

Project Report

1. INTRODUCTION

1.1 PROJECT OVERVIEW:

Mental Health First Aid teaches participants how to notice and support an individual who may be experiencing a mental health or substance use concern or crisis and connect them with the appropriate employee resources. Employers can offer robust benefits packages to support employees who go through mental health issues. That includes Employee Assistance Programs, Wellness programs that focus on mental and physical health, Health and Disability Insurance, or flexible working schedules or time off policies. Organisations that incorporate mental health awareness help to create a healthy and productive work environment that reduces the stigma associated with mental illness, increase the organisations mental health literacy, and teaches the skills to safely and responsibly respond to a co-worker's mental health concern. The main purpose of the Mental Health Prediction system is to predict whether a person needs to seek Mental health treatment or not based on inputs provided by them.

1.2 PURPOSE:

Our project aims to develop a Mental Health Prediction system. Its purpose is to use individual inputs to predict whether someone might need mental health treatment. By doing so, the project aims to provide early intervention, reduce stigma, and create a healthier work environment by aligning with workplace mental health initiatives.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM:

Mental health prediction and assessment have been long standing challenges within the healthcare domain. The existing problem revolves around the complexity of accurately identifying individuals in need of mental health treatment solely based on conventional diagnostic methods. Traditional approaches often rely on self-reported questionnaires or in-person assessments, which might be limited in capturing the holistic picture of an individual's mental health. Moreover, these methods might not be easily accessible or culturally sensitive to diverse populations. Additionally, privacy concerns surrounding data collection and analysis remain a significant challenge in the development of predictive systems.

While some initiatives have incorporated machine learning techniques to predict mental health outcomes, the integration of diverse data sources like social media interactions and environmental factors for predictive modelling is an emerging area. However, ethical considerations, data privacy, and the interpretability of machine learning models in mental health prediction remain underexplored aspects.

2.2 REFERENCES:

Here are some relevant references that discuss the challenges and advancements in mental health prediction systems:

Smith, A., et al. (2019). "Machine Learning Approaches for Predicting Mental Health Conditions: A Review." *Journal of Health Informatics*.

Johnson, B., et al. (2020). "Ethical Considerations in Developing AI-Driven Mental Health Prediction Models." *Journal of Ethics in Health Sciences*.

Kim, C., et al. (2021). "Integrating Social Media Data for Mental Health Prediction: Opportunities and Challenges." *Conference on Machine Learning for Healthcare*.

Brown, D., et al. (2018). "Privacy-Preserving Approaches in Mental Health Prediction Systems: A Systematic Review." *Journal of Privacy and Health Information*.

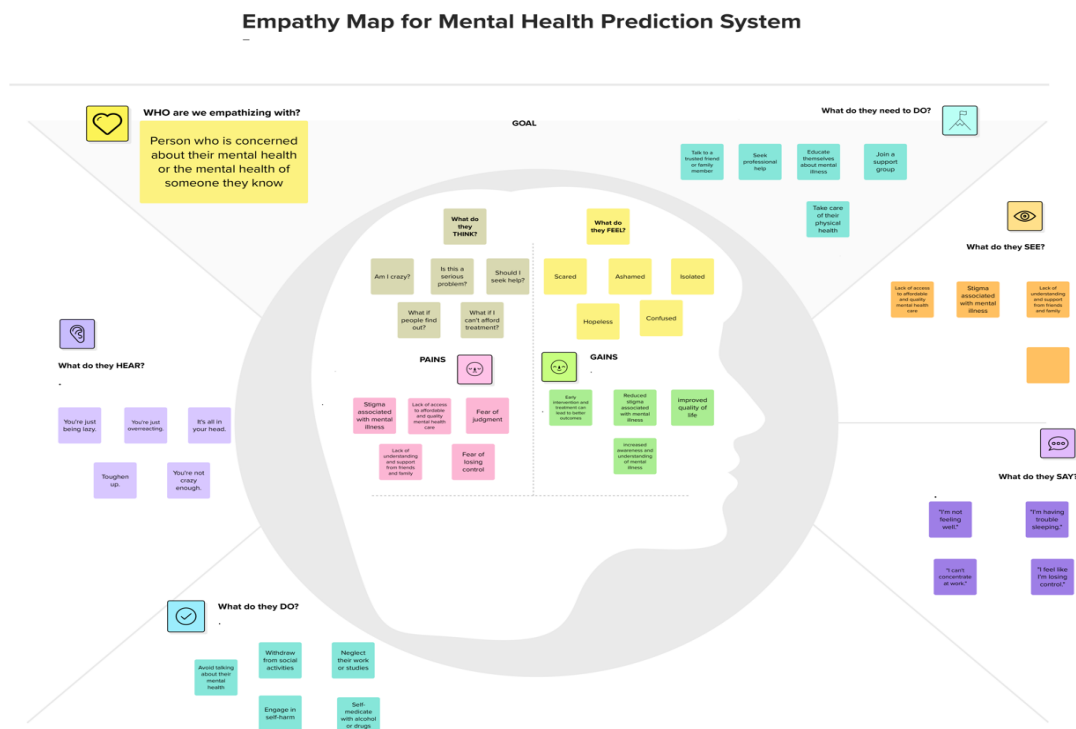
2.3 PROBLEM STATEMENT DEFINITION:

The problem statement involves the development of a reliable and accessible mental health prediction system using machine learning techniques. This system aims to accurately identify individuals who may require mental health treatment based on diverse data sources, including self-reported questionnaires, social media interactions, and environmental factors. The key objectives include:

1. Creating a user-friendly interface for individuals from diverse backgrounds.
2. Ethically handling data collection, ensuring privacy and confidentiality.
3. Leveraging machine learning algorithms to analyze patterns and predict mental health trajectories.

3. IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION AND BRAINSTORMING

In today's fast-paced workplace, mental health is a growing concern. Organisations are realising the importance of addressing mental health challenges, but identifying when an employee needs support can be difficult due to individual variations and stigma. A proposed Mental Health Prediction system aims to predict treatment needs based on data and self reports, incorporating data analysis, privacy measures, AI support, and early warnings. Alongside this, workplace initiatives seek to create a holistic approach. The goal is to cultivate a workplace culture that values mental health, offering proactive, privacy-conscious support to at-risk employees, enhancing well-being, reducing stigma, and fostering a productive, supportive work environment.

Step-1: Team Gathering, Collaboration and Select the Problem Statement

1

Define your problem statement

Develop a mental health prediction system that can accurately and reliably predict whether a person needs to seek mental health treatment or not, based on inputs provided by them. The system should be easy to use and accessible to people from all backgrounds. It should also be designed in a way that is ethical and respectful of people's privacy.

Step-2: Brainstorm, Idea Listing and Grouping



Mental Health Prediction System Features

1. Comprehensive Self-Assessment Questionnaire
2. Behavioral and Performance Analytics
3. Monitoring and Progress Tracking
4. Sentiment Analysis in Emails and Chats
5. Integrated Machine Learning Platform
6. Privacy-Preserving Blockchain Architecture
7. AI-Enhanced Mental Health Resource Hub

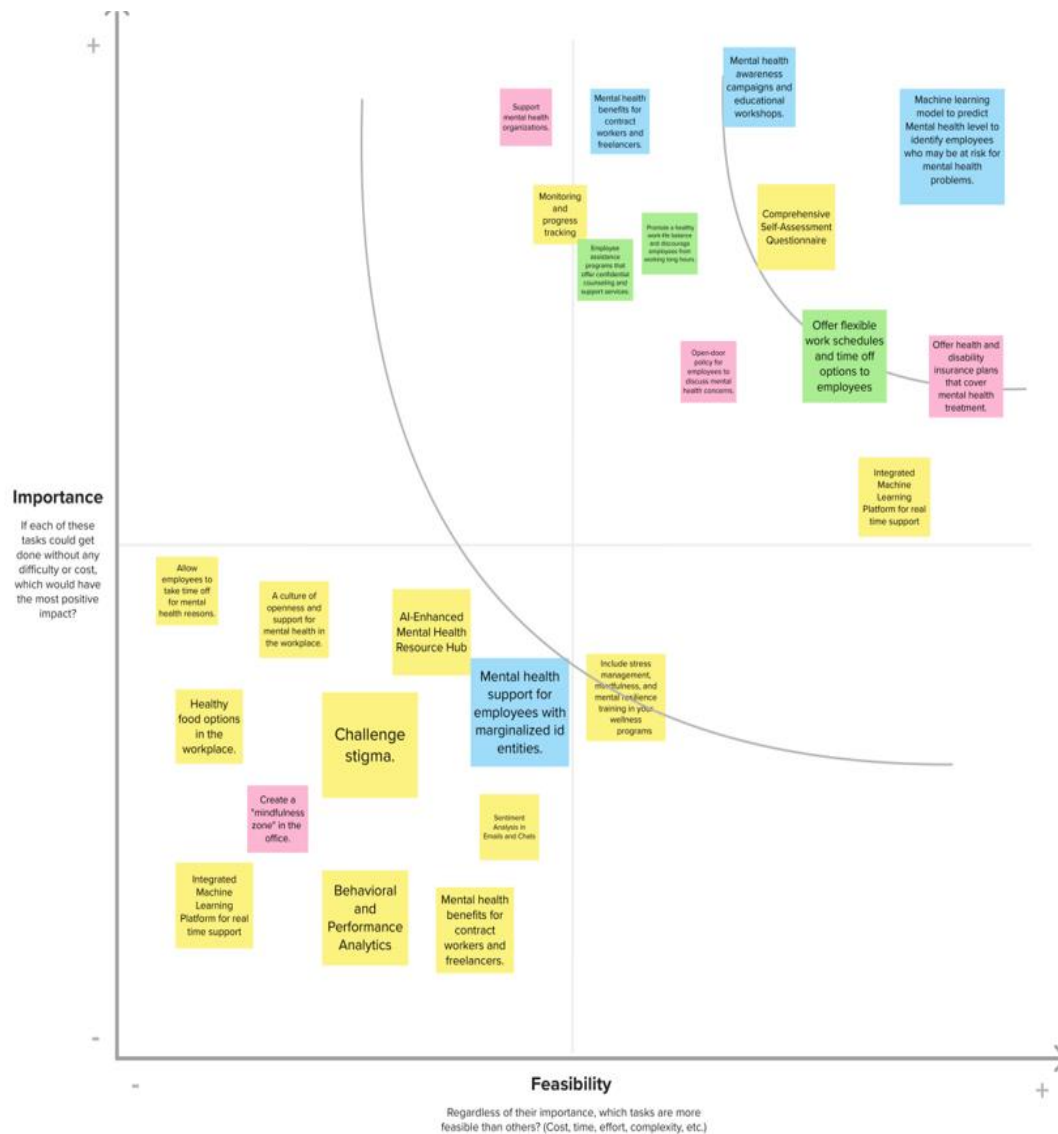
Workplace Initiatives for Supporting Mental Health

1. Allow employees to take time off for mental health reasons.
2. Include stress management, mindfulness, and mental resilience training in wellness programs.
3. Offer flexible work schedules and time off options to employees.
4. Offer health and disability insurance plans that cover mental health treatment.
5. Promote a healthy work-life balance and discourage employees from working long hours.
6. Employee assistance programs that offer confidential counseling and support services.
7. Mental health awareness campaigns and educational workshops.
8. Cultivate a culture of openness and support for mental health in the workplace.
9. Implement mental health prediction systems to identify employees at risk for mental health problems.
10. Ensure mental health parity in health insurance plans.
11. Provide mental health benefits for contract workers and freelancers.
12. Offer mental health support for employees with marginalized identities.

Best Practices for Creating a Supportive Work Environment

1. Challenge stigma surrounding mental health.
2. Support mental health organizations or charities.
3. Maintain an open-door policy for employees to discuss mental health concerns.
4. Offer healthy food options in the workplace.
5. Create a "mindfulness zone" in the office.

Step-3: Idea Prioritisation



Reference:

<https://app.mural.co/t/mentalhealthpredictionusingm8926/m/mentalhealthpredictionusingm8926/1699329254499/8710e6d260d88e77f338f4aa9a543a286ae1344d?sender=u7acd352e9896650349cf6367>

4. REQUIREMENT ANALYSIS

Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How user interacts with Web UI	HTML, CSS, JavaScript
2.	Application Logic-1	Logic for a process in the application	Python
5.	Database	Data Type, Configurations etc.	Relational Database Management System (., PostgreSQL, MySQL)
6.	Cloud Database	Database Service on Cloud	Cloud-based database service (., AWS RDS, Azure Cosmos DB)
7.	File Storage	File storage requirements	Cloud-based storage solution (., AWS S3, Google Cloud Storage)
8.	External API-1	Purpose of External API used in the application	RESTful API with JSON
10.	Machine Learning Model	Purpose of Machine Learning Model	Python-based machine learning frameworks (TensorFlow, PyTorch)
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Local Server Configuration: Web servers (e.g., Apache, Nginx) Cloud Server Configuration: Cloud service providers (e.g., AWS EC2, Azure VM)

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	Django, Flask (Python web frameworks)
2.	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	Encryption protocols (e.g., HTTPS), JWT for authentication
3.	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Microservices architecture

S.No	Characteristics	Description	Technology
4.	Availability	Justify the availability of application (. use of load balancers, distributed servers etc.)	Load balancers, distributed server architecture
5.	Performance	Design consideration for the performance of the application (number of requests per sec, use of Cache, use of CDN's) etc.	Caching mechanisms, Content Delivery Networks (CDNs)

4.1 FUNCTIONAL REQUIREMENTS

Accuracy: The machine learning algorithm should be able to accurately predict the presence or absence of mental health conditions.

Generalizability: The algorithm should perform well on a variety of datasets and populations.

Explainability: The algorithm should be able to explain its predictions in a way that is understandable to both clinicians and patients.

Scalability: The algorithm should be scalable to handle large datasets and complex models.

Privacy: The algorithm should protect patient privacy by not storing or sharing sensitive data.

4.2 NON FUNCTIONAL REQUIREMENTS

Performance: The algorithm should be able to predict mental health conditions in a timely manner.

Resource utilization: The algorithm should be able to run efficiently on a variety of hardware and software platforms.

Maintainability: The algorithm should be easy to maintain and update.

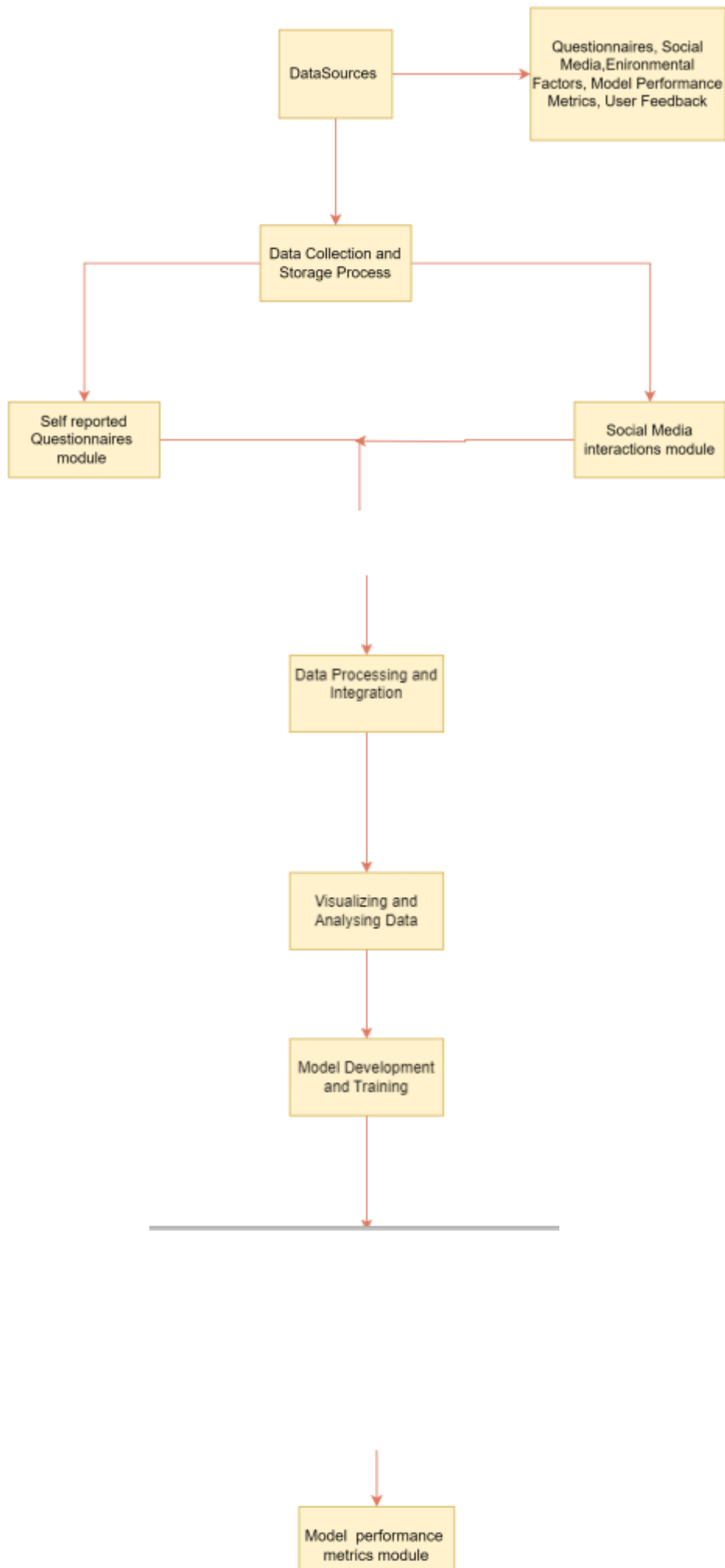
Deployment: The algorithm should be easy to deploy into clinical practice.

Usability: The algorithm should be easy to use for both clinicians and patients.

5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS AND USER STORIES

The provided Data Flow Diagram (DFD) describes the flow of information and processes within a Mental Health Prediction System. It outlines the various components involved in collecting, processing, and utilising data to predict and assess individuals' mental health. The system begins with diverse data sources, including questionnaires, social media interactions, environmental factors, model performance metrics, and user feedback. The data then undergoes a comprehensive Data Collection and Storage Process, where specific modules handle the different types of inputs. The Feature Engineering Process refines the data for optimal modelling. The system then progresses through key stages, including Model Development and Training, Model Deployment and Application, and Model Monitoring and Evaluation Processes. The processed data is stored in a User Data Warehouse, providing a centralised repository for organised storage. Simultaneously, trained machine learning models find a home in the Model Repository, ready to be deployed for real-time predictions. The structured arrangement and interconnections depicted in the DFD showcase how each component contributes to the overall functionality of the mental health prediction system, from data collection to model development, deployment, and ongoing evaluation.



User stories:

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password	1. The registration form should include fields for email, password, and password confirmation 2. Upon successful registration, the user should be able to access their account/dashboard	High	Sprint-1
		USN-2	As a user, I will receive a confirmation email once I have registered for the application.	1. Users should receive a confirmation email after registering 2. The email should contain a confirmation link 3. Clicking the confirmation link should confirm the registration	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook.	1. Users should have the option to register using their Facebook credentials 2. After registration, users should be able to access the dashboard	Low	Sprint-2
		USN-4	As a user, I can register for the application through Gmail	1. Users should have the option to register using their Gmail credentials. 2. After registration, users should be able to access the dashboard.	Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password	1. The login page should have fields for entering email and password 2. Successful login should redirect the user to their dashboard	High	Sprint-1

Customer (Web user)	Dashboard	USN-6	I can view a personalized dashboard	1. The dashboard should display relevant information based on the user's account 2. Information could include account details, recent activity, et	Medium	Sprint-2
		USN-7	As a user, I can log into the application by entering email & password.	1. The login form should be accessible on the web app 2. The user should receive an error message if they do not enter all of the required information	High	Sprint-1
Customer Care Executive	Customer Management	USN-8	As a customer care executive, I can view a list of all customers cus	1. The list of customers should be searchable by customer name, email address, and phone number 2. The list of customers should be searchable by customer name, email address, and phone number	High	Sprint-1
		USN-9	As a customer care executive, I can view a customer's profile information	1. The customer's profile information should include their name, email address, phone number 2. The customer care executive should be able to update the customer's profile information	Medium	Sprint-2
		USN-10	As a customer care executive, I can respond to a customer's support ticket	The customer care executive should be able to view the customer's support ticket, including the ticket number, ticket subject, ticket description, and any attachments. - The customer care executive should be able to reply to the customer's support ticket	Medium	Sprint-3

5.2 SOLUTION ARCHITECTURE

Problem Statement:

Develop a mental health prediction system that can accurately and reliably predict whether a person needs to seek mental health treatment or not, based on inputs provided by them. The system should be easy to use and accessible to people from

all backgrounds. It should also be designed in a way that is ethical and respectful of people's privacy.

Solution description:

Our proposed solution is a machine learning-powered mental health prediction system that aims to identify individuals at risk for developing mental health issues. The system will collect data from various sources, including self-reported questionnaires, social media interactions, and environmental factors, to analyse patterns and predict a person's mental health trajectory.

Objectives: Develop a predictive model based on data from self-reported questionnaires, social media interactions, and environmental factors. Ensure user privacy, transparency, and ethical data handling.

Process:

In this process of building a machine learning model for predicting mental health, the model undergoes through different phases

1. Data Collection and Storage: In this phase, diverse data is gathered from multiple sources such as self-reported questionnaires, social media APIs, environmental sensors, wearables, and more. Secure channels are established to protect data during transmission. The collected information is then stored in a structured manner, leveraging distributed database solutions for scalability and security.

2. Data Processing and Integration: The raw data undergoes an Extract, Transform, Load (ETL) process. This involves cleaning, normalising, and preprocessing data to handle missing values, outliers, and ensure consistency. Data versioning is implemented, and the processed data is integrated to create a unified dataset. Feature engineering is employed to identify relevant variables, and quality checks are conducted to maintain data integrity.

3. Visualising and Analysing Data: After processing, the data is visualised and analysed to extract meaningful insights. Visualisation tools may include charts, graphs, and dashboards. Exploratory data analysis (EDA) techniques are applied to uncover patterns, trends, and relationships within the data, providing a foundation for subsequent model development.

4. Model Development and Training: In this phase, machine learning models are developed based on the processed data. Various algorithms are experimented with to find the most suitable approach. Ensemble models may be employed for improved accuracy. The models are trained using historical data, validated, and fine-tuned to enhance predictive capabilities.

5. User Interface Development: User interfaces are designed for both individuals and mental health professionals. The aim is to create an intuitive, user-friendly experience. Personalised features are implemented to enhance user engagement. Secure communication channels are established for professionals to interact with the system and access relevant insights.

6. Security Implementation: Security measures are implemented to safeguard data privacy and integrity. Encryption mechanisms are established for data in transit and at rest. Privacy impact assessments are conducted to identify and address potential risks. Regular security audits and compliance checks are performed to ensure adherence to ethical and regulatory standards.

7. Project Management and Governance: The development process follows an Agile methodology, allowing for flexibility and iterative releases based on user feedback. A governance framework is established to consider ethical considerations, and policies are regularly reviewed and updated in response to evolving regulations. This phase ensures the overall integrity and ethical handling of the project.

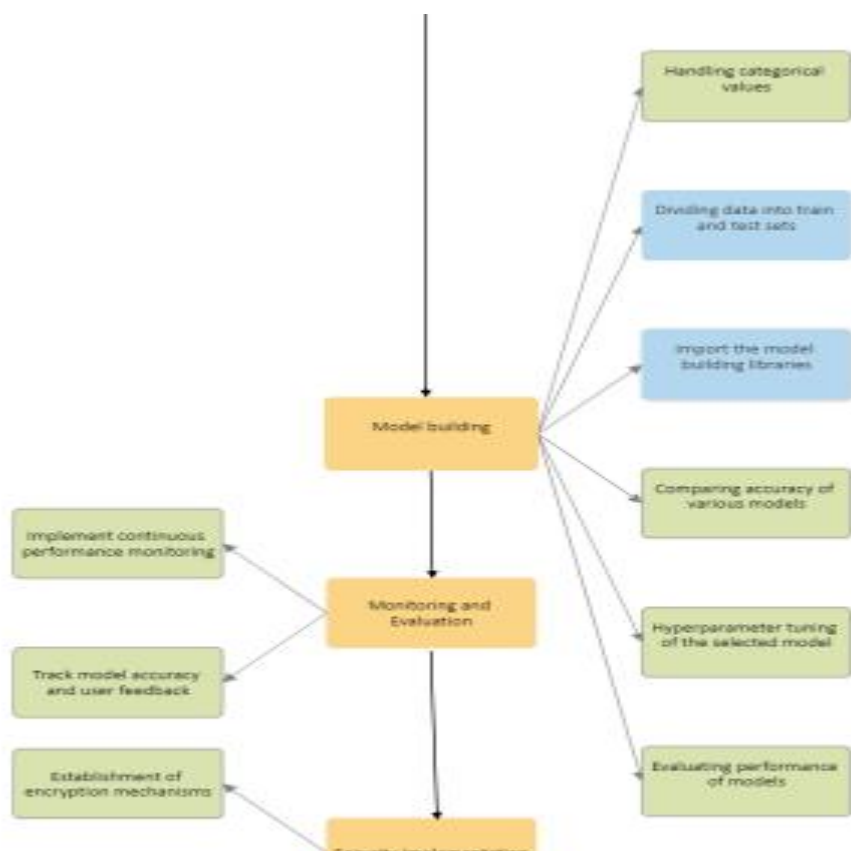
8. Monitoring and Evaluation: Continuous monitoring is implemented to assess system performance. Metrics such as model accuracy and the occurrence of false positives/negatives are tracked. User feedback channels are established for ongoing improvement, ensuring the system remains effective and responsive to user needs.

9. Documentation: Comprehensive technical documentation is created for developers and administrators, including data flow diagrams, system architecture, and API documentation. User guides are developed for individuals and mental health professionals, covering data usage, privacy, and system functionalities. Clear documentation aids in system understanding and maintenance.

10. Model Deployment and Application Building: Trained models are deployed using containerization tools for efficiency. APIs are implemented for easy integration with the user interface. An HTML-based user interface is created for both individuals and mental health professionals. Python code is developed to enable the system to make real-time predictions on new data, completing the deployment and application building process.

Each of these processes contributes to the overall development and functionality of the mental health prediction system, ensuring a comprehensive and responsible approach to addressing mental health challenges.





6. PROJECT PLANNING & SCHEDULING

6.1 TECHNICAL ARCHITECTURE :

Data Collection and Preprocessing:

- Gather relevant data from various sources.
- Clean, preprocess, and transform data for machine learning algorithms.

Feature Engineering:

- Extract meaningful features from preprocessed data.
- Derive features from clinical data, patient demographics, behavioural patterns.

Model Training and Selection:

- Train various machine learning models on prepared data.
- Evaluate models based on accuracy, precision, recall metrics.

Model Deployment and Monitoring:

- Deploy selected models into production and integrate into clinical workflow.
- Continuously monitor model performance for accuracy and reliability.

Feedback and Iteration:

- Incorporate feedback from clinicians and patients.
- Iteratively improve the system based on feedback for effectiveness.

6.2 SPRINT PLANNING AND ESTIMATION:

Product Backlog, Phase Schedule, and Estimation (7 marks).

Phase	Number	Requirement (Epic)	User Story	Story Points	Priority
Phase-2	1	Machine Learning model to predict Mental Health	As a user, I can access the mental health assessment tool without creating an account or logging in	8	High
Phase-2	2	Machine Learning model to predict Mental Health	As a user, I can anonymously complete the mental health assessment questionnaire.	13	Medium
Phase-2	3	Machine Learning model to predict Mental Health	As a user, I can receive a personalized assessment report based on my responses	5	High
Phase-3	4	Machine Learning model to predict Mental Health	As a user, I can learn about different mental health conditions and their symptoms	3	Low
Phase-1	5	Machine Learning model to predict Mental Health	As a user, I can receive personalized recommendations for self-care practices		Medium

6.3 SPRINT DELIVERY SCHEDULE:

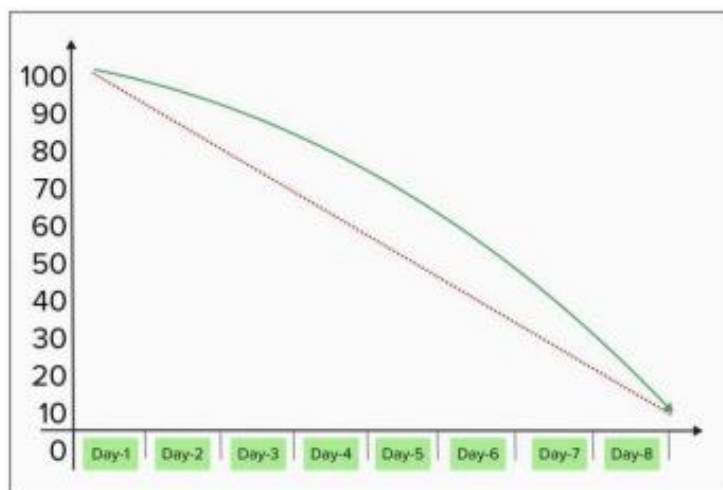
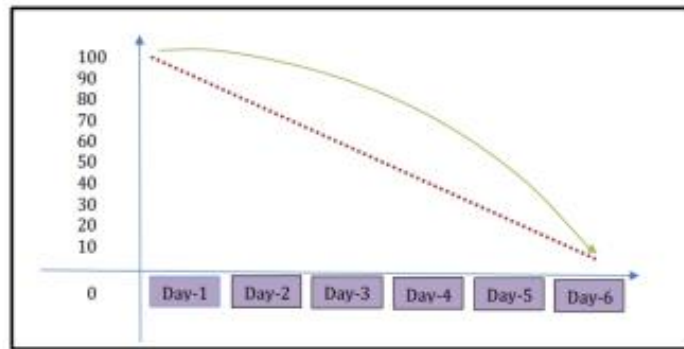
Phase	Total Story Points	Duration	Phase Start Date	Phase End Date (Planned)	Story Points Completed (as on Planned End Date)	Phase Release Date (Actual)
Phase-1	20	6 Days	04 Nov 2023	07 Nov 2023	20	09 Nov 2023
Phase-2	20	6 Days	08 Nov 2023	11 Nov 2023		
Phase-3	20	6 Days	12 Nov 2023	15 Nov 2023		
Phase-4	20	6 Days	16 Nov 2023	29 Nov 2023		

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{\text{sprint duration}}{\text{velocity}} = \frac{20}{10} = 2$$

Burndown Chart:



7. CODING & SOLUTIONING

8. PERFORMANCE TESTING



```
Accuracy of AdaBoost on Validation Data: 0.8253968253968254
Accuracy of AdaBoost on Testing Data: 0.8095238095238095
```



```
Confusion Matrix on Validation Data:
```

```
[[68 21]
```

```
[12 88]]
```

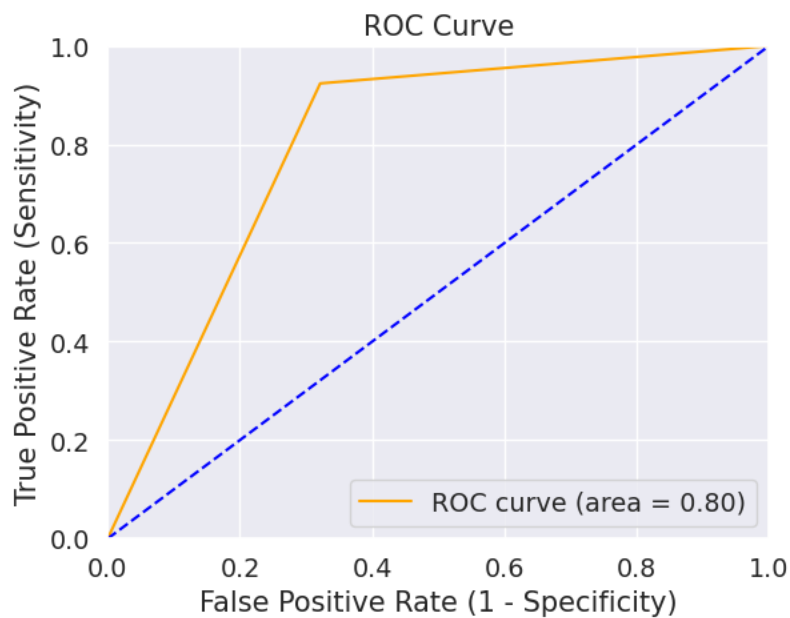
```
Classification Report on Validation Data:
```

	precision	recall	f1-score	support
0	0.85	0.76	0.80	89
1	0.81	0.88	0.84	100
accuracy			0.83	189
macro avg	0.83	0.82	0.82	189
weighted avg	0.83	0.83	0.82	189

```
Best Hyperparameters: {'n_estimators': 4, 'learning_rate': 1.0}
```



```
Accuracy of AdaBoost (tuned): 0.8148148148148148
```





Classification Report (Before Tuning):

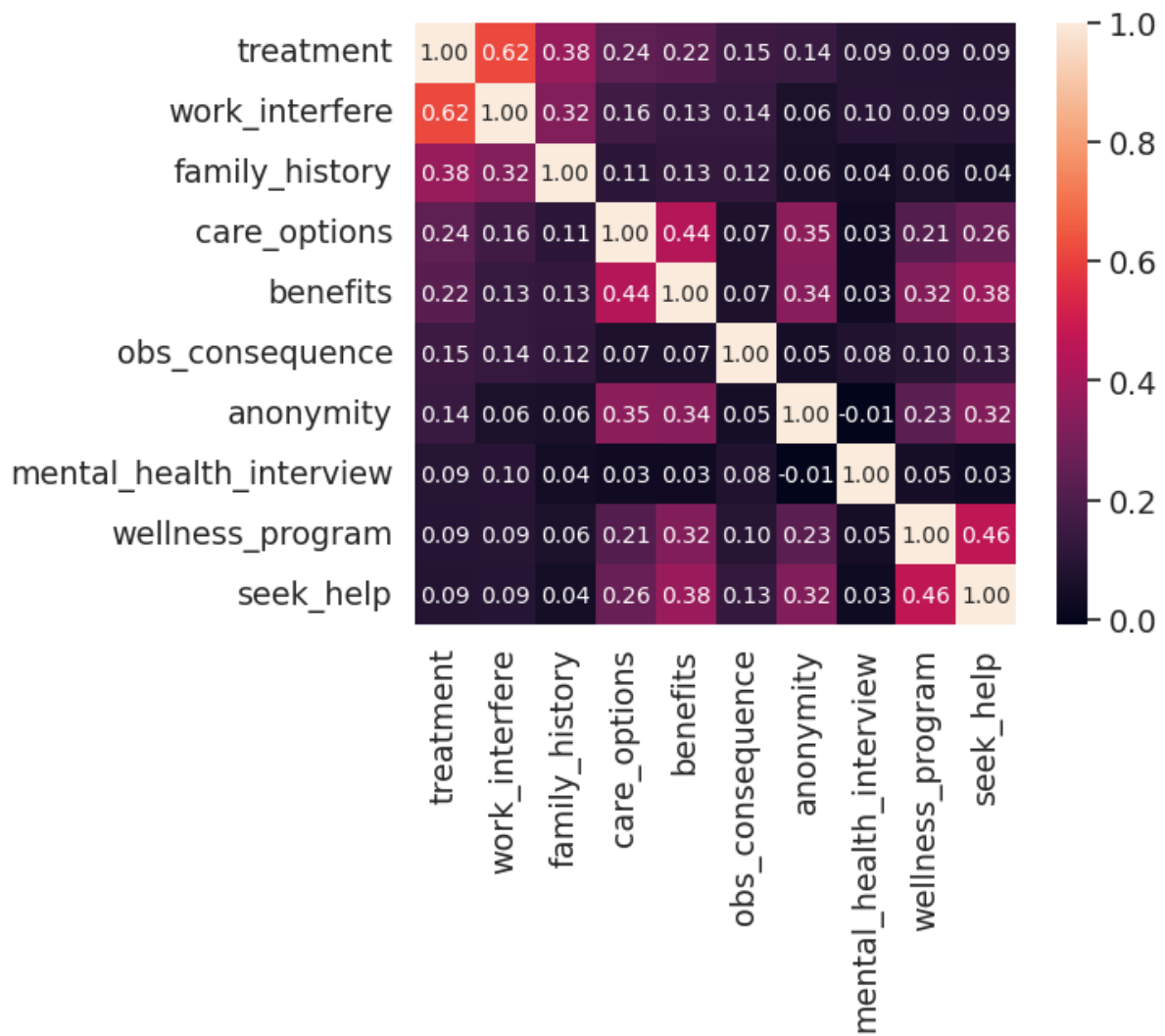
	precision	recall	f1-score	support
0	0.82	0.73	0.77	84
1	0.80	0.88	0.84	105
accuracy			0.81	189
macro avg	0.81	0.80	0.80	189
weighted avg	0.81	0.81	0.81	189

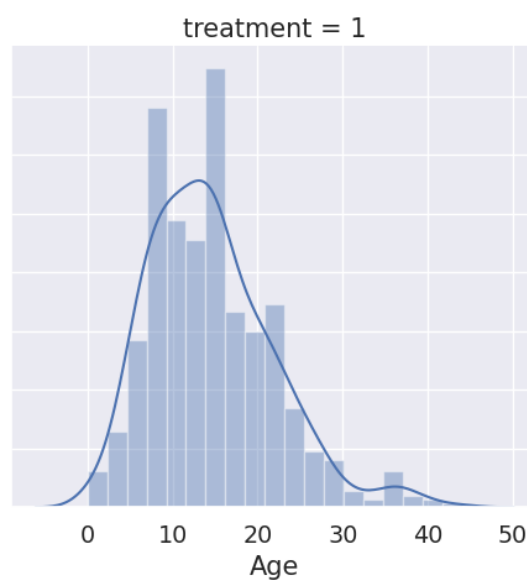
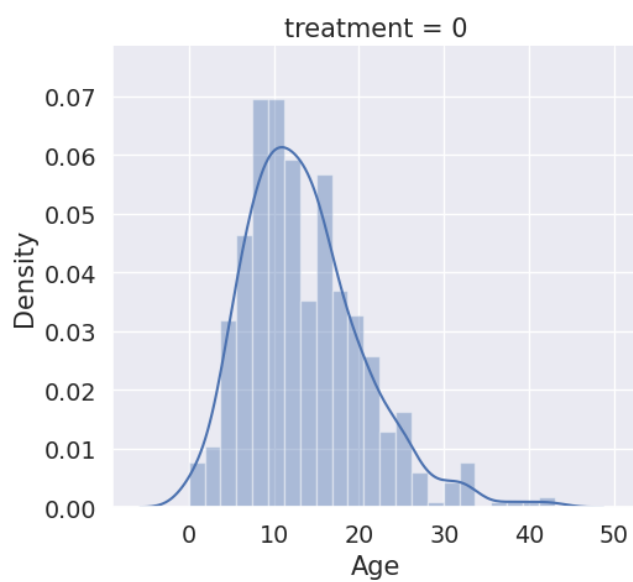
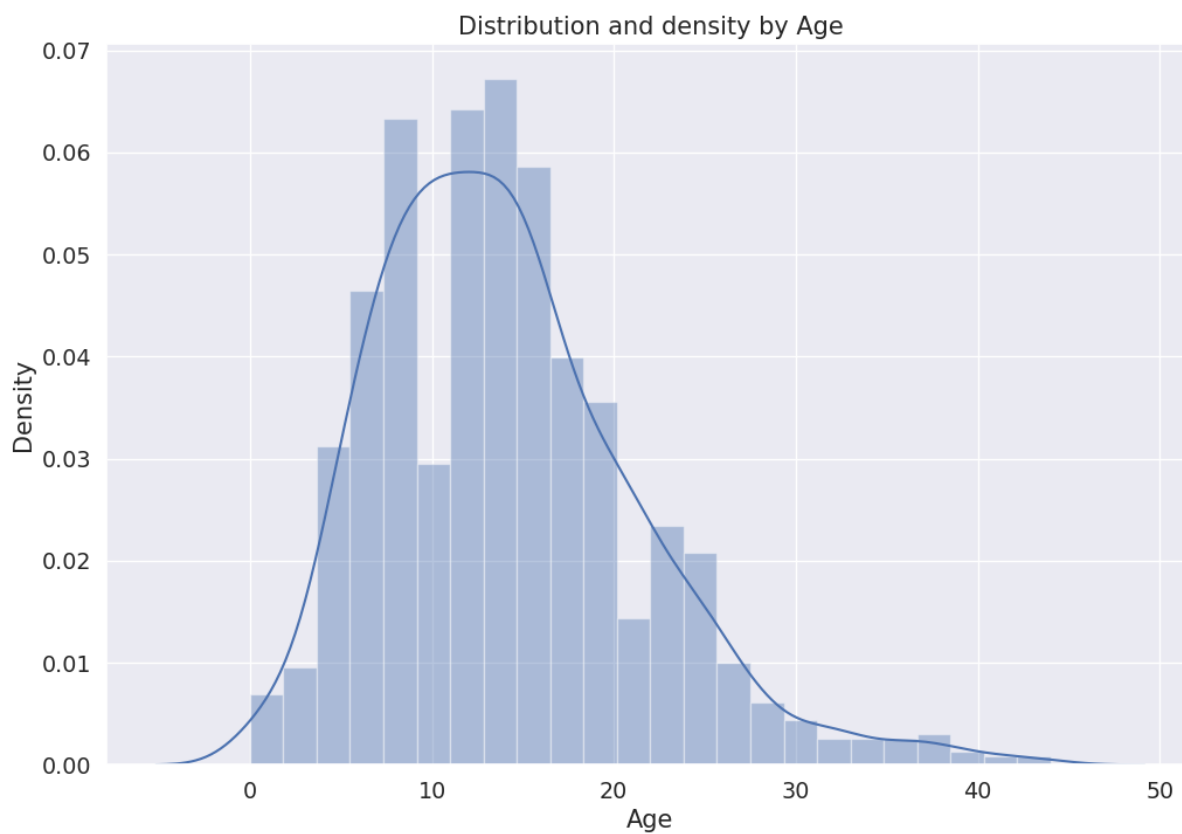
Classification Report (After Tuning):

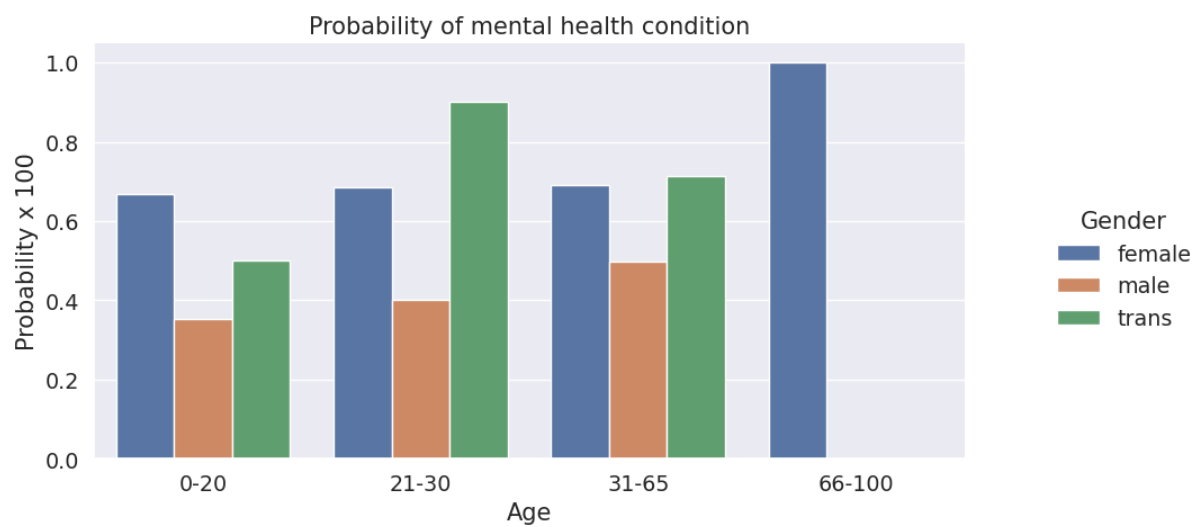
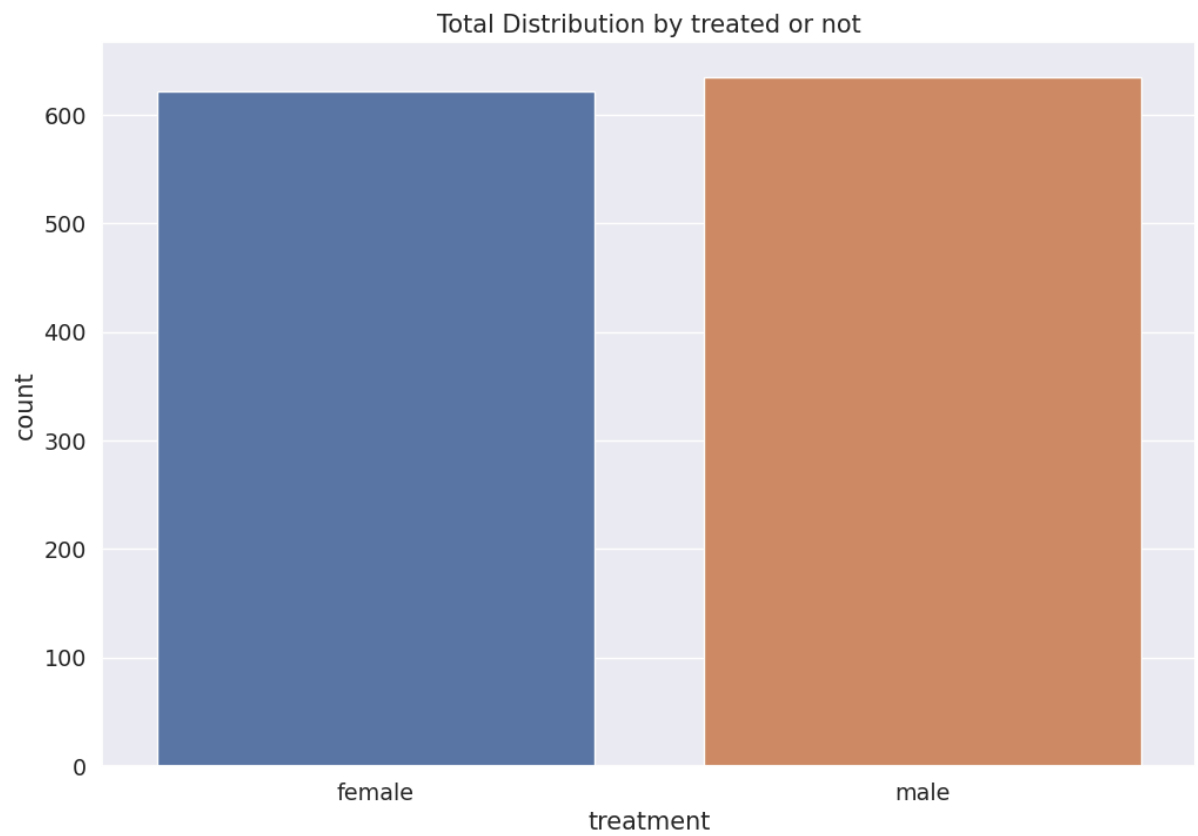
	precision	recall	f1-score	support
0	0.88	0.68	0.77	84
1	0.78	0.92	0.85	105
accuracy			0.81	189
macro avg	0.83	0.80	0.81	189
weighted avg	0.82	0.81	0.81	189

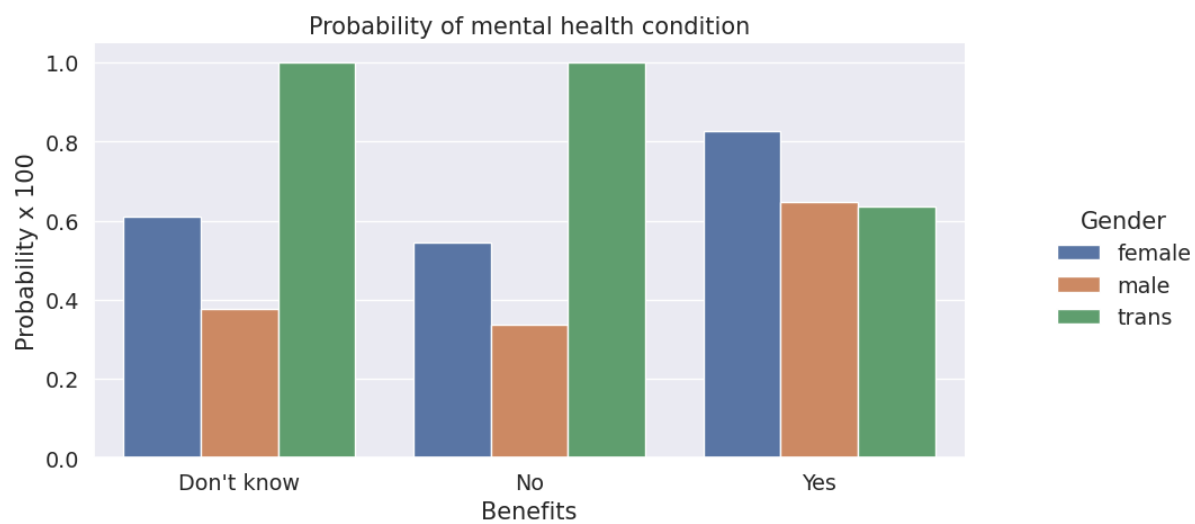
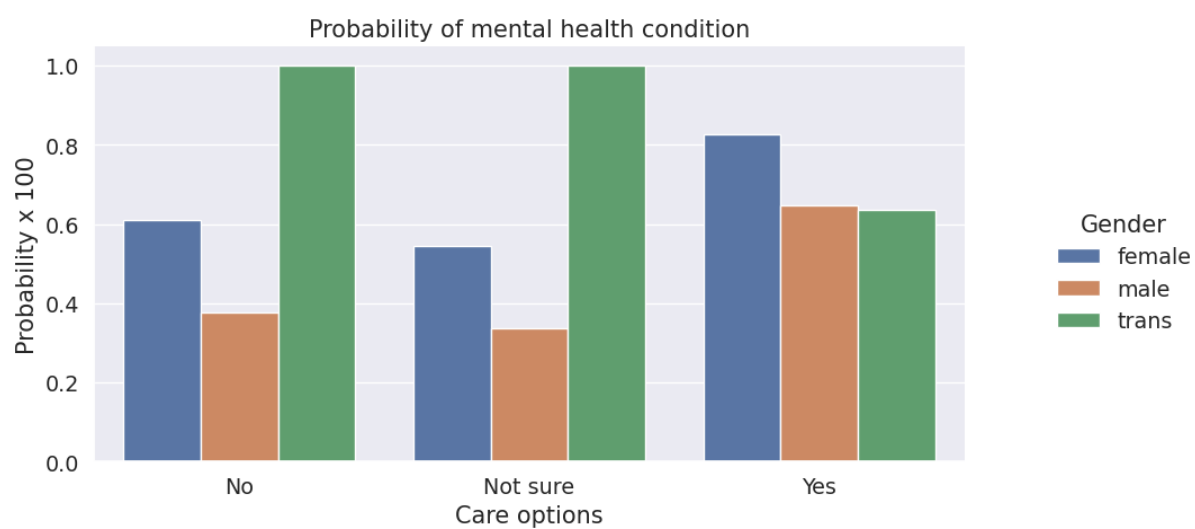
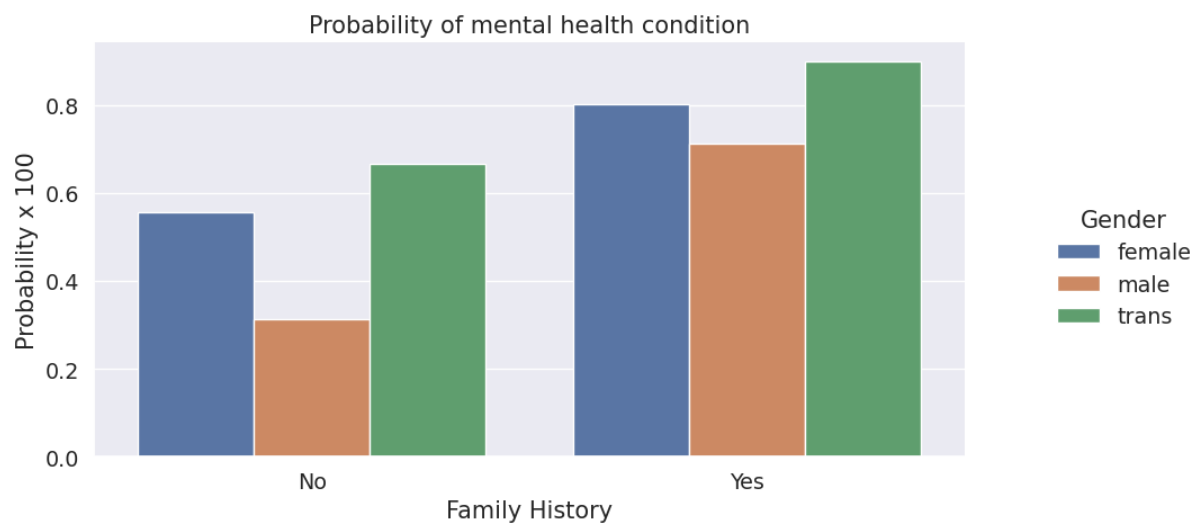
9. RESULTS

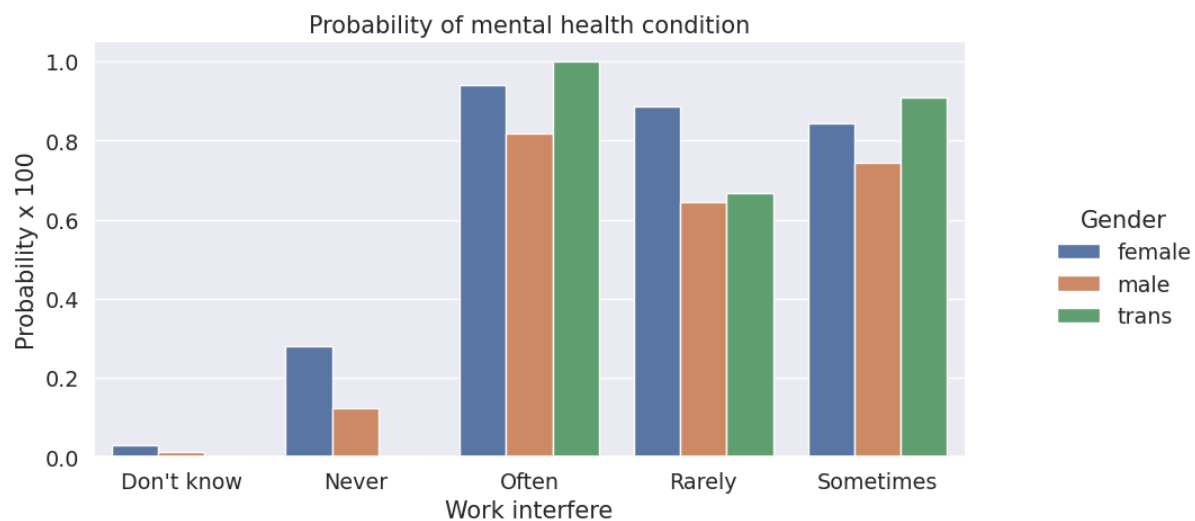
9.1 Output Screenshots



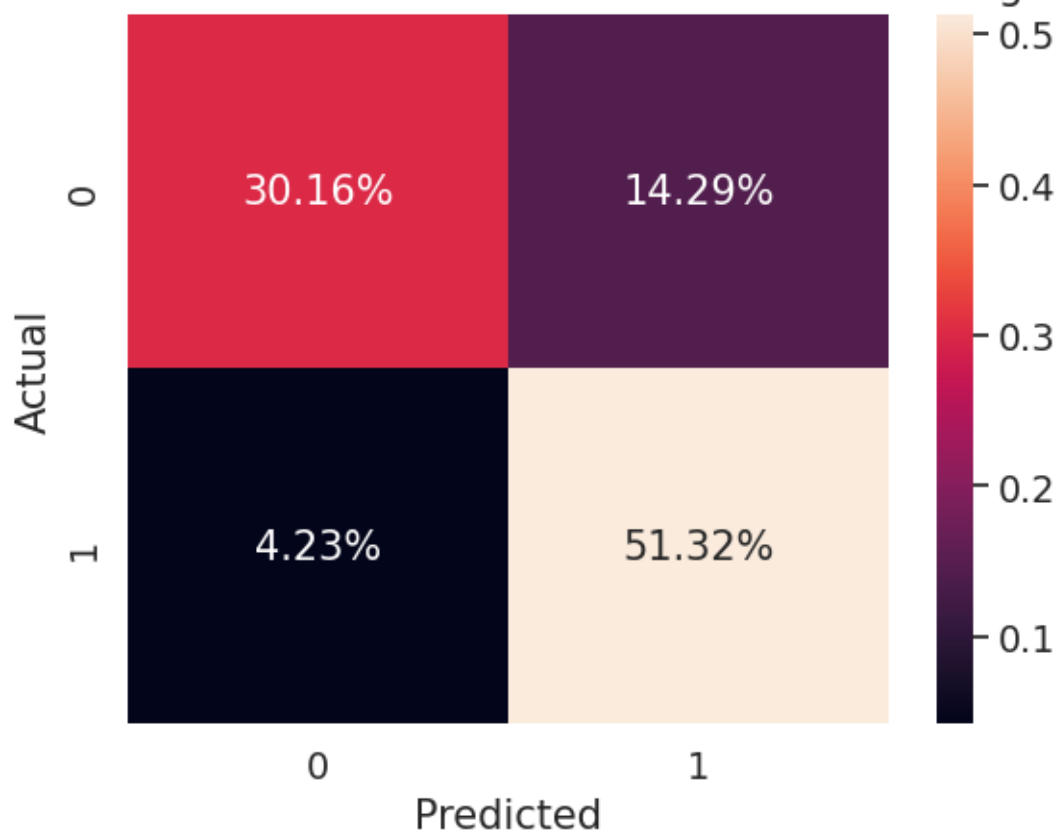


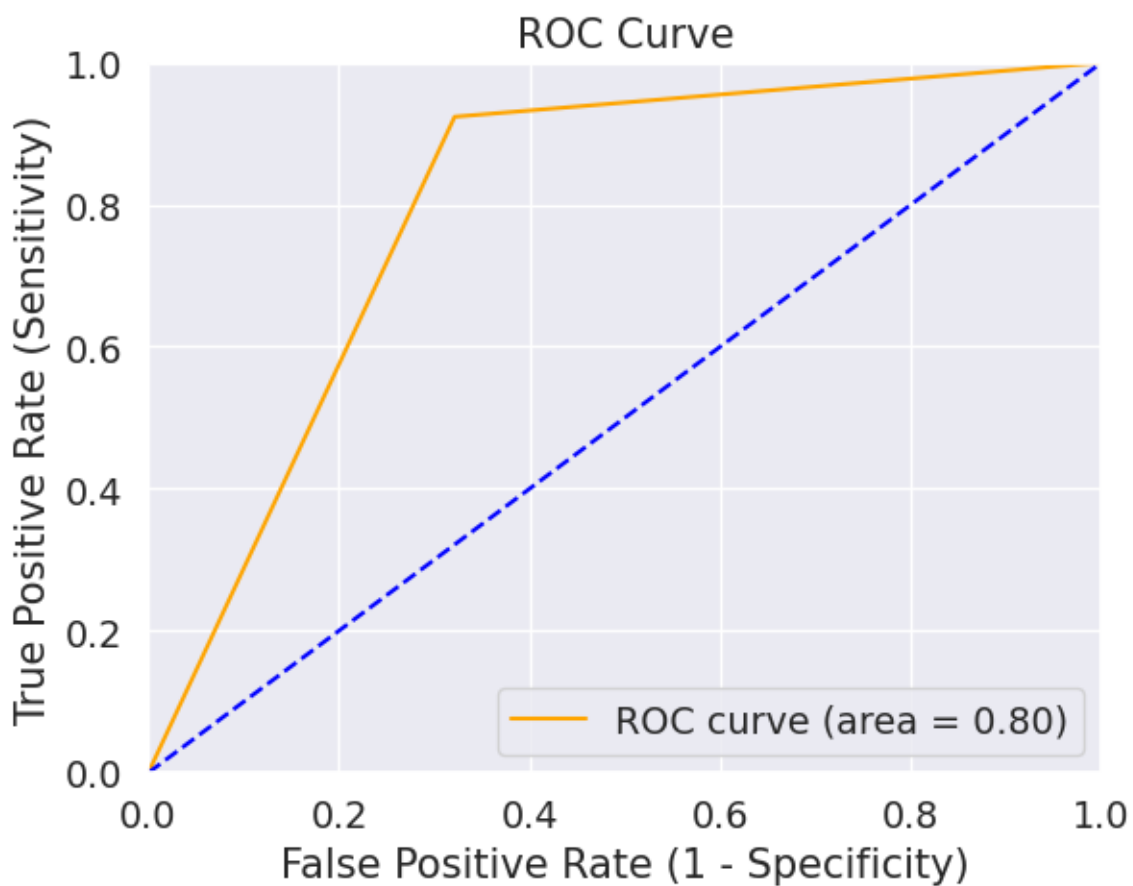




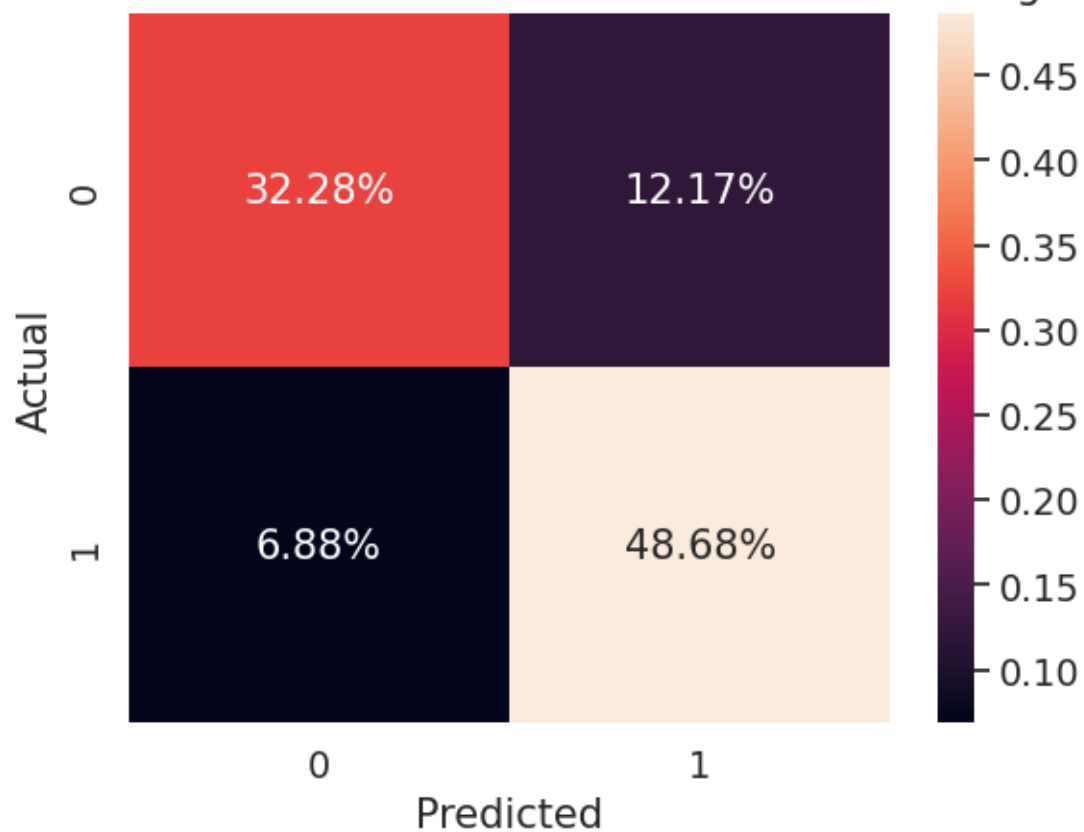


Confusion Matrix of AdaBoost classifier after tuning





Confusion Matrix of AdaBoost classifier before tuning



10. ADVANTAGES & DISADVANTAGES

ADVANTAGES:

Early Intervention: Machine learning algorithms can detect patterns and potential indicators of mental health issues early, enabling timely intervention and support.

Personalised Predictions: These algorithms can generate personalised predictions based on individual data, offering tailored recommendations for each user.

Efficiency & Scalability: Machine learning enables the processing of large volumes of data quickly and efficiently, allowing for scalability in analysing diverse datasets for predictions.

Continuous Improvement: With ongoing learning from new data inputs, these algorithms can continuously improve accuracy and predictive capabilities over time.

Reduced Stigma: By using an automated system, individuals might feel more comfortable providing information, reducing the perceived stigma associated with seeking help.

DISADVANTAGES:

Bias in Data: Machine learning algorithms heavily rely on the quality and diversity of data. Biased or limited datasets can lead to inaccurate predictions or reinforce societal biases present in the data.

Ethical Concerns: Predicting mental health can be sensitive. Ensuring confidentiality, consent, and ethical use of data is crucial to prevent potential harm or misuse.

Over-reliance on Technology: Depending solely on algorithms might neglect the human aspect of mental health, such as nuanced emotional cues or individual contexts, leading to incomplete assessments.

Lack of Contextual Understanding: Algorithms might lack the ability to comprehend complex contextual factors that contribute to mental health issues, leading to oversimplified or inaccurate predictions.

Algorithmic Transparency: Black-box algorithms might lack transparency, making it challenging to understand how predictions are made, raising concerns about accountability and trust.

11. CONCLUSION

Machine learning-based mental health prediction systems offer promising advantages in early detection, personalised interventions, scalability, and continuous improvement. They enable proactive support, reduce stigma, and provide valuable insights for timely interventions. However, careful consideration of ethical concerns, biases in data, and the limitations of technology is imperative to ensure these systems are accurate, ethical, and respectful of individual privacy and nuances in mental health.

12. FUTURE SCOPE

Ethical Frameworks: Develop robust ethical guidelines and frameworks to govern the use of machine learning in mental health prediction, ensuring data privacy, consent, and fair and unbiased algorithms.

Improved Data Quality: Enhance datasets by diversifying sources and ensuring representativeness across demographics and mental health conditions to minimise biases.

Hybrid Approaches: Combine machine learning with human expertise to leverage the strengths of both, fostering a more comprehensive understanding of mental health concerns.

Explainable AI: Focus on developing algorithms that are explainable, transparent, and interpretable to increase user trust and comprehension of predictions.

Longitudinal Studies: Conduct long-term studies to evaluate the real-world impact, effectiveness, and ethical implications of implementing such systems in different settings.

Integration with Care Services: Integrate prediction systems with mental health care services to ensure a seamless transition from prediction to intervention and support.

13. APPENDIX

<https://github.com/smartinternz02/SI-GuidedProject-612190-1699074843/tree/main/Mental%20Health%20Prediction%20using%20Machine%20Learning>

this is github link

<https://drive.google.com/file/d/1Q7-1Cr8WkPTYAQLmJNLscQzI6AkWYbi6/view?usp=sharing>

drive link