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

Akriti Kumari akkumari@iu.edu
Prathamesh More ppmore@iu.edu
Sarvesh Soni sarvsoni@iu.edu
Vedika Halwasiya vhalwasi@iu.edu
*Indiana University Bloomington*
*Bloomington, IN 47405, USA*

**Abstract**

Depression is a widespread mental illness that has a significant effect on people's lives. With actigraph data, MADRS scores, and patient details included, this dataset provides a groundbreaking chance to study depression. It makes a variety of applications possible, such as the prediction of MADRS scores and the classification of depressive states driven by machine learning. It also makes it easier to analyze sleep patterns in both depressed and non-depressed people in detail. The dataset is methodologically flexible enough to support a wide range of methodologies, including deep learning and supervised machine learning as well as regression modeling. Additionally, it facilitates the management of class imbalance by utilizing oversampling and cost-sensitive classification. This dataset facilitates comparative analyses, which in turn promotes greater comprehension of successful approaches. Essentially, it has the potential to automate the diagnosis of depression, which could lead to better treatment and a deeper understanding of this complex mental health issue.

**Keywords**– Deep learning, time series analysis, depression, actigraph data, MADRS scores, machine learning, classification, sleep patterns, sensor data

## 1. 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.

Bipolar disorder, another serious mental illness, commonly co-occurs with depression, proving that depression is not an isolated ailment. The primary difference is in the cyclical nature of bipolar disorder, which is typified by manic episodes that are characterized by elevated activity levels, impulsivity, decreased sleep needs, and goal-directed behaviors. It is known that there is a hereditary component to both diseases, indicating that mood episodes are caused by a genetic predisposition that combines with environmental circumstances.

The start of depressive symptoms can be attributed to a variety of environmental variables, including lifestyles that are out of sync with natural daylight cycles, alterations in social patterns (such as traveling across time zones or working shifts), and biological rhythms that are upset by seasonal variations in daylight. Furthermore, a person's depressive symptoms may be triggered by physical health conditions, drug side effects, life events, social factors, alcohol and substance addiction, and other reasons. With an estimated 15% lifetime prevalence worldwide, depression is a startling statistic. Notably, a large number of people have depressive episodes that don't quite fit the diagnostic parameters but yet have a substantial negative impact on their well-being.

Presented by Enrique Garcia-Ceja et al [1]., this dataset offers a multitude of actigraph data monitoring motor activity together with demographics, patient characteristics, and MADRS ratings. This dataset offers a foundational resource for a variety of applications, such as machine learning-based depression state classification, MADRS score prediction based on motor activity data, and the examination of sleep patterns in depressed versus non-depressed individuals. It has the potential to revolutionize the field of depression research. Additionally, it makes it easier to assess a variety of machine learning strategies, including cost-sensitive classification, methods for oversampling classes that are imbalanced, and feature-based and deep learning approaches like recurrent neural networks and convolutional neural networks for time series data. The creation of automated depression detection systems is made possible by this extensive dataset, which could further our knowledge of this complicated illness and lead to better diagnostic and therapeutic approaches.

## 2. Previous Work

[2] The authors of this work tackle the important problem of utilizing machine learning algorithms based on actigraphy data from wearable devices to classify depression levels. A common mental health problem is depression, for which early diagnosis is essential to successful treatment. The study offers a framework for classifying depression levels by utilizing actigraphy data, which offers continuous monitoring of physical activity. Survey factors are coupled with 14 circadian rhythm traits that were taken from actigraphy data to model depressive status. After evaluating a number of machine learning methods, the study concludes that the XGBoost classifier is the best at identifying depression levels. This study offers a viable method for the identification and categorization of depression and offers insightful information on the connection between depression and physical activity.

[3] This work offers a novel method for utilizing actigraphy data from non-invasive sensing devices to monitor and categorize illness states. The study focuses on ongoing patient monitoring for a range of chronic illnesses, offering insightful information for management and treatment. The authors extract structural information from actigraphy data using mathematical time-series modeling, namely Autoregressive (AR)–Generalized Conditional Heteroskedasticity (GARCH) models. They suggest a three-step process that entails locating structural fractures, creating time-series models for every segment, and using feature analysis to categorize the severity of the condition. The experimental results show that variables collected using this modeling approach are useful in differentiating between disease severities

in patients with depressive disorder. This strategy presents a viable way for ongoing observation and early illness state identification in a home-based environment.

[4] In the literature related to bipolar disorder and circadian rhythm disruptions, research has shown that early identification of prodromal symptoms and the stabilization of patients are crucial goals in managing the condition. Actigraphy, a technology capable of continuously monitoring movement activity and sleep patterns, presents an advantage over traditional methods like polysomnography. This paper explores a feasibility study that combines long-term actigraphy monitoring with a self-assessment of the patient's mood. By analyzing data from Actiwatch devices and patient questionnaires, the study aims to identify relapse events related to bipolar disorder, particularly mania or depression. The research indicates that the most promising parameter for detection is Interdaily Stability, and future studies plan to incorporate real-time actigraphy monitoring for improved performance.

[5] This research examines the important problem of depression and the possibility of early identification utilizing motor activity sensor data along with demographic data. The severe repercussions of depression, such as mental disease and suicide, highlight how vital it is to have efficient detection systems. Based on motor sensor readings and demographic information, the study uses machine learning methods, namely Random Forest, AdaBoost, and Artificial Neural Networks, to categorize sadness. The results are remarkable since they exhibit strong agreement and an astounding accuracy of 98%, as seen by the correlation coefficients of Matthew and Cohen. This study demonstrates how machine learning may be used to improve early detection techniques and provides insightful information about the use of sensor data for mental health analysis.

[6] This research addresses privacy concerns and presents a unique method for predicting depression using motor activity data from ActiGraph wearable wristbands. Depression is a common health problem worldwide, and early identification is crucial. In order to diagnose mental health issues, wearable technology provides a platform for ongoing health monitoring. However, issues with privacy prevent data exchange. In this study, a data augmentation strategy that greatly enhances depression detection performance is proposed. In addition, it investigates the use of privacy-preserving data analysis to guarantee the protection of patient data when it comes to mental health issues. This work offers a viable path for improving the diagnosis of depression and promoting privacy in the analysis of health data.

[7] This paper presents a real-time activity classification system using inertial sensors that is based on a semi-supervised Hidden Markov Model (HMM). Many applications find traditional approaches to activity classification impractical due to their large training dataset requirements. The suggested system makes it possible to recognize both long-term complex activities and short-term events. In order to reduce the amount of training data required, a general model is customized to each subject using Bayesian adaptation techniques. The framework exhibits efficient activity classification even with limited training data and offers insightful information about the health and behavioral patterns of the subjects.

## 3. Data Collection

The dataset utilized in this analysis was obtained from Kaggle, a platform hosting various public datasets. Specifically, the data is from the "The Depression Dataset" compiled by Möbius. The dataset was originally collected for the study of motor activity in schizophrenia and major depression. Motor activity was monitored with an actigraph watch worn on the right wrist (Actiwatch, Cambridge Neurotechnology Ltd, England, model AW4). The actigraph watch measures activity levels. The sampling frequency is 32Hz and movements over 0.05 g are recorded. A corresponding voltage is produced and is stored as an activity count in the memory unit of the actigraph watch. The number of counts is proportional to the intensity of the movement. Total activity counts were continuously recorded in one-minute intervals.

The dataset contains two folders (one containing the data for the control group and the other for the condition group) and a scores.csv file. The two folders contain CSV files for each patient's actigraph data collected over time. The columns are:
timestamp (one-minute intervals), date (date of measurement), and activity (activity measurement from the actigraph watch).

The scores file contains the following columns; number (patient identifier), days (number of days of measurements), gender (female or male), age (age in age groups), afftype (bipolar II, unipolar depressive, bipolar I), melanch (melancholia or no melancholia), inpatient (inpatient, or outpatient), edu (education grouped in years), marriage (married or cohabiting, or single), work (working or studying, or unemployed/sick leave/pension), madrs1 (MADRS score when measurement started), madrs2 (MADRS scores when measurement stopped).

## 4. Exploratory Data Analysis

We examined scores and actigraphy data in our thorough Exploratory Data Analysis (EDA) of the depression dataset. We found patterns in the MADRS1 and MADRS2 scores using pairsplots, boxplots, and barplots, and we also found subtle correlations in the condition group-specific correlation heatmap. Actigraphy offered insights into temporal trends through hourly mean activity (figure 1) and zero count (figure 2) analysis. When the aggregate mean activity was calculated over different time intervals, different patterns were seen in both groups. By breaking the mean activity by day of the week, additional granularity was obtained. These results provide a strong basis for further modeling and help identify features for efficient pattern detection in depression-related data.

The EDA results emphasize the importance of categorical and temporal dimensions in actigraphy data and ratings. Patterns provide useful insights for model building, such as patterns and different mean activity levels throughout the week. These results will direct the choice of features and shape the modeling approach, helping to create reliable predictive models that help analyze depression-related patterns in the dataset.

**Mean Activity**

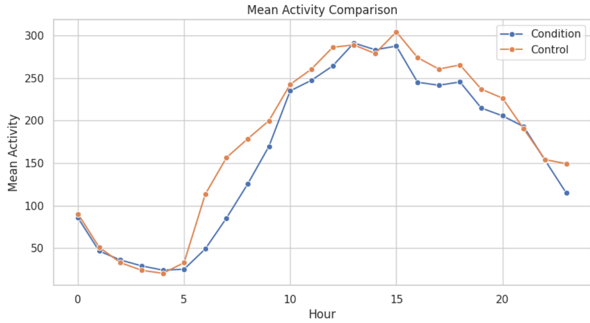


Fig 1.a: Hourly distribution of mean activity

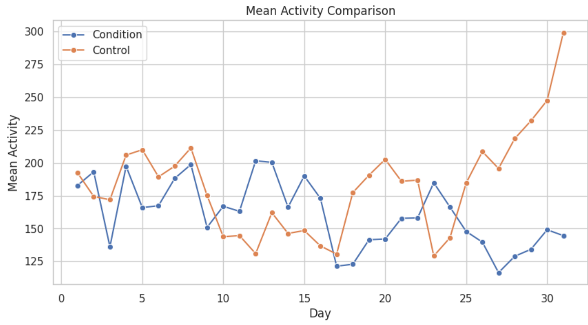


Fig 1.b: Daily (month) distribution of mean activity

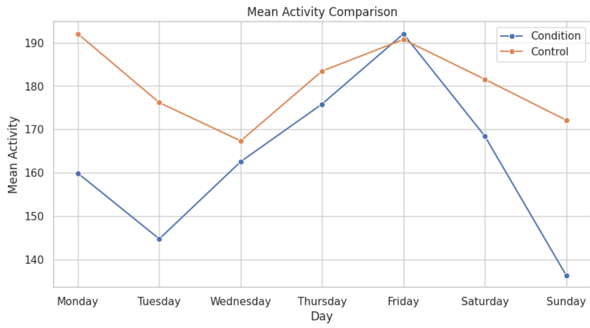


Fig 1.c: Daily (week) distribution of mean activity

**Zero Count**

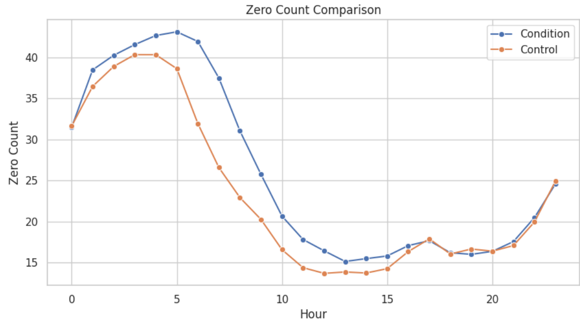


Fig 2.a: Hourly distribution of zero count

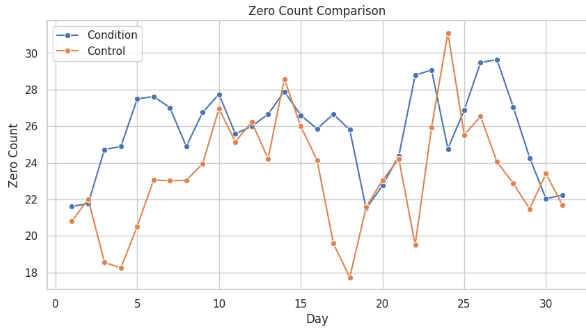


Fig 2.b: Daily (month) distribution of zero count

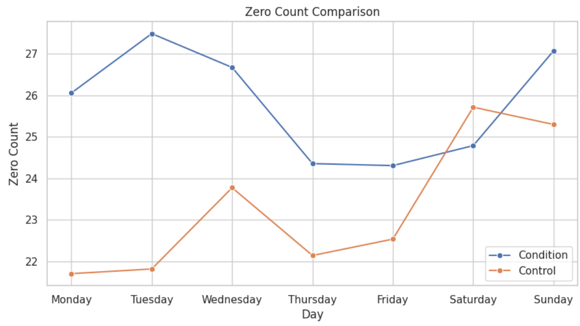


Fig 2.c: Daily (week) distribution of zero count

## 5. Methods

1. Depression State Classification using Machine Learning:
   In the classification task, the focus is on employing supervised machine learning techniques to categorize patients into different depression states, such as unipolar depression or bipolar I/II. The choice of algorithms includes Support Vector Machines, Random Forests, and Neural Networks. To ensure the models make accurate predictions, feature engineering is crucial. This involves selecting and transforming relevant attributes from actigraph data, demographic information, and MADRS scores. The goal is to extract meaningful patterns that can aid in distinguishing between various depression states.

2. Machine Learning to Predict MADRS Scores:
   Moving beyond classification, the analysis extends to regression models designed to predict MADRS scores at the beginning and end of the measurement period. Patient

characteristics and motor activity data serve as the input features for these models. Various regression approaches, such as ensemble methods, decision tree regression, and linear regression, are explored to identify the most effective method for predicting the severity of depression based on the provided features.

3. Comparative Analysis:
   To determine the most effective machine learning classification technique, a thorough comparative analysis is undertaken. This involves evaluating the performance of different algorithms in categorizing depressive states. Additionally, a comparison is made between deep learning approaches, such as Artificial Neural Networks (ANNs), and traditional feature-based methods for time series data. This comparison aims to identify the strengths and weaknesses of each approach and determine the most suitable methodology for the given dataset.

4. Feature-Based Classification:
   A feature-based classification approach is implemented using pertinent demographic and actigraph data features. Classic machine learning techniques are employed for this task. To improve model performance, dimensionality reduction techniques like Principal Component Analysis (PCA) are applied. By reducing the number of features while retaining the most critical information, PCA enhances the efficiency of the machine learning models in classifying depression states.

5. In-Depth Approaches:
   In-depth analysis delves into advanced techniques, particularly deep learning methods, for the time series data generated by actigraphs. Artificial Neural Networks (ANNs) are investigated for their ability to capture complex patterns in the temporal data. This exploration aims to uncover whether deep learning models can outperform traditional machine learning approaches in the context of depression analysis, providing a more nuanced understanding of the temporal dynamics associated with depressive states.

6. Metrics for Evaluation:
   The effectiveness of the classification and regression models is rigorously assessed using appropriate metrics. These metrics include the Area Under the ROC Curve (AUC) for discrimination ability, accuracy for overall correctness, precision for the proportion of true positive predictions, recall (sensitivity) for correctly identifying actual positives, and F1-score for balancing precision and recall. This comprehensive evaluation ensures a thorough understanding of the models' performance across various dimensions and aids in selecting the most suitable approach for the specific task at hand.

## 6. Results

### 6.1. Prediction of whether 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 |

Table 1

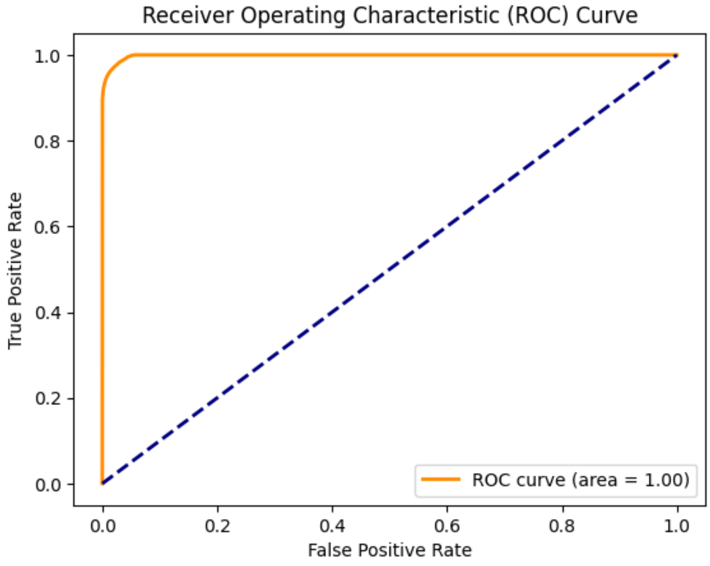


Fig 3: ROC Curve for ANN

In our exploration of predicting depression based on Actigraph Readings, we employed various machine learning models, each leveraging Principal Component Analysis (PCA) for feature reduction. Among the models considered, the Artificial Neural Network (ANN) and Random Forest exhibited exceptional performance. The ANN achieved an impressive accuracy of 97.8%, outperforming the other models. Furthermore, the ANN demonstrated superior precision (98.0%) and recall (96.0%), resulting in a balanced F1-score of 97.0%. The Random Forest model, incorporating cross-validation, also delivered compelling results with an accuracy of 97.6%. While all models demonstrated proficiency, the ANN stands out as the most robust and accurate classifier for predicting depression based on Actigraph Readings in our study. These findings underscore the efficacy of complex non-linear relationships captured by neural networks in handling the intricacies of the dataset, making the ANN the preferred choice for our predictive modeling task. Additionally, the Receiver Operating Characteristic (ROC) curve analysis for the ANN revealed an area under the curve (AUC) of

1.00, further affirming its exceptional discriminative power and suitability for distinguishing between individuals with and without depression based on Actigraph Readings.

### 6.2. Prediction of MADRS score (recorded post-treatment) 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% |

Table 2

In our investigation of predicting the Montgomery-Åsberg Depression Rating Scale (MADRS) scores post-treatment for patients with depression, we evaluated several regression models, including Linear Regression, Ridge Regression, Lasso Regression, Gradient Boosting, XGBoost, and K-Nearest Neighbors (KNN). The Mean Squared Error (MSE) and R-squared (R2) were employed as performance metrics. Among the models examined, K-Nearest Neighbors (KNN) emerged as the most effective predictor, achieving the lowest MSE of 5.712 and the highest R2 of 59.9%. KNN demonstrated superior accuracy and a higher proportion of explained variance, highlighting its capability to capture complex relationships within the data and provide accurate predictions for post-treatment MADRS scores. These results underscore the efficacy of KNN in this predictive modeling task, suggesting its suitability for assessing the mental health outcomes of patients with depression based on the recorded MADRS scores post-treatment.

## 7. Conclusion

To sum up, we have made great progress in our investigation into the prediction of depression using Actigraph Readings and post-treatment MADRS scores. An important milestone has been reached with the successful construction of predictive models, which include the use of K-Nearest Neighbors (KNN) for post-treatment MADRS scores and an Artificial Neural Network (ANN) for Actigraph Reading predictions. A thorough and flexible approach to modeling is shown in the application of a variety of machine learning techniques, from conventional methods like Random Forest to sophisticated deep learning approaches like ANN. Together, these initiatives represent a significant advancement in our knowledge of and ability to forecast depression, demonstrating the potential of these models to make a significant contribution to mental health research.

The remarkable outcomes of the ANN in Actigraph Reading predictions and the KNN model in MADRS score predictions highlight the efficacy of both approaches and validate their significance in enhancing our comprehension and forecasting of depression states. All things considered, this research is a noteworthy advancement in the automation of depression diagnosis, providing opportunities for enhanced therapeutic approaches and a more thorough understanding of this intricate mental health problem in a clear and meaningful way.

## 8. Future Work

The augmentation of the dataset by adding more patients to the disease and control groups should be the top priority for future research. A thorough examination of depression and its complex subfactors, as well as more accurate forecasts of Montgomery-Åsberg Depression Rating Scale (MADRS) scores, depend on this augmentation. A larger and more varied dataset will improve the prediction models' generalizability and robustness, offering a more detailed understanding of the issues surrounding depression. This phase is essential to improving the models' dependability and guaranteeing their efficacy for a wider range of patient characteristics.

In order to provide more thorough research, future work must focus on growing the dataset and incorporating cutting-edge methods. Time series analysis will improve our comprehension of the dynamics of mental health by highlighting complex temporal patterns. Slight changes in patient circumstances can be identified by utilizing Long Short-Term Memory (LSTM) models, which are adept at capturing long-term dependencies and enhancing temporal connection representation. Improved understanding of the relationship between sleep habits and mental health should result from model improvements, which include feature engineering, fine-tuning, and investigating specialized models. This multimodal method seeks to improve prediction power and expand understanding of depression dynamics in a succinct framework.

## 9. Author Contributions

***Akriti Kumari (akkumari)*** and ***Prathamesh More (ppmore)*** collaborated to investigate early research, comprehended the dataset, and carried out exploratory data analysis (EDA). They documented the most important findings from the EDA and conducted a thorough literature analysis, laying the foundation for the next stages.

***Sarvesh Soni (sarvsoni)*** and ***Vedika Halwasiya (vhalwasi***) took the responsibility of developing prediction models, drawing significant conclusions from the EDA, and directing the project's completion and subsequent efforts. They were important in improving the models, guaranteeing precision, and offering insightful analyses of the results.

The project's success was largely due to the authors' cooperative efforts. The authors' distinct areas of competence, which ranged from the literature review to the development of models and conclusion formulation, enabled a comprehensive investigation of depression analysis and laid the groundwork for future developments in the subject.

## 10. References:

[1] E. Garcia-Ceja, M. Riegler, P. Jakobsen, J. T\o rresen, T. Nordgreen, K. J. Oedegaard, and O. B. Fasmer, "Depresjon: A Motor Activity Database of Depression Episodes in Unipolar and Bipolar Patients," in Proceedings of the 9th ACM on Multimedia Systems Conference (MMSys'18), Amsterdam, The Netherlands, 2018, pp. 1-6, doi: 10.1145/3204949.3208125.

[2] J. -G. Choi, I. Ko and S. Han, "Depression Level Classification Using Machine Learning Classifiers Based on Actigraphy Data," in IEEE Access, vol. 9, pp. 116622-116646, 2021, doi:10.1109/ACCESS.2021.3105393.

[3] H. -Y. Kim et al., "Monitoring for disease progression via mathematical time-series modeling: Actigraphy-based monitoring patients with depressive disorder," 2013 IEEE 10th Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, 2013, pp. 56-61, doi: 10.1109/CCNC.2013.6488425.

[4] D. Nová, F. Albert and F. Šniel, "Analysis of actigraph parameters for relapse prediction in bipolar disorder: A feasibility study," 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 2014, pp. 4972-4975, doi: 10.1109/EMBC.2014.6944740.

[5] M. Raihan, A. K. Bairagi and S. Rahman, "A Machine Learning Based Study to Predict Depression with Monitoring Actigraph Watch Data," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-5, doi: 10.1109/ICCCNT51525.2021.9579614.

[6] A. Aminifar, F. Rabbi, V. K. I. Pun and Y. Lamo, "Monitoring Motor Activity Data for Detecting Patients' Depression Using Data Augmentation and Privacy-Preserving Distributed Learning," 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Mexico, 2021, pp. 2163-2169, doi: 10.1109/EMBC46164.2021.9630592.

[7] M. Xu, L. Zuo, S. Iyengar, A. Goldfain and J. DelloStritto, "A semi-supervised Hidden Markov model-based activity monitoring system," 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 2011, pp. 1794-1797, doi: 10.1109/IEMBS.2011.6090511