**Project Report Documentation**

|  |  |
| --- | --- |
| Date | 9 November 2023 |
| Team ID | 593151 |
| Project Name | Predicting Mental Health Illness Of Working Professionals Using Machine Learning |

1. **INTRODUCTION**
   1. Project Overview
   2. Purpose
2. **LITERATURE SURVEY**
   1. Existing problem
   2. References
   3. Problem Statement Definition
3. **IDEATION & PROPOSED SOLUTION**
   1. Empathy Map Canvas
   2. Ideation & Brainstorming
4. **REQUIREMENT ANALYSIS**
   1. Functional requirement
   2. Non-Functional requirements
5. **PROJECT DESIGN**
   1. Data Flow Diagrams & User Stories
   2. Solution Architecture
6. **PROJECT PLANNING & SCHEDULING**
   1. Technical Architecture
   2. Sprint Planning & Estimation
   3. Sprint Delivery Schedule
7. **CODING & SOLUTIONING (Explain the features added in the project along with code)**
   1. Feature 1
   2. Feature 2
8. **PERFORMANCE TESTING**
   1. Performace Metrics
9. **RESULTS**
   1. Output Screenshots
10. **ADVANTAGES & DISADVANTAGES**
11. **CONCLUSION**
12. **FUTURE SCOPE**

**13.APPENDIX**

Source Code

GitHub & Project Demo Link

**1.INTRODUCTION;**

**1.1 project overview:**

Predicting mental health is one of the most important aspects of lowering the likelihood of severe mental disease. In the meanwhile, the public health department might use mental health prediction as a theoretical foundation to develop psychological intervention strategies for healthcare professionals.

The project's goal is to create a machine learning model that, given specific input features, can forecast mental health issues. A dataset comprising details on people's demographics, lifestyle choices, and psychological health will be used to train the algorithm. The program will be able to forecast a person's mental health state by examining these features. This entails gathering and examining information about the different aspects of mental health in the workplace. The goal is to develop a tool that would enable early detection of any mental health problems and prompt assistance and intervention.

**1.2 purpose:**

The purpose of this project is to provide a tool that can assist in the early detection and prediction of mental health conditions. By leveraging machine learning techniques, the model can analyze various factors and identify patterns that might contribute to mental health issues. This can help healthcare professionals and individuals in making informed decisions regarding mental health interventions and treatments.

The project aims to benefit both individuals and the healthcare community by:

**Early Detection**: By predicting mental health conditions, the model can aid in the early detection of potential problems, allowing individuals to seek help and support at an earlier stage.

**Personalized Interventions:** The model's predictions can enable healthcare professionals to provide personalized interventions based on an individual's risk factors, improving the effectiveness of treatments and support services.

**Resource Allocation:** By identifying individuals at a higher risk of mental health conditions, resources can be allocated more efficiently, ensuring that those who need support the most receive it in a timely manner.

**Public Health Planning:** Aggregated data from the model can be used to identify trends and patterns in mental health conditions, helping policymakers and researchers in planning public health strategies and interventions.

**Customized Support:** Tailoring support systems based on individual needs is crucial. The project aims to provide personalized recommendations and resources to help professionals manage stress, maintain a healthy work-life balance, and address specific challenges they may be facing.

**Organizational Impact:** Improved mental health among employees can lead to increased productivity, reduced absenteeism, and a healthier workplace culture. The project aims to provide insights that organizations can use to create a more supportive and conducive work environment.

**Destigmatizing Mental Health:** By incorporating mental health prediction into the workplace, the project also seeks to contribute to destigmatizing mental health issues. Normalizing conversations around mental well-being can encourage individuals to seek help without fear of judgment.

**Data-Driven Decision Making:** Utilizing data analytics and machine learning, the project aims to provide valuable insights for employers and professionals alike. This data-driven approach can inform decision-making processes related to employee well-being initiatives and resource allocation.

**2.LITERATURE SURVEY**

**2.1 Existing problem**

Predicting mental health issues using machine learning in the context of working professionals faces several challenges:

**1. \*Data Privacy Concerns:\***

**-** Gathering sensitive mental health data raises significant privacy concerns. Employees may be hesitant to share personal information, and ensuring the secure handling of such data becomes crucial.

**2. \*Subjectivity and Diversity:\***

- Mental health is subjective and varies widely among individuals. Creating a universal model that accommodates diverse experiences and cultural nuances is challenging.

**3. \*Dynamic Nature of Mental Health:\***

- Mental health conditions can be dynamic, with fluctuations over time. A static model may struggle to capture the evolving nature of an individual's mental well-being.

**4. \*Bias in Data and Models:\***

**-** If the training data used for the model is biased, the predictions can also be biased. Existing biases in workplace practices, such as promotion patterns or task assignments, may inadvertently influence the model's predictions**.**

**5. \*Lack of Standardized Metrics:\***

- There is no universal metric for measuring mental health. Different individuals may exhibit varying symptoms for the same condition, making it challenging to establish standardized evaluation criteria.

**6. \*Ethical Implications:\***

- Using predictive models for mental health raises ethical questions. Determining who has access to the predictions, how they are used, and the potential consequences for employees are critical considerations.

**7. \*Limited Understanding of Mental Health Factors:\***

**-** While machine learning models can analyze data, understanding the complex interplay of psychological, social, and environmental factors contributing to mental health remains a significant challenge.

**8. \*Employee Trust and Acceptance:\***

**-** Employees may resist being monitored for mental health, fearing potential repercussions. Building trust and ensuring transparency about the purpose and use of the predictive model is crucial.

**9. \*Unforeseen External Factors:\***

- External events, such as personal life changes or global crises, can significantly impact mental health. Predictive models may struggle to account for these unforeseen variables.

**10. \*Integration with Human Support:\***

- Machine learning models should complement, not replace, human support systems. Ensuring that the predictions are used to enhance, rather than replace, the role of mental health professionals is essential.

Addressing these challenges requires a thoughtful and multidisciplinary approach, involving not only data scientists but also mental health professionals, ethicists, and workplace experts. It's crucial to strike a balance between technological innovation and ethical, human-centered practices to create a meaningful and effective mental health prediction system.

**2.2 References**

* Reddy, U. S., Thota, A. V., & Dharun, A. (2018, December). Machine learning techniques for stress prediction in working employees. In *2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (pp. 1-4). IEEE.
* Katarya, R., & Maan, S. (2020, December). Predicting mental health disorders using machine learning for employees in technical and non-technical companies. In *2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE)* (pp. 1-5). IEEE.
* Wang, X., Li, H., Sun, C., Zhang, X., Wang, T., Dong, C., & Guo, D. (2021). Prediction of mental health in medical workers during COVID-19 based on machine learning. *Frontiers in public health*, *9*, 697850.
* Srividya, M., Mohanavalli, S., & Bhalaji, N. (2018). Behavioral modeling for mental health using machine learning algorithms. *Journal of medical systems*, *42*, 1-12.
* Vaishnavi, K., Kamath, U. N., Rao, B. A., & Reddy, N. S. (2022). Predicting mental health illness using machine learning algorithms. In *Journal of Physics: Conference Series* (Vol. 2161, No. 1, p. 012021). IOP Publishing.
* Mutalib, S. (2021). Mental health prediction models using machine learning in higher education institution. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *12*(5), 1782-1792.
* Awal, G. K., & Rao, K. (2021). Can Machine Learning Predict an Employee’s Mental Health?. In *Information, Communication and Computing Technology: 6th International Conference, ICICCT 2021, New Delhi, India, May 8, 2021, Revised Selected Papers 6* (pp. 235-247). Springer International Publishing.
* Sairam, U., & Voruganti, S. Mental Health Prediction Using Deep Learning. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, *10*.
* FITRIYANTI, A. N. *PREDICTING TECH WORKERS’MENTAL HEALTH CONDITION USING MACHINE LEARNING TECHINQUES* (Doctoral dissertation, tilburg university).
* Narayanrao, P. V., & Kumari, P. L. S. (2020, March). Analysis of machine learning algorithms for predicting depression. In *2020 international conference on computer science, engineering and applications (iccsea)* (pp. 1-4). IEEE.
* Stevie Chancellor, Eric PS Baumer, and Munmun De Choudhury. Who is the” human” in human-centered machine learning: The case of predicting mental health from social media. Proceedings of the ACM on Human-Computer Interaction, 3(CSCW):1–32, 2019.
* Sarah Graham, Colin Depp, Ellen E Lee, Camille Nebeker, Xin Tu, Ho-Cheol Kim, and Dilip V Jeste. Artificial intelligence for mental health and mental illnesses: an overview. Current psychiatry reports, 21(11):1–18, 2019.
* Theodoros Iliou, Georgia Konstantopoulou, Mandani Ntekouli, Christina Lymperopoulou, Konstantinos Assimakopoulos, Dimitrios Galiatsatos, and George Anastassopoulos. Iliou machine learning preprocessing method for depression type prediction. Evolving Systems, 10(1):29–39, 2019.
* T Nagar. Prediction of mental health problems among children using machine learning techniques. [5] Adrian BR Shatte, Delyse M Hutchinson, and Samantha J Teague.
* Machine learning in mental health: a scoping review of methods and applications. Psychological medicine, 49(9):1426–1448, 2019.
* M Srividya, S Mohanavalli, and N Bhalaji. Behavioral modeling for mental health using machine learning algorithms. Journal of medical systems, 42(5):1–12, 2018.
* Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms (2020). https://doi.org/10.1016/j.procs.2020.03.442
* Stress Detection Using Machine Learning Algorithms(2020) .
* Mental Stress Level Prediction and Classification based on Machine Learning (2020) .

**2.3 Problem Statement Definition:**

In contemporary professional landscapes, the well-being of working professionals is a critical concern, with mental health challenges posing a significant threat to individual performance and organizational productivity. The absence of proactive tools to identify and address these issues in the early stages contributes to heightened instances of burnout, stress, and diminished overall mental well-being among the workforce.

This project seeks to address this gap by formulating and implementing a machine learning-based solution for predicting mental health illnesses among working professionals. The absence of a preemptive system often leads to undetected issues, impacting both individual job satisfaction and organizational efficiency. The objective is to develop a robust predictive model that analyzes relevant data points to identify patterns indicative of potential mental health concerns.

**\*Key Components of the Problem:\***

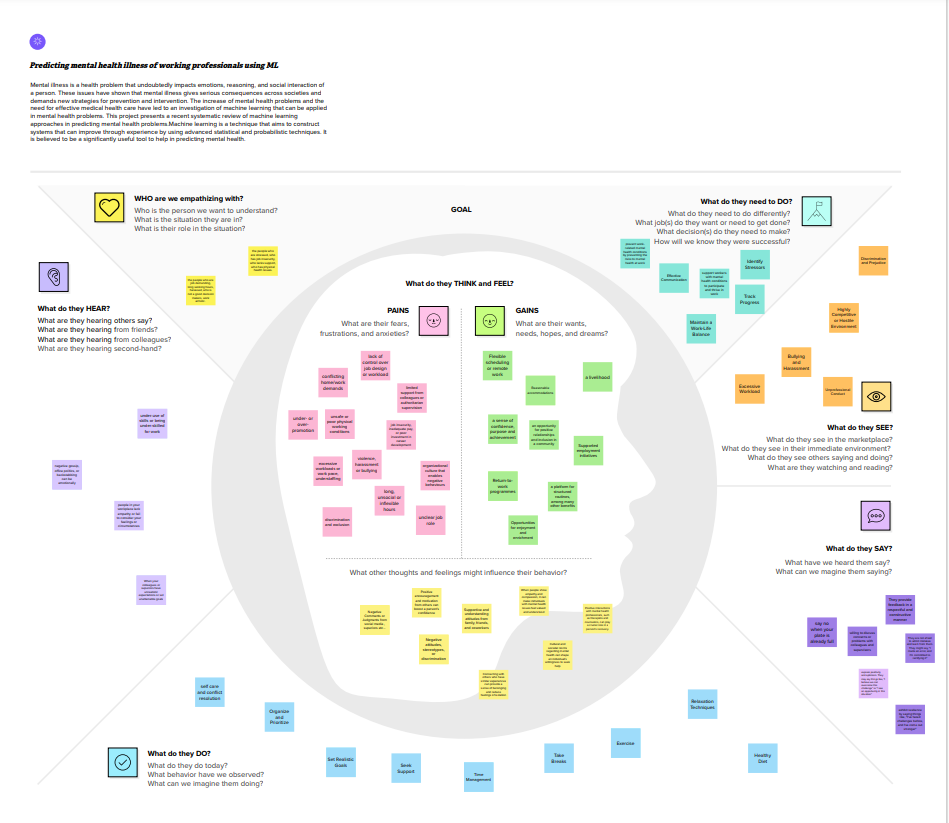
1. **\*Undetected Mental Health Issues:\*** Current workplace environments lack efficient mechanisms to identify and address mental health challenges, leading to undetected and unaddressed issues among working professionals.

2. **\*Impact on Individual Performance:\*** Unattended mental health issues contribute to decreased work performance, increased absenteeism, and a decline in overall job satisfaction, negatively impacting both the individual and the organization.

3. **\*Lack of Proactive Intervention:\*** The absence of a preemptive system means that individuals are often left without the necessary support until their mental health concerns reach critical levels, hindering their ability to thrive in the workplace.

**3.IDEATION & PROPOSED SOLUTION**

**3.1 Empathy Map Canvas:**

****

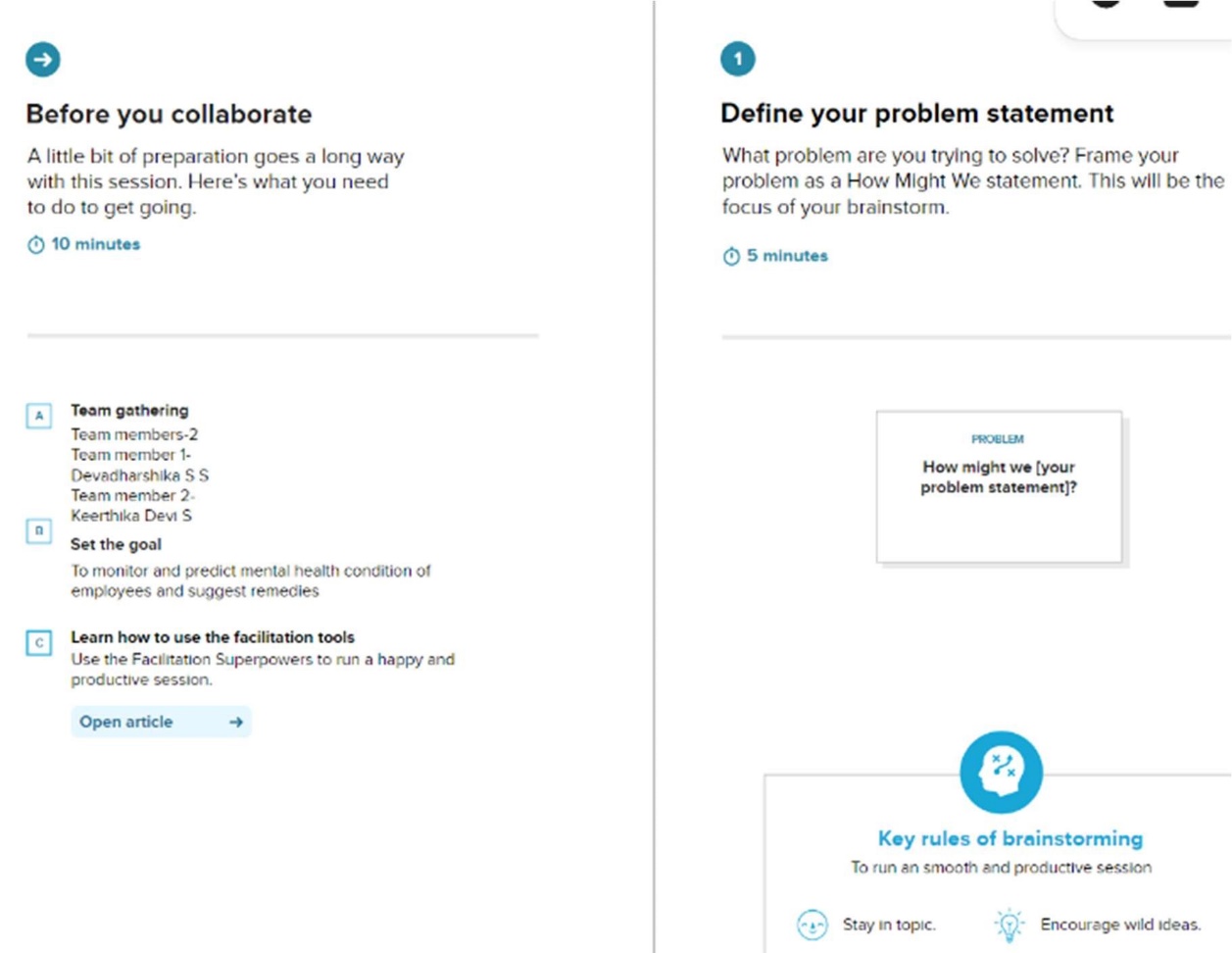
**3.2 Ideation & Brainstorming:**

Brainstorm & Idea Prioritization Template:

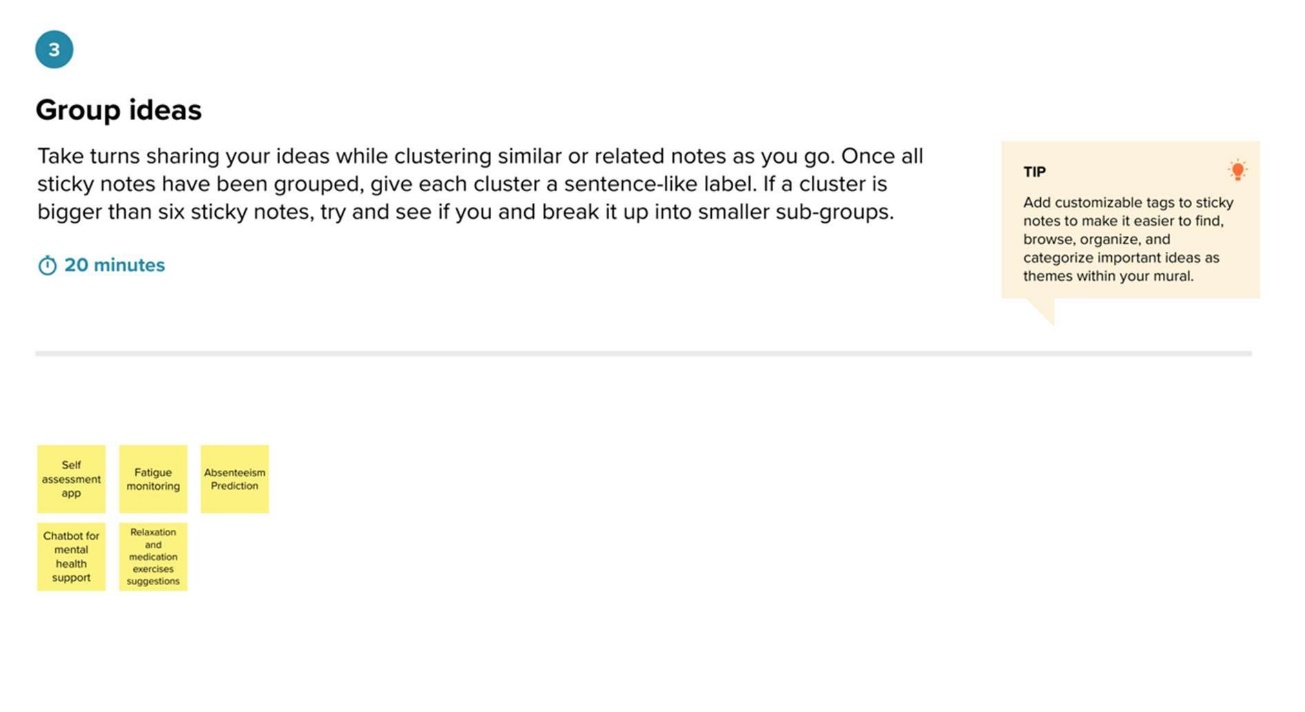
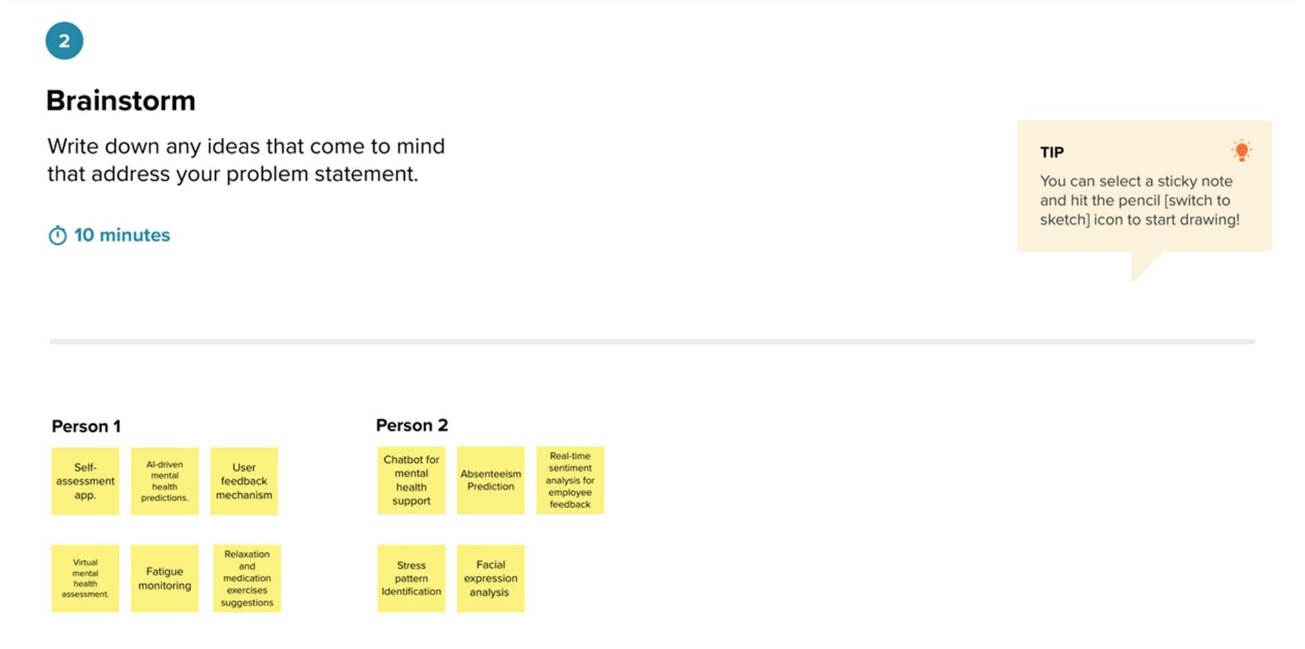
Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

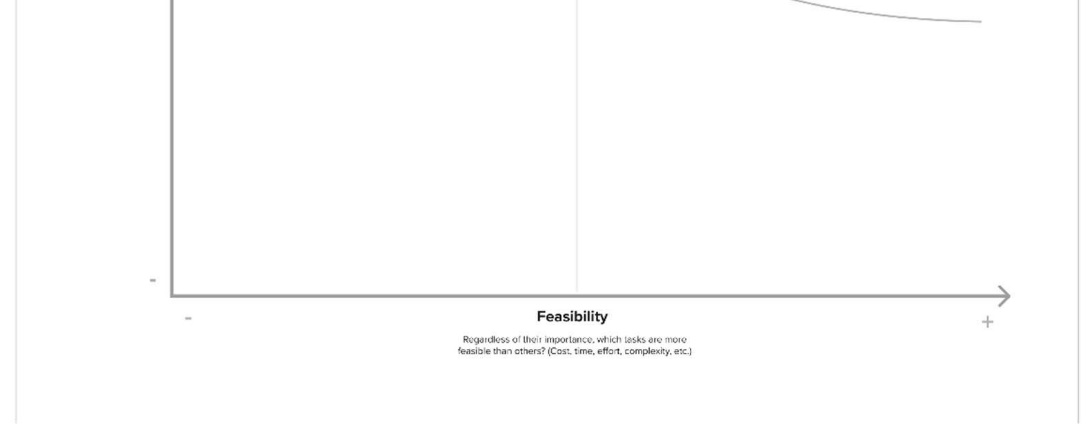
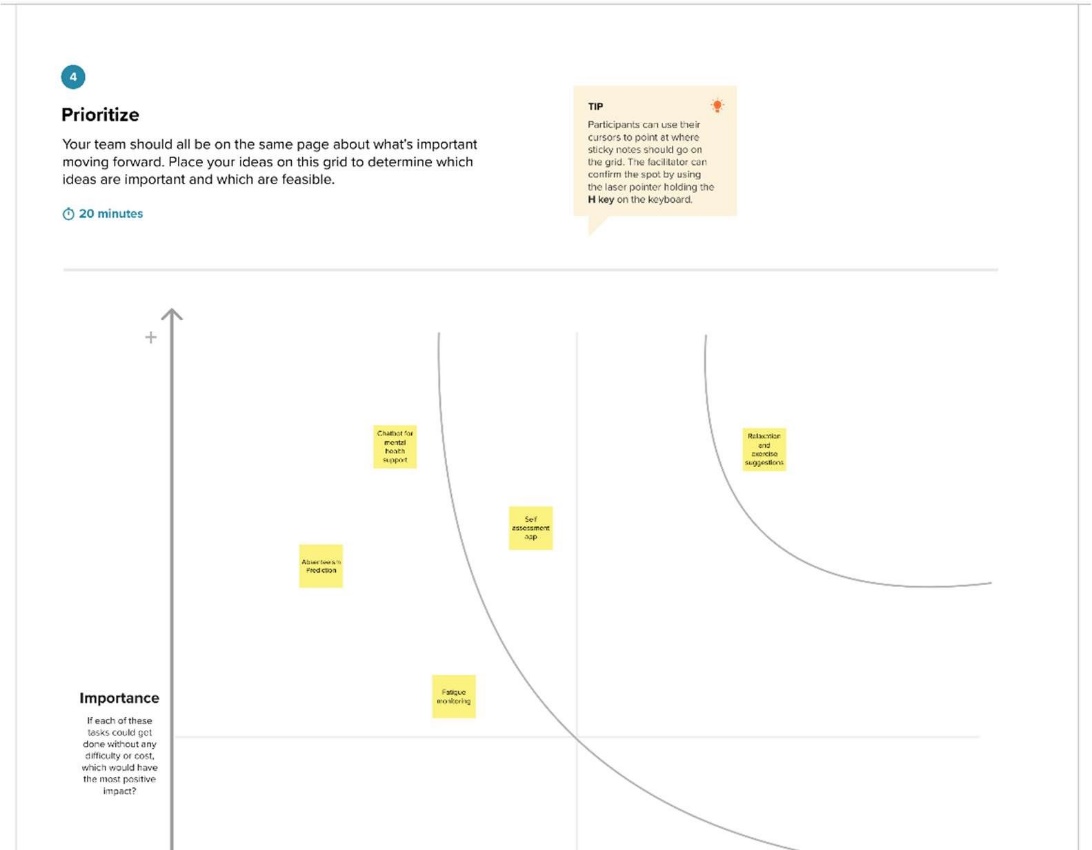
Step-1: Team Gathering, Collaboration and Select the Problem Statement



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



**4. REQUIREMENT ANALYSIS**

* 1. **Functional requirement**

User Authentication:

Users should be able to securely register and log in to the system.

Intuitive Questionnaire:

The system must present an easy-to-use questionnaire for users to provide details about their employment status and relevant attributes.

Machine Learning Integration:

The system must integrate machine learning algorithms (Random Trees and Logistic Regression) for mental health prediction.

Results Presentation:

The system should present clear and concise mental health prediction results to users.

Customizable Notifications:

Users should have the ability to customize the frequency and format of notifications related to their mental health results.

Data Privacy Features:

The system should allow users to delete or anonymize their data, ensuring privacy and compliance with data protection regulations.

Professional Collaboration:

Users should be able to share their mental health results with healthcare professionals or counselors for further guidance.

Educational Resources:

The system should provide educational resources or tips related to mental well-being.

* 1. Non-Functional requirements

Security:

The authentication process and user data must be secured to protect user privacy.

Performance:

The system must provide a responsive and efficient user experience, especially during the integration of machine learning algorithms.

Scalability:

The system should be designed to handle an increasing number of users and data inputs.

Compatibility:

The application should be compatible with various devices and browsers to ensure accessibility.

Reliability:

The mental health predictions must be reliable and consistent with recognized standards.

User-Friendliness:

The interface of the questionnaire and result presentation should be user-friendly and easy to navigate.

**5.PROJECT DESIGN**

**5.1** **Data Flow Diagrams & User Stories:**

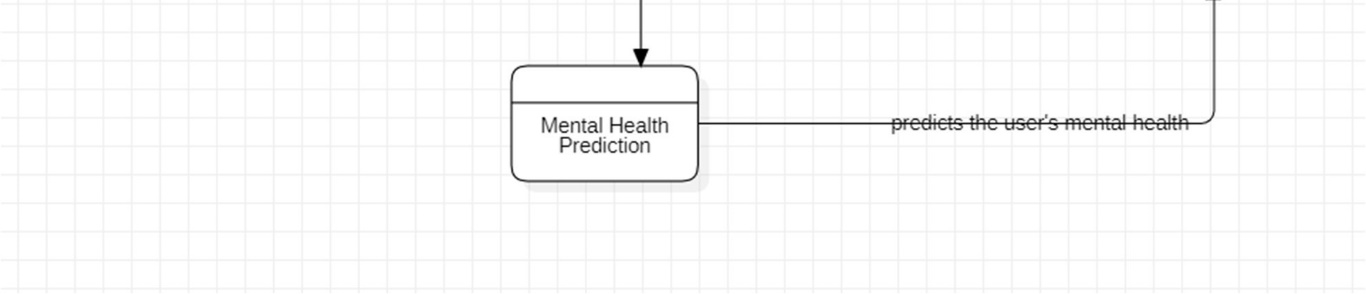
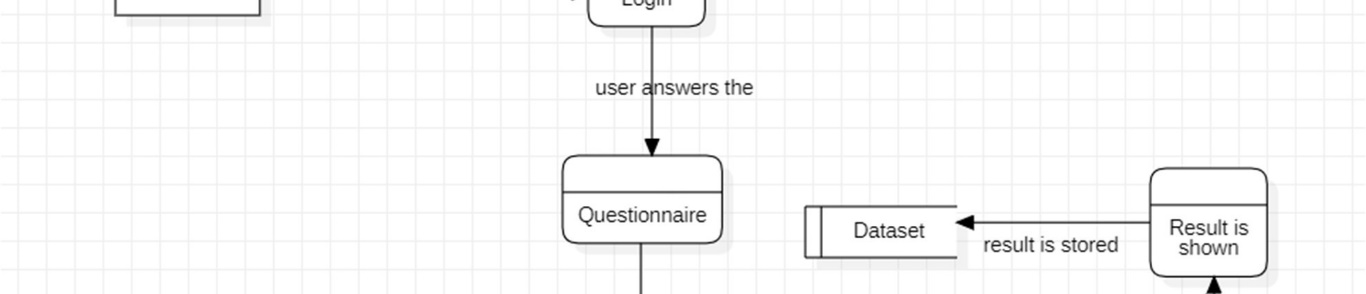
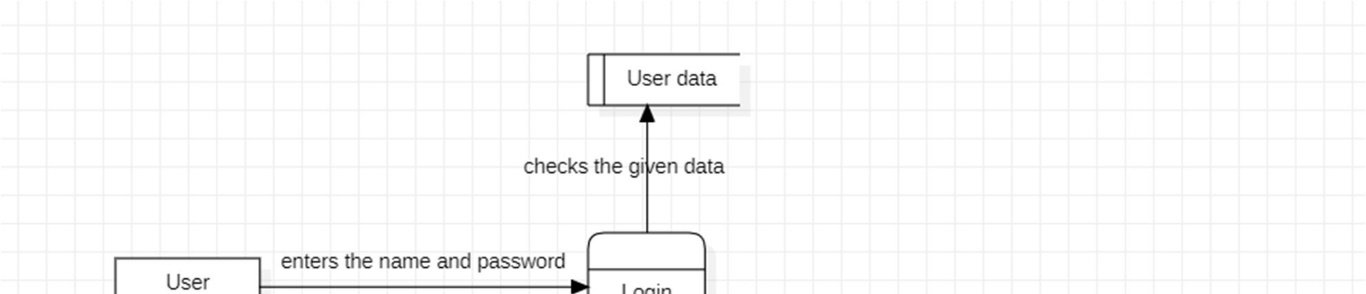
Data Flow Diagrams:

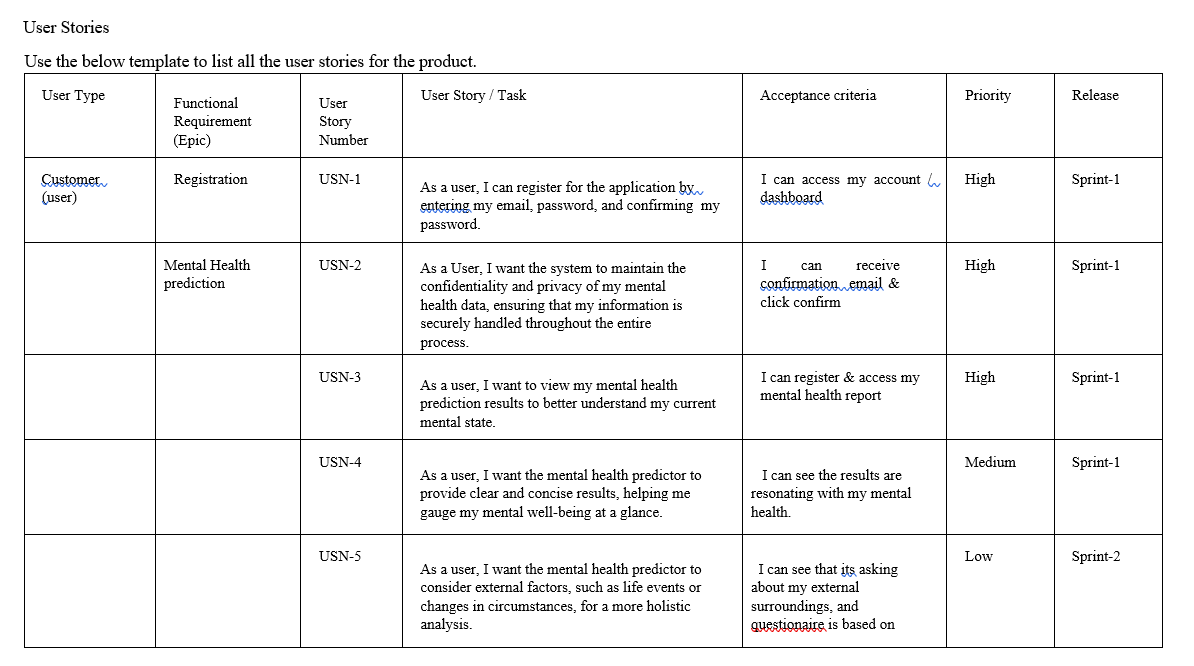
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

By initiating the system by logging in, triggering a secure authentication process. Subsequently, they navigate to a detailed questionnaire where they input key attributes, such as employment status. The gathered data then flows seamlessly into a robust mental health predictor, employing advanced algorithms like Random Trees and Logistic Regression. This predictive analysis yields valuable insights into the user's mental health.

To culminate this process, the results are systematically stored for future reference and analysis. This comprehensive approach ensures a user-centric, data-driven methodology for mental health assessment, with a focus on accuracy, security, and meaningful insights.

Data Flow Diagram – DFD 0 level





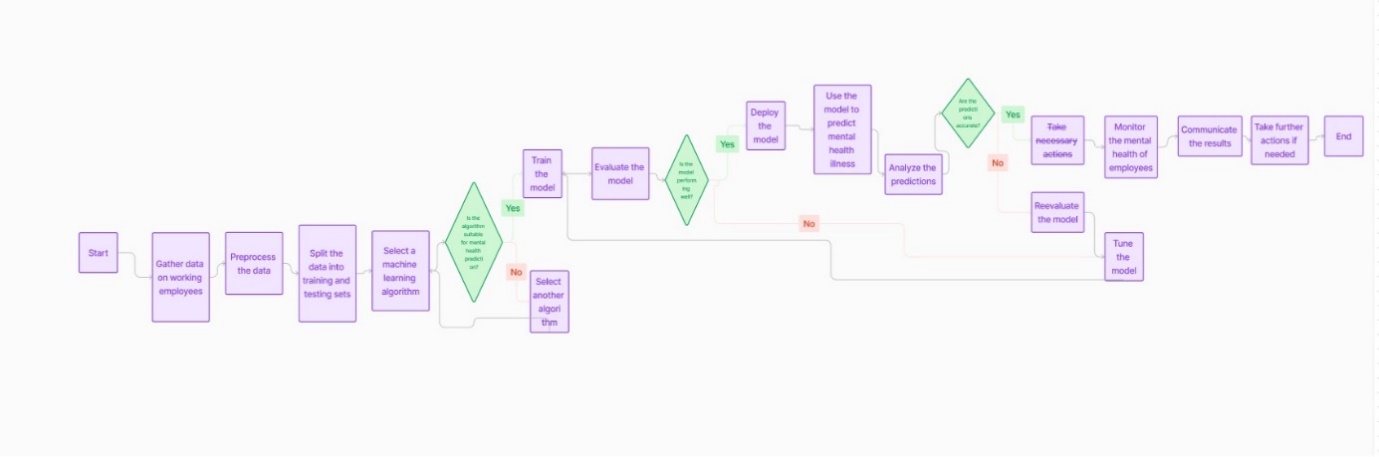
**5.2 Solution Architecture:**

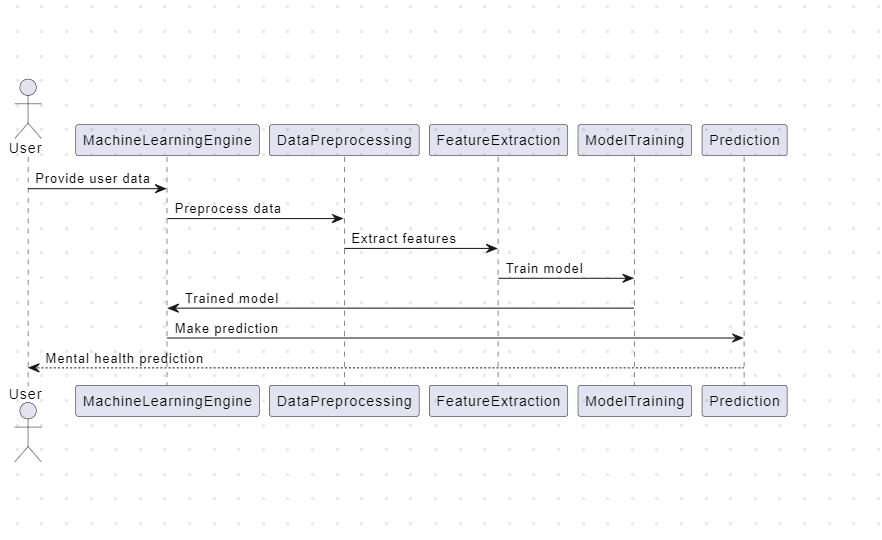
**Solution Architecture:**

Solution architecture is a complex process – with many sub-processes – that bridges the gap between business problems and technology solutions. Its goals are to:

* Find the best tech solution to solve existing business problems.
* Describe the structure, characteristics, behavior, and other aspects of the software to project stakeholders.
* Define features, development phases, and solution requirements.
* Provide specifications according to which the solution is defined, managed, and delivered.

**Example - Solution Architecture Diagram:**

****

****

Designing a solution architecture for predicting mental health illness in working employees using machine learning involves several components and considerations. Here's a high-level architecture for such a system:

1. Data Collection:

- Employee Data: Collect information about the employees, including demographics, work history, job roles, and any self-reported mental health information (e.g., surveys, self-assessment).

- Work Environment Data: Gather data about the work environment, such as workload, working hours, workplace stress factors, and company policies related to mental health.

2. Data Storage:

- Store the collected data in a secure and scalable data storage solution, such as a relational database or a data warehouse. Ensure that the data is anonymized and complies with data privacy regulations.

3. Data Preprocessing:

- Perform data cleaning and preprocessing to handle missing values, outliers, and standardize data formats. Normalize or scale features as needed.

4. Feature Engineering:

- Create relevant features from the collected data that can help in predicting mental health issues. This may include creating variables based on historical data or aggregating certain attributes.

5. Machine Learning Model:

- Develop and train machine learning models to predict mental health issues based on the prepared data. You can use various algorithms such as logistic regression, decision trees, random forests, or neural networks.

- Consider using natural language processing (NLP) techniques if there are textual data, such as employee comments or feedback.

6. Model Evaluation:

- Split the dataset into training, validation, and test sets to evaluate the model's performance. Common evaluation metrics include accuracy, precision, recall, F1-score, and ROC-AUC.

- Fine-tune hyperparameters and perform cross-validation to optimize model performance.

7. Model Deployment:

- Once you have a well-performing model, deploy it as a web service or API using containerization tools like Docker. You can use platforms like AWS, Azure, or GCP for deployment.

- Implement security and access control measures to protect sensitive employee data.

8. Real-time Data Collection:

- Set up real-time data collection mechanisms to continuously monitor employee data. This can include integrating with HR systems or periodic employee surveys.

9. Monitoring and Alerts:

- Implement monitoring and alerting systems to detect anomalies or potential mental health issues in real-time. Notify relevant stakeholders, such as HR or managers, when a concerning pattern is identified.

10. Visualization and Reporting:

- Create dashboards and reporting tools for HR and management to visualize and interpret the model's predictions and insights.

11. Privacy and Ethics:

- Ensure the system complies with data privacy regulations, such as GDPR or HIPAA, and maintains ethical considerations for handling sensitive mental health data.

12. Continuous Improvement:

- Regularly retrain and update the model using fresh data to improve accuracy and adapt to changing workplace conditions.

13. User Feedback Integration:

- Encourage employees to provide feedback on the predictions and system usability to make continuous improvements.

14. Employee Support:

- Offer resources and support to employees who are identified as at risk of mental health issues, such as counseling services, mental health awareness programs, or access to EAP (Employee Assistance Program).

15. Legal and Compliance:

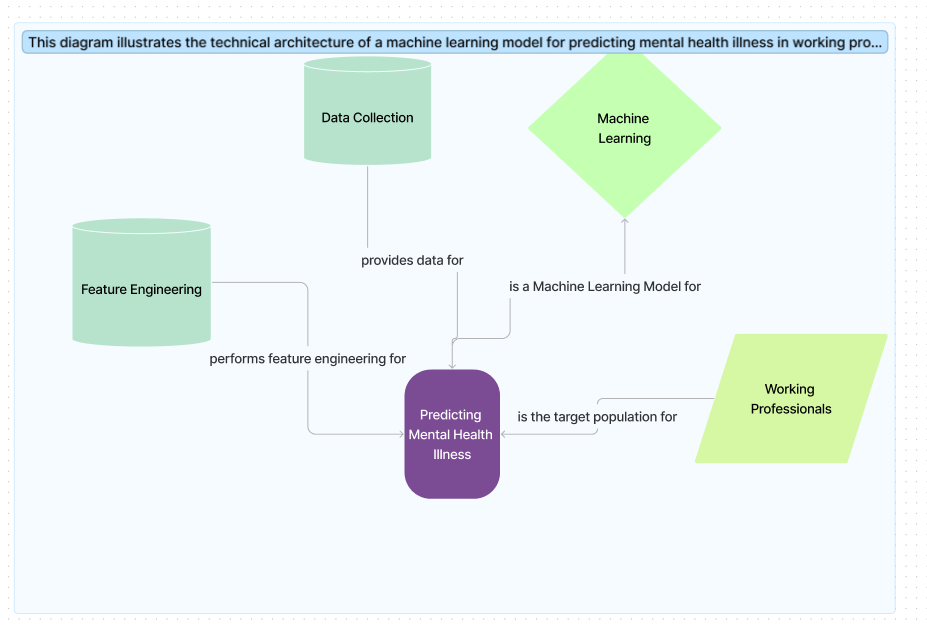
- Work with legal and compliance teams to ensure that the system aligns with all relevant employment and privacy laws and regulations.

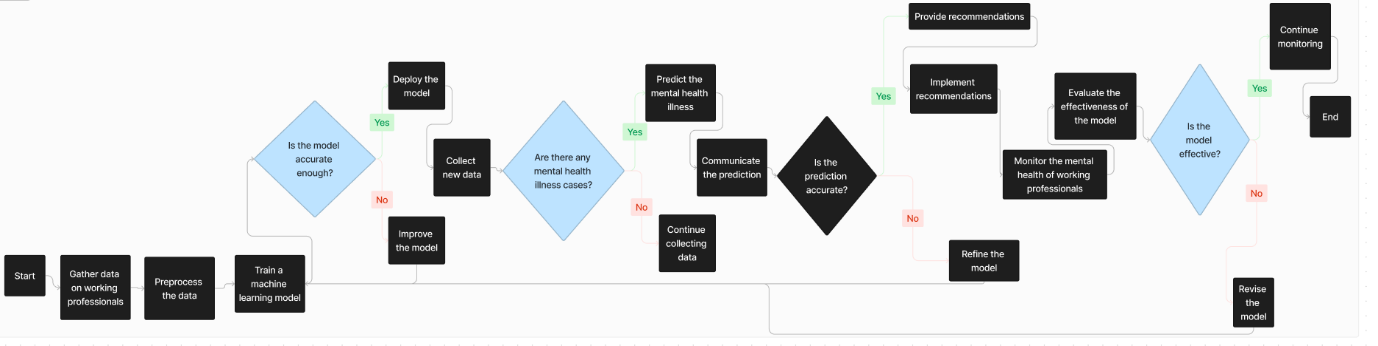
It's essential to involve mental health professionals, data scientists, and HR experts in the design and implementation of this system to ensure its effectiveness, accuracy, and ethical handling of sensitive data. Additionally, strong communication and transparency with employees about the purpose and use of the system are crucial for its success.

**6.PROJECT PLANNING & SCHEDULING**

**6.1 Technical Architecture:**

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table 2





**Table-1 : Components & Technologies:**

|  |  |  |
| --- | --- | --- |
| **S.No** | **Component** | **Technology** |
| 1. | **Programming Language** | Python: Widely used for machine learning and has extensive libraries and frameworks. |
| 2. | **Data Collection and Storage** | SQL or NoSQL databases (e.g., PostgreSQL, MongoDB) for storing structured or unstructured data related to professionals' mental health. |
| 3. | **Data Preprocessing** | * Pandas: For data manipulation and cleaning. * NumPy: For numerical operations on the data. |
| 4. | **Machine Learning Framework** | * Scikit-learn: Offers a variety of tools for data mining and data analysis. * TensorFlow or PyTorch: Deep learning frameworks for building and training neural networks. |
| 5. | **Model Selection** | * Logistic Regression, Decision Trees, Random Forests: For simpler models. * Neural Networks: For complex patterns in the data. |
| 6. | **Model Evaluation** | * Cross-validation techniques to assess model performance. * Metrics like accuracy, precision, recall, and F1 score to evaluate the model's effectiveness. |
| 7. | **Hyperparameter Tuning** | GridSearchCV or RandomizedSearchCV in Scikit-learn for optimizing model parameters. |
| 8. | **Deployment** | * Flask or Django: For building a web application. * Docker: Containerization for easy deployment and scalability. * Cloud services (e.g., AWS, Azure, Google Cloud) for hosting the application. |
| 9. | **Monitoring and Logging** | * Implement logging mechanisms to track model performance and user interactions. * Set up monitoring tools to identify issues and ensure the system's reliability. |
| 10. | **Security** | * Implement secure coding practices. * Protect sensitive user data and comply with data protection regulations. |
| 11. | **User Interface** | * HTML, CSS, JavaScript for creating a user-friendly interface. * Visualization libraries like D3.js or Chart.js for displaying insights. |
| 12. | **Collaboration and Version Control** | * Git: For version control. * Platforms like GitHub or GitLab for collaboration. |

**Table-2: Application Characteristics:**

|  |  |  |
| --- | --- | --- |
| **S.No** | **Characteristics** | **Description** |
| 1 | **Data Privacy and Security** | * Implement robust data encryption and ensure compliance with data protection regulations. * Prioritize user consent and clearly communicate how their data will be used. |
| 2 | **User-Friendly Interface** | * Design an intuitive and user-friendly interface to encourage regular usage. * Include simple and clear instructions for users to input relevant data. |
| 3 | **Comprehensive Assessment** | * Incorporate a diverse range of factors such as work-related stressors, sleep patterns, physical activity, and social interactions. * Utilize validated mental health assessment tools to ensure accuracy. |
| 4 | **Real-time Monitoring** | * Enable continuous monitoring to detect changes in mental health over time. * Provide timely alerts or recommendations based on the analysis of user data. |
| 5 | **Personalized Insights** | * Tailor recommendations based on individual profiles and preferences. * Consider factors like personality traits and coping mechanisms in the analysis |
| 6 | **Integration with Wearables** | * Allow users to connect their wearable devices to provide additional data for analysis. * Incorporate features that leverage biometric data for a more holistic understanding |
| 7 | **Education and Resources** | * Offer educational content on mental health, stress management, and coping strategies. * Provide links to relevant resources and support networks. |
| 8 | **Multilingual Support** | * Ensure the application is accessible to a diverse user base by offering multiple language options. * Consider cultural nuances in the design and recommendations. |
| 9 | **Machine Learning Algorithms** | * Use advanced machine learning algorithms for accurate prediction. * Continuously update and refine the algorithms based on user feedback and evolving research |
| 10 | **Feedback Mechanism** | * Include a feedback system to gather user input and improve the application over time. * Encourage users to report any concerns or inaccuracies in predictions. |
| 11 | **Collaboration with Professionals** | * Provide an option for users to share insights with mental health professionals. * Collaborate with mental health experts to enhance the accuracy of predictions and recommendations |
| 12 | **Scalability** | * Design the application to handle a growing user base and evolving technology. * Ensure the infrastructure can support increased data volume without compromising performance. |
|  |  |  |

**6.2 Sprint Planning & Estimation:**

Use the below template to create product backlog and sprint schedule

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sprint | Functional Requirement (Epic) | User Story  Number | User Story / Task | Story Points | Priority | Team  Members |
| Sprint-1 | User Authentication | USN-1 | As a user, I want a secure authentication process to ensure the confidentiality of my mental health data. | 3 | High | Keerthika |
| Sprint-1 | Intuitive Questionnaire | USN-2 | As a user, I want an intuitive questionnaire interface to easily provide details about my employment status and other relevant attributes. | 5 | Medium | Devadharshik a |
| Sprint-2 | Machine Learning  Integration | USN-3 | As a user, I want the mental health predictor to employ advanced machine learning algorithms like Random Trees and Logistic Regression for accurate predictions. | 8 | High | Keerthika |
| Sprint- | Results Presentation | USN-4 | As a user, I want the mental health predictor to present clear and concise results, helping me gauge my mental wellbeing at a glance. | 5 | Medium | Devadharshik a |
| Sprint-3 | Customizable Notifications | USN-5 | As a user, I want the ability to customize the frequency and format of notifications related to my mental health results. | 3 | Low | Keerthika |
| Sprint-3 | Data Privacy Features | USN-6 | As a user, I want control over my data, with the option to delete or anonymize it from the system when discontinuing use. | 5 | Medium | Devadharshik a |
| Sprint-4 | Professional  Collaboration | USN-7 | As a user, I want the option to share my mental health results with healthcare professionals or counselors for further guidance and support. | 8 | High | Keerthika |

**6.3 Sprint Delivery Schedule:**

Project Tracker, Velocity & Burndown Chart: (4 Marks)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Sprint | Total Story Points | Duration | Sprint Start Date | Sprint End Date (Planned) | Story Points Completed (as on Planned End  Date) | Sprint Release Date (Actual) |
| Sprint-1 | 20 | 7 Days | 20 Oct 2023 | 27 Oct 2023 | 13 | 29 Oct 2023 |
| Sprint-2 | 21 | 7 Days | 27 Oct 2023 | 03 Nov 2023 | 17 | 03 Nov 2023 |
| Sprint-3 | 18 | 7 Days | 03 Nov 2023 | 10 Nov 2023 | 12 | 05 Nov 2023 |
| Sprint-4 | 5 | 7 Days | 10 Nov 2023 | 17 Nov 2023 | 5 | 10 Nov 2023 |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 16 (points per sprint). Let’s calculate

the team’s average velocity (AV) per iteration unit (story points per day)

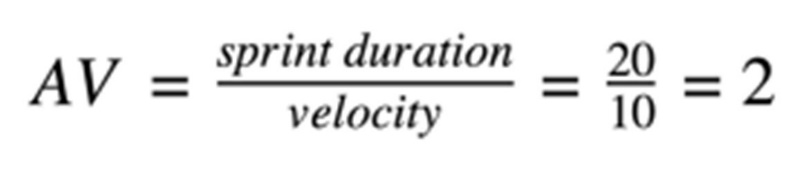
AV= Sprint Duration

\_\_\_\_\_\_\_\_\_\_\_\_\_

Velocity

= 16 / 7

= 2.28



Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.

**7.CODING & SOLUTIONING (Explain the features added in the project along with code)**

7.1 Logistic Regression model

def logisticRegression():

logreg = LogisticRegression()

logreg.fit(X\_train, y\_train)

y\_pred\_class = logreg.predict(X\_test)

accuracy\_score = evalClassModel(logreg, y\_test, y\_pred\_class, True)

#Data for final graph

methodDict['Log. Regression'] = accuracy\_score \* 100

logisticRegression()

This code defines a logistic regression model using scikit-learn's `LogisticRegression` class. It fits the model to the training data (`X\_train`, `y\_train`), predicts the target variable for the test data (`X\_test`), and evaluates the model using a function called `evalClassModel`. The accuracy score is then stored in a dictionary (`methodDict`) for later comparison. The final accuracy score, converted to percentage, is added to the dictionary as 'Log. Regression'.

7.2 TreeClassifier

def treeClassifier():

# Calculating the best parameters

tree1 = DecisionTreeClassifier()

featuresSize = feature\_cols1.\_\_len\_\_()

param\_dist = {"max\_depth": [3, None],

"max\_features": randint(1, featuresSize),

"min\_samples\_split": randint(2, 9),

"min\_samples\_leaf": randint(1, 9),

"criterion": ["gini", "entropy"]}

tuningRandomizedSearchCV(tree1, param\_dist)

tree1 = DecisionTreeClassifier(max\_depth=3, min\_samples\_split=8, max\_features=6, criterion='entropy', min\_samples\_leaf=7)

tree1.fit(X\_train, y\_train)

y\_pred\_class = tree1.predict(X\_test)

print(y\_pred\_class)

accuracy\_score = evalClassModel(tree1, y\_test, y\_pred\_class, True)

#Data for final graph

methodDict['Decision Tree Classifier'] = accuracy\_score \* 100

This code defines a decision tree classifier using scikit-learn's `DecisionTreeClassifier`. It initializes a decision tree (`tree1`), determines the best parameters using hyperparameter tuning via `tuningRandomizedSearchCV`, and then fits the model to the training data (`X\_train`, `y\_train`). The predicted values for the test set (`X\_test`) are printed, and the model is evaluated using the `evalClassModel` function. The resulting accuracy score is stored in a dictionary (`methodDict`) for later comparison, labeled as 'Decision Tree Classifier'.

**8.PERFORMANCE TESTING**

**8.1 Performace Metrics**

Project team shall fill the following information in model performance testing template.

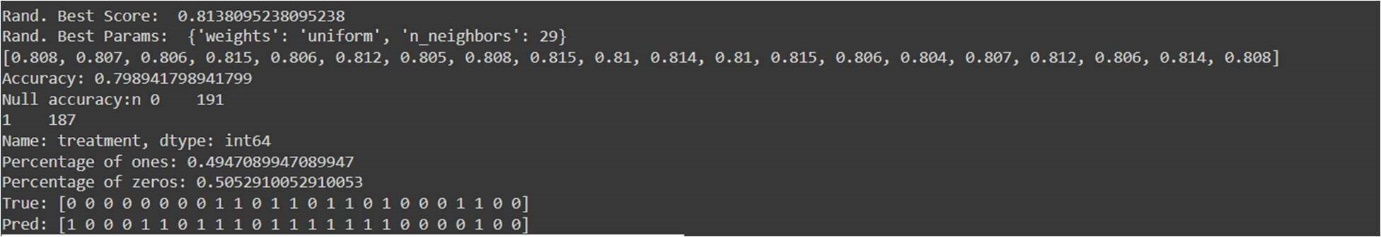
|  |  |  |  |
| --- | --- | --- | --- |
| S.No. | Parameter | Values | Screenshot |
| 1. | Metrics | Regression Model:  MAE - , MSE - , RMSE - , R2 score -    Classification Model:  Confusion Matrix - , Accuray Score- & Classification Report - | Classification Model:  Logistic Regression:  Accuracy Score-  0.7962962962962963        Classification Report:  Classification Accuracy:  0.7962962962962963  Classification Error:  0.20370370370370372  False Positive Rate:  0.25654450261780104  Precision:  0.7644230769230769 AUC Score:  0.7968614385306716  Crossvalidated AUC values:  0.8721058623942408  First 10 predicted  responses:n [1 0 0 0 1 1 0 1 0 1] |

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | First 10 predicted  probabilities of class members:n [[0.08039445  0.91960555]  [0.95925545 0.04074455]  [0.96155434 0.03844566]  [0.76950203 0.23049797]  [0.36949375 0.63050625]  [0.05229329 0.94770671]  [0.72663965 0.27336035]  [0.28650524 0.71349476]  [0.58390697 0.41609303]  [0.44894488 0.55105512]]  First 10 predicted  probabilities:n [[0.91960555]  [0.04074455]  [0.03844566]  [0.23049797]  [0.63050625]  [0.94770671]  [0.27336035]  [0.71349476]  [0.41609303]  [0.55105512]]    Decision Tree:  Accuracy :  0.8068783068783069 Confusion Matrix:      Classification Report:  Classification Accuracy:  0.8068783068783069  Classification Error:  0.19312169312169314  False Positive Rate:  0.3193717277486911  Precision:  0.7415254237288136 |

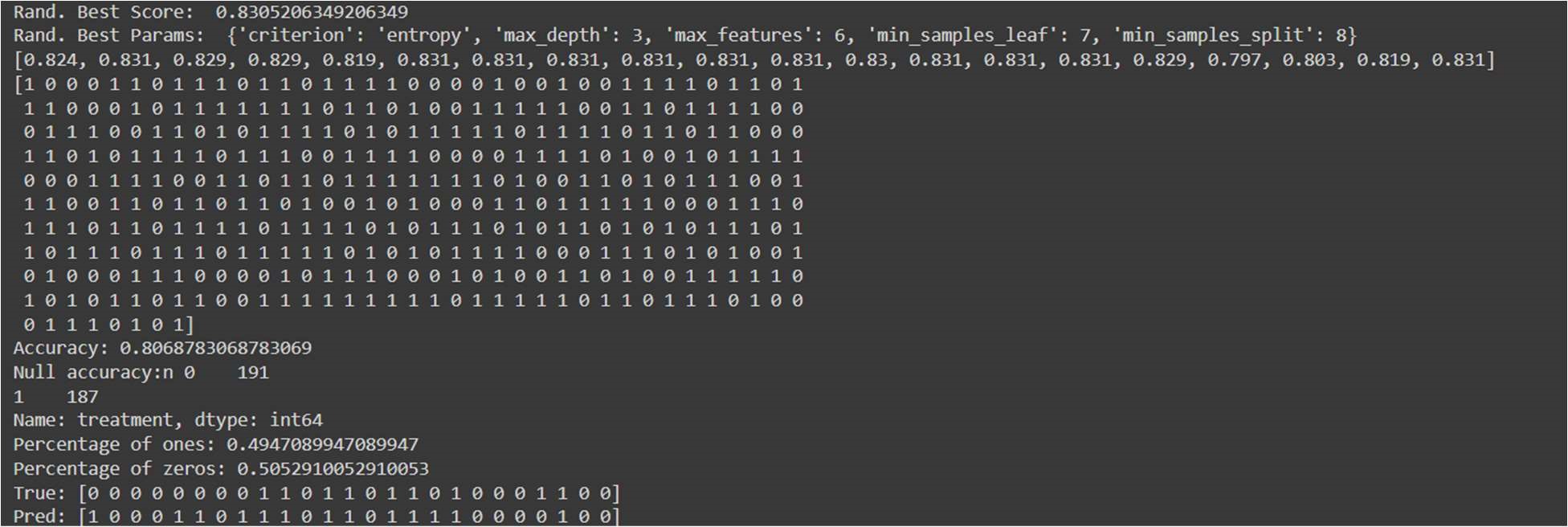
|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  | AUC Score:  0.8082285746283282  Crossvalidated AUC values:  0.8841930613718008  First 10 predicted  responses:n [1 0 0 0 1 1 0 1 1  1]  First 10 predicted  probabilities of class members:n [[0.29918033  0.70081967]  [0.96969697 0.03030303]  [1. 0. ]  [0.90517241 0.09482759]  [0.29918033 0.70081967]  [0.18181818 0.81818182]  [0.90517241 0.09482759]  [0.29918033 0.70081967]  [0.22018349 0.77981651]  [0.22018349 0.77981651]]  First 10 predicted  probabilities:n [[0.70081967]  [0.03030303]  [0. ]  [0.09482759]  [0.70081967]  [0.81818182]  [0.09482759]  [0.70081967]  [0.77981651]  [0.77981651]]    Knn:  Accuracy:  0.798941798941799 Confusion Matrix:      Classification Report:  Classification Accuracy:  0.798941798941799 |
|  |  |  | Classification Error:  0.20105820105820105  False Positive Rate:  0.2931937172774869  Precision:  0.7488789237668162 AUC Score:  0.7999272055323796  Crossvalidated AUC values:  0.8653315335991026  First 10 predicted  responses:n [1 0 0 0 1 1 0 1 1  1]  First 10 predicted  probabilities of class members:n [[0.25925926  0.74074074]  [1. 0. ]  [0.88888889 0.11111111]  [0.7037037 0.2962963 ]  [0.48148148 0.51851852]  [0.2962963 0.7037037 ]  [0.55555556 0.44444444]  [0.11111111 0.88888889]  [0.25925926 0.74074074]  [0.33333333 0.66666667]]  First 10 predicted  probabilities:n [[0.74074074]   1. ]   [0.11111111]  [0.2962963 ]  [0.51851852]  [0.7037037 ]  [0.44444444]  [0.88888889]  [0.74074074]  [0.66666667]] |
| 2. | Tune the Model | Hyperparameter Tuning - knn  Validation Method – Cross  Validation Score,  Grid Search |  |

Tune the Model:

Screenshot for GridSearch: Knn:

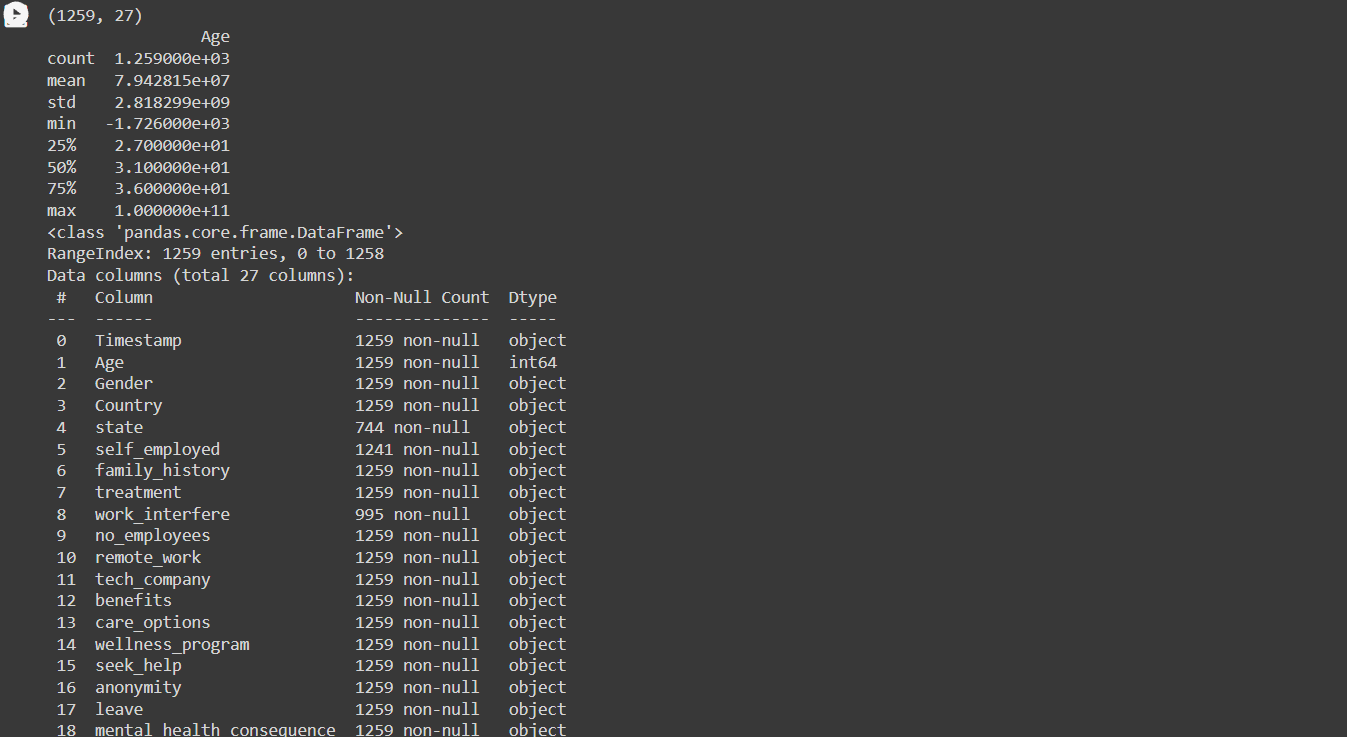


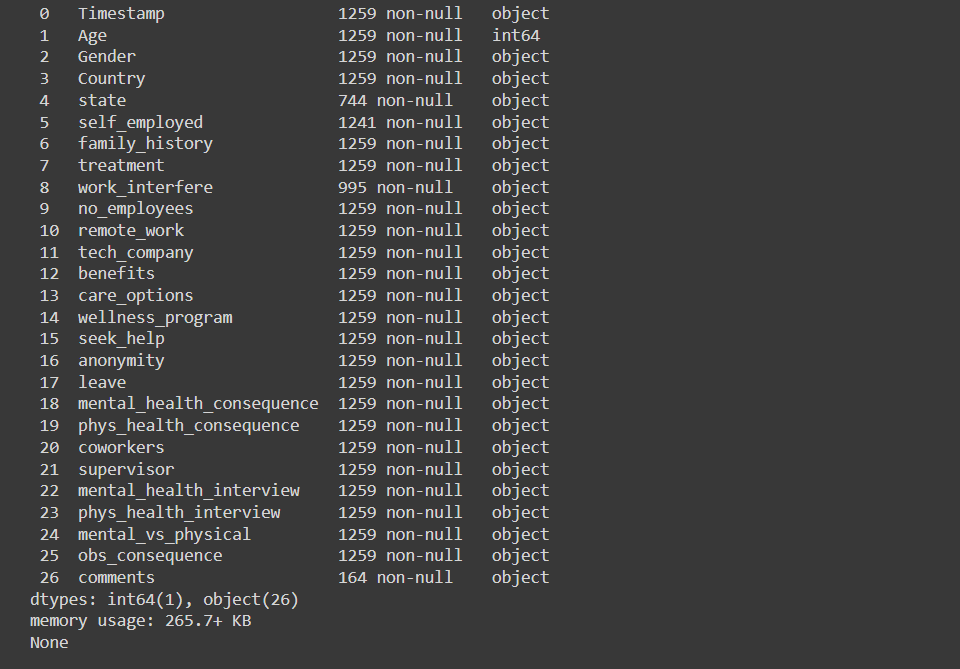
Tree Classifier:

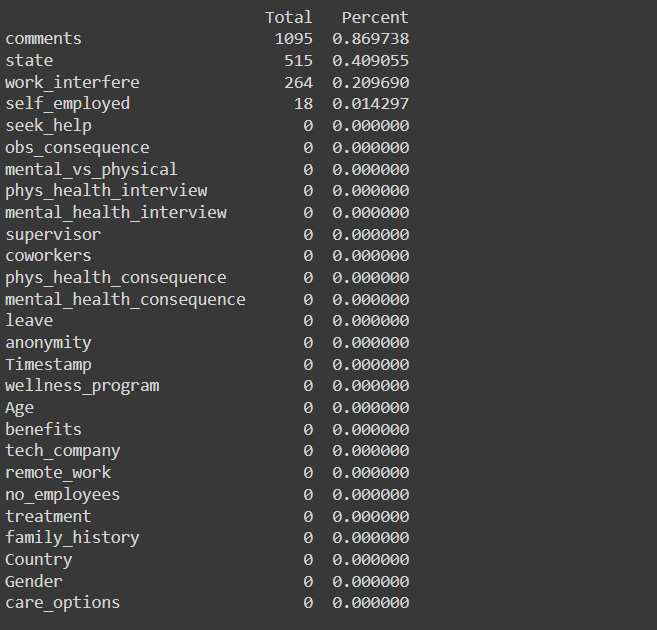


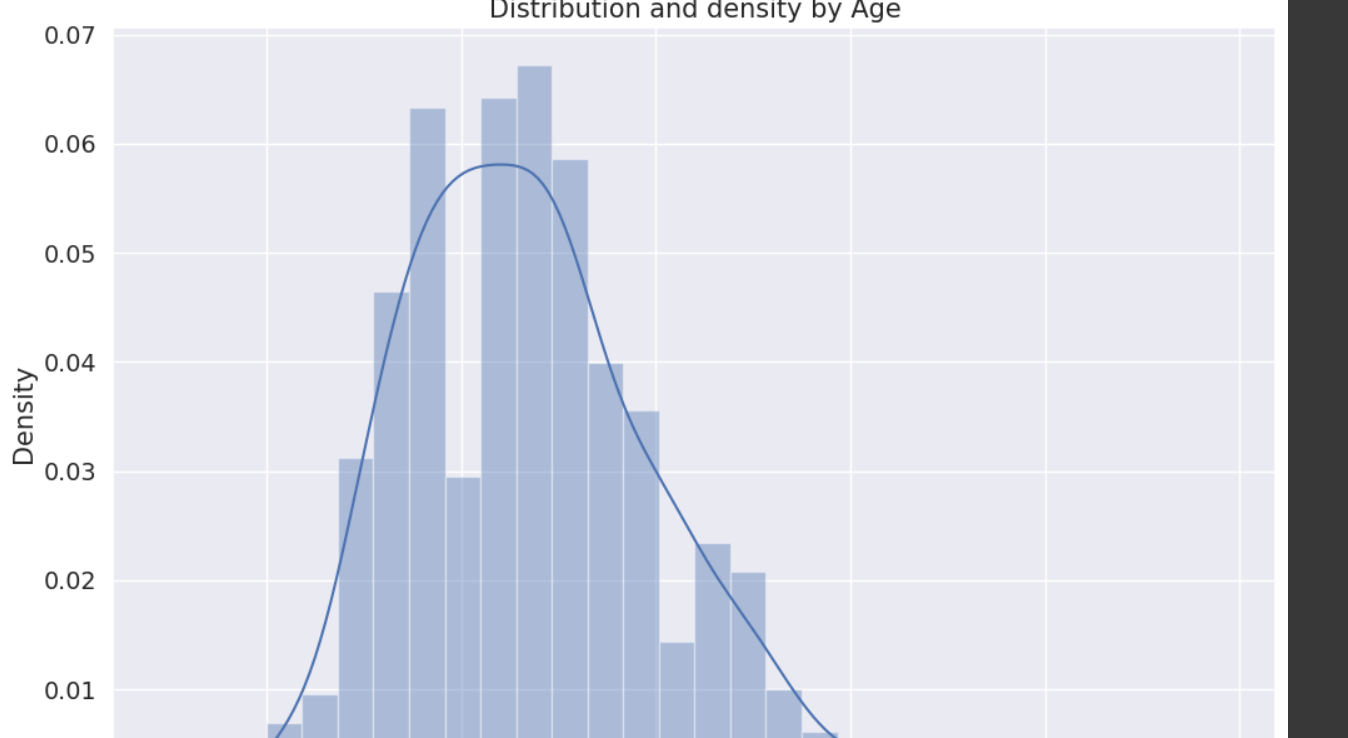
**9.RESULTS**

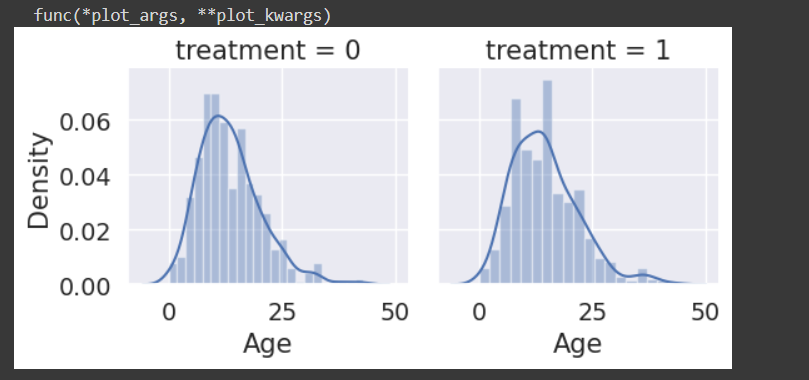
**9.1 Output Screenshots**

****

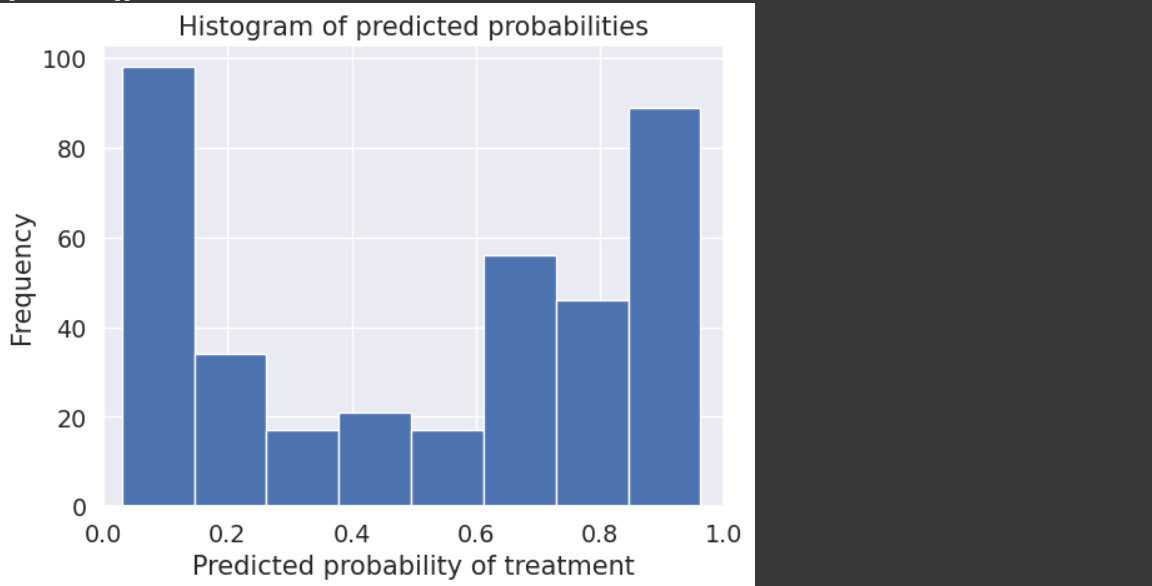
****

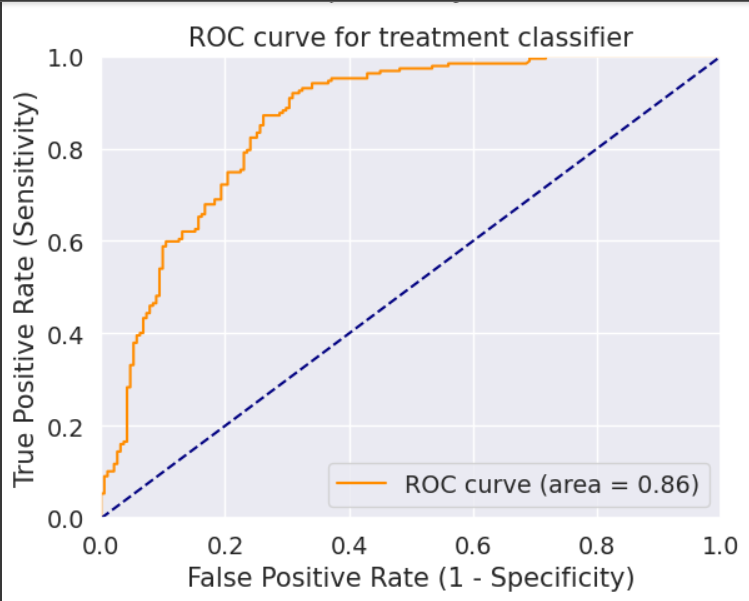
****

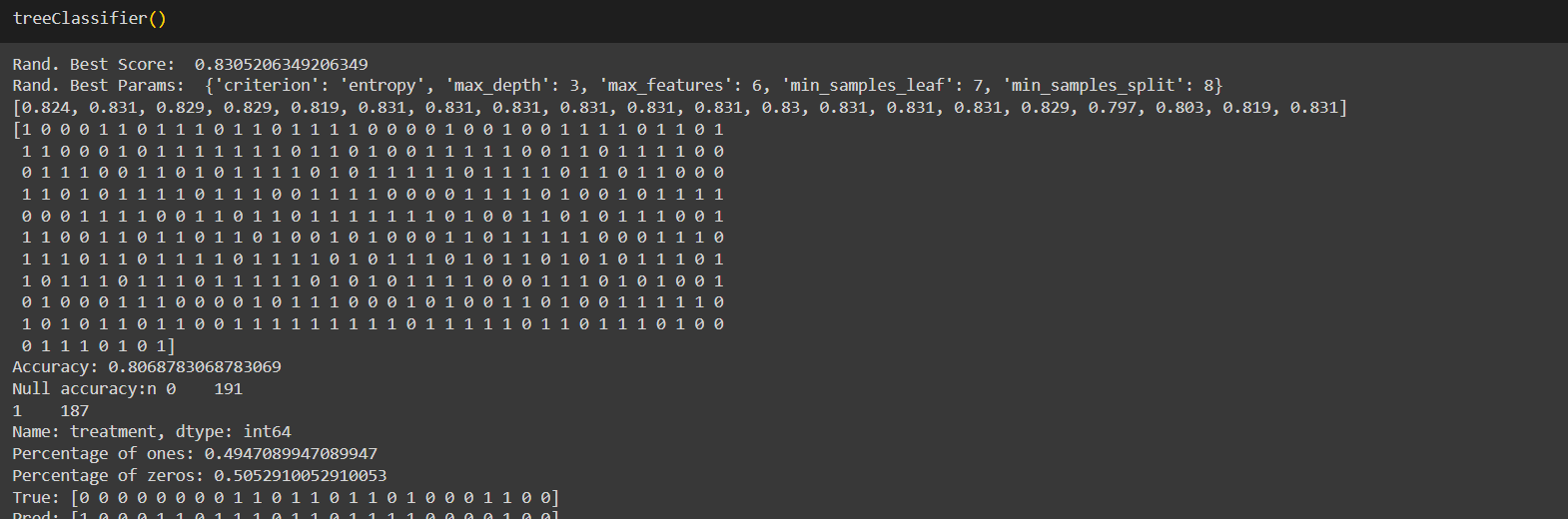
****

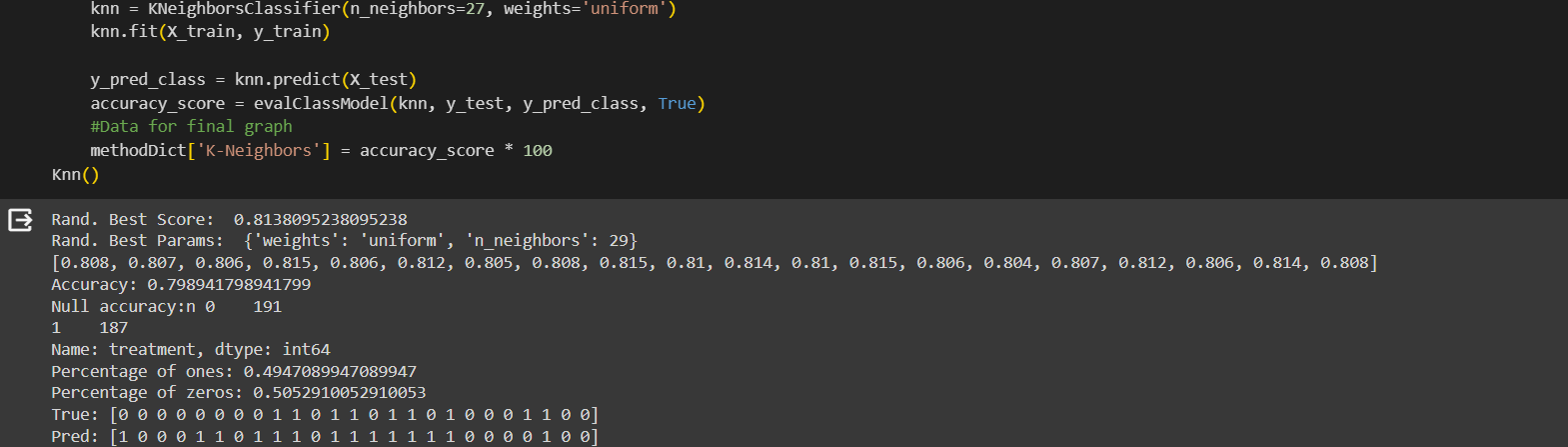
****

****

****

****

****

****

**10.ADVANTAGES & DISADVANTAGES**

**Advantages of predicting mental health illness of working professionals using machine learning:**

* Early detection: Machine learning models can analyze large amounts of data and identify patterns that may indicate the presence of mental health issues. Early detection can lead to timely intervention and support, potentially improving outcomes for individuals.
* Objective assessment: Machine learning algorithms can provide an objective assessment of mental health conditions, reducing the potential for bias and subjective judgments. This can help overcome the stigma associated with mental health and encourage individuals to seek appropriate help.
* Scalability: Machine learning models can be scaled to analyze large datasets, making it possible to predict mental health illness on a broader scale. This scalability can facilitate population-level interventions and resource allocation.
* Personalized interventions: By analyzing individual data, machine learning models can help tailor interventions and treatments to the specific needs of individuals. This personalization can lead to more effective and targeted interventions, potentially improving outcomes for individuals.
* Personalized Support: Predictive models can provide insights into individual needs, enabling employers to offer more personalized support tailored to each employee.
* Reduced Stigma: By framing mental health discussions in a data-driven context, the stigma around seeking help can be reduced, fostering a more open and supportive workplace culture.
* Resource Optimization: Organizations can allocate resources more efficiently, directing support where it is needed most, which can lead to cost savings and improved overall well-being.
* Improved Productivity: By addressing mental health issues proactively, employees may experience improved focus, job satisfaction, and productivity.

**Disadvantages of predicting mental health illness of working professionals using machine learning**:

* Data limitations: Availability of comprehensive and accurate mental health data is a significant challenge. Limited or biased datasets can impact the performance and generalizability of machine learning models.
* Privacy concerns: Collecting and analyzing personal mental health information raises privacy concerns. Proper safeguards must be in place to protect the confidentiality and anonymity of individuals.
* Ethical considerations: The use of machine learning in mental health prediction raises ethical questions. Responsible handling of sensitive information, informed consent, and potential implications on individuals' employment or insurance can be complex issues to navigate.
* Interpretability and trust: Machine learning models often lack interpretability, making it challenging to understand the reasoning behind predictions. This can hinder the trust and acceptance of these models by mental health professionals and individuals seeking help.
* False positives and false negatives: Machine learning models may produce false positives (identifying mental health issues where none exist) or false negatives (failing to identify actual mental health issues). This can have significant consequences, including unnecessary interventions or missed opportunities for support.
* Over-Reliance on Technology: Relying solely on machine predictions may lead to a neglect of the human aspect. Emotional intelligence and human judgment should complement the insights provided by the models.
* Bias in Data: Machine learning models are only as good as the data they are trained on. If the data used to train the model is biased, it may result in biased predictions, potentially exacerbating existing disparities.

In conclusion, while machine learning can offer valuable insights into predicting mental health issues in working professionals, it's crucial to approach it with a careful balance of ethical considerations, transparency, and a human-centered approach to ensure the well-being and privacy of the individuals involved.

**11.CONCLUSION:**

As there are many available techniques of machine learning, it is very important to compare those techniques and then identify the best among them that will suit the domain of interest. Nowadays, we have many special programs in the medical field that predict disease very accurately in advance so that treatment can be done effectively and efficiently. In this proposed work we have compared five different techniques of machine learning which are used to classify the dataset on various problems of mental health. It is very clear from the results that all the five machine learning techniques give more accurate results. The accuracy of all the classifiers are above 80%. The data set used in the research is very minimal and in the future, a large data set can be used and the research can be applied on the same for more accuracy.

**12.FUTURE SCOPE**

The future scope for predicting mental health issues among working professionals using machine learning is vast and promising. Here are some potential avenues for growth and development:

**1. \*Refinement of Models:\***

- Continuous improvement of machine learning models through feedback loops and incorporating more diverse datasets can enhance the accuracy of predictions.

- Integration of advanced algorithms, such as deep learning, for more nuanced analysis of complex patterns and subtle indicators of mental health.

**2. \*Personalized Interventions:\***

- Develop strategies to provide personalized interventions based on individual profiles. This could include tailoring support mechanisms, resources, and wellness programs to address specific needs.

**3. \*Real-time Monitoring:\***

- Move towards real-time monitoring of mental health indicators, allowing for immediate intervention when signs of distress are detected.

- Integration with wearable devices and IoT technology to capture and analyze real-time data for a more comprehensive understanding of an individual's mental well-being.

**4. \*Incorporating Multimodal Data:\***

- Combine various data sources, such as text analysis of communication patterns, physiological data from wearables, and behavioral analytics, to create a more holistic view of mental health.

**5. \*Human-AI Collaboration:\***

- Explore ways in which machine learning models can collaborate with mental health professionals to provide a more comprehensive assessment and support system.

- Develop tools that augment rather than replace human judgment, fostering a synergistic approach to mental health care.

**6. \*Ethical Considerations and Privacy:\***

- Strengthen ethical guidelines and privacy measures to ensure that the use of machine learning in mental health prediction remains transparent, secure, and respects individuals' privacy rights.

**7. \*Global Implementation:\***

- Expand the application of these models globally, considering cultural and regional differences in how mental health is perceived and addressed.

- Collaborate with international organizations and healthcare systems to create standardized, yet adaptable, models.

**8. \*Longitudinal Studies:\***

- Conduct longitudinal studies to track the long-term impact of interventions and refine models based on ongoing feedback and real-world outcomes.

**9. \*Public Awareness and Education:\***

- Promote awareness about the benefits and limitations of using machine learning in mental health. Educate both professionals and the general public to reduce stigma and foster trust in these predictive systems.

**10. \*Integration with Corporate Culture:\***

- Work closely with organizations to seamlessly integrate mental health prediction models into their corporate culture, ensuring that employees feel comfortable and supported.

The future of predicting mental health issues using machine learning involves a dynamic and evolving landscape, with a focus on improving accuracy, personalization, and ethical considerations. It holds the potential to revolutionize how we approach mental health care in the workplace and beyond.

**13.APPENDIX**

Source Code

GitHub & Project Demo Link