PREDICTING MENTAL HEALTH ILNESS OF WORKING PROFESSIONALS

MACHINE LEARNING PROJECT REPORT

TEAM ID || 594263

PRANJAY SEKSARIA || 21BCB0156

ESHA KASHYAP || 21BCE0875

ANIKET KAILAS KHAKRE || 21BCE2636

ANII SHAKYA || 21BCE3062

ACKNOWLEDGEMENTS

A successful project is a culmination of the efforts by many people, some directly and others who have encouraged and supported us from the background. We are deeply indebted to these people.

We would like to express our gratitude to VIT University and Smart Internz for this opportunity. We would also like to thank the Smart Internz counselors and teachers for their constant guidance and patient help.

Throughout this project, they have answered our queries and have been instrumental in its completion.

INDEX

INTRODUCTION

In a world that full of bright distractions and the deafening cacophony of notifications, our mental well-being often takes a back seat. But what if technology could be a bridge to a healthier mind? This is the driving force behind our journey—a mission to bring mental health to the forefront of our digital landscape.

Behind every virtual screen, there is a real person with a real story.

1.1] Project Overview:

The Corporate World and large Tech Giants are infamous for their treatment of employees that might be the leading cause of mental health degradation. The sudden surge in depression, alcoholism, eating disorders, etc cannot all be attributed to them, but a better workplace and more sensitive colleagues and management could definitely go a long way.

Our model uses Machine Learning Concepts such as supervised learning, classification, Decision Tree, KNN, etc and integrates this into a lucrative website that presents a questionnaire based on which a prediction is made.

1.2] Purpose:

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.

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. Organizations that incorporate mental health awareness help to create a healthy and productive work environment that reduces the stigma associated with mental illness, increase the organizations' mental health literacy, and teaches the skills to safely and responsibly respond to a co-worker's mental health concern.

LITERATURE SURVEY

2.1] Existing problem:

Mental health has become a forefront concern in contemporary society, with an increasing recognition of its impact on overall well-being. The ubiquity of smartphones and the growing role of technology in our lives offer an unprecedented opportunity to leverage digital tools for mental health support. However, despite the progress made in understanding and destigmatizing mental health issues, there remains a considerable gap in the timely and accurate prediction of these conditions.
The significance of mental health prediction lies in its potential to revolutionize the way we approach mental well-being. Unlike traditional methods that often rely on self-reporting or clinical assessments, predictive technologies can offer early insights, allowing for proactive interventions and personalized support. By harnessing the power of machine learning and mobile applications, we aim to create a bridge between individuals and the mental health resources they need, fostering a culture of proactive mental health management.
In this digital age, where our smartphones are extensions of ourselves, the integration of technology into mental health care is not just logical but imperative. A mental health prediction app can act as a vigilant companion, leveraging advanced algorithms to analyze subtle patterns in user behavior, emotions, and habits. By transforming smartphones into personalized mental health monitors, we aspire to empower individuals to take an active role in their emotional well-being.
Despite the potential benefits, existing barriers hinder the seamless integration of technology into mental health care. Stigma, concerns about data privacy, and the need for robust and accurate predictive models are among the challenges that must be addressed. Through this project, we aim to contribute to the ongoing discourse surrounding mental health, emphasizing the importance of technology as an ally rather than a replacement for traditional care.
As we embark on this journey, we recognize the critical need to bridge the gap between technological innovation and mental health support. Our project seeks to go beyond the surface, delving into the complexities of mental well-being to offer not just predictions but personalized insights and resources. By fostering a holistic approach to mental health care, we aspire to contribute to a future where technology is seamlessly integrated into our lives, supporting our emotional resilience and well-being.

2.2] References

1. https://www.ibm.com/topics/supervised-learning#:~:text=Supervised%20learning%2C%20also%20known%20as,data%20or%20predict%20outcomes%20accurately.

2. https://www.techtarget.com/searchenterpriseai/definition/unsupervised-learning#:~:text=Unsupervised%20learning%20is%20a%20type,are%20neither%20classified%20nor%20labeled.

3. https://www.javatpoint.com/regression-vs-classification-in-machine-learning#:~:text=The%20main%20difference%20between%20Regression,Spam%20or%20Not%20Spam%2C%20etc.

4. https://www.sciencedirect.com/topics/computer-science/logistic-regression#:~:text=Logistic%20regression%20is%20a%20process,%2Fno%2C%20and%20so%20on.

5. https://www.javatpoint.com/machine-learning-decision-tree-classification-algorithm

6. https://www.ibm.com/topics/random-forest#:~:text=Random%20forest%20is%20a%20commonly,both%20classification%20and%20regression%20problems.

7. https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

8. https://www.nvidia.com/en-us/glossary/data-science/xgboost/#:~:text=XGBoost%20is%20a%20scalable%20and,model%20performance%20and%20computational%20speed.

9. https://www.almabetter.com/bytes/tutorials/data-science/adaboost-algorithm

10. https://www.analyticsvidhya.com/blog/2021/09/gradient-boosting-algorithm-a-complete-guide-for-beginners/#:~:text=Gradient%20boosting%20is%20a%20method,has%20produced%20the%20best%20results.

11. https://www.analyticsvidhya.com/blog/2021/05/know-the-best-evaluation-metrics-for-your-regression-model/

12. https://www.youtube.com/watch?v=lj4I_CvBnt0

13. https://www.mentalhealthfirstaid.org/

14. https://www.ncbi.nlm.nih.gov/books/NBK401817/

15. https://recruitinginnovation.com/front-end-technologies/#:~:text=HTML%20is%20the%20first%20layer,stand%20for%20Hypertext%20Markup%20Language.

2.3] Problem Statement Definition

In the realm of mental health prediction, a critical void exists in the web-based landscape. Current solutions predominantly cater to mobile apps, leaving a significant portion of users untapped. This project addresses this gap by envisioning and developing a comprehensive web platform for predicting mental health conditions.

Our primary challenge involves adapting sophisticated supervised machine learning algorithms, including regression and classification models, to the unique constraints and features of web environments. One standout algorithm in our approach is the AdaBoost algorithm, chosen for its ability to enhance predictive accuracy by combining multiple weak learners.

To seamlessly integrate predictive models into the web, we leverage Flask mechanisms, ensuring a secure, responsive, and user-friendly interface. Flask enables the incorporation of advanced machine learning capabilities, empowering users to receive personalized mental health insights, resources, and support directly through their web browsers.

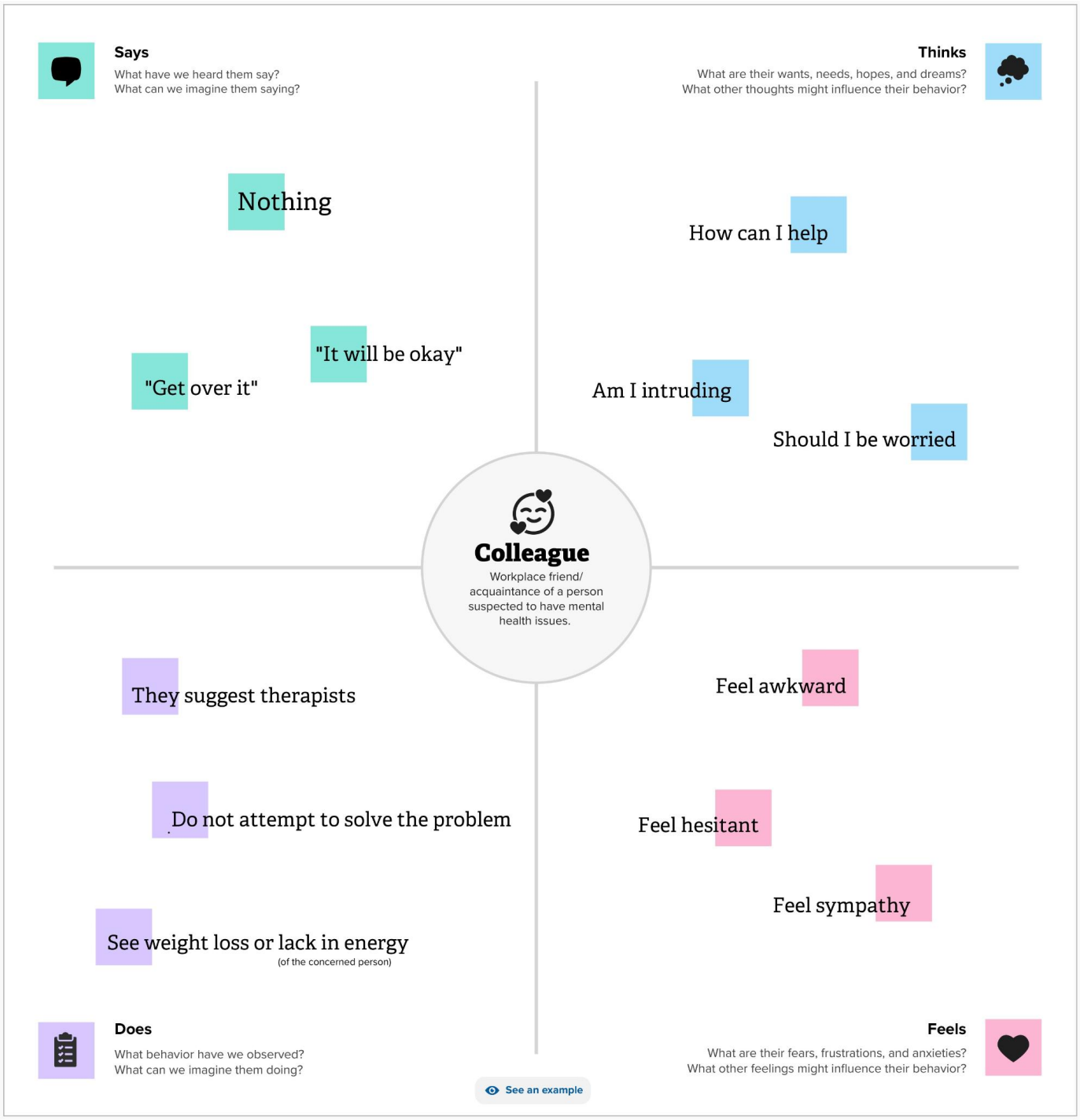
The frontend is designed using HTML, offering an intuitive and inclusive user experience. We prioritize user-centric design, aligning with ethical standards to address privacy concerns and data security. The platform utilizes a basic dataset with 21 data columns, capturing a comprehensive range of factors influencing mental health. Through this initiative, we aim to revolutionize how individuals engage with mental health prediction, fostering a proactive and accessible approach to emotional well-being on the web.

Final Problem Statement: A model to predict the risk level (regression) of mental health problems in a colleague and provide subsequent solutions for each range.

IDEATION & PROPOSED SOLUTION

3.1] Empathy Map Canvas

An empathy map is a simple, easy-to-digest visual that captures knowledge about a user's behaviours and attitudes.
It is a useful tool to helps teams better understand their users.
Creating an effective solution requires understanding the true problem and the person who is experiencing it. The exercise of creating the map helps participants consider things from the user's perspective along with his or her goals and challenges.



3.2] Ideation & Brainstorming

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.

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

Problem Statement: 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. Organizations that incorporate mental health awareness help to create a healthy and productive work environment that reduces the stigma associated with mental illness, increase the organizations' 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.

Step-2: Brainstorm, Idea Listing and Grouping

1. Employee Mental Health Monitoring: A machine learning system to continuously monitor and assess the mental health of employees within an organization, identifying early warning signs of stress or burnout.
2. Mental Health Resource Recommendation: A recommendation system that suggests relevant mental health resources (e.g., articles, self-help tools, or support groups) to employees based on their self-reported mental health status or specific concerns.
3. Mental Health Chatbot: A conversational AI chatbot that provides employees with a confidential and supportive platform to discuss their mental health concerns and receive guidance or referrals to appropriate resources.
4. Mood Prediction for Workforce Scheduling: Use of machine learning to predict employee mood and productivity based on factors like workload, work hours, and personal preferences, enabling optimized workforce scheduling to reduce stress.
5. Mental Health Awareness Training: A machine learning-enhanced training program that identifies employees who would benefit from mental health awareness training and tailors the training to their specific needs.
6. Peer Support Matching: A system that matches employees in need of mental health support with peer mentors or supporters based on compatibility, ensuring a network of emotional support within the organization.
7. Anonymous Reporting System for Mental Health Issues: A platform that allows employees to anonymously report mental health concerns and incidents to HR or management, with machine learning-based analytics to identify trends and patterns in reported issues.
8. Mental Health Feedback Loops: A system that collects feedback from employees about their experiences with mental health support programs and uses machine learning to iteratively improve and customize these programs.

Step-3: Idea Prioritization

Final Problem Statement: A model to predict the risk level (regression) of mental health problems in a colleague and provide subsequent solutions for each range.

REQUIREMENT ANALYSIS

4.1] Functional requirement

1. User Registration and Authentication: Users should be able to create accounts securely.

2. User Profile: Users should be able to create and update their profiles. The profile should include relevant information for mental health prediction.

3. Input Data Collection: The app must collect user input data related to emotions, behaviors, and habits. Data collection should be user-friendly and non-intrusive.

4. Machine Learning Model: Implement a machine learning algorithm to analyze user input. The model should predict potential mental health issues based on input data.Educational Resources.

5. Feedback Mechanism: Allow users to provide feedback on the accuracy of predictions. Use feedback to continually improve the machine learning model.

4.2] Non-Functional requirements

1. Security: Implement robust security measures to protect user data. Comply with data protection regulations.

2. Scalability: Design the app to handle a growing user base. Ensure scalability for increased data and user activity.

3. Performance: The app should respond promptly to user actions. Minimize latency in data processing and prediction.

4. Usability: Create an intuitive and user-friendly interface. Ensure accessibility for users with different abilities.

5. Reliability: The app should be available and reliable under normal operating conditions. Implement backup and recovery mechanisms.

6. Regulatory Compliance: Comply with relevant regulations and standards in the mental health and healthcare industry.

7. Data Privacy: Safeguard user privacy and ensure compliance with data protection laws.Clearly communicate the app's privacy policy to users.

8. Performance Monitoring: Implement tools for monitoring app performance. Track system usage and identify areas for improvement.

9. Documentation: Provide comprehensive documentation for developers, administrators, and users.

PROJECT DESIGN

5.1] Data Flow Diagrams & User Stories

DATA FLOW DIAGRAMS:

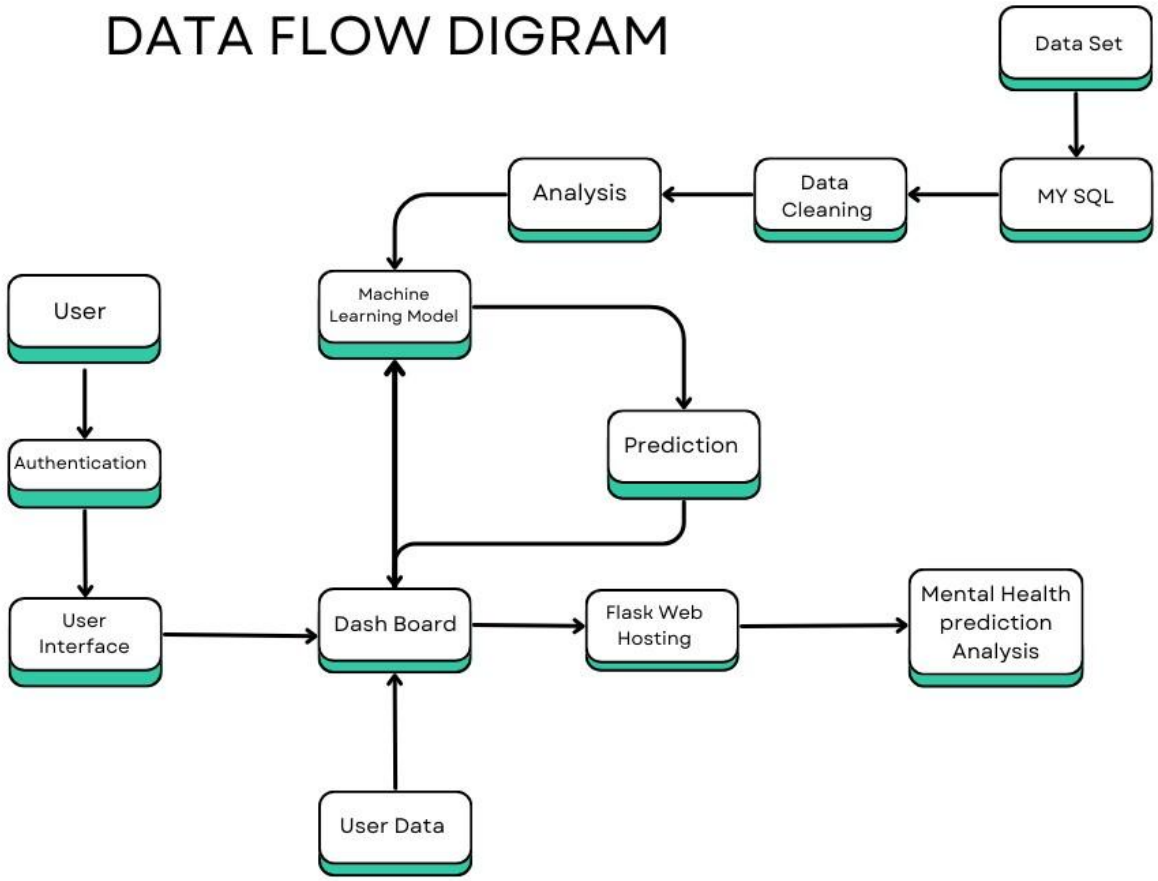
A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right

amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is

stored.

FLOW:

1. User configures credentials for login and starts the application.
2. On opening the app, our attractive user interface keeps the user engaged.
3. User inputs the data based on the questionnaire.
4. The predefined dataset is loaded into SQL.
5. Data cleaning and analysis is performed on the dataset and fed into the machine learning algorithm.
6. The trained algorithm receives new information inputted by the user.
7. The model makes a prediction based on previous learning
8. This prediction is displayed on the web application.



USER STORIES:

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
	Dashboard	USN -1	As a user I can navigate through and provide input in the questionnaire.	I can successfully submit my answers	High	Sprint -1
		USN -2	As a user, I can see solutions for my mental health based on the level predicted by the application	I can view an accurate solution.	High	Sprint -1
Counselor	Dashboard	USN-1	As a counselor, I can see relevant user data and review their reports.	I can access verified user data.	High	Sprint -2
			As a counselor, I can provide solutions based on score predictions	I can input effective solutions into the application.	Medium	Sprint - 2

5.2] Solution Architecture

Solution architecture is a complex process – with many sub-processes – that bridgesthe 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 thesoftware to project stakeholders.

Define features, development phases, and solution requirements.

Provide specifications according to which the solution is defined, managed,and delivered.

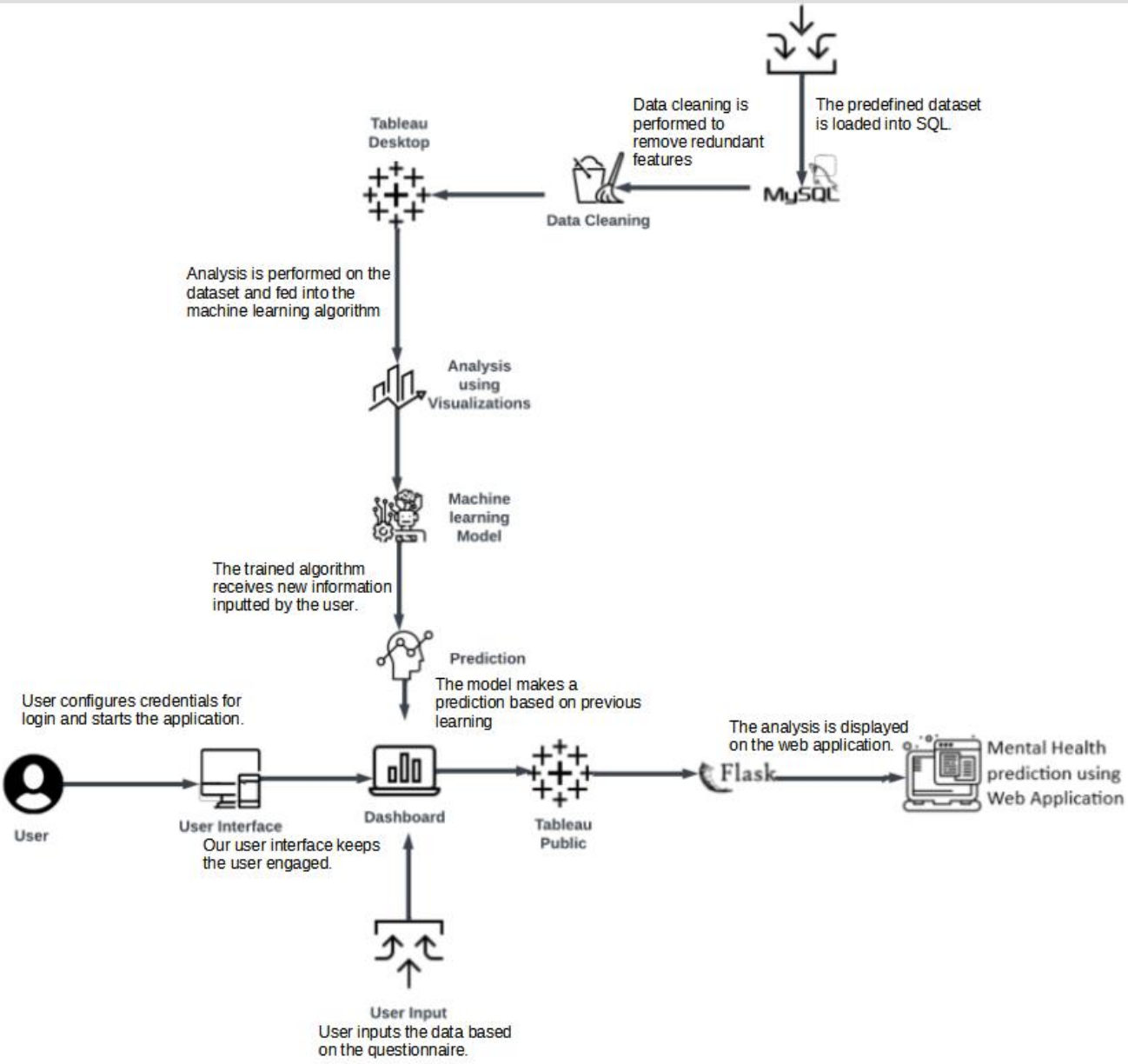


Figure 1: Architecture and data flow of the psycare sample web application

PROJECT PLANNING & SCHEDULING

6.3] Technical Architecture

Technical architecture is a foundational and systematic approach to designing and organizing complex systems, including software applications
and network infrastructures, with the primary objective of achieving specific technical goals. It is a crucial component in the realm of information
technology, playing a pivotal role in ensuring the successful implementation of projects. Technical architects, often serving as the guiding
visionaries, are responsible for overseeing the design and construction of these systems, ensuring they operate efficiently, maintain a high level

of security, and are in perfect alignment with the broader objectives of the project.

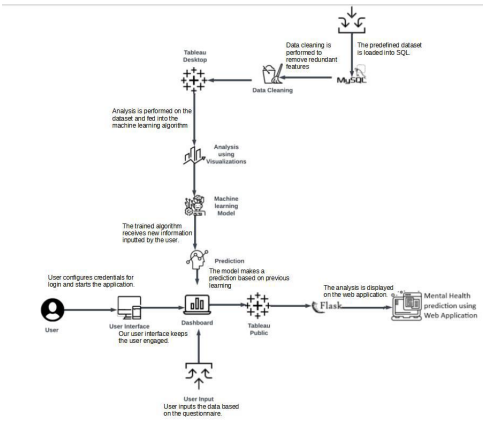


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	How the user interacts with web application.	HTML, CSS, JavaScript, React Js etc.
2.	User Authentication	Authentication and security	Firebase, OAuth2
3.	Database	Data Type, Configurations etc.	MySQL
4.	Infrastructure (Server)	Application Deployment on Local System	Local, Cloud Foundry, Kubernetes, etc.
5.	Data collection	Gather a diverse dataset of 1200 employees from different regions of the country across various age groups.	Dataset collection
6.	Data preprocessing	Cleaning, visualization, removal of nulls and redundant data. Removing stop words, tokenizing data and handling special characters.	NumPy, seaborn, etc.
7.	Model Architecture	Utilize word embeddings to convert words into numerical representations suitable for deep learning. Implement an RNN-based model which can capture sequential information in the text.	Word2Vec, GloVe, PyTorch
8.	Training	Train the model on the preprocessed dataset, utilizing GPU/CPU resources, and fine-tune model weights to minimize classification errors.	GPU/CPU, Deep Learning Frameworks
9.	Evaluation	Assess the model's performance and accuracy through metrics.	Metrics such as F1 score/ correlation
10.	Deployment	Integrate the trained model into applications for solutions, involving web app development for user accessibility.	Flask, python

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	List the open-source frameworks used	React JS, Node JS, flask
2.	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	Firebase, OAuth2
3.	Scalable Architecture	Designing the app's architecture for scalability, easy maintenance, and future updates.	Model-View-ViewModel (MVVM) architecture: Separation of concerns for easier scalability. Repository pattern: Isolating data access for flexibility and testability. Dependency Injection: Dagger Hilt for managing dependencies and scalability
4.	Availability	Ensuring the web app is available to users with minimal downtime	Load Balancing: Distributing incoming network traffic to maintain availability. Server redundancy: Multiple server instances to ensure availability in case of server failures.
5.	Performance	Optimizing the web app's speed and responsiveness	Caching: Use of local caching to reduce network requests. Image Compression: Compressing and optimizing images for faster loading. Background Tasks: Implementing background processing to keep the app responsive. Profiling and optimization tools: Android Profiler and tools for performance analysis and improvement

6.4] Sprint Planning & Estimation

Product Backlog, Sprint Schedule, and Estimation:

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password	4	High	Anii Shakya
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	4	High	Aniket Khakre
Sprint-2	Dashboard	USN-3	As a user I can navigate through and provide input in the questionnaire.	6	High	Esha Kashyap
Sprint-2		USN-4	As a user, I can see solutions for my mental health based on the level predicted by the application application through Gmail	10	High	Pranjay Seksaria
Sprint-3	Dashboard	USN-5	As a counselor, I can see relevant user data and review their reports.	4	Medium	Pranjay Seksaria
Sprint-3			As a counselor, I can provide solutions based on score predictions	4	Medium	Anii Shakya

6.5] Sprint Delivery Schedule

Project Tracker, Velocity & Burndown Chart:

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	8	4 Days	24 Oct 2023	27 Oct 2023		
Sprint-2	16	8 Days	28 Oct 2023	04 Nov 2023		
Sprint-3	8	2 Days	05 Nov 2023	06 Nov 2023		

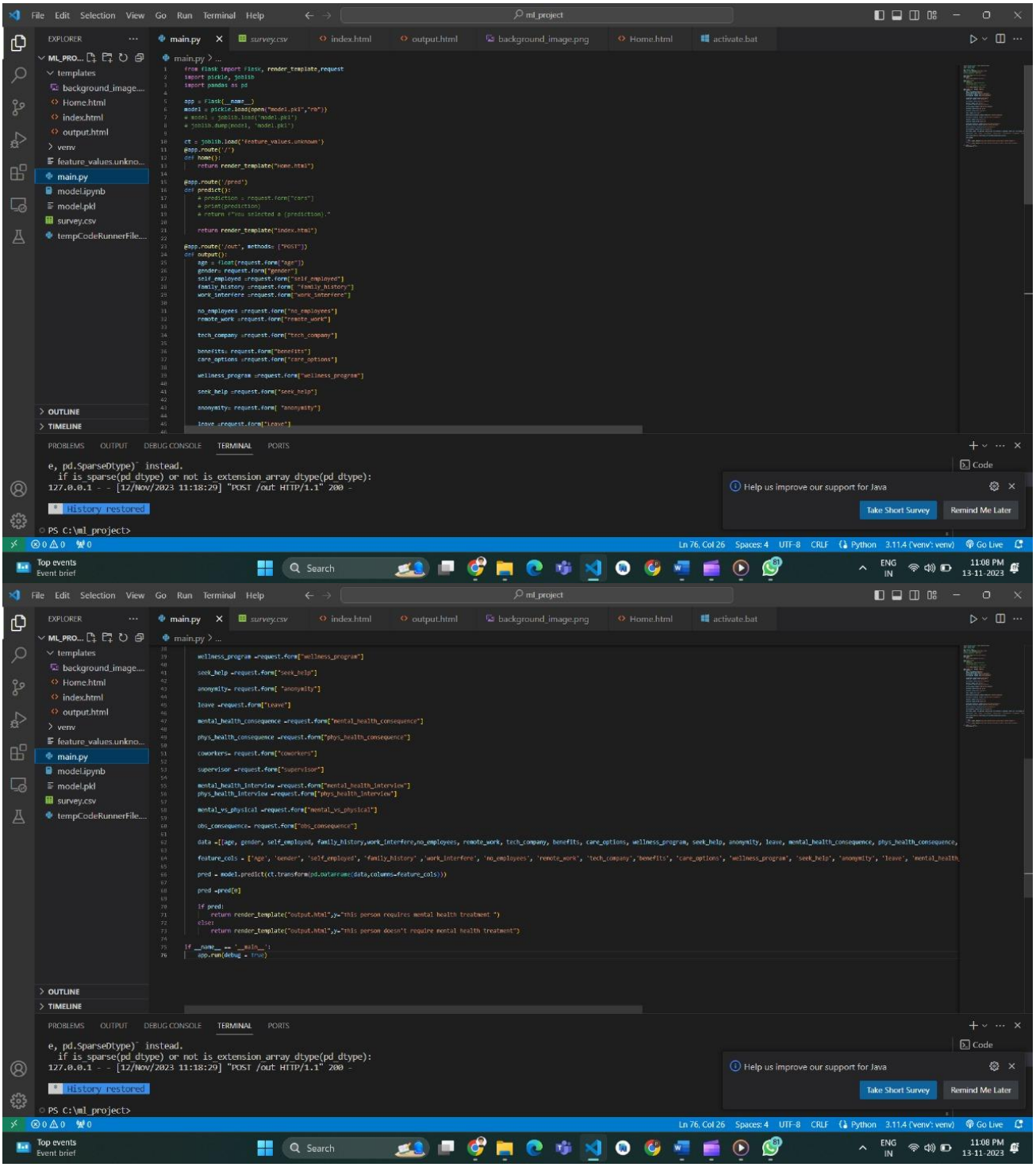
Velocity:
AV = 32/14 = 2.3

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

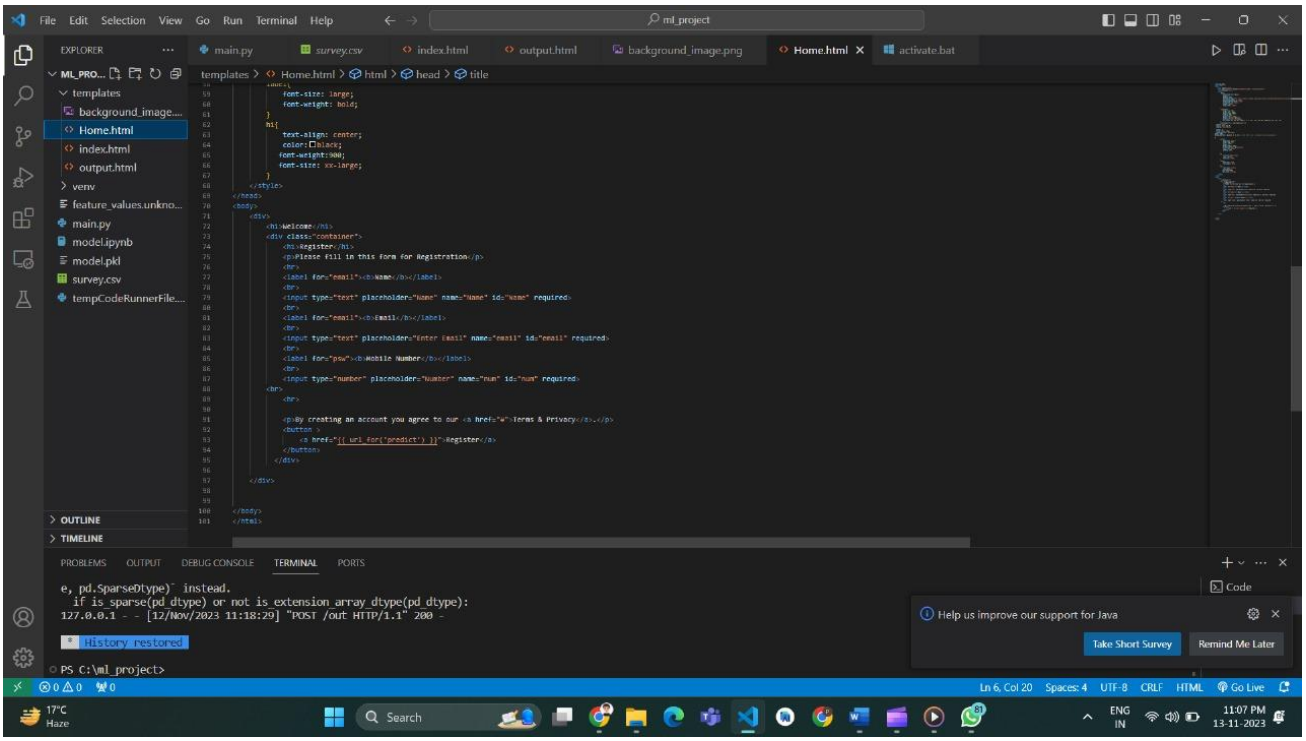
as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.

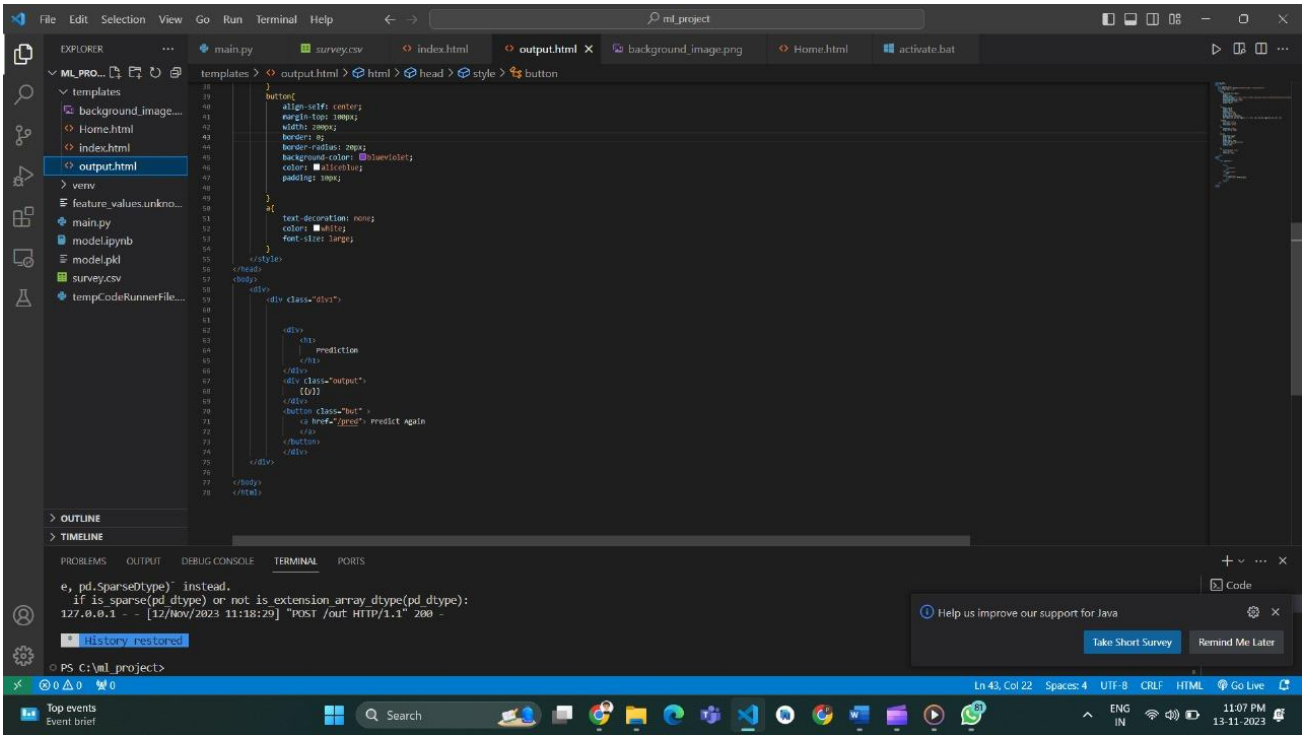
CODING & SOLUTIONING

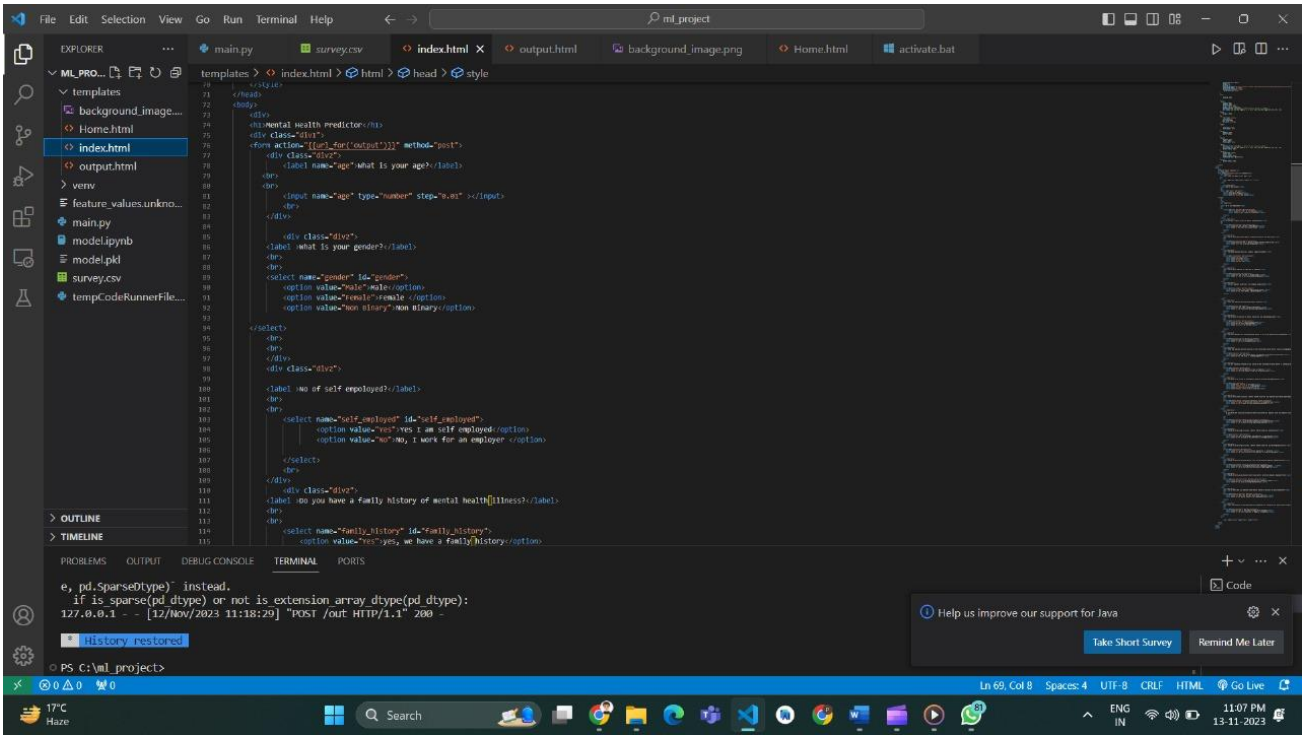
7.1] Feature 1:

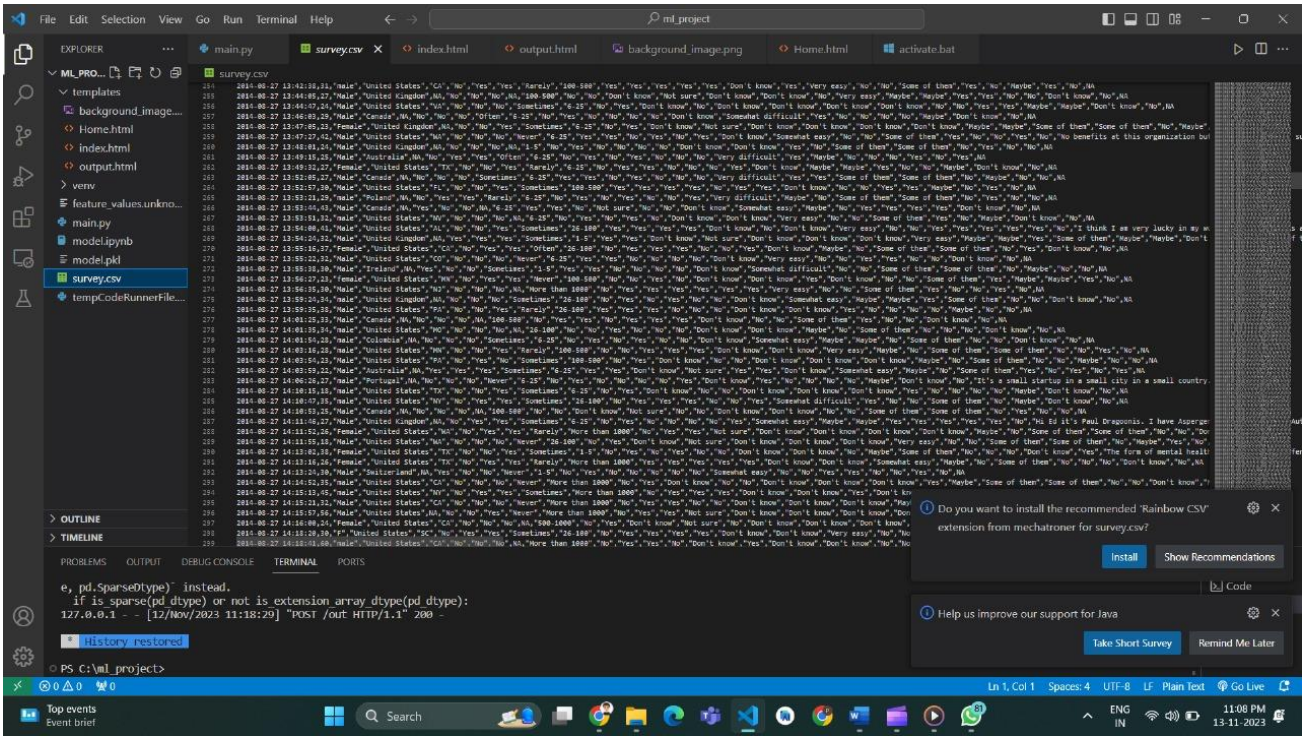


7.2] Feature 2:


File Edit Selection View Go Run Terminal Help
ml_project
EXPLORER
main.py
survey.csv
index.html
output.html
background_image.png
Home.html
activate.bat
templates > Home.html > html > head > title
templates
background_image....
Home.html
index.html
output.html
venv
feature_values.unkno...
main.py
model.ipynb
model.pkl
survey.csv
tempCodeRunnerFile....
OUTLINE
TIMELINE
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
e, pd.SparseDtype)" instead.
if is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
127.0.0.1 - - [12/Nov/2023 11:18:29] "POST /out HTTP/1.1" 200 -
History restored
PS C:\ml_project>
Help us improve our support for Java
Take Short Survey
Remind Me Later
Ln 6, Col 20 Spaces: 4 UTF-8 CRLF HTML Go Live
17°C
Haze
Search
ENG
IN
11:07 PM
13-11-2023


File Edit Selection View Go Run Terminal Help
ml_project
EXPLORER
main.py
survey.csv
index.html
output.html
background_image.png
Home.html
activate.bat
templates > output.html > html > head > style > button
templates
background_image....
Home.html
index.html
output.html
venv
feature_values.unkno...
main.py
model.ipynb
model.pkl
survey.csv
tempCodeRunnerFile....
OUTLINE
TIMELINE
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS
e, pd.SparseDtype)" instead.
if is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
127.0.0.1 - - [12/Nov/2023 11:18:29] "POST /out HTTP/1.1" 200 -
History restored
PS C:\ml_project>
Help us improve our support for Java
Take Short Survey
Remind Me Later
Ln 43, Col 22 Spaces: 4 UTF-8 CRLF HTML Go Live
Top events
Event brief
Search
ENG
IN
11:07 PM
13-11-2023


File Edit Selection View Go Run Terminal Help
ml_project
EXPLORER
main.py
survey.csv
index.html
output.html
background_image.png
Home.html
activate.bat
templates > index.html > html > head > style
ML_PRO...
templates
background_image....
Home.html
index.html
output.html
venv
feature_values.unkno...
main.py
model.ipynb
model.pkl
survey.csv
tempCodeRunnerFile....
OUTLINE
TIMELINE
PROBLEMS
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
e, pd.SparseDtype)" instead.
if is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
127.0.0.1 - - [12/Nov/2023 11:18:29] "POST /out HTTP/1.1" 200 -
History restored
PS C:\ml_project>
Help us improve our support for Java
Take Short Survey
Remind Me Later
Ln 69, Col 8
Spaces: 4
UTF-8
CRLF
HTML
Go Live
17°C
Haze
Search
ENG
IN
11:07 PM
13-11-2023


File Edit Selection View Go Run Terminal Help
ml_project
EXPLORER
main.py
survey.csv
index.html
output.html
background_image.png
Home.html
activate.bat
survey.csv
ML_PRO...
templates
background_image....
Home.html
index.html
output.html
venv
feature_values.unkno...
main.py
model.ipynb
model.pkl
survey.csv
tempCodeRunnerFile....
OUTLINE
TIMELINE
PROBLEMS
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
e, pd.SparseDtype)" instead.
if is_sparse(pd_dtype) or not is_extension_array_dtype(pd_dtype):
127.0.0.1 - - [12/Nov/2023 11:18:29] "POST /out HTTP/1.1" 200 -
History restored
PS C:\ml_project>
Do you want to install the recommended 'Rainbow CSV' extension from mechatroner for survey.csv?
Install
Show Recommendations
Help us improve our support for Java
Take Short Survey
Remind Me Later
Ln 1, Col 1
Spaces: 4
UTF-8
LF
Plain Text
Go Live
Top events
Event brief
Search
ENG
IN
11:08 PM
13-11-2023

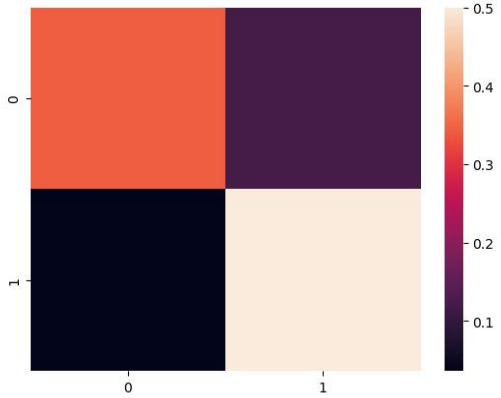
PERFORMANCE TESTING

8.1] Performace Metrics:

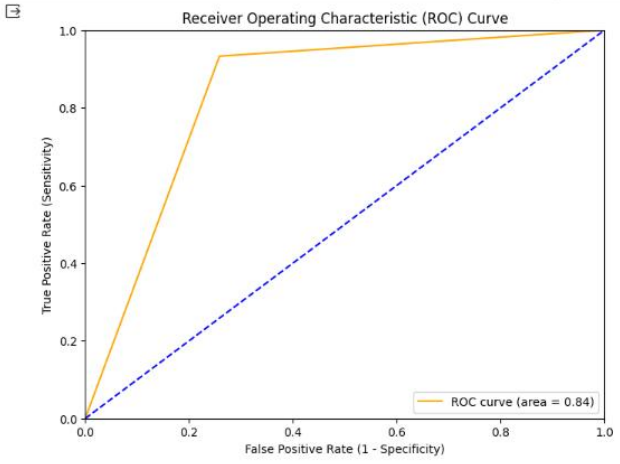
S.No.	Parameter	Values	Screenshot
1.	Metrics	**Classification Model:** Confusion Matrix - , Accuray Score- & Classification Report -	
1.	Tune the Model	Hyperparameter Tuning - Validation Method -	86.8% accuracy

```
import seaborn as sns
cf_matrix= confusion_matrix(y_test, pred_abc)
sns.heatmap(cf_matrix/np.sum(cf_matrix))
```

```
<Axes: >
```




```
from sklearn.metrics import roc_curve, auc
import matplotlib.pyplot as plt

fpr_abc, tpr_abc, thresholds_abc = roc_curve(y_test, pred_abc)
roc_auc_abc = auc(fpr_abc, tpr_abc)

plt.figure(figsize=(8, 6))
plt.plot(fpr_abc, tpr_abc, color='orange', label='ROC curve (area = {:.2f})'.format(roc_auc_abc))
plt.plot([0, 1], [0, 1], color='blue', linestyle='--')
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.0])
plt.title('Receiver Operating Characteristic (ROC) Curve')
plt.xlabel('False Positive Rate (1 - Specificity)')
plt.ylabel('True Positive Rate (Sensitivity)')
plt.legend(loc='lower right')
plt.show()
```



```
[64] from sklearn.model_selection import RandomizedSearchCV
     import numpy as np

     params_abc = {
         'n_estimators': [int(x) for x in np.linspace(start=1, stop=50, num=15)],
         'learning_rate': [(0.97 + x/100) for x in range(0, 8)],
     }

     abc_random = RandomizedSearchCV(estimator=abc, param_distributions=params_abc, n_iter=50, cv=5, n_jobs=-1, random_state=49)
```

```
[65] abc_random.fit(x_train, y_train)
```

RandomizedSearchCV
estimator: AdaBoostClassifier
AdaBoostClassifier

```
abc_random.best_params_
```

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

```
[67] from sklearn.ensemble import AdaBoostClassifier
     from sklearn.metrics import accuracy_score
     abc_tuned = AdaBoostClassifier(random_state=49, n_estimators=11, learning_rate=1.02)
     abc_tuned.fit(x_train, y_train)
     pred_abc_tuned = abc_tuned.predict(x_test)
     print('Accuracy of AdaBoost (tuned) =', accuracy_score(y_test, pred_abc_tuned))

     Accuracy of AdaBoost (tuned) = 0.868
```

```
[77] print(classification_report(y_test, pred_abc))
```

	precision	recall	f1-score	support
0	0.91	0.74	0.82	116
1	0.81	0.93	0.87	134
accuracy			0.84	250
macro avg	0.86	0.84	0.84	250
weighted avg	0.85	0.84	0.84	250

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations - 4
2.	Data Responsiveness	High
3.	Amount Data to Rendered (DB2 Metrics)	21
4.	Utilization of Data Filters	4 columns were filtered
5.	Effective User Story	3 pages

6.	Descriptive Reports	Data visualizations, bivariate, univariate and multivariate analysis were performed

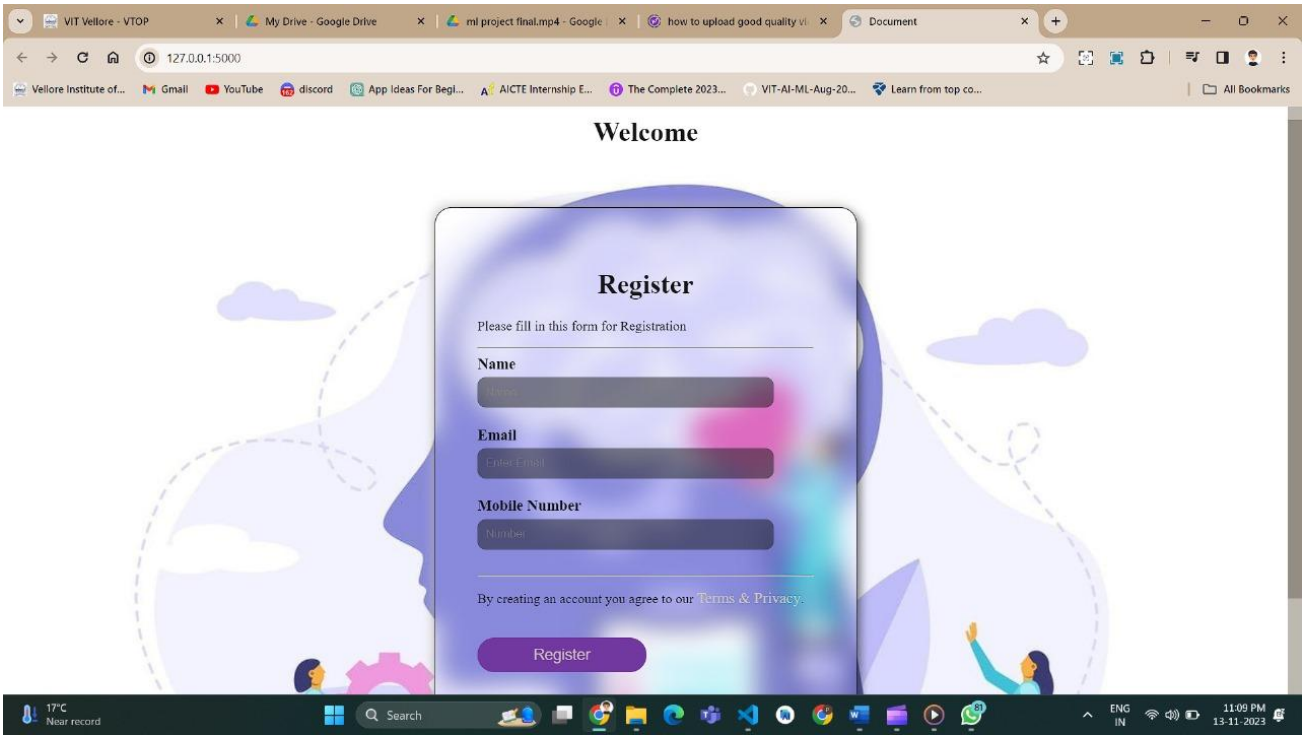
S.No.	Parameter	Values	Screenshot
1.	Model Summary	AdaBoost classifier	
2.	Accuracy	Training Accuracy - Validation Accuracy -	84.66 86.7

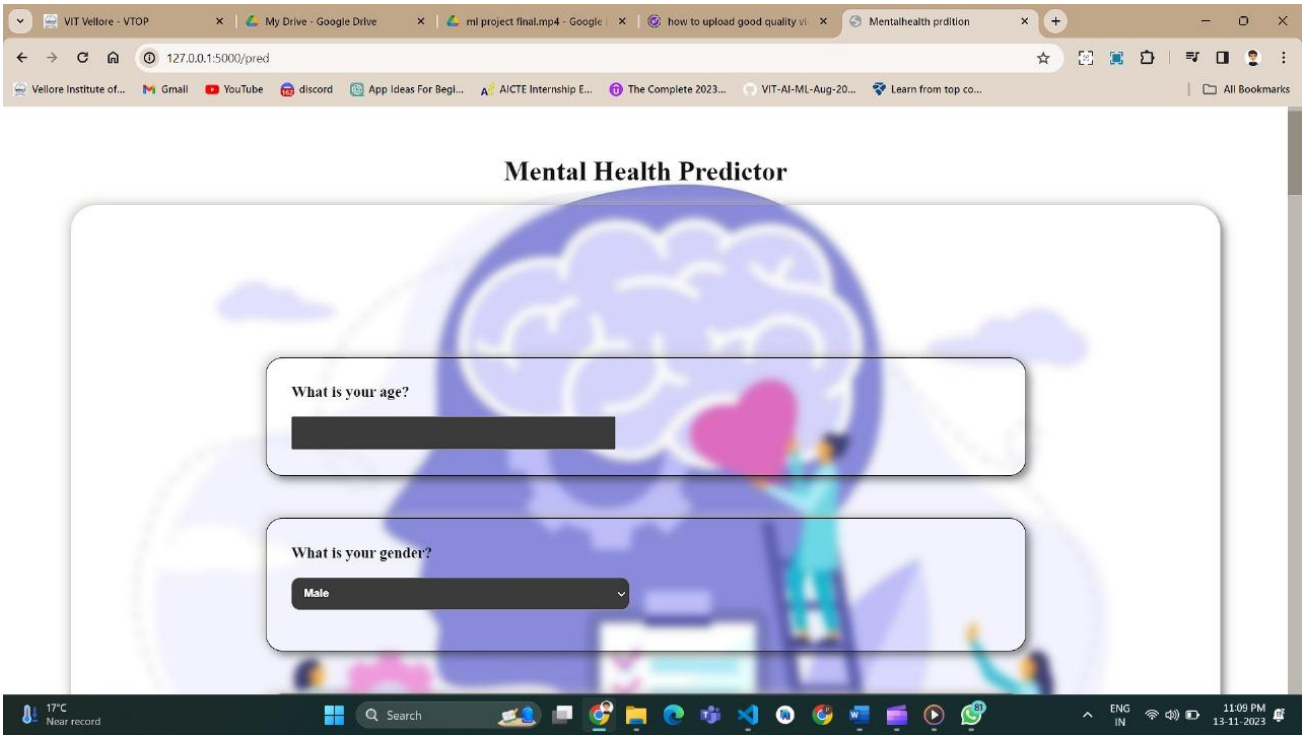
```
[63] from sklearn.ensemble import AdaBoostClassifier
     from sklearn.metrics import accuracy_score

     # Assuming you have X_train and y_train for training and x_test and y_test for testing
     abc = AdaBoostClassifier(random_state=99)
     abc.fit(x_train, y_train)
     pred_abc = abc.predict(x_test)
     print(f'Accuracy of AdaBoost = {accuracy_score(y_test, pred_abc) * 100}%', )
```

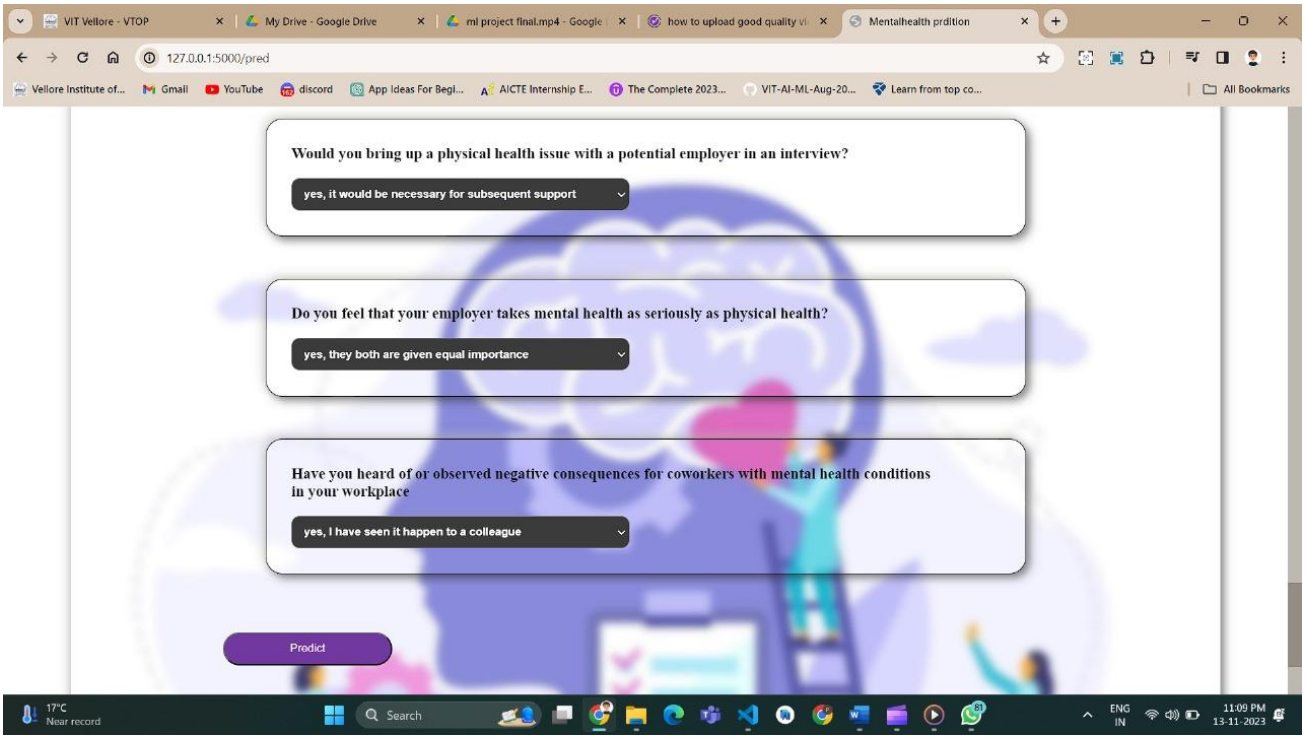
```
Accuracy of AdaBoost = 84.39999999999999%
```

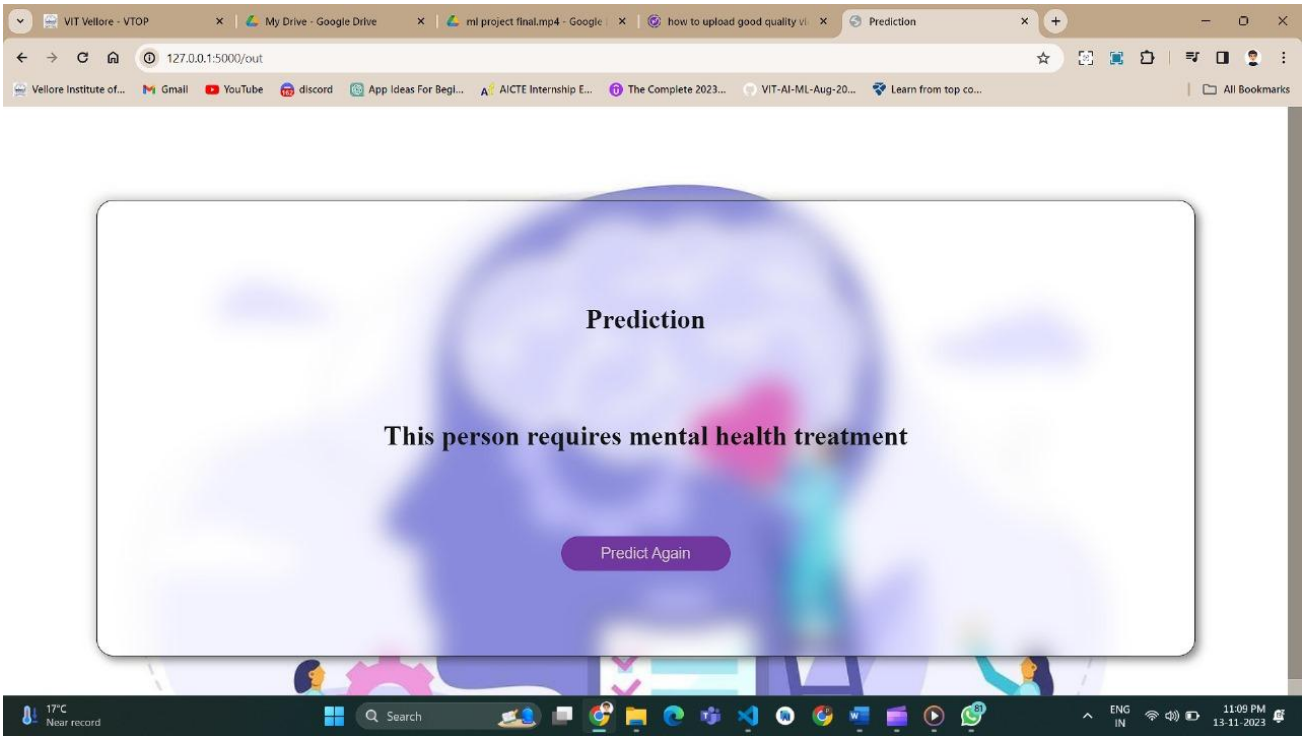
RESULTS

9.1] Output Screenshots:






Mentalhealth prdition
127.0.0.1:5000/pred
Would you bring up a physical health issue with a potential employer in an interview?
yes, it would be necessary for subsequent support
Do you feel that your employer takes mental health as seriously as physical health?
yes, they both are given equal importance
Have you heard of or observed negative consequences for coworkers with mental health conditions in your workplace
yes, I have seen it happen to a colleague
Predict


Prediction
127.0.0.1:5000/out
Prediction
This person requires mental health treatment
Predict Again

ADVANTAGES & DISADVANTAGES

10.1] Advantages:

1. Accessibility: Advantage: Web-based platforms are accessible from various devices with internet connectivity, allowing users to engage with mental health support using their preferred devices, including computers and tablets.
2. Inclusivity: Advantage: A web-based approach caters to a broader audience, reaching individuals who may not use or have access to mobile applications, thereby promoting inclusivity in mental health care.
3. User Engagement: Advantage: Web platforms, with their larger screens and interactive interfaces, can facilitate a more immersive and engaging user experience, potentially enhancing user engagement with mental health features.
4. Scalability: Advantage: Web-based solutions are often easier to scale and update, allowing for continuous improvements in predictive models, user interfaces, and additional features.
5. Integration with Existing Services: Advantage: Integration with existing web-based services and platforms becomes more seamless, enabling collaboration with other healthcare providers, telehealth services, and support networks.
6. Flexible User Interface: Advantage: HTML-based frontend allows for a flexible and customizable user interface, ensuring adaptability to different user preferences and needs.
7. Global Reach: Advantage: Web platforms have the potential for a global reach, transcending geographical boundaries and making mental health prediction accessible to a diverse international audience.

10.2]Disadvantages:

1. Limited Offline Functionality: Disadvantage: Web platforms may have limited functionality in offline environments, potentially restricting user access in areas with poor or no internet connectivity.
2. Device Compatibility: Disadvantage: While web platforms are versatile, ensuring compatibility across various devices and web browsers can be challenging, potentially leading to inconsistencies in user experience.
3. Security Concerns: Disadvantage: Web-based platforms may face security challenges, necessitating robust measures to protect sensitive mental health data and maintain user confidentiality.
4. Learning Curve: Disadvantage: Users who are not familiar with web-based technologies may experience a learning curve, potentially impacting their ability to navigate and utilize the platform effectively.
5. Dependence on Internet Connectivity: Disadvantage: The effectiveness of the platform relies heavily on consistent internet connectivity, limiting its usability in areas with unreliable or no internet access.
6. Potential for Information Overload: Disadvantage: The web environment may present the risk of information overload, where users may be overwhelmed by the volume of data and insights provided, affecting the overall user experience.
7. Ethical and Privacy Concerns: Disadvantage: The collection and storage of sensitive mental health data raise ethical and privacy concerns. Ensuring compliance with privacy regulations and maintaining user

trust is paramount.

8. Technical Challenges: Disadvantage: Developing and maintaining a robust web-based platform with advanced machine learning capabilities can pose technical challenges, requiring ongoing updates and technical support.

CONCLUSION

In the culmination of our project, we've charted new territory in the intersection of mental health care and digital innovation. By pioneering a web-based mental health prediction platform, we've addressed the critical gap in the accessibility of predictive tools for a diverse audience.

Harnessing the power of supervised machine learning, including regression and classification models, and deploying the potent AdaBoost algorithm, we've elevated predictive accuracy and personalized insights. The integration of Flask mechanisms ensures a secure, responsive, and user-friendly interface, seamlessly delivering advanced machine learning capabilities through web browsers.

Our commitment to user-centric design, as embodied in the HTML frontend, ensures an inclusive and intuitive experience. Throughout development, we've adhered to ethical standards, prioritizing user privacy and data security. The utilization of a basic dataset with 21 data columns allows for a comprehensive understanding of the myriad factors influencing mental health.

As we conclude, this project signifies not just a technological evolution but a paradigm shift in mental health care. It marks the inception of a platform that empowers individuals to proactively manage their emotional well-being in the digital age. We anticipate that our web-based solution will not only redefine how users engage with mental health prediction but also set a precedent for the ethical, secure, and user-centric development of digital mental health solutions. In the journey ahead, we remain dedicated to refining and expanding our platform, contributing to a future where mental health support is as accessible and personalized as the click of a button.

FUTURE SCOPE

The future scope of a machine learning or data analysis project can be quite extensive, offering opportunities for scaling and diverse applications.

1. Scalability for Big Data: As data volumes grow, adapting the project to handle big data scenarios becomes crucial. Employing distributed computing frameworks like Apache Spark or integrating with big data platforms can enhance scalability.

2. Advanced Machine Learning Models: Explore more sophisticated machine learning models and algorithms. As new methods and techniques emerge, integrating them into your project can improve prediction accuracy and robustness.

3. Real-Time Analytics: Extend the project to provide real-time analytics. This could involve implementing stream processing systems to analyze and derive insights from data as it arrives.

4. Integration with IoT Devices: If applicable, consider integrating with Internet of Things (IoT) devices. This can provide real-time data from sensors and devices, enhancing the project's capabilities and use cases.

5. Multi-Modal Data Analysis: Expand the project to analyze multiple types of data, such as text, images, and audio. This could involve implementing multi-modal machine learning models for a more comprehensive understanding.

6. Enhanced User Interaction: Improve the project's user interface and interaction features. Consider implementing features like interactive visualizations, natural language processing for user queries, or incorporating virtual/augmented reality elements.

7. Cross-Platform Compatibility: Make the project more versatile by ensuring compatibility across different platforms (web, mobile, desktop). This allows users to access and interact with the project seamlessly from various devices.

8. Collaboration with Other Systems: Explore opportunities for collaboration with other systems or platforms. This could involve integrating with existing enterprise solutions, databases, or other analytics tools.

APPENDIX

13.1] Glossary

ML: Machine Learning

HTML: Hypertext Markup Language

CSS: Cascading Style Sheets

13.2] Data Dictionary

The dataset used in this project consists of 21 columns capturing various aspects related to mental health. The data dictionary provides a detailed description of each column, including data type and potential values.

Column Name	Description	Data Type
User_ID	Unique identifier for each user	Integer
Age	Age of the user	Integer
Gender	Gender of the user	String (Male/Female/Other)

13.3] Machine Learning Model Details

1. Supervised Learning:

I. AdaBoost Algorithm
II. KNN Classifier
III. Random Forest
IV. Decision Tree
V. Gradient Boosting Classifier

Models Used:

Regression model for continuous prediction.
Classification model for categorical prediction.

Training Data: Dataset with labeled examples for supervised training.

2. Flask Integration

Endpoints:

/predict: Endpoint for making predictions.
/feedback: Endpoint for user feedback.

Security Measures:

HTTPS encryption for secure data transmission.
User authentication mechanisms.

13.4] Frontend Design

1. HTML Structure

Homepage: Overview of platform features and navigation.

Prediction Interface: User interface for input and displaying predictions.

Community Section: Forum, discussion groups, and community events.

2. CSS Styling

Color Palette: Calming and inclusive color scheme. - purples and light pastels

Responsive Design: Ensures usability across different screen sizes - drop down boxes and prewritten answers for ease of use.

Source Code:

** attached in the GitHub repository for this project

GitHub & Project Demo Link:

GitHub Link: https://github.com/smartinternz02/SI-GuidedProject-591336-1697650886.git

Project Demo Link (video explanation):
https://drive.google.com/file/d/1k9hFug87KLGTvcXKQBk5l-W5wbk6I7-Z/view?usp=sharing