Exploring the Dynamics of Depression with Actigraphy based Time-Series Data and Demographic Factors

Data Mining B-565

Akriti Kumari, Prathamesh More, Sarvesh Soni, Vedika Halwasiya

Table of Contents

Introduction

- A severe and widespread mental illness, depression has a substantial negative impact on people's lives all over the world

- Numerous symptoms, such as ongoing melancholy, a feeling of emptiness, worry, sleep difficulties, and a deep loss of interest in once-pleasurable activities, are indicative of this illness. Other signs of depressive episodes include low energy, thoughts of suicide, feelings of guilt or worthlessness, and even psychotic symptoms.

- The number, intensity, and duration of these symptoms, as well as how they affect social and professional functioning, are used to determine the severity of depression.

Project Objectives

Develop predictive model

Develop a predictive model utilizing actigraph-based time series data, MADRS scores, and demographic factors to forecast depression severity.

Implement ML Techniques

Implement machine learning techniques, including deep learning and supervised learning, for the classification of depressive states.

Investigate sleep patterns

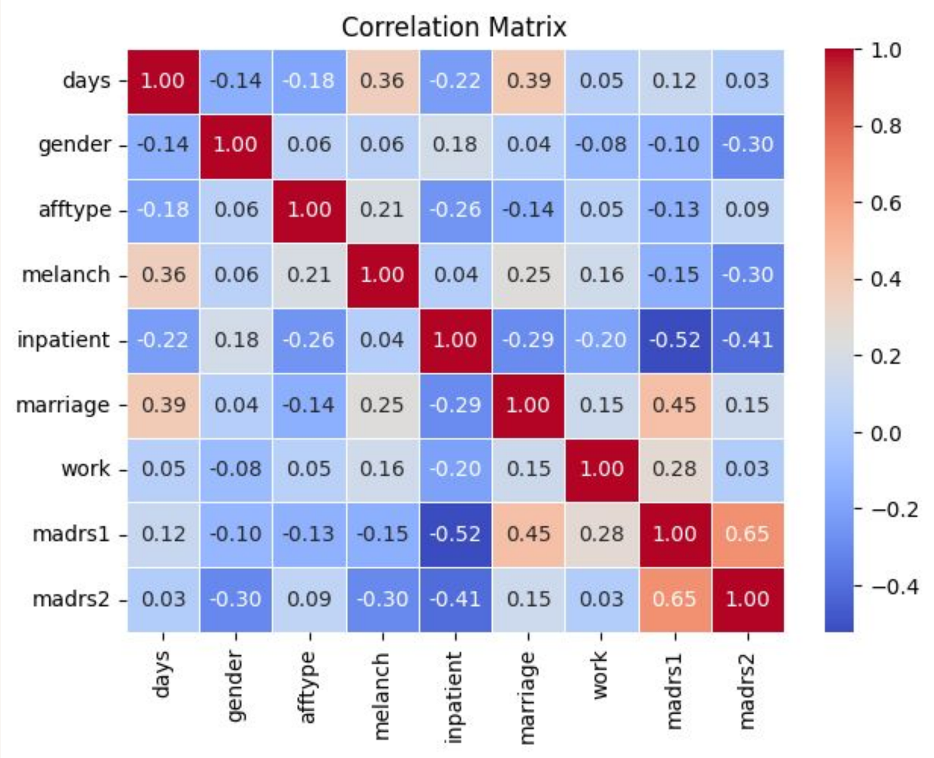
Investigate sleep patterns in depressed and non-depressed populations, aiming to automate depression diagnosis for enhanced treatment and understanding.

Dataset Overview

- **"The Depression Dataset"** from Kaggle
  https://www.kaggle.com/datasets/arashnic/the-depression-dataset/code

- **Contents of Dataset:**
  - ***Condition Data:*** 23 datasets containing minute-wise actigraphy readings across several days for 23 patients affected by depression
  - ***Control Data:*** 32 datasets containing minute-wise actigraphy readings across several days for 32 patients not suffering from depression.
  - ***Scores.csv:*** A dataset containing patient number, demographic attributes and clinical attributes for all 55 patients (23 of condition and 32 of control group)

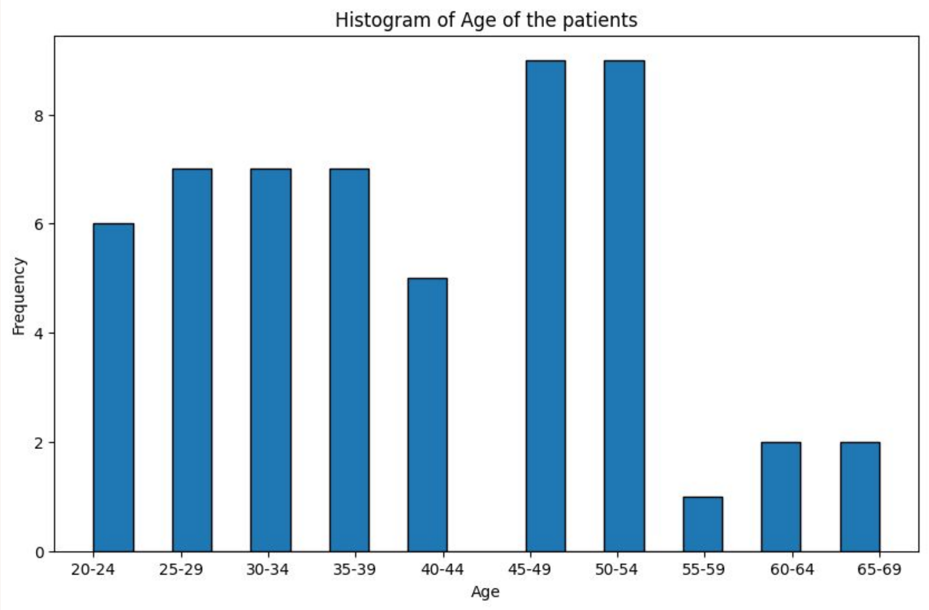
EDA (Scores Data)



Scores Data

- MADRS2 being the target has strong features :
    - Gender
    - Melanch
    - Inpatient
    - MADRS1

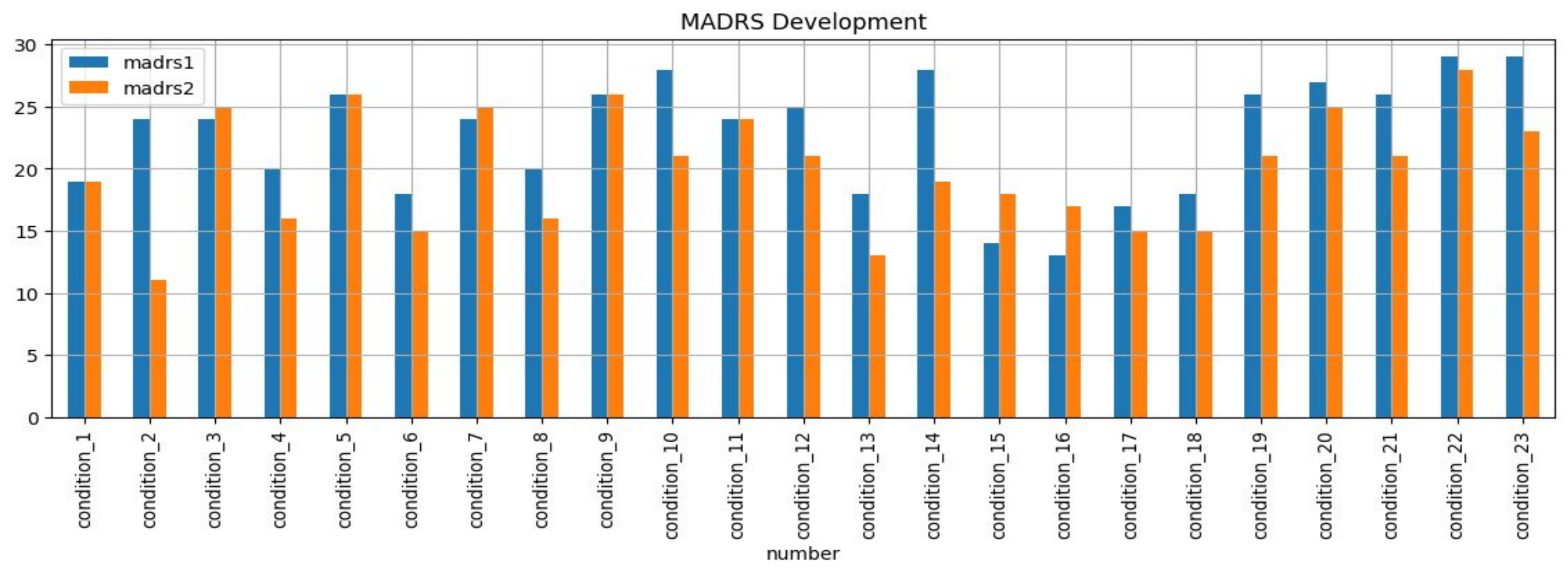
EDA (Scores Data)



Scores Data

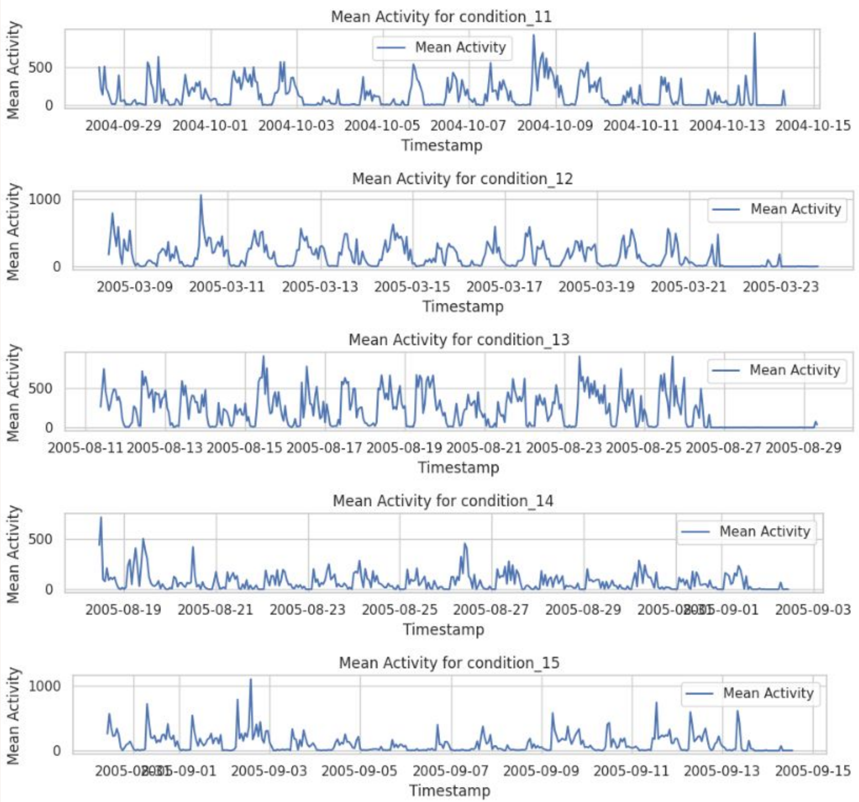
- Depression is observed as
    - Most prominent in age groups 45-54
    - Least in age group 55 and above

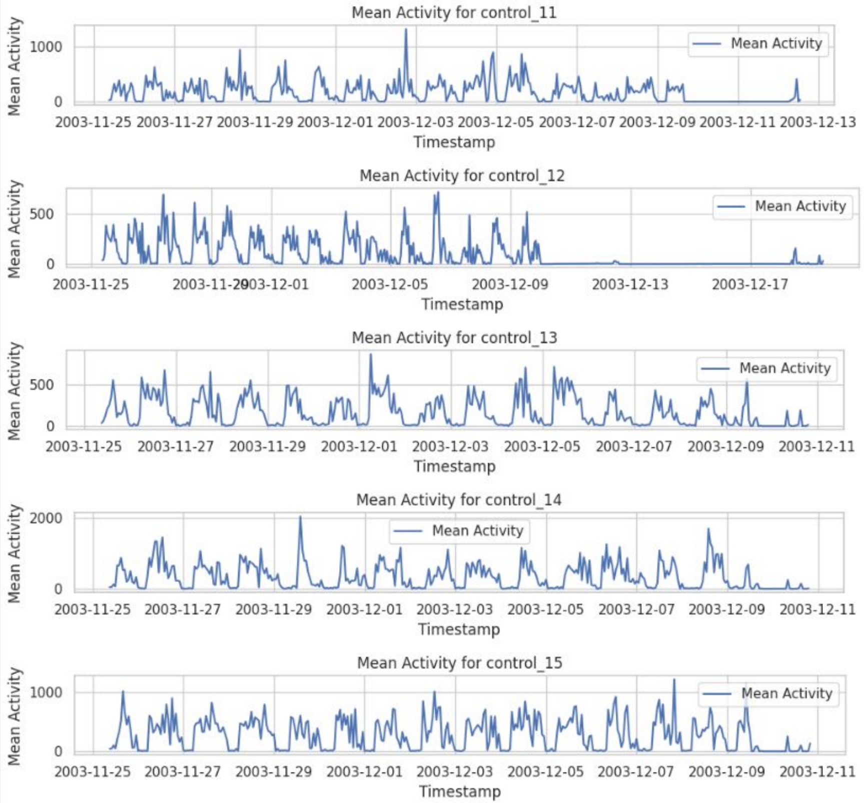
EDA (Scores Data)



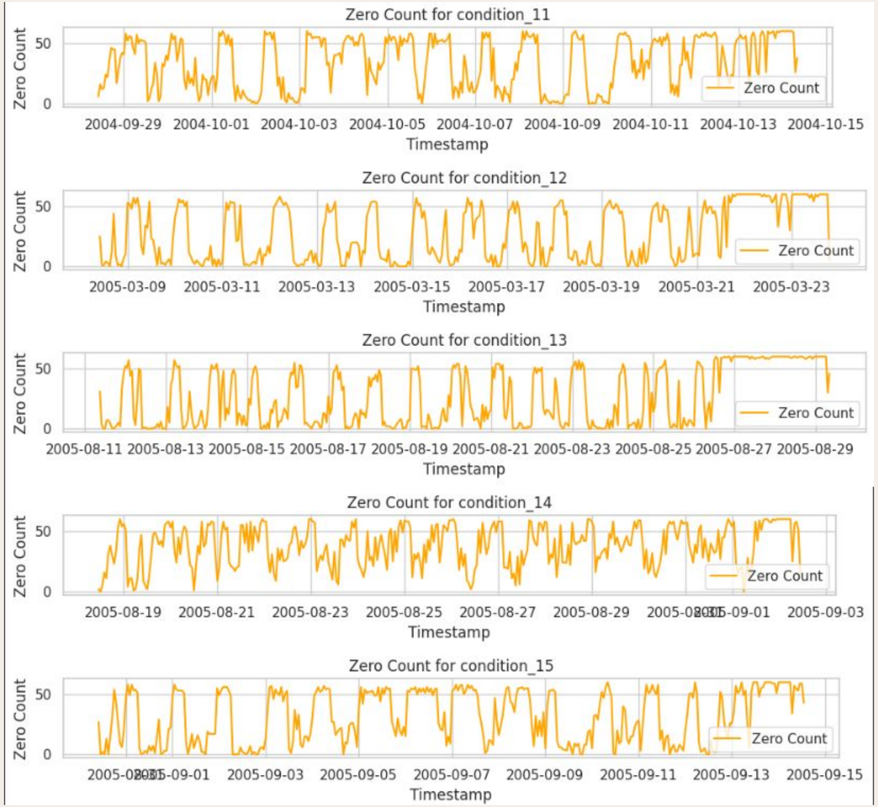
- MADRS1: MADRS value at the start of Actigraphy measurement
- MADRS2: MADRS value at the end of Actigraphy measurement
- Majority of the patients show a drop in MADRS value by the end of measurement

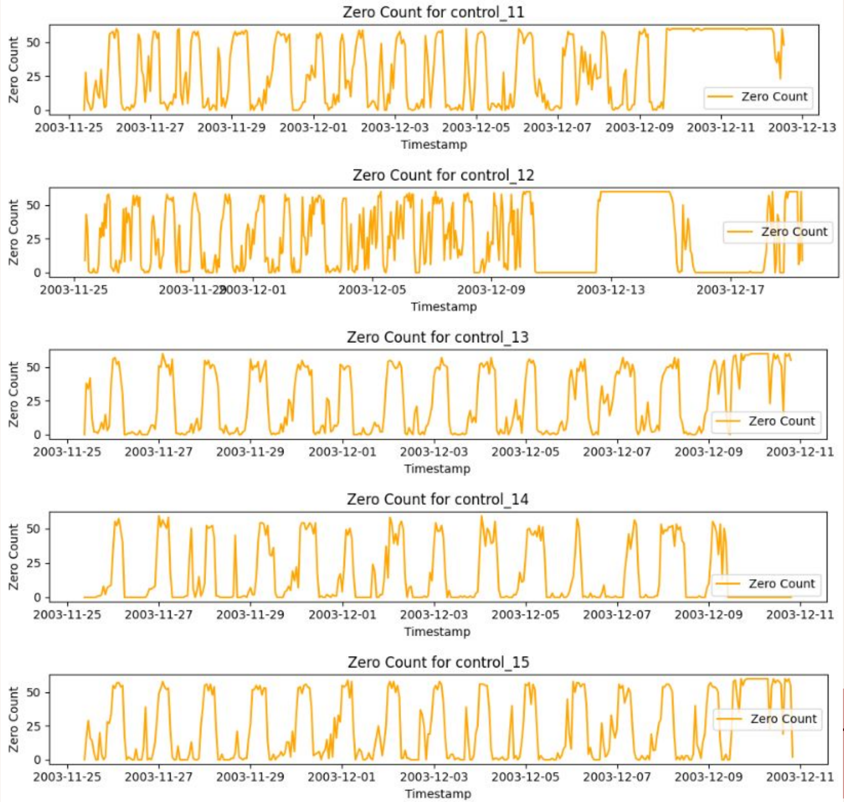
EDA (Actigraph Data)



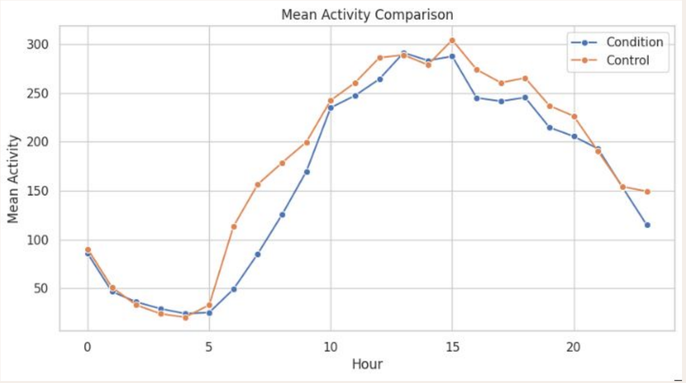


EDA (Actigraph Data)





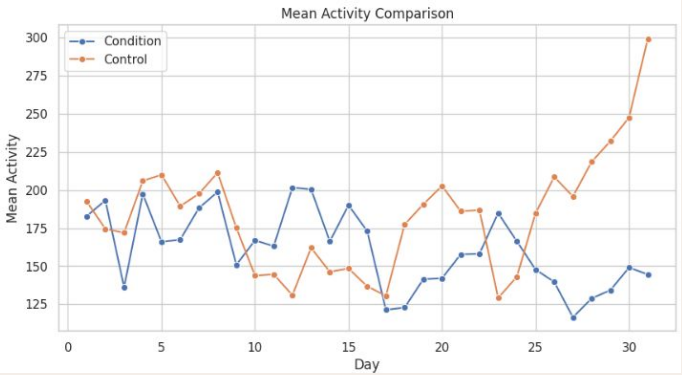
EDA (Actigraph Data)

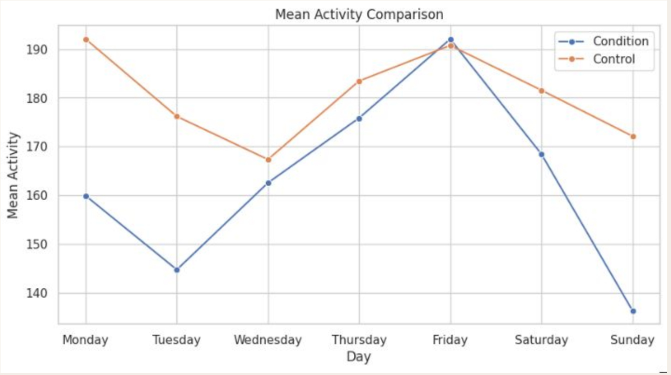


Comparison between mean activity for condition and control group:

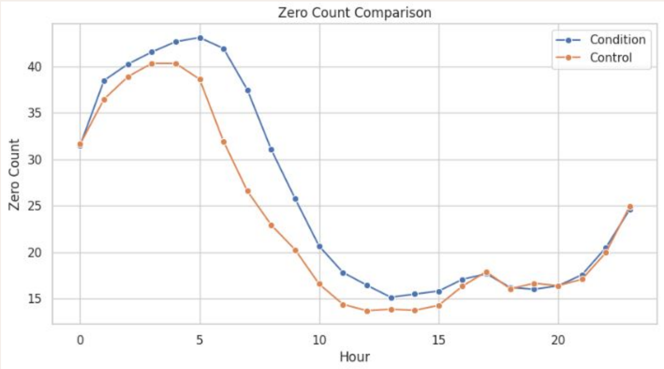
1. Hour of Day
2. Day of month
3. Day of week

We can observe clear difference between the values for both groups.





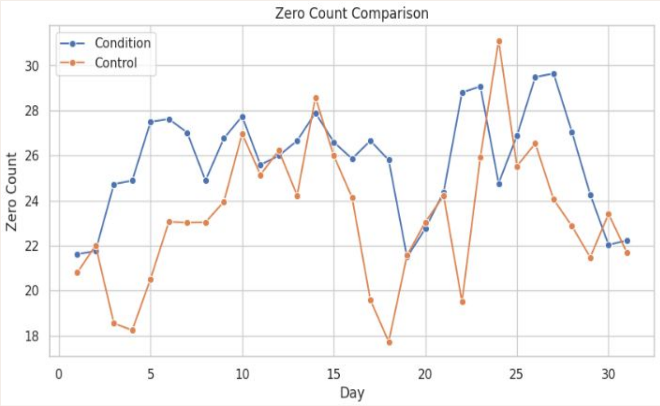
EDA (Actigraph Data)

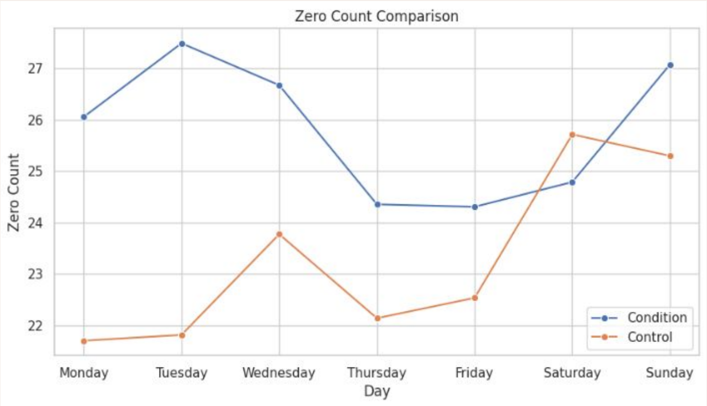


Comparison between Zero Activity for condition and control group:

1. Hour of Day
2. Day of month
3. Day of week

The difference suggests how actigraph readings can be useful for prediction of depression





Models Used

Prediction of if a person has depression or not based on Actigraph Readings

	Logistic Regression (with PCA)	Naive Bayes (with PCA)	KNN (with PCA)	ANN (with PCA)	Random Forest (with cross-validation
Accuracy	0.730	0.718	0.974	0.978	0.976
Precision	0.88	0.72	0.97	0.98	0.97
Recall	0.27	0.33	0.96	0.96	0.97
F1-score	0.42	0.46	0.96	0.97	0.97

Models Used

Prediction for MADRS2 value of a patient having Depression

	Linear Regression	Ridge Regression	Lasso Regression	Gradient boosting	XGBoost	KNN
MSE	7.174	7.243	8.454	8.252	8.430	5.712
R^2	49.6%	49.1%	40.6%	42.1%	40.8%	59.9%

Current Progress

- Two out of three objectives have been achieved:
  - Preprocess the data
    For instance, transforming the timestamps of the actigraph data from per-minute intervals to hourly mean values and subsequently averaging them to obtain weekly values.
  - Developed a Predictive Model
    - Implemented various machine learning techniques to create a robust predictive model.
    - Explored models such as K-Nearest Neighbor, XG-Boost, Naive Bayes, ANN and Random Forest.

- Investigation of sleep patterns in progress

Conclusion & Future Work

Conclusion:

- Successful development of a predictive model.
- Implementation of diverse machine learning techniques.
- Significant progress in understanding and predicting depression.

Future Work:

- Incorporate advanced techniques like time series analysis.
- Integrate LSTM models for improved temporal dynamics understanding.
- Refine models for a deeper comprehension of the interplay between sleep patterns and mental health.

THANK YOU