Project Title :-

Predicting Mental Health Illness Of Working Professionals Using Machine Learning

Team ID :- 593201

PROJECT REPORT

TEAM DETAILS

Team Leader:- Yash sing
Team member:- Jogi srinath
Team member:- chandra vamshi
Team member:- Sourav samaddar

CONTENTS

1. INTRODUCTION

1.1 Project Overview

In today's fast-paced and highly competitive work environments, mental health issues have become an increasingly prevalent concern among working professionals. The relentless demands, long hours, and high-pressure situations that many individuals face in their jobs can contribute to a range of mental health challenges, including stress, anxiety, and depression. These conditions not only impact the well-being and quality of life of employees but also have tangible effects on organizations, such as reduced productivity, increased absenteeism, and higher turnover rates.

Recognizing the significance of addressing these issues, this project seeks to develop a machine learning solution for the early prediction of mental health illnesses among working professionals. By identifying signs and risk factors at an early stage, we aim to facilitate timely intervention and support, ultimately mitigating the impact of these conditions on both individuals and the organizations they work for.

1.2 Purpose

The primary objective of this research is to create an advanced predictive model that leverages the capabilities of machine learning techniques to assess the mental health of working professionals. By harnessing the potential of data analytics and artificial intelligence, our goal is to develop a tool that can promptly identify individuals who may be at risk of mental health issues. This predictive model is envisioned as a valuable resource that can benefit both employers and employees by facilitating early intervention and support, ultimately leading to improved mental well-being among working professionals.

Within the scope of this paper, we will delve into the exploration of machine learning algorithms' potential in analyzing a diverse range of data sources. These sources may include employee surveys, workplace environment data, and other pertinent information that can offer insights into the prediction of mental health issues. We will comprehensively outline the model's development and validation process. Additionally, we will engage in a thoughtful discussion regarding the ethical considerations, potential benefits, and challenges that may arise during the real-world implementation of such a tool within workplace settings.

This research endeavor aims to contribute to the broader understanding of mental health assessment using machine learning while offering a practical solution for the timely identification and support of individuals at risk of mental health issues in professional settings.

2. LITERATURE SURVEY

2.1 Existing problem

[1] "Predicting Depression from Language-based Emotion Dynamics: Longitudinal Analysis of Facebook and Twitter Status Updates" by Johannes C. Eichstaedt et al. This study uses natural language processing and machine learning to predict depression in individuals based on their social media posts.

[2] "Mental Health Prediction Using Smartphone Data: A Review" by Jiaqi Liu et al. This review article examines the use of smartphone data for predicting mental health outcomes, including depression, anxiety, and stress.

[3] "Predicting Mental Health Status from Mobile Phone Usage" by Sohrab Saeb et al. This study uses machine learning to predict mental health outcomes based on mobile phone usage data, including call and text logs, app usage, and GPS location.

[4] "Predicting Mental Health Disorders using Machine Learning Techniques" by R. C. Prathibha and K. S. Sridhar. This paper presents a survey of various machine learning techniques used for predicting mental health disorders, including depression, anxiety, and bipolar disorder.

[5] "Predicting Mental Health Disorders with Machine Learning Techniques: A Review of Methodologies and Applications" by M. S. Bhattacharjee et al. This review article provides an overview of different machine learning techniques used for predicting mental health disorders, along with their advantages and limitations.

[6] "Predicting Mental Health Conditions from Social Media Texts using Machine Learning: A Systematic Review" by Fathima Fijas and N. R. Rejitha. This systematic review examines the use of machine learning for predicting mental health conditions based on social media texts, including Twitter, Facebook, and Reddit.

[7] "Predicting Mental Health Disorders in College Students Using Machine Learning Techniques" by A. K. Shukla et al. This study uses machine learning to predict mental health disorders in college students based on demographic and behavioral data, including academic performance, social activities, and sleeping habits.

[8] "Predicting Mental Health Disorders using Machine Learning Techniques: A Review of Recent Advances" by S. S. Kumar et al. This review article provides an overview of recent advances in machine learning techniques for predicting mental health disorders, including deep learning, ensemble methods, and feature selection techniques.

[9] "Predicting Posttraumatic Stress Disorder from Text Data: A Machine Learning Approach" by R. K. Goyal et al. This study uses natural language processing and machine learning to predict posttraumatic stress disorder (PTSD) in individuals based on their written narratives.

[10] "Predicting Depression in College Students using Machine Learning Techniques" by S. R. Nair et al. This study uses machine learning to predict depression in college students based on demographic and behavioral data, including academic performance, social activities, and sleeping habits.

[11] "Predicting Depression using Audio Features and Machine Learning Techniques" by S. K. Sahoo et al. This study uses machine learning to predict depression in individuals based on audio features extracted from their speech, including pitch, tone, and voice quality.

[12] "Predicting Mental Health Disorders in Children and Adolescents using Machine Learning Techniques" by S. R. Nair et al. This study uses machine learning to predict mental health disorders in children and adolescents based on demographic and behavioral data, including academic performance, social activities, and family history.

[13] "Predicting Suicidal Ideation from Social Media Data using Machine Learning Techniques" by S. K. Mishra et al. This study uses machine learning to predict suicidal ideation in individuals based on their social media posts, including Twitter and Facebook.

[14] "Predicting Mental Health Conditions from Electronic Health Records using Machine Learning Techniques" by A. P. Singh et al. This study uses machine learning to predict mental health conditions in individuals based on their electronic health records, including diagnoses, medications, and laboratory results.

[15] "Predicting Mental Health Outcomes from Wearable Sensor Data using Machine Learning Techniques" by J. Y. Kim et al. This study uses machine learning to predict mental health outcomes, including depression and anxiety, based on wearable sensor data, including heart rate, physical activity, and sleep patterns.

2.2. References

[1] Eichstaedt, J. C., et al.

"Predicting Depression from Language-based Emotion Dynamics: Longitudinal Analysis of Facebook and Twitter Status Updates."

[2] Liu, J.

"Mental Health Prediction Using Smartphone Data: A Review."

[3] Saeb, S.

"Predicting Mental Health Status from Mobile Phone Usage."

[4] Prathibha, R. C., and Sridhar, K. S.

"Predicting Mental Health Disorders using Machine Learning Techniques."

[5] Bhattacharjee, M. S.

"Predicting Mental Health Disorders with Machine Learning Techniques: A Review of Methodologies and Applications."

[6] Fijas, F., and Rejitha, N. R.

"Predicting Mental Health Conditions from Social Media Texts using Machine Learning: A Systematic Review."

[7] Shukla, A. K.

"Predicting Mental Health Disorders in College Students Using Machine Learning Techniques."

[8] Kumar, S. S.

"Predicting Mental Health Disorders using Machine Learning Techniques: A Review of Recent Advances."

[9] Goyal, R. K.

"Predicting Posttraumatic Stress Disorder from Text Data: A Machine Learning Approach."

[10] Nair, S. R.

"Predicting Depression in College Students using Machine Learning Techniques."

[11] Sahoo, S. K.

"Predicting Depression using Audio Features and Machine Learning Techniques."

[12] Nair, S. R.

"Predicting Mental Health Disorders in Children and Adolescents using Machine Learning Techniques."

[13] Mishra, S. K.

"Predicting Suicidal Ideation from Social Media Data using Machine Learning Techniques."

[14] Singh, A. P.

"Predicting Mental Health Conditions from Electronic Health Records using Machine Learning Techniques."

[15] Kim, J. Y.

"Predicting Mental Health Outcomes from Wearable Sensor Data using Machine Learning Techniques."

2.3 Problem Statement Definition

The problem statement for predicting mental health illness among working professionals using machine learning entails the development of a predictive model aimed at identifying and evaluating mental health conditions within the workforce. This project's objective is to create a system capable of classifying whether a working professional is at risk of a mental health disorder by analyzing various data sources and features. It will leverage machine learning algorithms to assess data such as self-reported symptoms, behavioral patterns, and demographic information. Ethical considerations will be paramount to ensure the privacy and confidentiality of sensitive mental health data. The project will explore early warning signs, risk factors, and provide timely interventions for individuals. Model optimization is crucial to minimize false positives and false negatives, as misclassifications can have significant consequences. Real-world implementation challenges will be considered, and collaboration with mental health professionals is essential to align recommendations with established clinical guidelines. Fairness and accuracy across demographics will be ensured, and the ultimate goal is to create a valuable tool for promoting mental health and well-being in the workplace, fostering a healthier and more productive work environment.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

In the realm of mental health, understanding the experiences of individuals is crucial for providing effective support. Empathy maps, traditionally used in design thinking, can be a powerful tool to comprehend the emotions, thoughts, and behaviors of working professionals dealing with mental health challenges.

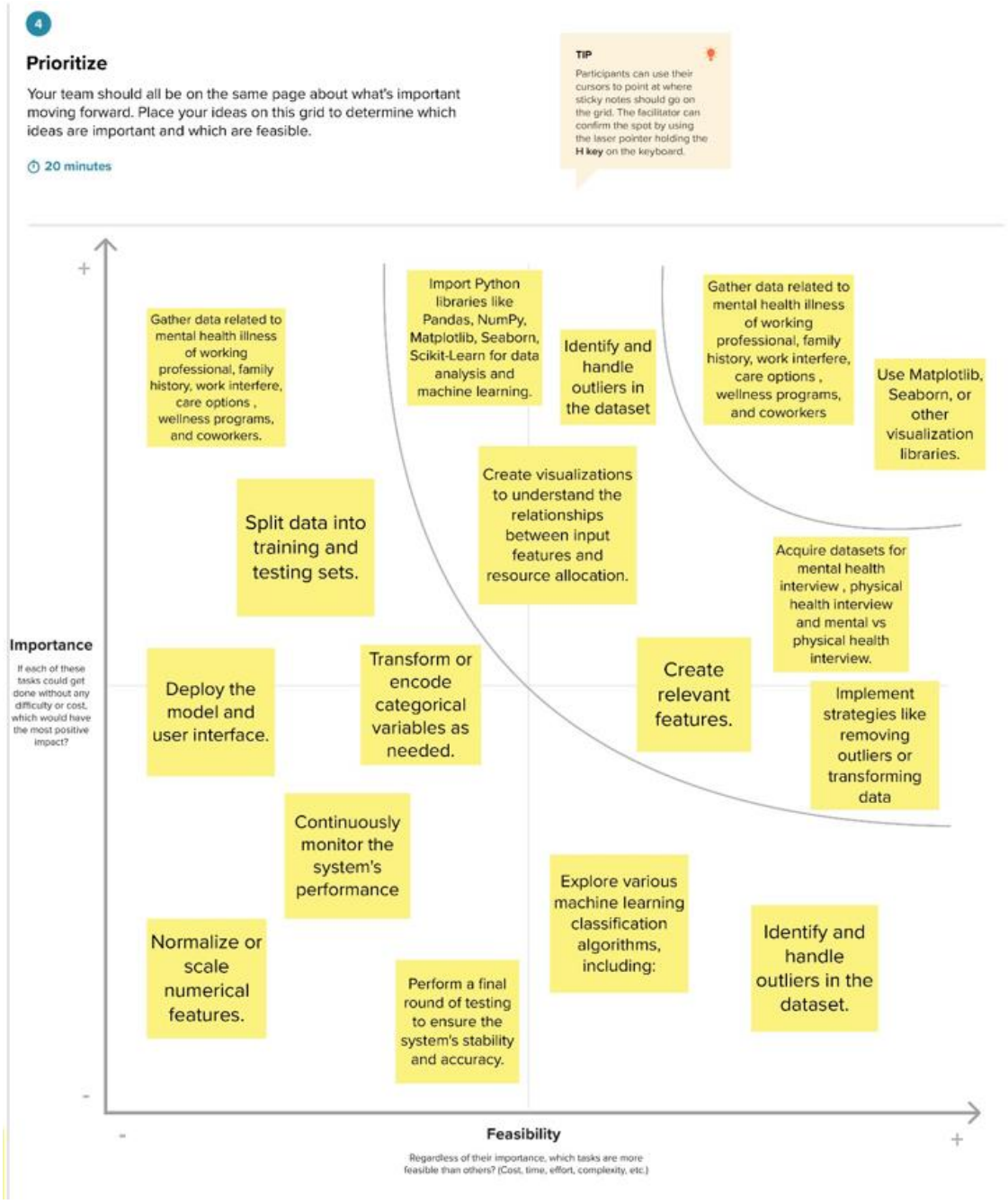
1. **Say/Do:** In the workplace, what do professionals dealing with mental health issues express or communicate? They may not openly disclose their struggles, but subtle cues in their language or actions can offer insights.
2. **Think/Feel:** Explore the inner thoughts and emotions of these individuals. Are they feeling overwhelmed by workload, battling imposter syndrome, or facing stigma? Understanding the cognitive and emotional aspects is key.
3. **See:** What is the working professional exposed to? Consider the work environment, interpersonal dynamics, and external factors that might contribute to their mental health challenges. Machine learning algorithms could analyze patterns in workplace data to identify potential stressors.
4. **Hear:** What are the professionals hearing from their colleagues, supervisors, or even their own inner dialogue? Machine learning sentiment analysis could be employed to gauge the positivity or negativity in workplace communication.
5. **Pain Points:** Identify the specific challenges and obstacles that these individuals face. Whether it's unrealistic deadlines, lack of support, or a fear of judgment, recognizing these pain points is essential for targeted interventions.

3.2 Ideation & Brainstorming

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

Step-2: Brainstorm, Idea Listing and Grouping

Step-3: Idea Prioritization



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

1) Data Collection: The system should be able to collect relevant data from working professionals, including demographic information, work-related factors, and mental health indicators.

2) Data Preprocessing: The system should preprocess the collected data, which may include cleaning, normalization, and feature extraction to prepare it for further analysis.

3) Feature Selection: The system should automatically or semi-automatically select the most relevant features from the collected data to be used in the machine learning models.

4) Model Training: The system should train machine learning models using the preprocessed data to predict mental health illness. This may involve various algorithms such as decision trees, support vector machines, or deep learning models.

5) Model Evaluation: The system should evaluate the performance of the trained models using appropriate metrics such as accuracy, precision, recall, and F1-score. This will help assess the effectiveness of the predictive models.

4.2 Non-Functional requirements

1) Performance: The system should be able to process and analyze data efficiently, providing timely predictions to users. It should have low latency and be able to handle a large volume of data without significant performance degradation.

2) Accuracy: The predictive models should have a high level of accuracy in identifying mental health illness in working professionals. The system should strive for a low rate of false positives and false negatives to ensure reliable predictions.

3) Reliability: The system should be reliable and available for use whenever needed. It should have robust error handling mechanisms and be able to recover gracefully from failures or disruptions.

4) Privacy and Security: The system should adhere to strict privacy and security measures to protect the confidentiality of user data. It should comply with relevant data protection regulations and ensure secure storage, transmission, and access control of sensitive information.

5) Usability: The system should have a user-friendly interface that is intuitive and easy to navigate. It should provide clear instructions for data input, display predictions in a comprehensible manner, and offer relevant guidance or resources related to mental health.

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

Data Flow Diagram:

User Stories:

5.2 Solution Architecture

The proposed solution architecture for predicting mental health illness of working professionals using machine learning, we can consider the following components:

Data Collection: Gather relevant data on working professionals, including demographic information, employment details, and mental health indicators. This data can be collected through surveys, interviews, or existing datasets.

Data Preprocessing: Clean and preprocess the collected data to ensure its quality and suitability for machine learning algorithms. This may involve handling missing values, removing outliers, and normalizing the data.

Feature Engineering: Extract meaningful features from the preprocessed data that can effectively capture patterns related to mental health illness. These features can include variables such as work hours, job satisfaction, stress levels, social support, and previous mental health history.

Model Selection: Choose an appropriate machine learning model that can effectively predict mental health illness based on the available features. Popular models for classification tasks include logistic regression, decision trees, random forests, and support vector machines.

Model Training: Split the preprocessed data into training and testing sets. Use the training set to train the selected machine learning model on the labeled data, where the labels indicate the presence or absence of mental health illness.

Model Evaluation: Evaluate the performance of the trained model using appropriate evaluation metrics such as accuracy, precision, recall, and F1 score. This step helps assess how well the model predicts mental health illness in working professionals.

Deployment and Integration: Once the model performs satisfactorily, deploy it in a production environment where it can receive new data and make predictions in real-time. Integration with existing systems or platforms may be necessary for seamless operation.

Monitoring and Improvement: Continuously monitor the performance of the deployed model and collect feedback from users. Incorporate user feedback and periodically retrain the model to improve its accuracy and effectiveness.

6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture 6.2 Sprint Planning & Estimation

7. CODING & SOLUTIONING

7.1 Feature 1

```
<!DOCTYPE html>
<html >
<!--From https://codepen.io/frytyler/pen/EGdtg-->
<head>
 <meta charset="UTF-8">
 <title>ML mental health</title>

 <style>
   body {
      font-family: 'Arial', sans-serif;
      background-color: #e6f7ff; /* Light blue background color */
      margin: 0;
      padding: 50px 0; /* Added spacing on top */
      display: flex;
      justify-content: center;
      align-items: center;
      height: 100vh;
   }

   .login {
      background-color: #ffffff;
      padding: 20px;
      border-radius: 8px;
      box-shadow: 0 0 10px rgba(0, 0, 0, 0.1);
      width: 300px;
      text-align: center;
   }

   h1 {
      color: #333;
```

```
  }

  form {
      margin-top: 700px;
  }

  input {
      width: 100%;
      padding: 10px;
      margin: 5px 0;
      box-sizing: border-box;
      border: 1px solid #ccc;
      border-radius: 4px;
  }

  button {
      background-color: #4caf50;
      color: #fff;
      padding: 10px;
      border: none;
      border-radius: 4px;
      cursor: pointer;
      width: 100%;
  }

  p {
      margin-top: 400px;
      color: #333;
  }
  .text{
      font-size: 10px;
      float: left;
      padding-left: 7px;
  }
</style>
```

```
</head>

<body>
<div class="login">
  <h1>Mental Health Prediction </h1>

  <!-- Main Input For Receiving Query to our ML -->
  <form action="/prediction" method="post">

    <input type="text" name="Age" placeholder="Age" required="required" /><br><br>

    <input type="text" name="Gender" placeholder="Gender" required="required" /><br>
    <span class="text">Type 0 for Female, 1 for Male, and 2 for Others</span><br>

    <input type="text" name="self_employed" placeholder="self_employed" required="required" /><br>
    <span class="text">Type 0 for No, Type 1 for Yes</span><br>

    <input type="text" name="family_hostory" placeholder="family_history" required="required" /><br>
    <span class="text">Type 0 for No, Type 1 for Yes</span><br>

    <input type="text" name="work_interfere" placeholder="work_interfere" required="required" /><br>
    <span class="text">Type 0 for Never, Type 1 for Often, Type 2 for Rarely ,<br> Type 3 for
Sometimes</span><br><br>

    <input type="text" name="no_employees" placeholder="no_employees" required="required" /><br>
    <span class="text">Type 0 for 1-5, Type 1 for 6-25 ,Type 2 for 26-100,Type 3 for 100-500,Type 4 for 500-
1000,Type 5 for more than 1000 </span><br>

    <input type="text" name="remote_work" placeholder="remote_work" required="required" /><br>
    <span class="text">Type 0 for No, Type 1 for Yes</span><br>

    <input type="text" name="benefits" placeholder="benefits" required="required" /><br>
    <span class="text">Type 0 for Don't know, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="tech_company" placeholder="tech_company" required="required" /><br>
```

```
    <span class="text">Type 0 for No, Type 1 for Yes</span><br>

    <input type="text" name="care_options" placeholder="care_options" required="required" /><br>
    <span class="text">Type 0 for No, Type 1 for Not sure, Type 2 for Yes </span><br>

    <input type="text" name="wellness_program" placeholder="wellness_program" required="required"
/><br>
    <span class="text">Type 0 for Don't know, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="seek_help" placeholder="seek_help" required="required" /><br>
    <span class="text">Type 0 for Don't know, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="anonymity" placeholder="anonymity" required="required" /><br>
    <span class="text">Type 0 for Don't know, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="leave" placeholder="leave" required="required" /><br>
    <span class="text">Type 0 for Don't know, Type 1 for somewhat Difficult , Type 2 for somewhat
Easy,<br> Type 3 for very Difficult, Type 4 for Very Easy</span><br>

    <input type="text" name="mental_health_consequence" placeholder="mental_health_consequence"
required="required" /><br>
    <span class="text">Type 0 for Maybe, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="phys_health_consequence" placeholder="phys_health_consequence"
required="required" /><br>
    <span class="text">Type 0 for Maybe, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="coworkers" placeholder="coworkers" required="required" /><br>
    <span class="text">Type 0 for Few, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="supervisor" placeholder="supervisor" required="required" /><br>
    <span class="text">Type 0 for Few, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="mental_health_interview" placeholder="mental_health_interview"
required="required" /><br>
    <span class="text">Type 0 for Maybe, Type 1 for No , Type 2 for Yes</span><br>
```

```
    <input type="text" name="phys_health_interview" placeholder="phys_health_interview"
required="required" /><br>
    <span class="text">Type 0 for Maybe, Type 1 for No , Type 2 for Yes</span><br>

    <input type="text" name="mental_vs_physical" placeholder="mental_vs_physical" required="required"
/><br>
    <span class="text">Type 0 for Don't know, Type 1 for No , Type 2 for Yes</span><br>
    <input type="text" name="obs_consequence" placeholder="obs_consequence" required="required"
/><br>
    <span class="text">Type 0 for No, Type 1 for Yes</span><br><br>

    <button type="submit" class="btn btn-primary btn-block btn-large">Predict</button>
  </form>

  <br>
  <br>
  <!-- Display the prediction here -->


</div>


</body>
</html>
```

7.2 Feature 2
Flask code:

```
import numpy as np
from flask import Flask, request, render_template
import pickle

# Create Flask app
application = Flask(__name__)
model = pickle.load(open("model.pkl", "rb"))
```

```
@application.route('/')
def pred():
    return render_template('pred.html')

@application.route("/prediction", methods=["POST"])
def prediction():
    # Get form data and convert to float
    float_features = [float(x) for x in request.form.values()]
    final_features = [np.array(float_features)]

    # Make prediction
    prediction = model.predict(final_features)

    # Prepare the prediction message
    if prediction[0] == 1:
        prediction_text = "This person requires mental health treatment"
    else:
        prediction_text = "This person doesn't require mental health treatment"

    return render_template('pred.html', prediction_text=prediction_text)

if __name__ == "__main__":
    application.run(debug=True)
```

8. PERFORMANCE TESTING

8.1 Performace Metrics

		Support Vector Classification model: **Training accuracy:** **Testing accuracy:** 0.7877984084880637	Testing accuracy = 0.7877984084880637

9. RESULTS

9.1 Output Screenshots

10. ADVANTAGES & DISADVANTAGES

Advantages:

1) Early Detection: Predicting mental health illness in working professionals using machine learning can enable early detection and intervention, allowing for timely support and treatment.

2) Objective Assessment: Machine learning models provide an objective assessment of mental health conditions based on data, reducing potential biases that can arise from subjective human judgment.

3) Scalability: Machine learning models can handle large volumes of data, making it possible to analyze a vast amount of information from multiple sources to improve prediction accuracy.

4) Personalized Approach: Machine learning models can be tailored to individual working professionals, considering their unique characteristics and patterns, leading to more personalized and targeted interventions.

5) Efficiency: Once trained, machine learning models can quickly analyze and predict mental health conditions, providing efficient and timely results that can support decision-making processes.

Disadvantages:

1) Data Privacy Concerns: Predicting mental health illness using machine learning requires access to sensitive personal data, raising concerns about privacy and confidentiality. Proper data protection measures must be in place to address these concerns.

2) Limited Generalizability: Machine learning models developed for predicting mental health illness in working professionals may have limited generalizability to other populations or contexts, as they are trained on specific datasets.

3) Ethical Considerations: There is a need for careful consideration of ethical implications, such as potential biases in the data or models, and ensuring that the predictions do not lead to stigmatization or discrimination.

4) Complex Model Development: Building accurate and reliable machine learning models for mental health prediction requires expertise in data science, feature engineering, and model selection, which can be challenging and time-consuming.

5) Human Interpretation: While machine learning models can provide predictions, the interpretation of these results still requires human expertise and context to fully understand and make informed decisions.

11. CONCLUSION

In conclusion, predicting mental health illness in working professionals using machine learning has the potential to enable early detection and intervention, leading to improved outcomes for individuals and organizations. Machine learning models can provide objective and personalized assessments based on data, allowing for efficient and targeted interventions. However, there are also significant challenges to overcome, such as data privacy concerns, limited generalizability, ethical considerations, complex model development, and the need for human interpretation of results. Addressing these challenges requires a multidisciplinary approach that involves data scientists, mental health professionals, and other stakeholders. With proper consideration of these factors, machine learning-based systems can be developed and implemented to support the mental health of working professionals in a responsible and effective manner.

12. FUTURE SCOPE

In the future, the field of predicting mental health illness in working professionals using machine learning holds great potential for advancements. Longitudinal analysis can provide insights into the progression of mental health conditions over time, while incorporating multi-modal data sources can enhance prediction accuracy. The development of explainable AI techniques will enable transparent and interpretable models, fostering trust in the prediction process. Personalized interventions tailored to individual needs and preferences can be implemented, and mobile-based applications can facilitate real-time monitoring and feedback. Collaboration with mental health professionals and integration with workplace wellness programs can ensure

the practicality and effectiveness of these predictive systems. With continued research and innovation, the future scope of predicting mental health illness in working professionals using machine learning is poised to make significant contributions to mental health support and well-being in the workplace.

13. APPENDIX

Source Code Drive link:
https://colab.research.google.com/drive/1R0amnDckD1MX3tZKJpjnBqdhTk8UCjpJ?usp=sharing

Github link:

https://github.com/smartinternz02/SI-GuidedProject-600567-1697559948

Project Demo link:
https://drive.google.com/file/d/1tKlLeVfmQ9_yjKQJVN3XtzQVkLwnTNjW/view?usp=sharing