Intelligent Mood Monitoring with Voice Companion

P. Logaiyan[1], V. Chandru[2*]

[1]Associate Professor, Department of Master Computer Application, Sri Manakula Vinayagar Engineering College, Pondicherry, India

[2]Student, Department of Master Computer Application, Sri Manakula Vinayagar Engineering College, Pondicherry, India

How to Cite this Article:

Chandru, V. (2026). Intelligent Mood Monitoring with Voice Companion. International Journal of Creative and Open Research in Engineering and Management, 02(6). https://doi.org/10.55041/ijcope.v2i6.131

https://doi.org/10.55041/ijcope.v2i6.131

***Abstract*:** In the contemporary world of constant hustle and bustle, mental issues like stress, anxiety, and emotional imbalance are rising rapidly. Stress, if continuously exposed to it, can adversely impact physical as well as psychological well-being, hampering productivity and causing various diseases. Thus, detecting and monitoring mood and other emotions becomes a necessity. Conventional approaches like using questionnaires and self-reporting methods may give unreliable results due to subjectivity and personal bias. To address these problems, the idea of creating an AI-Based Mood Monitoring and Voice Companion system is suggested in a given project. The system will use Artificial Intelligence and Machine Learning algorithms to determine user's emotion based on their vocal inputs and messages they type in a chat. Identifying patterns in these interactions and comparing them with pre-established ones, the program categorizes the current mood into various states: happy, sad, stressed, neutral, etc. Accordingly, based on the determined state of emotions, the voice companion offers real-time response and gives appropriate suggestions to the user.

KEY WORDS: Artificial Intelligence, Mood Monitoring, Emotion Detection, Voice Companion.

1. Introduction

In today's fast-paced world, people face stress, pressures, and emotional challenges all the time. Because of high academic and work-related stresses and changes in personal lifestyles, the prevalence of mental disorders like stress and anxiety has become rampant. Moreover, it is often hard for some individuals to understand their emotions, making it more challenging for them to convey their feelings. Hence, monitoring of emotion has become necessary in order to live a healthier life. The traditional approaches to gauging human emotions through questionnaires and self-reporting techniques are unreliable in many ways. For one thing, these traditional approaches hinge on a person's reliability and ability to understand themselves. Furthermore, they are inefficient since they do not provide a continuous stream of information about a person's emotional status. In the modern world, due to advancements in technology,
we have intelligent machines that can interpret human moods effectively. An Intelligent Mood Monitoring with Voice Companion is a proposed

system that can identify a person's mood based on vocal and written expressions of thoughts. The mood will then be classified according to its nature such as happiness, sadness, stress, or neutrality. It will communicate with user to offer support during their tough times. The goal of the intelligent mood monitoring system is to provide support to users without needing human intervention. In summary, this project will help in develop smarter and user-friendly approaches towards emotional well-being. The project utilizes the approach of mood monitoring combined with communication to create a more interesting interaction for the user. The project can be further developed and incorporated into smartphones or other smart gadgets. Moreover, modern advancements in intelligent systems facilitate more sophisticated approaches to interpreting human behavior and emotions. With the use of linguistic and para linguistic cues, it becomes possible to detect nuances in people's moods which would be hardly noticeable otherwise. Consequently, a whole new range of applications appears on the market, enabling users not only to perform complex tasks but to receive valuable information about their psychological condition. In turn, such tools may be used to predict future emotional problems and help avert them before they become critical. Another advantage of the technology is that it provides an improved way of communicating. Thanks to the voice component, a person feels much more involved when talking to a machine. The voice companion serves as a conversational partner who listens to what one has to say and gives a relevant reply. Furthermore, the suggested system is highly flexible and can be used in different fields including health care, students' well-being assessment, and office stress management. The system provides a cost-efficient approach that can be utilized by any user group. With technological advancement, these types of systems may become an integral part of people's lives and contribute to better awareness of psychological problems.

2. Literature Review

Sharma et al. (2020): The scholars focused on emotion recognition by analyzing emotions using

text-based approaches. To detect emotions, NLP was combined with machine learning algorithms like Naïve Bayes, Support Vector Machine, and others. The study proved that emotions could be detected based on text-based analysis; however, it did not include any other sources except text and therefore had low practical use. Rao and Mehta (2021): Rao and Mehta analyzed the emotion recognition process with the help of acoustic characteristics like pitch, tone, and speech speed. As part of their research, the scholars used CNN as part of deep learning technology to increase accuracy. The algorithm worked well in case the input was of good quality, but poor quality decreased its effectiveness significantly. Kumar et al. (2022): The research was based on integrating both textual and vocal inputs for detecting emotions. Hybrid Machine Learning models were employed to increase the efficiency of classifying data. The system showed a high level of efficiency but had no mechanism for interaction with users and could not provide real-time emotional support. Patel and Singh (2023): Researchers designed a conversational chat bot to help users with their mental health problems. This is done through AI technology. The chat bot interacted with users and provides them with some kind of emotional support. The problem with the system was that it could not detect emotions. Zhang et al. (2022): In their work, the scientists presented an emotional recognition method based on deep learning algorithms. The authors used LSTM networks to recognizing emotions by analyzing patterns of behavior over time.

While the algorithm was accurate, it required significant computing power and large data sets, which made it less suitable for light applications. From the above researches, it can be seen that considerable development has been achieved in the area of emotion detection and stress identification with the use of intelligent techniques. Different methodologies including physiological signals, text analysis, voice recognition, and multi modal system have been used in the researches. Although these techniques have proved successful, they have their drawbacks, such as being expensive, having no interactive process, depending on specific kinds of information, and raising problems concerning privacy and scalability. The current systems depend mostly on one input device and are not comprehensive or user-friendly. It is therefore evident that there is a need for a technology that utilizes multiple means of input together with the provision of interactivity. It is for this reason that the proposed system intends to address the challenges posed by using mood recognition systems through incorporating the provision of an interactive voice assistant.

3. Problem Statement

In today's contemporary lifestyle, the human being is often faced with stress, anxiety, and emotional imbalances because of academic pressures, work-related demands, and other personal problems. Such emotional disorders affect not only the mental well-being of a person but are also related to physical well-being and productivity. In spite of the increasing relevance of mental health issues, people tend to ignore or overlook their emotional state in order to prevent any negative results. The methods traditionally applied in the evaluation of human mood include surveys and various self-reporting systems, which are rather subjective because they require people to be honest and sincere about their mood and emotions. Besides, a human being should clearly know his/her current state, and it cannot be ensured all the time. However, existing technologies, which include wearable devices and physiological monitoring tools, produce more factual data by assessing heart rate and sleeping patterns. Nevertheless, the use of such technologies is rather expensive, requires certain hardware installations, and is limited to people who can afford them. Moreover, their usability in daily life can cause discomfort for users along with raising issues connected with privacy protection and data security. The other limitation is the lack of personalization of existing emotion recognition systems. People's emotional states can greatly differ depending on different factors and thus a universal tool cannot guarantee the same level of precision. Besides, most of the available emotion recognition algorithms use only one input data source, while the inclusion of various input data would improve the system significantly. There is, thus, a dire need for an intelligent, automated, and user-friendly system that will constantly monitor emotional states and offer appropriate help. This paper proposes the creation of such a solution Intelligent Mood Monitoring with Voice Companion System. It is intended to utilize both mood monitoring technology and voice assistance technology.

4. Existing System

The existing technologies for mood monitoring and emotional analysis involve only traditional means of evaluation like surveys, questionnaires, and self-reports. These techniques are convenient and easy-to-use; however, they are very subjective. The reliability of the results depends directly on user's honesty and awareness. Consequently, these technologies do not allow obtain consistent results. Alongside traditional means, there are complex technological solutions which use bioinformation to detect emotional states. In particular, these technologies include measuring the heart rate, brain waves, sleeping and awakening states, and skin reaction using special wearables. Despite the increased objectivity of these technologies, their implementation is complicated because of the necessity of additional hardware and equipment. The other type of emotion detection that uses Natural Language Processing has also been created. It detects emotion from the messages of users or information posted on social media sites. Though easier to use, these systems work only with text inputs and cannot detect the actual emotional state of the user due to language ambiguity. Another method that is used to detect emotions is voice analysis. It works by detecting the pitch, tone, and speed of speech to determine what emotions the person speaking is expressing. Though this system works quite well, it is highly dependent on environmental factors such as noise. The current systems rely only on one form of data, whether it is physiological data, text data, or vocal data. Such an approach makes it difficult for the system to integrate different forms of data, which may lead to a decline in efficiency and effectiveness during the process of emotion recognition. In addition, most systems lack any post-emotion recognition interaction and feedback, which further diminishes their usefulness

in continuous emotional monitoring. In general, there are numerous limitations in the current systems, including high costs, poor accessibility, low accuracy, and lack of user interaction, among others.

5. Proposed System

The suggested system named as Intelligent Mood Monitoring with Voice Companion is intended to offer a practical way out of detecting the user's emotion in a real time setting. Different from previous approaches, this approach utilizes different inputs like voice and text to analyzing the user's emotional state by making use of some intelligent processes to detect user's behavior. Inputs will be taken using both voice and text. Data taken from users will be processed and analyzed. Some relevant features will be identified to detect the emotions of the user. These features will assist in categorizing users into moods. The user's emotional state can be classified into one of several states, such as happy, sad, stressful, and neutral among others. It is a system that has been designed to be easy to use and affordable. There is no need to use sophisticated equipment like wearable devices. Rather, it can be adopted by use of devices that are readily available. Examples include smartphones and personal computers. Adaptability is yet another quality of the recommended system. This is because it will learn from the interaction process and thereby become more efficient in generating better results. This means that it will be able to suit individuals with unique emotional traits. In general, the suggested system can solve the problems inherent in traditional approaches due to its integrative, real-time, and interactive nature, making it feasible and highly beneficial for managing moods and psychological states. Besides detecting emotions, it can also provide prompt assistance by interacting with a virtual assistant.

6. System architecture

The AI-Based Personal Productivity and Task Management Assistant has layers that work together. These layers help with productivity management.

A. Presentation Layer : The Presentation Layer is the user interface. Users interact with the application through this layer. They can do things here. These include registering, logging in creating tasks setting goals, managing reminders, keeping notes viewing analytics generating reports and talking to the AI Assistant. The interface works on desktop and mobile devices. It is easy to use and understand.

B. Application Layer : The Application Layer is built using Python and the Django Framework. This layer handles tasks. These tasks include user authentication, task management, goal tracking, reminder scheduling, notes management, calendar operations, productivity analytics and report generation. It processes user requests. It also helps parts of the application communicate with each other. These parts include the frontend, AI services and database components.

C. AI Processing Layer : The AI Processing Layer provides help and productivity suggestions. This layer uses AI models and Gemini API integration. It analyzes user activities. It also generates productivity suggestions answers user queries. Provides task-related guidance. The AI Assistant helps users make decisions. It provides recommendations to improve productivity.

D. Database Layer : The Database Layer uses Supabase PostgreSQL. It. Manages application data securely. It keeps records of users, tasks, goals, reminders, notes, productivity reports, analytics data and user preferences. The database helps store, retrieve, update and manage information. It makes sure data is safe and secure.

E. Analytics and Reporting Layer : The Analytics and Reporting Layer looks at user productivity data. It generates insights. It creates reports on task completion goal achievement, productivity trends, reminder effectiveness and overall performance. The generated analytics help users track progress. They also help users find areas, for improvement. Users can make decisions for better productivity management.

7. Methodology

The methodology that has been suggested for designing the proposed application "Intelligent Mood Monitoring with Voice Companion" consists of different stages beginning from the data collection up to the process of mood monitoring and generating a voice response. It has to be noted that there are five steps in the system, namely data collection, data preprocessing, feature extraction, model training, and mood classification with generating a voice response. All these steps are connected with each other in a row that allows converting users' input data into the output emotional status.

Data Collection: Data concerning the emotion of the user is obtained through different methods such as data input via voice and text. Voice input may include voice

parameters such as speech features, tonality, and others. As for text data, it may consist of user messages, their inputs in conversations, and similar. Typing behavior may also be taken into account when analyzing users' emotions.

Preprocessing: Data collected is preprocessed before conducting any analysis. The process entails eliminating noise in voice data, rectifying spelling mistakes in texts, and formatting the data. Unnecessary data is eliminated to ensure accuracy to the results. Texts are tokenized and normalized during data preprocessing while voice data is converted to appropriate format. The data is split into a training set and test set in the proportion of 70% and 30%, respectively.

Features Selection: In this phase, significant features associated with emotions are selected from the data set. In case of voice input, features like pitch, tone, energy, and speech rate are taken into account. While in case of text input, features like sentiment score, keywords, and emotions are selected. These features are then mapped to numbers so that they can be utilized by machine learning algorithm.

Training and Classification of Model: The set of features obtained can be used for training machine learning models such as Decision Trees, Random Forests, and Support Vector Machines. The pattern recognition algorithm is used for recognize various emotional states based on labeling the data. After training the system, the emotional state of the user is classified into one of the following classes such as happiness, sadness, stress, and neutrality.

Prediction and Voice Response Stage: Once the process of classification has been completed, the system gives out corresponding outputs based on the identified mood state. A voice companion component is incorporated into the system for outputting the information in an interactive form. The system is designed to give feedback or suggestions to the user in real time.

In general, the suggested approach is an effective solution to detect moods and interact with humans. Due to the involvement of several steps ranging from collecting data to generating a response, the system can perform its tasks effectively.

8. Working Process

The method involves capturing the user input via voice or text communication and analyzing the information in order to identify emotions. After cleaning the input data and preparing it for analysis, certain features, including the tone and sentiment, are extracted from it. The features are later fed into a trained model that identifies the mood of the user in various states, including happy, sad, stressed, or neutral. A response to the user is then provided based on the identified state through the use of a voice companion. The system constantly analyzes the input and provides responses for the best results. The system captures user input via voice or text. The input data is cleaned and analyzed. Certain features are extracted from the input. The mood can be classified as happy, sad, stressed, or neutral. A response is provided using a voice companion.

9. Dataset Description

The dataset for the Mood Monitoring with Voice Companion system has voice recordings from users. These recordings show emotional states. The main goal of the dataset is to train and test the emotion recognition model. This model identifies emotions from speech.

The dataset has samples of various emotions. These emotions are happiness, sadness, anger, fear, stress and neutral mood. Each audio sample is labeled with a category. It also has information like recording duration, timestamp and confidence score. The dataset was prepared using techniques. These techniques include noise removal, normalization, feature extraction and speech segmentation. Important speech features like Mel Frequency Cepstral Coefficients (MFCC) pitch, energy and spectral characteristics were extracted. These features help the machine learning model understand emotions in speech. The prepared dataset was used to train and test the emotion detection model. This model is part of the voice companion system. The dataset helps the system recognize moods in time and give personalized AI responses. The system works better with voice samples. It is more reliable and adaptable, for users and emotional conditions.

ame	escription
ser_ID	nique identifier assigned to eac er
udio_ID	nique identifier assigned to eac ice recording
ice_Input	ser's recorded speech or text inpu ed for emotion analysis
notion_Label	etected emotional category such a appy, Sad, Angry, Neutral, c ressed

motion_Score	onfidence score representing th curacy of the detected emotion
peech_Duration	uration of the recorded voice inpu seconds
esponse_Text	I-generated response based on th tected emotional state
lert_Status	dicates whether an emergenc ert has been triggered
mestamp	ate and time when the interactio curred
ser_Feedback	edback provided by the use garding the system response

10. Implementation

Data Acquisition: Data is gathered via voice and text interaction. User emotions can be identified using the gathered data. It is used as a foundation for the next process.

Data Preprocessing; It involves cleaning and normalization of the data. For voice-based data, noise removal is done, while text-based data involves eliminating irrelevant words.

Feature Extraction & Feature Selection: Critical features such as tone, pitch, and sentiment are extracted. Relevant features are selected in order to minimize complexities.

Model Selection and Training: A Machine Learning algorithm like Decision Tree or SVM is chosen. The model is then trained through labeled data in order to learn the patterns of various emotions.

Model Validation: The model undergoes validation through the use of evaluation criteria such as accuracy and precision. This will help ensure that the model works properly.

Sentiment Prediction and Response Generation: The validated model uses the input data to predict the sentiment of the user. Depending on the prediction, responses are generated by the system. The voice companion responds in real-time.

System Implementation: The system is implemented on platforms such as mobile or Web. The system can be conveniently accessed by the users. Real-time usage is facilitated in the system for better user experience .

In general, there is a systematic process in which the implementation process begins with collecting data, followed by the training process of machine learning models until deploying them into the application itself.

Every step serves its purpose in helping to detect the mood accurately and interact smoothly. Furthermore, implementing Machine Learning models in this case increases prediction accuracy, and adding the voice companion creates a more pleasant experience. Thus, the whole implementation process serves to help develop mood monitoring system..

11. Tools and Technology

In order to facilitate the process of creating the Intelligent Mood Monitoring with Voice Companion system, different technologies and tools are employed in order to guarantee efficiency. Specifically, Python will serve as the primary coding language due to its ease of use and excellent support for data processing. For data manipulation and processing, Pandas and NumPy will be employed. Meanwhile, Scikit-learn can be utilized for machine learning models, and Tensorflow can be employed for deep learning operations. In addition, NLP and speech recognition technologies will be employed in order to analyze the textual and vocal input provided by the user. Finally, the system will be created through the use of platforms such as Jupyter Notebook and Visual Studio Code.

12. Challenges

Security Concerns with the Data: Personal and emotional data is considered sensitive data and needs proper protection. The risks are related to possible misuses or leaks of user data.

Data Integrity: Different emotions may be expressed by different people in a unique way. Input data does not always correspond to the real feelings of a person. It might result in inaccurate results.

Challenges of Voice Recognition: Background noise may influence voice recognition. People's voice characteristics differ according to accent, for instance. It might affect results and generate inaccuracies in emotion detection.

Text Analysis Obstacles: The problem of ambiguity hampers interpretation since sarcasm and complicated structures are challenging to identify. This will also affect the efficiency of sentiment analysis.

Inadequate Data: It is not easy to access huge datasets with equal numbers of classes. This will lower the performance of the algorithm.

Efficiency in Processing Time: Processing must be efficient and done instantly. Processing time will affect the performance of the system.

Personalization Difficulties: Each individual has his or her own emotional patterns. The same model cannot apply to all individuals. Continuous learning is necessary for personalization.

13. Future Work

Integration with Wearable Devices: Integration with wearable devices like tracker may help improve the functionality of collecting physiological information like heartbeat and sleep pattern.

Implementation of Deep Learning Models: Deep learning models could further boost the performance and efficiency of the system since they will improve prediction through recognition of emotions.

Multilingual Version of the Application: This software can be developed in different languages. It will increase its usability and accessibility to a wider number of users.

Voice Interactions: Responses can also be personalized and made to sound more like the human voice. With improved speech synthesis, communications can be better. This will improve user engagement.

User-Specific Recommendations: The system can offer suggestions depending on how the user behaves. It can suggest relaxation techniques or any activity that will make him feel happy.

Mobile App Design: The system can also be created as a mobile application which would make it more practical to use.

The Personal Productivity and Task Management Assistant that uses Artificial Intelligence can be made better by adding new technologies and features that help people get things done. In the future the system might let people create and manage tasks by talking to it.

The system can automatically decide which tasks are most important guess when they are due and suggest the schedule based on what the user has done before. These new Artificial Intelligence features can help people get more done in time and make better choices about how they use their time. The system can sync everything across all devices and send notifications in time so people always know what is going on.

The Personal Productivity and Task Management Assistant that uses Artificial Intelligence can also get features that give people more information about how they get things done. It can track habits monitor wellness and even predict how well people will do in the future. All these new features can make the system more useful, for both work tasks. The Personal Productivity and Task Management Assistant that uses Artificial Intelligence can really help people get more done and be more productive.

14. Conclusion

In this research, Intelligent Mood Detection Using Text and Voice Companion is a simple yet efficient solution that enables the detection and management of emotions of a human being using voice and textual input. It assists in determining whether someone is stressed, sad, or happy. The system is not only more accurate than the previous solutions but also easier to use. Furthermore, the system allows for continuous monitoring, which was a challenge in earlier technologies.

In addition the system uses Artificial Intelligence and Natural Language Processing to analyze both voice and text inputs. This helps it understand how people are feeling. By using both voice and text the system gets better at recognizing emotions. Gives more accurate results. The system looks at voice and text to understand users emotions in a way. This makes the mood detection process work better.The Voice Companion module is like a friend that talks to users in a friendly way. It gives responses based on how the user's feeling. For example if a user is feeling down it might send a message. If a user is feeling stressed it might suggest ways to relax. This makes users feel like someone cares and is listening. The system keeps track of users moods over time. This helps users see how their emotions change. It also helps them find patterns in their stress. The system keeps mood history records.

It generates insights based on user interactions. The system also keeps user data safe. Only people who are allowed to can see user information. Overall the system helps people feel better emotionally. It helps people and technology work together in a way. The Artificial Intelligence and Natural Language Processing techniques are key, to making this happen. They help the system understand users and provide support.

15. Reference

[1]Saini, R., et al., "Stress Detection using Machine Learning Techniques," 2019.

[2]Healey, J., & Picard, R. W., "Detecting Stress during Real-World Driving Tasks," 2010.

[3]Canzian, L., & Musolesi, M., "Trajectories of Depression using Smartphone Data," 2015.

[4]Kaur, P., & Sharma, M., "Stress Prediction using Deep Learning Methods," 2021.

[5]Zhou, X., et al., "Deep Learning Framework for Stress Detection using Wearable Data," 2022.

[6]Hossain, M. S., & Muhammad, G., "Emotion Recognition using Deep Learning," 2019.

[7][7] Schmidt, P., et al., "A Multimodal Dataset for Stress