NILKAMAL SCHOOL OF MATHEMATICS, APPLIED STATISTICS & ANALYTICS

Exploring Technological Solutions to Improve Accessibility and Affordability of Mental Health Support Services, Through Facial Expression Analysis

Bhavesh Pashte | Vinayak Khithani | Purva Burundkar

A001 | A019 | A018

Guided By Prof. Dr. Yogesh Naik

Abstract

The research seeks to enhance the accessibility and affordability of mental health support services by utilizing facial expression analysis. Through the fusion of technology and empathy, the study aims to dismantle barriers associated with stigma and cost, fostering a robust support network to improve individuals' well-being. Positioned within the realm of applied research, the study devised five pairs of alternative hypotheses to explore the impact of high stress on individuals' performance and physical health, as well as financial constraints as a hindrance. Data collection involved a survey distributed via Google Forms, yielding 51 responses. Statistical analyses and visualizations were performed using SPSS and PowerBI, encompassing feasibility, validity, and significance testing. Findings reveal that financial constraints and social stigma act as barriers to accessing mental health support, while high stress significantly affects individuals' performance and physical health, as supported by Pearson's Chi-square Asymptotic significance levels are below 0.05 for all null-alternative hypothesis pairs. In summary, there is a clear need for affordable and easily accessible mental health support. To address this need, the study developed a web application utilizing a VGG16 pre-trained model for facial expression analysis, followed by a PHQ9 questionnaire.

1 Introduction

Mental health disorders pose a significant global burden, with one in four adults experiencing a mental health condition annually. Despite recent technological advancements contributing to our understanding of mental health, a crucial gap persists in translating these insights into accessible and affordable support systems, bridging identification and intervention. This research proposal presents a multi-faceted platform integrating evidence-based practices and personalized resources to empower individuals on their journey toward psychological well-being.

Recent advancements in AI-powered emotion recognition models have enhanced comprehension of mental health by decoding facial expressions. However, translating these insights into accessible and affordable support remains a missing link. The proposed platform offers tailored self-soothing techniques as calming strategies, coupled with personalized evidence-based resources to navigate emotional landscapes. For additional guidance, multiple support levels, including access to mental health professionals through affordable subscription plans, are envisioned.

A core tenet is providing an accessible and affordable technological solution, collaborating with healthcare providers, insurers, and employee wellness programs to offer the platform as a service, transcending financial limitations. The methodology encompasses surveying challenges faced by those with mental health issues, reviewing literature to refine the approach, and pilot testing for accuracy and effectiveness.

Ultimately, this research aims to amplify the voices of those struggling with mental health issues, transforming them into a narrative of hope. By integrating technology with empathy, the aspiration is to foster an environment where seeking support is collaborative and accessible, breaking down barriers of stigma and cost, promoting psychological well-being.

2 Literature Review

The literature on emotional health ("Facial Expression Recognition Utilizing Local Direction-Based Robust Features and Deep Belief Network") [1], particularly in the context of the elderly population, underscores the critical role it plays in enhancing the overall quality of life. Negative emotional states among individuals, if left unaddressed, can contribute to various social and mental health challenges. Facial expressions serve as a key medium for daily communication, and recent years have witnessed significant efforts in developing reliable FER systems. The FER system comprises three fundamental components: face image pre-processing, feature extraction, and recognition. The literature also highlights the significance of FER as the most natural form of human emotion expression, making it a prominent research area for computer vision and image processing applications.

Methodology used is a novel method for feature extraction known as the local directional position pattern (LDPP). This method utilizes depth video data for person-independent expression recognition, where pixel intensities are distributed based on their distances to the camera. To further enhance the capabilities of LDPP, the authors integrate principal component analysis (PCA) and generalized discriminant analysis (GDA). These additional steps contribute to better illustrating facial characteristics in expression. The proposed features, enriched through LDPP, PCA, and GDA, are subsequently applied to a deep belief network (DBN) for expression training and recognition.

Potential research gaps could be real world implementation by exploring the practicality of deploying the system in healthcare settings, gauging user satisfaction. Exploring the ethical dimensions, including issues related to consent, data security, and potential misuse of facial expression data, especially in healthcare contexts.

The paper ("Deep Convolution Network Based Emotion Analysis Towards Mental Health Care") [2] unveils an innovative emotion analysis framework rooted in deep convolution networks, aiming to craft mental health systems that are not only effective but also user-friendly and budget-friendly. By tapping into the power of the AlexNet architecture, especially honing in on the nitty-gritty of the Fully Connected Layer 6, the proposed method stands out by beating the competition in both accuracy and efficiency. Thorough testing across diverse databases like JAFFE, KDEF, CK+, FER2013, and AffectNet solidifies the model's reliability. Despite these achievements, the existing literature highlights the ongoing need for studies that dive into cultural and individual differences, model clarity, and the ethical dimensions of rolling out automated facial expression analysis tools for mental health evaluation.

Depression is a significant mental psychological disorder with significant physical and social-economic impacts. A recent study ("An Automated and Real-time Approach of Depression Detection from Facial Micro-expressions") [3] by Ghulam Gilanie et al. proposed an automated and real-time approach for depression detection from facial micro-expressions. The light-weighted CNN-based model demonstrated robust performance, even with low-resolution video data. This approach presents a promising avenue for early intervention and support for individuals experiencing depression.

The study uses head movements, facial dynamics, and vocal parody to assess depression severity. The use of Z-Face technology for 3D registration from 2D video and stacked Denoising Auto-encoders (SDAE) for encoding facial and head movements showcases the evolving technological landscape in depression detection research. The proposed approach aligns with the global public health concern regarding depression, as highlighted by Marcus et al. and Zhu et al.'s improved classification model for depression detection using EEG and eye tracking data. The research contributes to the growing body of literature on automated and real-time depression detection, offering a promising avenue for advancing mental health research and clinical practice

The literature ("Personalized models for facial emotion recognition through transfer learning") [4] highlights the advancements in facial emotion recognition through the integration of transfer learning and subject-specific data. The study provides insights into the differential impact of valence and arousal dimensions and explores the potential of active learning algorithms for efficient model training.

The methodology of this study proposes a transfer learning approach for regression, utilizing a CNN (AlexNet) trained on a large in-the-wild dataset (AffectNet). The results emphasize the necessity of both transfer learning and subject-specific data, showcasing improved Root Mean Square Error (RMSE) values compared to alternative methods such as evaluating pre-trained networks or training with random weight initialization. Transfer learning has emerged as a promising solution to address the challenges associated with personalized emotion recognition models. This approach involves leveraging knowledge gained from a large multi-subject dataset through a deep-convolutional neural network (CNN) to extract emotional content in single-subject scenarios.

The landscape of mental health diagnosis is evolving with the integration of artificial intelligence (AI), particularly in the context of diagnosing depressive disorders. The existing literature ("Diagnosis of Depressive Disorder Model on Facial Expression Based on Fast R-CNN") [5] underscores the significance of exploring innovative approaches to supplement traditional self-report questionnaires, enhancing the diagnostic process by bridging the gap between subjective and objective assessments. Smartphone-based interventions emerge as a pivotal tool, offering swift identification of depressive disorders and generating valuable data for timely interventions. The study introduces a novel methodology employing fast region-based convolutional neural networks (R-CNN) within the domain of deep learning. This method focuses on tracking positional changes in key facial features, specifically the eyes and lips, to infer emotions from accumulated participant photos. The literature supports the exploration of such AI-driven models, as they hold the potential to revolutionize depressive disorder diagnostics, ensuring prompt and reliable assessments that contribute to effective intervention strategies.

The identified research gap in the above-mentioned papers revolves around mental health intervention. While these papers touch upon various aspects of mental health, they lack in-depth exploration or comprehensive analysis of effective strategies and interventions aimed at addressing mental health issues. There is a need for more robust investigations into the practical implementation of mental health interventions, evaluating their efficacy, and understanding the factors that contribute to successful mental health outcomes. This research gap highlights the necessity for studies that not only recognize the importance of mental health but also delve into the development, application, and assessment of intervention programs to support individuals facing mental health challenges.

3 Methodology

This research is classified as Applied Research as it aims to develop a practical solution – a web application – to address the problem of barriers in seeking professional help for depression. The application will leverage data analysis and machine learning techniques to provide insights and potentially bridge the gap between struggling individuals and appropriate resources. Based on the identified barriers, the research will formulate specific hypotheses about the relationship between these barriers and the likelihood of seeking professional help. Statistical analysis will be employed to test these hypotheses and assess their validity.

3.1 Convolutional Neural Network (CNN)

A convolutional neural network (CNN) falls under the category of machine learning models, specifically a type of deep learning algorithm designed for the analysis of visual data. Also known as convnets, CNNs leverage principles from linear algebra, particularly convolution operations, to extract features and recognize patterns within images. While their primary application is image processing, CNNs can be modified to handle other types of data, including audio and various signal data. The architecture of CNNs draws inspiration from the connectivity patterns observed in the human brain, especially the visual cortex, which is responsible for perceiving and processing visual stimuli. The artificial neurons within a CNN are strategically organized to efficiently interpret visual information, allowing these models to process entire images. Due to their proficiency in object identification, CNNs are commonly employed in computer vision tasks such as image recognition and object detection.

CNNs employ several layers, each of which picks up unique characteristics from an input image. A CNN may have hundreds, thousands, or even more layers, depending on how complicated the task for which it is designed. Each layer builds on the outputs of the one before it to identify intricate patterns. Three general categories can be used to aggregate CNN layers: convolutional, pooling and full connected layers. The CNN's complexity rises as data moves through these layers, enabling it to detect progressively more abstract characteristics and greater areas of a picture.

The convolutional layer is the core component of a CNN and is where most of the processing takes place. In order to find particular features in an input image, this layer moves across its receptive field using a filter, or kernel, which is a small matrix of weights. Initially, the kernel is slid across the width and height of the image, and after several rounds, it is finally swept across the complete image. A dot product is computed at every place between the weights of the kernel and the image's pixel values underneath the kernel. This process converts the original image into a collection of convolved features, also known as feature maps, each of which shows the existence and intensity of a different feature at different locations throughout the image. In many CNNs, there is a configuration of numerous stacked convolutional layers. This layered structure enables the CNN to systematically analyse the visual details present in the original image data. In the initial layers, the CNN recognizes fundamental features like edges, textures or colours. As the data progresses through the network, deeper layers receive input from the feature maps of preceding layers, empowering them to identify more intricate patterns, objects, and scenes.

Following the convolutional layer in a CNN, the pooling layer plays a crucial role. While it shares a sweeping process across the input image with the convolutional layer, its primary function differs. The pooling layer is designed to decrease the dimensionality of the input data while preserving essential information, thereby enhancing the overall efficiency of the network. This reduction is commonly achieved through downsampling, involving a decrease in the number of data points in the input, often translating to a reduction in the pixels used to represent the image. A prevalent technique within pooling is max pooling, where the maximum value within a specified window (i.e., the kernel size) is retained, and other values are discarded. Another common technique, knowns as average pooling, takes a similar approach but uses the average value instead of the maximum. Down sampling results in a significant reduction of the total number of computations and parameters. This enhances efficiency and fortifies the model's capacity for generalization. Higher-level features in simpler models tend to make them less prone to overfitting. There is a possible drawback to shrinking the representation's spatial size, which is the possibility of losing some information. For tasks like object detection and image classification, however, it is frequently sufficient to learn only the salient aspects of the input data.

In the last stages of a CNN, the fully connected layer is essential because it classifies images by using the attributes that were retrieved from the earlier layers. When a neuron in one layer is said to be fully linked, it signifies that every other neuron in the layer below it is connected as well. The different features that were extracted from the preceding convolutional and pooling layers are integrated and mapped to particular classes or results by the fully connected layer. In the fully connected layer, every input from the preceding layer is coupled to every activation unit, allowing the CNN to evaluate all features at once before classifying a sample. Reducing the total number of fully linked layers strikes a compromise between the capacity to learn intricate patterns and computational efficiency and generalization.

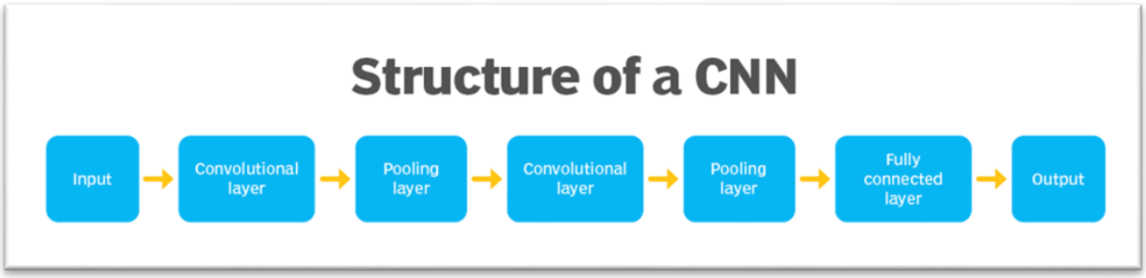

Figure 1: Structure of CNN

In the initial phase of our research, we constructed a seven-layered convolutional neural network that gave us a validation accuracy of 65% and exhibited inconsistency during the testing phase. Hence, we recognized the need for a more robust solution and we moved to the VGG16 pre-trained model, which is a 16-layer convolutional neural network. This re-usage of weights from a pre-trained model for our own model is called Transfer Learning. Briefly, we will discuss the key characteristics of VGG16 model. VGG16 is characterized by its depth, consisting of 16 convolutional and fully connected layers. The extensive depth of the network enables it to capture intricate features in images, allowing for high-quality classification results. The network primarily comprises a series of convolutional layers, each equipped with 3x3 filters and a stride of 1. This consistent filter size and stride contribute to the network's simplicity and ease of interpretation. After each set of convolutional layers, VGG16 incorporates max-pooling layers with 2x2 filters and a stride of 2. Pooling layers serve to down sample the feature maps, reducing spatial dimensions while retaining essential information. Towards the end of the architecture, VGG16 includes three fully connected layers with 4096 neurons each, followed by a final output layer with 1000 neurons corresponding to the number of classes in the ImageNet dataset, on which it was originally trained. Throughout the network, rectified linear unit (ReLU) activation functions are used to introduce non-linearity, enabling the model to learn complex patterns in the data. To mitigate overfitting, dropout layers are incorporated after the fully connected layers, randomly dropping a fraction of neurons during training to encourage the network to learn more robust features.

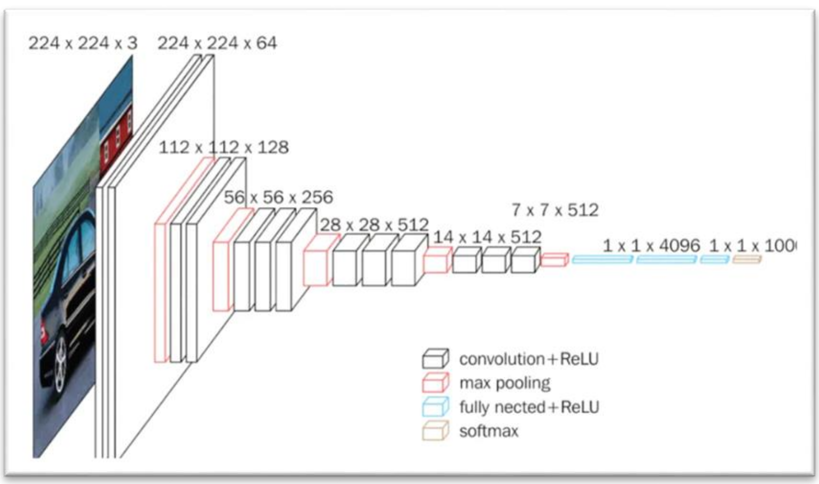

Figure 2: Architecture of VGG16 model

3.2 PHQ-9 Questionnaire

The PHQ-9 is a brief, 9-question self-assessment tool specifically designed to screen for and measure the severity of depression. Developed by Dr. Robert K. Spitzer, Dr. Janet B.W. Williams, and Dr. Kurt Kroenke in 1999, it has become a widely used and important tool in both clinical and research settings. Compared to other questionnaires, the PHQ-9 is shorter and easier to administer than many other depression assessments, it focuses specifically on depression symptoms, while other questionnaires may include questions about other mental health conditions, the PHQ-9's scoring system is straightforward and readily interpretable.

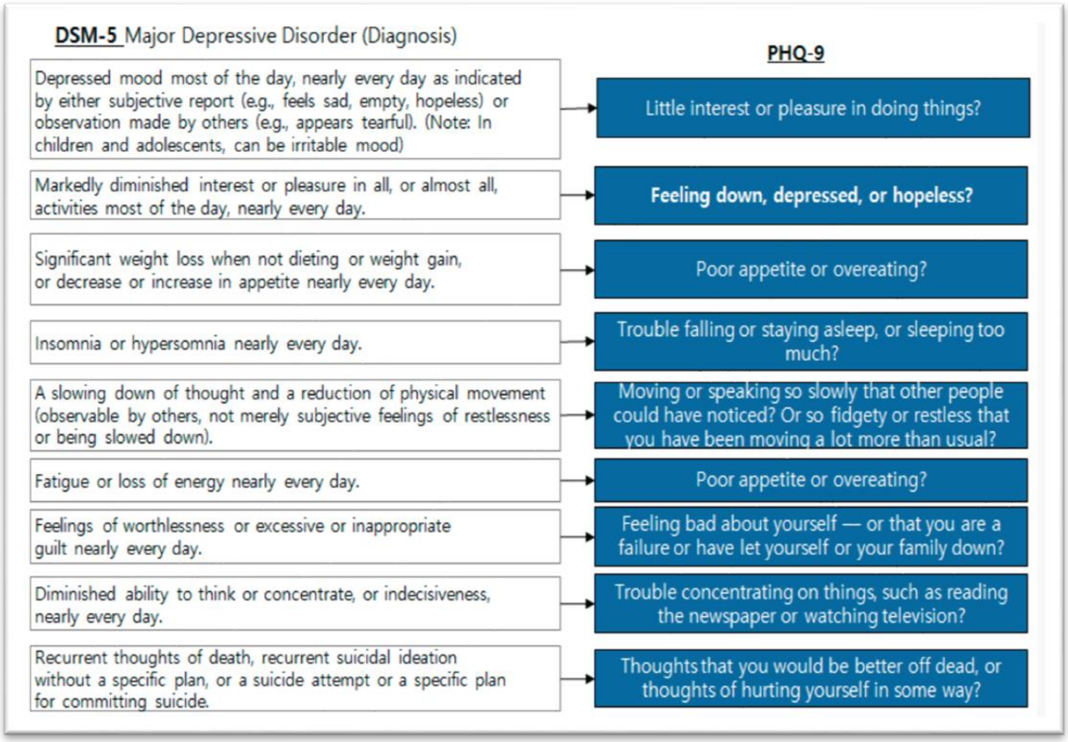

Figure 3: PHQ-9 Questionnaire

3.3 Web Application

Following hypothesis testing, the research will shift towards developing the web application. On landing the homepage of web application, the application will capture the user's photo and utilize deep learning techniques Convolutional Neural Networks (CNNs) to analyse facial expressions and provide an emotional assessment. The publicly available dataset FER2013 [7] will be utilized for the purpose of training the model. After the application has captured users' facial emotion the users will be given a PHQ-9 [6] questionnaire. The application will combine the PHQ-9 score and the facial expression analysis to provide a comprehensive output for the user. This output could include:

- Depression Severity Indication: Providing a clear and understandable indication of the user's potential depression severity based on both questionnaire and facial analysis data.
- Resource Recommendations: Based on the user's output, the application will recommend relevant resources and support options, such as hotlines, mental health professionals, or self-help materials.
- Personalized Guidance: The application could offer personalized guidance and tips based on the user's specific needs and identified barriers, potentially including motivational messages, coping strategies, and steps to overcome help-seeking obstacles.

The developed application will undergo rigorous testing with target users to assess its usability, effectiveness, and potential impact. User feedback and data collected from the application's use will be continuously analysed to refine the features and algorithms, ensuring the application remains relevant and helpful for its intended audience. The research will adhere to strict ethical guidelines regarding informed consent, data privacy, and confidentiality.

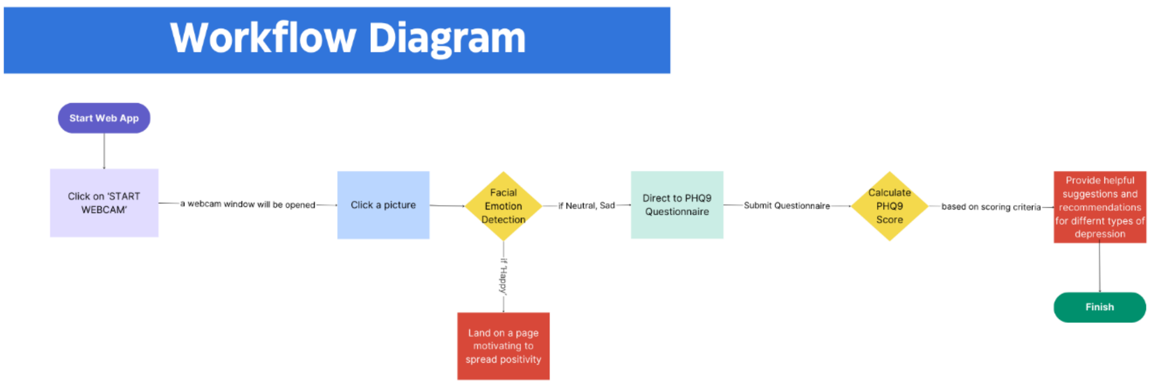

Figure 4: Workflow of the web application

4 Hypotheses Development

4.1 Hypothesis 1: Assessing the impact of high stress levels on the performance of individuals.

The aim of the hypothesis is to rigorously examine whether elevated stress levels have an impact on performance across various domains such as academics, office and other working environments.

- Null hypothesis (H0): High stress levels do not affect our performance in academics, office or any other working environment.
- Alternate hypothesis (H1): High stress levels affect our performance in academics, office or any other working environment.

4.2 Hypothesis 2: Assessing the fear of social judgment on individuals.

The primary objective is to investigate whether individuals experience hesitation in seeking help for mental health issues specifically due to a fear of social stigma.

- Null hypothesis (H0): An individual does not hesitate to seek help for mental health issues due to fear of social stigma.

- Alternate hypothesis (H1): An individual hesitates to seek help for mental health issues due to fear of social stigma.

4.3 Hypothesis 3: Assessing the impact of financial constraint as a barrier.

The aim of the hypothesis is to explore whether financial constraints pose a significant obstacle to individuals seeking support from professionals.

- Null hypothesis (H0): Financial constraint does not become a hurdle for finding support from professionals.
- Alternate hypothesis (H1): Financial constraint becomes a hurdle for finding support from professionals.

4.4 Hypothesis 4: Assessing the impact of high stress levels on physical health.

The primary objective is to investigate whether elevated stress levels have a substantial impact on physical health.

- Null hypothesis (H0): High stress levels do not affect our physical health.
- Alternate hypothesis (H1): High stress levels affect our physical health.

4.5 Hypothesis 5: Assessing the impact of soothing activities on mental health.

The objective of the hypothesis is to thoroughly investigate and determine whether engagement in soothing activities or hobbies exerts a substantial influence on aiding an individual in transitioning out of a high stress period.

- Null hypothesis (H0): Engaging in soothing activities or hobbies does not assist a person in transitioning out of the high stress period.
- Alternate hypothesis (H1): Engaging in soothing activities or hobbies assists a person in transitioning out of the high stress period.

5 Data Collection

In the pursuit of hypothesis testing, a meticulous approach to data collection was adopted through the utilization of Google Forms surveys. This online survey methodology was selected for its user-friendly interface, accessibility, and robust features that facilitate streamlined data gathering. The survey instrument was meticulously designed, employing a Likert scale to capture responses from participants. The formulated questionnaire represents a comprehensive and carefully crafted set of inquiries aimed at extracting detailed insights relevant to the research hypotheses. The Likert scale, chosen for its ability to measure degrees of agreement or disagreement on a continuum, provides a nuanced and graded understanding of participants' perspectives. This strategic approach is poised to yield a wealth of detailed information essential for robust hypothesis testing and the generation of meaningful insights in the research domain. The following is the drafted questionnaire:

Q1. Does emotional turmoil have a serious impact on an individual’s performance in academics/office/other working environments?
(1- Strongly disagree, 5 - Strongly agree)

Q2. Does fear of social stigma become a barrier to people seeking mental health support?

Q3. How significant is the impact of financial constraints on people’s ability to access adequate professional support?

Q4. Does high stress have a vital effect on an individual’s physical health?

Q5. Do you think stress levels decrease noticeably after participating in soothing activities or hobbies?

Q6. How would you rate your perceived stress level on a daily basis?

These were the six different questions that were formulated based on all the hypotheses. The objectives behind these questions were to comprehensively assess the multifaceted influences on an individual’s well-being, particularly focusing on their emotional, social, financial and physical dimensions. The first question aimed to gauge the correlation between emotional turmoil and academic or professional performance, emphasizing the potential impact of emotional states on daily functioning. The second question delved into the social aspects, exploring whether the fear of social stigma acts as a barrier to seeking essential mental health support. The third question addressed the influence of financial constraints on an individual’s access to professional support, acknowledging the potential challenges associated with seeking assistance. The fourth question explored the critical connection between high stress levels and their adverse effects on physical health, recognizing the intricate relationship between mental and physical well-being. The fifth question sought to understand the potential impact of engaging in soothing activities or hobbies on stress levels, exploring the role of personal pursuits in stress management. Together, these questions aim to provide a comprehensive understanding of the factors influencing stress levels, encompassing emotional, social, financial and physical dimensions.

5.1 Demographics of the data.

From the Google Forms survey distributed, a total of 51 responses were collected, reflecting the valuable insights shared by diverse participants. The demographic breakdown of the collected data is presented below, offering a comprehensive overview of the characteristics of the respondents.

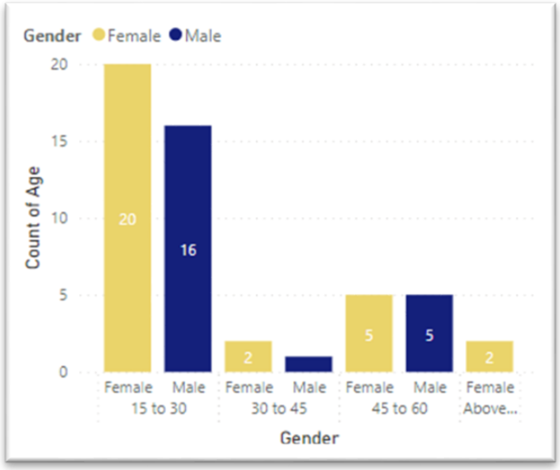

Figure 5: Demographic Visualisation of Responses

The above graph illustrates the distribution of responses across various age groups, further segmented based on gender, specifically distinguishing between male and female participants. This comprehensive categorization enables a nuanced examination of the survey data, offering insights into diverse perspectives and experiences within different demographic cohorts. By analysing the number of responses within each age group and gender category we gain a more detailed understanding of how perceptions and experiences may vary across distinct life stages and between genders.

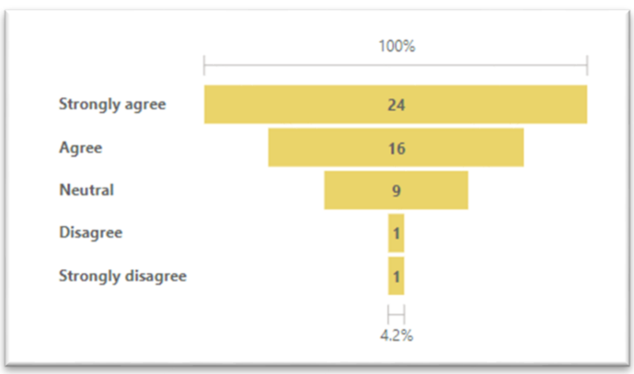

Figure 6: Responses for fear of social stigma

The above figure meticulously portrays the distribution of responses regarding the query on whether the fear of social stigma acts as a barrier to individuals seeking mental health support. This visual representation serves as a valuable tool for dissecting the nuances embedded within the participant's perspectives on the interplay between social stigma and mental health seeking behaviours. The distribution is stratified across distinct categories shedding light on the varying degrees of agreement or disagreement among respondents. Similar visual presentations for other questions are provided below.

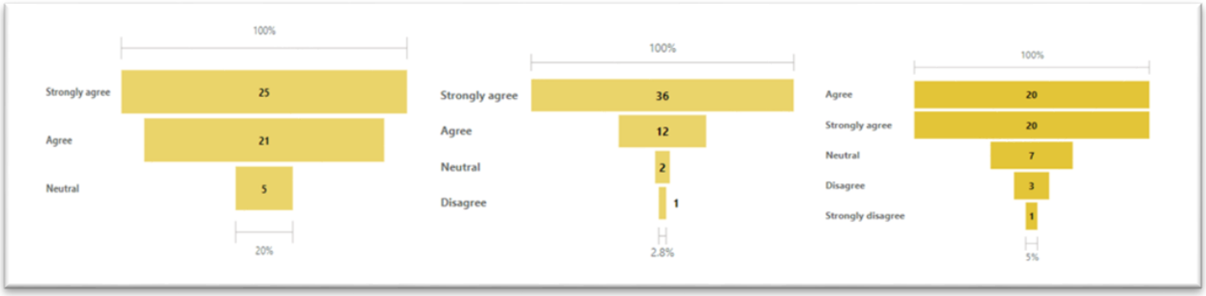

Figure 7: Responses for impact of stress on performance, physical health and impact of hobbies in decreasing stress levels

6 Results

This section encompasses discussions on various statistical analysis techniques, along with the presentation of their results in tabular form. The analyses conducted include tests for validity and reliability, feasibility of questions, and correlation between variables.

6.1 Test of Reliability

Reliability testing holds paramount importance in product development and quality assurance endeavours, aiming to assess and ensure the consistency, dependability, and stability of a product or system's performance across different conditions and timeframes. Cronbach’s alpha was employed as a metric to gauge consistency in the reliability test. The obtained results fall within the range of 0.7 to 0.8, indicating satisfactory reliability for each item investigated in the research.

Reliability Statistics

Cronbach's Alpha	N of Items
.751	6

Figure 8: Test of Reliability

6.2 Test of Validity

This study employs correlation testing to evaluate the validity of its research questions. The provided table presents the results of the validity test conducted on each item, based on data collected from 51 respondents. Red ovals indicate a strong correlation between respective variables with a 99% confidence interval, while green ovals signify a significant correlation with a 95% confidence interval. The validity coefficient (r-value) is compared against its reference criterion, derived from the r table, confirming the validity of each question.

Correlations

		Office_Performance	Physical_health	Hobbies	Social_stigma	Finance	variable
Office_Performance	Pearson Correlation	1	.565**	.229	.259	.562**	.380**
	Sig. (2-tailed)		.000	.107	.067	.000	.006
	N	51	51	51	51	51	51
Physical_health	Pearson Correlation	.565**	1	.263	.216	.497**	.478**
	Sig. (2-tailed)	.000		.063	.127	.000	.000
	N	51	51	51	51	51	51
Hobbies	Pearson Correlation	.229	.263	1	.397**	.149	.125
	Sig. (2-tailed)	.107	.063		.004	.295	.380
	N	51	51	51	51	51	51
Social_stigma	Pearson Correlation	.259	.216	.397**	1	.538**	.326*
	Sig. (2-tailed)	.067	.127	.004		.000	.019
	N	51	51	51	51	51	51
Finance	Pearson Correlation	.562**	.497**	.149	.538**	1	.377**
	Sig. (2-tailed)	.000	.000	.295	.000		.006
	N	51	51	51	51	51	51
variable	Pearson Correlation	.380**	.478**	.125	.326*	.377**	1
	Sig. (2-tailed)	.006	.000	.380	.019	.006	
	N	51	51	51	51	51	51

Figure 9: Test of Validity

6.3 Test of Feasibility

The correlation between questions was assessed through Bartlett's test and the Kaiser-Meyer-Olkin (KMO) test, which are employed to evaluate the feasibility of questions undergoing factor analysis. The table displays that Bartlett's test of sphericity yields a significant p-value of 0.000, falling below the 0.05 threshold. This result indicates a meaningful relationship between the items in the questions. The Kaiser–Meyer–Olkin (KMO) value, standing at 0.701, falls within the range of 0.5 to 1, suggesting homogeneity among the questions. In essence, both tests affirm the viability and coherence of the questions subjected to factor analysis.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.701
Bartlett's Test of Sphericity	Approx. Chi-Square	106.004
	df	15
	Sig.	.000

Figure 10: Test of Feasibility

6.4 Chi-square Test

The chi-square test serves as a statistical technique to assess the relationship between categorical variables within a dataset. It quantifies the disparity between observed and expected frequencies and is commonly employed to examine independence or assess goodness of fit. Compute the Chi-square test statistic using the following formula:

The Formula for Chi Square Is

$$\chi^2_c = \sum \frac{(O_i - E_i)^2}{E_i}$$

where:

c = degrees of freedom

O = observed value(s)

E = expected value(s)

Figure 11: Chi-square test formula

7 Hypotheses Testing

7.1 Impact of high stress levels on performance of individuals.

- Null hypothesis (H0): High stress levels do not affect our performance in academics, office or any other working environment.
- Alternate hypothesis (H1): High stress levels affect our performance in academics, office or any other working environment.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	13.373[a]	4	.010
Likelihood Ratio	10.160	4	.038
Linear-by-Linear Association	7.234	1	.007
N of Valid Cases	51		

Figure 12: Result of Chi-square test for 1st hypothesis

Result: The p-value is 0.010 which is less than 0.05 so we reject the null hypothesis and accept the alternate hypothesis.

7.2 Fear of Social judgement.

- Null hypothesis (H0): An individual does not hesitate to seek help for mental health issues due to fear of social stigma.
- Alternate hypothesis (H1): An individual hesitates to seek help for mental health issues due to fear of social stigma.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	23.990[a]	8	.002
Likelihood Ratio	23.924	8	.002
Linear-by-Linear Association	5.327	1	.021
N of Valid Cases	51		

Figure 13: Result of chi-square test for 2^{nd} hypothesis

Result: The p-value is 0.02 which is less than 0.05 so we reject the null hypothesis and accept the alternate hypothesis.

7.3 Impact of Financial Constraint.

- Null hypothesis (H0): Financial constraint does not become a hurdle for finding support from professionals.
- Alternate hypothesis (H1): Financial constraint becomes a hurdle for finding support from professionals.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	11.434[a]	4	.022
Likelihood Ratio	8.295	4	.081
Linear-by-Linear Association	5.740	1	.017
N of Valid Cases	51		

Figure 14: Result of chi-square test for 3^{rd} hypothesis

Result: The p-value is 0.022 which is less than 0.05 so we reject the null hypothesis and accept the alternate hypothesis.

7.4 Impact of high stress levels on physical health

- Null hypothesis (H0): High stress levels do not affect our physical health.
- Alternate hypothesis (H1): High stress levels affect our physical health.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	14.310[a]	6	.026
Likelihood Ratio	12.611	6	.050
Linear-by-Linear Association	11.434	1	.001
N of Valid Cases	51		

Figure 15: Result of chi-square test for 4th hypothesis

Result: The p-value is 0.026 which is less than 0.05 so we reject the null hypothesis and accept the alternate hypothesis.

7.5 Impact of soothing activities on mental health

- Null hypothesis (H0): Engaging in soothing activities or hobbies does not assist a person in transitioning out of the high stress period.
- Alternate hypothesis (H1): Engaging in soothing activities or hobbies assist a person in transitioning out of the high stress period.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	14.106[a]	6	.028
Likelihood Ratio	9.761	6	.135
Linear-by-Linear Association	5.960	1	.015
N of Valid Cases	51		

Figure 16: Result of chi-square test for 5th hypothesis

Result: The p-value is 0.028 which is less than 0.05 so we reject the null hypothesis and accept the alternate hypothesis.

8 User Interface of Web Application

Home Page:

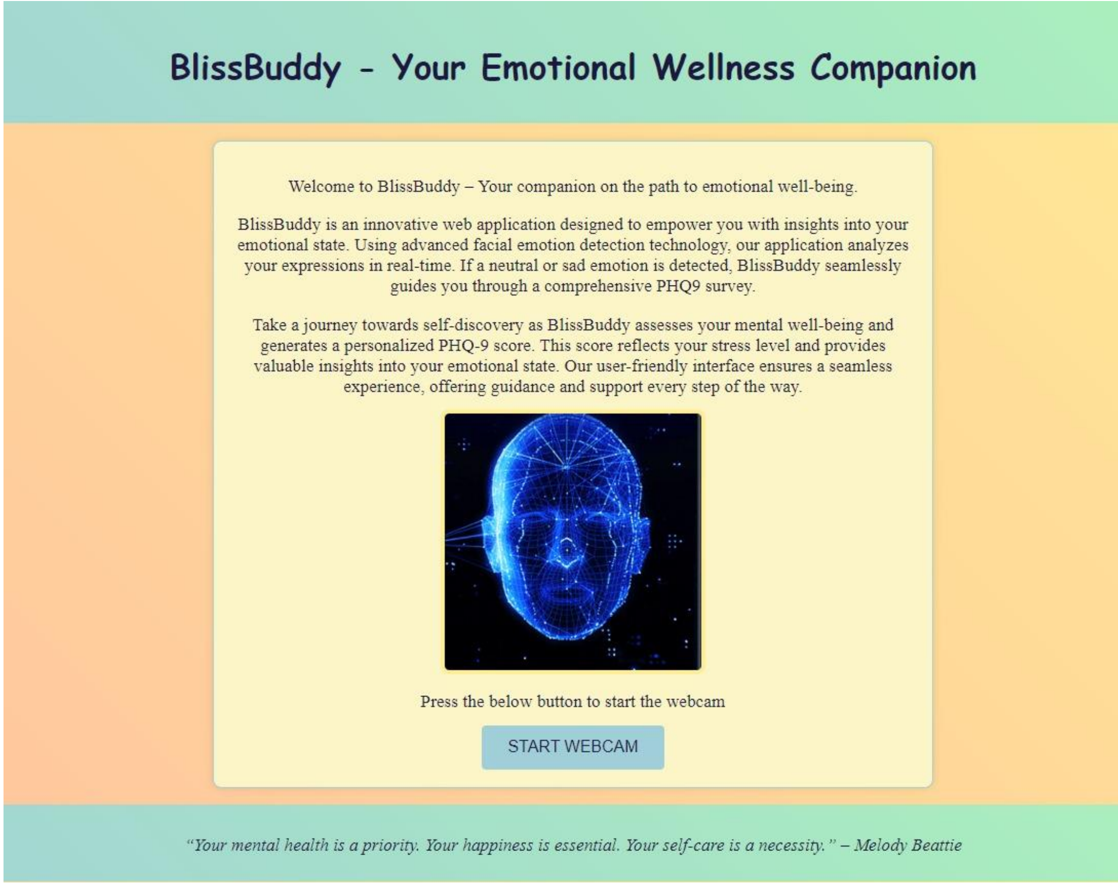

Figure 17: Home Page

After clicking on START WEBCAM button, a webcam window will pop up as shown below:

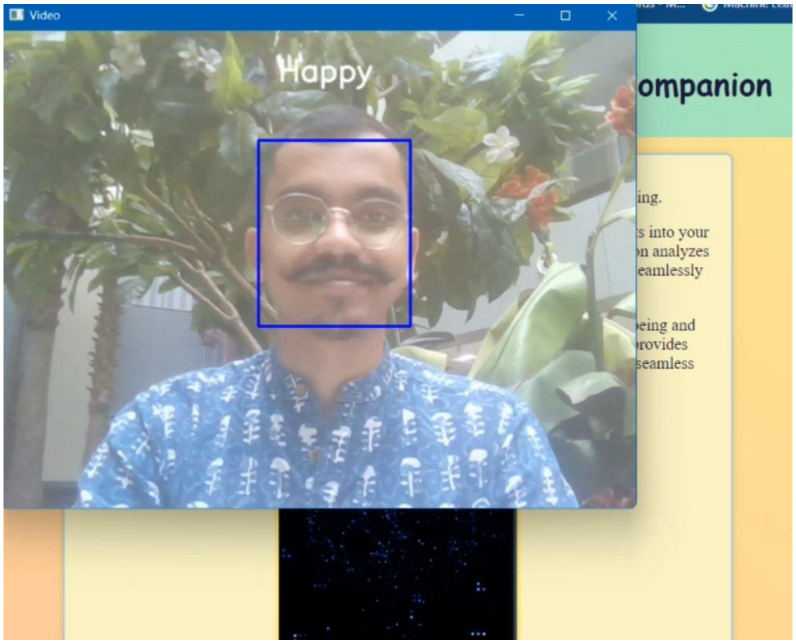

Figure 18: Webcam Window

As the above image shows, the detected emotion is ‘Happy’, so click picture and proceed further. Press ‘S’ from the keyboard, and it’ll redirect to the following page:

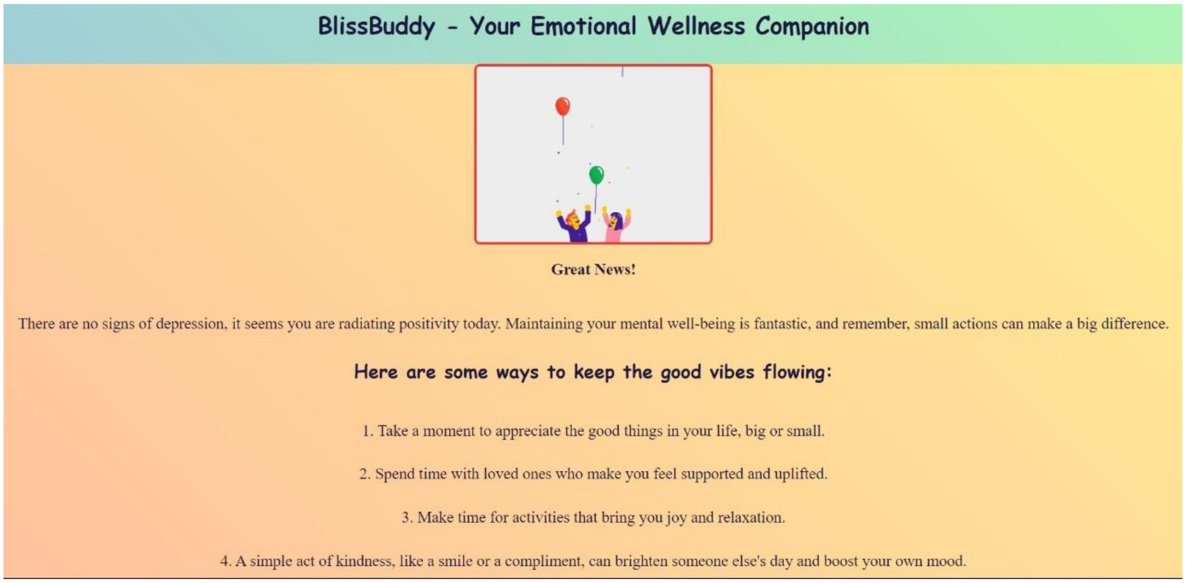

Figure 19: Happy Emotion Page

Similarly, if the detected emotion is ‘Sad’:

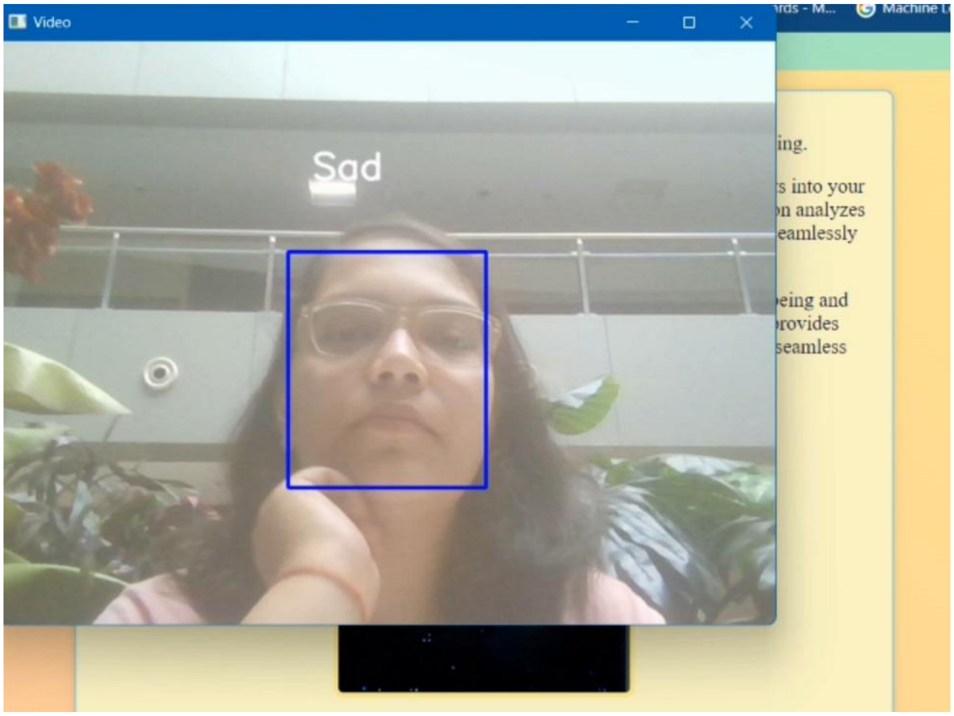

Figure 20: Webcam Window

It will take user to PHQ9 Questionnaire:

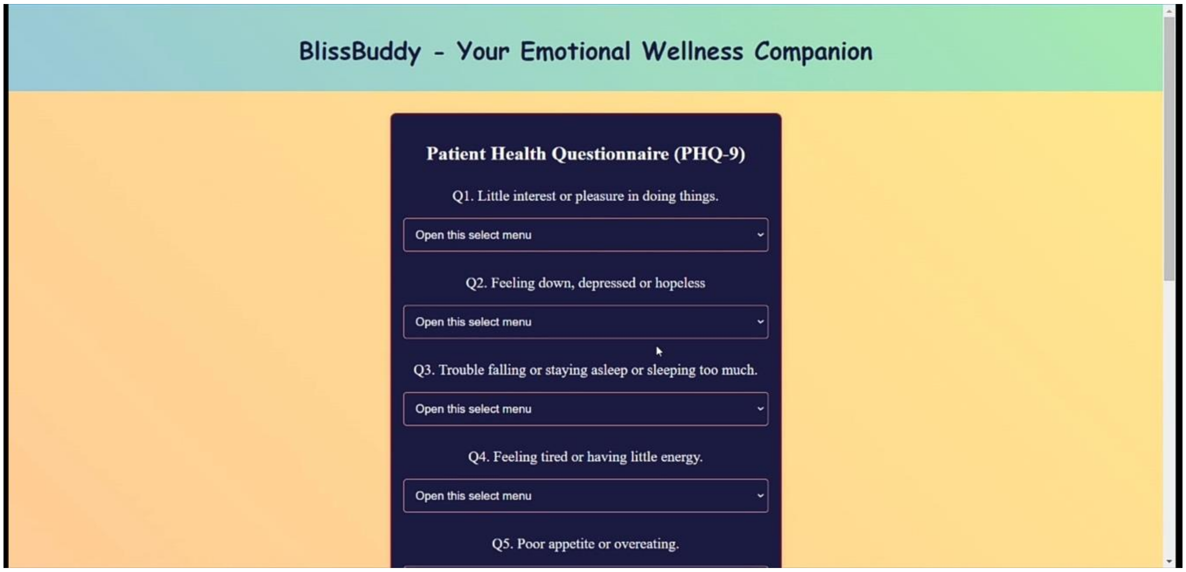

Figure 21: PHQ9 Questionnaire

After submitting, based on the questionnaire score, suggestions and recommendations for mindfulness exercises, different breathing techniques and useful links will be provided:

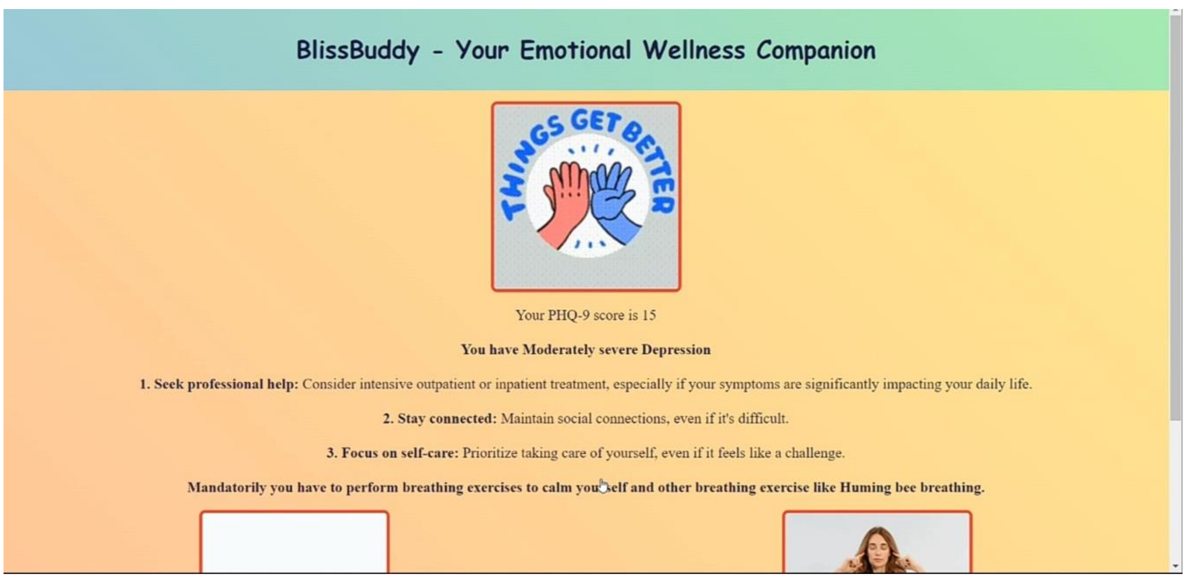

Figure 22: Recommendation Page

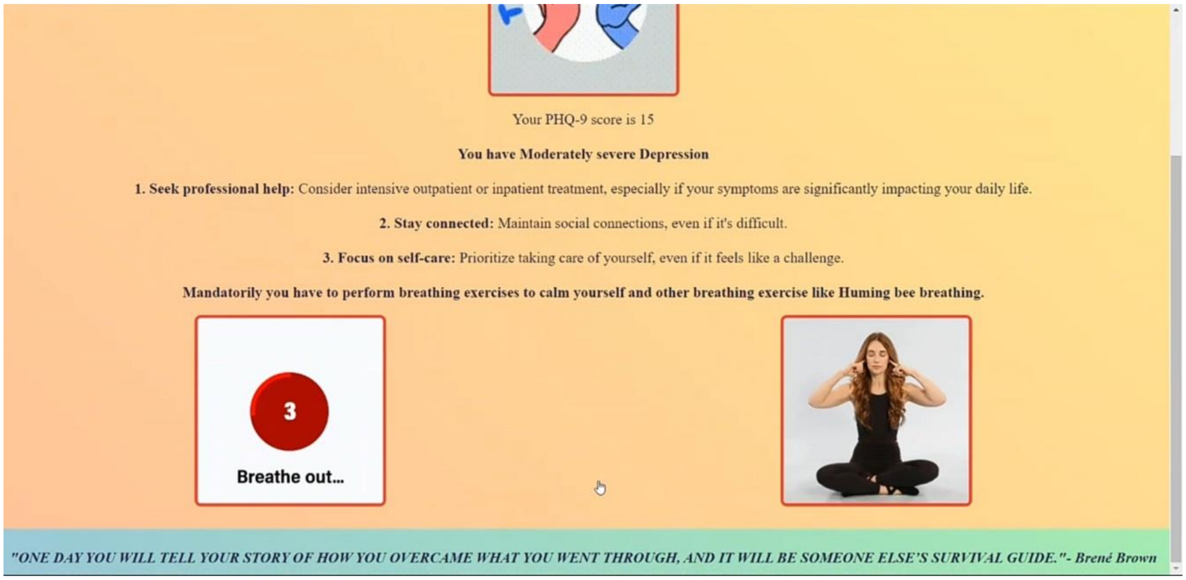

Figure 23: Recommendation Page

If detected emotion is 'Neutral':

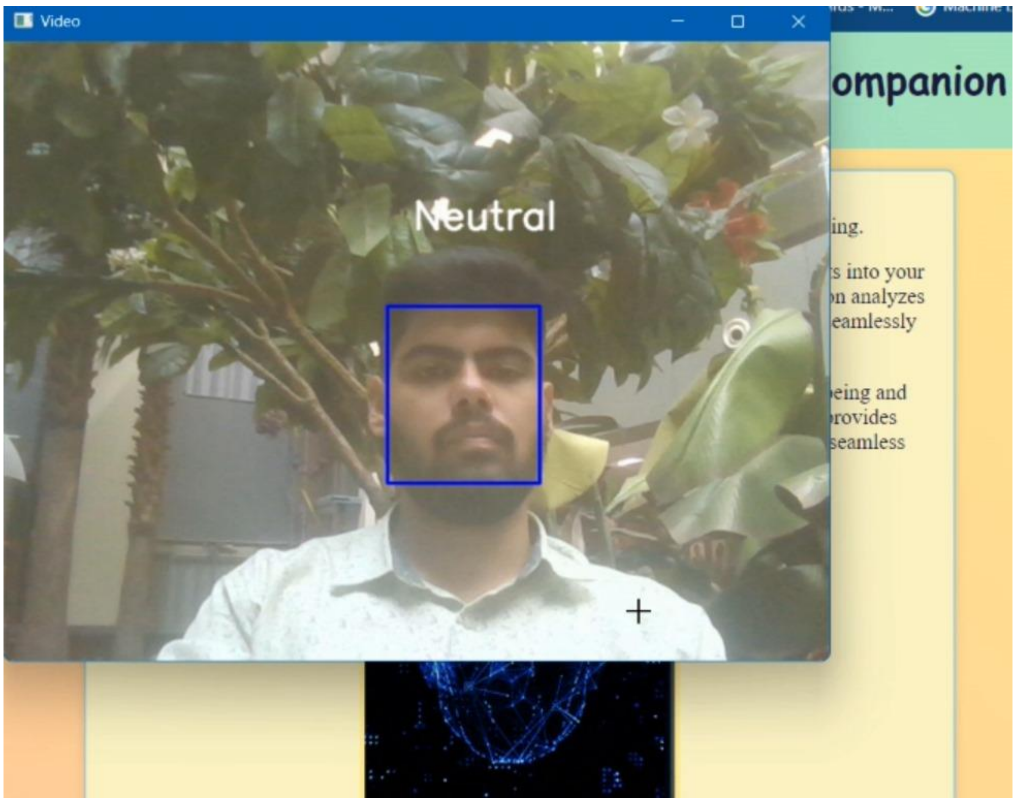

Figure 24: Webcam Window

It will follow same workflow as above and provide suggestions and recommendations of mindfulness exercises, different breathing techniques and useful based on the questionnaire score:

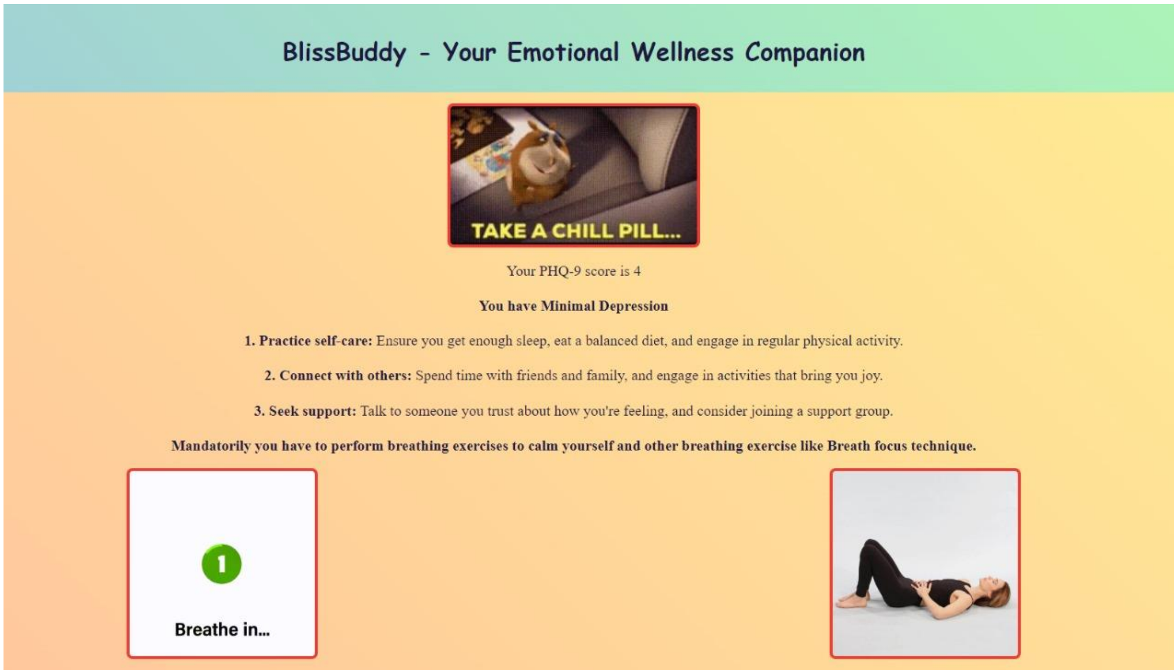

Figure 25: Recommendation Page

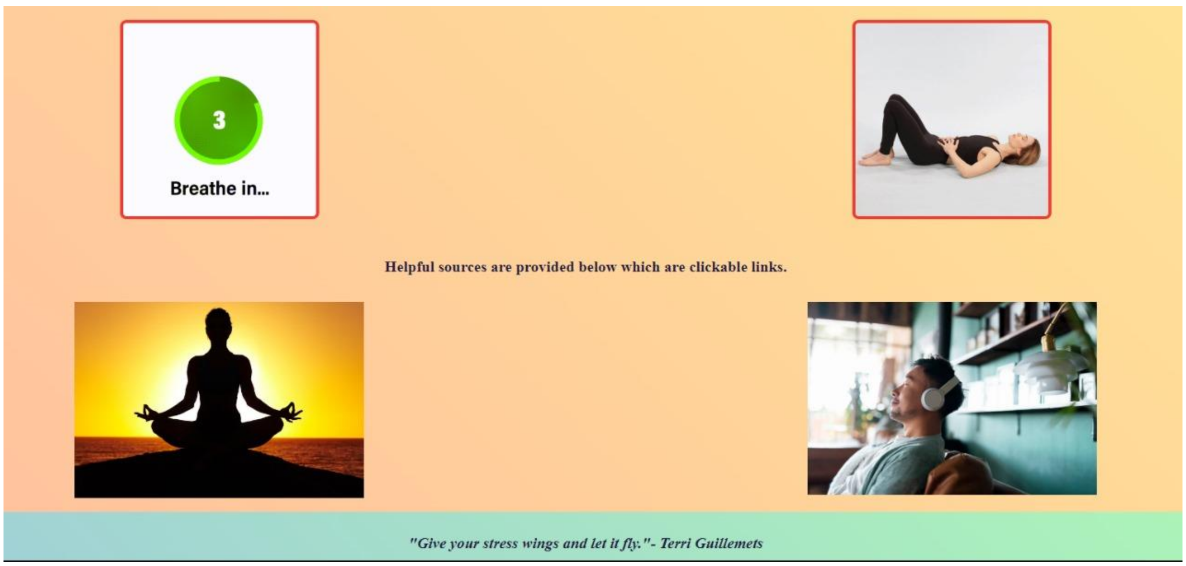

Figure 26: Recommendation Page

9 Conclusion

By combining facial expression analysis technology with empathy, this research project successfully aimed to increase the accessibility and affordability of mental health support services. The reliability, validity, and feasibility of the research questions are confirmed by the statistical analyses that were carried out. Chi-square test results on five hypotheses consistently reject null hypotheses, showing the importance of financial constraints, fear of social judgment, and high stress levels in different aspects of people's lives. The findings underlined the importance of treating mental health concerns since they showed a substantial impact of high stress on people's performance and physical health.

According to the study, social stigma and financial limitations are major obstacles to receiving mental health care, underscoring the need for readily available and reasonably priced solutions. These findings provide important information that may direct therapies and networks of support for individuals managing stress and related problems. By incorporating additional visual cues, conducting real-world testing, enhancing interpretability, and improving user experience, future iterations of the model can greatly impact the diagnosis and management of depression on a global scale.

With a Home Page and a Webcam Window for emotion detection ('Happy,' 'Sad,' 'Neutral,' etc.), the web application boasts an easy-to-use interface. After completing the PHQ9 Questionnaire, users are directed to a recommendation page with tailored breathing exercises and mindfulness exercises. The web application is good for promoting well-being because it places a high priority on user engagement, mental health assessment, and useful recommendations.

10 Future Scope

The proposed machine learning model for depression detection using facial expressions and PHQ-9 responses presents a ground-breaking advancement in the field of mental health diagnosis and intervention. While the initial study showcases the tremendous potential of this approach, there are numerous exciting avenues for further exploration and development that could significantly enhance its effectiveness and applicability.

One promising direction for future research involves expanding the scope of visual cues analysed by the model to include not only facial expressions but also other non-verbal cues such as body language and speech patterns. Studies have shown that individuals experiencing depression often exhibit distinct speech characteristics, including a slower speech rate, a monotone voice, and reduced gross body movements. Integrating computer vision and speech analysis modules into the model could provide a more comprehensive understanding of depressive symptoms, thereby improving accuracy and diagnostic precision.

Moreover, while the current testing has primarily relied on pre-collected datasets, conducting additional real-world testing and data collection is imperative for validating the model's performance in diverse clinical settings. Deployment of the model within the healthcare system or as an at-home monitoring tool would enable the collection of real-time data, facilitating the development of more robust and generalizable models. Furthermore, the integration of intervention mechanisms into the model could allow for timely detection of warning signs and prompt delivery of support services to individuals at risk.

Addressing mental health, a sensitive subject, involves partnering with mental health professionals to offer diverse mindfulness exercises and calming techniques. To address financial considerations, economical subscription plans can be introduced on the web application, providing personalized mental health support from professionals.

Nevertheless, a significant challenge in the domain of machine learning-driven depression detection involves the interpretability of model decisions. It is essential to enhance the clarity of machine learning models on a case-by-case basis to establish trustworthiness and improve diagnostic capabilities. Developing methods to clarify and interpret model behaviour empowers clinicians and individuals, enabling them to better comprehend and respond effectively to the model's predictions.

Furthermore, enhancing the user interface and overall user experience of the system is vital to boosting engagement and participation in the mental health assessment process. Crafting user-friendly interfaces that facilitate seamless interaction with the model can motivate individuals to willingly interact with the system, thereby fostering more reliable and comprehensive mental health assessments.

11 Acknowledgements

The research project would like to express its gratitude to all of the people who helped make it happen. The authors would like to thank the research advisors for their invaluable advice and assistance, whose knowledge and mentoring greatly enhanced the study. Additionally, thanks are given to colleagues whose cooperative efforts improved the caliber of the study. The involvement of study participants was essential to the comprehensiveness of the research.

References

1. Uddin, M. Z., Hassan, M. M., Almogren, A., Alamri, A., Alrubaian, M., & Fortino, G. (2017). Facial expression recognition utilizing local direction-based robust features and deep belief network. IEEE Access, 5, 4525-4536. [1]
2. Fei, Z., Yang, E., Li, D. D. U., Butler, S., Ijomah, W., Li, X., & Zhou, H. (2020). Deep convolution network based emotion analysis towards mental health care. Neurocomputing, 388, 212-227. [2]
3. Gilaine, G., Hassan, M., Asghar, M., Qamar, A. M., Ullah, H., Rehan, U. K., Aslam, N., Irfan, U. K.. (2022). An Automated and Real-Time Approach of Depression Detection from Facial Micro-expressions. CMC, 73, 2514-2527. [3]
4. Rescigno, M., Spezialetti, M., & Rossi, S. (2020). Personalized models for facial emotion recognition through transfer learning. Multimedia Tools and Applications, 79, 35811-35828. [4]
5. Lee, Y. S., & Park, W. H. (2022). Diagnosis of depressive disorder model on facial expression based on fast R-CNN. Diagnostics, 12(2), 317. [5]
6. https://med.stanford.edu/fastlab/research/imapp/msrs/_jcr_content/main/accordion/accordion_content3/download_256324296/file.res/PHQ9%20id%20date%2008.03.pdf [6]
7. https://www.kaggle.com/competitions/challenges-in-representation-learning-facial-expression-recognition-challenge/data [7]
8. https://medium.com/analytics-vidhya/car-brand-classification-using-vgg16-transfer-learning-f219a0f09765#:~:text=VGG16%20is%20a%2016%20layer,3%20represents%20the%20RGB%20channel.
9. https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network
10. https://www.youtube.com/watch?v=lPJAtHWq08k
11. https://www.youtube.com/watch?v=YRJ6xoiRcpQ
12. https://www.youtube.com/watch?v=nZP088xSDeQ
13. https://www.youtube.com/watch?v=fb3lDTS5IS4