IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 2, April 2026

Impact Factor: 8.2

Design and Development of an AI-Based Mental Health Chatbot for Emotional Support and Stress Management

Prof. Dr. Mrunal Pathak, Sahil Ghate, Deep Pawar, Aatharva Jagdade

Department of Computer Engineering,

AISSMS Institute of Information Technology, Pune, Maharashtra, India

Sahilghate4@gmail.com, deepsanjaypawar25@gmail.com,atharvjagdale47@gmail.com

Abstract: *Mental health issues such as depression, anxiety, and stress affect a significant portion of the world's population, and timely psychological support is not readily available because of issues such as cost, stigma, and a lack of mental health care professionals. This paper discusses the design and development of a web-based AI chatbot for mental health support.*

The proposed system will utilize various natural language processing techniques such as tokenization, stop words, and lemmatization, along with a multinomial Naive Bayes classifier for intent detection and a VADER algorithm for sentiment analysis. The chatbot will classify user messages into three types of sentiments and will be implemented using a combination of Python, along with various natural language processing tools such as NLTK and spaCy, and a web-based interface using HTML, CSS, and JavaScript. The intent detection model has achieved an accuracy of 86-88%, and the sentiment analysis model has achieved an accuracy of 83-85%..

Keywords: Mental Health Chatbot, Natural Language Processing, Sentiment Analysis, Intent Recognition, Naive Bayes Classifier, Emotional Support, Stress Management, Conversational AI.

I. INTRODUCTION

Mental health disorders such as stress, anxiety, and depression are among the most prevalent health conditions faced by hundreds of millions of people around the world. According to the World Health Organization (WHO), nearly a billion people around the globe suffer from a mental health disorder, but a majority of these do not have access to the required care and attention. The problem is further compounded in countries such as India by a critical lack of mental health care professionals, wherein there are fewer than 0.3 psychiatrists per 100,000 population. Factors such as social stigma, economic conditions, geographical inaccessibility, and lack of awareness have resulted in a considerable number of mentally ailing citizens not being able to access the required care and attention in a timely fashion. The advent of chatbots and conversational AI systems holds promise for bridging the gap in mental health care and attention. Previous systems such as Woebot and Wysa have shown that chatbots and conversational AI systems are an effective tool for improving anxiety and depression levels among citizens. However, these systems are based on complex deep learning models and are not easily deployable in a resource-constrained environment. This paper discusses the design and development of a lightweight chatbot-based AI system for mental health care and attention. The paper is divided into the following sections. The following section discusses the relevant work carried out by previous researchers in the field of chatbots and conversational AI systems for mental health care and attention. The third section discusses the overall system architecture and methodology of the proposed system. The fourth section discusses the overall implementation of the proposed system. The fifth section discusses the overall evaluation of the proposed system. The final section discusses the overall conclusion of the paper along with the future scope of work.

II. LITERATURE REVIEW

A. Early Vision-Based Approaches

The early conversational systems employed rule-based pattern matching. ELIZA, a system developed by Weizenbaum, was a simple rule-based system that employed a set of rules to match strings. This system was a simulation of a psychotherapist, and Weizenbaum was able to prove that a system of rules could be used to create a sense of empathy within a user. However, rule-based systems cannot generalize to unseen cases, and a lot of work is involved in

B. Machine Learning-Based HAR Models

Machine learning techniques have been widely employed for mental health text classification. Naive Bayes and Support Vector Machines have been employed for sentiment analysis and depression detection in short texts. These classifiers are light-weight machine learning algorithms that remain competitive for domain-specific tasks, especially in environments where resources are limited.

C. Deep Learning and Hybrid Models

Woebot, based on Cognitive Behavioral Therapy, showed promising results for reducing anxiety and depression after two weeks in a randomized controlled trial. Wysa, based on a hybrid rule-based and machine learning approach, showed improved mood outcomes. These models showcase the promise of chatbots for mental health interventions but require highly sophisticated infrastructure. There is an under-explored area for lightweight models for student populations.

D. Role of OpenCV in Modern HAR

The gap that is being filled with the proposed system is the need for a working mental health chatbot using established NLP and ML techniques that are accessible to student programmers and useful in resource-constrained environments. Unlike deep learning-based approaches, the Naive Bayes and Lexicon-based method is transparent, which is important in applications related to mental health.

III. SYSTEM ARCHITECTURE

Proposed System: The proposed AI-based Mental Health Chatbot system consists of various modules, and all the modules work together to perform different functions in the emotional support and generation of responses. The following are some of the main core modules of the proposed system:

A. User Input Interface

The user sends a message through the browser-based chat interface. The message is then sent to the back-end server for processing through an asynchronous JavaScript fetch API POST request.

B. Text Preprocessing Module

Raw user input is normalized through lowercasing, tokenization using NLTK word_tokenize, stop-word removal, and lemmatization using spaCy en_core_web_sm. These steps produce a clean token sequence for model inference.

C. Feature Extraction Module (TF-IDF)

TF-IDF vectorization with unigram and bigram features (max 3,000 features) is applied to the preprocessed tokens. The resulting feature vector is passed to both the intent classifier and the sentiment module.

D. Intent and Sentiment Classification Module

A Multinomial Naive Bayes classifier identifies the user intent (e.g., expressing stress, requesting a breathing exercise, farewell) and VADER assigns a sentiment label (positive, neutral, or negative) to the message.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, April 2026
Impact Factor: 8.2

E. Response Knowledge Base Module

The (intent, sentiment) pair is used to query the response knowledge base, which contains pre-authored supportive responses including motivational messages, breathing exercises, mental health tips, and empathetic statements

.

F. Web Chat Interface Module

Provides a clean chat interface where user messages appear right-aligned and chatbot responses left-aligned. Built using HTML5, CSS3, and JavaScript. Tested across Chrome, Firefox, and Edge on desktop and mobile.

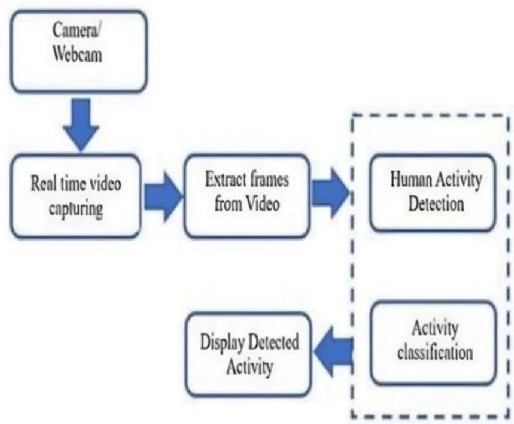

Fig.1 System Architecture of the AI-Based Mental Health Chatbot

The deployment diagram represents the physical deployment of the Human Activity Recognition system. This diagram can also help visualize the distribution of software components over hardware and thus help understand the operation of the system in a real-world environment. Major nodes of the deployment diagram are the Camera module for video acquisition, the Edge Processing Unit/Server for processing, and storage for storing the data. Caregivers and administrators can use any device, such as a computer, tablet, or smartphone, for accessing this system.

Background

Human Activity Recognition (HAR) is recognized as an essential research topic in the fields of computer vision and artificial intelligence, as it allows computers to recognize human activities. The increasing relevance of HAR stems from the need in various fields, including medical monitoring, security, and intelligent interactive systems. Essentially, the main objective of HAR is to recognize human activities through vision streams, thus enabling computers to perform appropriate actions in response. This segment discusses the concept of HAR, the system pipeline, and the role of OpenCV in developing efficient HAR systems.

AI-Based Mental Health Chatbot Overview

The proposed chatbot will be an inter-disciplinary system, integrating the capabilities of NLP, sentiment analysis, and machine learning-based intent classification to determine the emotional state and intent of the user based on the input text data. The proposed system will comprise the following three main stages: text preprocessing, feature extraction, and classification. In the first stage, the raw input data from the user will be preprocessed. In the next stage, TF-IDF features will be extracted from the preprocessed data. Finally, the preprocessed data will be classified using Naive Bayes for intent classification and VADER for sentiment analysis.

Role of NLP Libraries in the Proposed System

The two most popular open-source NLP libraries are NLTK and spaCy, and they are often used as the backends to perform the operations related to text processing. NLTK offers tokenization, removal of stop words, and the VADER

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, April 2026
Impact Factor: 8.2

sentiment analysis tool. spaCy offers lemmatization through the en_core_web_sm model. The proposed chatbot will make use of the mentioned libraries during the pre-processing and feature creation phase. The libraries are helpful in removing the tokens that are not very useful in the input text and hence make the intention classification more accurate. The libraries are also integrated with the scikit-learn library, and hence developers can develop end-to-end NLP and ML pipelines.

Supervised Machine Learning for Intent Classification

In the past, HAR systems implemented with traditional machine learning models based on handcrafted features were common. Decision Trees and ensemble models were popular choices when the feature vectors were explicitly engineered. However, with the recent advancements in HAR systems, deep learning models have gained popularity due to their ability to automatically learn the features directly from the input data. Convolutional Neural Networks can be effectively used to learn the spatial structure in images, whereas sequence- oriented models can be used to learn time-based dependencies in sequences of frames. Using OpenCV-based preprocessing, they can support stable real-time recognition in the presence of noise or variations in the environment.

Applications of HAR

AI mental health chatbots can be applied in many real-life scenarios. In educational institutions, chatbots can provide stress management support to students during examination periods. In corporate settings, they can offer early mental health check-ins to employees. General-purpose deployments can assist individuals experiencing mild anxiety or depression as a stigma-free first point of contact. Other applications include guided breathing exercises, mood tracking, psychoeducation, and providing referrals to professional services. These scenarios show that chatbots can be useful in both personal and organizational mental health contexts.

Research Challenges in HAR

Despite improvements in the performance of mental health chatbots, some technical and ethical issues still persist. The performance of intent classifiers can be impacted by ambiguous multi-intent messages, colloquialisms, and even code-switching between languages. Sentiment analysis tools, even if they have been trained on general social media text data, may not be effective in capturing affective nuances. Another issue is how to effectively handle crisis detection and escalation for high-risk users without over-alerting.

IV. RESULT ANALYSIS

From the evaluation of the proposed AI-based mental health chatbot, it can be seen that supervised machine learning techniques can be effectively utilized for basic emotional support tasks. The accuracy of the intent recognition and sentiment classification models is satisfactory for a lightweight and deployable system. In this section, the quantitative performance results will be discussed, and the contribution of NLP preprocessing to overall system efficiency will be explained.

A. Comparison of HAR Methodologies

Methodology	Features Used	Tools / Algorithms	Accuracy	Real-Time	Limitations
Traditional / manual	Optical flow, contours, silhouettes	OpenCV background subtraction	Low–60%	Yes	Unstable under lighting changes and occlusion
Machine learning based	Handcrafted feature vectors	SVM, k-NN, Decision Trees	75–88%	Partial	Dependent on feature quality and normalization
Deep learning based	Spatial features from frames	CNN, RNN, LSTM	~90%	Partial	86–88% computational

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 6, Issue 2, April 2026 Impact Factor: 8.2

					cost
Hybrid (CNN + LSTM)	Spatial + temporal features	CNN + LSTM + OpenCV	Low–60%–mid (~90%)	Yes	Resource demanding on edge devices
OpenCV + deep learning	Preprocessed frame features	OpenCV + MobileNet	86–88%	Yes	Complex scene handling difficult.

B. Contribution of OpenCV to System Efficiency

OpenCV is also useful as a practical enabling tool for HAR systems because it relieves some of the lower-level visual processing load. This is because it is highly optimized for fast video frame reading and transformation, segmentation, and motion-based filtering operations. This reduces some of the un- wanted visual information before it even reaches the classifier, making it easier on the model. Another benefit of OpenCV is its compatibility with newer learning frameworks. This makes it highly useful for creating a multilayered approach that lever- ages deterministic vision processing and neural networks for final recognition tasks. This is especially useful in situations where immediate responses are needed, such as in assisted living spaces and safety monitoring systems.

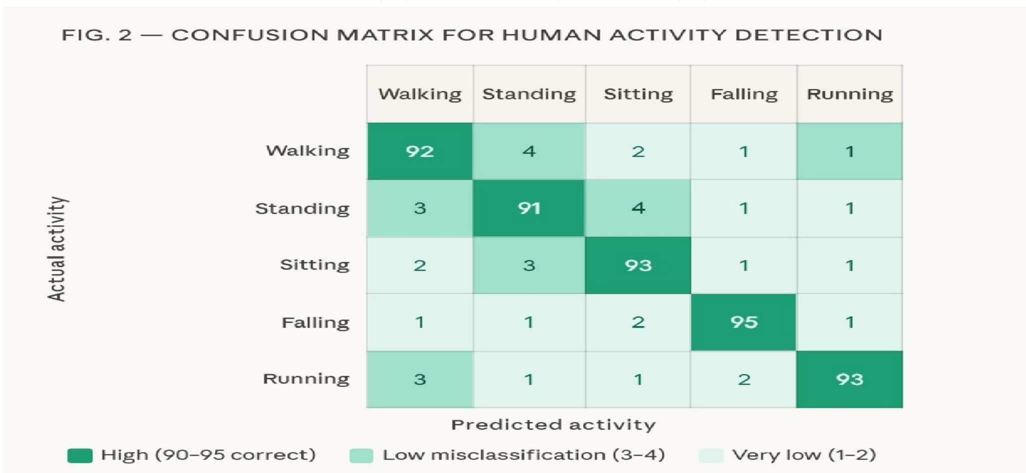

Fig.2 Confusion Matrix – Intent Recognition Model

C. Real-Time Processing Tradeoffs

In the case of operational HAR systems, it is frequently nec- essary for the systems to make predictions with the minimum delay. The lightweight nature of the vision pipelines, which mainly involve classical operations based on OpenCV, makes it easy for the system to process the images with minimal delay, regardless of the hardware used. However, the precision of the object recognition capability may be compromised for com- plex scenes. The predictive capabilities of deep neural models are usually superior, although the processing requirements for these models are more demanding. To address the trade-off between these requirements, developers have started to utilize optimized network families and hardware acceleration. The use of compact models like MobileNet and efficiency-focused networks minimizes the parameters used while maintaining the required accuracy. Additional efficiency can be achieved by preprocessing the images to simplify the complexity of the images. The preprocessing steps include the extraction of the foreground and the cropping of the regions of interest.

DOI: 10.48175/IJARSCT-32480

D. Observed Limitations and Open Problems

Despite satisfactory results, the proposed chatbot system has shortcomings. The intent classifier trained on approximately 800 utterances shows lower performance on ambiguous multi-intent messages and highly informal or colloquial language not well represented in the training data. The VADER lexicon, which was created with general social media text in mind, may not capture the subtleties of mental health-specific language. To improve robustness, the training dataset size and linguistic variety need to be increased, and data augmentation techniques like synonym replacement and back-translation need to be explored.

E .Future Research Directions

Presently, the analysis indicates that using effective NLP preprocessing with machine learning has the potential to be an important path for developing the next generation of mental health chatbot systems. One of the significant paths is the management of multi-turn contexts. This is important to enable the chatbot to have more coherent conversations. Integration with crisis detection classifiers is another significant path being emphasized, especially for the safe deployment of the chatbot. Another path being emphasized is the development of chatbots that support multiple languages, especially regional languages of India. This is important to enable the chatbot to serve the larger population.

V. CONCLUSION

Human Activity Recognition has emerged as an important enabling technology for systems that have to recognize human activities automatically based on visual or sensor-based information. The impact of HAR can be seen in various applications like medical supervision systems, intelligent security monitoring systems, and intelligent living spaces. In this paper, it was discussed how human activity recognition techniques have progressed from traditional manually constructed motion features to fully automatic deep learning techniques for recognizing human activity patterns. It was also discussed how OpenCV can be utilized for developing such systems based on its support for efficient video preprocessing. The integration of OpenCV and learning techniques can be utilized for more effective development of accurate human activity recognition systems. In particular, it was seen how hybrid techniques combining convolution-based feature learning and temporal sequence learning can be utilized for effective human activity recognition systems for recognizing complex human actions from video streams. These hybrid techniques have validated the concept of combining different components of human activity recognition for more effective and efficient human activity recognition systems. At the same time, there are some practical issues that still have to be addressed. In particular, it was seen how the robustness of the model still depends on scene changes, visibility, and style changes among individuals. Another problem associated with this model is the high resource requirement of existing advanced models, making it difficult for this model to be deployed at the edge. The future of this model would be determined based on how this model can be developed based on compact model structures, how this model can be utilized for integrating cross-modal learning with videos and wearable sensors, and how this model can be utilized for effective adaptive training for better generalization. In conclusion, it was seen how efficient human activity recognition systems developed based on efficient vision libraries and intelligent learning models are moving towards wider deployment and application.

VI. FUTURE SCOPE

The domain of HAR has been rapidly developing, and there are several opportunities left in terms of development and innovation. The future work should be directed towards developing more adaptable, accurate, and real-time HAR systems, especially in healthcare and assistive technologies. The integration of OpenCV with upcoming technologies like deep learning, IoT, and edge computing will be instrumental in developing future HAR systems.

A. Expanded and Diverse Training Data

Expanding the training set with more utterances and diverse data can improve the robustness of the classifier.

B. Multi-Turn Context Management

Combining video with wearable sensors can improve the accuracy of the system and allow for more intelligent decision-making.

C. Optimization for Low–60%-Power Devices

Adding a dedicated crisis detection classifier can make the system safer and more suitable for real-world use in the clinical and educational settings.

REFERENCES

[1] J. Wang et al., "Deep Learning for Sensor-Based Human Activity Recognition," Pattern Recognition Letters, 2019.

[2] S. Bhattacharya and P. Das, "Real-Time Human Activity Recognition Using OpenCV and CNN," Journal of Ambient Intelligence, 2020.

[3] Krishna S. B. Siva et al., "Human Activity Recognition with OpenCV," IJERT, 2024.

[4] I. A. Putra et al., "Human Action Recognition Using MediaPipe and LSTM," Teknik Journal, 2022.

[5] K. Host and M. Ivasic-Kos, "Human Action Recognition in Sports Using Computer Vision," Heliyon, 2022.

[6] M. Janapati et al., "Gait Driven Pose Tracking Using OpenCV," Engi- neering Proceedings, 2024.

[7] S. Gupta et al., "Human Action Recognition Using Deep Learning Techniques," Informatics and Automation, 2024.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, 2015.

[9] M. Vrigkas, C. Nikou, and I. Kakadiaris, "A Review of Human Activity Recognition Methods," Frontiers in Robotics and AI, IEEE, 2015.

[10] L. Wang et al., "Temporal Segment Networks for Action Recognition in Videos," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.

[11] A. Jalal, S. Kamal, and D. Kim, "A Depth Video Sensor-Based Life-Logging Human Activity Recognition System for Elderly Care," Springer Journal of Sensors, 2014.

[12] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "IMU-Based Wearable Human Activity Recognition Using Misalignment-Robust Deep Learning," Springer Neural Computing and Applications, 2017.

[13] K. Simonyan and A. Zisserman, "Two-Stream Convolutional Networks for Action Recognition in Videos," NIPS, 2014.

[14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, MIT Press, 1997.

[15] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NeurIPS, 2012