Smart Smartphones for Mental Health

Aninda Roy Ani[1], Anik Alvi[2], Sabir Saheel[3], Tarem Ahmed[4], Daniel L. Chen[5], Zaki Wahhaj[6], Njogu Mbau[7]
[1]St. Mary's University Twickenham, London, UK
2313005@live.stmarys.ac.uk
[2]New Mexico State University, Las Cruces, New Mexico, USA
aalvi@nmsu.edu
[3]University of Minnesota Duluth, Duluth, Minnesota, USA
sahee002@d.umn.edu
[4]Independent University, Bangladesh (IUB), Dhaka, Bangladesh
tarem@iub.edu.bd
[5]Toulouse School of Economics, Toulouse, France
daniel.chen@iast.fr
[6]King's College London, London, UK
zaki.wahhaj@kcl.ac.uk
[7]Strathmore Business School, Strathmore University, Nairobi, Kenya
njogu.benson@strathmore.edu

***Abstract*—Mental health disorders among university students are a growing concern, particularly in regions with limited access to professional support. This study explores a machine learning-driven approach to predicting mental health status using smartphone sensor data collected via the Beiwe mobile app. Data from 79 university students in Bangladesh, including GPS, power state, Bluetooth, and accelerometer signals, were analyzed to identify behavioral patterns linked to mental well-being. Three regression-based machine learning models, Decision Tree, Support Vector, and Random Forest Regressors, were evaluated, with hyperparameter tuning performed using grid search and 5-fold cross-validation. Results indicate that the Random Forest Regressor achieved the best performance (RMSE = 1.303, R^2 = 0.425), demonstrating its potential in capturing meaningful insights from digital phenotyping data. While findings highlight the feasibility of passive mental health monitoring, the modest predictive accuracy underscores the need for incorporating additional behavioral variables and advanced modeling techniques. This research contributes to the growing field of digital mental health diagnostics, paving the way for scalable, ethical, and data-driven intervention strategies.**

***Index Terms*—K-means Clustering, Random Forest Regressor, Support-Vector Regressor, Decision Tree Regressor**

I. Introduction

This paper will introduce and evaluate an innovative method for identifying people experiencing mental health problems in Bangladesh. A key element of this innovation is the collection and analysis of smartphone phenotype data from university students (sensor data) to make predictions about their mental health status using Machine Learning algorithms. Our objective is to be able to find patterns within the datasets using various clustering algorithms that would allow us to make predictions about a student's mental health status within the context of Bangladesh.

Smartphone usage has increased significantly in recent years, particularly among younger populations. Therefore, we used a research platform in the form of a smartphone application called Beiwe [1], designed to collect and analyze raw data from smartphone sensors and usage patterns. Beiwe has been used successfully in psychiatric research to study mental health of participants based on their smartphone data. We utilize Beiwe's ability to collect and analyze sensor data, administer questionnaires, and communicate with participants over an extended period without face-to-face interaction.

Students in our social environment often keep their mental health problems to themselves and assume that conditions will improve over time, which is not always the case. This issue has been further intensified by the COVID-19 pandemic, particularly in low-resource settings where access to mental health services is limited. A large proportion of university students experience mental health challenges related to academic pressure. Therefore, early identification of such issues is critical.

Our approach would allow us to identify students who suffer from mental health issues by their smartphone usage data. Machine Learning algorithms would allow us to make accurate predictions on a person's mental health status by just analyzing their smartphone usage patterns. Being able to identify Mental Health issues would allow a student to be more confident on whether or not they should be seeking therapy. Which would inevitably result in improved mental wellness all around and promote greater productivity.

The combined analysis of power state and GPS data provides insights into participants' behavioral tendencies, offering a foundation for further exploration of mental health patterns. These findings are consistent with prior research linking mobility and activity patterns with mental health outcomes [2].

A. Our Contribution

This study makes several unique contributions to the growing field of digital mental health research:

- **Geographic Novelty:** To our knowledge, this is the first study using the Beiwe digital phenotyping platform in Bangladesh, and among the first in any low- or middle-income country. Prior Beiwe-based research has mostly occurred in high-income contexts, such as the United States, focusing on clinical populations [1]. This geographic expansion is significant because it tests the feasibility and performance of the Beiwe platform and predictive algorithms in a new cultural and socioeconomic setting [3].
- **Mental Health Focus:** While earlier studies using Beiwe have largely centered on psychiatric conditions such as depression, bipolar disorder, or schizophrenia [4], this study shifts the focus toward general mental health state prediction in a non-clinical population. Specifically, it targets outcomes like stress, mood, and anxiety levels in university students. This work moves beyond passive monitoring into proactive prediction, marking a shift in how Beiwe data is used.
- **Larger Sample Size:** With 79 participants, our dataset is larger than those used in many prior Beiwe studies, which often included only a few dozen individuals [5]. For example, early schizophrenia and depression studies using Beiwe had sample sizes as low as 15 to 16 [4]. A cohort of 79 participants enhances the statistical power and reliability of our findings, especially for single-site data collection efforts [6].
- **Methodological Contributions:** Although Decision Trees, SVMs, and Random Forests are established machine learning models [7]–[9], applying them comparatively on Beiwe-generated data in a low-resource setting is novel. Many earlier studies relied on simpler techniques like penalized regression or clustering with limited predictive success [10]. Our work benchmarks multiple machine learning regressors on the same dataset, offering insights into model performance and laying the foundation for more accurate mental health predictions. This comparative, regression-based modeling approach builds upon and extends prior work, which often focused on classification or basic correlations.

B. *Outline of Paper*

The remainder of this paper is organized as follows. Machine learning algorithms are described in Section II. Section III describes our experimental setup and results. Section IV presents the conclusion and future work.

II. Machine Learning Algorithms

A. *Decision Tree Regressor*

There exists many supervised machine learning algorithm. One of the most basic is the decision tree regressor. This model works recursively by dividing the input space into smaller regions, each of which is linked to a straightforward linear model that forecasts the target variable in light of the input features. The final prediction is the weighted average of the predictions made by all the linear models associated with the leaf nodes of the tree [7].

In order to divide the input space into smaller regions appropriately with the most informative features, there is a need for an objective function to maximize the information gain at each split and so is defined as:

$$IG(D_p, f) = I(D_p) - (\frac{N_left}{N_p} I(D_{left}) + \frac{N_{right}}{N_p} I(D_{right}) \quad (1)$$

where f is the feature to perform split, D_p, D_{left}, D_{right} are the datasets of the parent and child nodes, I is the imputrity measure, N_p is the total number of samples at the parent node, whereas N_{left} and N_{right} are the number of samples in the child nodes.

B. *Support Vector Regressor*

Support Vector Regressor (SVR) is another machine learning algorithm used for non-linear regression analysis. It is a variant of the classification-focused Support Vector Machine (SVM). SVR adopts a similar methology as SVM, but instead of identifying a linear boundary between two classes, it seeks a hyperplane that minimizes the error and best matches the data points [8].

C. *Random Forest Regressor*

Another machine learning algorithm that can be used for regression analysis is the Random Forest Regressor. This algorithm is also an extension of the Random Forest algorithm that is used for classification. The Random Forest Regressor works by building a large number of decision trees and aggregating their predictions to produce a single final prediction [9].

In a random forest regressor, each decision tree is built on a random subset of training data and a random subset of features which reduces overfitting and improves model generalization performance. The final prediction is made by averaging the predictions made by all the decision trees. This makes this algorithm very flexible and can easily handle non-linear relationships between input and target variables.

III. Experimental Setup & Results

A. *Data Collection*

Students at the university were informed about the study and given the option to voluntarily participate at their own discretion. Once we had enough participants, we began emailing invitation links in order for them to download the Beiwe smartphone application. Beiwe began to run in the background once it had been downloaded. Due of Biewe's restricted functionality on iOS, we only incorporated data gathered from Android-powered smartphones.

We collected GPS, WiFi, Powerstate, Bluetooth, Gyroscope and Accelerometer data from each participant. In addition to data from the Beiwe application, we also prompted our participants to respond to online questionnaires pertaining to their mental health, their experiences with aggression (including domestic abuse and gender-based violence), and their attitudes toward domestic and gender-based violence (among both men and women). In order to determine if the smartphone phenotype data and the online mental health surveys are correlated, our next steps will involve employing machine learning techniques.

B. Research Methodology

Our total dataset consisted of 79 participants and have some missing values in the input parameters. To fill these missing points in the dataset, we have computed the values based on the 4-degree nearest neighbor technique.

This sample size is relatively large compared to many earlier Beiwe studies. Early digital phenotyping research often included only a few dozen participants. For example, one pilot study using Beiwe for schizophrenia tracking involved just 16 participants, and another study on depression included around 15 individuals. These small samples were common in initial feasibility studies. A cohort of 79 participants, therefore, represents a larger-than-typical dataset for a single-site Beiwe study, indicating a robust data collection effort. Only in recent years have smartphone-based mental health studies begun to scale to hundreds of participants, often through multi-site or alternative platforms. As such, the 79-person sample in this study enhances the statistical power and reliability of its findings relative to earlier pilot efforts.

To train the regression based machine learning algorithms, we randomly split the dataset into two parts: training (70%) and testing (30%) dataset. Moreover, we used the grid search cross-validation technique to determine the optimum values for the different parameters of the machine learning model. For all of our machine learning models, we implemented 5-fold cross validation grid search technique [11].

This study applied machine learning regressors—specifically Decision Tree, Support Vector Machine (SVM), and Random Forest models—to predict mental health scores from smartphone data. These techniques are well-established in machine learning and have been used in various forms in prior digital mental health research, though not always within the Beiwe platform. Earlier studies often relied on simpler statistical correlations or basic models. For example, some studies used penalized regression (elastic net) or clustering to relate phone sensor features to mood, generally achieving modest accuracy (e.g., AUC ~ 0.65 in one case). Others have tried supervised learning on smartphone data: one team reported using machine learning models (including tree-based algorithms) to predict daily mood, though with mixed success and only modest correlations for most users.

Thus, the use of Decision Trees, SVMs, and Random Forests in this study is not entirely novel—these or similar algorithms have been tested before in mobile sensing contexts. However, this study's contribution lies in applying a suite of machine learning models to Beiwe data in a new population, and in refining the approach. By comparing multiple algorithms (Decision Tree vs. SVM vs. Random Forest) on the same dataset, the study provides insight into what techniques might work best for smartphone-based mental health prediction. This comparative machine learning approach and the focus on regression of well-being scores build upon prior work, which often focused on classification or simple correlations, marking an incremental methodological improvement in the Beiwe literature.

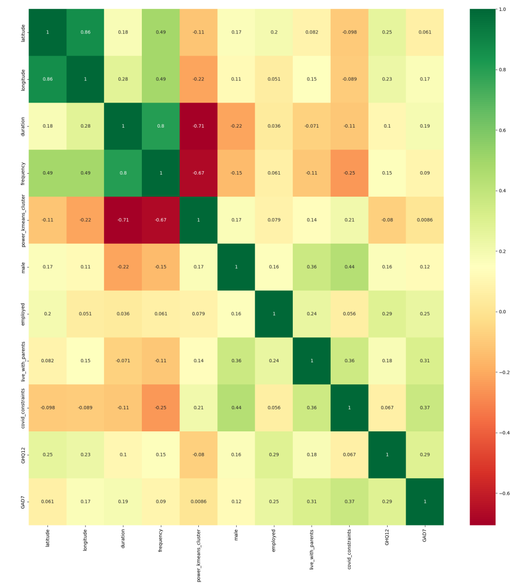

Fig. 1. Correlation Matrix.

C. Feature Analysis

Feature analysis which is also known as feature selection is an important step in any machine learning which helps discover the most relevant and informative features that needs to be selected from any dataset. Feature analysis helps improve the performance and increase efficiency of machine learning models by removing irrelavant features which ends up overfitting the machine learning models.

We have implemented the Correlation matrix to check if any of the input parameters depends upon one another. Figure 1 shows the correlation matrix for our given dataset where it can seen that there exists not much correlation between any of the variables. As a result, we have taken into account all the parameters while running the machine learning models.

D. Results and Analysis

The goal of this research is to find correlation between digital phenotype data and survey data for each participant to check if it matches with his/her mental health condition. To evaluate the performance of the machine learning algorithms, we have calculated the Root Mean Square Error (RMSE) and R^2 for each of these algorithms.

The analysis revealed distinct patterns in participants' smartphone usage and mobility behavior. For the power state data,

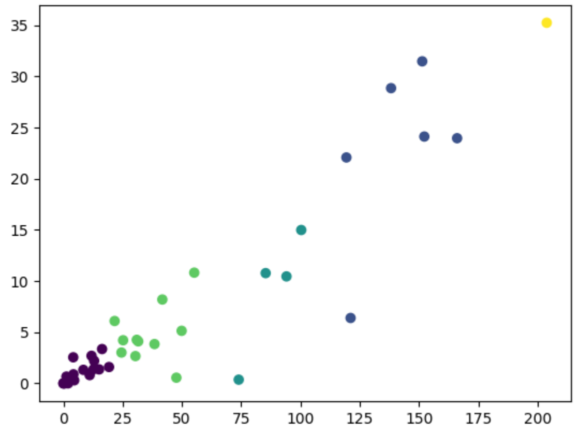

Fig. 2. K-Means clustering for power state data showing smartphone usage patterns.

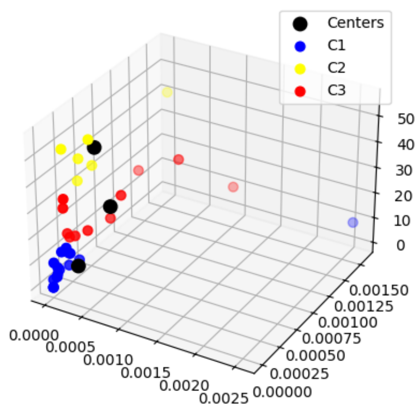

Fig. 3. 3D clustering for GPS data illustrating movement patterns.

the K-Means clustering algorithm identified four clusters, distinguishing participants based on their overall usage duration and night-time phone activity. In the 2D plot in Figure 2, participants are closer to the X-axis exhibited higher day-time usage, while those closer to the Y-axis showed higher night-time usage. Those farther from the origin had significantly higher overall smartphone usage.

The GPS data, analyzed in 3D space in Figure 3, grouped participants into three clusters based on their mobility patterns, with outliers indicating participants with unusually high movement levels. Further analysis of night-time mobility employed standard deviation to detect participants with irregular movement patterns, enhancing the precision of outlier identification.

The combined analysis of power state and GPS data provides insights into participants' behavioral tendencies, offering a robust foundation for further exploration of mental health

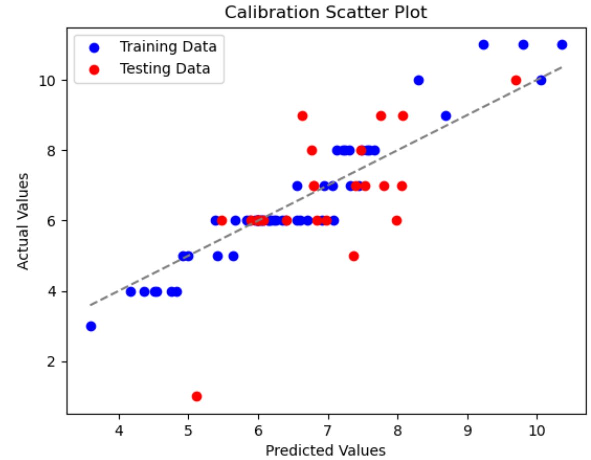

Fig. 4. Calibration plot of the total actual values and prediction values.

and lifestyle traits.

Table I shows the results obtained by using the machine learning algorithms. Based on the results, Random Forest Regressor performs the best among the other machine learning algorithms which RMSE of 1.303 and R^2 of 0.425 which shows the algorithm is able to predict 42.5 % of the data considering predicting mental health can be challenging as it depends on diverse number of factors.

Moreover, we have implemented calibration plot for our dataset to evaluate the obtained results shown in Figure 4. The calibration plot in the Figure 4 shows how close the data points are from the line of best fit. The figure also shows it fits the training data better than the testing data and therefore requires further study to be able to fit the data better.

IV. Conclusion and Future Work

In this study, we explored the potential of leveraging smartphone sensor data collected via the Beiwe app to predict mental health status using various machine learning approaches. Our experimental results indicate that while ensemble methods such as the Random Forest Regressor showed promise—with an RMSE of 1.303 and an R² of 0.425—the challenge of accurately modeling mental health from digital behavioral data remains significant. The modest predictive performance suggests that although meaningful patterns exist within the smartphone data, the complex nature of mental health likely requires the incorporation of additional variables and more sophisticated analytical techniques. Future work should focus

TABLE I
Evaluation of Regression-Based Machine Learning Algorithms for Mental Health Prediction.

Machine Learning Algorithms	**RMSE**	R^2
Decision Tree Regressor	1.744	0.029
Support Vector Regressor	1.583	invalid
Random Forest Regressor	1.303	0.425

on expanding the dataset to include a larger and more diverse participant pool, which would help in generalizing the findings and improving model robustness. Exploring advanced machine learning frameworks, including deep learning approaches, may also provide a better understanding of the nonlinear relationships inherent in digital phenotyping data. Finally, addressing ethical considerations such as data privacy and participant consent will be critical as this research moves toward real-world applications.

REFERENCES

[1] J. Torous, J.-P. Onnela, and M. Keshavan, "New dimensions and new tools to realize the potential of rdoc: digital phenotyping via smartphones and connected devices," *Translational Psychiatry*, vol. 6, no. 3, p. e1053, 2016. [Online]. Available: https://doi.org/10.1038/tp.2016.43

[2] L. Canzian and M. Musolesi, "Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis," *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 1293–1304, 2015.

[3] A. Cohen, J. F. Huckins, J. L. Schleider, and J.-P. Onnela, "Forecasting mental states in schizophrenia using digital phenotyping data across the us and india," *PLOS Digital Health*, vol. 2, no. 1, p. e0000144, 2023. [Online]. Available: https://doi.org/10.1371/journal.pdig.0000144

[4] I. Barnett, J. Torous, P. Staples, L. Sandoval, M. Keshavan, and J.-P. Onnela, "Relapse prediction in schizophrenia through digital phenotyping: a pilot study," *Neuropsychopharmacology*, vol. 43, no. 8, pp. 1660–1666, 2018. [Online]. Available: https://doi.org/10.1038/s41386-018-0030-z

[5] N. C. Jacobson, S. M. Summers, and J. Wilhelm, "Digital biomarkers of social anxiety severity: digital phenotyping using passive smartphone sensors," *Journal of Medical Internet Research*, vol. 21, no. 7, p. e13204, 2019. [Online]. Available: https://doi.org/10.2196/13204

[6] J. Meyerhoff, S. Liu, M. Kaimal, and J. Torous, "Correlations between digital phenotyping data and mental health symptoms in college students: analysis and replication," *BJPsych Open*, vol. 7, no. 6, p. e190, 2021. [Online]. Available: https://doi.org/10.1192/bjo.2021.996

[7] S. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 3, pp. 660–674, 1991. [Online]. Available: https://doi.org/10.1109/21.97458

[8] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in Neural Information Processing Systems*, M. Mozer, M. Jordan, and T. Petsche, Eds., vol. 9. MIT Press, 1996. [Online]. Available: https://rb.gy/cfi67

[9] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001. [Online]. Available: https://doi.org/10.1023/a:1010933404324

[10] A. Pratap, D. C. Atkins, B. N. Renn, J. Volponi, S. D. Mooney, and J. A. Anguera, "The accuracy of passive phone sensors in predicting daily mood," *npj Digital Medicine*, vol. 2, p. 7, 2019. [Online]. Available: https://doi.org/10.1038/s41746-019-0078-3

[11] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014. [Online]. Available: https://arxiv.org/abs/1412.6980