EAI Endorsed Transactions
on Mobile Communications and Applications Research Article **EAI.EU**

Enhancing Mental Health With PHILOI: A Comprehensive Analysis of Mood Music and Chatbot Module

K.S. Kumavat[1,*], D. Gatagat[2] , K. Wakode[3] , S. Gundawar[4] and V. Jain[5]

[1,2,3,4,5] Vishwakarma University, Pune

Abstract

The project aim was to develop an app that would enable the recording and monitoring of behaviour related to specific aspects of wellness, as well as support those aspects of wellness that are entertainment-related. Our main goal was to envision and develop an app with the well-being of users in mind. People's moods can be improved upon or changed by music, with music and mental health tightly intertwined. Music is frequently used to complement or change an individual's mood. While there are advantages to mood-appropriate music, it may cause us to remain in a depressed, angry, or nervous state. A survey was conducted to examine these aspects. After performing a lot of research and interviews in this area, we found 68% of those surveyed listen to music according to their mood or to change their mood. This inspired us to build an application that not only plays music but also recommends songs to users, eliminating the daily nuisance of selecting the right music, which can waste valuable time. As mental balance is an essential component of healthy existence in today's hectic world, to enhance the practicality of our app, as icing on the cake, we included an AI chatbot that not only converses with the user but also provides them with suitable advice on their concerns.

Keywords: Mood music, chatbot, Artificial intelligence, Android application, Face detection.

Received on 20 03 2024, accepted on 10 06 2024, published on 22 07 2024

doi: 10.4108/eetmca.5146

1. Introduction

The nexus of mental health and digital innovation has emerged as a crucial route for offering accessible and efficient support to those dealing with emotional and psychological issues in a time of fast technology growth. 'Philoi' distinguishes out among the plethora of mental health applications as a light of hope, fusing cutting-edge technology with sympathetic design to meet the various requirements of its users. The two main components of the mental health counselling app Philoi that are extensively examined in this study report are Chatbot and Mood Music. One cannot overestimate the importance of mental health as it is a crucial component of total wellbeing. The urgent need for accessible and efficient interventions is highlighted by the prevalence of mental health illnesses around the world, which is compounded by several socio-cultural and environmental factors. With its cutting-edge methodology, Philoi aims to close the gap between conventional therapeutic techniques and modern technology, providing people seeking direction, comfort, and empowerment in their mental health journeys with a vibrant platform. The deep influence of music on emotional well-being is encapsulated in the Mood Music module, a crucial part of Philoi. Utilizing the healing power of sound, Philoi uses a carefully chosen collection of songs that are designed to suit a range of moods and emotional states. Based on an empirical study on the psychological impacts

*Corresponding author. Email: kavita.kumavat@vupune.ac.in

of music, this module offers users a customized auditory experience that aims to research the psychological effects of music, providing users with a personalized auditory experience designed to alleviate stress, anxiety, and depression. The Chatbot function complements the Mood Music module and enhances the user experience by offering a private and easily accessible channel for communication. The chatbot uses natural language processing algorithms to recognize and address a variety of emotions and worries. It is intended to engage users in compassionate dialogues. The chatbot provides direction, coping mechanisms, and resources through an easy interface, establishing a sense of connection and support for people navigating the complexity of mental health.

The goal of this study is to thoroughly examine Philoi in order to clarify the effectiveness of its Mood Music and Chatbot modules in enhancing mental health outcomes. We want to present a nuanced perspective of the app's potential as a revolutionary tool in the field of mental health intervention by synthesizing empirical studies, user feedback, and expert evaluations. The theoretical foundations of the mood music module will be covered in more detail in later sections when we look at the psychological processes by which music has a therapeutic effect. Additionally, we will carefully examine the Chatbot feature's operation and efficacy, gauging its ability to offer sympathetic and fact-based support. In conclusion, Philoi is a prime example of how innovation and compassion can coexist, and the incorporation of technology into mental health care signifies a fundamental paradigm change. We want to present a nuanced perspective of the app's potential as a revolutionary tool in the field of mental health intervention by synthesizing empirical studies, user feedback, and expert evaluations. The theoretical foundations of the mood music module will be covered in more detail in later sections when we look at the psychological processes by which music has a therapeutic effect. Additionally, we will carefully examine the Chatbot feature's operation and efficacy, gauging its ability to offer sympathetic and fact-based support. In conclusion, Philoi is a prime example of how innovation and compassion can coexist, and the incorporation of technology into mental health care signifies a fundamental paradigm change.

The paper is arranged in the following manner-

Section 2 focuses on the literature survey, section 3 contains a system overview, section 4 elaborates requirements of the system, section 5 focuses on results and implementation details, and section 6 concludes the paper.

2. Literature Survey

AI chatbots are increasingly being used to provide consumers with mental health advice. The use of mood music by chatbots to aid users in managing stress and enhancing their mental health is one area where they are demonstrating promise [1]. An increasing corpus of studies suggests that music may benefit mental health. Music has been shown to alleviate stress, anxiety, and depression, as well as improve mood and emotional well-being [2]. While research on the effectiveness of AI chatbots for mood music is limited, preliminary findings suggest that they can be effective in improving mood and reducing stress. According to one study, users who listened to mood music suggested by an AI chatbot reported a significant reduction in stress and anxiety [3]. A promising and creative strategy for mental health care is the employment of chatbot therapists. They provide convenient, accessible, and cost-effective support for those suffering from mental illness. However, further research is needed to explore their effectiveness in comparison to traditional therapy methods and to improve their design and implementation [4].

The study examined the prevalence of mental health problems, the use of mental health services, and the global burden of illness to assess the Region of the Americas' lack of mental health care. Mental and substance use disorders accounted for 10.5% of the global illness burden in the Americas, with a 65.7% treatment gap for moderate to severe disorders. More than half of children and adolescents with serious mental illnesses were not receiving treatment, and a sizable proportion of indigenous peoples in the Americas were not either. The study emphasizes the need for improved access to mental health services in the area [5]. To manage employee mental health, a VR-based mental healthcare training program has been developed that reduces stress. A new version of the course employing chatbots and cell phones is being considered to make the course more convenient and to keep users motivated. In an experiment, chatbots were employed in the self-guided mental healthcare course to increase user motivation and reduce stress. In addition to the study's limited sample size, there was no assessment of long-term use. Additional research is needed to address these restrictions [6]. In underprivileged regions where people might not have access to conventional mental health therapy, mind-body practices have been shown to help treat mental health problems. Racial/ethnic minorities also use supplementary health practices, such as prayer and herbal medicines, to cure their illnesses. Evidence shows that mind-body therapies can enhance mental and physical health symptoms, self-care, functioning, and general quality of life in disadvantaged groups. However, larger studies are needed to investigate the efficacy of mind-body interventions in this population [7]. According to research, analyzing social media material might help increase the success rate of mental health diagnosis, with an algorithm intended to identify depression from Instagram postings outperforming clinicians. Chatbots may eventually utilize user data to notify mental health professionals of more effective treatments. Furthermore, by offering users coping mechanisms and even warning specialists of suicidal inclinations, chatbots might help to alleviate the difficulty in receiving mental health care in remote places where there is a scarcity of mental health practitioners. To better detect people's emotions, future chatbots may employ advanced sentiment analysis algorithms [8]. The study proposes a typology for chatbots based on the duration of the user's interaction with the chatbot and the locus of control for user involvement. The typology is used to assess high-level interface design for four sample chatbot usage. The typology was demonstrated to be thorough and to contain just certain types, and the categorization dimensions were determined to be generic and pertinent. This work is considered a step toward improving the utility and user experience of chatbots [9].

The study highlights the increasing use of motivational interviewing (MI) for mental health concerns, in addition to addiction and substance abuse. It also emphasizes the importance of incorporating both technical and relational components of MI in designing conversational sequences. The presented case study suggests that such sequences can facilitate stress management conversations and encourage self-reflection. However, more diverse sequences and contextualized feedback are needed to improve conversational experiences and confirm empirical effects [10]. The article discusses the use of machine learning algorithms for detecting human emotions from facial expressions captured in videos, EEG signals, or images. The implementation is divided into three parts, including face detection, feature extraction, and classification using machine learning algorithms. The study explores various machine learning algorithms and feature extraction techniques to achieve accurate emotion identification. The results show promising accuracy rates for the proposed method, outperforming previous studies that used ORB feature descriptors [11]. The research offers a chatbot typology based on the length of the user's relationship with the chatbot and the placement of the user's control. Four sample chatbot purposes are examined using the typology to examine high-level interaction design. Typologies were revealed to be comprehensive and to include exclusive types, while categorization dimensions were universal and pertinent. As a result of the work presented, chatbots are expected to become more useful and enjoyable to use [12]. The article proposes an emotion-aware personalized music recommendation system (EPMRS) that can identify the association between user data and music using a deep convolutional neural network (DCNN) and weighted feature extraction (WFE) technique. The system uses implicit user ratings generated by the term-frequency and inverse document frequency (TF-IDF) approach to provide music suggestions to the user on how they are feeling at the moment. Based on electroencephalography input, the EPMRS performs better than the PMRSE and the content similarity music recommendation system (CSMRS). The automatic extraction of the user's current emotion from social media data will be combined with the user's data from other sources, such as YouTube, Facebook, and Twitter, in future work [13]. Recent years have seen a rise in interest in the study of emotion identification from facial expressions, with many studies concentrating on increasing the precision of existing models. Several studies have looked into deep learning techniques to increase accuracy, including convolutional neural networks (CNNs). However, one challenge that has been identified is the age bias problem in the training data, which can affect the accuracy of emotion recognition, particularly for children. Several studies have investigated the impact of age bias on emotion recognition and proposed various solutions to address this issue. Overall, there is a need for further research to improve emotion recognition accuracy and address the age bias problem in training data [14]. According to the literature review, there has been a significant recent study on face recognition, with applications in security, surveillance, identity verification, and more. The suggested method uses computer vision and deep learning algorithms to simplify the process of facial identification. The three objectives of the study are face detection, identification, and emotion classification. The suggested system's performance metrics are verified with an accuracy of 88%, and it is intended to detect, recognize, and categorize human faces in real-time. To classify and identify faces, the VGG 16 and KDEF dataset are utilized. According to the study, this application may be extensively employed in the fields of education, business, medicine, and electronics [15]. The experimental results presented in this study showcase several key findings. Firstly, the proposed approach achieves a high Correct Recognition Rate (CRR), indicating its effectiveness in recognizing facial expressions.

Secondly, the consideration of facial elements and muscle movements leads to significant performance improvements, highlighting the importance of incorporating dynamic features. Additionally, the approach demonstrates promising results even in the presence of face registration errors, which enhances its practical utility. Lastly, a comparison with state-of-the-art methods validates its superiority by achieving the highest CRR on the JAFFE database and ranking among the top performers on the Cohn-Kanade (CK) database [16]. Firstly, it presents the "Integral Image" representation, enabling rapid feature computation. Secondly, it employs a simple yet effective classifier using the AdaBoost algorithm for feature selection. Lastly, it introduces a cascading technique to swiftly discard non-promising regions, enhancing overall computational efficiency. The approach achieves face detection results comparable to previous state-of-the-art systems, running at an impressive 15 frames per second on a standard desktop. Importantly, this work has broader implications for computer vision and image processing, offering potential applications in diverse domains beyond face detection. The utilization of the Integral Image for feature extraction holds promise for tasks requiring scale invariance, and Haar-like features may find applications in various other domains. The paper's second contribution, an efficient classifier, is expected to benefit object detection tasks like automobile or pedestrian detection. Furthermore, its feature selection approach with AdaBoost has broader implications for improving efficiency in various machine learning applications. The third contribution, the cascade of classifiers, stands out for its simplicity and homogeneity in structure, making it both comprehensible and adaptable. Unlike other approaches proposing more complex and heterogeneous mechanisms, this system offers straightforward trade-offs between processing time and detection accuracy [17]. Recent advancements in facial expression recognition (FER) have primarily revolved around deep neural network (DNN)-based approaches, showcasing their potential to overcome limitations of traditional machine learning-based FER methods. However, the high memory requirements and processing costs of DNNs restrict their practical applications, particularly in low-specification devices like those found in vehicles. This paper introduces a fast FER algorithm tailored for monitoring a driver's emotions in resource-constrained environments. It employs a hierarchical weighted random forest (WRF) classifier trained based on sample data similarity, enhancing accuracy. Geometric features extracted from facial landmarks are input into the hierarchical WRF classifier, yielding competitive performance with deep learning-based FER methods while drastically reducing processing costs. The study not only presents an innovative FER method based on geometric features and hierarchical WRF but also addresses the need for real-world driving datasets with varying illumination conditions. The newly created KMU-FED dataset serves this purpose and validates the proposed method's effectiveness. Additionally, the research highlights the adaptability of this FER method in embedded systems for intelligent vehicles and envisions its broader applications in fields like entertainment, education, virtual reality, and gaming [18]. Facial expression recognition in videos is a complex challenge in computer vision and human-computer interaction. Texture features have been a popular choice for capturing skin deformation-related intensity changes. However, they often struggle with issues related to albedo and lighting variations. To address these challenges, this paper introduces a novel texture feature known as "image ratio features." Image ratio features outperform traditional texture features such as high gradient component features in terms of robustness to albedo and lighting variations. Furthermore, the study combines these image ratio features with facial animation parameters (FAPs), which describe the geometric movements of facial feature points, to improve facial expression recognition accuracy.

The evaluation, which was carried out on multiple databases, including the Cohn-Kanade dataset, reveals that image ratio features outperform albedo and lighting variations. Furthermore, the combination of image ratio features and FAPs is a powerful approach, outperforming either feature set alone. The study also looks into asymmetric facial expressions, demonstrating the effectiveness of the combined recognition system on their own facial expression database [19]. Facial expression recognition plays a crucial role in nonverbal communication and has applications in diverse fields like gaming, criminal interrogations, psychiatry, and animations. Existing techniques for facial expression recognition primarily rely on either appearance features, which capture texture changes in the face, or geometric features, which analyze facial shape and components. In this context, the paper proposes an efficient and faster approach for facial expression recognition using deep convolutional neural networks (DCNN) with the Caffe framework on CUDA-enabled GPU systems. The method achieves state-of-the-art results on publicly available datasets, benefiting from the computational power of GPUs for feature extraction. The proposed model offers a versatile solution applicable to various facial expression recognition datasets without the need for retraining or extensive pre-processing. Future work involves exploring other pre-trained DCNN models like GoogLeNet for further advancements in this field [20]. This paper highlights the increasing importance of human emotion detection, particularly from facial expressions, in the context of modern artificial intelligence systems. Recognizing human emotions, whether from images, videos, or EEG signals, holds significant potential for applications ranging from identifying intent to enhancing security measures. The paper focuses on machine learning algorithms and feature extraction techniques as essential tools for accurate emotion identification. One notable aspect of the experiment is the emphasis on feature extraction. The addition of distance and area features led to promising results, particularly achieving an accuracy of 89% on the CK+ database. However, cross-database experiments revealed that raw features performed better with Logistic Regression, yielding accuracies of 66% and 36% for different datasets. The study's findings suggest that the proposed emotion detection algorithm demonstrated good generalization from training to testing sets, outperforming SVM and other algorithms. The achieved average accuracies of 86% for the RaFD database and 87% for the CK+ database in cross-validation scenarios indicate the effectiveness of the model [21]. This work focuses on the detection of user emotions based on facial expressions, utilizing either live camera feed or existing images. Emotion recognition from facial expressions is a well-established area in computer vision, and this study employs Python (2.7), OpenCV, and NumPy for implementation. By comparing the scanned testing image to a training dataset, the system predicts the person's emotion. The primary objective is to develop an image analysis system for accurate emotion prediction, demonstrating the feasibility of this approach and its validity. The study underscores the potential of Artificial Intelligence in addressing complex tasks like emotion detection, even though it can be challenging, especially with a large number of images. It acknowledges the inherent complexities and potential errors in recognizing emotions, with an achieved optimum accuracy of approximately 83%. This work discusses recommendation systems, a technology used by various companies like Facebook, Netflix, and Amazon to provide personalized recommendations to users. Two primary techniques for recommendation systems are collaborative filtering, which relies on user preferences and behavior, and content-based filtering, which uses extracted content information.

Another paper [22] presents a preliminary approach to Hindi music mood classification, focusing on the extraction of simple audio features encompassing rhythm, intensity, and timbre. The study employs the MIREX mood taxonomy and employs a decision tree classifier (J48) for classification. The reported average accuracy achieved with a 10-fold cross-validation is 51.56%. The paper suggests several future directions for research, including the incorporation of additional audio features to enhance mood classification results, the integration of song lyrics for multimodal mood classification, and the preparation of a larger audio dataset along with the collection of lyrics for further investigations in this domain. The authors [23] discuss personalized music recommendation services focusing on predicting user-favorite songs through the management of user preferences and music genre classification. The system described extracts user preference data from brainwaves and audio features from music, aiming to enhance recommendation accuracy. The study employs a short feature vector for genre classification, utilizing dimensionality reduction techniques. The proposed user's preference classifier achieves an 81.07% accuracy rate in binary preference classification for the KETI AFA2000 music corpus, demonstrating potential applications in audio devices, apps, and services. Deep convolutional neural networks (DCNN), a technology previously successful in bridging the semantic gap in visual tasks, are used in this paper to address the challenge of speech emotion recognition. The method entails extracting log Mel-spectrograms from speech signals, similar to how RGB images are represented, and then using a pre-trained AlexNet DCNN model to learn high-level features. These segment-level features are then aggregated for emotion classification using a discriminant temporal pyramid matching (DTPM) strategy. The study shows promising results across a variety of public datasets, highlighting the efficacy of the DCNN model and the DTPM strategy in speech emotion recognition. Notably, the pre-trained DCNN model from image applications extracts effective speech features well, with fine-tuning improving recognition performance even further [24].

Automated facial expression recognition systems have gained significant importance in various applications, including human-computer interaction and psychological research. These systems play a crucial role in assessing a person's affective state, cognitive activity, personality, intention, and psychological condition, enhancing interpersonal communication. This particular project focuses on real-time emotion analysis, with an emphasis on identifying depression among the six main emotions. It employs convolutional neural networks (CNNs) for emotion detection and includes a chatbot component to engage users and assess their emotional state, differentiating between sadness and depression. Detecting and addressing depression early is vital for preventing potentially harmful situations [25]. The work presented in [26] highlights the significance of music in influencing and enhancing human mood. It emphasizes the role of computer vision in automating the process of selecting music based on a user's emotional state, saving time and effort. The system utilizes facial expression recognition through algorithms like Haar Cascade and CNN to detect the user's mood in real time. It then suggests a playlist tailored to the detected emotion, providing an instant and appropriate music selection. This approach offers convenience and mood improvement through music, eliminating the need for manual song selection based on one's mood. This paper addresses the contemporary challenges in the field of face detection and emotion recognition, particularly in uncontrolled environments with varying poses, lighting conditions, and expressions. It emphasizes the crucial role of databases in facilitating face and emotion recognition systems. The proposed method efficiently creates feature databases for both face and emotion recognition. It employs the Viola-Jones face detection algorithm to detect faces in input images and extracts Mel frequency components from a human voice for emotion detection. The KNN classifier is utilized for face and emotion recognition, achieving high efficiency with reported recognition rates of 94.5% to 97% in experimental results [26]. The proposed emotion detection algorithm demonstrated good generalization from training to testing sets, outperforming SVM and other algorithms. The achieved average accuracies of 86% for the RaFD database and 87% for the CK+ database in cross-validation scenarios indicate the effectiveness of the mode [27].

Research [28-30] introduces a unique approach termed deep comprehensive multipatches aggregation convolutional neural networks (CNNs) to address the difficult job of facial expression recognition (FER) in computer vision. Using a deep framework and two branches of CNNs, this method pulls holistic information from the complete image of the face expression, while the other retrieves local features from image patches. Through pre-classification aggregation of both kinds of information, the model fully captures expressions at various scales. To make the model resilient to nuisance changes, a novel expressional transformation-invariant pooling technique is presented for training. Promising recognition results are obtained when the approach is tested on two datasets: the Japanese Female Facial Expression (JAFFE) database and the Extended Cohn-Kanade (CK+) dataset.Automated facial expression recognition systems have gained significant importance in various applications, including human-computer interaction and psychological research. These systems play a crucial role in assessing a person's affective state, cognitive activity, personality, intention, and psychological condition, enhancing interpersonal communication. This particular project focuses on real-time emotion analysis, with an emphasis on identifying depression among the six main emotions. It employs convolutional neural networks (CNNs) for emotion detection and includes a chatbot component to engage users and assess their emotional state, differentiating between sadness and depression. Detecting and addressing depression early is vital for preventing potentially harmful situations.

The existing techniques support either mood music or a chatbot individually, whereas we propose a combined application that supports both a chatbot and mood music.

3. System Overview

The suggested solution combines the Chat Bot and Mood Music modules to create an AI-enabled assistance system. Each module has a specific function, and together they provide a complete user experience.
Figure 1 shows a system overview which contains 2 main modules mood music and chatbot.

A. Mood Music Module: Based on the user's emotional state, the Mood Music module recommends music based on advanced facial recognition and mood categorization algorithms. The following are the main elements and features:
Component for facial recognition:
A live video feed or a photograph of the user's face taken with a camera serves as the input.
Processing: uses cutting-edge facial recognition algorithms (like OpenCV and Dlib) to find and examine facial characteristics, expressions, and landmarks. Identifies and extracts important face traits, such as brow position, mouth curve, and eye movement.
Component of mood categorization:
The user's emotional state is represented by the retrieved facial features as input.
Processing: the classification of the user's mood into predefined emotional states (such as happy, sad, or amorous) using machine learning models (e.g., SVM, Neural Networks) trained on labeled facial expression data.
Component for music recommendations:
The mood label from the Mood Categorization Component is an input.
Processing: accesses a collection of well-crafted music playlists that have been mood-tagged. Choose a playlist that is suited to the user's recognized emotional state. A suggested playlist for the user's listening is the output.

B. The Chat Bot module:
Acts as a conversational agent who is clever and capable of comprehending and resolving user issues in a variety of contexts. For efficient communication, it makes use of cutting-edge Natural Language Processing (NLP) algorithms. The following are the main elements and features: A component of natural language processing (NLP). Text or spoken user input is referred to as input.
Processing: uses NLP models to analyze and comprehend user questions and statements, such as Transformer-based models like BERT or GPT. From the input, extract pertinent data, intents, and entities. The parsed user query and related intents are output.
Component of issue categorization:
Input: The NLP Component's parsed user query and related intents.
Processing: Categorises user issues into distinct domains using established rules or machine learning models (such as educational, economic, and interpersonal concerns).
Contextual Guidance Component:
The issue domain is from the issue categorization component as input.
Processing: Accesses a knowledge base or database with information on recommendations and solutions for the selected problem domain. Depending on the user's unique problem offers recommendations and solutions that are contextually relevant.
Output: Customised solutions and suggestions for the user.

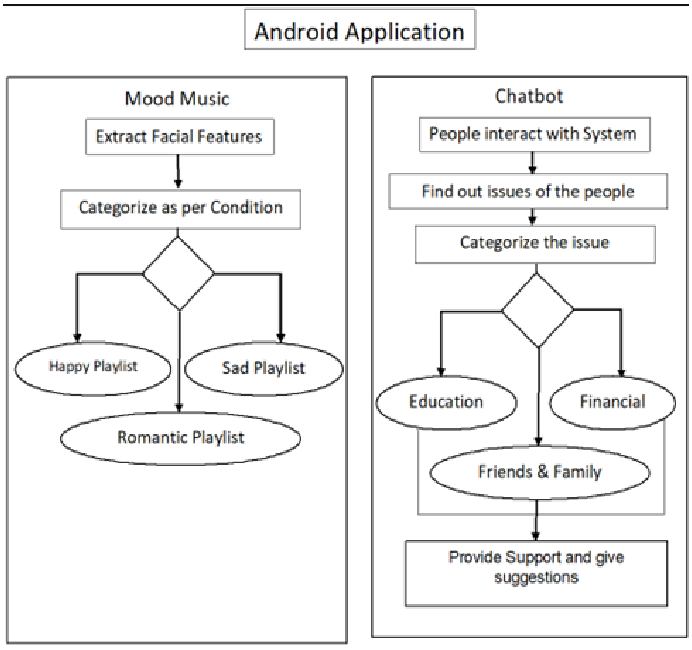

Figure 1. System Overview

4. Requirement Specification

The main goal of the application known as the mood-based music recommendation system is real-time mood recognition. It's a product prototype with two main modules: music suggestion and facial expression recognition/mood detection.

A. Mood Recognition Software

This Module is divided into two sections:

• Face Detection: The ability to locate a face in any input frame or image. The output is the bounding box coordinates of the discovered faces. For this assignment, the OpenCV Python Library was initially considered. However, because integrating it with an Android app would be difficult, the Java FaceDetector class was considered. This library counts the number of faces in a Bitmap visual object by identifying the people's faces within it.

• Detecting mood: classifying facial expressions as joyful, furious, depressed, neutral, surprised, afraid, or disgusted. The traditional Python Keras package was used for this assignment, but the survey revealed that this method works slowly when connected to Android apps and takes a long time to train and validate. As a result, MobileNet was used, which is a CNN architecture model for mobile vision and image classification. While other models exist, MobileNet is distinguished by its extremely low processing power requirements for operation and transfer learning applications. As a result, it can be used with mobile devices, embedded systems, and PCs that have low computing efficiency or no GPU without sacrificing accuracy. It constructs lightweight deep neural networks from depth-wise separable convolutions.

The training dataset was created by combining Kaggle's MMA Facial Expression Recognition dataset [7] and FER 2013 dataset [6]. The grayscale photos with 48x48 pixels in the FER 2013 collection. The MMA Facial Expression Recognition dataset contained images with a variety of characteristics. To create an even larger dataset with 40,045 training photos and 11,924 testing images, all of these images were converted by the images in the FER 2013 dataset and combined. Our model was trained and tested for seven classes using MobileNet and Keras: happy, furious, neutral, sad, surprised, afraid, and disgusted. After 25 epochs of training, we were able to get an accuracy of about 75%.

B. The Module for Recommending Music

You can find the Hindi and English versions of the mood-categorized music dataset on Kaggle. To store, retrieve, and query this song data upon user request, research was done to find a reliable cloud storage platform. Although alternatives such as AWS, Google Cloud, and so on were discovered, they were turned down because of their high cost and meagre nature of free storage offerings. Next, look into free and open-source streaming services such as Ampache, Restream.io, etc. To hold the labels.txt and.tflite files, an assets folder was made in Android Studio. The model's class labels are contained in the labels.txt file. For loading the model, executing the interpreter, and getting the results, all the necessary methods were developed. MP3 tracks were uploaded to the storage part of the Firebase project. These songs are arranged in the real-time database section based on language and mood. The Firebase database was then connected to Android Studio. The tflite model techniques were integrated with the Firebase music, and a suitable user interface was developed for the Android application. Ultimately, the program was tested to ensure that any errors were fixed.

The physical computer resources, also known as hardware, are the most common set of requirements defined by any operating system or software application. The following hardware is required for this project: • A minimum of 4 Gigabytes (GB) of RAM (for processing) • Webcam (for laptop/desktop testing) • Camera with a minimum resolution of 16 Megapixels (MP) (for testing on Android devices) • 30 MB (approximate) memory space.

Software Requirements are concerned with defining the software resource requirements and prerequisites that must be installed on a computer in order for an application to function properly. These prerequisites or requirements are typically not included in the software installation package and must be installed separately before the software can be installed. The following software requirements are required for this project: With Python 3.6 mood music module, the moods are categorized into various mood like happy, sad, and romantic. The mood is recognized according to the facial expressions and then will match with the appropriate mood and the music will be suggested accordingly. In the second module chatbot, it is based on the interaction with the system. The user communicates with the system, discuss their issues with the chatbot and then the issues are categorized in various sectors such as educational, financial, or friends and family. The issues are analyzed and appropriate advice suggested to the user.

5. Results and Discussion

The Philoi app features a simple, clear, and user-focused interface created to encourage an easy and sympathetic user experience. The interface is distinguished by its serene color scheme, with subtle, gentle blues and greens that evoke a sense of peace and tranquillity.

The home page:
When users start the app for the first time, they are welcomed by a welcoming home screen. Customers are presented with two noteworthy options in this case: the two primary Philoi modules, "Mood Music" and "Chatbot" as shown in figure 2.

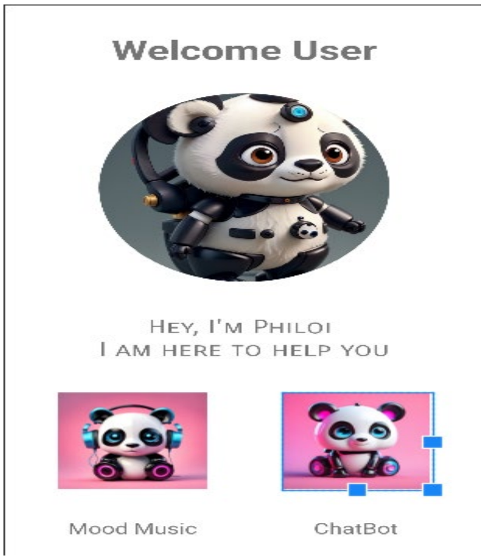

Figure 2. Home screen

The Philoi app boasts a clean, intuitive, and user-centric interface designed to promote a seamless and empathetic user experience. The interface is characterized by its calming color palette, with soft, muted tones of blues and greens, evoking a sense of tranquillity and serenity.

I. Home Screen:

Upon launching the app, users are greeted by a welcoming home screen. Here, they are presented with two prominent options: 'Mood Music' and 'Chatbot', the core modules of Philoi. Figure 3 shows the Philoi User Interface of the application.

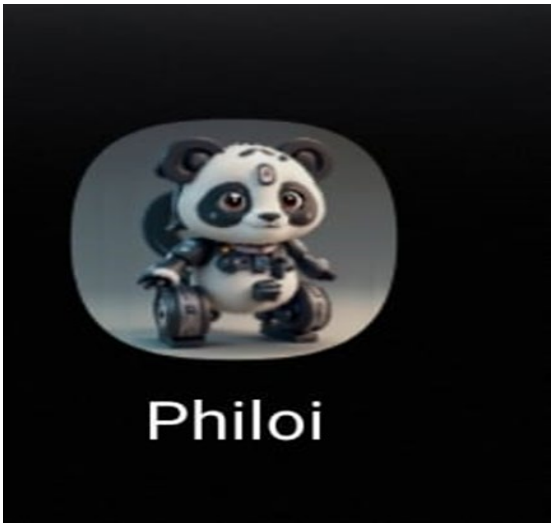

Figure 3. UI of the application

Figure 4 depicts the User Interface of Mood Music, a system that assesses the user's mood based on an emoji input and plays a song corresponding to the user's mood.

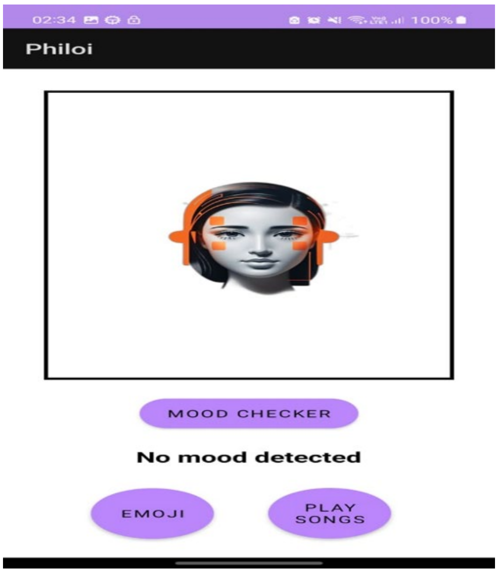

Figure 4. UI of mood music

Figure 5 illustrates the facial expression associated with anger, triggering the Philoi app to play appropriate songs.

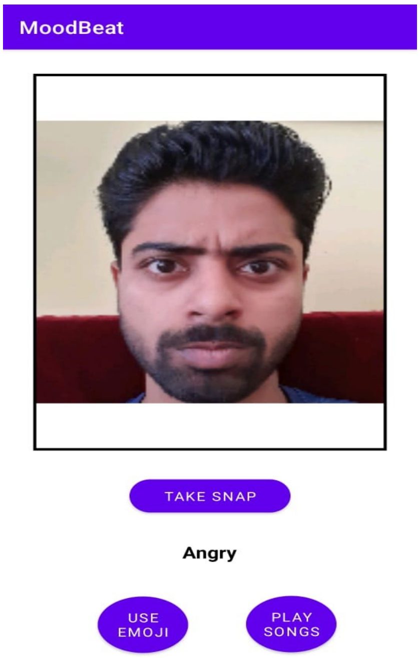

Figure 5. Angry mood

Figure 6 demonstrates the detection of a happy mood and the subsequent music playback based on that mood.

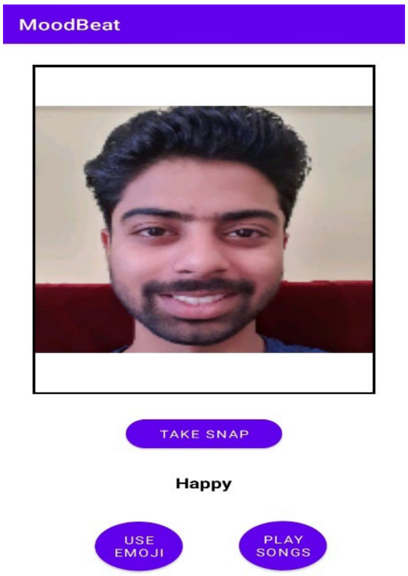

Figure 6. Happy mood detection

II. Mood Music Module
When a user chooses a playlist, the Now Playing page displays the selected music along with options for volume adjustment, playing controls, and a progress meter.

Following real-time mood identification, the system presents a playlist for each mood with accuracy. The "happy" mood playlist is displayed in Figure 8. It is capable of downloading and playing the music from the recommended playlist if there is a consistent internet connection.

Figure 7. Happy mood playlist

The user may examine the songs in the playlist in list view, and when they choose a song to play, it opens in a media player with all the standard controls—play, pause, shuffle, next, and previous—as well as a seek bar as shown in figure 7. The user is presented with devotional, motivating, and patriotic music suggestions for the following moods: furious, fearful, disgusted, and surprised. Therefore, the user also receives mood enhancement.

III. Chatbot Module

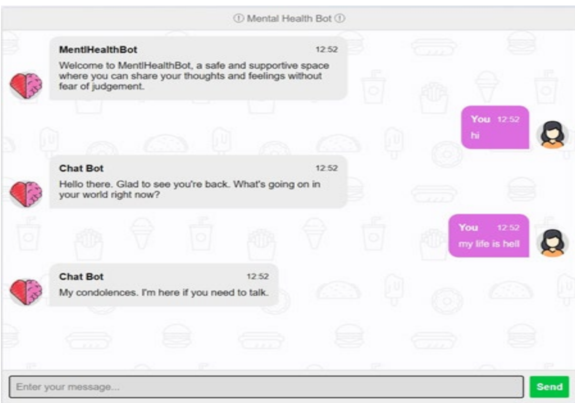

Figure 8. Chatbot Module

Figure 8 displays the chatbot module, where users can engage in conversation and share information about themselves, with the Philoi app offering suitable solutions after analyzing the problem.

Figure 9. Screen for a user to generate emotions

Figure 9 shows different emotions that the user selects to express his or her emotions

The user can click the "Use Emoji" button and choose the emoji that best captures their current mood or the mood they want their playlist to be created from whenever they don't want to or are unable to take a snapshot of their mood for a variety of reasons, such as extremely high or low lighting, a malfunctioning camera, a lower resolution camera that makes it difficult to take a clear picture of their face or any other reason. The snapshot of the screen that appears to the user when they click the "Use Emoji" button is shown in Figure 9. The emoticon representing the mood is the first.

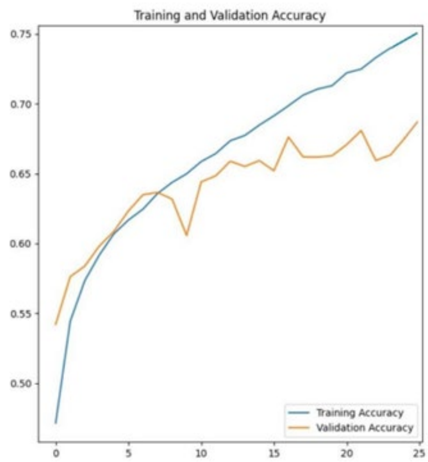

Figure 10. Graph for accuracy

The accuracy of our model is shown in graph Fig. 10, where the y-axis indicates accuracy and the x-axis indicates the number of epochs. The figure illustrates the nearly 75% accuracy that our model has attained. Being a completely computer-based system, it has been trained to comprehend emotions in a certain way.

6. Conclusion

With its seamless fusion of technology and compassion, the Philoi app represents a significant turning point in the evolution of mental health assistance. It offers a comprehensive and accessible platform for people navigating the complexity of their emotional well-being. The revolutionary potential of Philoi to improve mental health outcomes has been revealed through a careful analysis of its Mood Music and Chatbot modules. The Mood Music module, which was inspired by actual studies on the psychological impacts of music, is evidence of the healing potential of sound. Users of Philoi are given the ability to tap into the therapeutic potential of music, providing comfort and relief during trying times by providing expertly produced playlists catered to a range of emotional states. Additionally, the Chatbot module acts as a guiding light by offering a private and understanding environment for communication. The chatbot, which is based on cutting-edge natural language processing, engages users in deep dialogue while providing advice, coping mechanisms, and resources with a degree of understanding that extends beyond the boundaries of the digital world.

The user experience of the app, which is distinguished by its peaceful aesthetics and simple design, promotes a feeling of connection and trust. Users from diverse backgrounds can interact with Philoi in a way that best meets their unique needs and preferences thanks to the intuitive navigation and accessibility features. It is important to note, nevertheless, that while Philoi represents a tremendous advancement in mental health intervention, it is not a substitute for qualified counseling or medical care. It ought to be seen as an additional instrument in a wider range of care possibilities.

The Philoi app, which provides a light of hope and understanding to people in need, is a prime example of how technology has the power to completely transform the way that mental health treatment is provided. The app has the potential to revolutionize the field of mental health interventions as it develops and adapts. Philoi's integration of technology, empathy, and evidence-based strategies lays the groundwork for a time when mental health support is widely available and characterized by deep empathy. In the end, Philoi's influence goes beyond the limitations of a single application; it serves as an example of how human-centric technology has the potential to improve mental health, one person at a time.

References

1. Zhang, Ligang, and Dian Tjondronegoro. "Facial expression recognition using facial movement features." IEEE transactions on affective computing 2.4 (2011): 219-229.
2. Viola, Paul, and Michael J. Jones. "Robust real-time face detection." International journal of computer vision 57 (2004): 137-154.
3. Jeong, Mira, and Byoung Chul Ko. "Driver's facial expression recognition in real-time for safe driving." Sensors 18.12 (2018): 4270.
4. Song, Mingli, et al. "Image ratio features for facial expression recognition application." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 40.3 (2009): 779-788.
5. Mayya, Veena, Radhika M. Pai, and MM Manohara Pai. "Automatic facial expression recognition using DCNN." Procedia Computer Science 9.3 (2016): 453-461.
6. Pantic, Maja, and Leon JM Rothkrantz. "Facial action recognition for facial expression analysis from static face images." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 34.3 (2004): 1449-1461.
7. Pantic, Maja, and Leon JM Rothkrantz. "Expert system for automatic analysis of facial expressions." Image and Vision Computing 18.11 (2000): 881-905.
8. Rizwan Ahmed Khan, Alexandre Meyer, Hubert Konik, Saïda Bouakaz. Framework for reliable, realtime facial expression recognition for low-resolution images" Pattern Recognition Letters 34.10 (2013): 1159-1168.
9. Happy, S. L., and Aurobinda Routray. "Automatic facial expression recognition using features of salient facial patches." IEEE transactions on Affective Computing 6.1 (2014): 1-12.
10. Déniz, Oscar, et al. "Face recognition using histograms of oriented gradients." Pattern recognition letters 32.12 (2011): 1598-1603.
11. Raut, Nitisha. "Facial emotion recognition using machine learning." SJSU ScholarWorks. (2018).
12. Puri, Raghav, et al. "Emotion detection using image processing in python." arXiv preprint arXiv:2012.00659 (2020).
13. Kaufman, Jaime C. "A Hybrid Approach to Music Recommendation: Exploiting Collaborative Music Tags and Acoustic Features." (2014).
14. Patra, Braja Gopal, Dipankar Das, and Sivaji Bandyopadhyay. "Automatic music mood classification of Hindi songs." Proceedings of the 3rd Workshop on Sentiment Analysis where AI meets Psychology. (2013).
15. Lee, Jongseol, et al. "Music Recommendation System Based On Geenre Distance And User Preference Classification." Journal of Theoretical & Applied Information Technology 96.5 (2018):1-12.
16. Zhang, Shiqing, et al. "Speech emotion recognition using deep convolutional neural network and discriminant temporal pyramid matching." IEEE Transactions on Multimedia 20.6 (2017): 1576-1590.
17. Yang, Yi-Hsuan, and Homer H. Chen. "Ranking-based emotion recognition for music organization and retrieval." IEEE Transactions on audio, speech, and language processing 19.4 (2010): 762-774.
18. Xie, Siyue, and Haifeng Hu. "Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks." IEEE Transactions on Multimedia 21.1 (2018): 211-220.

19. Neha, S., et al. "Emotion recognition and depression detection using deep learning." (2020): 3031-3036.
20. Geetanjali Mate et.al., “Mood Detection with Chatbot using AI-Desktop Partner”, International Journal of Advanced Science and Engineering 1.1 (2023), 1-14.
21. Kamita, T., Ito, T., Matsumoto, A., Munakata, T., & Inoue, T. A chatbot system for mental healthcare based on SAT counseling method. Mobile Information Systems,1.2 (2019), 1–11.
22. Følstad, A., Skjuve, M., & Brandtzaeg, P. B. Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design. In Internet Science. Springer International Publishing 11.5 (2019). 145–156.
23. Park, S., Choi, J., Lee, S., Oh, C., Kim, C., La, S., Lee, J., & Suh, B. Designing a chatbot for a brief motivational interview on stress management: Qualitative case study. Journal of Medical Internet Research, 21.4, (2019). 55-72.
24. Raut, N. Facial emotion recognition using machine learning. San Jose State University. (2019).
25. Abdul, A., Chen, J., Liao, H.-Y., & Chang, S.-H. An emotion-aware personalized music recommendation system using a convolutional neural networks approach. Applied Sciences (Basel, Switzerland), 8.7, (2018) 1103-1112.
26. Kamita, T., Ito, T., Matsumoto, A., Munakata, T., & Inoue, T. (2019). A chatbot system for mental healthcare based on SAT counseling method. Mobile Information Systems,5.3 (2019), 1–11.
27. Sepahpour, T. Ethical considerations of chatbot use for mental health support. Johns Hopkins University. (2020).
28. Følstad, A., Skjuve, M., & Brandtzaeg, P. B. Different chatbots for different purposes: Towards a typology of chatbots to understand interaction design. In Internet Science, Springer International Publishing.36-4 (2019).,145-156.
29. Park, S., Choi, J., Lee, S., Oh, C., Kim, C., La, S., Lee, J., & Suh, B. (2019). Designing a chatbot for a brief motivational interview on stress management: Qualitative case study. Journal of Medical Internet Research, 21. 4,(2019), 1-14.
30. Amrita Nair, Smriti Pillai, Ganga S Nair, Anjali T, “Emotion Based Music Playlist Recommendation System using Interactive Chatbot”, 6th International Conference on Communication and Electronics Systems (ICCES), (2021).