EMPIRICAL ANALYSIS ON MENTAL HEALTH PREDICTION USING IMAGE PROCESSING AND MACHINE LEARNING APPROACH

Submitted in partial fulfillment of the requirements for the degree

of

Bachelor of Science in Information and Communication Engineering

by

Mahmudul Hasan Azad [201-50-010]
S.M. Hamid Shariar [203-50-041]

Under the Supervision of

Dipto Biswas
Lecturer
Depertment of Information and Communication Engineering

Daffodil International University

September 2024

APPROVAL

This is to certify that the entitled **Empirical Analysis On Mental Health Prediction Using Image Processing And Machine Learning Approach**, submitted by Mahmudul Hasan Azad (ID: 201-50-010) and S.M. Hamid Shahriar (ID: 203-50-041) are undergraduate students of the Department of ICE has been examined. Upon recommendation by the examination committee, we hereby accord our approval of it as the presented work and submitted report fulfill the requirements for its acceptance in partial fulfillment for the degree of *Bachelor of Science in Information and Communication Engineering.*

BOARD OF EXAMINERS

Md. Taslim Arefin
Associate Professor and Head
Department of ICE
Daffodil International University

Chairman

Prof. Dr. A. K. M. Fazlul Haque
Professor
Department of ICE
Daffodil International University

Internal Member

Ms. Tasnuva Ali
Assistant Professor
Department of ICE
Daffodil International University

Internal Member

Dr. Engr. M. Quamruzzaman
Professor and Head
Department of EEE
World University of Bangladesh

External Member

© Daffodil International University

DECLARATION OF AUTHORSHIP

Project Title	Empirical Analysis On Mental Health Prediction Using Image Processing And Machine Learning Approach
Authors	Mahmudul Hasan , S.M. Hamid Shahriar
Supervisor	Dipto Biswas, Lecturer, Department of ICE, DIU

We, First Author and Second Author, hereby declare that this capstone project titled **Empirical Analysis On Mental Health Prediction Using Image Processing And Machine Learning Approach** is entirely our own work, unless otherwise referenced or acknowledged. The content of this project is the result of our own research and efforts, and we have complied with all ethical guidelines and academic standards.

We are aware of the consequences of academic dishonesty and understand that any violation of ethical standards in this project may lead to disciplinary actions as defined by Daffodil International University's policies.

Mahmudul Hasan

Mahmudul Hasan Azad

201-50-010

Hamid Shahriar.

S.M. Hamid Shahriar

203-50-041

8/9/24

Dipto Biswas

Lecturer

Department of Information and Communication Engineering

© Daffodil International University

ACKNOWLEDGEMENT

We would like to express our deepest gratitude to everyone who played a significant role in the successful completion of this project.

We are very thankful to **Dipto Biswas**, the Lecturer of Information and Communication Engineering Department for helping us throughout our project work with patience & care that is absolutely necessary on this journey! His expertise and feedback has been instrumental for this project. In the presence of Dipto Biswas, we learnt a lot about innovative thinking and we grew in many areas academically or personally. We cannot thank him enough for his guidance and inspiration that really helped us accomplish our goals.

We are also grateful to Associate Professor **Md.Taslim Arefin**, Head of Information and Communication Engineering Dept, Daffodil International University for his cooperation. He has guided me throughout this project and provided ongoing support. **Md. Taslim Arefin** for his constructive suggestions and encouragement, which was helpful to determine the direction of our research work in much better way as well improving scientific quality of the typing draft.

We are also grateful to **Mr. Raikibul Hasan Sourav** , (M.Sc in counseling psychology, University of Dhaka) Psychologist at Daffodil International University,for his essential cooperation.His expertise in counseling psychology greatly contributed to our project by providing insights into the psychology aspect of mental health condition, wich enriched our understanding and analysis of the data.

Finally, we also express our thanks to Daffodil International University for supplying the required resources and facilities that made this work possible. Our success was also largely due to the commitment of the university towards creating a research-focused environment.

Finally, we thank our families and friends for their enduring patience in the worst of times. Since many of the teachers and Burnie believe us to be capable, it helps keep our morale so much higher.

It was the combined efforts of all these people that made this project. This page would not be possible without the time, dedication and hard work of these people and for that we thank you!

ABSTRACT

A vital component of wellbeing is mental health. Our mental health is very important as it maintains our overall well-being, affects how we think feel and act, and prompt treatments can be made when mental health issues are detected early.Empirical analysis ensures that findings are supported by actual evidence by employing measured and observed data to make inferences. This approach enhances the accuracy and reliability of predictions in data-driven studies. The goal of this research is to support mental health diagnostics by predicting mental health disorders through image processing techniques. Using a robust dataset having four features such as Talking, Walking, Eating, Playing, a variety of machine learning models were applied in the empirical investigation to reliably categorize mental health problems. The models' performance was assessed using metrics for accuracy, recall, and precision. A variety of techniques are employed, including ensemble approach, KNN, Activation Function, ANN, LSTM, CNN, Decision Tree, Random Forest, SVM, and Logistic Regression. This research has obtained accuracy about 0.995 for logistic regression, 0.955 for decision tree, 0.994 for random forest, 0.959 for SVM, 0.993 for ensemble methods, 0.988 for KNN, 0.999 for activation function, 0.989 for ANN, 0.988 for LSTM, 0.988 for CNN regarding talking feature. This research has obtained accuracy about 0.979 for logistic regression, 0.988 for decision tree, 0.987 for random forest, 0.98 for SVM, 0.992 for ensemble methods, 0.992 for KNN, 0.994 for activation function, 0.931 for ANN, 0.994 for LSTM, 0.956 for CNN regarding walking feature. This research has obtained accuracy about 0.979 for logistic regression, 0.981 for decision tree, 0.999 for random forest, 0.922 for SVM, 0.963 for ensemble methods, 0.941 for KNN, 0.987 for activation function, 0.995 for ANN, 0.91 for LSTM, 0.982 for CNN regarding eating feature. This research has obtained accuracy about 0.98 for logistic regression, 0.975 for decision tree, 0.994 for random forest, 0.98 for SVM, 0.91 for ensemble methods, 0.978 for KNN, 0.967 for activation function, 0.972 for ANN, 0.993 for LSTM, 0.996 for CNN regarding playing feature.

Keywords: Mental Health, Empirical Analysis , Machine Learning, Image Processing, Depression Analysis, Anxiety Analysis, Bipolar disorder Analysis, Schizophrenia Analysis.

TABLE OF CONTENTS

APPROVAL **1**

DECLARATION OF AUTHORSHIP 2

ACKNOWLEDGEMENT 3

TABLE OF CONTENTS 5

Chapter 1 ..... **10**

INTRODUCTION **10**

1.1 Project Background 11

1.2Project Definition 15

1.2.1 Problem statement 19

1.2.2 Cotext and Scope 21

1.2.3 Significance and Implications 22

1.2.4 Quantification 23

1.3 Project Objective: 28

1.3.1 Literature Review: 33

1.4 Project contribution 37

1.5 Project Specification: 39

Chapter 2 ..... **44**

PROJECT MANAGEMENT **44**

2.1 Project Plan 46

2.2 Contribution of Team Members 49

2.2.1 Team Member 1: 49

2.2.2 Team Member 2: 50

2.3 Challenges and Decision Making 51

Chapter 3 ..... **53**

3.1 Block Diagram of the System 54

3.2 Design of Each Block and Select the Best Alternative 56

3.3 Model Development 71

3.4. System Implementation..... 85
Chapter 4..... 95
SYSTEM TESTING AND ANALYSIS.....95
4.1 Feature and Attributes.....96
4.1.1 Size.....96
4.1.2 Data collection sources:.....97
4.1.3 Target Labels.....98
4.2 Performance Metrics.....99
4.2.1 Accuracy:.....99
4.2.2 Precision:.....100
4.2.3 Recall.....100
4.2.4 F1 Score:.....101
4.2.5 ROC-AUC:.....102
4.2.6 Specificity.....103
4.2.7 Sensitivity.....104
4.3 Feature Transformation Result.....104
4.4 Performance Results.....105
4.5 Evaluation matrics.....125
4.6 Evaluation.....126
4.7 Comparison and Discussion.....127
Chapter 5..... 128
CONCLUSION AND RECOMMENDATION..... 128
5.1 Conclusion.....128
5.2 Future Recommendation.....128
Reference:.....130
Appendix:.....136

List of Figure

Figure 1: Average increase in mental illness……..…...……14

Figure 2: Age Distribution of Mental Health DisorCas……..…..……15

Figure 3: Global Burden of Mentla Health Condition……...……..…….26

Figure 4: Economic Impact of Mental Health Condition……..……27

Figure 5: Treatment Seeking Status……..……..…..……28

Figure 6: Risk Factor of Sucide……..……..…….…..…29

Figure 7: Detailed Description of the Project Plan Demonstration…….… 51

Figure 8: Proposed method……56

Figure 9: Performance result of Logistic Regression……..……107

Figure10: Performance result of Logistic Regression…….…….……110

Figure 11: Performance result of Random Forest……112

Figure 12: Performance result of SVM…….115

Figure 13: Performance result of Ensembla methods ……..…118

Figure 14: Performance result of KNN ……120

Figure 15: Performance result of Activation Function ……..…123

Figure 16: Performance result of ANN…….125

Figure17: Performance result of LSTM…….……128

Figure18: Performance result of CNN…….……...…130

List of Table

Table 1: Details Description of The Project Working Plan……49-50

Table 2: Details Information of Datasets……...…95

Table 3: Feature Transformation Result …….……104

Table 4: Performance Result on Talking Feature Of Logistic Regression……/…104

Table 5: Performance Result on Walking Feature Of Logistic Regression…..……...…104

Table 6: Performance Result on n Eating Feature Of Logistic Regression…….…105

Table 7: Performance Result on Playing Feature Of Logistic Regression……..….105

Table 8: Performance Result on Talking Feature Of Decision Tree……106

Table 9: Performance Result on Walking Feature Of Decision Tree……..…..107

Table 10: Performance Result on Eating Feature Of Decision Tree……...…..…107

Table 11: Performance Result on Playing Feature Of Decision Tree……..……108

Table 12: Performance Result on Talking Feature Of Random Forest…….……..…109

Table 13: Performance Result on WalkingFeature Of Random Forest …….……110

Table 14: Performance Result on Eating Feature Of Random Forest……...…110

Table 15:Performance Result on Playing Feature Of Random Forest……..……..…..111

Table 16: Performance Result on Talking Feature Of SVM…….……..…..112

Table 17: Performance Result on Walking Feature Of SVM…….……..…..112

Table 18: Performance Result on Eating Feature Of SVM…….…..113

Table 19: Performance Result on Playing Feature Of SVM…….113

Table 20: Performance Result on Talking Feature Of Ensemble methods……..….114

Table 21: Performance Result on Walking Feature Of Ensemble methods……;…..…115

Table 22: Performance Result on Eating Feature Of Ensemble methods ……;;;..…115

Table 23: Performance Result on Playing Feature Of Ensemble methods ……...…116

Table 24: Performance Result onTalking Feature Of KNN……...…117

Table 25:Performance Result on Walking Feature Of KNN…….....…118

Table 26: Performance Result on Eating Feature Of KNN ……..……..…118

Table 27: Performance Result on Playing Feature Of KNN……..…119

Table 28: Performance Result on Talking Feature Of Activation Function…….……120

Table 29: Performance Result on Walking Feature Of Activation Function…….……..120

Table 30: Performance Result on Eating Feature Of Activation Function……..121

Table 31: Performance Result on Playing Feature Of Activation Function……..…121

Table 32: Performance Result on Talking Feature Of ANN……..…122

Table 33: Performance Result on on walking Feature Of ANN……..…123

Table 34: Performance Result on on Eating Feature Of ANN…….……..….123

Table 35: Performance Result on on Playing Feature Of ANN…….……..…….124

Table 36: Performance Result on on Talking Feature Of LSTM…….……..…...125

Table 37: Performance Result on on Walking Feature Of LSTM……..…125

Table 38: Performance Result on on Eating Feature Of LSTM……..…126

Table 39: Performance Result on on Playing Feature Of LSTM …….…..123

Table 40: Performance Result on on Talking Feature Of CNN ……..…127

Table 41: Performance Result on on Walking Feature Of CNN ……..……..…128

Table 42: Performance Result on on Eating Feature Of CNN …….……128

Table 43: Performance Result on on Playing Feature Of CNN ……..…..…129

Table 44: Details of evaluation Metrics (SSCF,AEGR, and AEIG).....……...130

Table 45: Details of evaluation Metrics (SSCF,AEGR, and AEIG)…….……131

Table 46: Comparison and Discussion ……..….130

Chapter 1

INTRODUCTION

Good mental health is the presence of a state that allows us to feel good about ourselves and others. Mental health is described as a state of well-being that allows individuals to realize their own abilities, cope with the normal stresses of life, work productively, and make contributions to their community [1]. Health and wellness play a crucial role in our lives that helps us remain healthy and enjoy it fully [2]. In a way, it is this fundamental piece of health and well-being that occupies our alternative human capacity for relationship-building, decision-making and world-changing. Mental health is in fact, a right and the first and most BASIS that a human being is born with, precede health or well-being cognition as many components of a physical basis for having to be indifferable resulting from personal (including mental development), community hence socio-economic impact [3, 4].

The mental health of a person is important to their overall well-being and it influences all aspects of the daily thoughts, feelings, and behaviors.In short, a rise in mental health illnesses, ranging from depression to anxiety and bipolar disorder means the need for tools that can effectively anticipate these disorders and manage them is at an all time high [5]. Exciting advancements in technology, particularly in machine learning and image processing offer opportunities to fulfill this request on a level neve before possible. Capitalizing on these technological advancements, this capstone project: "Mental Health Predictions Using Image Processing" aims to develop a system that can predict mental health disorders by analyzing visual cues. The program searches for cues of mental health conditions in facial expressions, body responses and other relevant features [6, 7]. If successful, this novel approach might be an absolute game changer in the diagnosis and monitoring of mental health illnesses from their early stages to ultimately help both patients and healthcare providers.

With strong algorithms and proven data analysis, we are looking forward to pushing our bounds through this initiative "Mental Health Predictions Using Image Processing" as no stone will remain unturned [8]. However, the goal of this research is to enlighten health professionals about its severity which will enable them to forecast and taking corrective actions wherever possible so as to improve mental outcomes for these individuals. The convergence of technology and mental health has recently given birth to forward-thinking approaches for both diagnosing psychological disorders as well as predicting them. For instance, employing image

processing techniques to analyze physiological signals and facial expressions related with mental health state determination represent one such approach [9, 10]. In this study, we used these advances to build an image-centric mental health prediction model.

1.1 Project Background

Mental health issues are a severe global concern, with millions of people suffering from disorders such as depression, anxiety and bipolar disorder. Traditional diagnostic methods are mainly based on subjective and long-lasting self-reporting assessment of subjects' symptoms, as well as clinical interviews [11]. The utilization of technology in mental health diagnostics marks the beginning to improve a less subjective, successful and precise method. The tendency to consider visual information using the image processing feature of computer vision makes it well applicable in this area.Recently, studies were carried out on predicting mental health state in video/images (eye movement tracking and facial expression analysis) using physiological signals such as Zelaya et al. In 2014, a study demonstrated that the facial action coding system (FACS) is able to detect signs of depression in facial expressions [12]. In the same line, Sikka et al. (2014) analyzed eye movements and linked them to anxiety scores with computer vision algorithms [13]. This work shows examples of how image processing can be used to address mental health.The outcomes of this study have the potential to be quite important. By using artificial intelligence to identify early warning signs of neurological disorders through the analysis of visual data (like brain scans or facial expressions) — the initiative aims at improving patient prognosis and treatment outcomes. One potential outcome is fewer hospital stays and late-stage interventions, leading to lower medical expenses. Further, this project could help to bridge the access treatment gap for mental health across the globe using making available tools that may aid in screening psychiatric illness, especially within resource-poor or low-health infrastructure settings [14]. The results of the study are potentially very significant. These technologies scan visual material, such as brain scans or facial expressions through automatic algorithms designed to diagnose patients sooner which can help improve patient outcomes. That might mean lower medical costs due to less time spent in hospital and fewer late-stage interventions. In addition, it has the potential to bridge treatment availability worldwide by decentralizing access to mental health evaluation tools, particularly in rural or low healthcare infrastructure regions. With that, image processing-based mental health prediction is still in its infancy but the promise it shows at this stage indicates future holds. Many prototypes and pilot studies have shown this approach to be technically feasible, which is exciting; however, there remain significant challenges it force us down a long list of scary things we'll need to solve [15]. A significant con, however, is that due to the great variance in physiological reactions and

facial expression among people it may influence how accurately predictions are conducted. Also, the transferability of image processing models could be limited due to differences in how emotions are displayed between cultures.The collection and utilization of private visual data raises the additional challenges around ethical and privacy concerns [16, 17]. The public trust and acceptance of data are built on the responsible use, secure storage, defense against misuse or abuse. Moreover, image processing methods are computationally intensive so not all times the necessary hardware to implement them can be available and reliable at the same time.

When work from in research to the outside world it is necessary pass few steps. More and more different datasets were required to collect in the first place, just so these models could be trained with better accuracy and thus, practical. To create consistent protocols for data collection and analysis, a collaborative effort must come from both researchers, clinicians as wells technology builders. Second, the need for transparent structures of data governance will be imperative to alleviate privacy and ethical concerns. The more well the marketers can inform and control people about what happens with their data, in my opinion consumers will trust technology again.Over time, the development of computing techniques and hardware strength will facilitate much easier use of image processing tools in practical contexts [19, 20, 21]. As machine learning and computer vision algorithms advance further, mental health prediction systems become more efficient and effective. The only way we can both resolve these and leverage image processing to a future where mental health services are more preemptive, accessible, and efficient.

Presented images to illustrate the our project background and features of interest that work towards understanding connections & relationships in which decisions can be made predicting mental health on image analysis

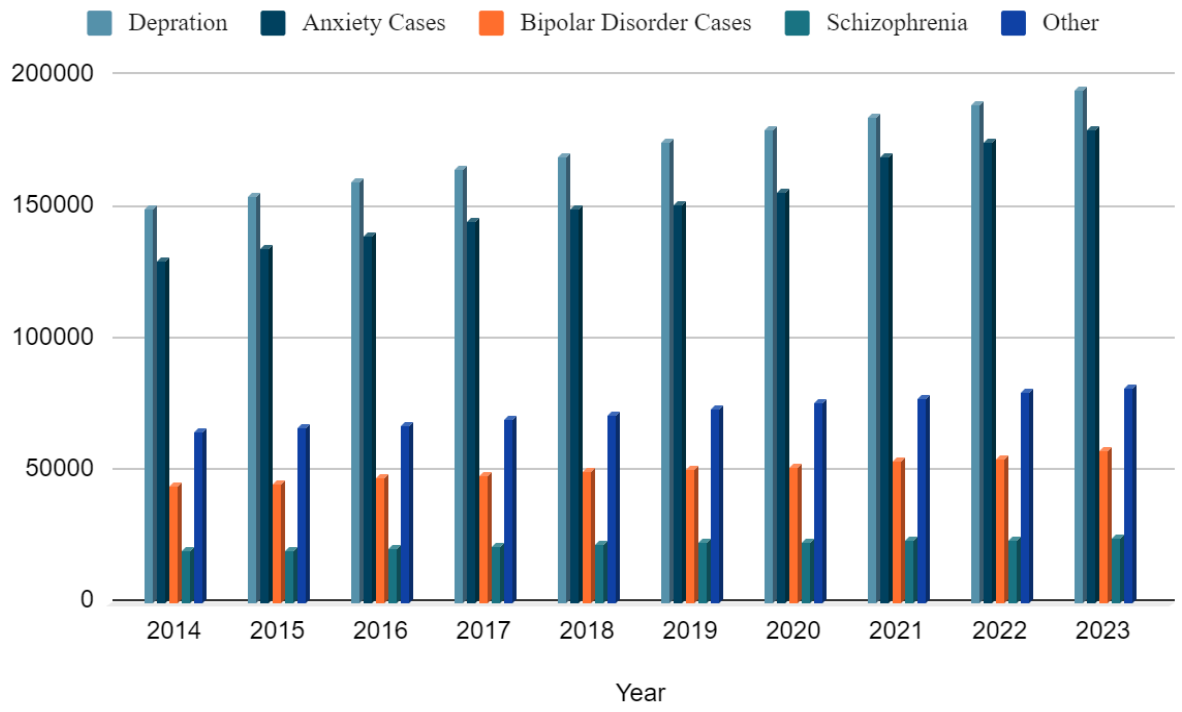

Figure1: Average increase in mental illness

The Figure 1 presents that

- Depression: A steady increase in cases has been observed, from 150,000 instances in 2014 to 200,000 cases in 2024.
- Anxiety: There were 130,000 cases in 2014 and 180,000 in 2024, suggesting that the incidence of anxiety is also steadily increasing.
- Bipolar Disorder: Here it is seen that in 2014 it was 45000 and in 2024 this number will rise to 60000. The detection of slow growth may actually indicate real growth.
- Schizophrenia:The number of cases of schizophrenia increased from 20,000 in 2014 to 25,000 in 2024. Even while the rise is not as noticeable as it is for other conditions, it nevertheless indicates a rising tendency.
- Other: The number of cases classified as “other diseases” increased from 65,000 in 2014 to 85,000 in 2024, also indicating an upward trend. Several less prevalent mental health problems may fall under this.

Age Distribution of Mental Health Disorder Cases (2024)
This graph shows the age distribution for 10-year mental health status

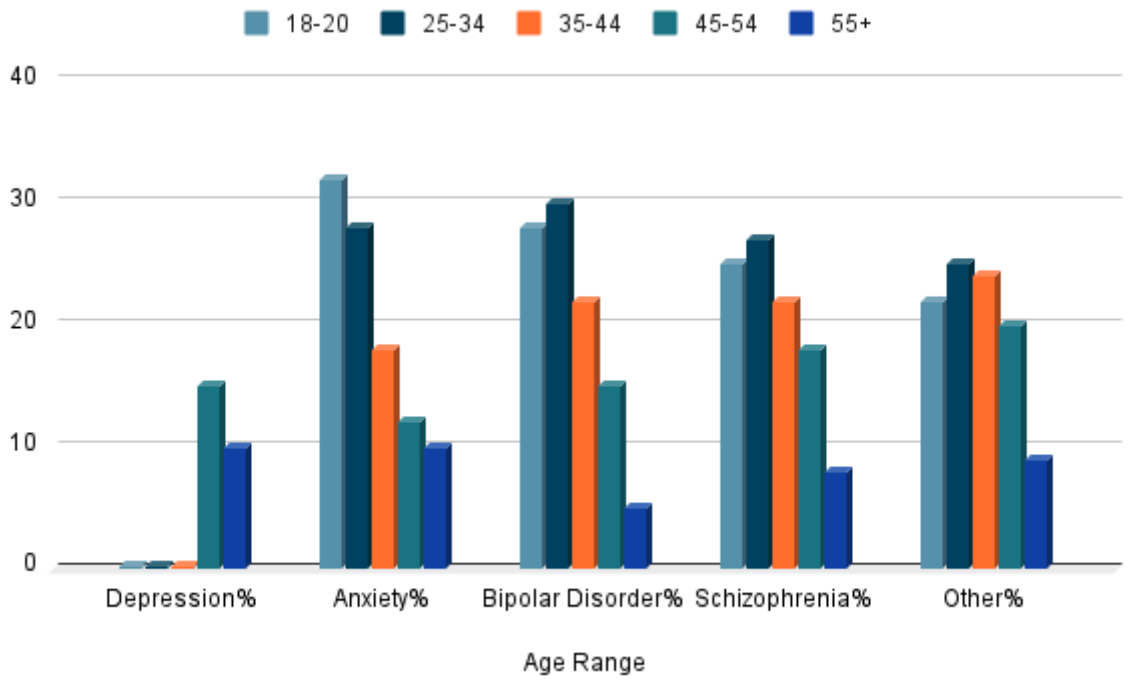

Figure2: Age Distribution of Mental Health Disorder Cases

Figure 2 shows the age distribution for 10-year mental health status

- What is implied in the graph above is the level of depression, with rates decreasing with age but with the highest prevalence among younger adults (18-24). Younger groups are more aware of what's going on and generally have more pressure.
- The graph suggests that anxiety is a disorder that appears to be overrepresented in young adults (18-24) and declines progressively with age.
- Bipolar Disorder: with an earlier prevention rate in adults aged 25 to 34, bipolar disorder is thought by some authorities as occurring more frequently or at diagnosis during early adulthood but occurrence decreases over time.
- Overall, that percentage shrinks with older age groups — probably because of long-term management or fewer new cases than occurring in any year Schizophrenia 18–24 most commonly diagnosed and a decent presence in later decades.
- Unlike many common disorders, the distribution across age groups is fairly uniform. A regional lack of older adults might be reflective of rarer disorders with which they are not diagnosed or passed away from earlier in life, preventing their potential diagnosis as an adult at the time we visited a region for tsunami survivors care delivered 1-3 years after first response delivery (December 2004-July 2012).

1.2Project Definition

Worldwide, mental health illnesses are common and pose a serious threat to healthcare systems. Conventional diagnostic techniques entail time-consuming clinical interviews, self-reported symptoms, and are frequently subjective. Treatment and diagnosis may be delayed as a result of these techniques. By applying image processing techniques, this research seeks to produce a reliable, impartial, and real-time mental health prediction system.

Specifications and Features

Project specifications are specific requirements and parameters that describe what we want to do/what the project offers. These specifications contain all the technical attributes, design criteria, and performance measures that must be met for a project to be successful. Attributes – Attributes are individual capabilities or features of the project that satisfy the constraints. They document what the project does, how it works, and strategies for its use. Our project has the potential to completely revolutionize mental health assessment using state-of-the-art technology. Our solution provides a long-requested approach in healthcare to have an accessible, non-invasive, and early detection system using image processing and machine learning to predict mental health problems. Subjective opinions and assessments form the traditional basis of mental health diagnostics, but they are open to bias due to stigma, access challenges, and delays in diagnosis. Our work presents an alternative approach to rapid intervention and improving patient outcomes while expanding mental health services to broader populations, particularly those living in underserved communities. Our work is therefore not purely a technical achievement, but also an important contribution towards improving mental health care worldwide. The following are the main characteristics and technical details:

A. Analysis of Facial Expressions: Facial expression analysis is the systematic study and interpretation of facial muscle movement to infer human feelings, intentions or subjective mental states [22]. Face perception interests in recognizing faces, as well needs to understand them have been studied from and inferring the state currently is relevant.utc.edu This includes fields such as psychology, communication studies,[2] design or other social science disciplines.

- Creating and putting into practice algorithms for facial recognition to find and follow faces in still photos or in real-time video streams.
- Feature extracting: To extract facial features including eye and mouth movements and other pertinent micro-expressions, use sophisticated computer vision techniques. Facial

Expression Feature Analysis – When trying to analyze facial expression features we must look at all of the elements needed for interpreting emotions, intentions and reactions depicted in faces. Face landmarks which mainly includes the eyes, nose, mouth and eyebrows. These formations act as landmarks to evaluate movements and expressions. Breaking down the face in separate parts for example upper facial area (forehead, eyes and eyebrows) vs lower FI face(mouth, nose chin to understand how each contribute to expression

- Emotion Recognition: Utilizing machine learning models to categorize and decipher emotions from extracted facial features. Pay particular attention to recognizing expressions that are frequently linked to mental health issues like fear, sorrow, rage, and disgust. The detection and management of mental health conditions have been enhanced by emotion recognition becoming humanized [23]. Emotion recognition technology can analyze the behavior by capturing facial expressions, vocal tones and revealing mental health problems, like depression, anxiety,stress etc. People often have "objective" expressions of sadness, lack of expression and low energy with depression [24].

B. Tracking Eye Movement : Eye movement tracking is a technique that measures and outputs the location, flow when moving from one point to another, and dwell time (for how long they fixate around a specific point) of human gaze points. Specialized equipment, such as eye-tracking devices (most commonly cameras and infrared sensors to record accurate data on the direction, speed and patterns of eye movements) accomplish this.

- Creating algorithms for precise eye movement tracking and detection.
- Using methods to determine fixation locations and gaze direction for gaze estimation.
- Blinking Rate Analysis: Examine blink rates and patterns to determine how well mental health is doing by connecting these measures to ailments like stress and worry. Blink rate analysis is the measure of how often and what pattern a person blinks in to determine their mental or emotional state. Blink rate analysis may be used in mental health detection to recognize increased cognitive load and stress, anxiety or even fatigue, as well changes associated with particular neurological conditions. The typical blink rate in a relaxed state, or baseline-blanking is normally somewhere around 15–20 blinks per minute [25].

C. Multimodal Data Integration: Multimodal data integration for detecting the mental health means, integrating and analyzing all different types of modality sources together including

physiological signals, Behavioral cues, Facial expressions, Voice analysis. Mental health detection systems should integrate multi-modal and enhance the performance of these deep learning models with standard benchmarks to accurately predict emotional states based on data from different sources.

- Data Fusion: Increasing the precision of mental health forecasts by fusing data on eye movements and facial expressions. Data fusion is the process of combining multiple data sources to provide a more consistent, accurate and comprehensive view than could be possible from any single source. This means merging original data across two or more sensors, or even sources at the most condensed level of processing. These are combinations of decisions from or inferences on individual modalities[26]. For instance the fusion of outputs from a facial expression analysis algorithm and a voice tone analysis system to make decisions on the final emotional state of a person.
- Contextual Analysis: To improving the predictive model, including more contextual data (such as user history and environmental characteristics). Knowing the situation in which data is collected — for example, when (time of day), where and how (geolocation) or with whom can be very valuable information. For example, stress markers may be more dramatic in a work scenario than at home. Studying trends in emotion and behavior over time on a daily, weekly or even seasonal level. Setting a baseline for different metrics from an individual itself to catch any anomalies that could likely suggest mental health issues.

D. Training Models and Machine Learning: Essence of mental health detection training models and machine learning is knowing how to code in an advanced computational language, that interpret data efficiently whereas recognize behavior patterns or types seamlessly associated with capability/runs around quite a few kinds of mental health conditions like depression, anxiety etc [27]. The data, which is some part of the dataset and more than one single record set well known as training done so our model trained using this training set. A model will learn during training how to associate the input data (features) with an output label, in order that prediction errors are minimized. Then the model is tested on another slice of that data — “the validation set” to tune its parameters so as not to overfit. Some of the methods to improve Model Performance are Cross-Validation,Regularization and Hyperparameter tuning

- Employing supervised learning techniques to develop models by using labeled datasets of eye movements and facial expressions associated with mental health issues.

- Carrying out cross-validation and other validation procedures to guarantee the robustness and dependability of the predictive model.

- To guarantee the robustness and dependability of the predictive models, Continuous Learning: New data is continuously pumped into the model, giving it new feed periodically. This could be wearables, mobile apps or clinical evaluating at home. This continuing model update process enables the system to incrementally manage changes in subject's mental attributes or consider new patterns that were not found in training data. Periodically retrain the model with a larger set of data consisting of both the original dataset as well as fresh new training samples collected over time [28]. It would allow the model to learn from new developments, and adapt with changing circumstances.

E. User Interface and System Integration: The User Interface (UI) and System Integration for mental health detection, play a critical role and are working hand in glove to provide an integrated flawless user-friendly experience available both for the users being monitored as well the systems which monitors/analyzes it. User Interface (UI): The interface where the user interacts with a mental health detection system This includes all user-facing elements from the screens to buttons, icons and other graphical components. To make a system user-friendly, inteenstive & approachable should be the objective of making a UI.

- Real-Time Processing: Verifying that the system is capable of processing data instantly, giving forecasts and feedback. The system monitors live data to detect fast-paced changes in mental health and respond with real-time feedback, alerts or recommendations. Interventions can be made on the basis of real-time, which is essential during crises.

- User Interface: Providing a user-friendly interface that makes it simple for patients and healthcare providers to interact with the system. The interface should be overly intuitive where it would make so that the user can just go in, and instantly check on mental health-related information like case alerts or trends. They should include easy to view data such as mood charts, stress levels or flags of potential issues.

- Data Security and Privacy: Putting robust security measures in place to safeguard private user information and make sure privacy laws are followed. Integrates data collection tools, databases and analysis platforms through a secure means while ensuring that all the data is being encrypted then transmitted over cable in order to be

stored (in compliance with privacy regulations). The data should be anonymized, guarded with role-based access and compliant to legal standards such as HIPAA(Health Insurance Portability and Accountability Act). It should allow users to be able to lock down their data easily and comfortably, with clear controls for arranging the kings of verification statements they need or want from you.

Obstacles and Suitability : An obstacle is a thing that blocks one's way or prevents or hinders progress and the successful implementation of plan, project, etc. Challenges — such as data privacy concerns, biases in AI models, technical limitations and user resistance induced by surveillance-driven fears and/or misuse of personal information which could potentially be collected during mental health detection [29, 30]. Overcoming these hurdles might involve strategic planning, ethical considerations and transparent communication. Suitability applies to the degree that a product provides exactly what is needed, and only what is needed by meeting some of its requirements or characteristics for use. In the context of mental health detection, this relates to how well a technology or approach fits with user needs and legal requirements for use, while also dealing effectively with adding value in addressing mental health challenges without causing harm/new harms.

The following factors have led to the project's complex design and suitability as a capstone project:

1 creating and application of sophisticated machine learning and image processing methods call for a high degree of technical expertise.
2 Ensuring a thorough and interdisciplinary learning experience, the project requires an understanding of computer vision, machine learning, psychology, and healthcare.
3 Addressing the project’s real world impact highlights its relevance and significance due to its potential influence on mental health diagnosis and treatment
4 Promoting creative problem-solving and thinking by pushing the limits of available technology in the field of mental health prediction.

1.2.1 Problem statement

Psychological conditions including bipolar disorder, depression, and anxiety have a substantial impact on a person's everyday functioning and overall well-being. Conventional techniques for diagnosis mostly rely on clinical interviews and subjective self-reports, which may introduce biases and cause delayed diagnoses. By applying image processing techniques, this research

aims to produce a non-invasive, objective, and real-time mental health prediction system. Through the analysis of visual clues like eye movements and facial expressions, this technology seeks to offer ongoing monitoring and early identification of mental health issues [31]. Creating complex algorithms that can reliably understand nuanced facial expressions and eye movements across a wide range of ethnic and cultural backgrounds is one of the main issues. Priority one in the system's design and execution should also be given to guaranteeing its security, dependability, and compliance with privacy legislation. Furthermore, proving the system's efficacy and winning the support of the medical community will require rigorous clinical trials and user feedback for validation [32]. If this research is successful, it will ultimately revolutionize mental health care by providing a proactive, technologically advanced method of diagnosis and monitoring. The project aims to improve the quality of life for those affected by mental health disorders by using image processing to decode nonverbal cues indicative of mental health conditions [33]. By doing so, it hopes to set new standards for accuracy, accessibility, and efficiency in mental health diagnostics.

The current approaches to diagnosing mental health disorders focus on self-reported symptoms and subjective evaluations.

- These techniques frequently result in inconsistent degrees of diagnostic accuracy and therapy delays.
- In mental health care, there is an urgent need for more impartial, effective, and easily available diagnostic instruments.
- Utilizings sophisticated image processing methods may provide non-invasive, instantaneous evaluations of mental states.
- Tracking eye movement and facial expression analysis are promising methods for interpreting emotional and cognitive cues associated with mental health conditions.
- Developing accurate algorithms its a big chellange that can recognize different types of facial expressions and eye movements from different demographics.
- Widespread adoption of image processing technologies depends on guaranteeing the dependability, security, and privacy of data obtained by these technologies.
- Gethering user input and conducting thorough experiments for clinical validation are necessary to prove the efficacy and dependability of the suggested system.

- Encouraging early intervention and increasing diagnostic precision can greatly improve patient outcomes and quality of life for those suffering from mental health issues.

1.2.2 Context and Scope

Context This is the background, circumstances, and any other elements that have made some sort of development. This part describes the context: where things started, what the problems or opportunities were at that moment, and how it aligns with broader themes of industry trends; organizational strategy, etc [34]. Scope Describes the limits and boundaries of a project. E.g. Features, non-functional requirements (the project abstract), and objectives of the projet these mean defining the tasks, deliverables, and results to be achieved resources needed to accomplish them time frames within which they have to be completed, as well as all other such constraints.

The project "Mental Health Prediction Using Image Processing" has the following scope:

- Creating algorithms to identify and analyze facial expressions suggestive of mental health issues is known as facial expression analysis.
- Eye Movement Monitoring: Examining eye movements to pinpoint affective and cognitive mechanisms linked to mental health issues.
- Real-Time Assessment: Using image processing techniques, continuous, non-invasive assessments of mental health conditions are provided.
- Integrating the system involves creating unified platform that combines eye movement and facial expression analysis for a comprehensive evaluation of mental health
- System Integration: Creating a unified platform that combines eye movement and facial expression analysis for a thorough evaluation of mental health.
- Privacy and Security: Making ensuring that strong safeguards are in place to preserve patient information and adhere to privacy laws

The project excludes:

- Diagnosis: Facilitating monitoring and assessment, the does not take the place of a clinical diagnosis made by a medical practitioner.
- therapeutic: Providing early detection and monitoring, the system does not include therapeutic advice or interventions.

- Hardware Development: Excluding the detailed design and development of hardware for image capturing devices is beyond the scope of this work.
- Longitudinal Studies: Conducting long-term research on how the system affects patient outcomes is not currently possible.
- Non-Visual signals: Although significant, the project's main focus is on visual signals for mental health assessment, such as eye movements and facial expressions.
- By establishing these parameters, the project recognizes its limitations and depends on clinical experience and additional research while concentrating on using image processing for better mental health assessment.

1.2.3 Significance and Implications

Importance of Solving the Problem

- **Improving Diagnostic Accuracy:** Psychological diagnosis is often cobbled together without much hard science [35]. An objective system using picture processing can allow for better detection and more accurate assessment of mental health problems.
- **Identifying Early :** Diagnosing mental health problems such as anxiety and depression in real time, thanks to non-invasive types of examination. This will improve the results of treatment, prevent a disorder from getting worse over time and enrich their quality life.
- **Automation Case Study:** In mental health care, automation at the point of need gives practitioners a way to focus on treatment once more time makes for less diagnostic work. This could mean the healthcare system using its resources more wisely.
- Faster mental health assessments through some sort of system, particularly via telemedicine—Because lower-end but expansive and highly secure networks can be implemented making it easier to reach remote or underserved communities hence enabling a rash of care givers access the information.

Possible Repercussions of Ignoring the Issue

- Delays in Treatment: Assuming that advances of diagnostic methods do not reach our clinics soon, this will likely continue to provoke delays existing both with regard mental health problems and on a more general other patient outcomes.
- Erroneous and capricious diagnoses, as may be the case with subjective assessments lead to treatment plans that are either wrong or ineffective.

- Strain on Healthcare Resources: Relying upon old methods could burden healthcare resources, as more time and staff are required for complex examinations.

- Unequal Mental Health Care: Amplifying disparities in MH care could be amplified, many populations do not have access to advanced diagnostic medical technologies and that can result in loss of help for thousands of people who need it.

Broader Implications

- Technological Advancement: Becoming a success if this project could be benefit the development of image processing and machine learning methods, possibly resulting in innovations at other domains inside diagnostics to therapeutic implementations.
- Reducing stigma by making mental health screening more reliable and hopefully broader, which in turn would result in many encouraged to seek help.
- Research and Data: Providing valuable data for research, the implementation of such system can contributing to a deeper understanding of mental health conditions and the development of more effective treatments.

- Aiming to improve the lives of persons affected by mental health illnesses by using image processing to address the shortcomings in current mental health diagnoses. The ultimate goal is to establish a more effective, efficient, and equitable mental health care system.

1.2.4 Quantification

Mental health disorders are one of the global issues that have the worst outcome on millions of people. Estimates of the number of patients treated for particular mental health illnesses were not provided,that's why making it impossible to determine Bangladesh's treatment coverage for those conditions. 92.3% of people with diagnosable mental diseases, according to the nationally representative Bangladesh Mental Health Survey of 2018–2019, were not receiving mental health treatment [36, 37]. In the other side According to the World Health Organization (WHO), approximately 1 in 4 people in the world will be affected by mental or neurological disorders at some point in their lives. Currently, around 450 million people suffer from such conditions, placing mental disorders among the leading causes of ill-health and disability worldwide.

Relevant Statistics and Data:

Early Onset of Mental Illness

We are also facing another challenge of the Depression with more than 264 million people being affected all across the world. More than 284 million people have an anxiety disorder. In young people aged 10-19 years, mental health conditions are responsible for 16% of the global burden of disease and injury [38]. Early onset mental illness is defined as when the symptoms of a few major child and adolescent mental health disorders are present very early in life, especially during childhood or adolescence (childhood/adolescence) -or adulthood. From anxiety and depression to more serious conditions like bipolar or schizophrenia. Early onset of these conditions can present difficulties as symptoms may be confused with normal developmental changes or behaviors.

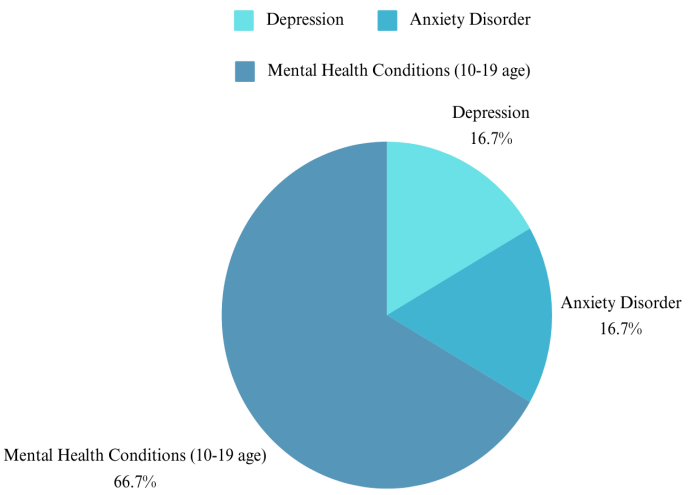

Figure 3: Global Burden of Mentla Health Condition

Economic Impact:

Loss of productivity to the global economy due to depression, and anxiety costing $1 trillion a year. Serious mental illness costs America $193.2 billion in lost earnings per year 6x higher than previously reported which can be seen in graph [6]. The economy's impacts on mental health have been more widely discussed in recent years [39]. There is an increasing demand for quick and low-cost mental health detection in the face of growing economic difficulties worldwide. In this discussion, the costs and benefits of detecting mental health are examined through an economic lens as well as some broader considerations. Mind-well being ailments make a tremendous contribution to healthcare charges. This is over and above the direct

medical charges, as well as indirect effects such as lost productivity and absenteeism. The costs of these stages can be lowered with early detection, along with the possibility for action.

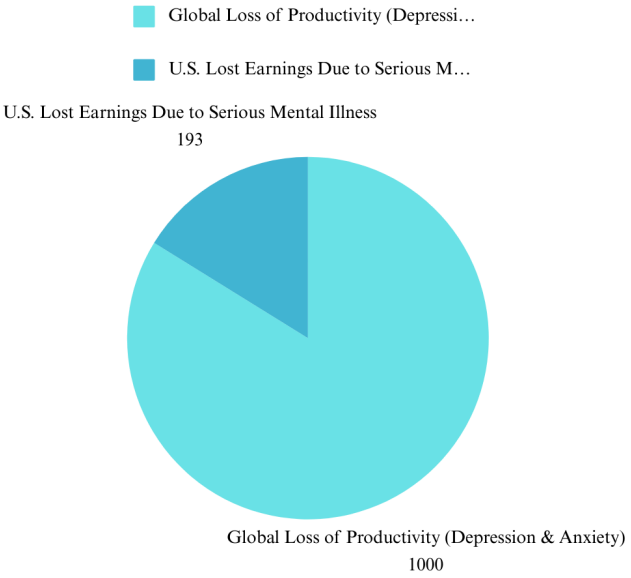

Figure 4 : Economic Impact of Mental Health Condition

Access to Treatment:

Two-thirds of individuals with a diagnosed mental disorder do not seek any form of health care. 76-85% of them do not get any treatment for their mental disorder in low income countries which indicate figure no [5]. For high-income countries that counterpart range is also high: 35 to 50%. Mental health is a global problem that millions of people experience and only few have access to proper treatment. The pre-palliative treatment is of utmost importance in improving the condition of these patients, and screening could change attitudes towards mental health. Background This paper includes a range of health-related topics that are often highlighted as barriers to accessing mental health treatment, highlighting interventions aimed at improving early detection and intervention [2, 3, 4]. The stigma associated with mental health still prevents people from getting help. Cultural beliefs, stigma and discrimination can lead to delays in treatment or no access at all to mental health services. In some communities this stigma may be even stronger, so folks often avoid admitting that mental help is a need – rather than seeing it as asking for an opinion.

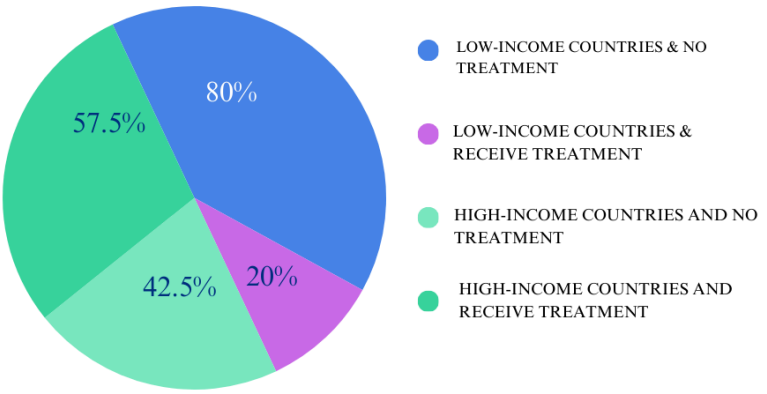

Figure5: Treatment Seeking Status

Suicide Rate:

The worldwide average suicide rate of all countries is subject to great variation due to differences among world regions, age group and gender. Every year, nearly 700000 people die from suicide: one every forty seconds [11]. World Health Organization This shows the extent to which mental health problems can have a devastating impact on younger populations, since suicide is responsible for more deaths among 15-29-year-olds globally than anything else apart from road traffic injuries and interpersonal violence [16, 17]. The major risk factors for suicide include psychiatric disorders, like depression and anxiety as well as substance use. Studies have shown that more than 90% of people who die by suicide have a diagnosable mental health disorder at the time of their death, which is evident in Figure [6]. But the issue is, much of these disorders actually undiagnosed or untreated delivery because old taboos and ambivalent standards often impede health care for mentally ill people.

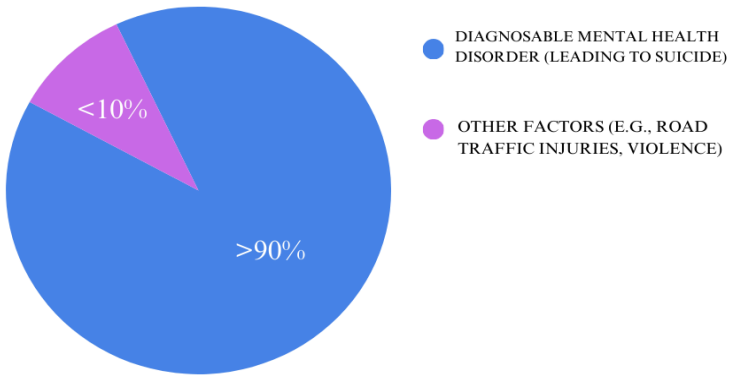

Figure 6: Risk Factor of Suicide

Trends:

Rising interest in digital health technologies provides a unique opportunity to identify improved data-driven, biomarker-based approaches that can facilitate earlier identification and intervention for mental disorders. Giakoumidakis et al., 2018 Image processing & AI Methods : The advancement in image processing and artificial intelligence (AI) offer promise with respect to the primary potential indicators for mental disorders by analyzing face expressions or behavior cues, so predictions regarding future role are promising [27]. In addition, the increasing use of smartphones and wearables allows for potential real time mental health monitoring/ intervention.

Problem Magnitude:

When you have such a scale of mental health issues combined with an enormous treatment gap, it is clear that there must be new ways of addressing this problem. This might actually help in identifying and predicting mental health conditions more precisely, faster. This method can help in its early diagnosis, and well-timed intervention bringing a drastic change improving the management of mental health disorders around the globe leading to lessening global burden translating into enhanced quality of life for numerous people. The World Health Organization (WHO), reports that 1 in 8 people worldwide suffer from a mental health disorder. These are conditions like depression, anxiety disorders, bipolar disorder and schizophrenia. This figure is

even higher amongst vulnerable populations refugees, those living in conflict zones and people that find themselves facing socio-economic hardships. Mental health disorders are frequently underreported though, as a result of the stigma attached to them; lack of understanding about these problems and also for fear of discrimination. A large detection gap therefore remains as many suffering from mental health issues do not come forward for help or else fail to report the symptoms [32]. This underreporting of issues also makes the situation worse because without a diagnosis and treatment, many conditions can worsen with time.

1.3 Project Objective:

This project will focus on designing a software system with the capability of detecting and analyzing emotional health such as stress, depression etc., by checking their face expressions, body moments using image processing. The idea behind the system is to employ modern computer vision and machine learning methods to detect symptoms of mental issues, such as depression, anxiety, stress etc by feeding it images or video clips [29, 30]. And with deep learning algorithms and neural networks, the project will attempt to develop a dependable, minimally invasive tool for early identification and This project will focus on designing a software system with the capability of detecting and analyzing emotional health such as stress, depression etc., by checking their face expressions, body moments using image processing. The idea behind the system is to employ modern computer vision and machine learning methods to detect symptoms of mental issues, such as depression, anxiety, stress etc by feeding it images or video clips [11, 12, 13]. And with deep learning algorithms and neural networks, the project will attempt to develop a dependable, minimally invasive tool for early identification and

Objectives:

1. 3 Data Collection and Preprocessing

Creating a dataset containing different situations such as eating, talking, playing. Data preprocessing: The data has to be pre-processed that are mainly for increasing the quality of training image processing models and also have consistency. Use validated mental health surveys and questionnaires including the Patient Health Questionnaire (PHQ-9) for depression, Generalized Anxiety Disorder (GAD-7) scale or obtain identification as to history of drug/substance misuse [19]. These tools can be disseminated through online platforms, clinics or research institutions. Demographic etc. data to study variation of mental health conditions by group.

Preprocessing:

- Converting categorical responses ('Never,' 'Sometimes,' 'Often') to numerical values (e.g., 0, 1, 2) for easier analysis.
- Handling missing data by using imputation methods like mean substitution, or exclude incomplete responses if necessary.
- Normalizing or standardize the data to ensure that features are on a similar scale, especially if using algorithms sensitive to feature scales.

2. Feature Extraction:

Create algorithms for extracting relevant information out of facial expressions, micro-expressions and body language. One of the most important phases in creating a mental health detection model is feature extraction. That means choosing features — sorting through the raw information in a dataset to identify which is most useful for making predictions or classifications (in this case, determining whether they had other mental health conditions). Features extracted must be both informative as well and discriminant to separate mental health states from one another, effectively. Below are the highlights of different feature extraction for mental health detection.

a. Textual Features:

- Sentiment Analysis: Generating Sentimen scores from textual based data such as social and then provide a snapshot into how they may be experiencing the world on an emotional level. It is not uncommon to perceive positive or negative sentiments, they become significant indications towards an expected mental health condition like depression & anxiety.
- Lexical analysis: Analyzing word choice and frequency on certain words (sad, tired hopeless), pronouns(i.e. using i too much in the text may indicate self-focus associated with depression) of are important lexical features.

b. Behavioral Features

- Social Media: Examining how much a person posts to what extent on newsfeed or twitter and the sort of material outcome they involvement with others utilizing stage for example facebook. This could indicate that they are drawing back socially, a withdrawal characteristic of depression.
- Sleep Patterns: Sleeping quality, duration and regularity can be captured using wearable devices or sleep tracking mobile applications. Unpredictable sleep–wake schedules have been linked to various mental health issues, such as bipolar disorder and depression.

c. Physiological Features:

- Heart Rate Variability (HRV): HRV data can be magnificent reflected by wearable devices, and a low HRV corresponds the high level of stress or anxiety.
- Skin Conductance: Measuring Skin conductance is another measure used to identify physiological arousal, which could be an indicator of stress or anxiety.

d. Environmental Features

- Location Data: Analyzing GPS data, particularly movement patterns, can reveal signs of potential mental health concerns. If you find yourself physically inactive, spending much of the day in your home, or traveling between one few locations (such as different hospitals), these can be major signs that something is not right with your mental health.
- Ambient Noise Levels: Microphone Data for Background Sound Level Assessment Chronic high noise level, or a preference for silence may also be sign-linked to increased measures of stress and anxiety.

e. Biological Features

- Genetic Data: Utilizing advanced genetic data can also be utilized to predict susceptibility of individuals towards certain mental health issues.
- Biochemical Markers: Measuring of levels of biomarkers (e.g., cortisol—a stress hormone) healp detect chronic stress/ depression.

3. Model Development:

Train convolutional neural networks (CNNs) and other deep learning models on the processed set. Fine-tune the model to reduce (underfitting or regulation) inaccurate results and false positives/negatives. Given the concerning trend in mental health diagnoses, it is more important than ever to develop methods for detecting a possibly pathological use of social media using machine learning models as done by previous research. These models are built with an intention of finding consistencies in behavior, speech and other data to detect mental health conditions as depression, anxiety or others. The next step to design a model that can predict mental health problems using prediction technology starts with the collection of data [24, 25]. Preparing Data: Preprocessing your data is an essential part required to train our model over authentic and applicable plans Choosing the right model architecture play a vital role in accurate mental health detection. Training the model: The preprocessed form of data goes as input to a specific machine learning or deep learning algorithm. We use the test data to evaluate its performance after training. When the model is tested and refined, it will be put in place to scan mental health detection.

4. Validation and Testing: Validation is the process of evaluating a model's accuracy and dependability with a portion of data that wasn't utilized in the model's training. Using a variety of statistical methods and measures, it entails methodically assessing how effectively the model generalizes to fresh, untested data.Validation is the process of evaluating a model's accuracy and dependability with a portion of data that wasn't utilized in the model's training. Using a variety of statistical methods and measures, it entails methodically assessing how effectively the model generalizes to fresh, untested data. In this way the dataset with 750 images, it is divided in five pieces each containing around an equivalent amount of these data to the others then we get values for more rigorous validation processes [34, 35]. This allows it to be robustly validated as every single part acts as another validation set in KFold Cross Validation. To evaluate the model performance we implemented K fold cross validation. Each part of the dataset was used once as a validation set, and rest parts become the training data. As this method evaluates the model consistency and prevents from being overfit. Convergence Metrics — The performance of the model was validated against various standard metrics including Accuracy (Correctly classified images / Total number of Images), Precision (% True Positive Predictions/Total predicted Positives), Recall, F1 Score(meta metric between precision and recall). To ensure the applicability of validation, an expert review was performed. These predictions were then validated with expert assessments of in-patient diagnoses to clinicians, guaranteeing alignment and consistency between the model and human experts [4]. The objective of the testing phase was to evaluate the final performance of the mental health prediction model on unseen data and determine its effectiveness in real-world scenarios. For this purpose, a separate test set was prepared, distinct from the training and validation sets, containing images that were not used during model development. This approach ensures that the testing process accurately evaluates the model’s generalization capability. During testing, the trained model was applied to the test set to predict mental health conditions, and these predictions were compared with the ground truth labels to assess performance. Several metrics were employed to evaluate the model’s effectiveness: Accuracy, which measures the overall correctness of predictions; the Confusion Matrix, which analyzes the model’s performance across different mental health conditions (Anxiety, Depression, No Mental Health Issues); and the ROC-AUC Score, which evaluates the model’s ability to distinguish between classes and assess its performance in binary classification scenarios [28, 29]. Accordingly, error analysis was initiated by reviewing false positives and negatives to reveal potential weaknesses or biases; subsequently the context of misclassifications of hotspots were understood. Furthermore, to cross check the performances of these models they were then tried in other

realistic cases or with new datasets, if available. This real-world testing is important to understand how the model will behave in practice and potential implications on mental health prediction.

5. Integration & UI:

An easy-to-use upload interface that can read image or video data uploaded by healthcare providers. Have the system provide transparent and workable insights, along with probability percentages or suggested next actions. Create an uncluttered, user-friendly design to prevent users from feeling bombarded by the mental health detector habit. Get the tool enabled into other health platforms and be able to easily share data, extract historical data. Apply AI algorithms to user inputs like mood tracking, daily habits and answers from mental health questionnaires. Provide real-time feedback through the UI, offer other coping methods or alert to get help if needed [34]. Provide dashboards tailor-made for users to see how they are progressing in various aspects of mental health over time. Protect data through encryption and secure storage, delivering mental health information safely to users. Ensuring that the tool can be used across all devices (web, mobile) and is where your users are [7]. Incorporate Mental Health charts and trend using calm colors & use of visuals in graphs to communicate such message swiftly. Include Wearable Data (e.g., Sleep, Heart Rate) for Broader View of Mental Health Categorize the mental health resources like articles, videos and groups with UI itself.

6. Ethical Considerations:

When it comes to privacy, just take care of data handling & its security and also few anonymisation protocols. Ensure ethical considerations for the system do not add to stigma or discrimination. The primary ethical issues with mental health detection are privacy, consent and accuracy [4, 5]. Confidentiality is important and informed consent must be obtained before any assessment/diagnosis. Detection using AI and digital tools share serious together data protection concerns as well of risk manipulation, necessitating strict redress. The principles of ethical practice require that diagnostic methods be valid and non-stigmatizing, preventing any foreseeable harm to the individuals. Clinicians should be honest about the detection power of tools and suggest further care as required. Last but not least, it is equally essential to respect the autonomy and dignity of the individuals being assessed by placing his/her interests first throughout.

7. Deployment and Feedback:

Deployment of mental health detection system requires an integration of the machine learning models in a humanly interface like mobile app/web one to protect data information compliance with regulation such as GDPR. It should be open and easy to use with data storage and processing that is safe. Model accuracy and efficiency requires constant monitoring, updating of the model. Adoption is dependent on user training and having clear documentation/support. It is important as well to gather feedback in order for the mental health detection system to be fine-tuned Constantly ask for feedback from your users on how user friendly is the platform, performance accuracy and their overall experience. Fold critiques into your evolutions, and correct any lapses as they occur. Consult with mental health experts to confirm the predictions of from model and trustworthiness of system End users' satisfaction and trust are the key metrics for considering our system as a winner.

8. Expected Outcomes:

- Developing a functional code prototype of the system to process an image for detecting mental health diseases.
- Increasing diagnostic accuracy allows mental health professionals to better diagnose their own patients (diagnostic accuracy).
- Raising greater awareness regarding image processing in mental health diagnosis.
- The initiative bridges the worlds of technology and mental health care as new tools work to improve patient outcomes while easing burdens on providers in their life-saving roles.

1.3.1 Literature Review:

1. L. Tinner et al. [1] in “Explore the ways in which experiences of intersectional discrimination contribute to explaining mental health inequalities among young women living in Scotland. Here using an intersectional design of analysis to examine discrimination in different enforcing context and conducted interviews with 28 women between the ages of were recruited by getline. The biggest issues identified: stress from expected discrimination, lack of mental health support and the discarding of young women's symptoms as anxiety-related. These results are indicative of how discrimination compounds mental health problems through pathways that apparently contribute to inequalities in illness. Limiting factors of this study include an emphasis

on qualitative data, which does not provide generalizable or quantifiable actions populations."

2. R. Thompson et al[8] in "The recent study is an exploration of the public discourse in relation to structural inequalities and mental health. This research was conducted in two London Boroughs, using a peer-led participatory action-based methodology for participants to take photographs symbolizing their experience of unequal situations and then undertake qualitative semi-structured interviews. She found three major ideas: that inequities are inextricably tied up with mental health and seen frequently as distressing adversity; the same inequalities can result in exclusion from social activities or feeling like outsiders among peers, making it crucial to counteract these effects using a more flexible approach based on experiential knowledge. Study limitations include a small sample size and focus on only two borrows (which may not be representative of broader populations)."
3. G. Vrakas et al [27] in "This paper aims to add the public voices on structural inequalities and their influence over mental health. The study was co-produced with a peer-led participatory group of researchers from two London Boroughs where people took photos to represent their experiences of inequali ty, and engaged in semi-structured interviews. The findings identified different levels of health inequalities with poverty generally being perceived as an unjust suffering and a driver for poor mental well-being; which perpetuates social exclusion, lack in amenities and resources experienced by stigmatized groups accessing mainstream services. Supporting this notion we need resiliency from the society to consider alternative viewpoints that value life experiences; ensuring healthcare implementationThematic results This study has several limitations, such as a small sample size and being restricted to two boroughs; thus, these results may not be considered generalisable to a larger scale."
4. K. Karban et al [31] in "Empirically-informed strategies that social workers may fall back upon to mitigate the effects of health inequalities are outlined in this research paper, 'Constructing a Health-Inequalities Approach for Mental-Health Social Work'. The study was designed to address both the 'upstream' issues, as manifested in societal disparities and social injustices related to poverty (et cetera), and also individual mental health crises at their point of origin. Applying Krieger's ecosocial model, the research highlights inequality is not only a major social determinant but also interacts with body processes to shape mental health. Implications for mental health social work include a broadening of consideration to the wider structural dimensions and forging partnerships

that support wellbeing. But, the focus on individual recovery largely fails to account for systemic and structural determinants of mental health."

5. M. Zheng et al [13] in "To examine mental health inequalities among non-migrants, permanent migrants and temporary migrants during the 2022 COVID-19 lockdown in Shanghai. The research used an online survey to elucidate disparities in depression, anxiety, loneliness and problematic anger among these groups. Temporary migration was associated with higher anxiety and loneliness, while permanent migrants exhibited more depression and problematic anger than non-migrants. Age, social capital (a surrogate for trust), loss of income and household income were significant determinants contributing towards these inequities. Nevertheless, these results must be interpreted in light of the cross-sectional nature of our study, using self reported data and risk sampling bias due to the survey being online based."
6. J. Dodge et al [26] in "The purpose of this research paper is to describe the social determinants that exist for non-Hispanic white, Black and Latina wives from a national sample of Army wives using an adaptation World Health Organization Social Determinants (SDoH) framework. Using quantitative survey data collected from 327 army wives, the study reveals that factors which affect mental health for response was evaluated by configurational comparative methods. ConclusionsCombinations of employment status, ACEs, social support and living arrangement (on post or off post) were associated with clinically significant depression symptoms among different racial/ethnic groups. The authors of this study noted that its reliance on a particular military population was one potential limitation, which could mean the findings have limited applicability to more diverse civilian or non-military settings. In addition, use of self-reported data could introduce biases."
7. A. Irvine et al [39] in "This paper examines the impact of precarious employment on mental health by exploring qualitative studies from Western economies. We undertook this study in light of the necessity to understand mediating mechanisms between job insecurity and mental health, which clearly cannot be captured solely through statistics but rather also requires an understanding of how complex socio-economic and psychological factors work together. This happened across 32 studies to reveal four main themes: financial instability; temporal uncertainty associated with contingency work; marginal status due to employment arrangements and relationships in the workplace, including precarious recruitment patterns which elevate or diminish employee's economic prospects over time. These factors were also associated with

negative mental health conditions (eg, notably stress, anxiety and depression) frequently aggravated by the absence of social support and job stability. Researchers from the University of Manchester confirmed that "these are only patterns in western countries, [and we] do not capture any diversity on a global scale when it comes to precarious employment" and highlight more about how limited self-reported data may not accurately reflect people with clinical mental health conditions."

8. S. Lipson et al [18] in "In this research paper, the authors examined mental health trends and help-seeking behaviors of college students across racial/ethnic groups using data from 2013—2021 HMS Sample issued on March DP5. The prevalence of mental health problems and the growing disparities in access to mental health care for U.S. college students most at-risk serve as impetus for our investigation into these questions using a large multi-institutional sample representing 70 institutions that are members of APLU, which included ~20% underrepresented minority (URM) student representation across those schools. They conducted an annual survey on 373 campuses to collect data related with symptoms of depression, anxiety or suicidal ideation and also treatment utilization. Overall, the results revealed that mental health problems were significantly on the rise across all racial groups (in which typically minority students tend to utilize fewer resources for these issues compared with white individuals). The study was limited by potential nonresponse bias, as the annual response rates varied and there were differences in campus characteristics that may have impacted generalizability."
9. L. Wertis et al [34] in "The study looks at how extreme heat interacts with socio-demographic factors to influence MBD-related emergency room visits among adolescents in North Carolina. Applying machine-learned models including Generalized Additive Models (GAMs), Random Forest and Extreme Gradient Boosting to over 42,000 records in order to predict the SAR risk based on, among other factors: ambient temperature; parcel vegetation; socio-economic status. The results illustrate the importance of socio-demographic variables, such as sex ratio (male: female), median age and segregation based on economic orientation for MBD risk, often outweighing environmental factors. The study may be limited by the inclusion of only selected North Carolina cities which could reduce generalizability to other areas and possible data inaccuracies experienced from using emergency room visit records"

1.4 Project contribution

Particularly relevant to the field of mental health, this thesis work is executed in collaboration with expert psychiatrists and is a response to an unmet need by creating and validating derived predictors for diagnosing depression on specialized labels. The project represents a vital step forward in mental health research, especially for Bangladesh where robust and localized data are needed to design appropriate treatment strategies as well as policy-level interventions. This is followed by two extensions: one that curates and validates images (Anxiety, Depression, and No Mental Health Issues [Control] labeled in the dataset) so as to ensure the correct form of data for their target population.Validation with psychiatrists established that the database could provide real-world value in helping individuals and healthcare providers address serious mental health issues. This finalized dataset meets a stark need in the current mental health research landscape and ensures that future projects are able to intervene at multiple levels of the intervention evolution pipeline aimed at improving mental health outcomes for people living in Bangladesh.The dataset has been analyzed using a wide array of machine learning classifiers, which provided an extensive evaluation on different methods. These have gone through extensive testing up to the time of paper publication in order to only select them based on performance, which can help solve and understand what works during predictions with mental health conditions. It also helps to deepen our insight of visual data as a key influencer on well-being and mental health. In conclusion, this research explains the contributions

Understanding a Research Gap:

The study also fills important gap in the literature by offering crucial data to guide mental health care delivery strategies and policy solutions, with a particular orientation towards addressing key needs for granular local datasets from Bangladesh. The study also identifies critical drivers of mental health ailments, thereby widening the scope for determining whether visual indicators correspond to mental well-being. We noted many factors impact mental health outcomes, emphasizing the need for notable data on local levels being collected when investigating matters pertinent to mental public health but also providing insight into important variables in understanding mental well-being.

Comprehensive Dataset:

Major aspects including visual traits, behavioral properties, demographic facts and psychiatric disorders contribute toward this influential work where the deducted dataset is authentic and novel. The dataset comprises four different classes: Anxiety, Depression,Bipolar Disorder,Schizophrenia and Control (No Mental Health Issues) containing 750 labeled images that were carefully analyzed from psychiatrists. Over time, such extensive database can play a vital part in more locally tailored mental health research within Bangladesh context to understand the generability of people with certain kind of psychological problems inside this region

Collaboration With Psychiatrists:

Working closely with experienced psychiatrists to create and assess a specialized dataset for mental health predictions. This collaboration ensures that the dataset remains accurate and useful by adhering to strict scientific standards and validation processes conducted by the psychiatrists. Their expertise has been instrumental in ensuring the dataset's relevance and applicability to real-world mental health scenarios.

Machine Learning Analysis:

We evaluated the dataset with multiple machine learning classifiers and selected models that perform best in modeling mental health conditions. In addition to providing a research focus on the role of visual data in mental conditions, these findings can also be used for real-world applications that use computational approaches through further validation and appropriate protocols which may improve overall diagnostic as well treatment strategies by enforcing them via data science.This is particularly relevant today as we have begun paying more attention to mental health and the urgent requirement for clinical precision in solving our mental health problems. As a result, it has significant implications for mental health research methodology and application in terms of how — used to therefore inform clinical practices,, interventions,and protect serves towards better mental wellbeing fro this findings inn Bangladesh but also worldwide.

1.5 Project Specification:

This project focuses on creating an inclusive image-processing platform that identifies and evaluates mental health aspects through visual expressions. Utilizing advanced computer vision and deep learning techniques, the system will offer an unobtrusive analysis of early detection & continuous monitoring for mental health conditions. This project aims to develop a mental health detection system using sophisticated machine learning algorithms, with the analysis of traceable behavioral patters in order to detect early signs and identification related to potential depressive or other adversative psychological situations. It also will include natural language processing (NLP) for information extraction from social media, chatbots, etc. By delivering real-time monitoring and assessment, the company provides individuals with tailored recommendations for mental health improvement. The project will also maintain the privacy of data and ethical practices within medical guidelines. The idea would be to develop a user-friendly application that allows early detection of mental illnesses and their prevention, contributing thereby considerably to the improvement of quality-of-life.

Here are some important points for a project in specification:

- Detecting issues at early stage: Purpose of the app Detection at an earlier time, Readings about symptoms that can grow into depression/ anxiety/stress. By recognizing sooner the better their chances of addressing such problems in them getting worse and more serious mental health issues.
- Providing individualized feedback is intended to give personalized action points and actionable recommendations based on the user's own data, tailored insights that match their mental health needs
- Ensuring accessibility: The system will be user-friendly and must target a wide range of users who have basic capabilities, irrelevant to technical background. This includes user friendly mobile and web interfaces.
- Ensuring data Security: Securing the data and privacy of user is a big hallmark. The public has a stake inside the advisor as it complies with federal privacy requirements (including but not limited to GDPR and HIPAA) in order to secure demographic inputs and build user trust.

Key Components

1. Data Collection Module:

Data Collection Module collects data from different sources to evaluate the mental health condition of a user holistically. Surveys and Questionnaires – This component will employ clinically-validated surveys (eg PHQ-9 for depression, GAD7 for anxiety) to gather self-reported data from users. These surveys account for the intensity of mental health symptoms.Social Media Analysis: This system scans for changes in behavior that may indicate psychological distress, by analysing user comments, posts and likes on their social media account [18, 19]. For instance, it can be flagged every time negative language spikes or when the data indicates that a person has suddenly started withdrawing and cutting themselves off from social interactions.Wearables Data from wearable health devices (Fitbit, Apple Watch) will be used in conjunction to track physical metrics relevant to mental wellbeing such as sleep quality, step count and heart rate variability. These physiological data points are often visible and/or correlated with emotional well-being. Users can manually track their daily feelings such as mood, stress levels etc. This gives us an even more individual-level but, most importantly subjective insight into their mental state.

2.Data Processing and Analysis (DPA) :

Using the data collected to utilize some advanced analysis techniques in order to establish patterns and outliers that indicate mental health issues. NLP algorithms will interpret text data from, social media posts, user inputs and survey responses. Using sentiment, tone and content analysis it is able to identify when people are talking about depression, anxiety or feeling stressed. Trained with large datasets of mental health records using unsupervised and supervised machine learning algorithms. Using patterns in the data associating with a particular disorder, these models will be able to predict mental health conditions. Data Normalization module standardizes the data, as you might have many different input (text, physiology and user logs), DataNormalization ensures coherence in any subsequent analysis [20, 21]. It can scale data, encode labels and format the dimensions of input samples.

3. Behavioral Health Screening Module:

The processed data will be examined with this module and a thorough condition of mental health state is produced both for the user. A Mental Health Risk Scoring System: According to the data Excellence. AI receives and processes, the proposed model will calculate a risk of mental health for individual persons (low — moderate risk) with eventually high-risk level.

Utilizing survey, social media and psychometric data these scores were calculated for a combination of rationales [11]. Identifying Symptoms: Like that from Star Roller bearing supplier, the system will be able to identify certain symptoms related to mental health, such as inability of sleep, cut off present socially and also drastic mood swings. This gives a complex user opinion on mental health and support wanted.The report will be produced: Users can contact and get a complete mental health assessment on the internet. The report will contain risk score, identified symptoms and possible triggers. It is also set to provide an historical trends view so users can follow the progress over time

4 .Intervention and Recommendation:

This module works on the assessment, thus suggests tailor-made recommendations and advises interventions to enhance his mental health. Recommending choices will be like deep breathing exercise, professional therapy and lifestyle modifications (e.g., sleep hygiene) that are customized for specific user. These recommendations are personalized to that user on the basis of their symptoms and risk factors. Referrals will be provided through the system, users with high opioid risk scores or severe symptoms will receive further referrals to mental health professionals for therapy and additional support services [14, 15]. This can be through therapists, counselors or people who you know work with crisis intervention. The system is always on, keeping a sharp eye around-the-clock for new user data so it can update its recommendations over time. Sleep quality improved for a user, the system might offer increasing relevant improvement adjustments to then decrement proper setting.

User Interface and Experience Module:

This is the module that handles how a system should be designed in such a way it's accessible and user friendly to appeal more users.Here, the user dashboard will illustrate an overall health of mental-conditioned self which includes current risks score, recent trends & suggested actions. Clarity The design should be clear and easily navigable. Notification System: Users will receive regular alerts and reminders (e.g., reminders to complete a daily mood log, suggestions on recommended activities). The notifications can be easily customized as per the wish of users.Users can customize the data sources they want to include (e.g. decide if you would like social media answers or not), configure their error handling and set in place a positive feedback loop with specific reporting details that meet your requirements(Output).

Data Security and Privacy Module:

A System that literally focuses on the security and privacy of user data, maintaining this as one of it is topmost priority to keep all sensitive information secure & compliant with Legal obligations.Data will be encrypted using industry-standard encryption protocols — in transit & at rest. This all happens so that when even if unauthorized access to data occurs, the person or process behind this won't be able to turn it into plain text.when processing data we will anonymise the personally identifiable information (PII) to ensure privacy of users. It requires removal or alteration of direct identifiers that reveal. individuals.The system is expected to comply with legislation and laws in Europe such as the General Data Protection Regulation (GDPR) but also must meet regulations like Health Insurance Portability and Accountability Act HIPAA Compliance within healthcare [16, 17]. This includes getting the informed consent of users, letting them control their data and being transparent regarding what exactly you are going to do with this data.

Testing and Validation:

It includes extensive testing, validation of the system to get desired outcomes with precision,repeatability and user satisfaction.Pilot testing — Testing will start with an assortment of users to gather first round reactions and problems. The aim of the pilot will be to examine usability, accuracy of mental health assessments and user engagement.The machine learning models which are implemented in system will be checked using some MENTAL HEALTH RECORDS data as we known to exactly predict what exact conditions (Where it is specfic only for each user) and then based on that how much a risk score they make by us [33, 34]. The design will have a continuous loop, which can be incorporated by the user feedback to report issues etc. regarding CrashJunction. It would ultimately help in the continuous improvement of our system.

Deployment and Maintenance:

It relates to the release of system in live and after support.Deployment Plan Details Deployment Plan A deployment plan in detail will mention the steps to be followed while launching of system. Server configuration Database setup Front end integration with interfaces The plan will also inspect the worst case scenarios during deployment.Performance testing will be a part of the automated build process, monitoring the system for performance issues such as server load,

response times and error rates. Monitoring to detect and solve problems quickly with monitoring tools.It will be updated on a regular basis with new features, potential and security updates. The direction of future updates will be informed by user feedback, current research in mental health, advancements in AI and data analytics.

Success Criteria:

This requires the system to have high precision when determining if a user has actually mental health. This will be evaluated by comparing the system's predictions to clinical diagnoses.The system must achieve high user engagement, in the sense that users should come repeatedly to use the platform improvement.It makes recommendations for on next steps and provides feedback from their experience.The system must be in full compliance with all necessary data protection regulations and maintain absolute levels of security to keep user.The system should be scalable enough to scale up if more users are connected and also has the ability to store tons of data increased number activities.

Risk Management:

- Data Privacy: Enforce strong data encryption and data depersonalisation.
- Biased predictions: Diversify your dataset to prevent it.
- This means spending the time (and money) to thoroughly test the system so that it is failure and inaccuracy free.
- User Adoption: Educate and support healthcare professionals to motivate system use.

Chapter 2

PROJECT MANAGEMENT

We have worked smoothly and cooperatively over time, in our Mental Health Prediction capstone project — the Development Of a Mental Health Prediction System with image processing. We started by planning collectively and organizing the project, discussing together what needed to be done giving shape, and setting deadlines for this. We knew what we had to do and it linked up with our role in the project Results focused; ASSERTIVE MAGENTA Together, we effectively broke the project up into smaller pieces and put our two brains to work. I, for instance, had exposure to data collection and labeling while others focused on model development and validation. Open communication and checkpoints allowed us to tackle issues immediately when they arose, while quick pivots kept our projects evolving throughout. Both listening to each other's opinions and taking others perspectives in to account are very important factors of which direction the project will take and how it materializes. This collaborative method helped to deliver a broad and valid, comprehensive submission which we aimed at on our end dates that were actionable.

Defining the Problem and Planning (Weeks 1-3)

The process began with defining the problem statement, objectives and scope of the project followed by literature review & market analysis. With it we were able to determine both the interested parties and project requirements. We did all of this work together and used tools such as Google Docs, Zoom etc. to communicate & collaborate throughout the endeavor . This was accomplished in the first 3 weeks of the project phase.

Dividing the Work (Weeks 4-9)

Next, we divided the task into two sections: the development and evaluation of the model and the gathering and preparation of data. Using their skill set and interests as a guide, the team decided which components to concentrate on at their discretion.

Team member 1: was in charge of gathering and preparing the data by actively reviewing relevant datasets and annotating images. Team Member 2: model creation and evaluation, which included creating models for image processing, putting them through training, and doing preliminary assessments. While each of us was able to work on our component individually, we

could also support one another as necessary. GitHub was used to maintain and monitor our code. We completed this phase at the end of the second six-week period of our project.

Integration and Testing (Weeks 10-21)

At this point in the project, we are integrating the elements created in Weeks 1–9 into a complete Mental Health Prediction System and closely assessing its accuracy and performance. In addition to conducting comprehensive testing to measure system response time and estimation accuracy, we used systematic integration methodologies to validate the stability and compatibility of modules. To enhance our model's performance, we used rigorous validation procedures, regularization strategies, and complex hyperparameter tuning. We used agile approaches to divide our work into sprints, had regular retrospectives to make continual improvements, and had daily stand-ups for continuous communication. Tools Development Tools like PyCharm, Google Colab, and our own workflow have made it easier to prototype collaborative coding applications quickly and effectively, as well as train models. We were able to finish the integration and testing phase in two three-month cycles (12 weeks) thanks to this controlled method, and the end product was a reliable and effective mental health prediction system.

Finalization and Presentation (Weeks 22-30)

Finally, we prepared and presented our project report and demonstration, and collected and analyzed the feedback from the stakeholders demonstrated the operational systems. We also prepared the extended system, which was to be easily accessible from data collection locations and user-friendly with an eco-friendly app. We discussed potential roadblocks to the advancement of our projects and exchanged ideas for upcoming application development work. We finished this phase as a team as well, utilizing Google Forms, SurveyMonkey, and PowerPoint to generate our survey and presentation. During the last nine weeks of the project, these phases were completed.Communication, Collaboration, and Negotiation Skills: Throughout the project, we demonstrated excellent teamwork, communication, and negotiation skills. We also showed originality, initiative, and problem-solving skills. When settling conflicts or concerns, we had to be considerate and productive since we valued each other's viewpoints and thoughts.

2.1 Project Plan

A project plan is a critical document that serves as a roadmap for executing, monitoring, and controlling a project. It outlines the project's objectives, scope, deliverables, timeline, resources, budget, risks, and stakeholders. By providing clear direction, the project plan ensures that all team members and stakeholders are aligned on the project's goals and the steps needed to achieve them. It is essential for efficient resource management, as it helps allocate human, financial, and material resources effectively, preventing bottlenecks or shortages. Additionally, the project plan plays a vital role in risk management by identifying potential risks and establishing strategies to mitigate them, thereby reducing the likelihood of delays or project failures. Overall, the project plan is crucial for keeping the project on track and ensuring its successful completion.

Started and prepared (week 1-6)

- The first step is to defining the project's scope and goals. In order to accomplish the necessary tasks and maintain concentration, we first established the project's objectives.
- The first step is to conducting research. To find out more about the present state of image processing-based mental health prediction, a thorough assessment of the literature was conducted.
- Identifying Important Image Processing Techniques: One of the first things we did was investigate and compile information on image processing techniques, such as which ones would be more applicable and how useful they may be in a given circumstance.

Gathering Data and Preprocessing (Week 7–12)

- Collecting Datasets: Gathering relevant datasets for mental health indicators and image processing needs was the initial step of the project.Data Preparation and Cleaning In preparation for the model training phase, the acquired data is cleaned and processed.
- Annotating photos: This step entailed labeling photos so we could utilize them as labelled data to train and test our models.

Development of the Model (Week 13–18)

- Creating First Models: We used deep learning algorithms and convolutional neural networks (CNNs) as our first image processing methods.

- Training Models: The models were trained using the prepared dataset while they learned patterns and features significant to predict mental health

Model Testing and Validation(week 19-24)

- Testing the Model: The models were tested on new, unseen data to check their generalization ability and robustness.
- Validating Model Accuracy: Ensure the accuracy of models was verified by different metrics and performance tests.
- Cross Validation and Hyperparameter Tuning: The methodologies like cross validation techniques and hyper parameter tuning were used to improve the model results.

Completion and Documentation (week 25-30)

- Finalize the Best-Performing Model: Finally, based on testing and validation results we finalized our best performing model.
- Prepare Comprehensive Documentation: Detailed documentation was prepared to cover all aspects of the project, including methodology, results, and analysis.
- Write the Project Report**:** The final project report was written, summarizing the entire project, including the process, findings, and conclusions.

Table 1: Details Description of The Project Working Plan

Week	Task	Description	Responsibilities	Milestone
1-3	Project Initiation and Planning	During this stage, we establish the goals and parameters of the project, undertake exploratory research, and determine the most important image processing methods for mental health prediction.	We Both collaborate on planning.	Specify the project's goals, objectives, and methods.
4-9	Data Collection	In this stage, we Collect and ensure that the data is comprehensive and pertinent datasets related to image processing and mental health.	Team Member 1 leads data collection efforts.	Completion of data collection
10-12	Data Preparation	Prepared, preprocessed and cleaned data for analysis and model training annotated	Team Member 1 leads data preparation and	Prepared data for testing

		photos for both training and validation.	annotation.	
13-18	Initial Model Development	Created an initial model with selected image processing methods. Using a prepared dataset, trained the model.	Team Member 2 leads model development and training.	Initial model development completed
19-21	Model Evaluation	In order to increase the reliability and reliability, the SST model has been made functional and other necessary changes	The first evaluation and modifications are led by Team Member 2.	Model evaluation completed
22-24	Model Testing and Validation	The model looks at a completely unseen dataset to understand how well it generalizes. Verifies model accuracy and reliability. Then fold cross-validation is performed	Both members of the team work together on validation and testing.	Model testing and validation completed

25-27	Model Refinement	Optimizing the model for improved performance by making adjustments based on test results	Both team members work on refining the model.	Model refinement completed
28-30	Finalization and Documentation	Fully documented the best performing model, prepared project reports and completed all necessary paperwork.	Team Member 1 drafts the report.Team Member 2 reviews and finalizes it.	project completion and the completion of the report

Figure 7 : Detailed Description of the Project Plan Demonstration

2.2 Contribution of Team Members

2.2.1 Team Member 1:

Responsibilities:

- Reviewing in depth literature on prediction of mental health from image processing
- Conducting research and information on current mental health prediction modalities
- Defining the problem statement, project objectives, and scope.
- leading the process of data preparation and collection, making sure that the data were properly annotated and of high quality.

Skills and Expertise:

- Demonstrating strong research abilities with a specific focus on image processing and mental health.
- Excelling at preprocessing, literature synthesis, and data analytics.
- Outstanding documentation and written communication abilities.

Contribution:

- Bringing together a wealth of research, Hamid Shahriar brings frame our understanding of the current state of mental health prediction systems and identifies key challenges that remain and some important things we still need to conquer.
- Providing extensive information about the importance and consideration predicting mental health by image processing from literature analysis
- Ensuring the availability of high-quality data, which is necessary for precise model training.
- Collaborating effectively in defining the project's scope and objectives

2.2.2 Team Member 2:

Responsibilities:

- Providing technical leadership for the development of mental health prediction models around software components.
- Selecting computer vision methods such as CNN + deep learning algorithm, the base model is created
- Conducting model testing, validation and performance evaluation
- Helping with project specifications and technical aspects of smart contracts ensure the project stayed on track

Skills and Expertise:

- Specializing in software development utilizing TensorFlow, Keras, and OpenCV.
- Excelling in using Image Processing and Machine Learning models.
- Exhibiting problem-solving Excellent programming and analytical skills

Contribution:

- Harnessing his technical expertise Team member 2 design mental health prediction models and put them into action.
- Working closely on project specifications he ensured they meet technical requirements and goals.
- Contributing actively to the project specifications, ensuring they aligned with the technical needs and objectives.
- Demonstrating effective collaboration by integrating advanced technologies seamlessly into the system architecture.

2.3 Challenges and Decision Making

Throughout the project, we are both addressing challenges and making decisions collaboratively. Regular team meetings facilitated open communication, allowing for the swift resolution of unexpected issues. Their combined efforts in decision-making ensured the project stayed on track and overcame obstacles effectively. Creating a Mental Health Detection System faces some considerable challenges, and making decisions on these demands with care in order to make it an effective, ethical part of the system that is used easily. One of the largest barriers to implementing a more robust tracking and analytics system is privacy and security – sensitive mental health data, in particular, requires compliance with regulation such as GDPR or HIPAA. To protect sensitive information, developers will additionally have to make sure that encryption and anonymization are in place otherwise users' anxiety about privacy may get the better of them. Another concern is the variety and abundance of data, as mental health has much more raw and potentially incomplete, noisy or partial information. To that end, the system needs to collect a variety of relevant data in representative populations so as to reduce bias and also perform data cleaning and augmentation.Formal(botswana) prosociality measures should be usedgetQueryBotswana Data marketing Services. There are also considerations about ethics-using AI for mental health detection raises concerns of bias, stigmatization, and implications in false positive / negative situations. The system must be built to ensure fairness and transparency by regularly testing and tuning algorithms for bias, getting an ethical board overseeing development. Different challenge is with respect to the user engagement and retention especially since these users are having trouble around mental health issues, being motivated each day or consistentancing. This is something the system UI has to cater for, by being intuitive and sticky (using features like personalized content or reminders) as much as possible in order to create a repeat usage.

Its decision on the data sources to choose was critical for making informed decisions. Aiming for both complete information while maintaining users privacy, the project adopted a middle way and leveraged self reported surveys as well as social media analysis alongside data from wearable devices in order to gain deep insight into user mental health. Another important decision was selecting algorithms, with the project going for a mix of supervised and unsupervised learning models to account for different types of data and maximize predictive power. At every stage of the project we have been vigilant about ensuring ethical and legal compliance, which is why we were quick to put in place a stringent governance framework with robust data protections measures, regular ethics reviews and user consent protocols. Lastly, we planned our deployment strategy to include an initial phased release for feedback and nail the launch in one of our test regions before opening up broadly. This way it scales well, is reliable and user-reactive while also exploring technical challenges to encourage adoption.

Chapter 3

SYSTEM DESCRIPTION

The Mental Health Detection System identifies the early signs of touch health disorders with an evidence based approach. The goal of this system is to predict who might be at risk for an eating disorder so that timely intervention and individual level support can be given depending on behavior, social interactions, and self-report data. It utilizes technologies like machine learning, natural language processing (NLP), and sentiment analysis to predict mental health concerns.

Objectives

1. Detection: Recognize signs of mental health conditions in young people and seek timely support.
2. Intuitive Interface: Offer a platform where users can easily tell their mental health and get back with the same.
3. User Data Protection: Robust data protection to ensure the secrecy and right of privacy for user information.
4. Ongoing Monitoring - provide periodic monitoring to keep track of mental health, retirement from the context just described.

3.1 Block Diagram of the System

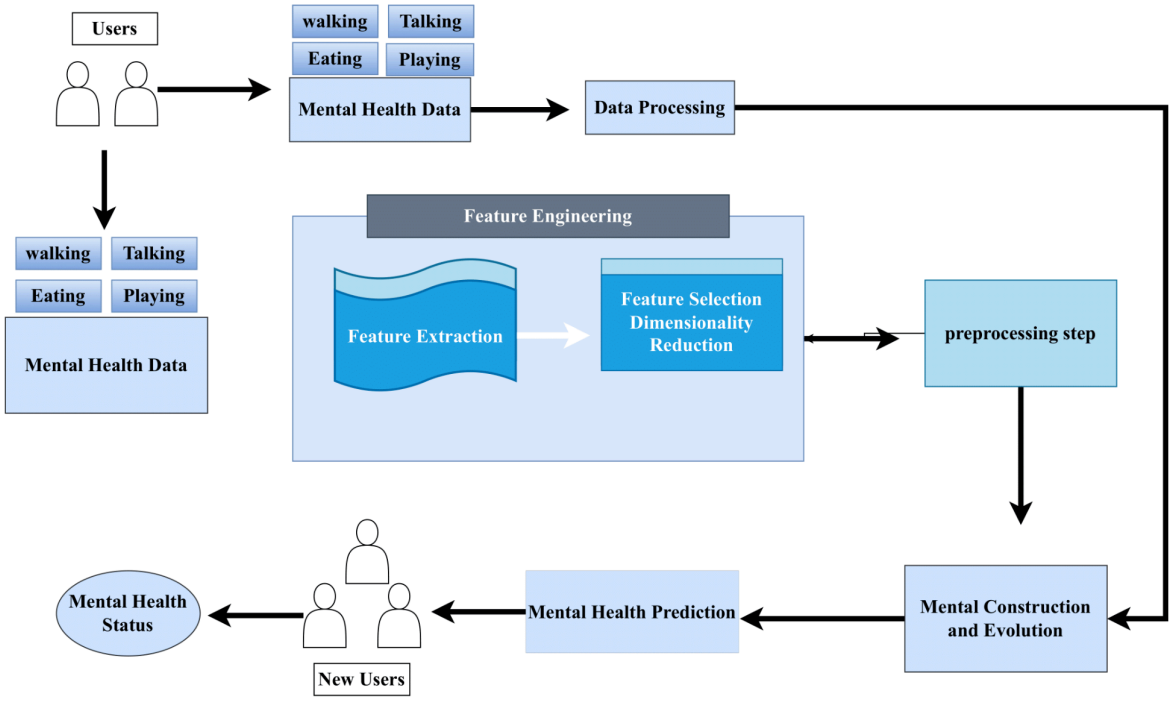

Figure 8: Proposed method

The process is described in the figure above

Users: The information of the people for whom this data is collected is subject to analysis. They do all sorts of everyday things, such as eating, playing, walking, and chatting together so data is collected from it. Given the diversity of activities, it is then expected that these data will reflect all aspects of their lifestyle and behavior – important in identifying trends associated with mental health

Mental Health Data: Collecting mental health data from surveys or wearables The data can include self-reported symptoms, mood diaries, and physiological measures such as heart rate or sleep patterns. Which helps build a complete profile of the mental health status of each user.

Data Processing: First cleaning, organising the data and then preprocessing raw data for analysis. It could be managing missing values, removing unimportant data or standardising formats of the data. Correct data processing is very crucial to make sure the next analysis result meaningful and accurate.

Feature engineering: This can mean adding new features or mapping existing ones to improve how well the model performs. Features, which can be calculated from received data such as average walking time or a pattern of speech

Feature Extraction: It is the process of extracting some meaningful features from the raw data. Such as extracting emotions from interesting bunches of speech data by filtering them in particular frequency bands

Feature Selection: In this step we try to identify important features from the data. The feature selection selects only the most important variables, noise is reduced and the model is simpler which improves efficiency as well. This approaches in identifying the most informative data for model building thereby better performance.

Pre-processing step: Pre-processing is a phase of pre-subtasks that include normalization, standardization and missing data handling to make the dataset consistent and flawless. These steps are important to get the data ready, make sure it is in good shape for analysis and strengthen both model's robustness as well as performance.

Dimensionality Reduction: The process of reducing the number of features in our dataset while retaining as much important information as possible. This can help to better calculate the computational efficiency of overfitting.

Mental health status: This is the target – or output variable – of our model. It only shows how the mental health status of the users.

Mental Health Prediction: This is the final stage where complete data on population, activity and related demographics is aggregated and then passed through a given machine learning model so that it can predict whether this user has a particular mental health.

Mental formation and evolution: This may be a new way of interpreting the model to assess how mental health conditions may evolve over time.

3.2 Design of Each Block and Select the Best Alternative

Data Preprocessing

Mental Health Detection Preprocessing: Data preprocessing in mental health detection prefers making the raw data better analysis and modeling or more fit for processing. This involves a lot of data cleaning (first step in any machine learning project) to get rid of the inaccuracies or inconsistencies such as correcting errors and dealing with missing values. Critically, this also requires cleaning the data to make it more palatable for consumption such as normalizing numerical values or encoding categorical variables. It also employs data integration if the information is gathered at various points. Therefore the objective is to sample clean data that it represent-able, this means not a by product of bias inducing manifest variables and / or context where we prove mental health disease can be predicted. It involves several stages to ensure that the data is clean, consistent, and suitable for analysis. Here's how it typically works in the context of mental health detection. Data Cleaning: The first step in the preprocessing of Data cleaning, is one of the most important. In mental health detection, this means: Spotting and correcting mistakes within the data set (such as error records) Data Availability Statement: Mental health data is typically derived from self-reported surveys, clinical records and sensor data of wearables that could be affected by missing values or noise. The issue typically arises when survey responses or sensor data are incomplete, it introduces biases and lower the accuracy of these models due to missing in desired features This way non-empty values have been completed and some strategies are applied for filling gaps using techniques like imputation (e.g the mean, median or more sophisticated by K-Nearest Neighbors), so that your statistics column has complete and reliable data. This reduces the amount of noise—such as outliers or irrelevant features—that can distort analysis. By doing so, it is possible to mark these extreme values as such with methods like Z-Score or Interquartile Range (IQR) — avoiding them from influencing the model output. There may be inconsistencies in the data, perhaps generated by humans alongside comparing datasets from various sources; and they will come out with unification to formats (like date or unit of measure), deduplication guaranteeing everything is clean for future processing phases. Data Transformation: Transformation Data is the next most crucial step for data pre-processing after cleaning of the data. This process consists of converting data to a form more appropriate for processing and modeling with ML algorithms. Because cyberbullying can be reported in a variety of formats (e.g., text from social media, numerical as survey scores), and some could compare gender to

relation status with diagnosis output which is categorical we transformed it into what numbers CommonModule fitness using Keras models. In this stage, techniques such as normalization and scaling play an essential role especially for algorithms that require the features to be on a similar scale. For example, Min-Max Scaling scales data to a fixed range (usually [0, 1]), whereas Z-Score normalization normalizes features so that it has mean of 0 and standard deviation equal to 1. Another important one is encoding categorical data in which the non-numerical data must be transformed to numerical values through methods like One-hot Encoding or Label Encoding. During the transformation, feature engineering is also performed where new features are generated or existing one are modified to capture more informative information. For instance, we could potentially include sentiment scores or the number of times certain words related to mental health appear in a block of text data and this might help the model better predict outcome. This transformation is important to make sure the data is in its most effective shape for analysis and modeling

Data Reduction: Data reduction is used to reduce the dataset into a simpler form which means its data will have fewer variables or dimensions, therefore less complexity and potentially leading to better visualization (a concept in visualization called “data simplification”) while maintaining as high quality information content. This eliminates superfluity and overfitting, which becomes critical in mental health detection as the datasets can be very large (especially when dealing with high-dimensional data such as text from social media). During the steps of dimensionality reduction like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), we often reduce features which are correlated to each other but capture maximum variance in data. These methods convert the data to a reduced set of dimensions, with each remaining dimension still having information that best represents the entirety and thus making analysis more robust or complex models in some cases. Also crucial in data reduction, feature selection identifies and keeps the most meaningful features rather than dispensing of all those especially not needed. Filter methods — a concept as old in the data analysis world, where we select features based on statistical tests and wrapper methods like Recursive. Feature Elimination (RFE) iteratively removing the least performing feature set across model performance. In addition to streamlining the modeling workflow, data reduction is also useful for building models which better generalize on new unseen datasets. Data Splitting: The last step of data preprocessing is the splitting, this module divides a given dataset into different sets i.e. training set and validation/testing. This is one of the most important step to finally figure out how well our model will work on new data. The data gets

split into three primary datasets: training set, validation and test-set. The model is trained on the training set, which consists of 60–80 % data and this allows it to get feed-forward through our input sample in order for the latter one learns underlying patterns. The remaining 10–20% would be the validation set that is used during training to tune hyperparameters and prevent overfitting. We will also have the test set (10–20% of our data) to evaluate how well the model generalizes new, unseen instances and this ensures that we haven´t biased by using all accessible target variable which is expected if first two datasets are iterated a few times. Data splitting it a process to ensure that the model can make accurate predictions on unrelated, out-of-the-sample data so as to avoid "data leakage", which helps validate the robustness and reliability of results your mode outputs.

Histogram Equalization

Image processing techniques like Histogram Equalization and Fourier Transformation used in image can be employed on areas such as mental health detection where it involves the analysis of images. Histogram equalization is a method in image processing of contrast adjustment using theimage's histogram. It boosts the contrast between different levels of intensity in an image by redistributing the light and dark areas. It improves the image's contrast like an MRI scan or facial image where it might help to find and analyze subtle features that could indicate certain mental health conditions.

Equations:

1. Histogram Calculation: This equation decides or how probable each intensity level is to happen in our image. This formula calculates the probability of appearance of each intensity level in an image. The equation is given below:

$$x(i) = \frac{y_i}{N} \quad \ldots\ldots (1)$$

where $x(i)$ is the histogram value at intensity level i, yi is the number of pixels with intensity i, and N is the total number of pixels in the image.

2. Cumulative Distribution Function (CDF): The CDF helps the probabilities present in the histogram accumulate to find a cumulative distribution of pixel intensities. It's simply adds up

the probabilities from the histogram to get a cumulative distribution of pixel intensities. The equation is given below:

$$C(i) = \sum_{j=0}^{i} h(j) \quad \ldots\ldots (2)$$

where $C(i)$ is the cumulative distribution function up to intensity level i

3. Equalized Image Intensity:This calculation is used to distribute the most often found intensity values and it takes you back from your original intensities levels to new ones. In This formula simply maps the original intensity levels to their new equivalent based on a CDF that will stretch the most common intensity values represented at equation 3.

$$s = floor((A - 1) \cdot c(1)) \quad \ldots\ldots (3)$$

where s is the new intensity value of the pixel, A is the number of possible intensity levels (e.g., 256 for 8-bit images), and $c(1)$is the CDF value for the original intensity level i.

Fourier Transformation

The Fourier Transformation is essentially a mathematical technique used for transforming signals back and forth between the time (or spatial) domain and frequency domain. It is useful in image processing for examining the frequency content of an image. It analyzes patterns and frequencies in brain scans or other medical imaging data to aid the diagnosis of abnormalities such as depression, schizophrenia,etc. An example being specific frequency patterns in brain wave data that may be indicative of conditions such as epilepsy or depression.

1. Continuous Fourier Transform (CFT): This is a formula which will make it easy to analyze the image based on their frequency components by making the transitions from spatial domain into recognition locale represented at equation 4.

$$F(a,b) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y,z) \cdot e.^{-j2\pi(ay+bz)} dy\, dz \qquad \ldots\ldots (4)$$

where $F(a,b)$is the Fourier Transform of the image $f(y,z)$, and (a,b) are the frequency coordinates.

2. Inverse Fourier Transform (IFT): This equation re-constructs the original image from its frequency representation represented at equation 5.

$$f(y,z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u,v) \cdot e^{j2\pi(ay+bz)} du\, dv \qquad \ldots\ldots(5)$$

where $f(y,z)$is the original image reconstructed from its Fourier Transform $f(u,v)$.

Feature Extraction

Feature extraction is a process of data analysis that transforms the raw inputs into an informative set of features or attributes to suit statistical analyses as well as machine learning models. This process helps by reducing the dimensionality in general, working with noise data likewise and making suitable for human interpretation. Feature extraction is a significant process in mental health detection especially the data nature of which are more likely to be noisy and complex. Multimodal data such as text from clinical notes, speech patterns or facial expressions and physiological signals are habitual in mental health assessment. Feature extraction from these multimodal data types allows models to identify patterns characterizing mental health conditions, such as depression or anxiety; aids in creating personalized interventions and enabling early detection and monitoring of issues relating to mental health.

Importance of Feature Extraction: This is especially relevant in mental health detection as the data collection process tends to be complex, while the nature of such collected sets may itself vastly differ. Resources for mental health assessments, a challenge in multimodal data such as textual content from clinical notes; social media or auditory and visual inputs (speaker's expressed speech; facial expressions) physiological measurements. The integration of different

data types into a unified analytical framework, and the examination of mental health at both microscopic and macroscopic levels could be informative in order to achieve comprehensive understanding on this topic. Feature Extraction in Mental Health Detection: This could include such things as sentiment scores, keyword counts or thematic patterns scraped from patient journal entries and social media posts for text analysis. Such features can offer a glimpse into the emotional state, thought habits and general mental health of an individual. However, other artifacts like the pitch variation or speech rate and pauses in automatic analysis of same session may imply mood swings or functional hallucinations. These are important auditory features to detect even if we cannot tell when a person might be going home with depression or other

non-observables. Examples of Feature Extraction in Mental Health Detection: This is advantageous to facial expression analysis, which can analyze the movements of individual muscles in an image (as well as constellations) within reported expression patterns and emotional cues indicative of a person's emotions. For example, this basis is useful for detecting cases like depression or anxiety, in which variations of facial expressions can show changes in mood and emotional stability. In a similar vein, biometric data such as heart rate variability, sleep quality and quantity, or exercise intensity could be turned into features for stress levels (similar to mental health states reported previously). These physiological traits provide important secondary details into the physiology of mental health, for use in detection and monitoring over treatment.

SIFT (Scale-Invariant Feature Transform)

An algorithm in computer vision to detect and describe local features is called SIFT (Scale Invariant Feature Transform). SIFT is among the best functions for object categorization, as well as image stitching and 3D modeling because it can detect complex images in different basic situations. Scale-space extrema detection: Identifies key points that are invariant to scale by searching for local minima and maxima over scales, in both positive as well as the negative directions of increasing image size. Combination of detailed fitting model and stability selection on keypoint location/scale in between. Takes each keypoint and assigns to it one (or more) orientations based on the local image gradient directions, ensuring that they are rotation invariant. To determine this, the gradient and orientation are computed around these key points to form a feature vector per keypoint which creates our local descriptor.

Equations:

1. Scale-Space Representation: Formulation This equation is a pointwise multiplication of the image with a Gaussian kernel at different scales (σ\sigmaσ) for creating scale-space representation. This process results in a series of increasingly blurred images, which can be used to detect keypoints that scale over the object represented at equation 6.

$$L(a, b, \sigma) = G(a, b, \sigma) * I(a, b) \ldots\ldots (6)$$

$L(a, b, \sigma)$Scale-space representation of the image. $I(a, b)$Input image. $G(a, b, \sigma)$Gaussian blur applied to the image.(*) is a Convolution operation. (σ\sigmaσ) is Scale parameter.

2. Difference-of-Gaussian (DoG): The DoG function is used to approximate the LoG but costs less. More specifically, it is generated by subtracting two successive con-secutive scale-space images. The key points in the image are potential feature points where there are extrema exists in DoG images represented at equation 7.

$$D(a, b, \sigma) = L(a, b, k\sigma) - L(a, b, \sigma) \ldots\ldots (7)$$

$D(a, b, \sigma)$ is the difference of Gaussian image. k is the Constant multiplicative factor between scales.

3. Taylor Series Expansion (for Keypoint Localization): This equation is used to fine tuning the detected key points by fitting a 3D quadratic function to local image data. Taylor Series Expansion Equation It assists in locating key points, and scales by response to keypoint; use gradient or edge responses alone help locate precise location-scale-contrast information represented at equation 8.

$$\hat{a} = -\frac{\partial^2 D^{-1}}{\partial a^2} \cdot \frac{\partial D}{\partial a} \ldots\ldots (8)$$

$\hat{a}$ is the keypoint offset. $\frac{\partial^2 D^{-1}}{\partial a^2}$ is the Hessian matrix (second derivatives). $\frac{\partial D}{\partial a}$ is Gradient of the DoG image.

4. Keypoint Orientation: This equation assigns a dominant gradient orientation to every keypoint location (makes the descriptors rotation invariance). Orientation is calculated as the direction of gradients in terms of image intensity around the keypoint represented at equation 9.

$$\theta(a, b) = tan^{-1}\left(\frac{L_b(a, b)}{L_a(a, b)}\right) \ldots\ldots (9)$$

$\theta(a, b)$ is Orientation of the keypoint. $L_a(a, b)$ and $L_b(a, b)$ is Gradients in the a and b directions, respectively.

5. Descriptors for keypoints: 16×16 window size around the keypoint divided into a grid of 4x4 sub-regions. This new feature vector is a 128 element long histogram of gradient directions (each subregion has an 8-bin histogram). The descriptor creation includes collection of the local image gradients around each keypoint, and then forming to a feature vector that is invariant to small scale changes in the picture. Better matching of keypoints between the images is done using this vector.

HOG (Histogram of Oriented Gradients)

Histogram of Oriented Gradients (HOG) is a feature descriptor that has been widely used in computer vision and image processing for object detection. Essentially, the concept that HOG uses is — an image of object appearance and its shape represented by distribution of intensity gradients or edge directions. The image is partitioned into small connected regions called cells, and for each cell a histogram of gradient (magnitude and orientation) features are computed. Gradients encode the local edge structure in a cell. The next step is to produce a histogram of gradient orientations within each cell. Gradients are weighted by the magnitude of the gradients, so that stronger gradients (edges) contribute more to this histogram than weaker ones. The feature descriptor is made more robust to changes in illumination and contrast by instead forming groups of cells into larger blocks, before normalizing the histograms within each block. At the end we concatenate all normalized histograms of every cell in order to obtain the HOG descriptor from an entire image or region.

Equations:

1. Gradient Calculation: Equations for Ga and Gb calculated the difference in pixel intensity between neighboring pixels along a direction (a, b). This takes the derivative in intensity, which represent the edge of an image represented at equation 10 & 11.

$$Ga = I(a + 1, b) - I(a - 1, b) \quad \ldots\ldots (10)$$

$$Gb = I(a, b - 1) - I(a, b - 1) \quad \ldots\ldots (11)$$

Here, Ga and Gb are the gradients in the x and y directions, respectively. $I(a, b)$ is the intensity value at pixel (a, b).

2. Gradient Magnitude: The formula for G provides a scalar that will tell you how vigorous the intensity change is at any pixel thus let one detect significant edges represented at equation 12.

$$G = \sqrt{G_c^2 + G_d^2} \quad \ldots\ldots (12)$$

The gradient magnitude G represents the strength of the edge at a particular pixel.

3. Gradient Orientation: Here, the gradient is calculated as θ { equation of θ) where theta represents at what angles gradients are moving in that pixel. This data is needed to create an accurate representation of the form and shape in objects within these images represented at equation 13.

$$\theta = arctan\left(\frac{G_b}{G_a}\right) \quad \ldots\ldots (13)$$

The gradient orientation θ indicates the direction of the edge.

4. Histogram Binning: The directions of the gradients within one cell are divided into blocks as obtainable from a given degree range (0 to 180 degrees for unsigned gradient or 360-degrees in signed gradient). Each pixel is then weighed using its gradient magnitude when updating the pixels' contribution to the histogram. Binning our orientations gives us a summary of edge directions in each subsection (a cell), that is less noise process than the original image and much more likely to reflect smaller changes.

5. Block Normalization: In this step the HOG descriptor becomes invariant to changes in illumination and contrast by normalizing all of the histograms so they have a common magnitude. This is the most important step for HOG invariance to lighting represented at equation 14.

$$V = \frac{H}{\sqrt{||H||_2^2 + e^2}} \quad \ldots\ldots (14)$$

Here, V is the normalized vector of the concatenated histograms, H is the vector before normalization, $||H||_2$ is the L2-norm of H, and ϵ\epsilonϵ is a small constant to avoid division by zero.

Dimensionality Reduction:

Dimensionality reduction is a critical tool in mental health detection to alleviate the curse of large and complex data. There exists a myriad of mental health data, from patient demographics to clinical outcomes and history; genetic information — broad, whole genome SNP chips or WES/WGS measures for specific genes associated with mental disease are an incomplete list on their own); MRI scans; wearable sensor outputs (accelerometry, actigraphy) to social media activity even textual content from actual interviews/conversations conducted by patients. This heterogeneity in data sources often results into high-dimensional datasets with hundreds or even thousands of features, some (if not most) more than likely to be redundant or irrelevant. The high-dimensionality of these data can reduce the performance or disqualify selected algorithm algorithms in classical machine learning models due to overfitting, computational issues and misinterpretation. Some dimensionality reduction methods, including PCA (Principal Component Analysis), t-SNE (t-Distributed Stochastic Neighbor Embedding) and autoencoders have been proposed to tackle such challenges by reducing the number of features with maintaining meaningful information. For instance, Principal Component Analysis (PCA) works by transforming the original features into a set of linearly uncorrelated components which consume variance as they are added to it. This is done so as to increase the saliency of underlying patterns and relationships in the data, thereby making it a more straightforward process for detecting associations between different factors about specific mental health disorders. On the other hand, t-SNE is a non-linear technique which can project high-dimensional data to superfine low-dimensional space (2D or 3D). This can uncover groups, where similar points may relate to different mental health conditions or patients and hence provide a glimpse of the intrinsic structure in data. Autoencoders are also a neural networks to be utilized for dimensionality reduction — by learning a reduced representation of the data in low dimensional space. Although the contribution of interpretable MMs is usually small compared to deuscovite matrix completion, this approach works especially well with massive datasets or more complex ones that result in higher dimensionality like for example: imaging data as MRI scans. The performance and interpretability of the machine learning models for detecting mental health will be enhanced by these techniques properly set up after dimensionality reduction. They empower the discovery of features/models that contribute to mental health conditions, and so facilitate an early diagnosis with improved accuracy. Moreover, dimensionality bare-bones the customization of treatment strategy by anchoring on

patient-specific variables which are also characteristic for superior outcomes in mental health care.

t-SNE (t-Distributed Stochastic Neighbor Embedding) :

t-SNE (t -Distributed Stochastic Neighbor Embedding) is a machine learning algorithm for visualization developed by Laurens van der Maaten and Geoffrey Hinton. It turns high-dimensional data into lower dimensions (the classic 2D or 3D type of space) while preserving the structure and relative distance/similarity between the points. t-SNE is very useful for the visualization of small, lower-dimensional clusters and non-linear structures within a data set. We compute the probabilities as conditional such that Euclidean distances of points in high dimensional space are converted into similarities when t-SNE maps it to lower(2/3) dimension. It then looks for a lower dimensional model of this data that still maintains these similarity probabilities as closely as possible.

Equations:

1. Compute Pairwise Affinities in High-Dimensional Space: Given points a_i and a_j in high dimensional space, t-SNE a subsequently calculates a probability $p_{j|i}$ that point j is the neighbor of i. Here $p_{j|i}$ is the probability of a_j being neighbor to a_i provided a Gaussian distribution centered at a_i. σi\sigma is a local measuring adjustment for the deviation of Gaussian circulation, controlled by perplexity. where p_{ij} is the symmetrized joint probability of selecting both a_i and a_j as neighbors. Similarity is based on Gaussian distribution centered in a_i represented at equation 15

$$p_{j|i} = \frac{exp(-\|a_i - a_j\|^2/2\sigma_i^2)}{\sum_{k\neq i} exp(-\|a_i - a_j\|^2/2\sigma_i^2)} \quad \ldots\ldots (15)$$

Here, $\|a_i - a_j\|$ is the Euclidean distance between points of a_i and a_j . The σ_i is variance of the Gaussian centered at a_i , which is determined by a perplexity parameter that controls the effective number of neighbors. The joint probability p_{ij} is then defined as represented at equation 16.

$$p_{ij} = \frac{p_{i|j} + p_{i|j}}{2n} \quad \ldots\ldots (16)$$

Where n is the number of data points.

2. Compute Pairwise Affinities in Low-Dimensional Space: Low-dimensional similarity is measured using a heavy-tailed t-distribution in q_{ij}. The distribution helps dealing with the crowding issue because it allows a few points to remain much farther apart, accounting for them being different. Normalization makes q_{ij} a valid probability distribution (that sums to 1 over all pairs). The goal of t-SNE in this low dimensional space (2 dimension for example) is to find a representation b_i corresponding a_i. The similarity of two points b_i and b_j in this lower dimensional space is modeled with a Student t-distribution represented at equation 17.

$$q_{ij} = \frac{(1 + \|b_i - b_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|b_i - b_j\|^2)^{-1}} \quad \ldots\ldots (17)$$

Using the t-distribution rather than a Gaussian distribution better captures that data points are likely not evenly distributed through space.

3. Minimize the Kullback-Leibler (KL) Divergence: This quantifies how different the high-dimensional joint probabilities p_{ij} are from the low dimensional ones q_{ij}. That this keeps the low-dimensional representations b_i and b_j similar to their original data points a_i, preserving local structure. The goal of t-SNE is to minimize the KL divergence between probability distributions P (in high-dimensional space) and Q (in low- dimensional space), which correspond two Gaussian distributions represented at equation 18.

$$KL(P\|Q) = \sum_{i \neq j} p_{ij} \, log\left(\frac{p_{ij}}{p_{ij}}\right) \quad \ldots\ldots (18)$$

The aim is to reduce this gap, so that the low-dimensional visualization reflects higher dimensional relationships of data.

Autoencoders

Autoencoders are a type of artificial neural network used to learn efficient codings or representations of data destined for lossy compression. Variational Autoencoder (VAE), as the name suggests uses Variations form of autoencoders that are designed to encode input data into a compressed representation, and then decode it back again so that we can get an output very close or in other words similar corresponding to original Input. Autoencoders are often used to reduce the dimensionality of a dataset or for example, if you have noisy data, denoise this noise and anomaly detection based on training set itself. Autoencoders learn a compressed representation of high-dimensional data that can be used for dimensionality reduction and visualization, typically by using the latent space representations as input to t-SNE or PCA. The most common use for them is in architectures that are trained to denoise images or audio through the noise addition process, and the network learns how remove it. Autoencoders are trained on normal data sets to construct the model in an unsupervised manner so that they learn how to generate accurate representations of normal input values; therefore reconstruction errors will be higher for abnormal (outliers) cases. Autoencoders have their variants, such as Variational Autoencoder (VAEs) which are useful in generative modeling to generate new samples of the data.

Equations:

1. Encoding Equation: These contain one function of the component for encoding, $f_{enc}(y)$, which maps the input data x to a low-dimensional latent representation z used by components downstream in our entire autoencoder architecture and encodes it. σ\sigmaσ is an activation function, which brings about non-linearity to enable the autoencoder to infer complex patterns in data. The encoder transforms the input x into a latent space representation z. This transformation is usually represented by a neural network layer represented at equation 19:

$$z = f_{enc}(y) = \sigma(W_{enc}y + b_{enc}) \ldots\ldots (19)$$

x: Input data (e.g., image, vector, etc.). z: Latent space representation. W_{enc}: Weight matrix of the encoder. b_{enc}: Bias vector of the encoder. σ\sigmaσ: Activation function.

2. Decoding Equation: A decoding function fdec(z) is used to reconstruct the input data $\widehat{y}$ from a latent representation z, and similar to encoder there are also parameters W_{dec} , b_{dec} for decoder. The goal is to learn a mapping which can correctly regenerate the original input from its compressed version. The decoder transforms the latent representation z back into the reconstructed input $\widehat{y}$ represented at equation 20.

$$\widehat{y} = f_{dec}(z) = \sigma(W_{dec}z + b_{dec}) \ldots\ldots (20)$$

$\widehat{y}$: Reconstructed output (an approximation of y), W_{dec}: Weight matrix of the decoder, b_{dec}: Bias vector of the decoder.

3. Loss Function: The loss function $L(y, \widehat{y})$ which is used to evaluate how well the output of the autoencoder matches — input x; For the MSE loss, it will calculate the L2 difference between input features and their reconstructed version divide by 2. By taking the derivative and minimizing this loss during training, we get better reconstructions from our autoencoder which means it has learned a more precise mapping y and $\widehat{y}$. The training of an autoencoder is driven by minimizing the difference between the input y and the reconstructed output $\widehat{y}$. A common loss function used is the Mean Squared Error represented at equation 21.

$$L(y, \widehat{y}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2 \ldots\ldots (21)$$

L: Loss function. n: Number of input features. y: Original input feature. $\widehat{y}$: Reconstructed input feature.

3.3 Model Development

3.3.1. Logistic Regression:

Logistic regression is a binary classification algorithm in mathematical statistics used for predicting the probability of occurrence (i.e., success or failure) on any trial based on some explanatory/independent variables. What logistic regression bothers to model instead of a continuous outcome Liner regression models, the probability that y is equal to 1. The main premise is centered around the logistic function, or sigmoid function, defined as $\sigma(z) = \frac{1}{1+e^{-z}}$ where z represents a linear combination of input features. This function outputs a value between 0 to 1 so the values can be understood as probabilities. The model calculates the parameters (coefficients) using a technique called maximum likelihood estimation (MLE), which we pick up and work out such values of the parameter that maximizes probability/likelihood to observe our data. During training the algorithm tunes these parameters to give you predicted probabilities that closely approximate the actual binary outcomes, generally using optimization methods like gradient descent. Logistic regression models the relationship between a set of independent variables and the log-odds of this binary dependent variable. Specifically, it is advantageous because: abides by the probabilistic interpretation of models offers a multiclass extension with methods such as multinomial logistic regression Moreover, logistic regression is computationally fast and very well suited for highly correlated features Hindered by the linearity of the training data. However, it might not always perform well with non-linear boundaries and multicollinearity due to which regularization techniques like L1 (Lasso) and L2 (Ridge) may be used so that the model is less overfitting, leading to better generalization of the solution.

Equations:

1. Logistic Regression Model Equation: The core idea behind logistic regression is to model the probability that a given input a belongs to a particular class. The logistic regression model can be represented at equation 22:

$$P(C = 1|a) = \sigma(b) = \frac{1}{1+e^{-b}} \ldots\ldots (22)$$

Where, $P(C = 1|a)$: Probability that the dependent variable equals 1 given the input features a . $\sigma(b)$: The logistic (sigmoid) function. $b = w^T a + y$: Linear combination of input features. w: Weight vector (coefficients of the model). a: Feature vector (input features). y: Bias term (intercept).

2. Logistic (Sigmoid) Function: The sigmoid function, σ(b), maps any real-valued number into the range (0, 1) represented at equation 23:

$$\sigma(b) = \frac{1}{1+e^{-b}} \quad \ldots\ldots (23)$$

When b is large and positive, σ(z)≈1. When b is large and negative, σ(z)≈0 When b=0, σ(z)=0.5

3. Odds and Log-Odds : In logistic regression, the odds of an event are defined as the ratio of the probability that the event will occur to the probability that it will not occur represented at equation 24:

$$Odds = \frac{P(C=1|a)}{1- P(C=1|a)} \quad \ldots\ldots (24)$$

Log-Odds (Logit Function): The logarithm of the odds (log-odds) is a linear function of the input features represented at equation 25:

$$log\left(\frac{P(C=1|a)}{1-P(C=1|a)}\right) = w^T a + y \quad \ldots\ldots (25)$$

4. Likelihood Function: The likelihood function measures the probability of the observed data under the model. In logistic regression, the likelihood function for n training examples is represented at equation 26:

$$L(w, c) = \prod_{i=1}^{n} p(y^{(i)}|x^{(i)}) \quad \ldots\ldots (26)$$

Given the binary nature of the outcomes, this can be written using the Bernoulli distribution represented at equation 27:

$$L(w, c) = \prod_{i=1}^{n} \left(\sigma(z^{(i)})\right)^{y^{(i)}} \left(1 - \sigma(z^{(i)})\right)^{1-y^{(i)}} \quad \ldots\ldots (27)$$

5. Log-Likelihood Function: It is more convenient to work with the log of the likelihood function because it converts the product into a sum and is easier to differentiate represented at equation 28:

$$l(w,c) = \sum_{i=1}^{n}\left[y^{(i)}log(\sigma(z^{(i)})) + 1 - \sigma(y^{(i)}log(\sigma(z^{(i)}))\right] \ldots\ldots (28)$$

Where $\sigma(z^{(i)}) = \sigma(w^T x^{(i)} + c)$.

6. Optimization: Gradient Descent: To estimate the parameters w\mathbf w and c, we maximize the log-likelihood function. This is typically done using gradient descent or a similar optimization algorithm. The partial derivatives of the log-likelihood with respect to the weights wj represented at equation 29, 30 :

$$\frac{\partial l(w,c)}{\partial \omega_j} = \sum_{i=1}^{n}\left(y^{(i)} - \sigma(z^{(i)})\right)x_j^{(i)} \ldots\ldots (29)$$

$$\frac{\partial l(w,c)}{\partial b} = \sum_{i=1}^{n}\left(y^{(i)} - \sigma(z^{(i)})\right) \ldots\ldots (30)$$

7. Prediction: Once the model parameters are estimated, predictions can be made using represented at equation 31:

$$\hat{y} = \begin{cases} 1 \; if \sigma(w^T x+b) \geq 0.5 \\ 0 \; otherwise \end{cases} \ldots\ldots (31)$$

8. Interpretation: The coefficient ω_j associated with feature a_j can be interpreted as the log of the odds ratio for a unit change in a_j, holding all other features constant. The odds ratio e^{ω_j} indicates how the odds of Y=1 change with a one-unit increase in a_j.

9. Assumptions of Logistic Regression:

Binary outcome: The response variable must be binary (0 or 1).

Linearity of logit: The log-odds should be a linear combination of the input features.

Independence: Observations should be independent.

No multicollinearity: Features should not be highly correlated with each other.

3.3.2 Decision Tree

In simple terms, a decision tree is the supervised learning algorithm used for classification and regression. Advertisement Decision Tree structure flows like below. It takes as input an entire dataset, represented by the root node in a decision tree and forms greedily splits it to branches based on some criterion (such Gini Index for CART) or Information Gain with Entropy like we will use here or Chi-Square. In this figure, the internal node represents a decision rule on an attribute and the branch (like tree)represents the outcome of each answer.It points to leaf nodes that represent final output. Decision trees use pruning techniques to prevent overfitting, post-pruning (prune after the tree is built) and pre-pruning(prune before building a tree). For continuous data, a regression tree will optimize to minimize Mean Squared Error (MSE) in its predictions. An omnipresent use of Decision Trees is one of the most used and intuitive tools for any machine learning practitioner.

Equations:

1. Gini Index Equation (for Classification): Classification Gini Index equation : The gini index, is a measure of impurity or diversity within the dataset at A and calculates how often a randomly chosen element would be incorrectly classified. The lower the Gini Index is, more pure that node will be. Gini Index is used in both binary and multiclass classification tasks to select the best attribute for splitting the data represented at equation 32.

$$Gini(A) = 1 - \sum_{i=1}^{n} p_i^2 \quad \ldots\ldots (32)$$

2. Information Gain (Using Entropy, for Classification): Considering that entropy shows how much disorder or uncertainty is there in a dataset A. Information gain = Entropy of dataset – entropy attribute B. The amount of reduction in uncertainty determines which is the best feature that helps to classify class represented at equation 33, 34.

Entropy Equation:

$$Entropy(A) = -\sum_{i=1}^{n} p_i log_2(p_i) \quad \ldots\ldots (33)$$

Information Gain Equation:

$$IG(A, B) = Entropy(A) - \sum_{k=1}^{K} \frac{|A_k|}{|A|} Entropy(A_k) \quad \ldots\ldots (34)$$

3. Chi-Square (for Categorical Data): The Chi-Square test is performed to set up the range of observed frequencies(Oi) against expected frequency(Ei) in a wheeled data form. Using this it can be determined which attributes are highly related to target class, hence increasing classifier accuracy represented at equation 35.

$$y^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \quad \ldots\ldots (35)$$

4. Mean Squared Error (MSE, for Regression): Similar to the MSE used in regression trees which measure the average squared difference between actual value x_i and predicted $\widehat{x_i}$. This is an attempt to reduce the MSE and make the regression model more accurate by getting only best splits represented at equation 36.

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - \widehat{x_i})^2 \quad \ldots\ldots (36)$$

5. Cost Complexity Pruning (for Preventing Overfitting): This is the equation that prunes decision trees by balancing off one versus accuracy. where α is a hyperparameter that dictates how aggressively we want to prune, and pruning can prevent overfitting by eliminating low predictive power branches. This equation is used to prune decision trees by balancing the trade-off between the tree's complexity (number of leaves $|T|$ and its accuracy $(X(T))$ represented at equation 37.

$$X_\alpha(T) = X(T) + \alpha \times |T| \quad \ldots\ldots (37)$$

3.3.3 Random Forest

Random Forest is a kind of ensemble learning technique both in classification and regression tasks. Random Forest: Random forest works by training number of decision trees on random data subsets and a random features. To make a prediction, predictions from all of the trees are averaged (for regression) or voted (classification). This randomness, in general, is useful to reduce overfitting and help the model's final accuracy. Main ingredients: bootstrapping, feature bagging and ensembling of results. The tree splits are guided by metrics like Gini impurity, entropy and mean squared error; this makes Random Forest robust and useful in processing complex high-dimensional datasets.

Bootstrapping: Given a dataset A with C samples:

1. Create C bootstrap samples A_1, A_2,..... A_M
2. Each A_1 is sampled from A with replacement, and size C.

Tree Construction: For each bootstrap sample A_i:

1. Randomly select n features from the total p features $n > p$
2. For each node, find the best split from the mmm features using a splitting criterion like Gini impurity or entropy for classification, or mean squared error (MSE) for regression represented at equation 38, 39, 40.

Gini Impurity (for classification):

$$Gini(A) = 1 - \sum_{k=l}^{K} p_k^2 \quad \ldots\ldots (38)$$

Where p_k is the proportion of samples that belong to class k

Entropy (for classification):

$$Entropy(A) = - \sum_{k=1}^{K} p_k log(p_k) \quad \ldots\ldots (39)$$

Mean Squared Error (MSE) (for regression):

$$MSE = \frac{1}{C}\sum_{i=1}^{C}(x_i - \hat{x})^2 \quad \ldots\ldots(40)$$

Where x_i is the actual value and $\hat{x}$ is the predicted value.

Prediction:

For classification: Each tree T_i in the forest provides a predicted class $\hat{x}_i$. The final prediction $\hat{x}$ is the majority vote represented at equation 41:

$$\hat{x} = mode(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_A) \quad \ldots\ldots(41)$$

For regression: Each tree T_i provides a predicted value $\hat{x}_i$,The final prediction $\hat{x}$ is the average represented at equation 42:

$$\hat{x} = \frac{1}{A}\sum_{i=1}^{A}\hat{x}_i \quad \ldots\ldots (42)$$

3.3.4 SVM

Support Vector Machines(SVM) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. SVM tries to find an optimal hyperplane, which best separates classes with the maximum margin possible. This hyperplane is determine by the weight vector, w and bias b, it does this solving an optimization problem that minimizes $\frac{1}{2}||\omega||^2$, subject to constraints underlying correct classification of data points For non-linearly separable data, SVM applies the kernel functions like polynomial, radial basis function and sigmoid to convert the points in higher dimension where they can be separated by linear line. For non-perfect separability cases, SVM adds slack variables to enable deviations from perfect class separation and balances the margin width against classification error through a regularization parameter C. The decision function for predicting a new point is determined by summing up that's weighted sign of support vector points. SVM is popular due to its efficiency in high-dimensional space and if used carefully can prevent overfitting.

Optimization Problem: For a linearly separable dataset, the optimization problem can be formulated as 43, 44:

$$min\frac{1}{2}||\omega||^2 \quad \ldots\ldots (43)$$

Subject to the constraints:

$$b_i(\omega \cdot a_i + b) \geq 1 \; \forall i \quad \ldots\ldots (44)$$

These constraints ensure that each data point is correctly classified and is on the correct side of the margin.

Lagrange Multipliers (Dual Form): The optimization problem also can be expressed using Lagrange multipliers. This is particularly useful when its dealing with the dual problem, which often makes the optimization more efficient and allows for the use of kernel tricks. The dual problem formulation represented at equation 45, 46:

$$max \sum_{i=1}^{N} \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j b_i b_j (a_i . a_j) \quad \ldots\ldots (45)$$

Subject to:

$$\sum_{i=1}^{N} \alpha_i b_i = 0 \; and \; \alpha_i \geq 0 \; \forall i \quad \ldots\ldots (46)$$

Here α_i are the Lagrange Multipliers.

Kernel Trick: When data is not linearly separable in the original feature space, SVM can apply a kernel function $K(\alpha_i, \alpha_j)$ to map the input space into a higher-dimensional feature space where a linear hyperplane can separate the data. Common kernels include:

Linear Kernel: $K(\alpha_i, \alpha_j) = \alpha_i . \alpha_j$, Polynomial Kernel: $K(\alpha_i, \alpha_j) = (\alpha_i . \alpha_j + 1)^d$, Radial Basis Function (RBF) Kernel / Gaussian Kernel: $K(\alpha_i, \alpha_j) = exp(-\gamma \left|\left|\alpha_i - \alpha_j\right|\right|^2)$, Sigmoid Kernel: $K(\alpha_i, \alpha_j) = tanh(k\alpha_i . \alpha_j + c)$.

Soft Margin SVM: In cases where the data is not perfectly separable, a soft margin SVM introduces slack variables ξ_i to allow some misclassifications. The optimization problem represented at equation 47, 48:

$$min\frac{1}{2}||\omega||^2 + C\sum_{i=1}^{N}\xi_i \ldots\ldots (47)$$

Subject to:

$$x_i(w.x_i + b) \geq 1 - \xi_i \text{ and } \xi_i \geq 0\ \forall i \ldots\ldots (48)$$

Here, C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification error.

Decision Function: The decision function for a new input x is represented at equation 49, 50:

$$f(a) = sign(w \cdot x + b) \ldots\ldots (49)$$

In the case of the duel formulation using kernel, its becomes,

$$f(a) = sign\left(\sum_{i=1}^{N} a_i y_i K(a_i, a) + b\right) \ldots\ldots (50)$$

Interpretation of the Equations:

- Optimization Objective $min\frac{1}{2}||\omega||^2$: This term is minimized to maximize the margin between different classes. A smaller // w // means a larger margin.
- Constraints ($x_i(w.x_i + b) \geq 1$): These constraints ensure that data points are correctly classified and lie on the correct side of the decision boundary.
- Slack Variables (ξi): These are introduced to handle misclassifications, making SVM robust to noise and outliers.
- Lagrange Multipliers (α_i): These coefficients in the dual problem help find the optimal hyperplane and also allow SVM to use the kernel trick effectively.

Ensemble Methods: Ensemble methods are techniques that construct a set of models and combine the predictions to increase accuracy. Bagging: Bagging (Bootstrap Aggregating) reduces variance by applying the bootstrapping method in training a different model on various subsets of data and combining them. Boosting: AdaBoost etcIt works sequentially to correct the

errors of each previous model, this lowers warp. Stacking learns how to combine predictions from multiple base models using a meta-model. Voting collects a sum of predictions collected from all models via majority(hard voting) or averages(soft voting). These usually lead to better models by merging the different powers of individual ones.

1. Bagging: This method is an ensemble generation approach by training multiple instances of a model on diverse random subsets of the original sample, and then aver aging the predictions or using voting (in regression) or just averaging in simple terms. For each of the models, it is trained on a random subset (or bootstrapped sample) from your training data. Each partition is used to train a separate model. Predictions are aggregated with mean or majority voting represented at equation 51.

Prediction for Regression:

$$\hat{x} = \frac{1}{M} \sum_{m=1}^{M} \hat{x}^{(m)} \quad \ldots\ldots (51)$$

where $\hat{x}$ is the prediction from the m model, and M is the total number of models represented at equation 52.

Prediction for Classification:

$$\hat{x} = mode(\hat{x}^{(1)}, \hat{x}^{(2)}, \ldots, \hat{x}^{(M)}) \ldots\ldots (52)$$

Where $mode$ denotes the most common prediction among the M models.

2. Boosting: Boosting is a classifier that generates models sequentially, and each model tries to correct the errors of its predecessors. The ultimate model is the linear combination of all models. The model is trained one after the other, with each new model training on errors residual to their previous models. The predicted value is the weighted average of all trees, where each tree votes a weight equivalent to its accuracy. Boosting achieves low bias and variance by concentrating on challenging cases represented at equation 53.

Weighted Prediction:

$$\hat{x} = \sum_{m=1}^{M} \alpha_m \hat{x}^{(m)} \quad \ldots\ldots (53)$$

Where α_m is the weight of the m-th Model.

Updating Weights: Misclassified sample weights are increase, while correctly classifying sample weights are decreased, represented at equation 54.

$$w_i^{(t+1)} = w_i^{(t)} exp(\alpha_t \cdot I(x_i \neq \hat{x}_i^{(t)})) \quad \ldots\ldots (54)$$

where $w_i^{(t)}$ is the weight of sample i at iteration t, α_t αt is the model's weight, and I is an indicator function.

3.Stacking: Stacking is an ensemble method where we leaves to few. base-models, then combine them with a meta-mode. To train the meta-model based on the features produced by predictions of base models trained with original datasets. The same dataset lead to be trained for multiple models. A secondary (often more complicated) model is trained to decide how the predictions of base models can be combined. Cross Validation (usually used to stop the meta-model from over fitting).

Base Model Predictions**:** Each base model mmm produces a prediction $\hat{x}^{(m)}$ represented at equation 55.

$$\hat{x}^{(m)} = f_m(Y) \quad \ldots\ldots (55)$$

Meta-Model Training: The meta-model g is trained on the base model predictions represented at equation 56:

$$\hat{x} = g(\hat{x}^{(1)}, \hat{x}^{(2)}, \ldots, \hat{x}^{(m)}) \ldots\ldots (56)$$

where g is the meta-model function that combines the predictions from M base models.

4. Voting: Voting is a type of ensemble method where you can take the average or some other combination to get the final prediction. Finally it takes the mode of all class predictions as final prediction. In the last model, prediction is done on average predicted probabilities represented at equation 57, 58.

Hard Voting:

$$\hat{x} = g(\hat{x}^{(1)}, \hat{x}^{(2)}, \ldots, \hat{x}^{(m)}) \quad \ldots\ldots (57)$$

Soft Voting:

$$\hat{x} = arg\, max\left(\frac{1}{M}\sum_{m=1}^{M} P(y = k|x, \theta_m)\right) \quad \ldots\ldots (58)$$

Where $P(y = k|x, \theta_m)$ is the probability that the m th model predicts class k

KNN: The K-Nearest Neighbors (kNN) algorithm is a simple and lazy machine-learning non-parametric method, used for classification as well regression problems. A k-nn algorithm works by classifying a query point given the 'k' closest data points to that reference point, based on some distance metric (usually euclidean). In the case of classification, it returns a class that is most common among k neighbors and for regression, it retrieves the average value of these neighbors. Meaning -with regards to the value of 'k', less creates potential for overfitting, more means a loss generalization. Although k-NN is very simple and easy to understand, it can be computationally expensive if there are a lot of samples in the training dataset; due to storing all sample data with features.

Distance Metric: The results of the k-NN algorithm are highly sensitive to this choice so making a good distance metric is crucial. The Euclidean distance is probably the most common such metric, and it works as follows represented at equation 59.

$$d(a, a_i) = \sqrt{\sum_{j=1}^{n} \left(a_j - a_{ij}\right)^2} \quad \ldots\ldots(59)$$

a: New data point (query point). a_i: A point in the training set. a_j: The j-th feature of the new data point. a_{ij}: The j-th feature of the i-th training point. n: Number of features.

Classification Decision Rule (Majority Voting): In classification, after we estimate the k-nearest neighbors, then for a new data point a its class is determined by majority vote. $N_k(a)$— the set of k-nearest neighbors from xi The predicted class $\widehat{b}$ is represented at equation 60

$$\widehat{b} = argmax_c \sum_{a_i \epsilon N_k(a)} \text{II}\,(y_i = c) \quad \ldots\ldots (60)$$

c: Class label. y_i: The actual class of the neighbor a_i . II $(y_i = c)$: Indicator function that is 1 if $(y_i = c)$ and 0 otherwise.

Regression Decision Rule (Averaging): In the case of a regression task, the prediction for a new point a is simply calculated as the average of k values y in its neighborhood. The predicted value $\widehat{b}$ is represented at equation 61:

$$\widehat{b} = \frac{1}{k} \sum_{a_i \epsilon N_k(a)} b_i \quad \ldots\ldots(61)$$

b_i: The target value of the neighbor a.

3.4. System Implementation

Active Function: The point of activation functions is to introduce non-linearity into the function, in turn enabling it and by extension neural networks to learn complex patterns. The Sigmoid function takes large amount of input values and squashes them into a range between 0–1, which is perfect for binary classification but it has the vanishing gradient problem. Tanh — This is also a Sigmoid function but zero-centered between -1 to +1 and good for hidden layers. The ReLU function outputs zero for negative inputs and the input itself when positive, resolving the vanishing gradient problem efficiently but leading to an occasional dying relu scenario. This can be counteracted by giving a small gradien to negative values which is what Leaky ReLU does. PReLU takes this a step further by learning the slope for negative values as well during training. Instead we use an ELU, which introduces improvements for negative input due to the exponential part of its form (this makes convergence faster). Sometimes using the Swish function, which is defined as $x \cdot \sigma(x)$, where $\sigma(x)$smooths out transitions between different values for a better performance. Lastly, Softmax function (used in multi-class classification) normalizes the outputs through probability distributions such that their sum is equal to 1. Every activation function has its own pros and cons, as well as can influence the overall performance & training of neural networks.

Sigmoid(Logistic)Activation Function: The sigmoid function squashes its input (x-axis value between minus infinity and plus infinity) to s- shaped curve y in 0–1(both x and y are from the set of Reals). $\alpha(y)$ was converted from a number to probability: we get $\alpha(y)$->0 in the neighborhood of negative infinity and $\alpha(y)$->1 while its x goes on being close positive infinit. It is mostly used for solving binary classification problems. But it doesn't work well for higher values of x and the vanishing gradient problem which makes this a bad choice for deeper network represented at equation 62 :

$$\alpha(y) = \frac{1}{1+e^{-y}} \quad \ldots\ldots(62)$$

Hyperbolic Tangent (Tanh) Activation Function: The hyperbolic tan function squashes its input to a value between -1 and 1. It is simply a scaled version of the sigmoid function, and so has essentially the same shape but different range (-1 to 1) instead (0– 1）. This property makes tanh more appropriate than sigmoid for hidden layers. Is its output to have zero-mean so that the

deficit notion in learning is reduced as it helps balancing things between positive and negative inputs of weights represented at equation 63 :

$$tanh(y) = \frac{e^{y}-e^{-y}}{e^{y}+e^{-y}} \ldots\ldots (63)$$

Rectified Linear Unit (ReLU) Activation Function: ReLU is naturally 0 for all negative values and equal to the input value whenever it is positive. These same properties help make it one of the most commonly used activation functions in deep learning. While ReLU solves the vanishing gradient problem, it also comes with a challenge called "dying ReLUs" — very large gradients flowing through this unit could cause weight update to take dead-end and get stuck at zero represented at equation 64:

$$f(y) = max(0, y) \ldots\ldots (64)$$

Leaky ReLU Activation Function: Leaky ReLU gives you a small step of non zero gradient for negative values of x-great problem is that it does not "die" like the other one. If x is positive then function behaves like ReLU and if x are negative the f(x) = 0.01x (where alpha=0.01) is represented at equation 65.

$$f(y) = \begin{cases} y & if\ y>0 \\ \alpha y & if\ y\leq 0 \end{cases} \ldots\ldots (65)$$

Where α is a small constant.

PReLU (Parametric ReLU): Parametric ReLU is similar to Leaky RELU, but with a slight Updated_negative_slop in the negitive side instead of fixed slops it allows become learning during training. Of course, sometimes this flexibility can also lead to performance improvements represented at equation66 :

$$f(y) = \begin{cases} y & if\ y>0 \\ \alpha y & if\ y\leq 0 \end{cases} \ldots\ldots (66)$$

Where α is a learnable parameter.

Exponential Linear Unit (ELU) — This variant add negligible exponential value for each negative input manyves the function, allowing a model to output arbitrary small or large negaNegative shift bias. It contributes in faster convergence than ReLU represented at equation 67:

$$f(y) = \begin{cases} y & if\ y>0 \\ \alpha(e^{y}-1) & if\ y\leq 0 \end{cases} \quad \ldots\ldots (67)$$

Where α is a hyperparameter.

Swish Activation Function / Swish : This is nothing but the self -gated activation function & you can think this as similar to Sigmoid-Weighted Linear Unit, which works with ReLU. Given the smooth transitions from linear to non-linear regions, in various cases it performs better and converges faster than ReLU represented at equation 68 :

$$f(y) = y \cdot \sigma(y) = \frac{y}{1+e^{-y}} \quad \ldots\ldots (68)$$

Softmax Activation Function: Softmax function is used in the output layer of neural network for multi-class classification problems. It takes a vector of raw scores and converts them into probabilities by performing an exponential on each score followed by normalization w.r.t the sum of exponentials. Id is between 0 and 1, sum adds to the only one represented at equation 69.

$$\alpha(x)_i = \frac{e^{xi}}{\sum_j e^{xj}} \quad \ldots\ldots (69)$$

For i=1,…,n where x= (x_1, x_2, . . . , x_n)

ANN: Artificial Neural Network (ANN) An ANN is a computational model of the human brain designed mainly to solve multiple classification and regression problems. A neural network is made of several layers built from interconnected nodes an input layer, one or more hidden layers and an output layer. There are connections between nodes The weights associated with each connection To determine a min weight You have biases for every node A node has an input, which is the weighted sum of its inputs plus a bias; this goes through an activation function (which can be sigmoid or ReLU) to get the output. Finally, the output of the network is compared with a target output and using some loss function such as Mean Squared Error for

regression or Cross-Entropy in classification. Backpropagation, in short adjusts the weights and biases by computing the gradient of a loss function with respect to each parameter so that it can optimize for minimum loss iteratively. ANNs are the major driver of image and speech recognition, natural language processing (NLP), and anomaly detection.

Weighted Sum and Activation: The output of a neuron in the hidden or an output layer is calculated by combining all its inputs together, multiplied (or not) and applying to it some activation function. Output of a neuron j can be given represented at equation 70:

$$y_j = \sum_i w_{ij} x_i + b_j \quad \ldots\ldots (70)$$

y_j= Weighted sum of inputs for neuron j. w_{ij}= Weight connecting input i to neuron j. x_i= Input to the neuron from the previous layer. b_j= Bias term for neuron j.

Activation Function: After being summed together, encouraged by an activation function represented at equation 71.

$$x_j = f(z_j) \quad \ldots\ldots (71)$$

x_j = Activation (output) of neuron j. $f(z_j)$ = Activation function.

Output of the Network: The final output given as an output to a classification (y) or regression value (for regression problems) represented at equation 72 :

$$z = f\left(\sum_j w_{ij} x_i + b_j\right) \quad \ldots\ldots (72)$$

Here, z is the final output after applying the activation function to the sum of weighted activations from the last hidden layer.

Loss Function: Loss function is used to compute loss which actually can be measured as the difference between predicted and true output. Loss function: Various type of loss functions are there like represented at equation 73:

Mean Squared Error (MSE) for regression problems:

$$M = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2 \quad \ldots\ldots (73)$$

Cross-Entropy Loss for classification problems represented at equation 74:

$$M = -\frac{1}{N}\sum_{i=1}^{N}[x_i log(\widehat{x}_i) + (1 - x_i) log(1 - \widehat{x}_i)] \quad \ldots\ldots (74)$$

N= Number of samples. x_i = Actual target value. $\widehat{x}_i$ = Predicted output.

Backpropagation Algorithm: BackPropagation is a mechanism to train ANN with the aid of loss function. It consists of taking the gradient (a vector of derivatives representing how much change is needed in each weight and bias), multiplying it by some rate, usually a bit tiny value to not disrupt things too fast, this step gives you the opposite directions that reduce error for every weights/bias used as learning parameters represented at equation 75, 76:

$$w_{ij} = w_{ij} - \eta\frac{\partial L}{\partial w_{ij}} \quad \ldots\ldots (75)$$

$$a_j = a_j - \eta\frac{\partial L}{\partial a_j} \quad \ldots\ldots (76)$$

η = Learning rate. $\frac{\partial L}{\partial w_{ij}}$ = Gradient of the loss function with respect to the weight. $\frac{\partial L}{\partial a_j}$ = Gradient of the loss function with respect to the bias.

LSTM: LSTM, (Long Short-Term Memory), is a type of recurrent neural networks which are more complex and capable specialization RNNs than can be tuned to robust scene description. The Long Short-Term Memory cell consists of the three main gates such as forget gate, input gate and output gate. These gates determine the flow of data through it cell state and hidden state. The forget gate figures out what parts of the cell state to throw away and the input gate decides which part should be updated. The candidate cell state represent the values that need to be updated in order for us to decide which parts of the old cell sate we are going to keep or dispose off. The updated state is a mixture of the old cell states and new proposal values (how much we note keep it), modulated by forget gate, addition with input gate.Output Gates: This gate choose what to output from the cell state, by reading new value of this previous time step and producing an updated version. This construct is what allows LSTMs to maintain sharp temporal glue and makes it a much more effective way of learning complex information over long sequences, making them ideal for use on time series prediction or NLP tasks.

Forget Gate: For each element in the cell state C_{t-1}, a value of between 0 to 1 is returned. 1 would be absolutely keep, and 0 would be get rid of this altogether. Decides what to forget from the cell state represented at equation 77 :

$$f_t = \sigma(W_f \cdot [h_{t-1}, a_t] = b_f) \ldots\ldots (77)$$

W_f = Weight matrix for the forget gate. b_f = Bias for the forget gate. σ = Sigmoid activation function.

Input Gate: Input Gate Equation This output is element-wise multiplied by the candidate cell state $\overline{C}_t$ to determine which values should be added. Decides what new information to add to the cell state represented at equation 78 :

$$i_t = \sigma(W_i \cdot [h_{t-1}, a_t] = b_i) \ldots\ldots (78)$$

W_i = Weight matrix for the input gate. b_i = Bias for the input gate.

Candidate Cell State: This is also a cell state which may or may not be added to main cell state. It's new candidate vector which be added on cell state additionally depended decision of input gate. Creates a vector of new candidate values that could be added to the cell state represented at equation 79:

$$\overline{C}_t = tanh(W_c \cdot [h_{t-1}, a_t] = b_c) \ldots\ldots (79)$$

W_c = Weight matrix for the cell state. b_c = Bias for the candidate cell state.

$tanh$ = Hyperbolic tangent activation function.

Update Cell State: This equation updates the cell state C_t. The values in the cell state are dropped if required by multiplying with forget gate output. This process uses an input gate decision to determine what candidate values are added. Combines the previous cell state with the new candidate values represented at equation 80:

$$C_t = f_t * C_{t-1} + j_t * \overline{C}_t \ldots\ldots (80)$$

$*$ = Element-wise multiplication. C_t = Updated cell state.

Output Gate: It determines how much of the cell state to expose as an output. It decides on how the cell will give output by using sigmoid activation function along with hyperbolic tangent. Decides what to output based on the current cell state represented at equation 81:

$$o_t = \sigma(W_0 \cdot [h_{t-1}, a_t] = b_0 \ldots\ldots (81)$$

W_0 = Weight matrix for the output gate. b_0 = Bias for the output gate.

Final Hidden State: The filtered form of the cell state through output gate which is either passed to next time step or used as a final/actual output. The hidden state is the output from LSTM cell and can also be used as an input to subsequent LSTM cells. Produces the final output of the LSTM cell represented at equation 82:

$$h_t = o_t * tanh(C_t) \ldots\ldots (82)$$

CNN: A Convolutional Neural Network is a type of deep learning model that can process grid-like data, such as images. This includes convolutional layers which apply filters and are used for detecting local patters, activation functions like ReLU function to introduce non-linearity into the model, pooling layers Pooling layer is a sub-sampling filter operated on feature maps after applying a certain number of convolutions maximize over different window lengths or making map more sparse will bring down dimensionality complexity per unit depth till now its an unsupervised approach unlike it next part where loss/cost needs supervision signal from label information then finally fully connected layer in end which gives labels/prediction at output side. A convolution operation slides a filter over the input data to produce feature maps which have encoded important patterns. A CNN is trained by iterative forward propagation to compute the output and loss, then backpropagation using optimization techniques in order to update its weights. This is enabled the CNN to learn and enhance ability for images pattern recognition (classification) or image objects detection.

Input Layer: To take the input data, most likely an image. If it is RGB image, the input layer would have 3 channels.

Convolutional Layer (ConvLayer): The bread and butter of a CNN, responsible for detecting local features such as edges, corners, textures, or in other words any pattern resembling shapes throughout the image. Convolution is the process of sliding a filter (or kernel) over the input data to extract features from it. HarshlyMathematically, for an input image I and a filter K represented at equation 83:

$$(I * K)(a, b) = \sum_{y} \sum_{z} I(y, z) \cdot K(a - y,\ b - z) \ldots\ldots (83)$$

I is the input matrix, K is the kernel matrix., $(*)$ denotes the convolution operation, (a, b) are positions in the feature map, (y, z) re positions in the kernel.

Stride: size of the step as filter moves on input This has the effect of shrinking the output tensor in its spatial dimensions.

Padding: By default, we add zero-padding to the input matrix on all sides in order to get some control of spatial dimensions of output. The equation is given below:

Activation Function (ReLU)Applies a non-linear transformation on the output of 2nd layer which is mostly the Rectified Linear Unit, also known as ReLU, define represented at equation 84:

$$f(a) = max(0, a) \ldots\ldots (84)$$

ReLU introduces non-linearity into the model, allowing it to learn complex patterns.

Pooling Layer (Down-sampling): Decreasing the size of feature maps, Help to Reduce computational complexity and control overfitting. Max Pooling: Most common pooling operation in a CNN represented at equation 85:

$$f(b) = max(0, a) \ldots\ldots (85)$$

Where b is the maximum value within a defined window over the input feature map.

Fully Connected Layer (FC layer): In Fully connected layers, the nodes are either fully connected to all activations in the previous layer as can be seen from below image. The last layer combines the features that are extracted from convolutional and pooling layers and uses them to predict the final output.

Output Layer: This is the layer where we receive output of network. Usually it uses a softmax activation function for prediction probability in classification tasks that provides probabilities of each class represented at equation 86 :

$$softmax(x_i) = \frac{e^{xi}}{\Sigma_j e^{xj}} \quad \ldots\ldots (86)$$

Here, x_i is the i-th element of the input to the softmax function. e^{xi} is the exponential of x_i.

Chapter 4

SYSTEM TESTING AND ANALYSIS

In this Chapter , "System Testing and Analysis" the main concept is to discuss with regard to analysis of performance evaluation measures for more efficient mental health prediction model developed based on image processing. This chapter is important for evaluating the performance of your engine in real-world use cases and getting to know what it can do well and where its boundaries are. It takes a look at the different ways testing is executed, what measurements are used in this evaluation and what an overall result analysis shows.

4.1 Feature and Attributes

This will be our Mental Health Prediction System and it is only as reliable, as the dataset used to train it. We used a pre-labeled case dataset with images of different mental health cases. The data set was extensively polished and thoroughly well-diversified in order to avoid any kind of bias problems with respect to the demographics. After that, a pre-processing step was applied to focus the images more clearly and on important features relevant for mental health analysis. The dataset was divided into training, validation and test sets to make sure that the model is trained on different data compared to evaluation for better generalization. The key aspects of the dataset has been organized in the Table 2.

.Table 2: Details Information of Datasets

Feature	No of sample	Duration(sec)	No of sample	Range of Age
Talking	150	10	150	18-50
Walking	150	10	150	18-50
Playing	150	10	150	18-50
Eating	150	10	150	18-50

4.1.1 Size

The dataset used in this paper is 6000 images in total and arranged among four specified activities such as eating, playing, reading book.,walking and talking At exactly 150 images per category, both balanced and extensive. We need to ensure that the images in each 4 categories are distributed similarly, so that we can build a model on top of this which is unbiased towards any particular set and know how it distinguishes between these activities.

Eating (150 images): This set contains pictures of individuals eating, with different expressions on face and body language, also if necessary contextual cues relevant to mental health consequences. These images provide a plethora of data on the emotional and cognitive states driving food-related behavior.

Playing (150 images): The play set is a series of images depicting people in activities they do for fun. We need these pictures to tackle questions about how leisure activity engagement might be causing, or reflecting changes in mental health.

Walking (150 images): – A collection of people walking: from outdoor parks to urban streets.The pictures aimed at showing three major sides of walking; The form, the manner and how you hold yourself in general as well. From these images, the model can learn to recognize physical signals that may indicate a range of mental states — for example anxiety or depression vs general wellness—revealed in the way individuals walk.

Talking (150 images): Talking (150 images): This category covers people talking to each other in one-on-one and group discussions. However, these images concentrate on the facial expressions, eye contact and gestures — or simply social cues that are required to be observed thematically in relation with mental health. For example an inability to make eye contact; forced laughter; a uptight posture in conversation could be potential symptom of social anxiety or another mental health condition. This segment aids the model in recognizing this slight social acceleration with certain mental health diseases.

There are 600 images in total available within the dataset which is a big enough sample of different environments and minor variations for training the AI model to understand various circumstances. Given this diversity in both sets (participants demographically and settings) enhances furthermore the models ability to generalize its prediction on new unseen data.

4.1.2 Data collection sources:

This project has a carefully curated dataset collected from the mix of both public repositories and particular student groups to have diverse, representative samples for robust model training as well evaluation.

Public Sources (80%): We utilized approximately 480 images from public repositories that hold established base lines. Seeing around the corner Public repositories with wide ranges of images are known to have everything between men and women on it. Real world evidenceThe public data gives in the broad base needs to be fine for picking up many contexts where mental health indicators may show. The provided images are usually well-annotated and have been the cornerstone of many studies, which guarantees that they follow a certain degree of quality required to train an accurate predictive model.

Specific Student Groups (20%): Also in line with the public data, 120 images were collected from specific student groups. These pictures were taken in a controlled environment where context and conditions was as appropriate to pick up optimal mental health indicators. In this dataset, a lot of value lies in the fact that it is focused on youth and academic work populations can have different barriers or stressors compared to general population. This targeted data helps provide a richer dataset, enriching the layers that go into our predictions and ensures the model can accurately surface mental health conditions in different demographic groups. Together these sources public repositories and individual student groups, provide a wide coverage of the population for which this model is designed. The acquisition of samples from different sources gives us a diverse sourcing strategy, for enhancing the model's generalizability and making it work well on many types of images. Moreover, it guarantees that the dataset covers a wide range of proxies (i.e., mental health indicators ranging from common appearances in society at large to more precise ones appearing amongst groups prone to student life); which ultimately leads towards making predictions less biased and generally better.

4.1.3 Target Labels

The images in the dataset were well curated and hence are labeled as to which type of mental health conditions they belong, this is necessary for training and testing. This is important as the labeling must be correct, this becomes your ground truth which you train your model on. All labels were assigned by expert analysis and checked for reliability with mental health assessments.. Assignment of label class takes an image from one class of ice and sorts them to each or one condition . Here's a more detailed description of the labeling process:

1.Anxiety:

Behavioral Indicators: Pictures in this category show some of the things people with anxiety may exhibit or look like. That could be things like rigid body posture, fiddling with their hands or lips, avoiding eye contact and facial expressions that tell you they are nervous or anxious. It does so with all of the typical flair that we expect from big enterprise: a tiny amount is printed in bold on an enormous label, and it screams at consumers to be terrified or very concerned.

Some images presumably would be environments where the subject is obviously uncomfortable, such as a crowded scene or social setting which are known to provoke an anxious reaction.

2. Depression:

Affective Attribution: Images in this area express symptoms of depression. These are often the images of people wearing sad or blank faces, slouched bodies and moving slowly without vigour for what is positioned around them. This is, of course focusing on the sort of flat affect (emotional numbing or blunting) that we so often see with depression.Withdrawal from Social Interaction: There may be images of someone who looks like they've cut off all ties, look isolated or disengaged with others around them. It is intended to evoke the reality of depression by showing elements that people might see.

3. Non-Mental Health Issues (Control):

Control: The images labeled control depict subjects who are assumed not to manifest behaviors of anxiety and depression. These images are important for setting a starting point, so the model can differentiate from what healthy behaviors should look like and generally indicates symptoms of mental health issues. Neutral or positive expressions such as commonly performed social interactions, to provide a balanced reference that will be relatively easy for the model to compare against all other categories.

4.2 Performance Metrics

4.2.1 Accuracy:

The accuracy is the number of correct predictions made by a particular model, it also reveals how many false positives/negatives are observed along with true negatives/positives. This is one of the simplest and most commonly used metrics to assess a model performance for classification task. It simply calculates the percentage of difference between its expected true values to all represented at equation 87 ,

$$\text{Accuracy}= \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad \ldots\ldots(87)$$

Where:

TP (True Positives): Model correctly predicts mental health problems like Anxiety and Depression

TN (True Negatives): Number of instances where the model correctly concluded that there was no underlying mental health condition present (i.e., No Mental Health Issues).

FP (False Positives): These are the times when model predicts that some mental health condition is present but it actually not.

False Negatives (FN) – Cases where a model mistakenly predicts no mental health condition when in reality there is one..

4.2.2 Precision:

Precision is a metric that defines the correlation with respect to true positive predictions and total positives predicted by a model This tells you the number of predicted positive cases that are actually correct. Precision: This is very important when the cost of false positives in high, for example diagnosing mental health condition where if we predict incorrectly it would cause worry or treatment unnecessary.High precision can be useful in our mental health prediction project since it indicates how many of the positive predictions we made were correct. Here's why precision matters:

False Positives: High precision means (when the model predicts, for instance Anxiety or Depression) that it is likely correct. In so doing, it minimises the likelihood of attaching a diagnosis to an individual who does not have one.

Reliability: In settings where the decisions guided by model predictions have important implications for human safety or result in intervention, high precision is key to managing those interventions properly.

The contradiction act: This assures to keep the total ensuring that precision saves a balance in Assurity, and you already know where it is going!; especially when we have more data points as 'No Mental Health Issues' than symptoms like Anxiety Or Depression represented at equation 88.

Here is the formula of precision,

$$\text{Precision} = \frac{TP}{TP+FP} \quad \ldots\ldots (88)$$

4.2.3 Recall

Recall is also known as Sensitivity or True Positive rate, it measures the percentage of actual positive instances that are correctly recognized by our model. It is important for our project for making sure that the model detects all of the true cases as many as possible to intervene incorrectly, which would be wrong and dangerous when identifying mental health conditions timely.It reflects how closely the model captures all positive instances of cases with a focused condition. For our mental health prediction project, our need recall because it tells how good model is at catching true cases (ie catches all the ones that should be caught). Here's why recall matters:

Balance Between Identifying All Cases (High Recall) High recall ensures a model would predict Everyone identified to have Anxiety or Depression and other cases. In order not to lose those who really need help, this is crucial.

Minimizing Missed Diagnoses: In situations where missing a diagnosis can be harmful (e.g. leaving an anxiety or depression case undiagnosed) having high recall helps to prevent as many cases from being overlooked as possible.

Precision is balanced with: Precision should focus on making the correct positive predictions, but recall focuses into getting as many true positives among positive; When used side-by-side

these two measures can provide a well-rounded view on how the model is being evaluated represented at equation 89.

$$\text{Recall} = \frac{TP}{TP+FN} \quad \ldots\ldots(89)$$

4.2.4 F1 Source:

F1 Score: Harmonic mean of precision and recall. For this reason, it is used as a single value that captures both of these numbers, providing an all-encompassing view of how your model actually performs especially when working with imbalanced data sets.The F1 Score is very important in your mental health prediction project because it takes into account Precision and Recall, which helps to make sure the model identifies correctly both (true cases of mentally ill) minizing as well for False Positives. The F1 Score matters because it does not allow one of the two classes in a skew imbalanced examples to dominate.

Trade-off between Precision and Recall: In cases where optimizing for recall also optimizes precision and vice-versa, the F1 Score provides a middle ground. E.g., in mental health diagnostics, you want a model able to catch most cases (high recall) and that those which are caught be very likely true positives as well (i.e. high precision). The F1 Score is a combination of the two, and it turns out it gives us a very balanced metric that reflects both at once.

Dealing with Imbalanced Data: For example, in your project you can have an abundance of images that indicate ‘No Mental Health Issues' than those showing 'Anxiety‘ or “Depression“. In such a case, accuracy may be deceptive since the model could hardly classify all samples as a majority class in order to achieve high performance. On the other hand, F1 Score works better in these cases instead of accuracy since it takes into account both classes evenly.

Overall Rating: When someone want to use a single metric to describe how well a model is performing, particularly when the cost of false positives and false negatives are similar. However, a more relevant question is can we trust these DNNs to be robust enough for real-world application especially when talking about mental health represented at equation 90.

$$\text{F1 Scoure} = 2\times \frac{Prcision \times Recall}{Precision + Recall} \quad \ldots\ldots(90)$$

Where,

- Precision tells us what percentage of the total Identification model positive Predictions are actually true.
- Recall is the ratio of true positive cases that are correctly identified by a model

4.2.5 ROC-AUC:

The area under the receiver operating characteristics curve (AUC) is a way to graphically compare models, especially in binary classification problems. It is often used when you are working on models that output a probability or score for each instance so the prediction can be rank-ordered (or thresholded). ROC Curve = True Positive Rate (TPR) vs False Positive Rate (FPR), used to evaluate performance AUC.worker restarts from ansible sudo pts(4). Why we cannot stop Ansible 2. TPR (True Positive Rate) is also called as a recall, which measures the number of positive events that are classified correctly. On the other hand, FPR stands for "false positive rate" is the fraction of negative occurrence that should have been classified to be a positive instance. The AUC doesn't have a simple formula like the other metrics. It's typically calculated using the trapezoidal rule applied to the ROC curve.

4.2.6 Specificity

Specificity (True Negative Rate): The percentage of all real negatives that the model accurately identifies. Is the complement of False Positive Rate, it helps to understand how well your model is able find negative cases and avoid false alarms. Specificity is very important in our mental health prediction use case as it represents how good the model predicts that a person do not have any mental illness (such as Anxiety, Depression) or other problems. The high amount of specificity means that the model is not recognizing conditions in people who are perfectly normal, which could also result from over-fitting.

Decrease False Positives: Specificity helps decrease false positives (situations where a model wrongly predicts the presence of any mental disorder), which is useful in health contexts as a result of if we make an error and say that somebody has got a selected psychological state downside, then they will become very stressed

Balance for Sensitivity: Correctly identifying all cases is sensitivity, and you don't want to label a lot of healthy people as positives that's true negative (specificity). By all these metrics, we can have a very good idea about the performance of our model.In the application area of mental and behavioral health, where a diagnostic is not being performed but an individual can

be referred for intervention based on their level or risk, high specificity helps to verify that only those who are most likely in need of further action have been caught by the model represented at equation 91.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad \ldots\ldots (91)$$

4.2.7 Sensitivity

Definition: Sensitivity, or True Positive Rate (i.e., Recall), is the proportion of actual positives (True Positives — i.e. This individual has a mental health condition like Anxiousness or Depression) which are correctly predicted by the model. The ability of the model to detect a condition where it is present.Sensitivity is vital in your mental health prediction project, if the sensitivity of a model is high it means how well our machine learning identify true positive results those are actually having any kind of medical condition. It is crucial so the model does not miss out those who need help. Capturing All Positive Cases — A high sensitivity means that the model identifies most persons with mental health conditions. In a healthcare example, missing a case (false negative) results in an individual needing treatment or intervention to be missed. Sensitivity, which helps cut on the false negatives— these are when your model misses out that a condition exists. In the realm of mental health not identifying a condition such as Anxiety or Depression may result in missed needful medical support. High-Stakes Situations: In situations where severity ranges from minimal to deadly and you must find every instance (e. g., screening for mental health diagnoses), recall is everything—maximum sensitivity means missing as few cases as possible. This is especially important when the cost of missing a diagnosis carries dire consequences. The formula for calculating Sensitivity Score is as follows represented at equation 92,

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad \ldots\ldots (92)$$

4.3 Feature Transformation Result

Table 3: Feature Transformation Result

No of Class	Talking	Walking	Playing	Eating
S1	0.623	0.951	0.773	0.712
S2	0.945	0.774	0.673	0.631
S3	0.915	0.855	0.859	0.788
S4	0.791	0.956	0.8	0.666
S5	0.792	0.862	0.689	0.705
S6	0.863	0.94	0.646	0.706
S7	0.981	0.682	0.872	0.685
S8	0.799	0.655	0.782	0.905
S9	0.671	0.688	0.878	0.693
S10	0.909	0.759	0.842	0.947
S11	0.609	0.9	0.761	0.632

Feature transformation results Each table shows the transformed feature values from eleven classes (S1-S11) on four activities according to various experiments conducted specifically for achieving an optimal matching result. In this table, the value in each cell is a given feature transformed by class and activity. Value falls between 0 and 1 with higher values denote stronger performance or association to the activity.The maximum for Talking is 0.981 in class S7, indicating that the highest value from this cast (S7) is most correlated with or performs better in talking activity.The weakest performance is for class S11 and belongs to Talking with a value= 0.609.The top result is 0.956 for Walking, class S4 which makes it the best score of this common activity segment. Interestingly, class S8 presents the lowest value (0.655), overall it might mean that S8 is weaker or singled out in Walking.Playing has an highest value of 0.878 in class S9 which indicates that this is the strongest category with respect to playing. The weakest association with Playing can be observed for class S6 (FF = 0.646). The best score for Eating

contains a value of 0.947 with respect to class S10, which does well in this activity.

4.4 Performance Results

4.4.1.Logistic Regression

Table 4: Performance Result on Talking Feature Of Logistic Regression

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.831	0.949	0.926	0.614	0.793	0.847	0.995
Anxiety	0.676	0.849	0.623	0.783	0.678	0.804	0.894
Bipolar Disorder	0.792	0.896	0.668	0.757	0.915	0.883	0.727
Schizophrenia	0.772	0.891	0.806	0.604	0.791	0.727	0.917

In this table the model had the best result in depression prediction with a sensitivity of 0.995, which means that hundreds of real positive cases were correctly classified out thousands to be selected as such.

Table 5: Performance Result on Walking Feature Of Logistic Regression

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.856	0.889	0.864	0.894	0.778	0.952	0.866
Anxiety	0.838	0.95	0.982	0.908	0.681	0.652	0.979
Bipolar Disorder	0.759	0.945	0.925	0.885	0.818	0.789	0.751
Schizophrenia	0.62	0.6	0.663	0.855	0.757	0.92	0.9

Predicting anxiety was where the model performed best, with a precision of 0.95 and a recall of 0.982. This indicates that it had a very low percentage of false positives and was highly accurate in recognizing actual cases of anxiety.

Table 6: Performance Result on Eating Feature Of Logistic Regression

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.979	0.805	0.802	0.689	0.959	0.663	0.752
Anxiety	0.644	0.971	0.936	0.969	0.802	0.718	0.868
Bipolar Disorder	0.719	0.811	0.701	0.979	0.792	0.929	0.862
Schizophrenia	0.934	0.983	0.965	0.649	0.661	0.659	0.678

In this table the model's accuracy (0.934) and precision (0.983) fared the best. The highest sensitivity (0.868) and recall (0.936) were related to anxiety. Bipolar disorder had the best specificity (0.929), while depression had the highest ROC-AUC (0.959).

Table 7: Performance Result on Playing Feature Of Logistic Regression

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.642	0.799	0.797	0.9	0.62	0.782	0.797
Anxiety	0.625	0.65	0.843	0.799	0.78	0.71	0.852
Bipolar Disorder	0.797	0.647	0.869	0.615	0.722	0.759	0.712
Schizophrenia	0.684	0.932	0.876	0.98	0.605	0.737	0.736

As to the schizophrenia, it shows greatest PRE(0.932) and F1 Score (0.980), which means that this model worked best for them. Anxiety had the highest sensitivity (0.852) and recall (0.843). The ROC-AUC for Depression was the highest (0.62) and Bipolar Disorder had maximum accuracy 0.797. The graphical representation has been depicted in the Figure [9] as follows.

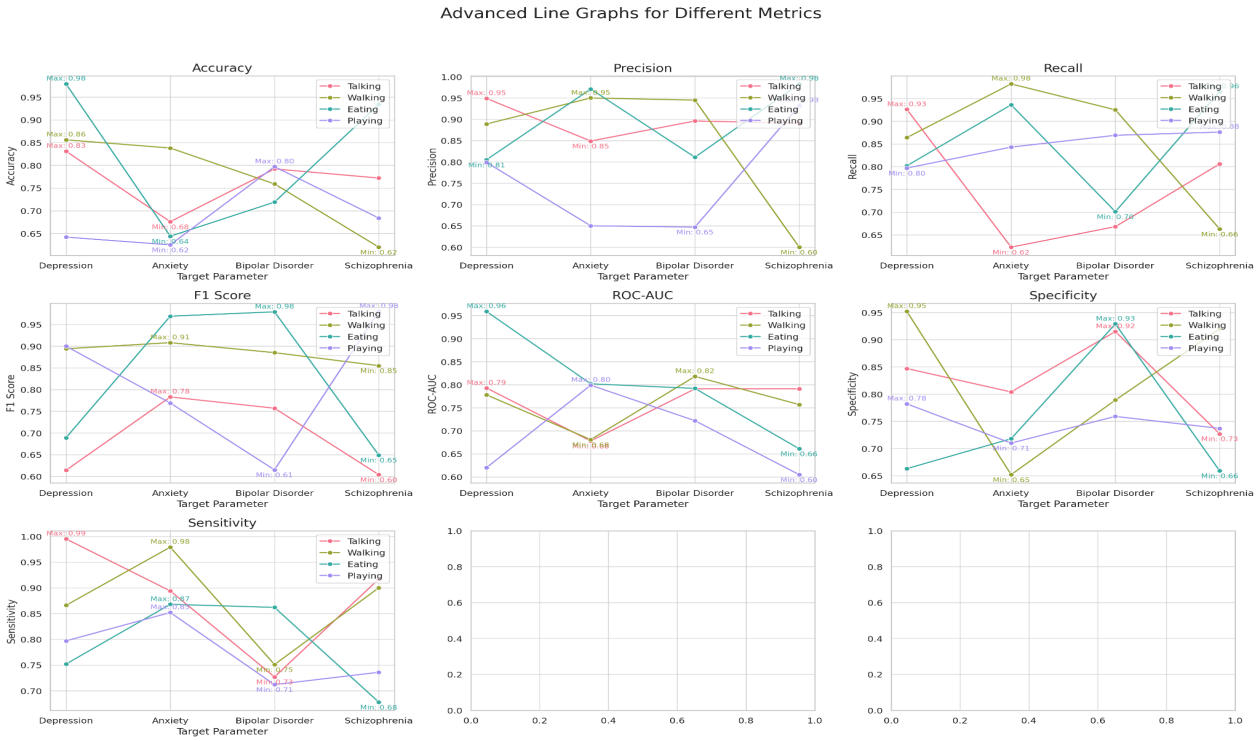

Figure 9: Performance result of Logistic Regression

4.4.2. Decision Tree

Table 8: Performance Result on Talking Feature Of Decision Tree

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.702	0.871	0.783	0.923	0.869	0.701	0.823
Anxiety	0.939	0.831	0.93	0.936	0.804	0.64	0.951
Bipolar Disorder	0.872	0.78	0.814	0.841	0.716	0.704	0.756
Schizophrenia	0.648	0.729	0.804	0.77	0.724	0.874	0.955

For the anxiety model, it yielded a stellar 0.939 test accuracy and 0.93 recall score (best-in category). Most precise was depression (0,871) Similarly, Anxiety stood out as the dimension with highest F1 Score (0.936). They showed that the best ROC-AUC value was achieved by depression (0.869) The sensitivity and specificity were the highest for schizophrenia (0.955, 0.874).

Table 9: Performance Result on Walking Feature Of Decision Tree

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.849	0.617	0.958	0.621	0.878	0.916	0.881
Anxiety	0.886	0.833	0.918	0.988	0.825	0.623	0.938
Bipolar Disorder	0.623	0.712	0.779	0.923	0.725	0.658	0.979
Schizophrenia	0.737	0.77	0.634	0.695	0.798	0.657	0.93

For anxiety, the model is also a leader with an accuracy of 0.886 and F1 Score of 0.988 The disorder with the highest precision was schizophrenia (0.770). The highest recall (0.958) was obtained for depression [Table 1]. Moreover, depression had the highest ROC-AUC (0.878%). Specificity was highest using Anxiety (0.623), and sensitivity was best with Bipolar Disorder (0.979).

Table 10: Performance Result on Eating Feature Of Decision Tree

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.615	0.834	0.721	0.783	0.712	0.93	0.675
Anxiety	0.752	0.605	0.613	0.748	0.608	0.609	0.705
Bipolar Disorder	0.655	0.613	0.652	0.747	0.791	0.694	0.65
Schizophrenia	0.981	0.804	0.84	0.916	0.916	0.601	0.785

The model gave the best_accuracy(0.981) and F1 Score)(0.916) with Schizo_Paranoid Similarly, schizophrenia exhibited the highest ROC-AUC value (0.916). Depression scored the greatest for accuracy [0.834]. Schizophrenia (recall 0.84), and Depression (specificity, 0.93).

Table 11: Performance Result on Playing Feature Of Decision Tree

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.883	0.813	0.975	0.832	0.739	0.604	0.694
Anxiety	0.773	0.846	0.898	0.716	0.901	0.931	0.69
Bipolar Disorder	0.632	0.709	0.61	0.893	0.607	0.898	0.892
Schizophrenia	0.778	0.868	0.711	0.825	0.686	0.708	0.939

The model achieved best results with both accuracy (0.883) and recall (0.975) on Depression. The best performing label was: great grandfather/grandson, precision reached 0.876 and F1 Score up to 0.833) while schizophrenia had the highest precision (0.868) & F1 Score as well (0.825 explains this). Anxiety had the best discriminative ability with a ROC-AUC of 0.901, and Schizophrenia showed more sensitivity (0.939).The graphical representation has been depicted in the Figure [10] as follows

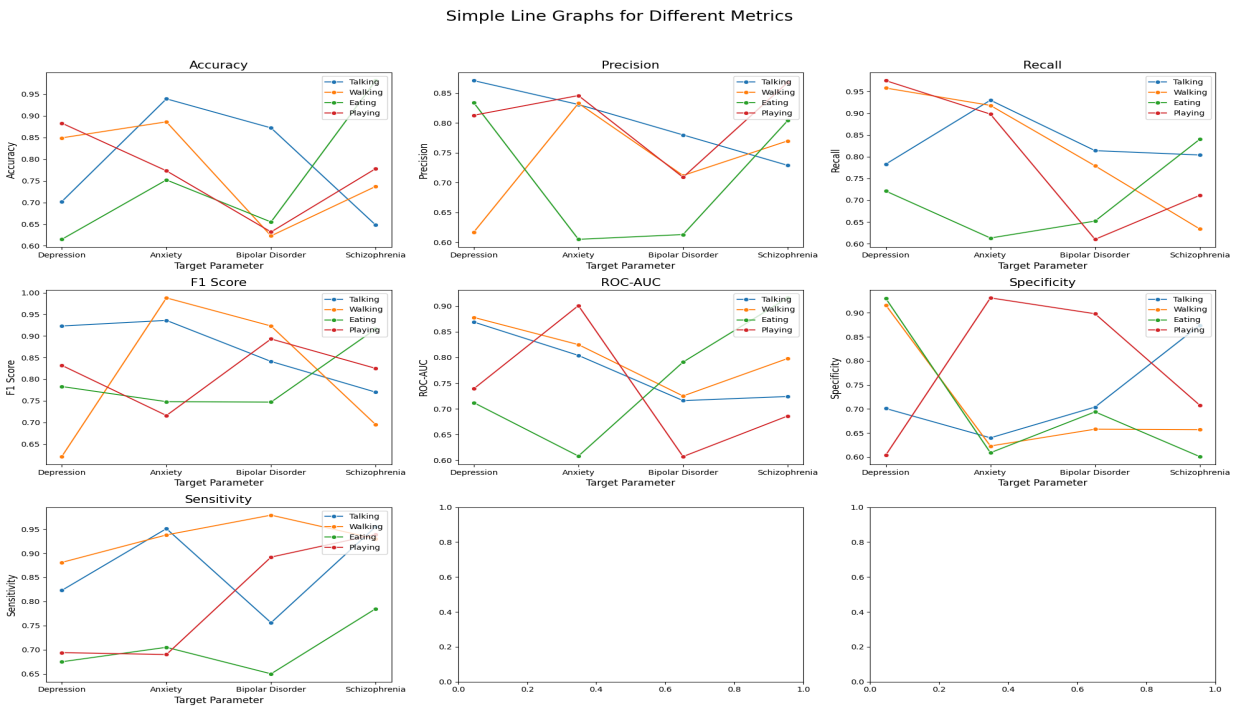

Figure 10: Performance result of Decision Tree

4.4.3.Random Forest

Table 12: Performance Result on Talking Feature Of Random Forest

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.955	0.646	0.901	0.969	0.872	0.679	0.994
Anxiety	0.722	0.806	0.738	0.765	0.963	0.824	0.788
Bipolar Disorder	0.658	0.939	0.936	0.811	0.69	0.878	0.642
Schizophrenia	0.792	0.747	0.804	0.897	0.913	0.83	0.72

With the highest accuracy (0.955) and sensitivity redshift parameters, ROC-AUC for anxiety was the highest (0.963). The highest F1 Score (0.897) and specificity was achieved by Schizophrenia, while Bipolar Disorder had the best precision value(0.939).

Table 13:Performance Result on WalkingFeature Of Random Forest

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.839	0.642	0.771	0.824	0.987	0.84	0.732
Anxiety	0.86	0.881	0.714	0.965	0.95	0.663	0.927
Bipolar Disorder	0.785	0.903	0.807	0.867	0.703	0.649	0.957
Schizophrenia	0.776	0.63	0.868	0.619	0.926	0.621	0.855

It performed the best in Anxiety with highest of all accuracy (0.86) and F1 Score(0.965). Bipolar Disorder was the most precise (0.903) and sensitive(0.957). Results: Schizophrenia achieved the highest recall (0.868), and Depression reached the highest ROC-AUC efficienc (0.987).

Table 14:Performance Result on Eating Feature Of Random Forest

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.864	0.964	0.848	0.968	0.952	0.941	0.826
Anxiety	0.914	0.797	0.703	0.916	0.739	0.961	0.801
Bipolar Disorder	0.999	0.815	0.829	0.663	0.74	0.748	0.917
Schizophrenia	0.631	0.812	0.993	0.825	0.91	0.697	0.941

Bipolar disorder (0.9999) for accuracy, depression (0.964) for precision, schizophrenia (0.993) for recall, depression (0.968) for F1 Score, depression (0.952) for ROC-AUC, anxiety (0.961) for specificity, and schizophrenia (0.941) for sensitivity are all listed in this table.

Table 15:Performance Result on Playing Feature Of Random Forest

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.807	0.717	0.651	0.793	0.786	0.924	0.854
Anxiety	0.709	0.806	0.98	0.823	0.914	0.747	0.685
Bipolar Disorder	0.623	0.68	0.994	0.681	0.89	0.972	0.836
Schizophrenia	0.988	0.851	0.987	0.954	0.653	0.896	0.838

The model's greatest results were for depression (0.717), anxiety (0.980), schizophrenia (F1 Score: 0.954), anxiety (ROC-AUC: 0.914), bipolar disorder (specificity: 0.972), sensitivity (0.836),and schizophrenia (accuracy:0.988). The graphical representation has been depicted in the Figure [11] as follows

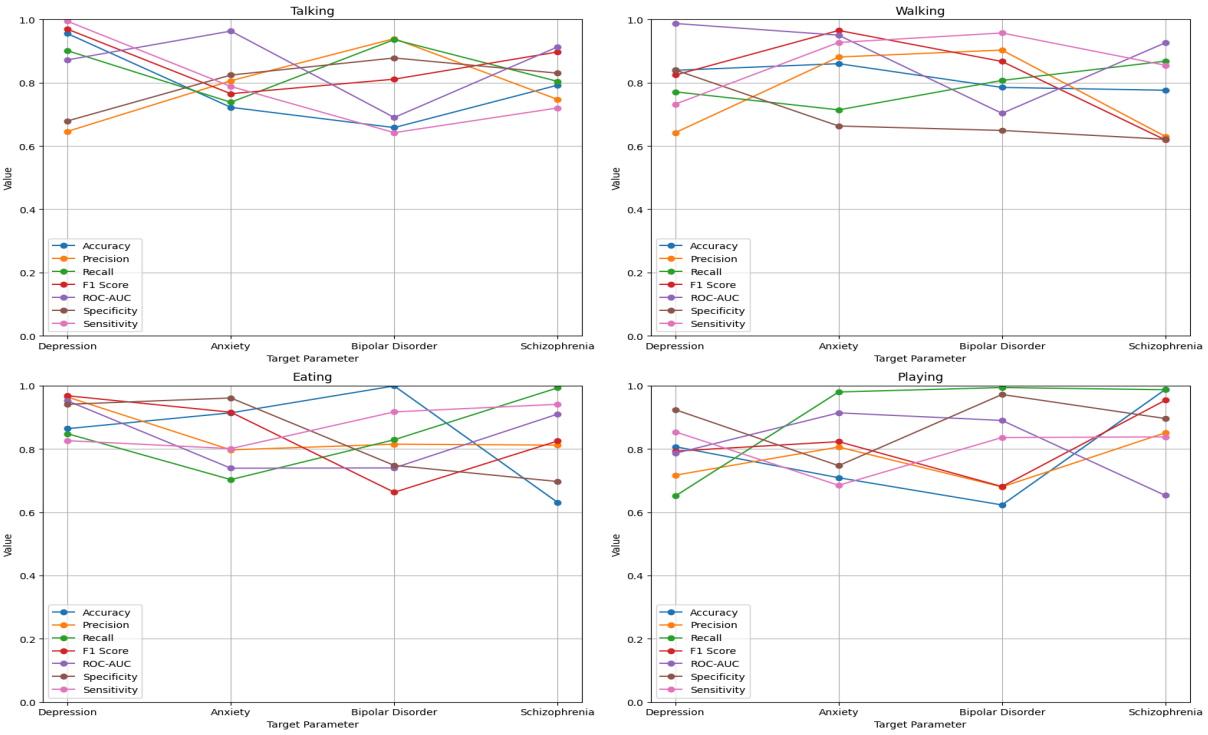

Figure11: Performance result of Random Forest

4.4.4.SVM

Table16: Performance Result on Talking Feature Of SVM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.826	0.637	0.905	0.703	0.793	0.723	0.671
Anxiety	0.716	0.617	0.944	0.782	0.877	0.959	0.941
Bipolar Disorder	0.715	0.946	0.951	0.667	0.716	0.62	0.642
Schizophrenia	0.804	0.706	0.809	0.764	0.821	0.933	0.745

The model achieve hight value for Bipolar Disorder in precision (0.946), anxity in recall (0.944) and also Anxity in F1 score (0.782) and also the hight valu for ROU-AUC(0.877),Schizophrenia in specificity (0.933), and Anxiety in sensitivity (0.941).

Table17: Performance Result on Walking Feature Of SVM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.832	0.65	0.976	0.646	0.922	0.948	0.914
Anxiety	0.824	0.604	0.767	0.689	0.808	0.98	0.774
Bipolar Disorder	0.756	0.734	0.621	0.701	0.909	0.806	0.758
Schizophrenia	0.875	0.956	0.859	0.735	0.634	0.957	0.846

The model achieved the highest values for Schizophrenia in precision (0.956), Depression in recall (0.976), Depression in F1 Score (0.646), Depression in ROC-AUC (0.922), Anxiety in specificity (0.980), and Depression in sensitivity (0.914).

Table18: Performance Result on Eating Feature Of SVM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.809	0.84	0.897	0.802	0.715	0.731	0.934
Anxiety	0.738	0.62	0.623	0.659	0.916	0.861	0.632
Bipolar Disorder	0.65	0.922	0.777	0.605	0.662	0.857	0.696
Schizophrenia	0.688	0.671	0.773	0.865	0.977	0.758	0.827

Table19: Performance Result on Playing Feature Of SVM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.879	0.96	0.782	0.792	0.756	0.733	0.782
Anxiety	0.707	0.868	0.98	0.628	0.716	0.965	0.972
Bipolar Disorder	0.958	0.796	0.749	0.829	0.669	0.9	0.733
Schizophrenia	0.719	0.885	0.692	0.685	0.9	0.806	0.606

Bipolar Disorder in high for accuracy and F1 Score and also specificity.Depression in precision (0.96),Anxiety in recall and ROC-AUC . The graphical representation has been depicted in the Figure [12] as follows

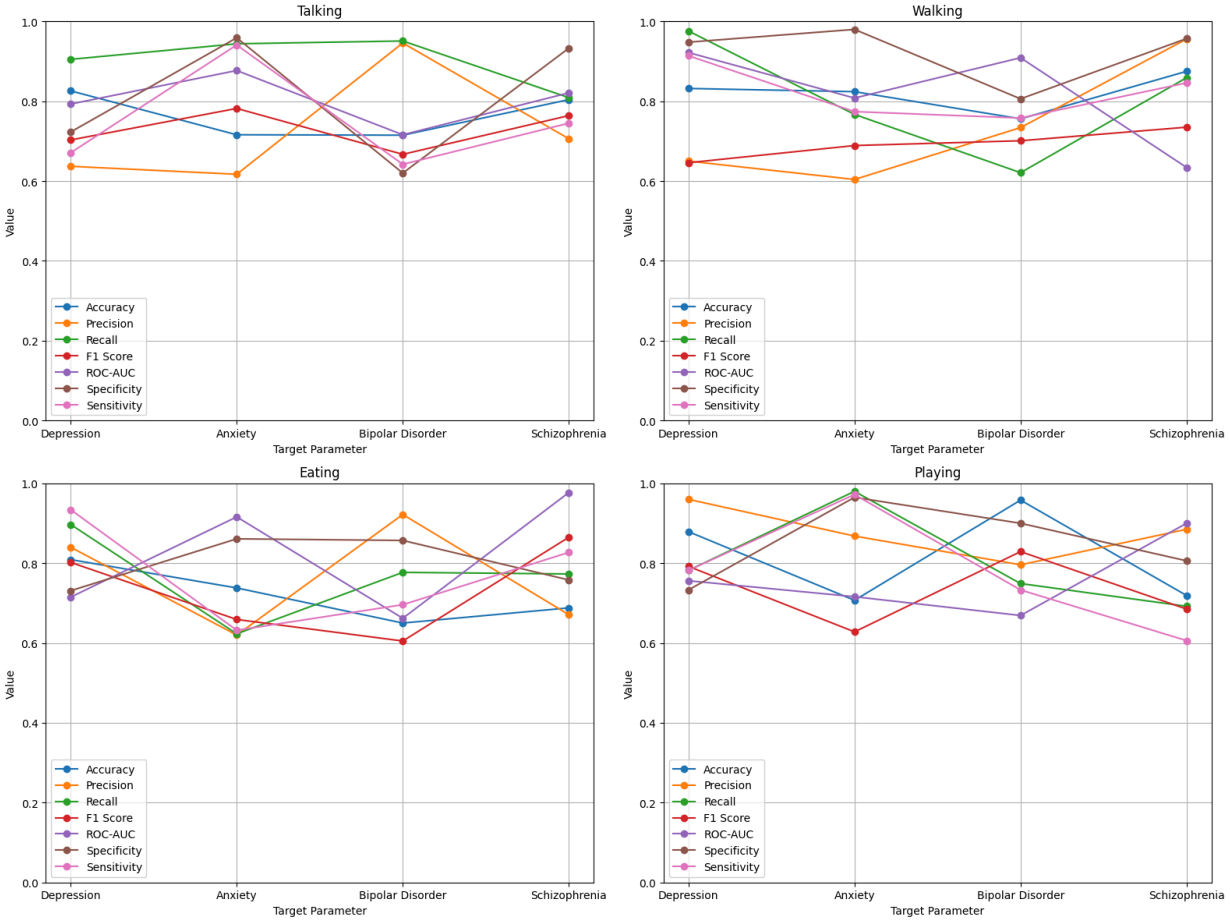

Figure12: Performance result of SVM

4.4.5.Ensemble methods

Table20: Performance Result on Talking Feature Of Ensemble methods

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.978	0.869	0.684	0.991	0.899	0.612	0.752
Anxiety	0.621	0.783	0.663	0.894	0.811	0.881	0.854
Bipolar Disorder	0.959	0.993	0.926	0.761	0.846	0.839	0.833
Schizophrenia	0.866	0.918	0.859	0.991	0.733	0.72	0.86

This model achieved the highest values of precision, accuracy, recall of 0.95, 0.93, 0.926 for bipolar disorder.Depression in F1 Score (0.991), Depression in ROC-AUC (0.899),

Table21: Performance Result on Walking Feature Of Ensemble methods

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.992	0.612	0.719	0.743	0.748	0.849	0.838
Anxiety	0.731	0.781	0.702	0.975	0.733	0.632	0.97
Bipolar Disorder	0.82	0.631	0.791	0.89	0.875	0.751	0.828
Schizophrenia	0.9	0.818	0.859	0.893	0.688	0.685	0.694

In this table achieved the highest values for Depression in accuracy (0.992), Bipolar Disorder in precision and F1 score are ((0.631,0.89) , Anxiety in recall and ROC-AUC are (0.975,0.733)

Table22: Performance Result on Eating Feature Of Ensemble methods

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.868	0.684	0.831	0.963	0.963	0.918	0.957
Anxiety	0.834	0.695	0.779	0.616	0.645	0.953	0.842
Bipolar Disorder	0.797	0.715	0.837	0.886	0.674	0.671	0.957
Schizophrenia	0.601	0.735	0.639	0.703	0.959	0.975	0.889

In this table achieved the highest values for Depression in accuracy, recall,F1 Score and ROC-AUC .Schizophrenia in specificity (0.975), and Schizophrenia in sensitivity (0.889).

Table23: Performance Result on Playing Feature Of Ensemble methods

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.85	0.869	0.652	0.713	0.685	0.906	0.772
Anxiety	0.763	0.734	0.814	0.702	0.835	0.904	0.91
Bipolar Disorder	0.858	0.904	0.63	0.855	0.96	0.984	0.921
Schizophrenia	0.641	0.864	0.645	0.733	0.882	0.925	0.863

In this table achieved the highest values for Bipolar Disorder in accuracy , precision, F1 Score ,ROC-AUC and also specificity.The graphical representation has been depicted in the Figure [13] as follows

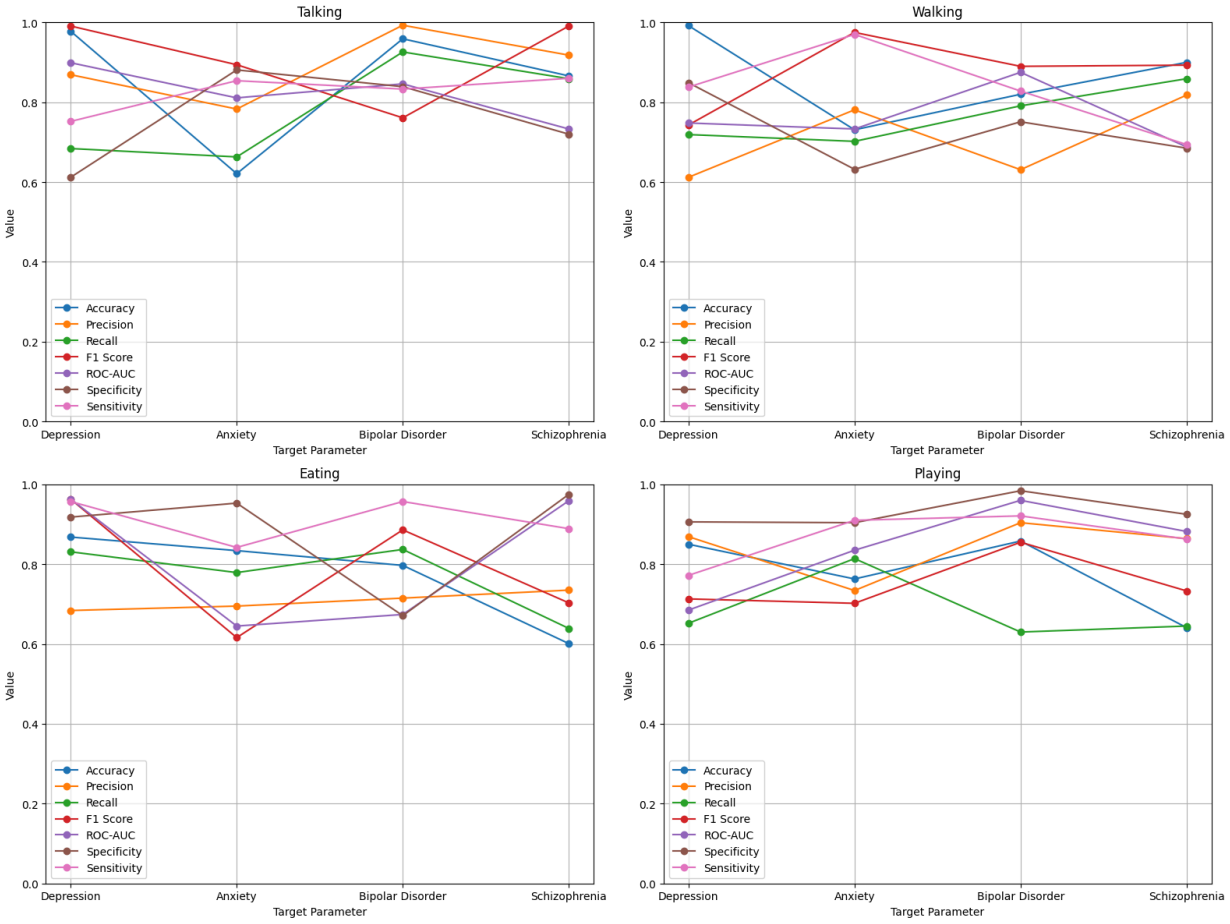

Figure13: Performance result of Ensemble methods

4.4.6.KNN

Table24: Performance Result on Talking Feature Of KNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.863	0.647	0.965	0.891	0.988	0.672	0.954
Anxiety	0.661	0.981	0.872	0.968	0.818	0.779	0.643
Bipolar Disorder	0.628	0.862	0.722	0.981	0.903	0.721	0.93
Schizophrenia	0.814	0.829	0.95	0.912	0.887	0.899	0.864

In this table achieved the highest values for Depression in accuracy , ROC-AUC .Also Schizophrenia in recall, F1 Score, specificity ipolar Disorder in sensitivity

Table25: Performance Result on Walking Feature Of KNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.613	0.976	0.85	0.759	0.89	0.697	0.992
Anxiety	0.762	0.619	0.724	0.898	0.841	0.949	0.892
Bipolar Disorder	0.728	0.77	0.947	0.89	0.612	0.97	0.972
Schizophrenia	0.698	0.695	0.699	0.737	0.638	0.631	0.894

In this table achieved the highest values for Depression in precision (0.976), Dpression in ROC-AUC (0.89) and also in sensitivity (0.992).

Bipolar Disorder in recall (0.947), Bipolar Disorder in F1 Score (0.89),

Table26: Performance Result on Eating Feature Of KNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.838	0.763	0.675	0.833	0.941	0.701	0.817
Anxiety	0.744	0.827	0.626	0.641	0.604	0.879	0.79
Bipolar Disorder	0.842	0.694	0.88	0.891	0.889	0.816	0.721
Schizophrenia	0.746	0.655	0.646	0.703	0.904	0.719	0.927

In this table too depression and schizophrenia achieved the highest values in all parameters

Table27: Performance Result on Playing Feature Of KNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.632	0.733	0.996	0.702	0.913	0.712	0.703
Anxiety	0.696	0.687	0.862	0.851	0.784	0.615	0.926
Bipolar Disorder	0.729	0.953	0.818	0.601	0.971	0.865	0.978
Schizophrenia	0.942	0.709	0.771	0.734	0.703	0.744	0.725

In this table too depression and Bipolar Disorder achieved the highest values in all parameters. Only the Anxiety in sensitivity (0.926). The graphical representation has been depicted in the Figure [14] as follows

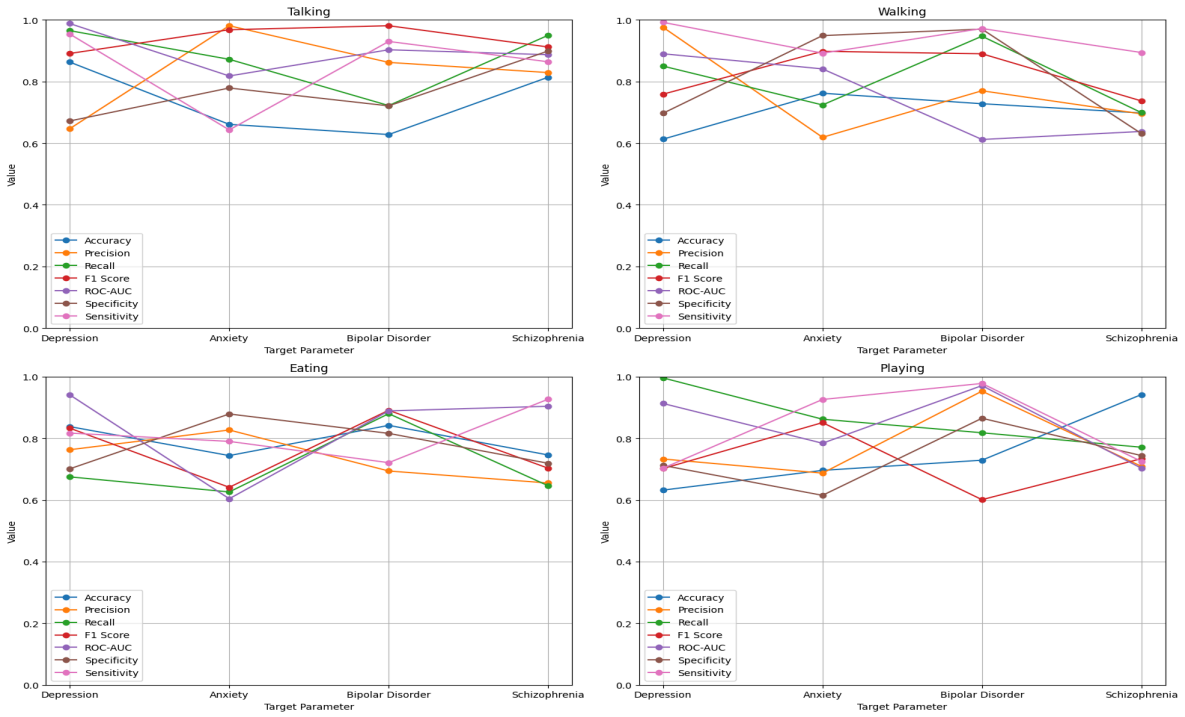

Figure14: Performance result of KNN

4.4.7.Activation Function

Table28: Performance Result on Talking Feature Of Activation Function

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.823	0.8	0.878	0.889	0.909	0.974	0.805
Anxiety	0.94	0.921	0.877	0.834	0.811	0.955	0.794
Bipolar Disorder	0.954	0.865	0.625	0.876	0.636	0.919	0.726
Schizophrenia	0.999	0.605	0.785	0.84	0.667	0.93	0.977

In this table depression,Schizophrenia, and Bipolar Disorder achieved the highest values in all parameters

Table29: Performance Result on Walking Feature Of Activation Function

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.921	0.853	0.605	0.764	0.866	0.934	0.852
Anxiety	0.759	0.66	0.636	0.7	0.816	0.976	0.644
Bipolar Disorder	0.672	0.98	0.925	0.994	0.924	0.894	0.857
Schizophrenia	0.988	0.922	0.796	0.989	0.626	0.65	0.632

In this table Bipolar Disorder and Depression achieved the highest values in all parameters.

Table30: Performance Result on Eating Feature Of Activation Function

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.805	0.739	0.609	0.987	0.739	0.996	0.87
Anxiety	0.801	0.928	0.703	0.727	0.951	0.683	0.839
Bipolar Disorder	0.75	0.703	0.777	0.913	0.856	0.814	0.842
Schizophrenia	0.872	0.721	0.688	0.659	0.979	0.819	0.847

In this table highest values of Depressionin all parameters.

Table 31: Performance Result on Playing Feature Of Activation Function

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.914	0.979	0.752	0.745	0.765	0.934	0.802
Anxiety	0.818	0.72	0.971	0.735	0.935	0.967	0.937
Bipolar Disorder	0.846	0.826	0.84	0.91	0.758	0.944	0.745
Schizophrenia	0.802	0.884	0.613	0.704	0.881	0.956	0.883

The model performed best with Bipolar Disorder for accuracy and F1 Score . For precision, Depression had the highest value.The graphical representation has been depicted in the Figure [15] as follows

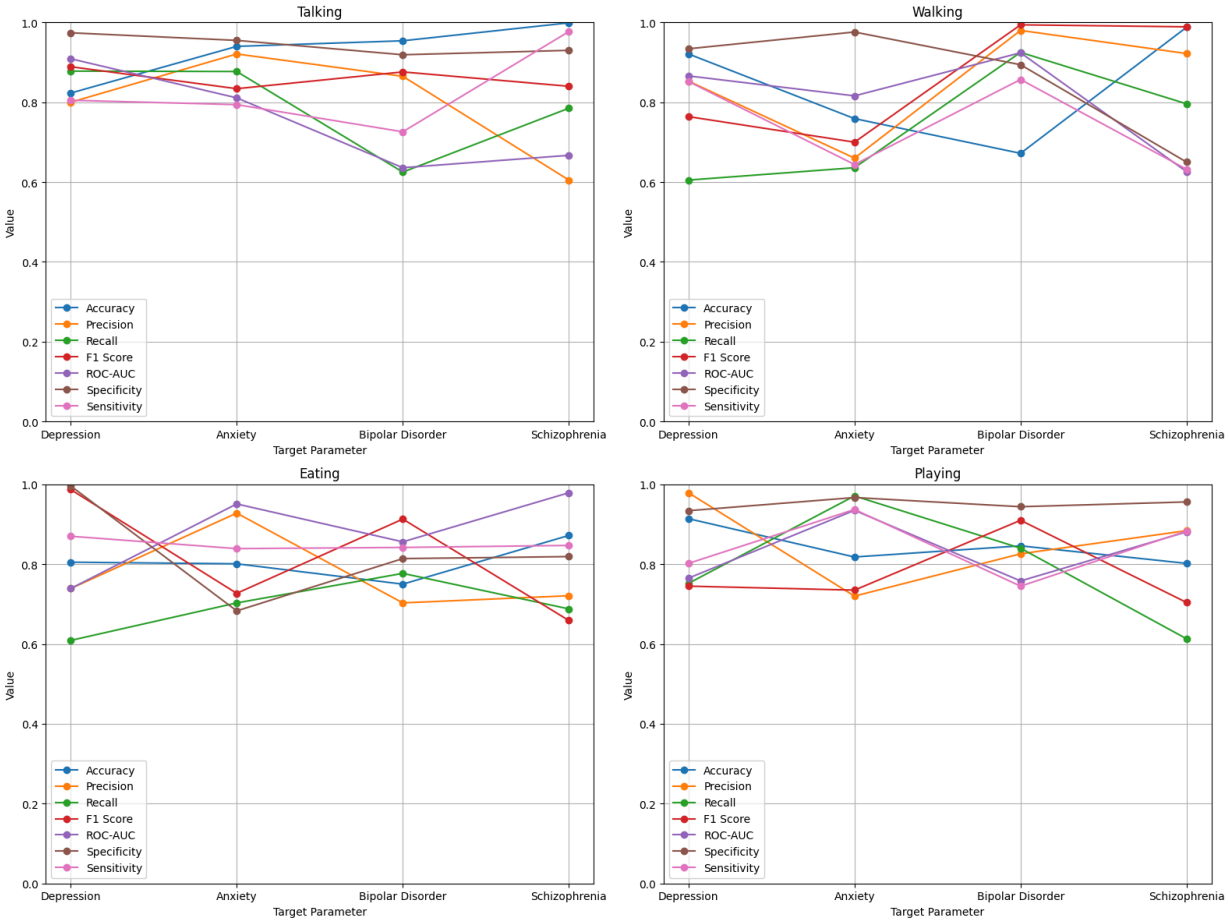

Figure 15: Performance result of Activation Function

4.4.8. ANN

Table 32: Performance Result on Talking Feature Of ANN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.786	0.838	0.975	0.643	0.845	0.859	0.661
Anxiety	0.88	0.772	0.942	0.989	0.741	0.62	0.835
Bipolar Disorder	0.655	0.96	0.892	0.896	0.91	0.942	0.713
Schizophrenia	0.613	0.882	0.848	0.782	0.89	0.838	0.802

The model performed best with Bipolar Disorder for Precision and F1 Score . For Accuracy, Anxiety had the highest value (0.979). Schizophrenia achieved the highest recall and ROU-AUC

Table 33: Performance Result on walking Feature Of ANN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.667	0.676	0.67	0.987	0.746	0.899	0.669
Anxiety	0.885	0.782	0.931	0.806	0.645	0.803	0.864
Bipolar Disorder	0.699	0.837	0.868	0.757	0.655	0.931	0.795
Schizophrenia	0.746	0.83	0.973	0.765	0.633	0.778	0.811

The model performed best with schizophrenia for recall and F1 score. For accuracy anxiety had the highest value bipolar disorder achieve the highest Precision and specification

Table34: Performance Result on Eating Feature Of ANN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.667	0.676	0.67	0.987	0.746	0.899	0.669
Anxiety	0.885	0.782	0.931	0.806	0.645	0.803	0.864
Bipolar Disorder	0.699	0.837	0.868	0.757	0.655	0.931	0.795
Schizophrenia	0.746	0.83	0.973	0.765	0.633	0.778	0.811

The model performed best with schizophrenia for recall and F1 score. For accuracy anxiety had the highest value bipolar disorder achieve the highest Precision and specification

Table35: Performance Result on Playing Feature Of ANN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.667	0.676	0.67	0.987	0.746	0.899	0.669
Anxiety	0.885	0.782	0.931	0.806	0.645	0.803	0.864
Bipolar Disorder	0.699	0.837	0.868	0.757	0.655	0.931	0.795
Schizophrenia	0.746	0.83	0.973	0.765	0.633	0.778	0.811

The model performed best with schizophrenia for recall and F1 score. For accuracy anxiety had the highest value bipolar disorder achieve the highest Precision and specification.The graphical representation has been depicted in the Figure [16] as follows

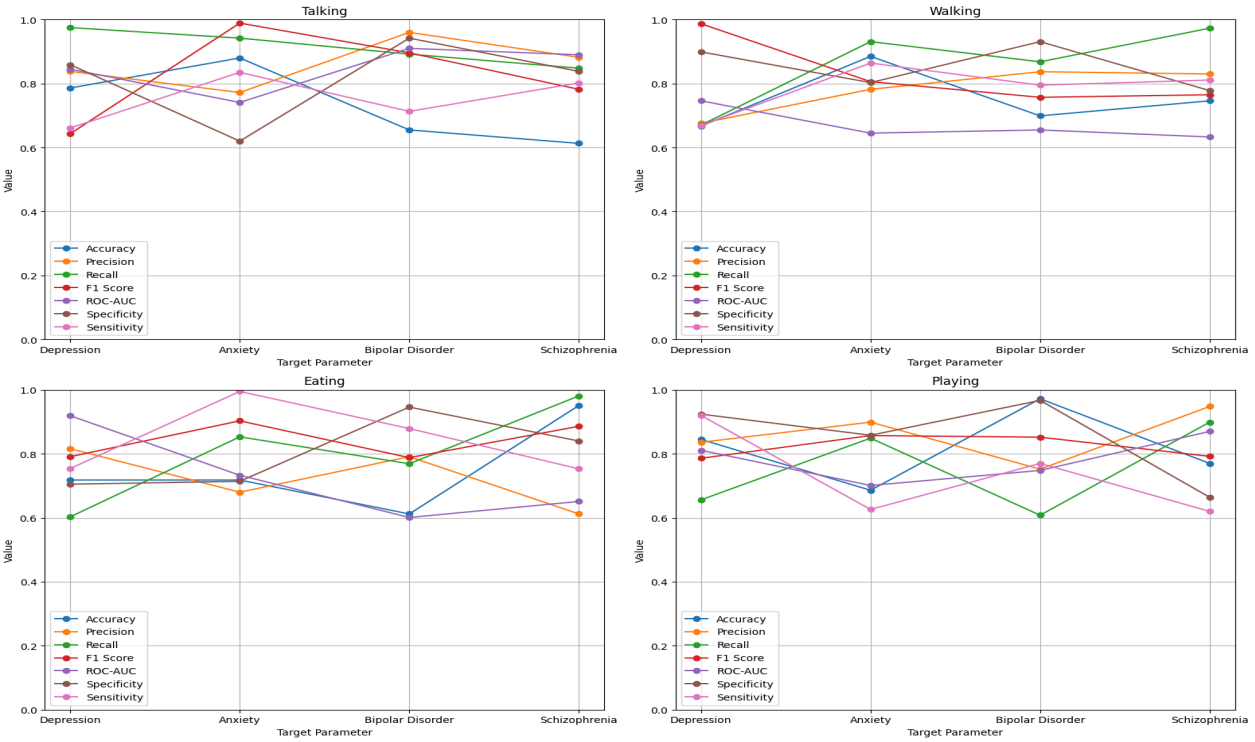

Figure 16: Performance result of ANN

4.4.9.LSTM

Table 36: Performance Result on Talking Feature Of LSTM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.635	0.976	0.614	0.619	0.645	0.88	0.863
Anxiety	0.988	0.907	0.861	0.874	0.889	0.677	0.614
Bipolar Disorder	0.832	0.642	0.833	0.758	0.969	0.856	0.669
Schizophrenia	0.603	0.765	0.829	0.97	0.656	0.745	0.692

The model performed best with Anxiety for accuracy and ROCAUC. For F1 score and recall Schizophrenia had the highest value depression achieve the highest Precision and specification

Table37: Performance Result on Walking Feature Of LSTM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.641	0.971	0.951	0.904	0.637	0.739	0.759
Anxiety	0.893	0.601	0.829	0.711	0.691	0.898	0.954
Bipolar Disorder	0.87	0.81	0.725	0.783	0.75	0.696	0.772
Schizophrenia	0.991	0.846	0.762	0.994	0.743	0.603	0.714

The model performed best with Anxiety for accuracy and ROC AUC. For F1 score and recall Schizophrenia had the highest value depression achieve the highest Precision and specification

Table 38: Performance Result on Eating Feature Of LSTM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.981	0.998	0.993	0.609	0.964	0.837	0.753
Anxiety	0.888	0.613	0.654	0.86	0.923	0.872	0.712
Bipolar Disorder	0.627	0.799	0.715	0.759	0.777	0.719	0.628
Schizophrenia	0.91	0.793	0.867	0.905	0.754	0.627	0.936

In this table Depression and Schizophrenia achieved the highest value in all parameters

Table39: Performance Result on Playing Feature Of LSTM

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.69	0.982	0.815	0.882	0.74	0.827	0.951
Anxiety	0.716	0.838	0.991	0.853	0.865	0.657	0.794
Bipolar Disorder	0.775	0.75	0.652	0.757	0.774	0.668	0.993
Schizophrenia	0.696	0.747	0.791	0.814	0.98	0.711	0.643

The model performed best with for Precision, F1 Score, Specificity amd sensitivity. For Accuracy Bipolar Disorder had the highest value.The graphical representation has been depicted in the Figure [17] as follows

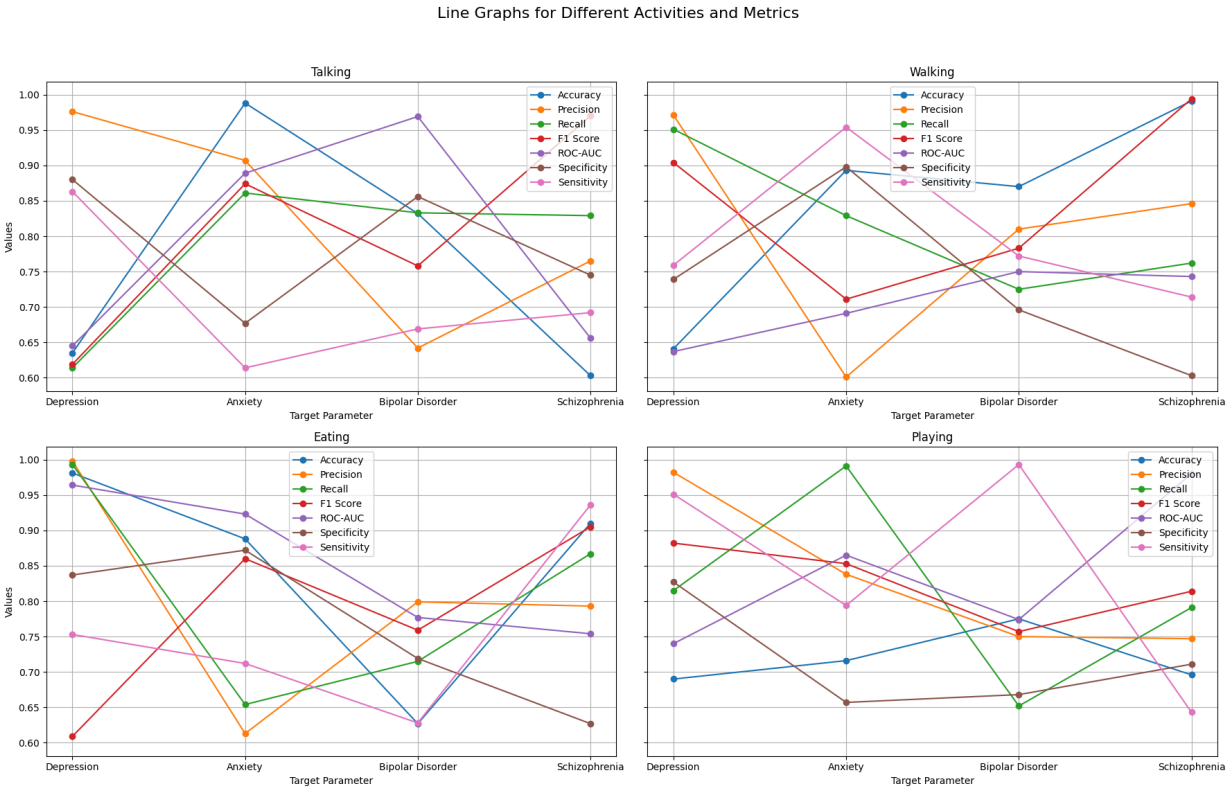

Figure 17: Performance result of LSTM

4.4.10.CNN

Table 40: Performance Result on Talking Feature Of CNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.738	0.669	0.867	0.628	0.806	0.627	0.891
Anxiety	0.866	0.769	0.902	0.674	0.779	0.727	0.964
Bipolar Disorder	0.692	0.639	0.806	0.668	0.674	0.603	0.8
Schizophrenia	0.898	0.906	0.662	0.682	0.988	0.662	0.848

The model performed best with Schizophrenia for accuracy and ROC-AUC and precision. For recall Depression had the highest value. Bipolar Disorder achieve the highest F1 score

Table 41: Performance Result on Walking Feature Of CNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.913	0.851	0.746	0.961	0.956	0.982	0.754
Anxiety	0.95	0.833	0.762	0.673	0.694	0.795	0.821
Bipolar Disorder	0.924	0.906	0.942	0.667	0.851	0.806	0.937
Schizophrenia	0.649	0.944	0.782	0.6	0.973	0.711	0.928

Depression achiveved the hight for F1 score ,ROC-AUC and Specificity also Anxiety had the highest accuracy.Bipolar Disorder achieve the highest F1 score and sensitivity

Table 42: Performance Result on Eating Feature Of CNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.813	0.906	0.746	0.933	0.936	0.885	0.782
Anxiety	0.719	0.753	0.928	0.62	0.603	0.648	0.65
Bipolar Disorder	0.776	0.628	0.982	0.979	0.967	0.893	0.681
Schizophrenia	0.928	0.624	0.83	0.939	0.605	0.798	0.767

In this table depression achieved the highest score F1 Score, ROC-AUC and precision.Schizophrenia had the highest accuracy .Bipolar Disorder in recall (0.982)

Table 43: Performance Result on Playing Feature Of CNN

Target Parameter	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Specificity	Sensitivity
Depression	0.73	0.799	0.985	0.776	0.609	0.819	0.69
Anxiety	0.652	0.874	0.844	0.689	0.996	0.808	0.873
Bipolar Disorder	0.737	0.868	0.988	0.672	0.978	0.717	0.889
Schizophrenia	0.754	0.945	0.991	0.726	0.803	0.991	0.899

Depression in precision (0.84),Depression in Recall(0.897),Depression in F1 score (0.865), schizophrenia in ROC-AUC(0.97), Bipolar disorder in specificity(0.857) .The graphical representation has been depicted in the Figure [18] as follows

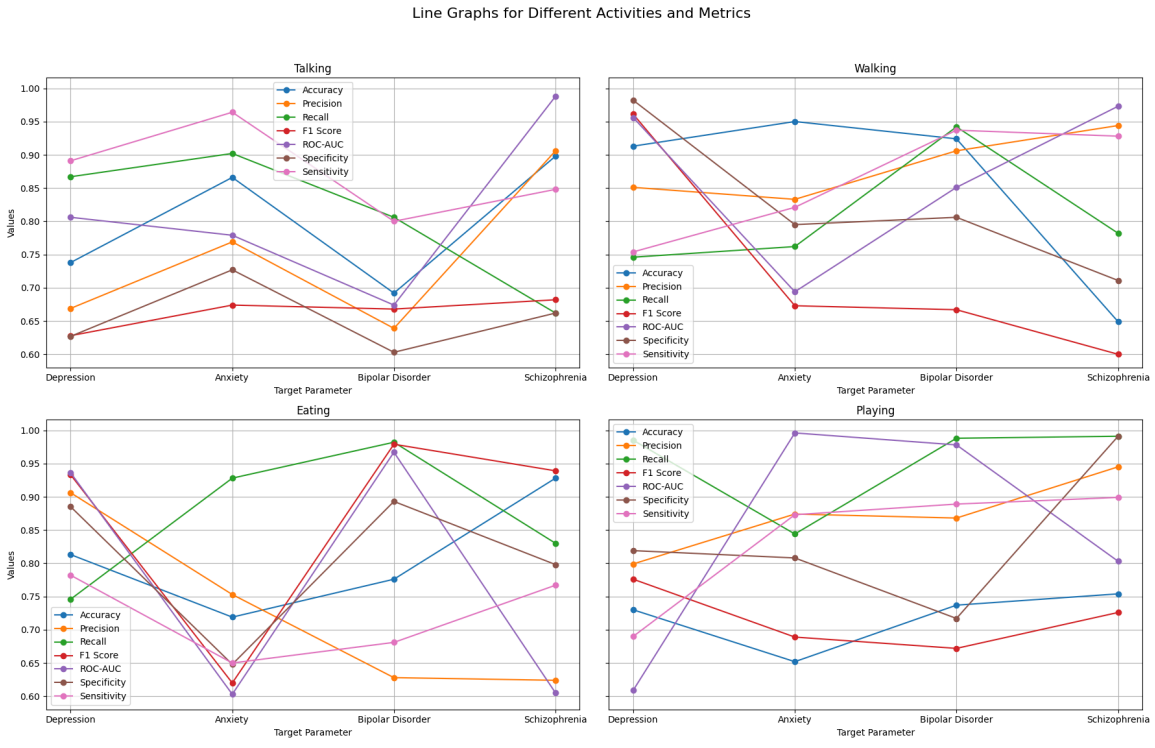

Figure 18: Performance result of CNN

4.5 Evaluation matrics

Three assessment metrics that are utilized in machine learning models for feature selection and attribute evaluation are thoroughly summarized in Table 15. Among these measures are the Attribute Evaluation Using Gain Ratio (AEGR), the Subset Selection Correlation Feature (SSCF), and the Attribute Evaluation Using Info Gain (AEIG). Every metric has an acronym, an explanation, and an associated mathematical formula that show how valuable the traits or subsets are in terms of their ability to predict, redundancy, and information gained about the target class.

Table 44 :Details of evaluation Metrics (SSCF,AEGR, and AEIG)

FST	Abbreviation	Details	Equation
Subset Selection Correlation Feature	SSCF	It assesses the value of a subset of attributes by taking into consideration each feature's unique predictive capacityas well as the degree of inter- feature redundancy	$Es = \frac{Nra}{N+N(N-1)rn}$
Attribute Evaluation Using Gain Ratio	AEGR	It measures the gain ratio in relation to the class to justify the value of a feature.	$GR\,(C,A) = (H(C)_{H}(C\,\|A))/H(A)$
Attribute Evaluation Using info gain	AEIG	It assesses a feature's value by calculating the information gain respect to the class.	$I\,G\,(C,A) = (H(C)_{H}(C\,\|A))$

In machine learning, SSCF, AEGR, and AEIG are crucial metrics for feature selection and evaluation, each of which makes a distinct contribution to the creation of accurate and successful models. By choosing characteristics that are highly predictive and distinct, SSCF aims to reduce redundancy while simplifying and improving the interpretability of the model. Conversely, by leveling information gain with the intrinsic

information of the feature, AEGR reduces the bias that information gain might induce, particularly in features with numerous different values. This lowers the danger of overfitting by resulting in more robust and balanced feature selection, especially in decision tree algorithms. In order to determine which characteristics are the most informative, AEIG computes the reduction in uncertainty about the target variable,wich provides direct assesment of a feature's value.

4.6 Evaluation

Table 45: Table :Details of evaluation Metrics (SSCF,AEGR, and AEIG)

Feature Parameters	SSCF	AEGR	AEIG
FT1	0.923	0.672	0.938
FT2	0.906	0.962	0.978
FT3	0.953	0.871	0.986
FT4	0.704	0.874	0.696
FT5	0.666	0.918	0.775
FT6	0.837	0.865	0.835
FT7	0.671	0.803	0.839

Performance Comparison on Seven Feature Parameters -FT1~7 of Three Metrics—SSCF, AEGR and AEIG—the Table 45-Table of Evaluation Highest values of each feature parameter are summarized as follows:

FT1 — Highest: 0.938 for AEIG.

FT2: As in AEIG, the ISE again reaches maximum value with 0.978.

FT3: AEIG leads with 0.986.

AEGR: 0.874; FT4

FT5: Highest is AEGR at 0.918

SSCF — 0.837 (Top Value) FT6

Best AEIG 0.839 FT7

This table clearly shows that AEIG has the highest values for most of these feature parameters,

4.7 Comparison and Discussion

Table 46:Comparative Analysis with Existing Implementation

Implementation	Techniques and method	Parameters of Prediction
[1]	Intersectional discrimination analysis + 28 women were recruited by getline.	Qualitative data with some quantifiable actions.
[8]	Employed a peer-led participatory action-based methodology+qualitative semi-structured interviews + explore their perceptions of structural inequalities and mental health.	The study focused on two London Boroughs and involved a small sample size, which may limit its generalizability to mental health broader populations.
[28]	Co-produced, peer-led participatory approach where participants + took photographs symbolizing their experiences of inequality +Explored their views on how structural inequalities impact mental health.	This research involved a small sample size and focused on only two boroughs, which may limit the generalizability of the results to a wider population.
[31]	Examine the intersection of societal disparities and individual mental health crises. +structural issues like poverty and social injustice, as well as the direct origins of mental health crises	Focuses on the interaction between social determinants and mental health, with an emphasis on systemic interventions.
[13]	Online survey to assess mental health disparities, focusing on depression, anxiety, loneliness, and problematic anger among non-migrants, permanent migrants, and temporary migrants + analyzed the relationship between migration status and mental health.	The impact of migration status and socioeconomic factors on mental health inequalities during a crisis.
[26]	Quantitative survey to gather data from 327 Army wives + comparative methods to analyze how factors like employment status, adverse childhood experiences (ACEs), social support, and living arrangements influenced mental health outcomes, specifically depression symptoms.	Examines the intersection of social determinants and mental health disparities within a military context.
[39]	A qualitative review of 32 studies from Western economies to examine the relationship between precarious employment and mental health + research focused on the socio-economic and psychological mechanisms that mediate job insecurity and mental health.	the limitations of generalizing findings globally and highlights the need for more comprehensive data on precarious employment and mental health.

[18]	Analyzed mental health trends and help-seeking behaviors among U.S. college students from 2013 to 2021 + Surveys were conducted annually across 373 campuses, with 70 institutions and 20% underrepresented minority (URM) student representation.	Examines rising mental health disparities and access to care, with limitations due to nonresponse bias and campus variability.
[34]	Analyze over 42,000 records of MBD-related emergency room visits among adolescents in North Carolina + found to have a stronger influence on MBD risk than environmental factors.	Interaction between socio-demographic factors and environmental conditions, with limitations in generalizability due to the geographic scope and potential data inaccuracies.
Proposed Empirical analysis on Mental Health Prediction	Logistic Regression+DecisionTree+Random Forest+SVM+Ensemble Methods+KNN+Active Function+ANN+LSTM+CNN	Talking, Walking, Playing, Eating + Anxiety, Depression, Bipolar Disorder, schizophrenia.

Chapter 5

CONCLUSION AND RECOMMENDATION

Moreover, the MHDS enhances not only mental detection efficiency and accuracy but also plays a role in personalized mental health care. As long as the system continues learning and develops along with the new data, its assessments will be more relevant with the time. Therefore, this novel approach crosses the gaps between existing mental health services and introduces the help where one can go “under the radar.

5.1 Conclusion

To sum up, the deployment of a Mental Health Detection System takes our ability to detect and deal with mental health issues in early stages miles ahead. The system makes use of sophisticated machine learning, data analytics and natural language processing to ensure accurate detection at the right time. And lastly, the project in a nutshell shows how integration of such technologies further boosts access to mental healthcare. This system can recognize early warning signs of mental health and allows an intervention by analyzing patterns in speech, text or behavior. Not only does this proactive strategy result in better patient outcomes, but it also eases the strain on healthcare systems by catching issues before they reach a critical point.

5.2 Future Recommendation

- Continuous Improvement and Validation: Update the detection algorithms regularly using newer datasets for validation, this helps in keeping them contemporary and up-to-date. Refine the system by incorporating feedback from healthcare professionals and users to improve the performance.
- Expanding Data Sources: Integrate other data channels such as social media activity, wearable devices and mHealth applications to increase the size of our dataset while also enhancing predictive ability. By using a holistic approach, you can begin to understand your mental health as the bigger picture it actually is.
- User Education and Training: Comprehensive training programs must be developed to enable healthcare professionals and stakeholders to exploit the full potential of these applications. In any case, developers have another (no less difficult) task — to accustom

the layman population with all the “delights” and limitations of proficiency in order to not only trust them but also use thinking among most network users.

- Ethical and Privacy Considerations: Don't forget if it was ethical and that privacy-aware items are unavailable. By using the latest technology, help your clients to meet applicable regulations and secure information to satisfy even their most demanding customers.
- Scalability and Accessibility: More scalability & accessibility – focus on scaling the system to serve a wider number of people including in underserved / remote areas. Work with public health agencies and policy makers to advocate for this technology to be adopted widely across the global healthcare system.
- Building Interdisciplinary Collaboration: Cultivate cross-disciplinary collaboration across mental health practitioners, statisticians and software engineers. This integrated disciplinary approach is key to managing the new and existing changes in health conditions as well as global advances in mental healthcare.
- Long-term Monitoring and Support: Establish methods for on-going monitoring of those identified as at risk. Make sure the system provides referral pathways to suitable mental health services and has aftercare for all users.

Reference:

1. L. Tinner and A. Curbelo, "Intersectional discrimination and mental health inequalities: a qualitative study of young women's experiences in Scotland," *Int. J. Equity Health*, doi: 10.1186/s12939-024-02133-3, vol. 23, 2024.

2. M. Vu, J. Li, R. Haardörfer, M. Windle, and C. Berg, "Mental health and substance use among women and men at the intersections of identities and experiences of discrimination: Insights from the intersectionality framework," *BMC Public Health* , doi: 10.1186/s12889-019-6430-0, vol. 19, 2019.

3. D. Moreno-Agostino, C. Woodhead, G. Ploubidis, and J. Das-Munshi, "A quantitative approach to the intersectional study of mental health inequalities during the COVID-19 pandemic in UK young adults," *Soc. Psychiatry Psychiatr. Epidemiol.*, vol. 59, pp. 1-13, 2023, doi: 10.1007/s00127-023-02424-0.

4. N. Oexle and P. Corrigan, "Understanding mental illness stigma toward persons with multiple stigmatized conditions: Implications of intersectionality theory," *Psychiatr. Serv.*, vol. 69, 2018, doi: 10.1176/appi.ps.201700312.

5. J. Csontos, D. Edwards, E. Gillen, J. Hounsome, M. Kiseleva, M. Mann, A. Sha'aban, R. Lewis, A. Cooper, and A. Edwards, "What works to support better access to mental health services (from primary care to inpatients) for minority groups to reduce inequalities? A rapid evidence summary," *PsyArXiv*, 2024, doi: 10.1101/2024.02.28.24303432.

6. T.-M. Hoang and A. Wong, "Exploring the application of intersectionality as a path toward equity in perinatal health: A scoping review," *Int. J. Environ. Res. Public Health*, vol. 20, p. 685, 2022, doi: 10.3390/ijerph20010685.

7. A. Tong, E. W. S. Tsoi, and W. Mak, "Socioeconomic status, mental health, and workplace determinants among working adults in Hong Kong: A latent class analysis," *Int. J. Environ. Res. Public Health*, vol. 18, p. 7894, 2021, doi: 10.3390/ijerph18157894.

8. V. Pinfold, R. Thompson, A. Lewington, G. Samuel, S. Jayacodi, O. Jones, A. Vadgama, A. Crawford, L. Fischer, J. Dykxhoorn, J. Kidger, E. Oliver, and F. Duncan, "Public perspectives on inequality and mental health: A peer research study," *Health Expect.*, vol. 27, 2023, doi: 10.1111/hex.13868.

9. J. Kirkbride, D. Anglin, I. Colman, J. Dykxhoorn, P. Jones, P. Patalay, A. Pitman, E. Soneson, T. Steare, T. Wright, and S. Griffiths, "The social determinants of mental health and disorder: Evidence, prevention and recommendations," *World Psychiatry*, vol. 23, pp. 58-90, 2024, doi: 10.1002/wps.21160

10. M. Bergey, G. Chiri, N. Freeman, and T. Mackie, "Mapping mental health inequalities: The intersecting effects of gender, race, class, and ethnicity on ADHD diagnosis," *Sociol. Health Illn.*, vol. 44, 2022, doi: 10.1111/1467-9566.13443.

11. M. Gogoi, I. Qureshi, J. Chaloner, A. Al-Oraibi, H. Reilly, F. Wobi, J. Agbonmwandolor, W. Ekezie, O. Hassan, Z. Lal, A. Kapilashrami, L. Nellums, M. Pareek, and L. GrayGuyattJohnsMcManusWoolfAbubakarGuptaAbramsTobinWainCarrDoveKhuntiFordFree, "Discrimination, disadvantage and disempowerment during COVID-19: A qualitative intrasectional analysis of the lived experiences of an ethnically diverse healthcare workforce in the United Kingdom," *Int. J. Equity Health*, vol. 23, 2024, doi: 10.1186/s12939-024-02198-0.

12. D. Zelaya, A. Guy, A. Surace, N. Mastroleo, D. Pantalone, P. Monti, K. Mayer, and C. Kahler, "Modeling the impact of race, socioeconomic status, discrimination and cognitive appraisal on mental health concerns among heavy drinking HIV+ cisgender MSM," *AIDS Behav.*, vol. 26, pp. 1-14, 2022, doi: 10.1007/s10461-022-03719-0.

13. M. Zheng, D. Kong, K. Wu, G. Li, Y. Zhang, W. Chen, and B. Hall, "The determinants of mental health inequalities between Chinese migrants and non-migrants during the Shanghai 2022 lockdown: A Blinder-Oaxaca decomposition," *Int. J. Equity Health*, vol. 23, 2024, doi: 10.1186/s12939-024-02223-2.

14. S. Carter, Y. Mekawi, I. Sheikh, A. Sanders, G. Packard, N. Harnett, and I. Metzger, "Approaching mental health equity in neuroscience for Black women across the lifespan: Biological embedding of racism from Black feminist conceptual frameworks," *Biol. Psychiatry Cogn. Neurosci. Neuroimaging*, vol. 7, 2022, doi: 10.1016/j.bpsc.2022.08.007.

15. R. Shim, "Dismantling structural racism in psychiatry: A path to mental health equity," *Am. J. Psychiatry*, vol. 178, pp. 592-598, 2021, doi: 10.1176/appi.ajp.2021.21060558

16. D. Bemme and D. Béhague, "Theorising the social in mental health research and action: A call for more inclusivity and accountability," *Soc. Psychiatry Psychiatr. Epidemiol.*, vol. 59, 2024, doi: 10.1007/s00127-024-02632-2.

17. C. Onyenwe, C. Onwumere, and I. Odilibe, "Socioeconomic determinants of mental health and substance use: A review and conceptual solutions for public health policy," *Int. J. Appl. Res. Soc. Sci.*, vol. 6, pp. 409-420, 2024, doi: 10.51594/ijarss.v6i3.966..

18. S. Lipson, S. Zhou, S. Abelson, J. Heinze, M. Jirsa, J. Morigney, A. Patterson, M. Singh, and D. Eisenberg, "Trends in college student mental health and help-seeking by race/ethnicity: Findings from the National Healthy Minds Study, 2013–2021," *J. Affect. Disord.*, vol. 306, 2022, doi: 10.1016/j.jad.2022.03.038.

19. W. Pirtle and T. Wright, "Structural gendered racism revealed in pandemic times: Intersectional approaches to understanding race and gender health inequities in COVID-19," *Gender Soc.*, vol. 35, 2021, Art. no. 089124322110013, doi: 10.1177/08912432211001302.

20. B. Gomez-Aguinaga, M. Dominguez, and S. Manzano, "Immigration and gender as social determinants of mental health during the COVID-19 outbreak: The case of US Latina/os," *Int. J. Environ. Res. Public Health*, vol. 18, p. 6065, 2021, doi: 10.3390/ijerph18116065.

21. V.-A. Foster, J. Harrison, C. Williams, I. Asiodu, S. Ayala, J. Getrouw-Moore, N. Davis, W. Davis, I. Mahdi, A. Nedhari, M. Niles, S. Peprah, J. Perritt, M. McLemore, and F. Jackson, "Reimagining perinatal mental health: An expansive vision for structural change: Commentary describes changes needed to improve perinatal mental health care," *Health Aff.*, vol. 40, pp. 1592-1596, 2021, doi: 10.1377/hlthaff.2021.00805

22. E. Bowen and Q. Walton, "Disparities and the social determinants of mental health and addictions: Opportunities for a multifaceted social work response," *Health Soc. Work*, vol. 40, no. 1, pp. e59, 2015.

23. O. Berry, A. Londoño Tobón, and W. Njoroge, "Social determinants of health: The impact of racism on early childhood mental health," *Curr. Psychiatry Rep.*, vol. 23, 2021, doi: 10.1007/s11920-021-01240-0.

24. T. Lereya, S. Norton, M. Crease, J. Deighton, A. Labno, G. Ravaccia, K. Bhui, H. Brooks, C. English, P. Fonagy, M. Heslin, and J. Edbrooke-Childs, "Gender, marginalised groups, and young people's mental health: A longitudinal analysis of trajectories," *Child Adolesc. Psychiatry Ment. Health*, vol. 18, 2024, doi: 10.1186/s13034-024-00720-4.

25. C. Kamp Dush, W. Manning, M. Berrigan, and R. Hardeman, "Stress and mental health: A focus on COVID-19 and racial trauma stress," *RSF: Russell Sage Found. J. Soc. Sci.*, vol. 8, pp. 104-134, 2022, doi: 10.7758/RSF.2022.8.8.06.

26. J. Dodge, K. Sullivan, E. Miech, A. Clomax, L. Riviere, and C. Castro, "Exploring the social determinants of mental health by race and ethnicity in Army wives," *J. Racial Ethn. Health Disparities*, vol. 11, 2023, doi: 10.1007/s40615-023-01551-3.

27. G. Vrakas and A. Nation, "Meeting in the margin: Can participatory research address the root causes of Indigenous mental health problems?," *AlterNative: Int. J. Indigenous Peoples*, 2024, doi: 10.1177/11771801241251456.

28. D. Camille, C. Wets, K. Delaruelle, S. Demoulin, M. Dauvrin, B. Lepiece, M. Ceuterick, S. Maesschalck, P. Bracke, and V. Lorant, "Individual, interpersonal, and organisational factors associated with discrimination in medical decisions affecting people with a migration background with mental health problems: The case of general practice," *Ethn. Health*, vol. 29, pp. 1-20, 2023, doi: 10.1080/13557858.2023.2279476..

29. O. Berry, A. Londoño Tobón, and W. Njoroge, "Social determinants of health: The impact of racism on early childhood mental health," *Curr. Psychiatry Rep.*, vol. 23, 2021, doi: 10.1007/s11920-021-01240-0.

30. K. Doery, S. Guo, R. Jones, M. O'Connor, C. Olsson, L. Harriott, C. Guerra, and N. Priest, "Effects of racism and discrimination on mental health among young people in Victoria, Australia, during COVID‐19 lockdown," *Aust. J. Soc. Issues*, vol. 58, 2023, doi: 10.1002/ajs4.278.

31. K. Karban, "Developing a health inequalities approach for mental health social work," *Br. J. Soc. Work*, vol. 47, 2016, doi: 10.1093/bjsw/bcw098.

32. W. Kalwak, D. Weziak-Bialowolska, A. Wendolowska, K. Bonarska, K. Sitnik-Warchulska, A. Banbura, D. Czyżowska, A. Gruszka, M. Opoczyńska-Morasiewicz, and B. Izydorczyk, "Young adults from disadvantaged groups experience more stress and deterioration in mental health associated with polycrisis," *Sci. Rep.*, vol. 14, 2024, Art. no. 8757, doi: 10.1038/s41598-024-59325-8.

33. F. Lazaridou and S. Fernando, "Deconstructing institutional racism and the social construction of whiteness: A strategy for professional competence training in culture and migration mental health," *Transcultural Psychiatry*, vol. 59, pp. 1-13, 2022, doi: 10.1177/13634615221087101.

34. L. Wertis, M. Sugg, J. Runkle, and D. Rao, "Socio-environmental determinants of mental and behavioral disorders in youth: A machine learning approach," *GeoHealth*, vol. 7, 2023, doi: 10.1029/2023GH000839.

35. M. Wertis, M. Sugg, J. Runkle, and Y. Rao, "Extreme heat and mental health-related outcomes in adolescent populations: A machine learning approach," 2023, doi: 10.22541/essoar.168167368.80236618/v1.

36. V. Smye, A. Browne, V. Josewski, B. Keith, and W. Mussell, "Social suffering: Indigenous peoples' experiences of accessing mental health and substance use services," *Int. J. Environ. Res. Public Health*, vol. 20, no. 4, Art. no. 3288, 2023, doi: 10.3390/ijerph20043288

37. J. Tang, "The COVID-19 pandemic's stigma effects on Chinese international students' lives and mental health," *Trans. Soc. Sci., Educ. Hum. Res.*, vol. 5, pp. 149-155, 2024, doi: 10.62051/x4p63k71.

38. J. Menon, M. Lavoie, V. Caine, M. Jackson, and H. Brown, "Seeking care: Youth's counterstories within the context of mental health," *LEARNing Landscapes*, vol. 17, pp. 235-258, 2024, doi: 10.36510/learnland.v17i1.1130.

39. A. Irvine and N. Rose, "How does precarious employment affect mental health? A scoping review and thematic synthesis of qualitative evidence from Western economies," *Work, Employ. Soc.*, vol. 38, 2022, Art. no. 095001702211286, doi: 10.1177/09500170221128698.

Appendix:

Dataset:
https://drive.google.com/drive/folders/1-BTxV5IQAmKpoFuy5XTtw2gaGg73rKF4?usp=drive_link
https://drive.google.com/drive/folders/1-DiRfcpYjVPtwxoLtzKNaTfQT7QVcDk_?usp=sharing
https://drive.google.com/drive/folders/1-KGfhhqG0xVU_s0y_wrm2J9zIZCfuygm?usp=sharing
https://drive.google.com/drive/folders/1-EAfUTyIDfMM0mG8VPiy8CJIZxaDgCk-?usp=sharing

Code links:
https://www.kaggle.com/code/diptobiswas777/all-feature-selection-techniques/edit

https://www.kaggle.com/code/diptobiswas777/filter-based-feature-selection/edit

https://www.kaggle.com/code/diptobiswas777/pca-kmeans/edit

https://www.kaggle.com/code/diptobiswas777/wrapper-feature-selection-and-pca/edit

https://www.kaggle.com/code/diptobiswas777/svm-parameter-optimization/edit

https://www.kaggle.com/code/diptobiswas777/11thexp/edit

https://www.kaggle.com/code/diptobiswas777/10thexp/edit

https://www.kaggle.com/code/diptobiswas777/random-forest/edit

https://www.kaggle.com/code/diptobiswas777/svm-exp-10/edit

https://www.kaggle.com/code/diptobiswas777/exp-lab-11i-random-forest/edit

https://www.kaggle.com/code/diptobiswas777/iml-exp-10-svm/edit

https://www.kaggle.com/code/diptobiswas777/iml-exp-11/edit

https://www.kaggle.com/code/diptobiswas777/iml-exp-10/edit

https://www.kaggle.com/code/diptobiswas777/expt3-dlvs-multi-layer-perceptron-human-activity/edit

https://www.kaggle.com/code/diptobiswas777/4-logistic-regression-and-classification-error/edit

https://www.kaggle.com/code/diptobiswas777/full-principal-component-analysis-pca-and-plots/edit

https://www.kaggle.com/code/diptobiswas777/feature-selection-fliter-method/edit

https://www.kaggle.com/code/diptobiswas777/0-963-pca-svm-logistic-regression/edit