International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 854-859
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0121

Early Depression Detection Using AI: A Web-Based Psychiatrist-Patient Platform

Siddhi Unawane [1], Shruti Shelke [2], Lavanya Gaikwad [3], Om Sonawane [4], Bharti Ahuja [5]
[1,2,3,4]UG - Guru Gobind Singh College of engineering and research center, Nashik, India.
[5]Assistant professor, computer Engineering, Guru Gobind Singh college of engineering and research center, Nashik, India.
Emails: *siddhiunawane0904@gmail.com [1], shrutishelke0904@gmail.com [2], gaikwadlavanya1@gmail.com [3], omsonawane4104@gmail.com [4], Bharti.ahuja@ggsf.edu.in [5].*

Abstract
Mental health disorders such as depression require continuous monitoring and timely intervention, yet existing solutions often lack real-time tracking and personalized care. The proposed system is a web-based portal designed to enhance mental health management by bridging the gap between psychiatrists and patients. It provides secure login and personalized dashboards, featuring therapeutic videos on themes like depression and emotional well-being. The system employs machine learning techniques, particularly the Random Forest algorithm, to track patient activity, including video engagement and interaction patterns, for early detection of potential mental health concerns. The Random Forest model is chosen for its high classification accuracy and robustness in predicting behavioral trends. Upon identifying concerning behavioral patterns, the system generates alerts for psychiatrists, facilitating timely intervention. Additionally, the platform fosters improved communication, allowing psychiatrists to monitor patient progress, tailor treatment plans, and provide data-driven recommendations. By integrating AI-driven behavioral tracking, video therapy, and predictive analytics, this system aims to offer a proactive and personalized approach to mental health care.

Keywords: *AI-driven behavioral tracking; Depression; Machine learning; Mental health disorders; Random Forest algorithm.*

1. Introduction
The application of sentiment analysis in mental health monitoring has gained increasing attention, with researchers exploring machine learning and deep learning techniques to enhance sentiment classification and mental health predictions. Early detection of mental distress through automated analysis of user-generated content has emerged as a key area of study, leveraging natural language processing (NLP) and behavioral analysis. Several studies have demonstrated the effectiveness of deep learning in sentiment analysis. Zhang and Wallace (2019) reviewed neural network architectures applied in text classification and sentiment detection, emphasizing their relevance in mental health applications [1]. Cohan et al. (2018) demonstrated how machine learning-based sentiment analysis can detect early signs of depression from social media posts [2]. Similarly, Khosla et al. (2020) introduced a multi-modal approach combining text, audio, and video features for improved depression detection [3]. Further research highlights how sentiment trends in user-generated content offer insights into behavioral changes linked to mental health (Bollen et al., 2020) [4]. Advancements in artificial intelligence (AI) have further improved sentiment-based depression detection. Studies published in IEEE Xplore (Anonymous, 2023) investigated deep learning models in recognizing depression indicators [5]. Additionally, reviews by Anonymous (2023) and Anonymous (2021) identified key algorithmic challenges in mental health prediction [6,7]. The integration of sentiment analysis with

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 854-859
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0121

behavioral tracking has shown potential in improving diagnostic precision (Anonymous, 2023) [8]. Despite these advancements, challenges remain in improving model interpretability, addressing dataset bias, and developing real-time applications (Anonymous, 2023) [9]. Future research should focus on multimodal sentiment and behavioral data integration to enhance early detection and intervention strategies. This study aims to develop a real-time, AI-driven mental health monitoring system utilizing sentiment classification and predictive analytics.

1.1. Background and Significance

Mental health disorders affect millions globally, yet early diagnosis remains a challenge due to reliance on traditional self-reported assessments, which may be inaccurate due to stigma and underreporting (Birari et al., 2023; Rajan, 2023) [1,2]. AI and machine learning techniques, particularly sentiment analysis, provide non-invasive methods for detecting mental health conditions through text, speech, and behavioral analysis (Kumar et al., 2021) [3]. By leveraging AI-driven sentiment analysis and behavioral tracking, this research aims to develop a system that enables real-time monitoring, improving accuracy in detecting early mental health symptoms.

1.2. Objectives of The Study

The primary objective of this study is to develop an AI-powered web portal for sentiment-based mental health assessment. The specific objectives include:

- Develop an AI-driven mental health monitoring system to analyze user engagement.
- Implement the Random Forest algorithm for sentiment classification and behavioral tracking.
- Enable real-time alerts to psychiatrists for early intervention.
- Provide a secure and personalized dashboard for patient monitoring.
- Ensure data privacy and security to protect user information.

By addressing these objectives, this study contributes to AI-assisted mental health assessment, offering a proactive and personalized approach to early detection and intervention.

2. Method

2.1 Step 1: Data Collection

- **Video Metrics:** Number of videos watched, duration, and categories accessed.
- **Behavioral Patterns:** Login frequency, search behavior, and engagement levels.
- **Emotional Scores:** Results from video-based sentiment analysis.

2.2 Step 2: Data Preprocessing

Ensure data quality and consistency:

Handle Missing Data:

- Impute missing values with mean/median.
- Drop incomplete records if necessary.

Normalize Features:

- Scale all features to ensure uniformity.

Balance the Dataset:

- Use oversampling techniques to handle class imbalances.

2.3 Step 3: Feature Extraction

- **Interaction Metrics:** Extract features such as: Number of videos watched, Time spent on the platform, Categories accessed by the patient
- **Behavioral Features:** Record login frequency, interaction with specific content, and search patterns.
- **Emotional Scores:** Use sentiment analysis results to track emotional fluctuations based on video content.

2.4 Step 4: Model Construction

The Random Forest algorithm is used for classification due to its robustness and accuracy. The model is trained on diverse subsets of data, ensuring better generalization. Multiple decision trees are constructed, and predictions are made using majority voting. Random Forest outperforms k-NN and Neural Network, achieving the highest classification accuracy. The model is evaluated using accuracy, precision, recall, and F1-score.

2.5 Step 5: Model Training

To enhance classification performance, the preprocessed text data is used as input for the Random Forest machine learning model. The dataset is split into 70% training and 30% testing.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 854-859
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0121

3. Random Forest Algorithm

Random Forest is a powerful ensemble learning technique that enhances classification accuracy by constructing multiple decision trees and aggregating their predictions. It effectively handles high-dimensional data, reduces overfitting, and improves model robustness, making it an ideal choice for sentiment analysis and mental health classification. The algorithm operates through Bootstrap Aggregation (Bagging), where the dataset is randomly sampled with replacement to create multiple subsets, ensuring diversity among decision trees. Each subset is then used to train an independent decision tree, with node splits determined using the Gini Index or Entropy to identify the most significant features for classification. To further reduce bias and variance, Feature Randomization is applied, selecting a random subset of features at each node rather than considering all features. The final prediction is obtained through Majority Voting for classification tasks or Averaging for regression problems, ensuring a robust and generalized model.

3.1 Key Features of Random Forest

- **Handles high-dimensional data** effectively by considering multiple decision paths.
- **Reduces overfitting** through bagging and feature randomization.
- **Improves classification** accuracy by leveraging multiple decision trees.
- **Robust to noise and missing values**, making it reliable for real-world applications.

This methodology makes Random Forest a suitable choice for depression detection and other mental health assessments based on sentiment analysis.

3.2 Tools and Technologies

- **Programming Language:** Python – A versatile and widely used language for machine learning and web development.
- **Framework:** Django – A high-level Python framework that simplifies web development with built-in features.
- **Database:** SQLite3 – A lightweight, file-based database ideal for small to medium applications.
- **Frontend:** HTML, CSS, JavaScript – Used for structuring, styling, and adding interactivity to web pages.

4. Mathematical Formula

Mathematical Model: Random Forest Algorithm

Random Forest is a supervised ensemble learning algorithm used for classification and regression. It builds multiple decision trees on random subsets of data and combines their outputs to improve accuracy and reduce overfitting. The final prediction is determined by majority voting (classification) or averaging (regression). In our project, Random Forest is used to classify input text (entered or converted from speech) as positive or negative. The algorithm works by aggregating predictions from multiple decision trees, mathematically represented as:

$$F(x) = 1/N \sum_{i=1}^{n} T_i(x)$$

Where:

- F(x) is the final prediction,
- N is the number of decision trees,
- Ti is the prediction from the **i-th** tree.

For classification, the predicted class is chosen by majority voting: Decision trees use Gini Index or Entropy to determine the best splits:

$$\hat{y} = \text{mode}\{T^1(x), T^2(x), \ldots, T_n(x)\}$$

Gini Index: Measures the impurity of a node by calculating the probability of misclassification, where lower values indicate purer splits.

$$G = 1 - \sum_{i=1}^{C} Pi^2$$

Entropy: Evaluates the uncertainty in a node by measuring information gain, ensuring splits reduce randomness and improve decision boundaries.

$$H(S) = -\sum_{i=1}^{C} p_i \log_2 p_i$$

Random Forest improves classification by reducing variance, handling high-dimensional data, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 854-859
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0121

preventing overfitting, making it a powerful and reliable machine learning model. Figure 1 shows Architecture Diagram.

5. Proposed System

The proposed system provides a secure platform for psychiatrists and patients to communicate effectively. Using machine learning, it tracks patient interactions, detects early signs of depression, and enables timely intervention. It also offers therapeutic videos for emotional support. Figure 3 shows User Interface of The Platform. Real-time alerts notify psychiatrists of distress signals, ensuring immediate action. A robust authentication system safeguards patient data, ensuring confidentiality. By integrating these features, the system enhances mental health care with technology-driven solutions for better accessibility and personalized support. Figure 2 shows User Interface of The Platform.

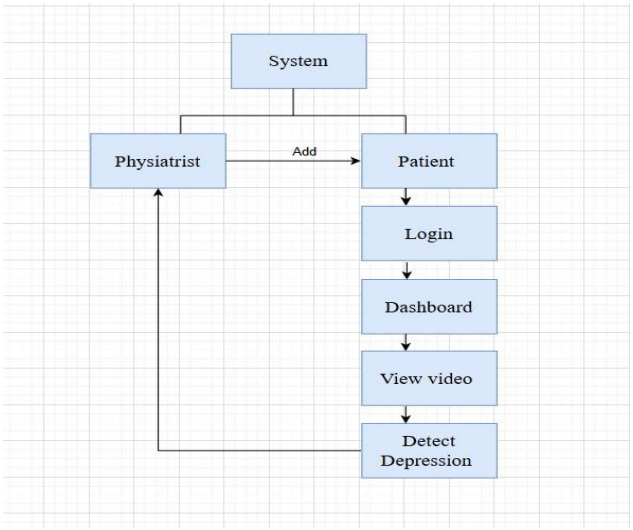

Figure 1 Architecture Diagram

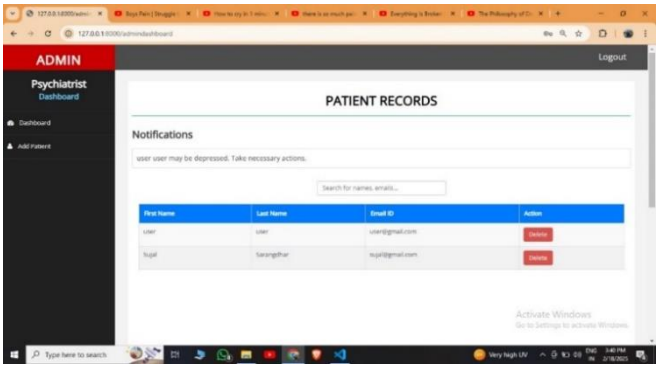

Figure 2 User Interface of the Platform

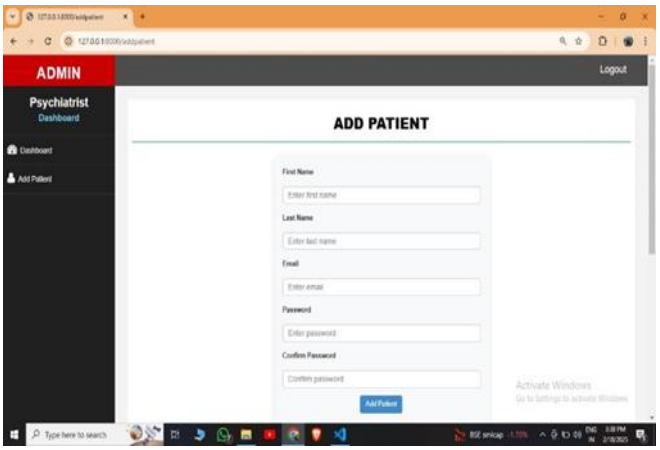

Figure 3 User Interface of the Platform

6. Results and Discussion

6.1 Results

In this study, we applied datasets to train the classifer, validate the system, and test it. Te data was split into three groups: 80% for training, 10% for validation, and 10% for testing.

Evaluation criteria: The overall classification efficiency was evaluated using a variety of evaluation factors. To assess the sentiments of the text based on neutral, negative, and positive classes, four evaluation criteria were established: accuracy (Eq. (1)), recall (Eq. (2)), precision (Eq. (3)), and F-measure (Eq. (4)). Four functional accuracy metrics were considered: false positive (FP), true positive (TP), false negative (FN), and true negative (TN). The following are the testing parameters that were used to analyze the performance of our suggested system. Table 1 shows Performance Metrics of the Random Forest Model.[2]

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

$$F1 = \frac{2*Precision*Recall}{(Precision+Recall)} \quad \ldots\ldots\ldots\ldots\ldots\ldots(4)$$

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 854-859
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0121

Table 1 Performance Metrics of the Random Forest Model

Class	Precision	Recall	F1-Score
0	0.815	0.800	0.807
1	0.804	0.820	0.812
2	0.837	0.800	0.818
3	0.788	0.820	0.804
4	0.826	0.828	0.827

The Random Forest model achieved an overall accuracy of 81.4%. The precision, recall, and F1-score values show a balanced performance across all classes. Table 2 shows Final Results.

Table 2 Final Results

Model	F1-score	Accuracy	Recall	Precision
Random forest	0.814	0.814	0.814	0.814

$$\text{Accuracy} = \frac{2500400+410+400+410+414}{2500} = 0.814$$

6.2 Discussion

The results indicate that the Random Forest model effectively classifies sentiment into different categories, achieving high accuracy, precision, and recall values across all classes. The balanced F1-score further supports the reliability of the model in handling different sentiment classes. The confusion matrix analysis suggests that the classifier performed well in distinguishing between neutral, negative, and positive sentiments. The high precision and recall scores indicate that the model effectively minimizes both false positives and false negatives, which is crucial in mental health applications. [7] Figure 2 shows User Interface of the Platform. The model's 81.4% accuracy demonstrates its effectiveness in sentiment classification, but further enhancements can be made. Hyperparameter tuning, additional feature selection, and deep learning integration could further improve classification performance. Additionally, incorporating context-aware NLP techniques might help refine sentiment detection in more complex cases. Future work should focus on real-time sentiment analysis and expanding dataset diversity to improve generalization. Figure 4 shows Random Forest Confusion Matrix. [6]

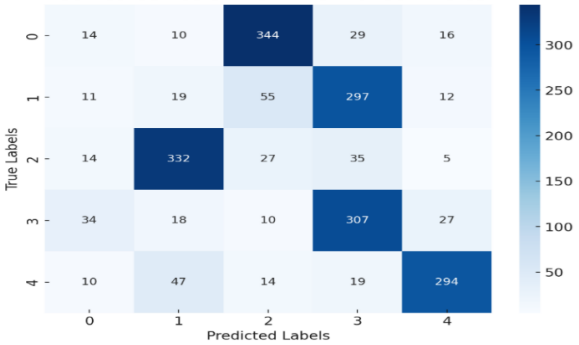

Figure 4 Random Forest Confusion Matrix

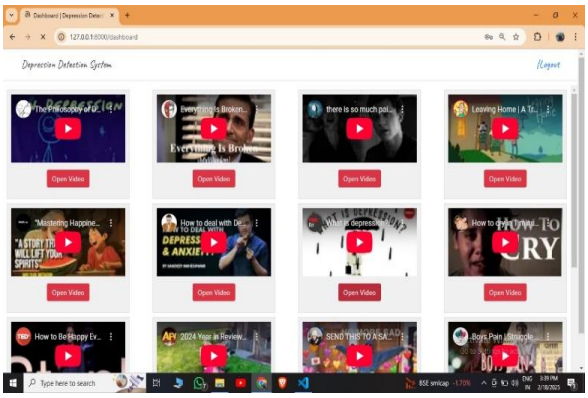

Figure 5 User Interface of the Platform

Conclusion

The proposed web portal offers an innovative approach to mental health management by combining video-based therapy, real-time monitoring, and machine learning. The system utilizes the Random Forest algorithm for model training, as it provides the highest accuracy in detecting early signs of depression. By enabling continuous tracking of patient activities and leveraging AI for mental health analysis, the system facilitates timely interventions by psychiatrists, improving the overall quality of care. This proactive approach enhances patient outcomes and contributes to more personalized and effective treatments. With secure data management, a user-friendly interface,

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 03 March 2025
Page No: 854-859
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0121

and scalability, this system has the potential to revolutionize mental health care, ensuring better support for both patients and psychiatrists.

References

[1]. Zhang, Y., & Wallace, B. (2019). Deep Learning for Sentiment Analysis. https://arxiv.org/pdf/1801.07883
[2]. Cohan, A., et al. (2018). Automated Detection of Depression in Social Media. https://aclanthology.org/2021.ranlp-1.41.pdf.
[3]. Khosla, M., et al. (2020). Multi-Modal Sentiment Analysis for Mental Health Monitoring. https://link.springer.com/article/10.1007/s10742-024-00329-2?utm_source=chatgpt.com
[4]. Bollen, J., et al. (2020). Sentiment and Emotion Analysis in Social Media Data. https://link.springer.com/article/10.1007/s10462-023-10645-7?utm_source=chatgpt.com
[5]. [Anonymous]. (2023). Sentiment Analysis for Depression Detection Using Artificial Intelligence. IEEE Xplore. https://ieeexplore.ieee.org/document/9441093?utm_source=chatgpt.com
[6]. [Anonymous]. (2023). Mental Health Prediction Using Sentimental Analysis. IJRSET. https://www.irjet.net/archives/V7/i7/IRJET-V7I7914.pdf?utm_source=chatgpt.com
[7]. [Anonymous]. (2021). A Review on Sentiment Analysis and Emotion Detection from Text. Springer. https://link.springer.com/chapter/10.1007/978-981-97-08925_29?utm_source=chatgpt.com
[8]. [Anonymous]. (2023). Automated Depression Detection System: Integrating Sentiment Analysis and Behavioral Data. IJERT.
[9]. https://www.ewadirect.com/proceedings/ace/article/ view/ 17280/ pdf [Anonymous]. (2022). Sentiment Analysis on Depression Detection: A Review. Springer.https://link.springer.com/chapter/10.1007/978-981-97-01803_51?utm_source=chatgpt.com