

# **Mental Health Treatment Predictor**

**PWSkills Mini-Hackathon: Healthcare Applications**

**Submitted by:**

**Prabal Roy**

prabalroy242@gmail.com

Mobile: 9679097884

**Date:** October 15, 2025

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Objective . . . . .	2
<b>2</b>	<b>Project Overview</b>	<b>3</b>
<b>3</b>	<b>Model Performance and Comparison</b>	<b>4</b>
<b>4</b>	<b>Setup Instructions</b>	<b>5</b>
4.1	Virtual Environment . . . . .	5
4.2	Install Dependencies . . . . .	5
4.3	Run the App . . . . .	5
<b>5</b>	<b>Project Structure</b>	<b>6</b>
<b>6</b>	<b>Batch Prediction</b>	<b>7</b>
<b>7</b>	<b>Troubleshooting</b>	<b>8</b>

# Chapter 1

## Introduction

The **Mental Health Treatment Predictor** is a machine learning-powered Streamlit application designed to predict whether an individual is likely to seek mental health treatment. This project aims to raise awareness and enable early intervention through data-driven insights.

### 1.1 Objective

The main objective of this project is to use predictive analytics to identify factors influencing mental health treatment decisions and assist organizations in developing targeted support strategies.

# Chapter 2

## Project Overview

- **Home:** Project overview and quick stats.
- **Model Information:** Displays model type, top feature importance, and comparison charts.
- **Make Prediction:** Interactive form with probabilities, confidence scores, and recommendations.
- **Batch Prediction:** CSV upload for multiple predictions with visual summaries.

# Chapter 3

## Model Performance and Comparison

The table below summarizes model evaluation metrics:

Model	Accuracy	Precision	Recall	F1 Score	ROC AUC
Naive Bayes	0.7421	0.7692	0.7031	0.7347	0.7983
SVM	0.7381	0.7627	0.7031	0.7317	0.7851
Logistic Regression	0.7302	0.7679	0.6719	0.7167	0.8218
Random Forest	0.7222	0.7500	0.6797	0.7131	0.7762
CatBoost	0.7222	0.7589	0.6641	0.7083	0.7888
XGBoost	0.7183	0.7395	0.6875	0.7126	0.7539
Gradient Boosting	0.7103	0.7350	0.6719	0.7020	0.8022
LightGBM	0.7103	0.7311	0.6797	0.7045	0.7864
Decision Tree	0.6548	0.6752	0.6172	0.6449	0.6515

**Best Model:** Naive Bayes (74.21% accuracy, 79.83% ROC AUC)

**Best ROC AUC:** Logistic Regression (82.20% ROC AUC)

**Ensemble Model:** Voting Classifier (72.62% accuracy, 80.89% ROC AUC)

# Chapter 4

## Setup Instructions

### 4.1 Virtual Environment

```
python -m venv venv  
./venv/Scripts/Activate.ps1
```

### 4.2 Install Dependencies

```
pip install -r requirements.txt
```

### 4.3 Run the App

```
streamlit run app.py
```

# Chapter 5

## Project Structure

```
.
project_root/
Notebooks
|   data_cleaning.ipynb
|   exploratory_analysis.ipynb
|   feature_engineering.ipynb
|   model_training.ipynb
src/
|   data/
|   models/
|   utils/
|   app.py
data/
|   raw/
|   |   mental_health_survey.csv
|   processed/
|       model_comparison_results.csv           # Shown in Model Information
|       cleaned_mental_health_data.csv
|       complete_processed_data.csv
|       selected_features.json
|       test_data.csv
|       train_data.csv
models/                                         # Trained artifacts used by predictor
|   best_model.pkl
|   feature_scaler.pkl
|   feature_selector.pkl
|   label_encoders.pkl
|   target_encoder.pkl
|   selected_features.json
readme.md
prediction.csv                                # Example input for Batch Prediction
requirements.txt
```

# Chapter 6

## Batch Prediction

Users can upload a CSV with the same input columns as the form and receive predictions for all records. The app also provides:

- Summary statistics
- Confidence distribution plots
- Downloadable prediction CSV file



# Chapter 7

## Troubleshooting

- **Missing metrics:** Ensure `data/processed/model_comparison_results.csv` exists.
- **Prediction error:** Check that all model artifacts are present in the `models/` folder.
- **New environment setup:** Use `pip freeze > requirements.txt` to lock dependencies.