Mental Wellness Toolkit

MANYA VAID
1BM22CS150

MARIA SAYEEMA
1BM22CS151

N PRIYANKA
1BM22CS167

NALLABOTHULA
SAI SRUTHI
1BM22CS170

Guide: DR. PALLAVI G B
Designation: ASSOCIATE PROFESSOR
Department of Computer Science & Engineering
B.M.S. College of Engineering

OUTLINE

1. Introduction
2. Problem Statement
3. Literature Survey
4. References

Introduction

Mental health disorders, such as depression, anxiety, and eating disorders, are increasingly becoming a global health crisis. The emergence of advanced technologies like machine learning (ML) and artificial intelligence (AI) has opened new avenues for detecting, diagnosing, and treating these conditions. Leveraging the vast data generated from social media, wearable devices, and clinical records, AI-driven tools are reshaping mental healthcare by providing more personalized, scalable, and accessible solutions. This project aims to develop a "Mental Wellness Toolkit," focusing on integrating AI and machine learning methodologies to enhance mental health support.

Problem Statement

Despite significant advancements in the understanding of mental health disorders, early detection and personalized treatment remain challenging. Traditional diagnostic methods are often subjective and time-consuming, leading to delayed or inadequate treatment. Additionally, the stigma associated with mental health prevents many individuals from seeking professional help. There is a critical need for accessible, efficient, and objective tools to diagnose and manage mental health conditions. This project addresses these challenges by developing an AI-powered system that utilizes various machine learning algorithms and natural language processing (NLP) techniques to provide real-time mental health support, enabling early diagnosis and personalized treatment.

Literature Survey

Title: Potential benefits and limitations of machine learning in the field of eating disorders: current research and future directions

Link:https://link.springer.com/article/10.1186/s40337-022-00581-2

Methodology:

The paper reviews existing applications of machine learning (ML) in detecting, preventing, and treating eating disorders.Supervised Learning: Primarily used for eating disorder detection (e.g., Logistic Regression, Support Vector Machines).Unsupervised Learning: Applied to explore hidden patterns in unlabeled data (e.g., clustering). Neural Networks: Used for complex tasks like processing neuroimaging data. Transformers (e.g., RoBERTa): For analyzing text-based content on social media to detect disordered behaviors.

Dataset:

The dataset employed combines multiple sources to enhance supervised machine learning (ML) models. It includes self-reported survey data on symptoms and eating behaviors, naturalistic social media content from platforms like Instagram and Twitter, and neuroimaging data such as fMRI scans for identifying biomarkers of eating disorders. Additionally, genetic data from genome-wide studies contributes to understanding risk factors for conditions like anorexia nervosa, providing a comprehensive foundation for ML analysis.

Pros:

- Early Detection
- Efficiency
- Scalability
- Personalization

Cons:

- Limited Generalizability
- Dropout and Engagement Challenges
- Privacy Concerns

Results:

The results indicate that machine learning (ML) methods can accurately predict eating disorder status, achieving accuracy rates between 70-91% from survey responses and social media data. Notably, early-warning systems for anorexia nervosa inpatients have reached predictive accuracies as high as 98%. Additionally, chatbots have demonstrated effectiveness in reducing symptoms and weight-related concerns, with significant improvements observed over six months.

Conclusion:

The use of machine learning has vast potential to revolutionize the detection, prevention, and treatment of eating disorders. However, significant challenges remain in making these tools accurate, unbiased, and generalizable.

Literature Survey

Title: An advanced Artificial Intelligence platform for a personalised treatment of Eating Disorders
Link:https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2024.1414439/full

Methodology:

The study introduces a Master Data Platform (MDP) integrated with artificial intelligence (AI) to tailor treatment for eating disorders (EDs). It employs various AI algorithms, including Machine Learning (ML) for identifying risk factors, analyzing symptoms, and predicting relapses. Additionally, Natural Language Processing (NLP) is utilized for chatbots that interact with patients, assess records, and monitor treatment adherence, while Deep Learning (DL) analyzes complex data like neuroimaging and behavioral patterns from wearable sensors.

Dataset:

The dataset comprises patient records from various care centers in the Campania region, including psychosocial variables, clinical assessments, and health records. Data from wearable devices is also included, which monitors physiological conditions and behaviors such as heart rate and sleep patterns.

Pros:

- ❏ Improved Accessibility
- ❏ Collaborative Care
- ❏ Scalable Infrastructure

Cons:

- ❏ Integration Complexity
- ❏ Cybersecurity Risks
- ❏ Technology Accessibility Issues

Results:

The expected results from implementing this AI-powered platform include enhanced patient satisfaction and adherence to treatment through personalized care. The system is anticipated to improve the management of waiting lists by streamlining care processes, enabling real-time interventions through AI-driven insights and monitoring. Moreover, it aims to increase awareness of eating disorders and improve healthcare resource allocation across the region.

Conclusion:

The AI-powered platform offers a promising solution for managing eating disorders by combining data-driven algorithms with a patient-centered approach. It enhances collaboration between healthcare providers and patients, ensuring timely and effective care.

Literature Survey

Title: Application of machine learning methods in predicting schizophrenia and bipolar disorders: A systematic review

Link: https://onlinelibrary.wiley.com/doi/full/10.1002/hsr2.962

Methodology:

This paper reviews studies that applied machine learning (ML) models to predict schizophrenia and bipolar disorder, focusing on articles from databases like PubMed, Scopus, and Google Scholar. The studies analyzed used ML techniques on data such as neuroimaging, genetic information, and clinical features. Inclusion criteria required studies to report model performance using metrics like accuracy, precision, recall, or area under the curve (AUC). The review followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for study selection and analysis.

Dataset:

The paper summarizes various datasets used in studies predicting mental disorders, including neuroimaging (MRI, fMRI), genomic data, and clinical records. Many studies relied on publicly available datasets, such as the ENIGMA Schizophrenia dataset, while others used custom data. However, the variation in datasets across studies presents a challenge in standardizing the performance of machine learning models for predicting mental disorders.

Pros:

- ❑ Potential to integrate a variety of data types
- ❑ Application of ML can lead to more personalized treatment

Cons:

- ❑ Struggle with the generalization problem
- ❑ Lack of standardization in ML techniques

Results:

Machine learning models, particularly support vector machines (SVM) and random forests, performed well in predicting schizophrenia and bipolar disorders, with several studies achieving AUC values above 0.80. Neuroimaging data, especially from functional and structural MRI, provided the most predictive features, while combining data from multiple domains, such as clinical and genetic information, further improved model performance in some cases.

Conclusion:

The review emphasizes the need for larger, standardized datasets and better integration of diverse data types to enhance the performance and generalizability of machine learning models in psychiatry. While the findings demonstrate the value of ML in aiding early diagnosis and personalized treatment, further research is required to standardize methodologies and validate results in larger, more diverse populations..

Literature Survey

Title: Automatic Diagnosis of Bipolar Disorder Using Optical Coherence Tomography Data and Artificial Intelligence

Link: https://www.mdpi.com/2075-4426/11/8/803

Methodology:

The study involved 17 patients with type I bipolar disorder (BD) and 42 age- and gender-matched healthy controls, analyzing nine retinal layers using the Early Treatment Diabetic Retinopathy Study (ETDRS) chart. Retinal thickness measurements were obtained using a SPECTRALIS HRA+OCT device, focusing on layers such as the retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), and inner plexiform layer (IPL). AI classifiers, including Gaussian Naive Bayes, K-nearest neighbors (KNN), and Support Vector Machines (SVM), were used to classify based on key retinal data. Statistical analysis included the Shapiro–Wilk test, t-Student, Mann–Whitney U-test, Pearson's correlation coefficient, and AUC for evaluating diagnostic capacity.

Dataset:

The study used OCT data from 59 individuals (17 with BD and 42 controls), with retinal thickness measurements across various retinal layers and regions.

Pros:

- ❑ Non-invasive and cost-effective diagnostic approach.
- ❑ AI classifiers achieved high accuracy
- ❑ OCT provides a safe and accessible biomarker

Cons:

- ❑ The study focused only on type I BD
- ❑ Lack of longitudinal data
- ❑ Small sample size, limiting the generalizability of the results.

Results:

The study found retinal thinning in bipolar disorder (BD) patients, especially in the ganglion cell and inner plexiform layers, with strong diagnostic potential (AUC > 0.75). AI classifiers performed well, with the linear Support Vector Machine (SVM) achieving the highest classification accuracy of 95%.

Conclusion:

The study suggests that structural retinal changes, as detected by OCT, can serve as a potential biomarker for BD. AI methods can improve diagnostic accuracy, but further studies with larger, multicenter datasets are needed to confirm these findings

Literature Survey

Title:A Novel Application for the Efficient and Accessible Diagnosis of ADHD Using Machine Learning

Link:https://ieeexplore.ieee.org/document/9311012

Methodology:

The system uses pupillometry to track changes in pupil size during a memory task as an indicator of ADHD. Several machine learning algorithms were applied, with an ensemble voting model providing the best performance. Data was validated using Leave-One-Out Cross-Validation (LOOCV).

Dataset:

The study involved 50 participants (28 with ADHD and 22 controls), aged 10-12. Pupillometric data was collected over 160 trials per subject during a visuospatial memory task.

Pros:

- High accuracy (sensitivity: 0.821, AUROC: 0.856)
- Objective and non-invasive
- Accessible, especially in resource-limited areas

Cons:

- Small dataset, possibly limiting generalizability
- Reliance on pupil dynamics alone might overlook other ADHD symptoms

Results:

The Ensemble Voting Model achieved 82.1% sensitivity, 72.7% specificity, and 85.6% AUROC, outperforming other models like Random Forest and SVM.

Conclusion:

This ADHD diagnostic system is better because it uses pupillometry for an objective diagnosis, offering more accuracy, affordability, and accessibility than traditional methods. Its real-time web app makes it a quick and reliable solution, particularly helpful in areas with limited healthcare

Literature Survey

Title: Neuro Intel: A System for Clinical Diagnosis of Attention Deficit Hyperactivity disorder (ADHD) using Artificial Intelligence

Link:https://ieeexplore.ieee.org/document/10218313

Methodology:

The Neuro Intel system uses a hybrid AI model combining machine learning (decision trees) and knowledge-based rules for ADHD diagnosis. When both models agree, the case is diagnosed automatically, and complex cases are referred to experts

Dataset:

Clinical data from an NHS mental health provider using tools like PHQ-9, GAD-7, and CAARS ADHD scores.

Pros:

- 92% accuracy in straightforward cases
- Combines AI and expert knowledge
- Reduces clinician workload

Cons:

- Limited dataset impacts generalization
- Complex cases still need expert input

Results:

The hybrid model achieved 92% accuracy, automating diagnosis for half the cases and referring complex ones to clinicians.

Conclusion:

Neuro Intel offers a highly accurate and efficient ADHD diagnostic tool by blending AI and expert knowledge, speeding up diagnosis while maintaining trust with explainable AI

Literature Survey

Title: Machine Learning in ADHD and Depression Mental Health Diagnosis: A Survey
Link:https://ieeexplore.ieee.org/document/10214293

Methodology:

The paper explores machine learning models like Support Vector Machines (SVM) and Neural Networks for ADHD and depression diagnosis. It uses wearable (EEG) and non-wearable (MRI, speech, and video) modalities to train models. These data sources help differentiate between control subjects and patients with ADHD or depression.

Dataset:

The survey discusses multiple datasets, including the ADHD-200 dataset, with 973 participants, and the DAIC-WOZ dataset for depression diagnosis

Pros:

- ❑ High accuracy in ADHD (97.6%) and depression (90%)
- ❑ Use of multi-modal data improves diagnosis reliability

Cons:

- ❑ Small datasets limit generalizability
- ❑ Data privacy concerns restrict access

Results:

SVM and Neural Networks showed high accuracy in detecting ADHD and depression. SVM models were particularly effective for EEG and MRI data, achieving up to 97.6% accuracy for ADHD detection

Conclusion:

Machine learning models, especially SVMs and Neural Networks, effectively diagnose ADHD and depression using multi-modal data like EEG, MRI, speech, and video. These methods outperform traditional approaches in accuracy, offering non-invasive and objective diagnostics. However, limited datasets and privacy concerns hinder broader applications.

Literature Survey

Title: Machine Learning based Detection of Post Traumatic Stress Disorder of Mental Health
Link:https://ieeexplore.ieee.org/document/10193036

Methodology:

The system uses various machine learning algorithms such as Random Forest, Gradient-Boosted Decision Trees, and Logistic Regression to predict Post-Traumatic Stress Disorder (PTSD). The data includes both social media posts (Twitter) and clinical data, with algorithms analyzing language patterns and psychological metrics

Dataset:

The study uses data from over 243,000 tweets from PTSD patients, as well as clinical datasets from military personnel and trauma-exposed individuals

Pros:

- ❑ High accuracy (97% for Random Forest)
- ❑ Utilizes real-time social media and clinical data
- ❑ Effective for early intervention in PTSD case

Cons:

- ❑ Sensitive to data quality
- ❑ Limited by dataset size for generalization across populations

Results:

The Random Forest algorithm showed the highest accuracy at 97%, outperforming other methods like Bagging (95%) and SVM (91%).

Conclusion:

The system demonstrates how machine learning, particularly Random Forest, can accurately detect PTSD using a combination of social media and clinical data. This approach offers high accuracy, making it a valuable tool for early detection and intervention in PTSD cases

Literature Survey

Title: Artificial-Intelligence based Prediction of Post-Traumatic Stress Disorder (PTSD) using EEG reports
Link: https://ieeexplore.ieee.org/document/10072671

Methodology:

This system uses EEG (Electroencephalography) reports to predict Post-Traumatic Stress Disorder (PTSD). The method involves collecting EEG data, pre-processing it for noise reduction, and extracting relevant features. A Convolutional Neural Network (CNN) model is then trained on this EEG data to classify individuals as either having PTSD or not

Dataset:

The dataset consists of EEG recordings from PTSD patients. EEG data is used to track brain activity, and specific channels from the EEG report are used for classification into PTSD or non-PTSD groups.

Pros:

- ❏ High accuracy (85%)
- ❏ Non-invasive (EEG-based)
- ❏ Early PTSD detection

Cons:

- ❏ Data-dependent
- ❏ Requires specialized EEG equipment

Results:

The CNN model used in this system showed high accuracy in predicting PTSD based on EEG reports. The performance of the model was validated using cross-validation, ensuring that the model generalized well across different datasets

Conclusion:

This system offers a highly accurate, non-invasive, and early diagnostic tool for PTSD using EEG data. By applying machine learning techniques such as CNN, it significantly improves upon traditional, subjective methods of PTSD diagnosis, offering a more reliable approach

Literature Survey

Title: Potential benefits and limitations of machine learning in the field of eating disorders: current research and future directions

Link:https://link.springer.com/article/10.1186/s40337-022-00581-2

Methodology:

The paper reviews existing applications of machine learning (ML) in detecting, preventing, and treating eating disorders.Supervised Learning: Primarily used for eating disorder detection (e.g., Logistic Regression, Support Vector Machines).Unsupervised Learning: Applied to explore hidden patterns in unlabeled data (e.g., clustering). Neural Networks: Used for complex tasks like processing neuroimaging data. Transformers (e.g., RoBERTa): For analyzing text-based content on social media to detect disordered behaviors.

Dataset:

The dataset employed combines multiple sources to enhance supervised machine learning (ML) models. It includes self-reported survey data on symptoms and eating behaviors, naturalistic social media content from platforms like Instagram and Twitter, and neuroimaging data such as fMRI scans for identifying biomarkers of eating disorders. Additionally, genetic data from genome-wide studies contributes to understanding risk factors for conditions like anorexia nervosa, providing a comprehensive foundation for ML analysis.

Pros:

- ❏ Early Detection
- ❏ Efficiency
- ❏ Scalability
- ❏ Personalization

Cons:

- ❏ Limited Generalizability
- ❏ Dropout and Engagement Challenges
- ❏ Privacy Concerns

Results:

The results indicate that machine learning (ML) methods can accurately predict eating disorder status, achieving accuracy rates between 70-91% from survey responses and social media data. Notably, early-warning systems for anorexia nervosa inpatients have reached predictive accuracies as high as 98%. Additionally, chatbots have demonstrated effectiveness in reducing symptoms and weight-related concerns, with significant improvements observed over six months.

Conclusion:

The use of machine learning has vast potential to revolutionize the detection, prevention, and treatment of eating disorders. However, significant challenges remain in making these tools accurate, unbiased, and generalizable.

Literature Survey

Title: An advanced Artificial Intelligence platform for a personalised treatment of Eating Disorders
Link:https://www.frontiersin.org/journals/psychiatry/articles/10.3389/fpsyt.2024.1414439/full

Methodology:

The study introduces a Master Data Platform (MDP) integrated with artificial intelligence (AI) to tailor treatment for eating disorders (EDs). It employs various AI algorithms, including Machine Learning (ML) for identifying risk factors, analyzing symptoms, and predicting relapses. Additionally, Natural Language Processing (NLP) is utilized for chatbots that interact with patients, assess records, and monitor treatment adherence, while Deep Learning (DL) analyzes complex data like neuroimaging and behavioral patterns from wearable sensors.

Dataset:

The dataset comprises patient records from various care centers in the Campania region, including psychosocial variables, clinical assessments, and health records. Data from wearable devices is also included, which monitors physiological conditions and behaviors such as heart rate and sleep patterns.

Pros:

- ❑ Improved Accessibility
- ❑ Collaborative Care
- ❑ Scalable Infrastructure

Cons:

- ❑ Integration Complexity
- ❑ Cybersecurity Risks
- ❑ Technology Accessibility Issues

Results:

The expected results from implementing this AI-powered platform include enhanced patient satisfaction and adherence to treatment through personalized care. The system is anticipated to improve the management of waiting lists by streamlining care processes, enabling real-time interventions through AI-driven insights and monitoring. Moreover, it aims to increase awareness of eating disorders and improve healthcare resource allocation across the region.

Conclusion:

The AI-powered platform offers a promising solution for managing eating disorders by combining data-driven algorithms with a patient-centered approach. It enhances collaboration between healthcare providers and patients, ensuring timely and effective care.

Literature Survey

Title: Application of machine learning methods in predicting schizophrenia and bipolar disorders: A systematic review

Link: https://onlinelibrary.wiley.com/doi/full/10.1002/hsr2.962

Methodology:

This paper reviews studies that applied machine learning (ML) models to predict schizophrenia and bipolar disorder, focusing on articles from databases like PubMed, Scopus, and Google Scholar. The studies analyzed used ML techniques on data such as neuroimaging, genetic information, and clinical features. Inclusion criteria required studies to report model performance using metrics like accuracy, precision, recall, or area under the curve (AUC). The review followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for study selection and analysis.

Dataset:

The paper summarizes various datasets used in studies predicting mental disorders, including neuroimaging (MRI, fMRI), genomic data, and clinical records. Many studies relied on publicly available datasets, such as the ENIGMA Schizophrenia dataset, while others used custom data. However, the variation in datasets across studies presents a challenge in standardizing the performance of machine learning models for predicting mental disorders.

Pros:

- Potential to integrate a variety of data types
- Application of ML can lead to more personalized treatment

Cons:

- Struggle with the generalization problem
- Lack of standardization in ML techniques

Results:

Machine learning models, particularly support vector machines (SVM) and random forests, performed well in predicting schizophrenia and bipolar disorders, with several studies achieving AUC values above 0.80. Neuroimaging data, especially from functional and structural MRI, provided the most predictive features, while combining data from multiple domains, such as clinical and genetic information, further improved model performance in some cases.

Conclusion:

The review emphasizes the need for larger, standardized datasets and better integration of diverse data types to enhance the performance and generalizability of machine learning models in psychiatry. While the findings demonstrate the value of ML in aiding early diagnosis and personalized treatment, further research is required to standardize methodologies and validate results in larger, more diverse populations..

Literature Survey

Title: Automatic Diagnosis of Bipolar Disorder Using Optical Coherence Tomography Data and Artificial Intelligence

Link: https://www.mdpi.com/2075-4426/11/8/803

Methodology:

The study involved 17 patients with type I bipolar disorder (BD) and 42 age- and gender-matched healthy controls, analyzing nine retinal layers using the Early Treatment Diabetic Retinopathy Study (ETDRS) chart. Retinal thickness measurements were obtained using a SPECTRALIS HRA+OCT device, focusing on layers such as the retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), and inner plexiform layer (IPL). AI classifiers, including Gaussian Naive Bayes, K-nearest neighbors (KNN), and Support Vector Machines (SVM), were used to classify based on key retinal data. Statistical analysis included the Shapiro–Wilk test, t-Student, Mann–Whitney U-test, Pearson's correlation coefficient, and AUC for evaluating diagnostic capacity.

Dataset:

The study used OCT data from 59 individuals (17 with BD and 42 controls), with retinal thickness measurements across various retinal layers and regions.

Pros:

- ❑ Non-invasive and cost-effective diagnostic approach.
- ❑ AI classifiers achieved high accuracy
- ❑ OCT provides a safe and accessible biomarker

Cons:

- ❑ The study focused only on type I BD
- ❑ Lack of longitudinal data
- ❑ Small sample size, limiting the generalizability of the results.

Results:

The study found retinal thinning in bipolar disorder (BD) patients, especially in the ganglion cell and inner plexiform layers, with strong diagnostic potential (AUC > 0.75). AI classifiers performed well, with the linear Support Vector Machine (SVM) achieving the highest classification accuracy of 95%.

Conclusion:

The study suggests that structural retinal changes, as detected by OCT, can serve as a potential biomarker for BD. AI methods can improve diagnostic accuracy, but further studies with larger, multicenter datasets are needed to confirm these findings

Literature Survey

Title: A Novel Application for the Efficient and Accessible Diagnosis of ADHD Using Machine Learning

Link: https://ieeexplore.ieee.org/document/9311012

Methodology:

The system uses pupillometry to track changes in pupil size during a memory task as an indicator of ADHD. Several machine learning algorithms were applied, with an ensemble voting model providing the best performance. Data was validated using Leave-One-Out Cross-Validation (LOOCV).

Dataset:

The study involved 50 participants (28 with ADHD and 22 controls), aged 10-12. Pupillometric data was collected over 160 trials per subject during a visuospatial memory task.

Pros:

- High accuracy (sensitivity: 0.821, AUROC: 0.856)
- Objective and non-invasive
- Accessible, especially in resource-limited areas

Cons:

- Small dataset, possibly limiting generalizability
- Reliance on pupil dynamics alone might overlook other ADHD symptoms

Results:

The Ensemble Voting Model achieved 82.1% sensitivity, 72.7% specificity, and 85.6% AUROC, outperforming other models like Random Forest and SVM.

Conclusion:

This ADHD diagnostic system is better because it uses pupillometry for an objective diagnosis, offering more accuracy, affordability, and accessibility than traditional methods. Its real-time web app makes it a quick and reliable solution, particularly helpful in areas with limited healthcare

Literature Survey

Title: Neuro Intel: A System for Clinical Diagnosis of Attention Deficit Hyperactivity disorder (ADHD) using Artificial Intelligence

Link: https://ieeexplore.ieee.org/document/10218313

Methodology:

The Neuro Intel system uses a hybrid AI model combining machine learning (decision trees) and knowledge-based rules for ADHD diagnosis. When both models agree, the case is diagnosed automatically, and complex cases are referred to experts

Dataset:

Clinical data from an NHS mental health provider using tools like PHQ-9, GAD-7, and CAARS ADHD scores.

Pros:

- 92% accuracy in straightforward cases
- Combines AI and expert knowledge
- Reduces clinician workload

Cons:

- Limited dataset impacts generalization
- Complex cases still need expert input

Results:

The hybrid model achieved 92% accuracy, automating diagnosis for half the cases and referring complex ones to clinicians.

Conclusion:

Neuro Intel offers a highly accurate and efficient ADHD diagnostic tool by blending AI and expert knowledge, speeding up diagnosis while maintaining trust with explainable AI

Literature Survey

Title: Machine Learning in ADHD and Depression Mental Health Diagnosis: A Survey
Link:https://ieeexplore.ieee.org/document/10214293

Methodology:

The paper explores machine learning models like Support Vector Machines (SVM) and Neural Networks for ADHD and depression diagnosis. It uses wearable (EEG) and non-wearable (MRI, speech, and video) modalities to train models. These data sources help differentiate between control subjects and patients with ADHD or depression.

Dataset:

The survey discusses multiple datasets, including the ADHD-200 dataset, with 973 participants, and the DAIC-WOZ dataset for depression diagnosis

Pros:

- ❑ High accuracy in ADHD (97.6%) and depression (90%)
- ❑ Use of multi-modal data improves diagnosis reliability

Cons:

- ❑ Small datasets limit generalizability
- ❑ Data privacy concerns restrict access

Results:

SVM and Neural Networks showed high accuracy in detecting ADHD and depression. SVM models were particularly effective for EEG and MRI data, achieving up to 97.6% accuracy for ADHD detection

Conclusion:

Machine learning models, especially SVMs and Neural Networks, effectively diagnose ADHD and depression using multi-modal data like EEG, MRI, speech, and video. These methods outperform traditional approaches in accuracy, offering non-invasive and objective diagnostics. However, limited datasets and privacy concerns hinder broader applications.

Literature Survey

Title: Machine Learning based Detection of Post Traumatic Stress Disorder of Mental Health
Link:https://ieeexplore.ieee.org/document/10193036

Methodology:

The system uses various machine learning algorithms such as Random Forest, Gradient-Boosted Decision Trees, and Logistic Regression to predict Post-Traumatic Stress Disorder (PTSD). The data includes both social media posts (Twitter) and clinical data, with algorithms analyzing language patterns and psychological metrics

Dataset:

The study uses data from over 243,000 tweets from PTSD patients, as well as clinical datasets from military personnel and trauma-exposed individuals

Pros:

- ❑ High accuracy (97% for Random Forest)
- ❑ Utilizes real-time social media and clinical data
- ❑ Effective for early intervention in PTSD case

Cons:

- ❑ Sensitive to data quality
- ❑ Limited by dataset size for generalization across populations

Results:

The Random Forest algorithm showed the highest accuracy at 97%, outperforming other methods like Bagging (95%) and SVM (91%).

Conclusion:

The system demonstrates how machine learning, particularly Random Forest, can accurately detect PTSD using a combination of social media and clinical data. This approach offers high accuracy, making it a valuable tool for early detection and intervention in PTSD cases

Literature Survey

Title: Artificial-Intelligence based Prediction of Post-Traumatic Stress Disorder (PTSD) using EEG reports

Link: https://ieeexplore.ieee.org/document/10072671

Methodology:

This system uses EEG (Electroencephalography) reports to predict Post-Traumatic Stress Disorder (PTSD). The method involves collecting EEG data, pre-processing it for noise reduction, and extracting relevant features. A Convolutional Neural Network (CNN) model is then trained on this EEG data to classify individuals as either having PTSD or not

Dataset:

The dataset consists of EEG recordings from PTSD patients. EEG data is used to track brain activity, and specific channels from the EEG report are used for classification into PTSD or non-PTSD groups.

Pros:

- ❏ High accuracy (85%)
- ❏ Non-invasive (EEG-based)
- ❏ Early PTSD detection

Cons:

- ❏ Data-dependent
- ❏ Requires specialized EEG equipment

Results:

The CNN model used in this system showed high accuracy in predicting PTSD based on EEG reports. The performance of the model was validated using cross-validation, ensuring that the model generalized well across different datasets

Conclusion:

This system offers a highly accurate, non-invasive, and early diagnostic tool for PTSD using EEG data. By applying machine learning techniques such as CNN, it significantly improves upon traditional, subjective methods of PTSD diagnosis, offering a more reliable approach

Literature Survey

Title: Stress Management System

Link:https://ieeexplore.ieee.org/abstract/document/9753121?casa_token=UwQyaR-k8mAAAAAA:RqRjpBFMTRw2h6KJoRDd_cLuBCUuXGZVdfCGa_J0llu7rVXF0Ig6gXe8EqICvuAGiWVOsnsK3EBcK-4.

Methodology Used in the Existing System

- **Data Preprocessing**: Data cleaning, handling missing values, normalization
- **Algorithms used**: KNN, Random Forest, Logistic Regression, Decision Tree, Bagging, Boosting
- **Chatbot Implementation**: Tokenization, stemming, SoftMax for real-time advice

Dataset Used

- **OSMI Dataset**: Survey responses on mental health from working individuals

Pros:

- Boosting has the highest accuracy (81.7%)
- NLP Chatbot provides real-time interaction with the users along with feedback

Cons:

- Bagging has the lowest accuracy (77.5%)
- Chatbot needs continuous data updates

Results Obtained

Boosting (81.7% accuracy) performed better than Bagging (77.5%) and Random Forest (81.2%)

Literature Survey

Title: Sentiment Analysis of Depression and Anxiety Social Media Tweets Using TF-IDF Weighting and Supervised Learning Algorithm
Link:https://ieeexplore.ieee.org/document/10687916

Methodology Used in the Existing System

- **Pre-processing**: Tokenization, lemmatization, and stopword removal for text cleaning.
- **TF-IDF Weighting**: Highlights important words based on their frequency.
- **Model**: Supervised learning using convolutional layers and softmax activation for classification.

Dataset

- Analyzed 68,228 tweets from the Twitter API related to depression and anxiety.

Pros

- High accuracy (98.01%) and F1-Score (98%).
- Outperformed Naïve Bayes and Random Forest.

Cons

- Limited to text: it may miss sarcasm
- Ethical concerns regarding data privacy..

Results

The model(supervised learning model based on convolutional neural networks (CNNs) with softmax activation for sentiment classification) performs better than Naïve Bayes and Random Forest due to its higher accuracy in sentiment analysis.

Literature Survey

Title: Artificial Intelligence based Anxiety Detection in Patients
Link: https://ieeexplore.ieee.org/document/9987310

Methodology Used in the Existing System

1. **Mental Health Prediction:**
   - **Preprocessing:** Cleaned and normalized data.
   - **Algorithms:** Decision Tree, Random Forest, Logistic Regression, KNN, Boosting
2. **Chatbot Development:**
   - **Preprocessing:** Used NLP techniques.
   - **Training:** PyTorch model on mental health queries.
3. **User Interface:**
   - Built with PyCharm and Streamlit for interaction and real-time predictions.

Datasets:

- **Prediction:** Mental Health in Tech (Kaggle, 1260 records).
- **Chatbot:** Mental health queries (Kaggle, 536 records).

Pros:

- Comprehensive anxiety detection.
- Engaging chatbot.

Cons:

- Limited dataset size. and basic NLP techniques.

Results:

- The system accurately classified users' mental health conditions and provided effective responses to user queries

Literature Survey

Title: Mind Relaxation Chatbot for University Students by Using Dense Neural Network
Link:https://ieeexplore.ieee.org/document/9664678

Methodology Used in the Existing System

- A chatbot designed for university students' mind relaxation utilizes NLP and a Dense Neural Network (DNN).
- Trained on a therapeutic dataset organized by intents and responses.

Dataset Used

- A CSV dataset of therapeutic responses, focusing on key attributes such as questionId and answerText.

Pros

- 24/7 Support
- Anonymity

Cons

- Limited Empathy
- Effectiveness relies on the quality of the training data.
- Struggles with variations in language like synonyms

Results Obtained

Achieved 99.99% accuracy in training, meaning it correctly identified user intents nearly all the time. However, if a user's input didn't match a known intent, the chatbot would provide a generic response.

https://ieeexplore.ieee.org/document/9664678

Literature Survey

Title: Exploring Machine Learning Algorithms for the Detection of Depression
Link:https://ieeexplore.ieee.org/document/10561447

Methodology Used in the Existing System

- The system employs Support Vector Machine (SVM) and Long Short-Term Memory (LSTM) models to detect depression in tweets.
- Tweets are collected from Twitter and Kaggle, cleaned of retweets and non-English content, then transformed into numerical features using TF-IDF

Dataset Used

- The dataset consists of 15,000 tweets, labeled by multiple annotators for accuracy.

Pros

- Identifies depression promptly.
- SVM offers clear decision-making.
- LSTM understands language context.

Cons

- SVM needs careful tuning.
- Social media data collection raises privacy concerns.

Results Obtained

- **SVM:** 81.79% accuracy and an F1 score of 81.27%.
- **LSTM:** 70% accuracy, with other models around 79-80%.

Literature Survey

Title: New breakthroughs in AI Chatbots and their potential in mental health services

Url:https://ieeexplore.ieee.org/document/10612516

Methodology Used in the Existing System

- Chatbots engage with users via text or voice to offer mental health support.
- Chatbots continuously gather information from users to personalize their care, using techniques like cognitive-behavioral therapy (CBT).

Dataset Used

- Data comes from mental health chatbot interactions, especially during the COVID-19 pandemic, targeting issues like depression, anxiety, and stress.Continuous data collection enables personalized mental health support.

Pros

- Provides support anytime, anywhere.
- Tailors care based on user data.
- Offers non-judgmental, anonymous support.

Cons

- Issues with data security and ethics.
- Not as effective in complex cases without human intervention.

Results Obtained

- Chatbots are useful for managing depression, anxiety, and stress.Help people who face stigma, low literacy, or communication difficulties.Chatbots improve emotional well-being and support self-care.

Conclusion:

- It highlights the benefits of chatbots, such as providing accessible and anonymous support, which can help overcome traditional barriers to therapy. The continuous data collection capabilities of chatbots enable personalized treatment approaches, improving diagnostic accuracy.

Literature Survey

Title: From Brain Waves to Diagnoses: AI's Role in Schizophrenia Detection

url:https://ieeexplore.ieee.org/document/10576096

Methodology Used in the Existing System

- Brain activity is recorded through EEG to detect abnormal patterns linked to schizophrenia.
- AI models like CNNs (Convolutional Neural Networks) analyze these brain patterns to identify schizophrenia.

Dataset Used

- EEG and fMRI data from people with and without schizophrenia.EEG and fMRI data from people with and without schizophrenia.

Pros

- AI models are very accurate at detecting schizophrenia.
- AI speeds up the brainwave analysis process.
- EEG is a non-invasive way to monitor brain activity.

Cons

- Issues with data security and ethics.
- Not as effective in complex cases without human intervention.

Results Obtained

- AI models, particularly CNNs, can detect schizophrenia with high accuracy. Some models have achieved accuracy rates as high as 98%.

Conclusion

- AI, especially deep learning, is very good at detecting schizophrenia by analyzing brain waves using EEG and fMRI. It offers a fast, non-invasive way to find patterns that might be missed by traditional methods. However, there are challenges like protecting patient data, making the AI easier for doctors to understand, and ensuring it works well for everyone.

Literature Survey

Title: AI-Powered Mental Health Diagnosis: A Comprehensive Exploration of Machine and Deep Learning Techniques

url:https://ieeexplore.ieee.org/document/10515610

Methodology Used in the Existing System

- Reddit dataset from mental health-related subreddits. Posts are cleaned, tokenized, and split into training (80%) and testing (20%) sets.
- CNN, random forest, support vector classifier, and decision tree are trained to classify mental health conditions.Uses layers like embedding, convolution, max-pooling, dense layers, and SoftMax for classification.

Dataset Used

- The Reddit dataset used in the study consists of posts from six mental health-related subreddits: r/depression, r/anxiety, r/bipolar, r/BPD, r/schizophrenia, and r/autism

Pros

- Reddit provides large, diverse data relevant to mental health discussions.
- CNN, random forest, support vector, and decision tree classifiers offer flexibility in selecting the best model.

Cons

- CNN architecture is computationally intensive, requiring significant resources.
- The model may struggle to generalize to data outside Reddit.

Results Obtained

- multi-class classification model using four subreddit (anxiety, depression, bipolar, schizophrenia) out of the six subreddits us. It is observed that the multi-class classification CNN model performs moderately with an accuracy of 81.5%.

Conclusion

- Performance metrics, such as accuracy, were used to evaluate the models. The random forest model achieved the highest accuracy of 0.907, outperforming the other models.This suggests that random forest is the most effective for classifying mental health conditions in this study.

Literature Survey

Title: Machine Learning Approaches for Obsessive Compulsive Disorder Detection

url:https://www.researchgate.net/publication/374719983_Machine_Learning_Approaches_for_Obsessive_Compulsive_Disorder_Detection

Methodology Used in the Existing System

- A statistical method used for binary classification that estimates the probability of a certain class based on input features.
- An algorithm that classifies data points based on the closest training examples in the feature space..

Dataset Used

- Electroencephalogram (EEG), Magnetic Resonance Imaging (MRI), Functional MRI (fMRI), and Diffusion Tensor Imaging (DTI)data from people with and without schizophrenia.

Pros

- Machine learning models can help identify OCD earlier,
- potentially leading to better treatment outcomes.
- Using biomarkers provides a more objective basis for diagnosis compared to traditional symptom analysis.

Cons

- Current models often struggle to achieve high prediction accuracy,
- can be complex and may require substantial computational resources and expertise..

Results Obtained

- The result of all four machine learning approaches is compared between each other, out of this the SVM technique gives the batter accuracy heaving with 83%. The limitation of this is they do not identify with OCD or healthy control.

Conclusion

- The research concludes that while machine learning approaches show promise in enhancing OCD detection through laboratory analysis of biomarkers, challenges remain in achieving satisfactory prediction accuracy.

Literature Survey

Title: A comprehensive review for machine learning on neuroimaging in obsessive-compulsive disorder
url:https://pmc.ncbi.nlm.nih.gov/articles/PMC10646310/pdf/fnhum-17-1280512.pdf

Methodology Used in the Existing System

- The review synthesizes findings from various studies that utilize different machine learning techniques, such as artificial neural networks and other AI methods, to interpret neuroimaging data, specifically focusing on MRI and EEG scans.

Dataset Used

- Primarily MRI and EEG scans. Utilized tools like the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) for assessing OCD severity.

Pros

- Machine learning can identify neurobiological markers associated with OCD.
- Tailoring treatment based on individual neurobiological profiles.

Cons

- Small or biased datasets can lead to inaccurate predictions.
- Results from machine learning models may be difficult for clinicians to interpret..

Results Obtained

- In distinguishing between healthy people and OCD, the best accuracy rate of the model is 71.64%; in distinguishing between healthy people and low level OCD, the best accuracy rate of the model is 77.12%; in distinguishing between healthy people and high level OCD, the best accuracy rate of the model is 71.86%; in distinguishing low level OCD

Conclusion

- The review highlights the potential of integrating machine learning with neuroimaging to advance understanding and treatment of OCD. It calls for further research with larger datasets to validate findings and improve clinical applications.

Comparison of the existing methods

Method	Pros	Cons
Supervised Learning (Logistic Regression, SVM)	- Early detection of disorders - High efficiency and accuracy (70-98%) - Scalability for large datasets	- Limited generalizability - Dropout and engagement challenges - Privacy concerns
Unsupervised Learning (Clustering)	- Explores hidden patterns in unlabeled data	- May not provide direct actionable insights - Requires extensive data preprocessing
Neural Networks	- Handles complex tasks (e.g., neuroimaging) - High accuracy in predictive tasks	- Requires large datasets - Computationally intensive

Comparison of the existing methods

Transformers (e.g., RoBERTa)	- Effective for text analysis on social media - Can detect disordered behaviors from language patterns	- Needs large training datasets - Computationally expensive
AI Chatbots (NLP, ML)	- Improved accessibility to healthcare - Convenient for users - Reduces clinician workload	- Limited scope and accuracy issues - Lack of voice interaction capabilities
Hybrid AI Models (Decision Trees + Rules)	- High accuracy in straightforward cases (92%) - Combines AI and expert input for complex cases	- Limited dataset impacts generalization - Complex cases still require expert input
EEG Data Analysis (CNN)	- Non-invasive and high accuracy (85%) - Early detection of PTSD	- Data-dependent on quality - Requires specialized equipment

Comparison of the existing methods

Random Forest for PTSD Detection	- High accuracy (97%) using social media data - Effective for early intervention	- Sensitive to data quality - Limited by dataset size for generalization
Boosting Algorithms	- Highest accuracy among traditional methods (81.7%) - Real-time interaction via chatbots	- Continuous data updates needed for chatbot effectiveness - Bagging has lower accuracy compared to boosting

Base Paper

Title: New breakthroughs in AI Chatbots and their potential in mental health services

Url:https://ieeexplore.ieee.org/document/10612516

Methodology Used in the Existing System

- Chatbots engage with users via text or voice to offer mental health support.
- Chatbots continuously gather information from users to personalize their care, using techniques like cognitive-behavioral therapy (CBT).

Dataset Used

- Data comes from mental health chatbot interactions, especially during the COVID-19 pandemic, targeting issues like depression, anxiety, and stress.Continuous data collection enables personalized mental health support.

Pros

- Provides support anytime, anywhere.
- Tailors care based on user data.
- Offers non-judgmental, anonymous support.

Cons

- Issues with data security and ethics.
- Not as effective in complex cases without human intervention.

Results Obtained

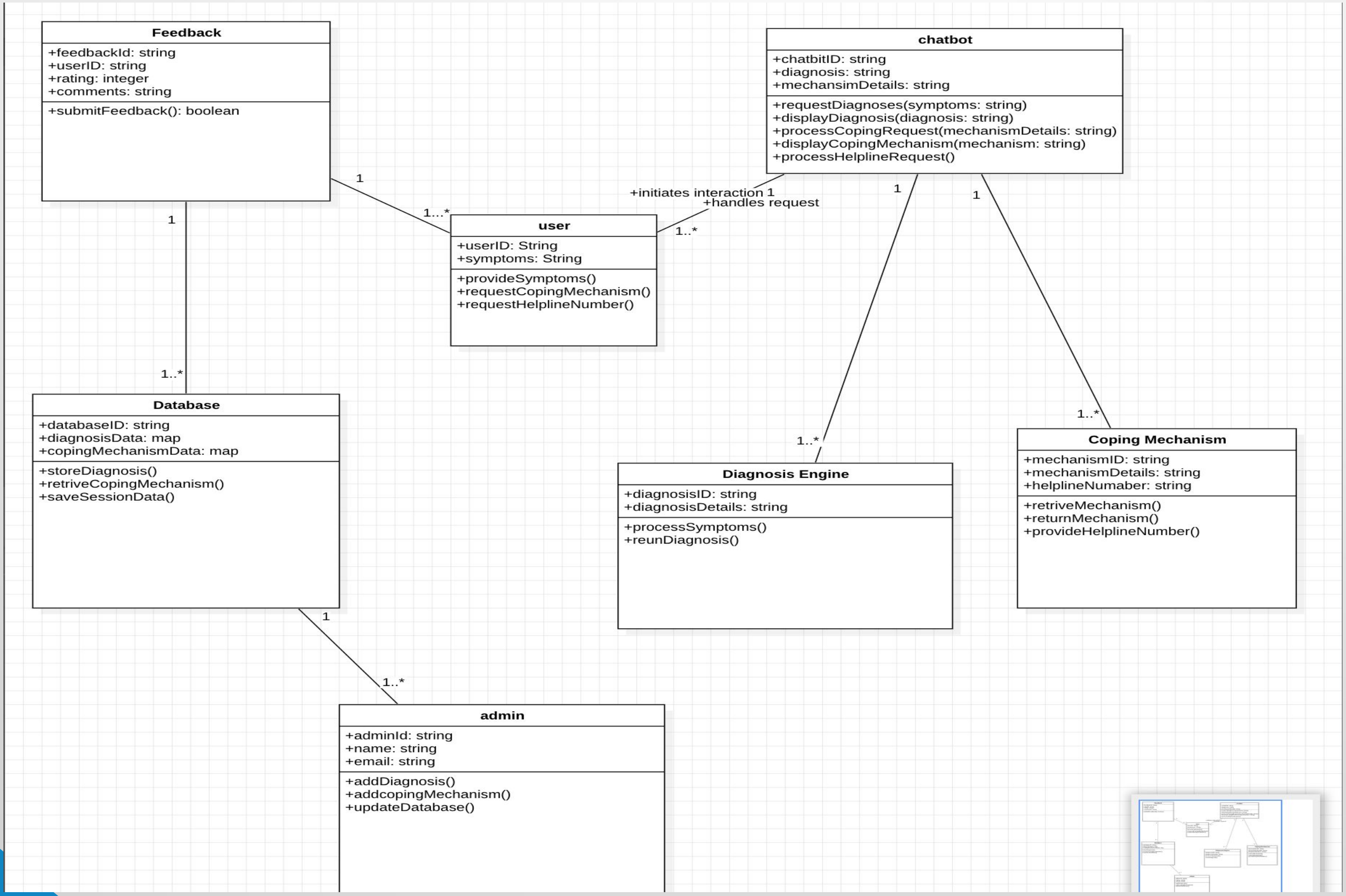
- Chatbots are useful for managing depression, anxiety, and stress.Help people who face stigma, low literacy, or communication difficulties.Chatbots improve emotional well-being and support self-care.

Conclusion:

- It highlights the benefits of chatbots, such as providing accessible and anonymous support, which can help overcome traditional barriers to therapy. The continuous data collection capabilities of chatbots enable personalized treatment approaches, improving diagnostic accuracy.

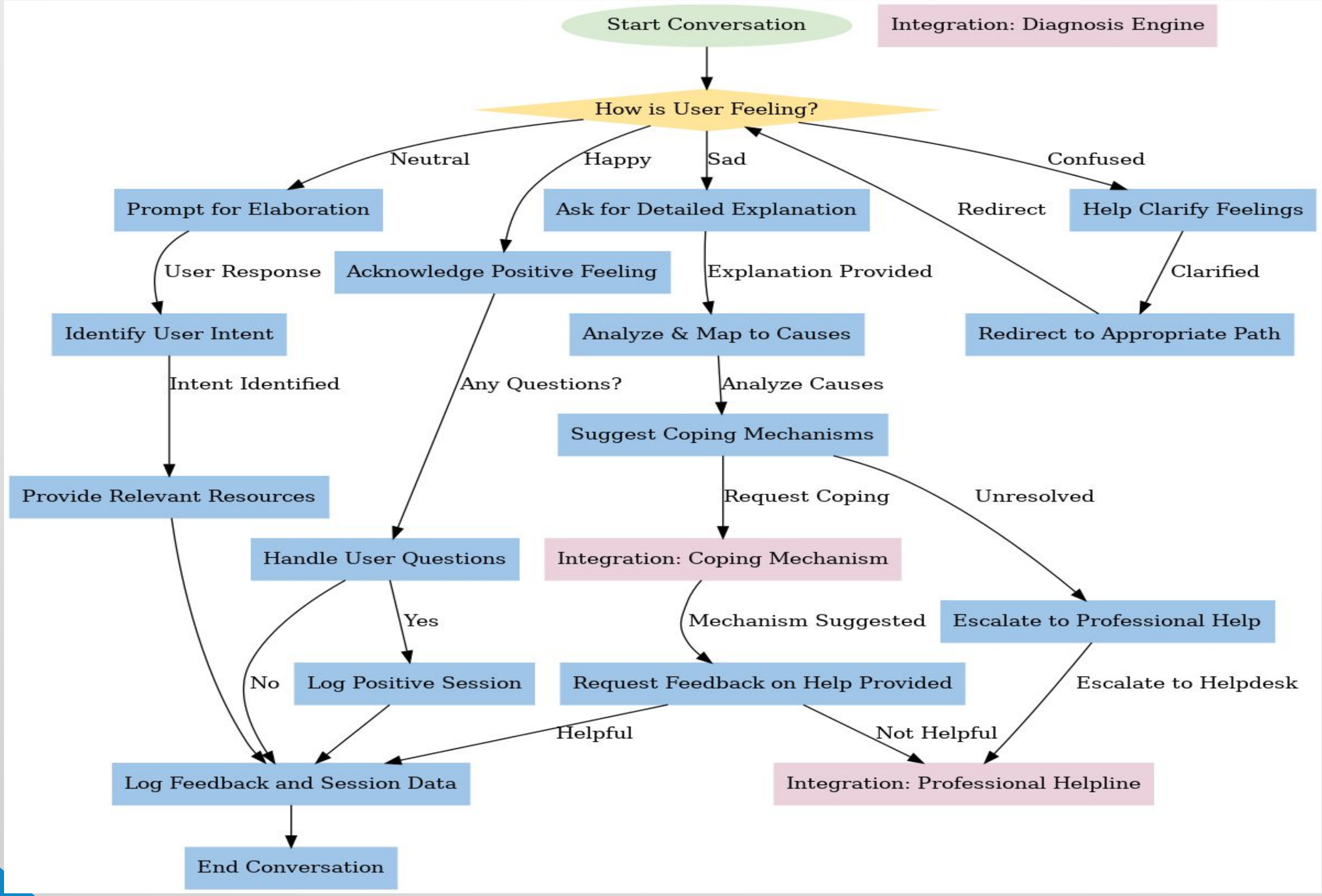
HIGH LEVEL DESIGN

1. CLASS MODEL



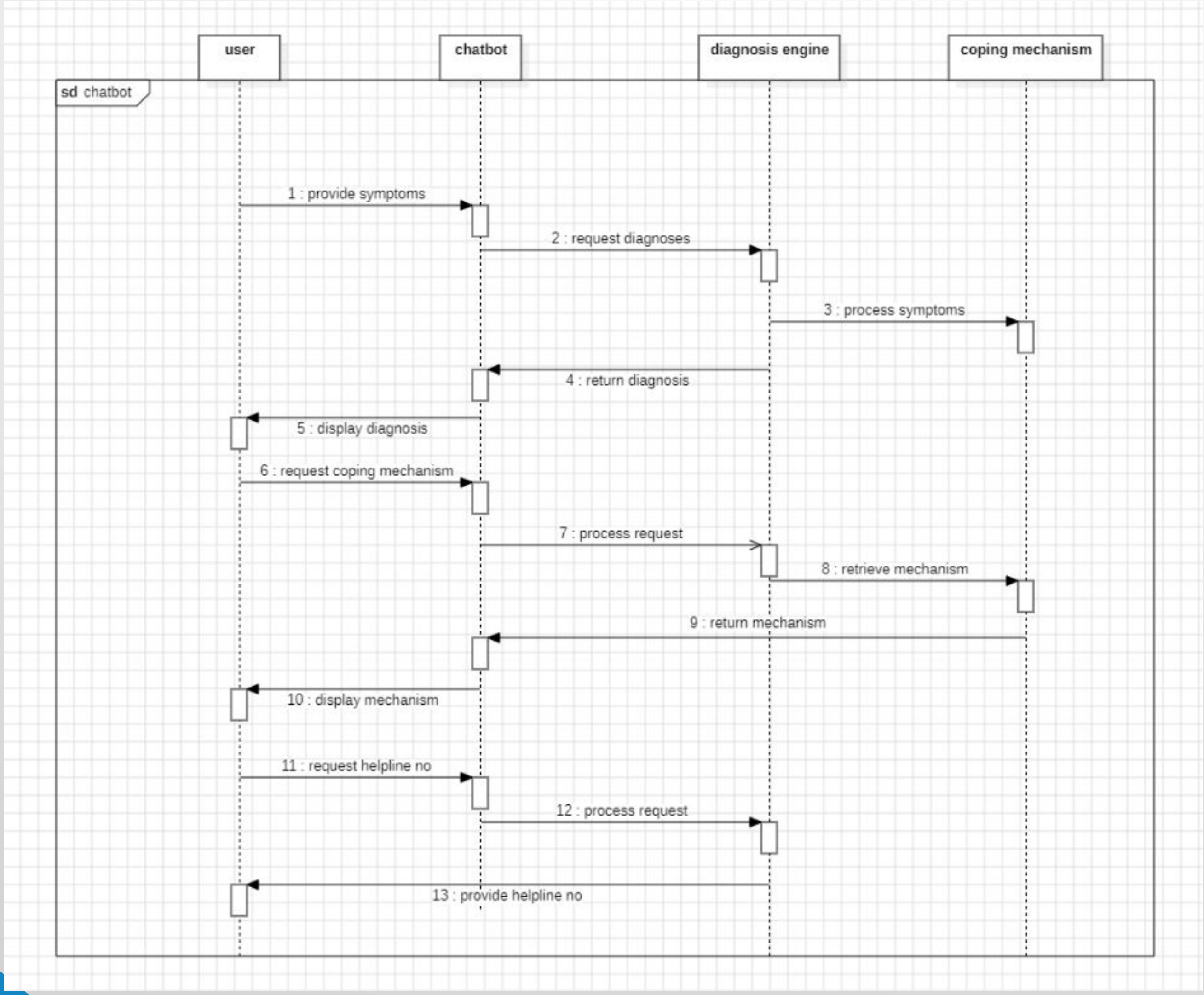
HIGH LEVEL DESIGN

2. ACTIVITY MODEL



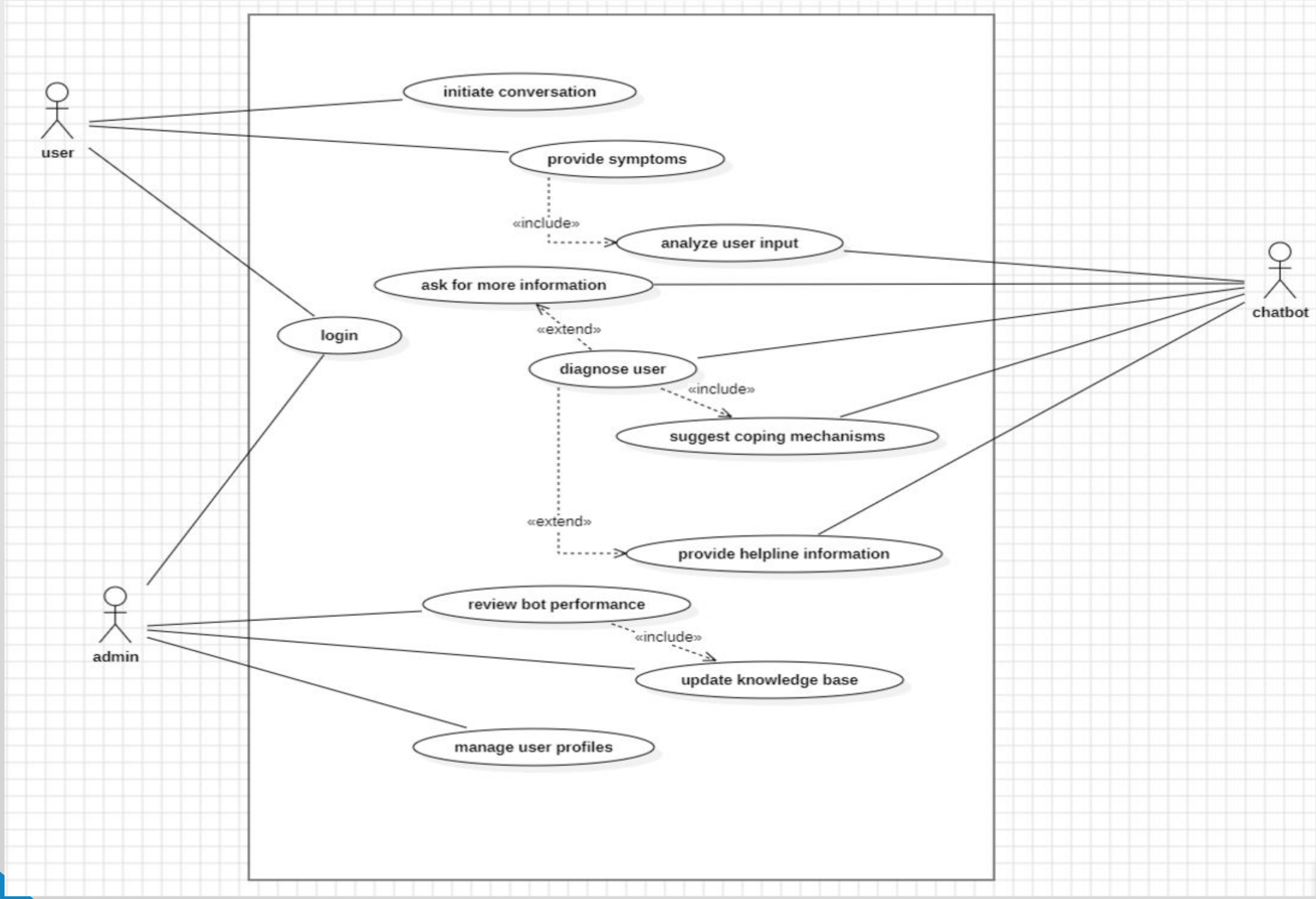
HIGH LEVEL DESIGN

3. SEQUENCE MODEL



HIGH LEVEL DESIGN

4. USE CASE MODEL



References

[1] Chakraborty, C., Gupta, A., Ghosh, S.K., and Zaman, S., "Machine Learning at the Core of Image and Video Analysis in the Healthcare Industry: An Outlook Beyond the COVID-19 Pandemic", SSRN, Page no. 1-25, 2022.

[2] Mahapatra, A., Sharma, P., and Tyagi, R., "Exploring Disordered Eating Behaviors through Social Media: A Machine Learning Perspective", Journal of Eating Disorders, Volume 10, Page no. 1-12, 2022.

[3] Amrani, A., Elharti, E., Zaki, S., and Zina, F., "Machine Learning Approaches for the Prediction of Depression: A Systematic Review", Frontiers in Psychiatry, Volume 15, Article 1414439, 2024.

[4] Fong, T., Li, H., and Sihota, M., "The Role of Artificial Intelligence in Cancer Diagnosis and Treatment: A Review of the Current State and Future Perspectives," Cancers, Volume 11, Page no. 803, 2021.

[5] Wang, Y., Zhang, Z., and Liu, C., "Artificial Intelligence in Health Care: Anticipating and Mitigating the Challenges", Health Services Research and Managerial Epidemiology, Volume 8, Page no. 1-9, 2021.

References

[6]Routhu, R. D., Likhita, C. K., Kranthi, G. K., & Lakshmi, A. (2022). Stress Management System. In 2022 IEEE Delhi Section Conference (DELCON) (pp. 1-5). IEEE. DOI: 10.1109/DELCON54057.2022.9753121.

[7]Nema, N. K., Shukla, V., Pimpalkar, A., & Tandan, S. R. (2024). Sentiment Analysis of Depression and Anxiety Social Media Tweets Using TF-IDF Weighting and Supervised Learning Algorithm. In 2024 OPJU International Technology Conference (OTCON) (pp. 1-6). IEEE. DOI: 10.1109/OTCON60325.2024.10687916.

[8]Koneru, L. C., Dedeepya, R. R., Lakshmi, A., & Kumar, G. K. (2022). Artificial Intelligence based Anxiety Detection in Patients. In 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (pp. 1-9). IEEE. DOI: 10.1109/I-SMAC55078.2022.9987310.

[9]Bopage, H., & Weerakoon, C. (2021). Mind Relaxation Chatbot for University Students by Using Dense Neural Network. In 2021 5th SLAAI International Conference on Artificial Intelligence (pp. 1-7).

[10]Bala Krishna, N., Bhanu, M. I., Anjali, R., Yaswanth Kumar, R., Omprakash Reddy, P., & Reethu Reddy, P. (2024). Exploring Machine Learning Algorithms for the Detection of Depression. In 2024 International Conference on Smart Systems for Applications in Electrical Sciences (ICSSES) (pp. 1-6). IEEE. DOI: 10.1109/ICSSES62373.2024.10561447.

References

[11]A. N. Vaidyam, H. Wisniewski, J. D. Halamka, M. S. Kashavan and J. B. Torous, "Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape", Can J Psychiatry, vol. 64, no. 7, pp. 456-464, Jul. 2019.

[12]C. L. Alves, A. M. Pineda, K. Roster, C. Thielemann and F. A. Rodrigues, "EEG functional connectivity and deep learning for automatic diagnosis of brain disorders: Alzheimer's disease and schizophrenia", Journal of Physics: complexity, vol. 3, no. 2, pp. 025001, 2022.

[13]H. Kour, "An hybrid deep learning approach for depression prediction from user tweets using feature-rich CNN and bi-directional LSTM", Multimedia Tools and Applications, vol. 81, no. 17, pp. 23649-23685, 2022.

[14]S. K. Kar, B. Kumari, and A. Singh, “Multimodal augmen-tation approach with transcranial direct current stimu-(a) Prediction accuracy percentagebar plot.(b) Precision bar plot. (c) Recall bar plot.Figure 4.Figure 4. OCD detection accuracy, Precision and Recalllation in management of obsessive-compulsive disorder with depression and comorbid seizure disorder: A case report,” Indian Journal of Psychological Medicine, vol. 44, no. 6, pp. 621–623, 2022

[15]Abd-Alrazaq, A., Alhuwail, D., Schneider, J., Toro, C. T., Ahmed, A., Alzubaidi, M., et al. (2022). The performance of artificial intelligence-driven technologies in diagnosing mental disorders: an umbrella review. NPJ Digit. Med. 5:87. doi: 10.1038/ s41746-022-00631-8

References

[16]Shaif, M. M., Rakhra, M., Singh, D., & Singh, A. (2022). AI-Based Prediction of PTSD Using EEG Reports. *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*, 1073-1077. IEEE.

[17]Batsakis, S., Papadakis, E., Tachmazidis, I., Chen, T., Antoniou, G., & Adamou, M. (2023). Neuro Intel: A System for Clinical Diagnosis of ADHD Using AI. *2023 IEEE Conference on ICT Solutions for eHealth (ICTS4eHealth)*, 1-6. IEEE

[18]Makhdoomi, S. M., Rakhra, M., Singh, D., & Singh, A. (2022). Artificial-Intelligence based Prediction of Post-Traumatic Stress Disorder (PTSD) using EEG Reports. In *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)* (pp. 1-6). IEEE. DOI: 10.1109/IC3I56241.2022.10072671.

[19]Nagarajaiah, K., Krishnappa, M. H., & Asha, K. R. (2023). Machine Learning based Detection of Post Traumatic Stress Disorder of Mental Health. In *2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 1-5). IEEE. DOI: 10.1109/ICESC57686.2023.10193036.

[20]Nash, C., Nair, R., & Naqvi, S. M. (2023). Machine Learning in ADHD and Depression Mental Health Diagnosis: A Survey. *IEEE Access*, 11, 82345-82357. Received July 11, 2023; accepted July 31, 2023; published August 10, 2023; current version August 17, 2023. DOI: 10.1109/ACCESS.2023.3304236.

THANK YOU