the model's accuracy and ensures that its actions are interpretable to medical professionals.

TABLE IV. *Comparative study analysis of feature selection strategies.*

Study	Feature Selection Technique	Justification	Consequence
Aleem et al., (2022) [18]	Information Gain, Chi-Square	Selected high-variance depression markers	Reduced training time
Vaishnavi et al., (2022) [3]	Recursive Feature Elimination (RFE)	Reduced overfitting from high-dim survey data	Improved SVM accuracy
Rayan & Alanazi (2023) [2]	Mutual Information Score	Evaluated non-linear feature relationships	Enhanced model interpretability
Abdul Rahman et al. (2023) [6]	Correlation Matrix + Expert Review	Combined statistical and clinical insight	Enhanced relevance of features
Garriga et al., (2022) [8]	Embedded (Tree-based feature ranking)	Native to the RF model	Built-in importance weights

Comparative study analysis of classification algorithm

Mental health prediction systems depend on classification algorithms which use the processed and selected traits as input to make helpful predictions. Table V compares the performance of various classifiers among the studies included. Many healthcare professionals choose these models due to their interpretable nature and resilience.

Deep learning models LSTM and hybrid CNN-LSTM delivered better results than traditional classifiers on data that comes in a sequence. According to [23] and [12], these approaches work better and give better results on social media text and mixed-media inputs. Also, recent studies that use BERT, for example, demonstrate how it effectively picks up on details in language [27]. It is clear from the comparison that data type, task complexity, target variable and both performance and interpretability should determine the chosen model.

TABLE V. *Comparative study analysis of the classification algorithm.*

Prior works	Classifier(s)	Performance	Superlative Prototypical
Vaishnavi et al., (2022) [3]	RF, SVM, LR	87% Accuracy	RF
Rayan & Alanazi (2023) [2]	RF, SVM, XGBoost	92.3% Accuracy	XGBoost
Al Banna et al., (2023) [12]	Hybrid CNN-LSTM	Accuracy: 93.1%	CNN-LSTM
Ameer et al., (2022) [27]	Transfer Learning (BERT)	F1-score: 0.91	BERT
Uddin et al., (2022) [23]	LSTM	F1-score: 0.89	LSTM

Research gap

While improvements in using machine learning (ML) for mental health prediction have been made, there is still a major lack of properly diverse and representative data [21]. The findings from these studies can't easily be applied to people outside of the study population because they use data collected only from university students or social media users in Western countries [23]. The fact that certain populations are usually overlooked means it's hard to roll out these models everywhere or use them reliably with people from different walks of life. Until now, there are no wide-scale, standardized datasets available to realistically show the variety of mental health experiences of different cultures.

Experts point out that dealing with mental health disorders that overlap or occur together remains an important gap in research [26]. A lot of models are set up to detect just depression or anxiety, but this doesn't reflect that many people present with both disorders combined. Focusing exclusively on one approach without sharing data oversimplifies mental health and could cause important distinctions in patient profiles to be overlooked [27]. Little research has been done on instruments that simultaneously predict several conditions and are able to consider how they work together.

Longer-term studies and models that account for changes through time are badly needed in mental health ML. Most current models use data from a single time point and believe mental states do not change over time, which isn't true for mental health [27]. A lack of recurrent inputs and special architectures prevents models from accurately monitoring symptoms or the risks of relapse [28]. More realistic and actionable predictions in mental health can be made by combining longitudinal observation, sequential learning, and data showing behavior changes over time.

In addition, the mysterious way some deep learning approaches work limits how doctors trust and use them. When you want a model to be accurate, it often turns out to be less interpretable. In healthcare, it becomes vital that all predictions are open and justified since decisions are based on them. SHAP and LIME tools are starting to emerge, yet most studies still do not use explainability as an important principle in their designs [29]. So far, not many XAI methods have been designed for

situations in psychology, where subjective and sensitive variables exist.

On another note, there are unnoticed social, legal, and ethical issues related to ML in mental health care [30]. Little attention is given in studies to informed consent, who owns data, security issues with surveillance, and forms of algorithmic bias. Because personal information about mental health is sensitive and can be stigmatized, it is very important to develop solutions that are both good at their job and respect user rights [32]. There is still not much research available on using ethical questions in model design, testing, and deployment.

While it's true that papers show that these systems work in laboratories, only a few have been used in real clinical or digital health environments [33]. Clinical use of medical technologies is rarely discussed, and there is little attention given to testing how easy they are to use, what clinicians think, or continuous tracking. As a consequence, a lot of good models are not put to proper use [34]. To improve, future efforts must check prototypes in the field, get feedback from all stakeholders during development, and ensure they are useful in actual mental health support.

Discussion and implications

The use of ML is transforming the way depression, anxiety, and stress are detected and managed. Classical algorithms such as RF, LR, and SVM remain valuable for structured datasets due to their interpretability and reliability, whereas deep learning approaches like CNNs, LSTMs, and their hybrids excel with unstructured inputs such as social media posts, spoken language, and wearable sensor data [35]. Together, these complementary strengths suggest that ML will play a central role in the future of diagnosis and treatment in mental health care.

A consistent lesson from the reviewed studies is the importance of data diversity and representation. Many investigations rely heavily on narrow populations—such as university students or English-language tweets—which limits generalizability [1]. Such practices risk embedding bias into models, making them less effective for culturally diverse or underserved groups. Addressing fairness and inclusivity is therefore essential to ensure that ML-driven mental health solutions are globally relevant and equitable.

Personalized approaches are emerging as another critical direction. By leveraging digital phenotypes and adaptive technologies, tailored ML models can align predictions and interventions with an individual's lifestyle and behavior patterns [36]. This personalization offers promise for preventive care, enabling early intervention before symptoms escalate. However, it simultaneously raises concerns regarding privacy and data protection, given the sensitive and continuous nature of behavioral data collection.

Despite encouraging results, translating ML models into clinical practice remains challenging. Reported issues include limited interpretability, poor clinician trust, and a lack of long-term, real-world validation [37]. To be clinically adopted, systems must comply with regulations such as GDPR and HIPAA while maintaining transparency. Explainability tools like SHAP and LIME are promising but remain underutilized and are not yet tailored for clinical psychology.

Beyond established frameworks such as GDPR and HIPAA, more recent regulatory developments are also shaping the landscape of AI in healthcare. The forthcoming EU AI Act introduces a risk-based framework that classifies clinical decision-support and predictive tools as high-risk AI systems. These systems must comply with strict obligations, including risk management, high-quality data governance, technical documentation and record-keeping, human oversight, robustness and cybersecurity safeguards, transparency to deployers, and post-market monitoring. For mental-health ML, this implies documenting the provenance and representativeness of training data, monitoring bias and subgroup performance, enforcing model change-control procedures, and maintaining clear human-in-the-loop escalation pathways in clinical practice. Aligning model development with harmonized standards such as ISO/IEC 42001 (AI management systems), ISO/IEC 23894 (AI risk management), and the NIST AI RMF 1.0 will further operationalize compliance across the model lifecycle. Importantly, EU AI Act compliance complements but does not replace health-data and privacy regulations such as GDPR, which remain central to protecting sensitive mental-health data.

However, there are still ongoing ethical issues. The examined literature seldom addresses topics like algorithmic fairness, data ownership, transparency, and informed permission [38]. Given that mental health data is both highly sensitive and stigmatized, even small errors can lead to serious consequences. To mitigate these risks, future research must embed ethical design principles from the outset and maintain them throughout deployment, ensuring that technological innovation progresses hand in hand with patient rights and societal trust.

Finally, significant methodological gaps are evident. Most ML systems are not equipped to handle comorbid or longitudinally evolving conditions (Quazi,) [39]. Many models rely on single time-point data, which fails to capture the dynamic nature of mental health. More advanced sequential modeling, together with real-time data from digital platforms, could enable more realistic and clinically actionable predictions.

Performance reporting across studies also reveals considerable heterogeneity. Some models trained on small,

homogeneous samples achieved very high accuracy but risk overfitting, while others using multi-source datasets with independent validation achieved more moderate yet reliable outcomes. Differences in validation strategies, ranging from simple train-test splits to rigorous k-fold cross-validation, further complicate comparison. Moreover, demographic factors such as age, gender, and cultural background are rarely analyzed despite their potential influence on prediction accuracy. This raises the risk of algorithmic bias, with models performing well on majority groups but poorly on marginalized populations.

Taken together, these findings show that reported performance metrics alone are not sufficient indicators of clinical readiness. Without addressing heterogeneity, validation rigor, population fairness, and ethical considerations, even the most accurate ML models risk failing when deployed in real-world healthcare environments.

Conclusion

The analysis has systematically addressed the use of ML for monitoring and predicting various mental health conditions, pointing out the main achievements as well as enduring issues. With an increase in global mental health problems, there is a strong need for helpful tools and approaches that complement traditional mental health diagnosis. ML offers a good solution, as it can find useful patterns in complex data such as questionnaire responses, biological information, health records, and online talks. Automated detection of depression, anxiety, and stress in these models allows for faster and more consistent identification of risk, which cannot be done as effectively using other approaches. One major observation from this study is the maturing suitability of ML algorithms used in classical and deep learning. RF, SVM, and LR are known to work well on structured data and are appreciated for their clear and simple way of working as well as being easy to compute. Likewise, on unstructured and multimodal data such as material from social networks and physiological signals gathered over time, CNNs, RNNs, and matching hybrid CNN-LSTM models have demonstrated exceptional performance. Likewise, these results prove that ML can handle multiple kinds of data in research on mental health.

Even so, there are big gaps that need to be filled. Most of the studies reviewed are held back by datasets made up mostly of individuals from a single type of population. Due to this issue, the usefulness of the models is harmed, and the risk exists that biases would continue to disadvantage already marginalized populations. Also, a few models are set up to notice comorbidities or look at mental health changes over a period, although mental illnesses tend to overlap and change with time. Future models are badly needed that can continuously adapt depending on how the user is doing. For clinical use, interpreting complex models is still a problem. Although these systems often work very well, little is known about how and why they decide. As treatment decisions in mental healthcare must be simple and logical, this difficulty with interpretability makes it hard to add the AI to regular workflows. SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are helpful, but they should be adapted more often for healthcare applications to be meaningful.

At the same time, what to do with mental health information is guided by both ethical and legal concerns. Privacy, deciding on informed consent, owning your data, fair use of algorithms, and possible surveillance must be handled through clear rules. To respect both a person's right and dignity, we need to design predictive technologies for mental health so that they are very accurate and secure. Principles of ethics need to be part of how data is collected, models are built, systems are used, and feedback is given to maintain both personal freedom and social trust. In summary, machine learning could change mental health prediction and care a lot if future research made inclusivity, transparency, interpretability, and ethical integrity a priority. Together, people from different fields, including technology, mental health, ethics, and policies, should join forces to produce tools that are both sound and helpful. Facing these many issues will allow us to come closer to offering prompt, personal, and open mental health care for everyone.

Conflict of interest statement

The author declare that they have no conflict of interest

Funding

The authors state that this work has not received any funding.

Ethical Approval

Institutional Review Board approval was not required.

Consent for Participation

All contributors agreed and gave consent to participate.

Consent for Publication

All contributors agreed and gave consent to publish.

Data availability

No data, models, or code were generated or used during the study.

Author Contribution

The authors confirm their contribution to the paper as follows and all authors reviewed the results and approved the final version of the manuscript.

References

1 Yeasmin S, Das S, Bhuiyan MFH, et al. Artificial intelligence in mental health: leveraging machine learning for diagnosis, therapy, and emotional well-being. J Ecohumanism. 2025;4(3):286-300. https://doi.org/10.62754/joe.v4i3.6640

2 Rayan A, Alanazi S. A novel approach to forecasting the mental well-being using machine learning. Alex Eng J. 2023;84:175-183. https://doi.org/10.1016/j.aej.2023.10.060

3 Vaishnavi K, Kamath UN, Rao BA, et al. Predicting mental health illness using machine learning algorithms. J Phys Conf Ser. 2022;2161(1):012021. https://doi.org/10.1088/1742-6596/2161/1/012021

4 Wahed SA, Wahed MA. AI-driven digital well-being: developing machine learning model to predict and mitigate internet addiction. LatIA. 2025;3:73. https://doi.org/10.62486/latia202573

5 Chung J, Teo J. Mental health prediction using machine learning: taxonomy, applications, and challenges. Appl Comput Intell Soft Comput. 2022;2022:9970363. https://doi.org/10.1155/2022/9970363

6 Abdul Rahman H, Kwicklis M, Ottom M, et al. Machine learning-based prediction of mental well-being using health behavior data from university students. Bioengineering (Basel). 2023;10(5):575. https://doi.org/10.3390/bioengineering10050575

7 Kumar K, Verma A, Verma P, et al. Unveiling mental stress: examining the impact on college students' well-being using machine learning based solutions. In: 2024 IEEE International Conference of Electron Devices Society Kolkata Chapter (EDKCON). IEEE; 2024. p. 179-184. https://doi.org/10.1109/EDKCON62339.2024.10870747

8 Garriga R, Mas J, Abraha S, et al. Machine learning model to predict mental health crises from electronic health records. Nat Med. 2022;28(6):1240-1248. https://doi.org/10.1038/s41591-022-01811-5

9 Oparina E, Kaiser C, Gentile N, et al. Machine learning in the prediction of human wellbeing. Sci Rep. 2025;15(1):1632. https://doi.org/10.1038/s41598-024-84137-1

10 Oyebode O, Fowles J, Steeves D, et al. Machine learning techniques in adaptive and personalized systems for health and wellness. Int J Hum Comput Interact. 2023;39(9):1938-1962. https://doi.org/10.1080/10447318.2022.2089085

11 Mishra AR, Rai A, Nandan D, et al. Unveiling emotions: NLP-based mood classification and well-being tracking for enhanced mental health awareness. Math Model Eng Probl. 2025;12(2). https://doi.org/10.18280/mmep.120228

12 Al Banna MH, Ghosh T, Al Nahian MJ, et al. A hybrid deep learning model to predict the impact of COVID-19 on mental health from social media big data. IEEE Access. 2023;11:77009-77022.

13 Wang AX, Nguyen BP, Elliott T, et al. Early detection of depression using machine learning and social well-being survey data. In: 2024 16th International Conference on Computer and Automation Engineering (ICCAE). IEEE; 2024. p. 181-186. https://doi.org/10.1109/ICCAE59995.2024.10569222

14 Squires M, Tao X, Elangovan S, et al. Deep learning and machine learning in psychiatry: a survey of current progress in depression detection, diagnosis and treatment. Brain Inform. 2023;10(1):10. https://doi.org/10.1186/s40708-023-00188-6

15 Nijhawan T, Attigeri G, Ananthakrishna T. Stress detection using natural language processing and machine learning over social interactions. J Big Data. 2022;9(1):33. https://doi.org/10.1186/s40537-022-00575-6

16 Sambandam RK, Vetriveeran D, Jenefa J, et al. Unlocking insights into mental well-being: a deep learning approach to depression detection. In: International Conference on Soft Computing and Signal Processing. Singapore: Springer; 2025. p. 535-545. https://doi.org/10.1007/978-981-97-9926-8_41

17 Kaushik K, Bhardwaj A, Dwivedi AD, et al. Machine learning-based regression framework to predict health insurance premiums. Int J Environ Res Public Health. 2022;19(13):7898. https://doi.org/10.3390/ijerph19137898

18 Aleem S, Huda NU, Amin R, et al. Machine learning algorithms for depression: diagnosis, insights, and research directions. Electronics. 2022;11(7):1111. https://doi.org/10.3390/electronics11071111

19 Castro Ribeiro T, García Pagès E, Ballester Coma L, et al. Design of a remote multiparametric tool to assess mental well-being and distress in young people (mHealth Methods in Mental Health Research Project): protocol for an observational study. 2024. https://doi.org/10.2196/51298

20 Rosaline S, Nasreen MA, Suganthi P, et al. Predicting melancholy risk among IT professionals using modified deep learning neural network (MDLNN). In: 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT). IEEE; 2022. p. 385-391. https://doi.org/10.1109/CSNT54456.2022.9787571

21 Oyeleye M, Chen T, Titarenko S, et al. A predictive analysis of heart rates using machine learning techniques. Int J Environ Res Public Health. 2022;19(4):2417. https://doi.org/10.3390/ijerph19042417

22 Massa SM, Porcu E, Riboni D. An initial investigation of mental well-being monitoring through personal healthcare devices. In: Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization. 2024. p. 410-414. https://doi.org/10.1145/3631700.3665240

23 Uddin MZ, Dysthe KK, Følstad A, et al. Deep learning for prediction of depressive symptoms in a large textual dataset. Neural Comput Appl. 2022;34(1):721-744. https://doi.org/10.1007/s00521-021-06426-4

24 Dhariwal N, Sengupta N, Madiajagan M, et al. A pilot study on AI-driven approaches for classification of mental health disorders. Front Hum Neurosci. 2024;18:1376338. https://doi.org/10.3389/fnhum.2024.1376338

25 Cheong J, Spitale M, Gunes H. Small but fair! Fairness for multimodal human-human and robot-human mental wellbeing coaching. arXiv. 2024:2407.01562.

26 Chang V, Ganatra MA, Hall K, et al. An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators. Healthc Anal. 2022;2:100118. https://doi.org/10.1016/j.health.2022.100118

27 Ameer I, Arif M, Sidorov G, et al. Mental illness classification on social media texts using deep learning and transfer learning. arXiv. 2022:2207.01012.

28 Bhatnagar S, Agarwal J, Sharma OR. Detection and classification of anxiety in university students through the application of machine learning. Procedia Comput Sci. 2023;218:1542-1550. https://doi.org/10.1016/j.procs.2023.01.132

29 Hinduja S, Afrin M, Mistry S, et al. Machine learning-based proactive social-sensor service for mental health monitoring using Twitter data. Int J Inf Manag Data Insights. 2022;2(2):100113. https://doi.org/10.1016/j.jjimei.2022.100113

30 Kour H, Gupta MK. An hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM. Multimed Tools Appl. 2022;81(17):23649-23685. https://doi.org/10.1007/s11042-022-12648-y

31 Koutsouleris N, Hauser TU, Skvortsova V, et al. From promise to practice: towards the realisation of AI-informed mental health care. Lancet Digit Health.

2022;4(11). https://doi.org/10.1016/S2589-7500(22)00153-4

32 Gupta P, Chouhan AV, Wajeed MA, et al. Prediction of health monitoring with deep learning using edge computing. Meas Sens. 2023;25:100604. https://doi.org/10.1016/j.measen.2022.100604

33 Tutun S, Johnson ME, Ahmed A, et al. An AI-based decision support system for predicting mental health disorders. Inf Syst Front. 2023;25(3):1261-1276. https://doi.org/10.1007/s10796-022-10282-5

34 Abdullah KH, Sofyan D. Machine learning in safety and health research: a scientometric analysis. Int J Inf Sci Manag. 2023;21(1):17-37. https://doi.org/10.22034/ijism.2022.1977763.0

35 Leghissa M, Carrera Á, Iglesias CA. Machine learning approaches for frailty detection, prediction and classification in elderly people: a systematic review. Int J Med Inform. 2023;178:105172. https://doi.org/10.1016/j.ijmedinf.2023.105172

36 Mittal S, Mahendra S, Sanap V, et al. How can machine learning be used in stress management: a systematic literature review of applications in workplaces and education. Int J Inf Manag Data Insights. 2022;2(2):100110. https://doi.org/10.1016/j.jjimei.2022.100110

37 Higgins O, Short BL, Chalup SK, et al. Artificial intelligence (AI) and machine learning (ML) based decision support systems in mental health: an integrative review. Int J Ment Health Nurs. 2023;32(4):966-978. https://doi.org/10.1111/inm.13114

38 Liu D, Feng XL, Ahmed F, et al. Detecting and measuring depression on social media using a machine learning approach: systematic review. JMIR Ment Health. 2022;9(3). https://doi.org/10.2196/27244

39 Quazi S. Artificial intelligence and machine learning in precision and genomic medicine. Med Oncol. 2022;39(8):120.

40 Savci M, Tekin A, Elhai JD. Prediction of problematic social media use (PSU) using machine learning approaches. Curr Psychol. 2022;41(5):2755-2764. https://doi.org/10.1007/s12144-020-00794-1

41 Arashpour M, Golafshani EM, Parthiban R, et al. Predicting individual learning performance using machine-learning hybridized with the teaching-learning-based optimization. Comput Appl Eng Educ. 2023;31(1):83-99. https://doi.org/10.1002/cae.22572

42 Solayman S, Aumi SA, Mery CS, et al. Automatic COVID-19 prediction using explainable machine learning techniques. Int J Cogn Comput Eng. 2023;4:36-46.

43 Revathi A, Kaladevi R, Ramana K, et al. Early detection of cognitive decline using machine learning algorithm and cognitive ability test. Secur Commun Netw. 2022;2022:4190023. https://doi.org/10.1155/2022/4190023

44 Javeed A, Dallora AL, Berglund JS, et al. Machine learning for dementia prediction: a systematic review and future research directions. J Med Syst. 2023;47(1):17. https://doi.org/10.1007/s10916-023-01906-7

45 Devi KJ, Alghamdi W, Alkhayyat A, et al. Artificial intelligence in healthcare: diagnosis, treatment, and prediction. E3S Web Conf. 2023;399:04043.

46 Hayat A, Morgado-Dias F, Bhuyan BP, et al. Human activity recognition for elderly people using machine and deep learning approaches. Information. 2022;13(6):275. https://doi.org/10.3390/info13060275

47 Amanat A, Rizwan M, Javed AR, et al. Deep learning for depression detection from textual data. Electronics. 2022;11(5):676. https://doi.org/10.3390/electronics11050676

48 Morid MA, Sheng ORL, Dunbar J. Time series prediction using deep learning methods in healthcare. ACM Trans Manag Inf Syst. 2023;14(1):1-29.

49 Zhang T, Schoene AM, Ji S, et al. Natural language processing applied to mental illness detection: a narrative review. NPJ Digit Med. 2022;5(1):46. https://doi.org/10.1038/s41746-022-00589-7

50 Lo-Ciganic WH, Donohue JM, Yang Q, et al. Developing and validating a machine-learning algorithm to predict opioid overdose in Medicaid beneficiaries in two US states: a prognostic modelling study. Lancet Digit Health. 2022;4(6). https://doi.org/10.1016/S2589-7500(22)00062-0

51 Ajagbe SA, Adigun MO. Deep learning techniques for detection and prediction of pandemic diseases: a systematic literature review. Multimed Tools Appl. 2024;83(2):5893-5927. https://doi.org/10.1007/s11042-023-15805-z

52 Javaid M, Haleem A, Singh RP, et al. Significance of machine learning in healthcare: features, pillars and applications. Int J Intell Netw. 2022;3:58-73. https://doi.org/10.1016/j.ijin.2022.05.002

53 Garg A, Parashar A, Barman D, et al. Autism spectrum disorder prediction by an explainable deep learning approach. Comput Mater Contin. 2022;71(1):1459-1471. https://doi.org/10.32604/cmc.2022.022170

54 AlShorman O, Masadeh M, Heyat MBB, et al. Frontal lobe real-time EEG analysis using machine learning techniques for mental stress detection. J Integr Neurosci. 2022;21(1):20.

55 Huang J, Galal G, Etemadi M, et al. Evaluation and mitigation of racial bias in clinical machine learning models: scoping review. JMIR Med Inform. 2022;10(5). https://doi.org/10.2196/36388

56 Zhou S, Zhao J, Zhang L. Application of artificial intelligence on psychological interventions and diagnosis: an overview. Front Psychiatry. 2022;13:811665. https://doi.org/10.3389/fpsyt.2022.811665

57 Raj M, Singh S, Solanki K, et al. An application to detect cyberbullying using machine learning and deep learning techniques. SN Comput Sci. 2022;3(5):401.

58 Leist AK, Klee M, Kim JH, et al. Mapping of machine learning approaches for description, prediction, and causal inference in the social and health sciences. Sci Adv. 2022;8(42).

59 Tejaswini V, Sathya Babu K, Sahoo B. Depression detection from social media text analysis using natural language processing techniques and hybrid deep learning model. ACM Trans Asian Low Resour Lang Inf Process. 2024;23(1):1-20. https://doi.org/10.1145/3569580

60 Balcombe L, De Leo D. Human-computer interaction in digital mental health. Informatics. 2022;9(1):14. https://doi.org/10.3390/informatics9010014

61 Solomon DD, Sonia, Kumar K, et al. Extensive review on the role of machine learning for multifactorial genetic disorders prediction. Arch Comput Methods Eng. 2024;31(2):623-640. https://doi.org/10.1007/s11831-023-09996-9