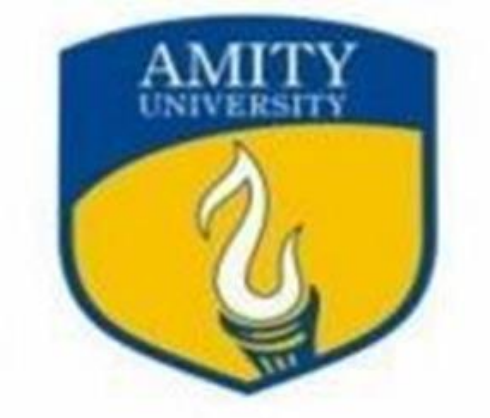
Project Report

on

Analyzing the Impact of Social Media on Mental Health

Submitted to

Amity University, Uttar Pradesh



In partial fulfilment of the requirements for the award of the degree of

Bachelor of Technology in Computer Science and Engineering

By

Naman Singhal

A2305222237

Under the guidance of

Dr. Vasudha Vashisht

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

AMITY SCHOOL OF ENGINEERING AND TECHNOLOGY

AMITY UNIVERSITY, UTTAR PRADESH

DECLARATION

I, Naman Singhal, hereby declare that the project work entitled "**Analyzing the Impact of social media on Mental Health**" submitted by me to Department of Computer Science and Engineering, Amity School of Engineering Technology, Amity University, Nodia, Uttar Pradesh, is submitted in partial fulfilment of the requirements for the award of the degree of Bachelor Technology in Computer Science and Engineering, has not been previously formed the basis for the award of any degree, diploma or other similar title or recognition.

The author attests that permission has been obtained for the use of any copyrighted material appearing in the report other than brief experts requiring only proper acknowledgement in scholarly writing and all such use is acknowledged.

Naman Singhal

A2305222237

5CSE-4X (2022-2026)

Date: 10/07/2024

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crown all efforts with success. I would like to thank **Prof (Dr) Sanjeev Thakur**, Head of Department-CSE, and Amity University for giving me the opportunityto undertake this project. I would like to thank my faculty guide **Dr Vasudha Vashisht** who is the biggest driving force behind my successful completion of the project. He has always been there to solve any query of mine and to guide me in the right direction regarding the project. I would also like to thank my team mates, Dhruv Jain and Raman Kalra who played their valuable part in completion of the project.

Naman Singhal

A2305222237

5CSE-4X (2022-2026)

Date: 10/07/2024

ABSTRACT

The increasing use of social media has raised questions about how it affects mental health. Despite all the advantages that social media provides, such as information access and social connections, overuse of the platform has been related to detrimental impacts on mental health. In-depth research on social media's influence on mental health is provided in this paper, which looks at both the advantages and disadvantages. In the report, the connection between social media use and mental health outcomes—such as anxiety, sadness, and low self-esteem—is examined. It also addresses the possible mechanisms—such as cyberbullying, sleep disruption, and social comparison—that may be responsible for these impacts. In order to encourage responsible social media use and lessen its detrimental impact on mental health, the research offers recommendations for individuals, legislators, and social media firms at the end.

CONTENT

INTRODUCTION

Overview

The surge of mental health issues and the requirement for better medical care has brought about a study on machine learning that can be employed in mental health problems. This report presents an up-to-date systematic review of machine learning and deep learning approaches to mental health problem prediction. Also, we will go through the challenges, limitations, and prospects applying machine learning in mental healthcare. We search reliable databases to find research articles and studies related to predicting mental health problems. Afterward, we order the collected research articles according to the type of psychological disorder like schizophrenia, bipolar disorder, anxiety and depression_, posttraumatic stress disorder, children's mental disorders. Bringing out these findings allow us to think about what are some important difficulties that researchers face when using machine learning technology in mental healthcare_. Moreover, there are some actionable guidelines as to where research may develop by considering advancement of deep learning technologies in the field of psychiatry. The major objective of Mental Health Prediction system is to decide whether a person should consult a psychiatrist or not based on provided inputs.

Purpose

Mental health first aid is meant to teach the attendants on how to recognize and support anyone suffering from mental problems or affected by substance abuse and providing them with links to concerned parties. Companies can provide comprehensive welfare packages that enable their workers to manage mental illnesses. These may include Employee Assistance Programs, Wellness programs addressing both physical and mental health, Health and Disability Insurance as well as flexible working arrangements or leave policies. Organizations, which have included mental health in their programs, help create a safe work climate that mitigates against discrimination experienced by people living with mental diagnostics, augmenting the overall knowledge pool of the organizations on such issues and teaching them how best to respond responsibly when colleagues demonstrate signs of poor mental health. Mental Health Prediction Program's main objective is to decide if an individual needs psychological advice from practitioners based on any data one inputs into it. Classification algorithms like Logistic Regression, KNN, Decision tree, Random Forest canl be used for this purpose. For this project we have used Random Forest algorithm and it has given us most accurate result.

LITERATURE SURVEY

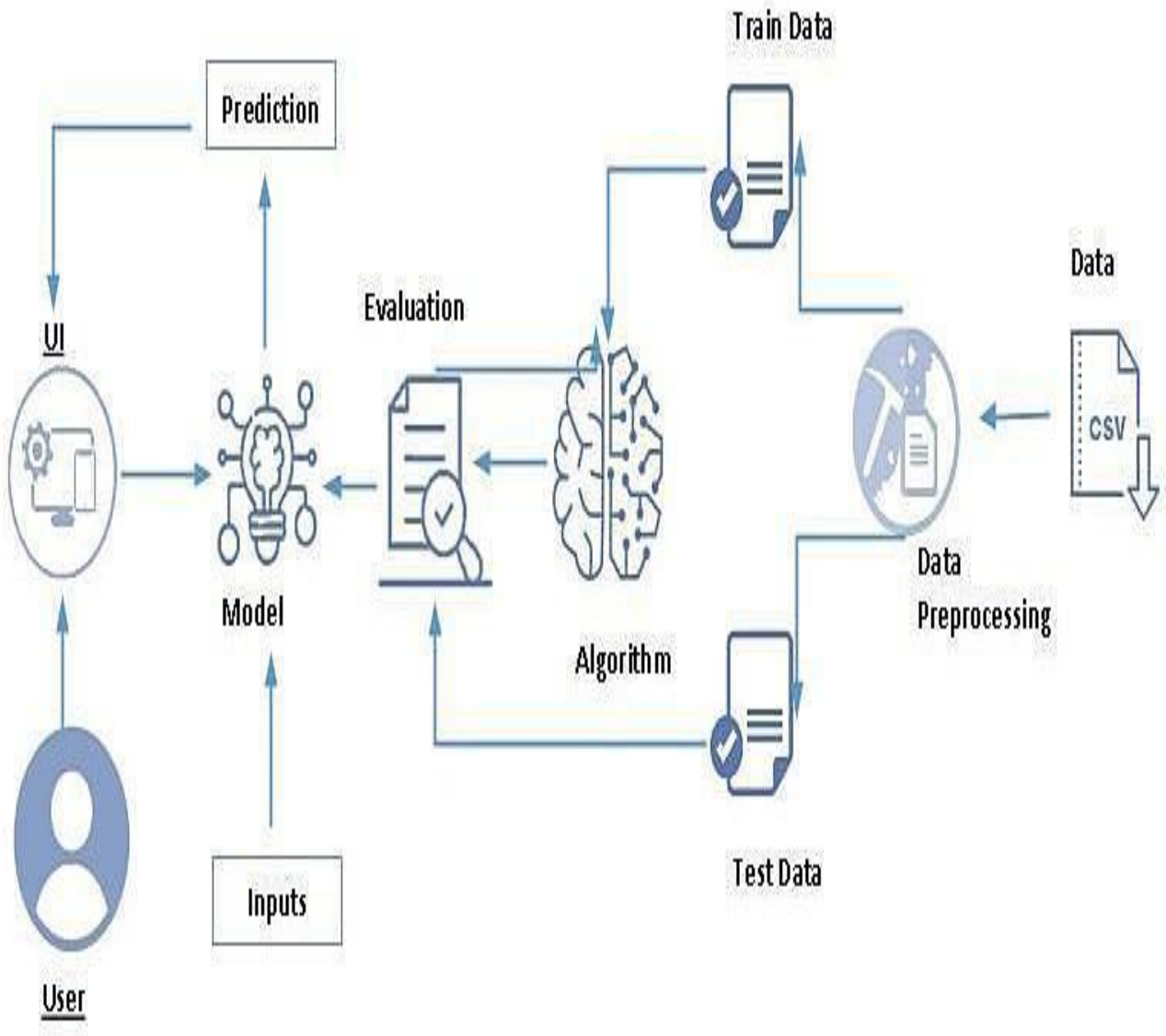
Existing Problem

Evidently, mental illness is a health problem that affects emotions, thinking and socialising in a person. Its effect on mental health is becoming more well acknowledged. Persistent exposure to carefully chosen pictures, perfected lives, and online social media can increase anxiety, depressive symptoms, and feelings of inadequacy. Self-esteem can be damaged and a sense of comparison can be fostered by the pressure to live up to the unattainable standards set by social media. Furthermore, the compulsive nature of scrolling through feeds and looking for approval from likes and comments can make offline feelings of loneliness and isolation worse.These have shown that mental illness has serious implications to societies and requires new approaches towards its prevention and intervention. Early identification of mental health is vital in achieving these strategies Numerous people should hopefully be able to regain their lives through appropriate care and treatment after suffering from any form of a mental disorder or emotional condition.

Proposed Solution

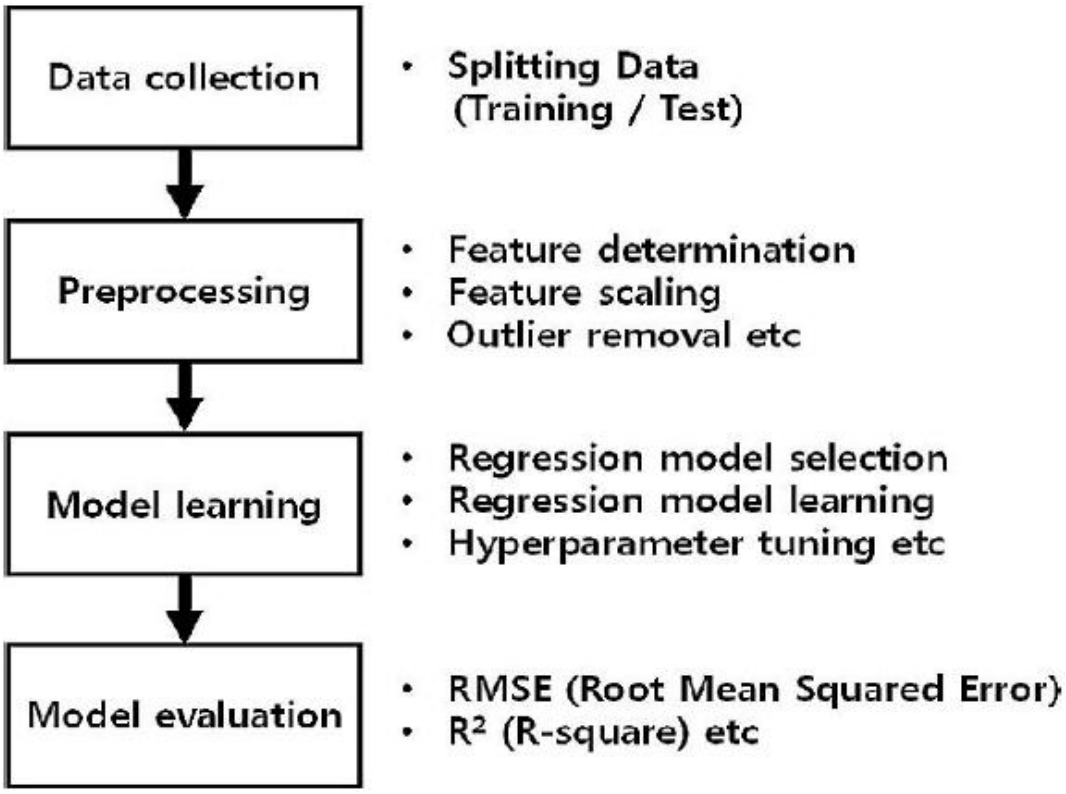
Machine learning is a statistical technique that uses sophisticated models, such as advanced statistics and probability techniques, with the goal of developing a system that learns and evolves as it gains experience. This leads to data mining for important insights, personalization experiences using personalized algorithms, and AI systems adoption in automating intelligence. Our model utilizes datasets gathered through questionnaires to train a machine learning algorithm aimed at predicting patients' mental health. This involves constructing an ensemble of decision trees (Random Forest) to aggregate predictions from multiple trained models. The algorithm considers various factors from the questionnaires, such as emotional well-being, stress levels, and coping mechanisms, to identify patterns and correlations indicative of different mental health outcomes.

MENTAL HEALTH PREDICTION ALGORITHM FLOWCHART



THEORETICAL ANALYSIS

Block Diagram



Hardware / Software Designing

Recommended System Requirements

● **Processors**: Intel® Core™ i5 processor 4300M at 2.60 GHz or 2.59 GHz (1 socket, 2 cores, 2 threads per core), 8 GB of DRAM Intel® Xeon® processor E5-2698 v3 at 2.30 GHz (2 sockets, 16 cores each, 1 thread per core), 64 GB of DRAM Intel® Xeon Phi™ processor 7210 at 1.30 GHz (1 socket, 64 cores, 4 threads per core), 32 GB of DRAM, 16 GB of MCDRAM (flat mode enabled)

● **Disk space**: 2 to 3 GB

● **Operating systems**: Windows® 10, macOS*, and Linux*

Minimum System Requirements

- **Processors**: Intel Atom® processor or Intel® Core™ i3 processor
- **Disk space**: 1 GB
- **Operating systems**: Windows* 7 or later, macOS, and Linux
- **Python:** versions: 3.10

Software requirements:

VS CODE: Visual Studio Code (VS Code) is a flexible code editor that was built by Microsoft and is popular for its strong functionality that includes things like IntelliSense, in-built debugging and Git support. It can be customized greatly with different themes and extensions which makes it compatible with various programming languages and frameworks. This allows developers to have an uninterrupted coding experience as VS Code comes with features such as integrated terminal and remote development capabilities. Developers rely on it due to its speed, adaptability, and good will of other programmers.

ALGORITHM USED

Random forest

Purpose: The random forest algorithm is designed to improve upon the limitations of individual decision trees by leveraging ensemble learning. It aims to enhance prediction accuracy and robustness in both classification and regression tasks by aggregating the outputs of multiple decision trees.

Application: Random forests find applications in various domains such as:

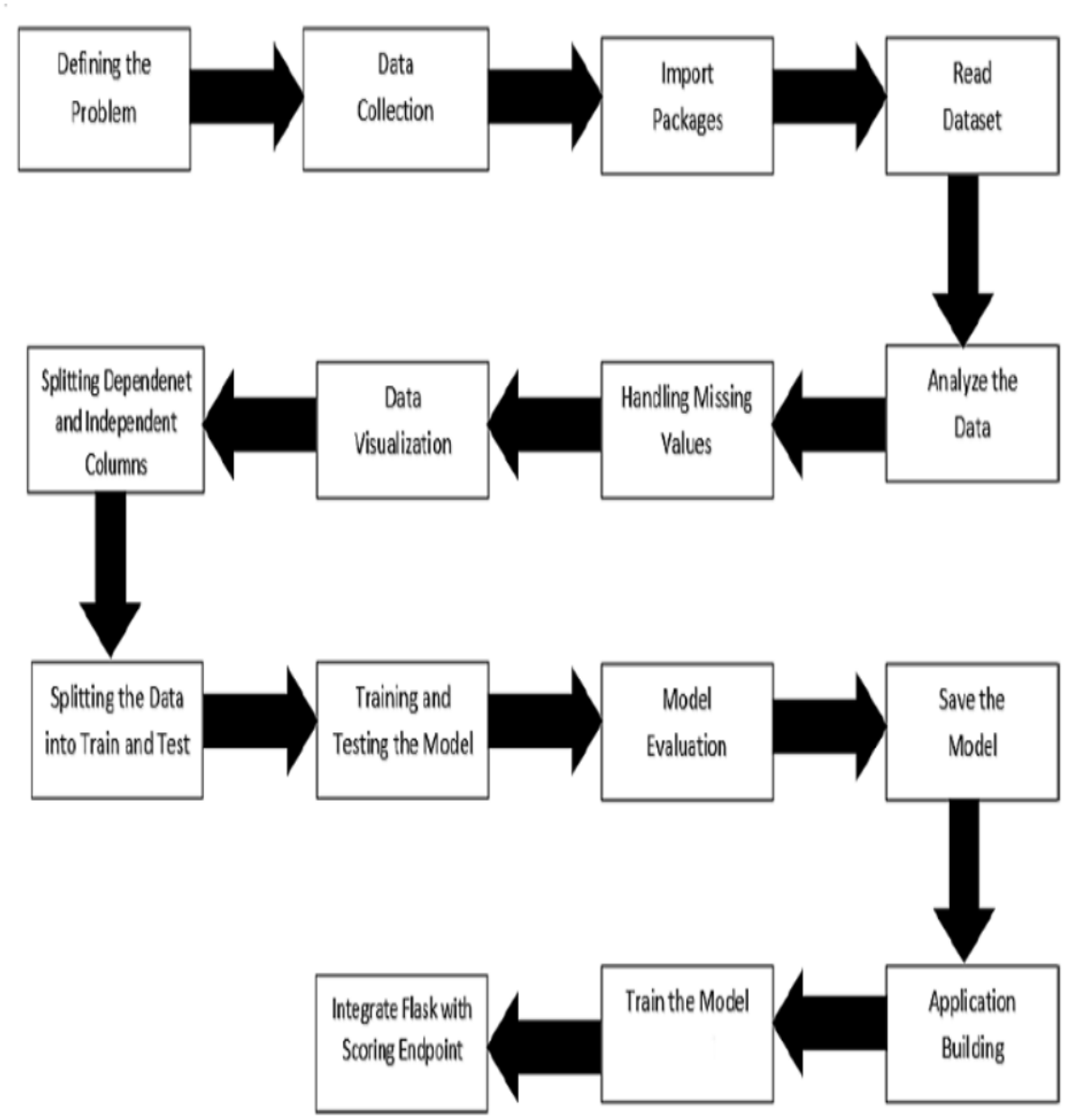
- **Classification:** Predicting the class of an observation based on its features, such as detecting spam emails or classifying diseases.
- **Regression:** Predicting continuous outcomes, like estimating housing prices or forecasting sales figures.
- **Anomaly Detection:** Identifying outliers or unusual patterns in data that deviate from normal behavior.
- **Feature Importance:** Assessing the relevance of features in predicting the target variable, aiding in feature selection.

Benefits:

- **Reduced Overfitting:** By averaging predictions from multiple trees trained on different subsets of data, random forests mitigate overfitting and generalize well to new, unseen data.
- **Robustness:** They handle noisy data and outliers effectively by aggregating predictions, which reduces the impact of individual erroneous predictions.
- **Feature Importance:** Random forests provide insights into which features are most influential in making predictions, aiding in understanding the underlying relationships within the data.
- **Scalability:** They can handle large datasets with high-dimensional feature spaces, making them suitable for real-world applications with complex data structures.

In essence, the random forest algorithm stands out for its versatility, accuracy, and ability to handle diverse data types and tasks, making it a popular choice across various fields of machine learning and data analysis.

Flowchart


Defining the Problem
Data Collection
Import Packages
Read Dataset
Analyze the Data
Handling Missing Values
Data Visualization
Splitting Dependenet and Independent Columns
Splitting the Data into Train and Test
Training and Testing the Model
Model Evaluation
Save the Model
Application Building
Train the Model
Integrate Flask with Scoring Endpoint

Code

Model

Import necessary libraries

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import LabelEncoder
from sklearn.ensemble import RandomForestClassifier
import joblib, pickle
```

Load the Dataset and handle missing values

```
def train_model():
    # Load the dataset
    df = pd.read_csv("C:\\Users\\dhruv\\Downloads\\project\\project\\dataset\\social_media_mental_health_7000.csv")

    # Remove any unnamed columns
    df = df.loc[:, ~df.columns.str.contains('^Unnamed')]

    # Handle missing values by filling with mode for categorical and median for numerical
    for column in df.columns:
        if df[column].dtype == 'object':
            df[column].fillna(df[column].mode()[0], inplace=True)
        elif df[column].dtype.name == 'category':
            df[column].fillna(df[column].mode()[0], inplace=True)
        else:
            df[column].fillna(df[column].median(), inplace=True)
```

Encode the dataset and split it into features and targets

```
    label_encoder = LabelEncoder()
    df['MentalStressLevel'] = label_encoder.fit_transform(df['MentalStressLevel'])

    # Encode categorical variables
    for column in df.columns:
        if df[column].dtype == 'object':
            df[column] = label_encoder.fit_transform(df[column])

    # Split data into features (X) and target (y)
    X = df.drop(columns=['MentalStressLevel'])
    y = df['MentalStressLevel']

    # Save feature names for later use
    joblib.dump(X.columns, 'feature_names.joblib')
```

Split the data into train and test sets

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Save the Model using pickle and create a predict function to be used on new data

```
def predict(data):
    # Load the trained model
    model = pickle.load(open('model.pkl', 'rb'))

    # Load feature names to ensure data columns are in correct order
    feature_names = joblib.load('feature_names.joblib')

    # Create a DataFrame with the input data, matching the original feature order
    df = pd.DataFrame([data], columns=feature_names)

    # Make predictions using the loaded model
    prediction = model.predict(df)
    return prediction[0]

# Train the model when this script is run
if __name__ == '__main__':
    train_model()
```

Initialize and fit the model using best parameters

```
best_params_rf = {
    'max_depth': 10,
    'max_features': 'sqrt',
    'min_samples_leaf': 1,
    'min_samples_split': 10,
    'n_estimators': 300
}

# Initialize the RandomForestClassifier with the best parameters
rf_classifier = RandomForestClassifier(random_state=42, **best_params_rf)

# Fit the model on the training data
rf_classifier.fit(X_train, y_train)
```

Flask App

Import necessary Libraries and Model

```
from flask import Flask, render_template, request
import model
```

Create Route for HomePage

```
@app.route('/')
def index():
    return render_template('index.html')
```

Create Route for handling form submission and output

```
@app.route('/out', methods=['POST'])
def out():
    if request.method == 'POST':
        # Collect form data from the POST request
        age = request.form['age']
        gender = request.form['gender']
        time_spend = request.form['Time_spend']
        feel = request.form['feel']
        felt_left = request.form['felt_left']
        avoid = request.form['avoid']
        affect = request.form['affect']
        amount = request.form['amount']
        pressure = request.form['pressure']
        experience = request.form['experience']
        worry = request.form['worry']
        distracted = request.form['distracted']
        others = request.form['others']
        share = request.form['share']
        anxious = request.form['Anxious']
        follow = request.form['Follow']
        interact = request.form['interact']
        impact = request.form['impact']
```

Create List of Data to pass to the Model

```
data = [age, gender, time_spend, feel, felt_left, avoid, affect, amount, pressure, experience, worry, distracted, others, share, anxious, follow, interact, impact]
```

Predict using the model and render the prediction result

```
        # Use the model to make a prediction
        prediction = model.predict(data)

        # Render the output.html template and pass the prediction result to it
        return render_template('output.html', y=prediction)

if __name__ == '__main__':
    app.run(debug=True)
```

HTML WEBPAGE

HTML STRUCTURE AND META DATA

```
<!DOCTYPE html>
<html>
<head>
  <meta charset="UTF-8">
  <title>Mental health prediction</title>
</head>
<style>
```

BODY CONTENT

```
    body {
        background-image: url("heart.jpg");  /* Sets the background image of the body */
        background-repeat: no-repeat;  /* Prevents the background image from repeating */
        background-size: cover;  /* Ensures the background image covers the entire viewport */
    }
</style>
<body>
```

Form submission via POST method to the /out endpoint

```
<div class="login">
    <center><h1>Mental health Prediction</h1></center> <!-- Centered heading -->

    <form action="/out" method="post">
```

Dropdowns

Gender selection

```
<p>
    <label for="gender">Gender</label><br>
    <select name="gender">
        <option value="Male"> Male </option>
        <option value="Female"> Female </option>
        <option value="Other"> Other </option>
    </select>
</p> <br>
```

Time spent on social media

```
<p>
    <label for="Time_spend">How often do you use social media?</label><br>
    <select name="Time_spend">
        <option value="1hr"> Less than 1 hour a day </option>
        <option value="2hr"> 1-2 hours a day </option>
        <option value="3hr"> 3-4 hours a day </option>
        <option value="4hr"> More than 4 hours a day </option>
    </select>
</p> <br>
```

- Similarly, there are dropdowns of various attributes such as feel, felt_left, avoid, affect, amount, pressure, experience, worry, distracted, others, share, anxious, follow, interact, impact

Submit button

```
    <a href="output.html" ></a><button type="submit" class="btn btn-primary btn-block btn-large"
    style="height:30px;width:200px">predict</button></a>
    <br>
</form>
```

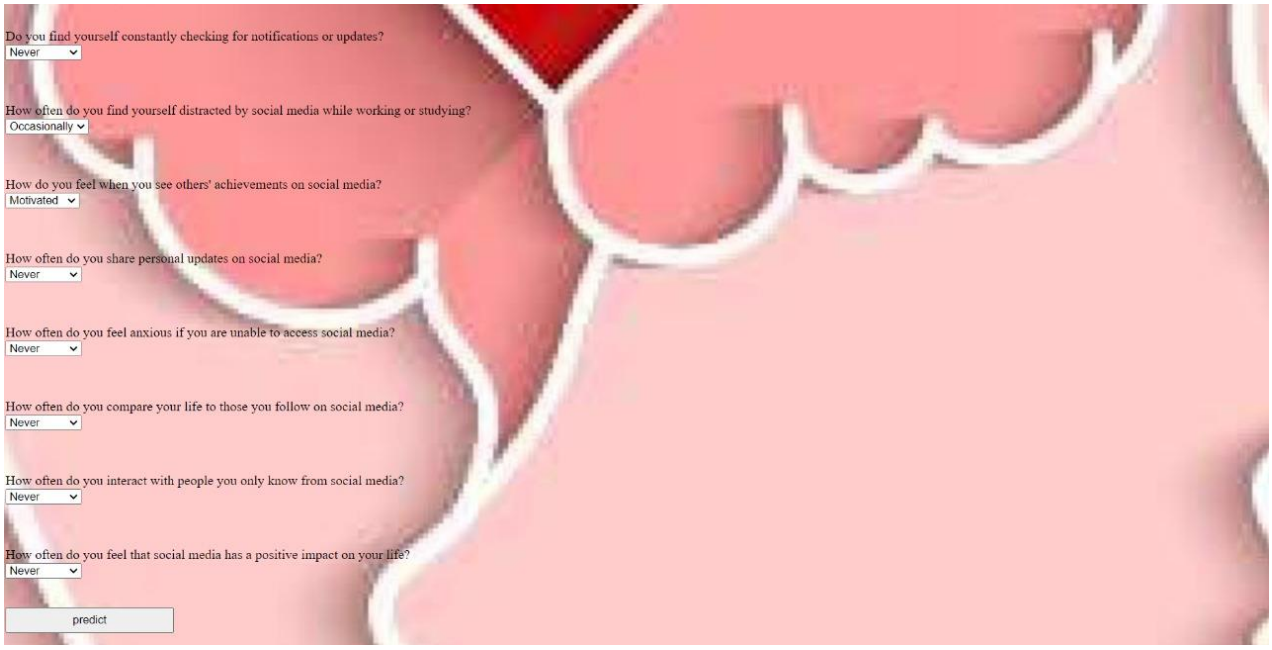
Placeholder for displaying prediction results

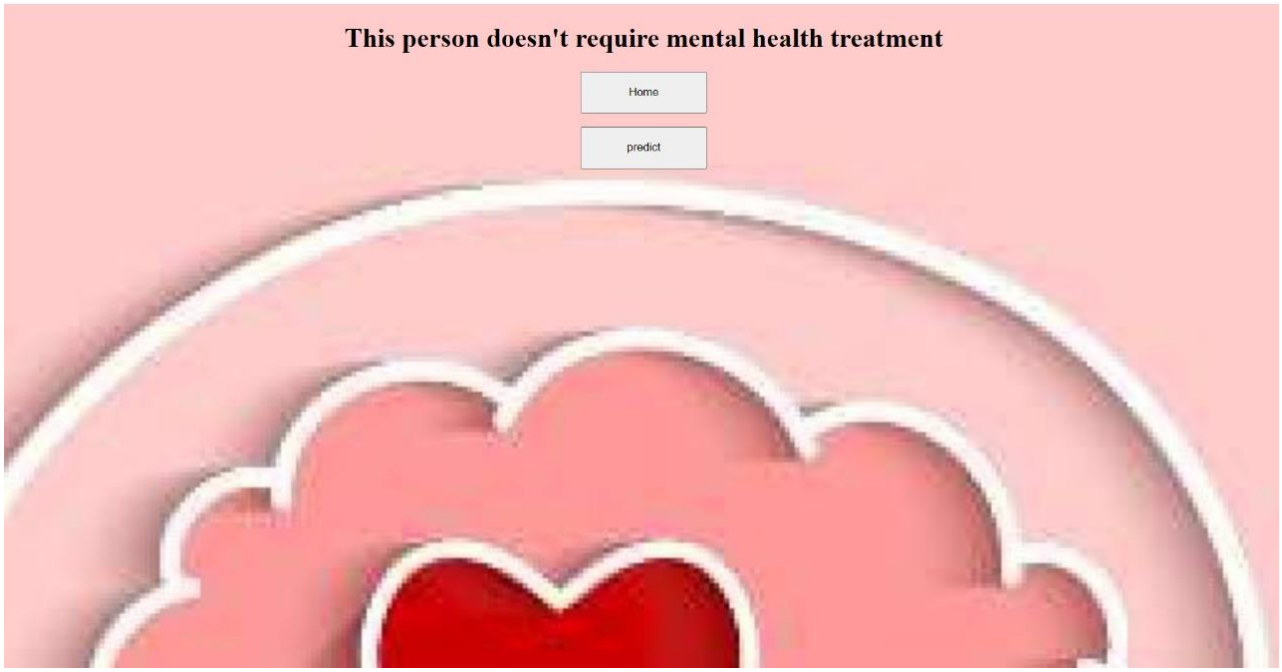
```
    <p> <b> {{y}} </b> </p>

</div>

</body>
</html>
```

RESULT


Mental health Prediction
Mental health affects your emotional, psychological and social well-being. It affects how we think, feel, and act. It also helps determine how we handle stress, relate to others, and make choices. In the workplace, communication and inclusion are keys skills for successful high performing teams or employees. The impact of mental health to an organization can mean an increase of absent days from work and a decrease in productivity and engagement. In the United States, approximately 70% of adults with depression are in the workforce. Employees with depression will miss an estimated 35 million workdays a year due mental illness.
Proceed


Mental health Prediction
Age
49
Gender
Male
How often do you use social media?
Less than 1 hour a day
How do you feel after spending time on social media?
Happy and relaxed
Have you ever felt left out or excluded after seeing posts on social media?
Never
Do you use social media to avoid real-life problems or responsibilities?
Sometimes
How does social media affect your sleep?
It slightly affects my sleep
How do you feel about the amount of time you spend on social media?
I'm satisfied


Do you find yourself constantly checking for notifications or updates?
Never
How often do you find yourself distracted by social media while working or studying?
Occasionally
How do you feel when you see others' achievements on social media?
Motivated
How often do you share personal updates on social media?
Never
How often do you feel anxious if you are unable to access social media?
Never
How often do you compare your life to those you follow on social media?
Never
How often do you interact with people you only know from social media?
Never
How often do you feel that social media has a positive impact on your life?
Never
predict


This person doesn't require mental health treatment
Home
predict

CONCLUSION

Various methodologies and algorithms have been proposed and implemented to assess and resolve mental health issues. Many of the solutions can yet be improved. Furthermore, there are still a lot of issues in the field of machine learning for mental health that need to be identified and investigated in a range of contexts. Since categorizing mental health data is typically an extremely difficult task, the features incorporated into machine learning algorithms will have a big impact on the classification's performance.

Research and studies now in existence indicate that machine learning can be a helpful tool in understanding mental illnesses. In addition, it might aid in identifying and categorising patients' mental health issues in order to provide them with additional care.

More recent methods that leverage data from the integration of multiple sensor modalities found in high-tech gadgets have shown to be a useful tool for identifying patient responses and mood state, among other things.

It is evident that the majority of research and studies still have difficulty validating their findings due to a lack of sufficient, accepted, validated data, particularly from outside sources. In addition, it's possible that different machine learning algorithms perform differently on different problems. Depending on the characteristics and data samples acquired, the machine learning models' performance will change.

Furthermore, preparatory tasks like data cleansing and parameter tuning might have an impact on machine learning models in order to get the best outcomes.

Therefore, it is crucial that academics look into and use different machine learning techniques to analyse the data.

algorithms to select the machine learning algorithm with the highest accuracy. In addition, the researchers' obstacles and constraints needed to be carefully handled in order to provide acceptable outcomes that may enhance clinical practice and decision-making.

REFERENCES

[1] Braghieri, Luca, Ro'ee Levy, and Alexey Makarin. "Social media and mental health." *American Economic Review* 112.11 (2022): 3660-3693.

[2] Huang, Chiungjung. "A meta-analysis of the problematic social media use and mental health." *International Journal of Social Psychiatry* 68.1 (2022): 12-33.

[3] Kaur, Simarjeet, Kamaljeet Kaur, and Rohan Verma. "Impact of social media on mental health of adolescents." *Journal of Pharmaceutical Negative Results* (2022): 779-783.

[4] Skaik, R. and Inkpen, D., 2020. Using social media for mental health surveillance: a review. *ACM Computing Surveys (CSUR)*, *53*(6), pp.1-31.

[5] De Choudhury, M., 2013, October. Role of social media in tackling challenges in mental health. In *Proceedings of the 2nd international workshop on Socially-aware multimedia* (pp. 49-52).