Journal of Information Systems Engineering and Management
2025, 10(8s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Detection of Postpartum Depression using Machine Learning

Smita Bhagwat[1], Ketki Kshirsagar[2], Alka Barhatte[3], Swati Patil[4], Jayashree Tamkhade[5], Dipti Pandit[6]
[1,2]Vishwakarma Institute of Technology Pune, India.
[3]Dr. Vishwanath Karad MIT World Peace University, Pune, India.
[4,5,6]Vishwakarma Institute of Information Technology Pune, India.
[1]smita.bhagwat@vit.edu, [2]ketki.kshirsagar@vit.edu, [3]alka.barhatte@mitwpu.edu.in, [4]swati.patil@viit.ac.in, [5]jayashree.tamkhade@viit.ac.in, [6]dipti.pandit@viit.ac.in

ARTICLE INFO

Received: 15 Oct 2024

Revised: 12 Dec 2024

Accepted: 25 Dec 2024

ABSTRACT

Postpartum depression (PPD) is a prevalent and serious mental health issue that affects a significant number of women after childbirth. Early detection of PPD is crucial for effective intervention, but traditional screening tools often fall short in terms of accuracy and timely prediction. This study focuses on extensive machine learning methods by using psychological and behavioural data from a large survey of postpartum women. Various traditional machine learning algorithms along with CatBoost are implemented in this research. The accuracy of 95.68% demonstrate that the CatBoost-based model outperforms other machine learning models, providing a robust and reliable method for the early detection of postpartum depression.

Keywords: PPD, CatBoost, Gradient Boosting, Women mental health, Machine Learning

INTRODUCTION

Postpartum depression (PPD) is a concerning mental health problem that affects 10-20% of women all over the world after childbirth, but actual count may be higher due to stigma, lack of awareness, and limited healthcare facilities [1]. In the postpartum time, new mothers face lots of emotional, physical, and psychological changes, which make them suffer from mental health problems. PPD hinders a mother's quality of life and impacts her capability to care for her newborn children, which causes long-term developmental problems for the child and causing bad family bonding. Classic methods for diagnosing PPD, such as Edinburgh Postnatal Depression Scale (EPDS) [2] and self-assessing questionnaires, mostly depends on a mother's willingness to reveal about her mental state, which can be manipulated by shame or fear [3]. These limitations create the need for more correct diagnostic tools. Machine learning methods provides a reliable solution for on-time detection and diagnosis of PPD [4]. Instead of classical methods that depends on static data assessed during doctor visits, machine learning models can analyse big and varied data in real time, such as electronic health records, wearable technologies, and handheld device usage [5]. Such algorithms can notice minor patterns and early symptoms of PPD, such as changes in sleep, physical activity, or mood fluctuations, which provides patient specific risk predictions for each mother [6]. This will prevent symptoms from going out of hand and make sure postpartum women receive the support they need.

LITERATURE REVIEW

The reviewed literature explores various studies using a variety of machine learning models intended for predicting mental health conditions in diverse populations [4]. Decision trees and Support Vector Machine with feature selection for prediction of stress, depression, and anxiety in students. Whilst study used a Random Forest model with data balancing techniques to improve prediction sensitivity and specificity for depression in women [5,6]. Mental health issues in employees is studied using Logistic Regression and Decision Tree models [7,8]. Few studies show Bayes Net for effective comparison of six classifiers for depression prediction [9-11]. Social media data is used to predict depression and anorexia using emotional patterns and lexicon-based approaches [8]. EMG signal classification [12], with the use of intensive feature extraction, and dimensionality reduction techniques. Their work stressed the fact that precise data processing is essential for improving diagnostic accuracy which is a principle that can be used to postpartum depression (PPD) prediction. Surveys and medical records [13] for feature-based depression and anxiety prediction, while custom CNN model [14] to detect mental health changes in students, revealing improved performance over standard CNN methods. For PPD prediction, Random Forest [15] targeting on maternal health features and SVM and Random Forest to a small dataset in China for prediction of PPD in women and minimized features while preserving accuracy [16]. Wearable technology data to capture real-time indicators like heart rate and sleep to predict PPD in women [17-19]. These studies mostly focus on binary classification for depression or PPD, and overlooked differentiation between levels of severity such as mild, moderate, or severe depression. The proposed study aims to develop a model capable of classifying depression into three categories “no depression, mild depression, and severe depression” while balancing classes, for enhancing diagnostic precision, and supporting more patient specific treatments.

METHODS

In this study, the extensive dataset is collected, which contains responses of various psychological and physical symptoms from postpartum women. Then we systematically preprocess this data, to enhance its quality, which makes it appropriate for accurate analysis and model training. We then perform an exploratory data analysis (EDA) to discover important patterns, correlations, and distributions of symptoms, for a thorough understanding of how different elements contribute to PPD. The following sections provide details for each stage of this methodology, from data gathering to feature selection and model development.

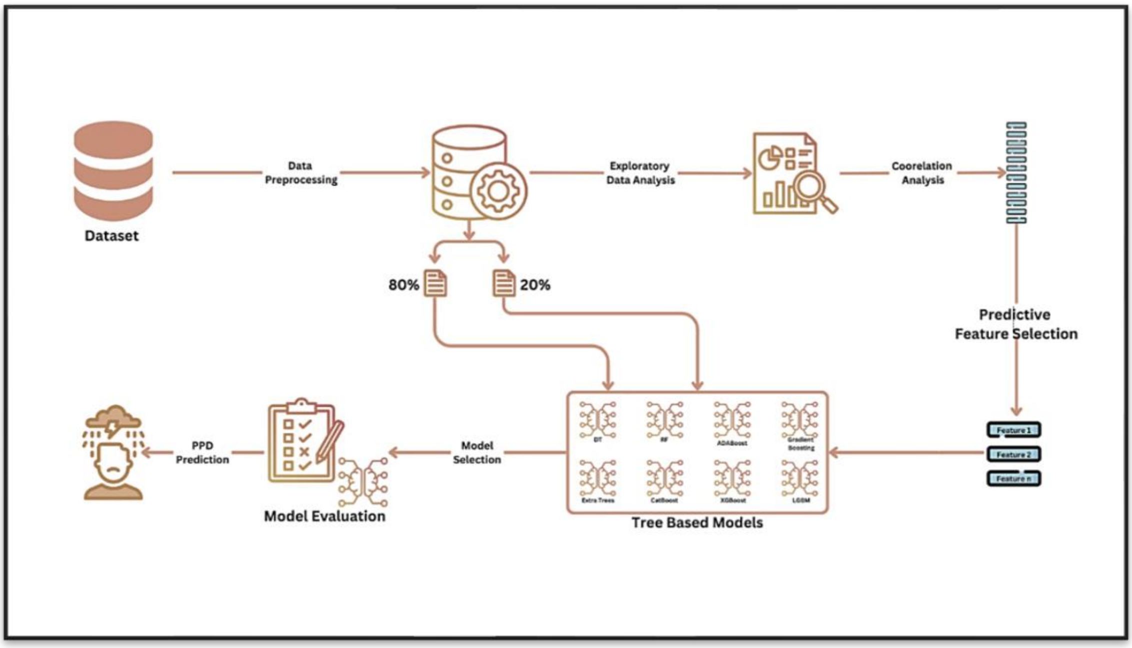


Figure 1: Proposed Method

3.1 Data Collection

In this study, the data of postpartum mental health symptoms were gathered from a dataset on Kaggle [20], which provides responses of 1,503 postpartum women. This dataset was originally collected via a Google Form survey and is publicly available. It includes multiple features that indicates various symptoms related with PPD, such as feelings of sadness, sleep difficulties, irritability, appetite changes, anxiety, guilt, and bonding issues with the baby.

3.2 Data Preprocessing

During preprocessing, the missing values, data formats, and data transformations were done to ensure the data is suitable for machine learning model. Removing the not relevant columns like 'Timestamp' and cleaning the column names to be more readable and consistent were some of the steps taken in ensuring data quality (for example, 'Feeling sad or Tearful" became 'feeling_sad_or_tearful'). The missing values in the categorical variables were treated with the mode and categorical variables were encoded numerically. The responses Yes/No for binary features were represented by 0 and 1 and ordinal features maintained their natural order. Continuous variables were rescaled in the case they were not balanced and the class imbalance, particularly in the variable of 'Suicide attempt,' was handled by the use of SMOTE.

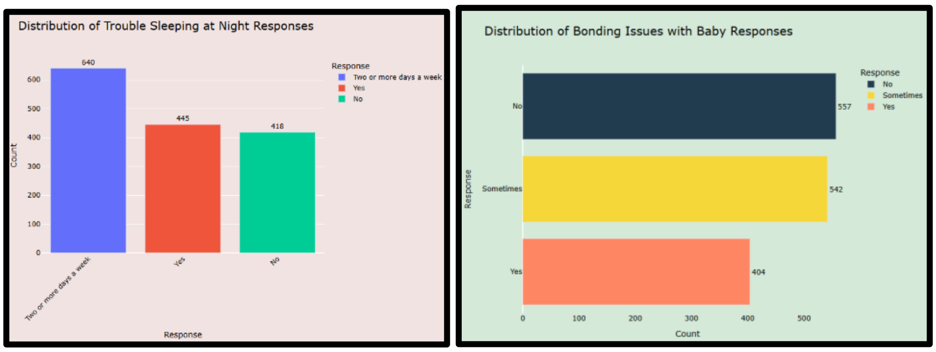
Table 1: Dataset Summary

Table 1 shows the summary of PPD dataset. The different symptoms with their respective number of samples are given.

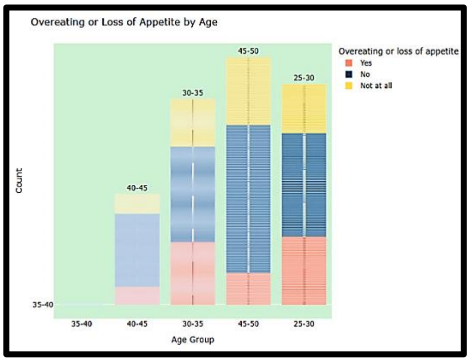
Symptom	Count	Unique	Most Common Value
Age	1503	5	40-45
Feeling sad or Tearful	1503	3	Yes
Irritable towards baby & partner	1497	3	Yes
Trouble sleeping at night	1503	3	Two or more days a week
Problems concentrating or making decision	1491	3	No
Overeating or loss of appetite	1503	3	No
Feeling anxious	1503	2	Yes
Feeling of guilt	1494	3	No
Problems of bonding with baby	1503	3	No
Suicide attempt	1503	3	No

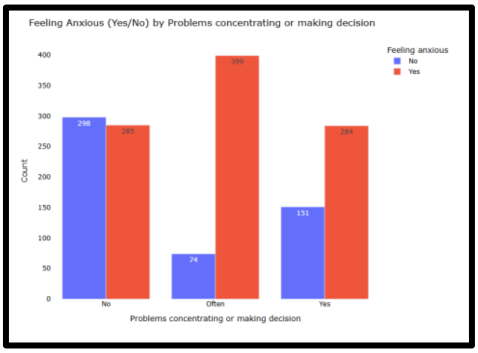
3.3 EXPLORATORY DATA ANALYSIS (EDA)

Exploratory Data Analysis (EDA) was performed to understand the distribution of symptoms, examine patterns in the dataset, and identify correlations between different variables. Various visualizations and statistical measures are used in order to gather perception into the frequency and relationships of various symptoms, to decide in feature selection for the machine learning model.

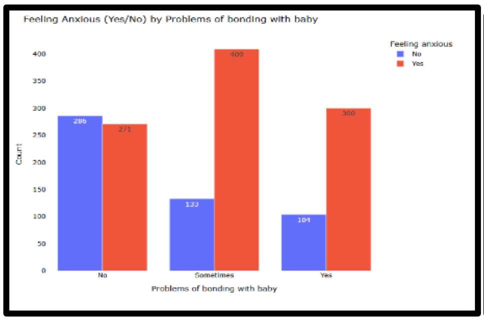


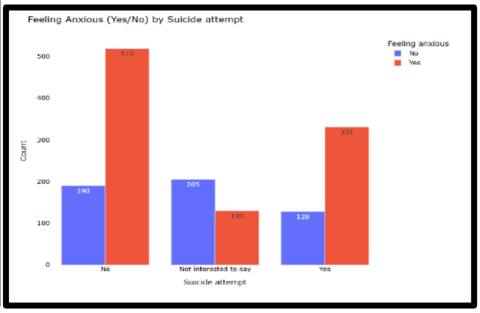
(a) **(b)**





(c) **(d)**





(e) **(f)**

Figure 2: Distribution of (**a**) trouble sleeping at night, (**b**) bonding issues with baby, (**c**) overeating or loss of appetite, (**d**) feeling anxious by making decisions, (**e**) feeling anxious by problems with bonding with baby, (**f**) feeling anxious by suicide attempt

In the figure 2 (a to f), the key symptoms related with postpartum depression such as sleep disturbances, bonding issues, appetite changes, anxiety, concentration problems, and suicidal thoughts is analysed. In fig.2(a), two or more days a week was the most common response for sleep disturbances between 640 number of women, indicating sleep issues as an important predictor. Also, 445 women had less frequent sleep issues, and 418 women had none, which provides a control group to understand sleep's impact on mental health. The fig.2(b) shows that the bonding issues with the baby were commonly not present in 557 women, but 542 shown infrequent problems and 404 shown consistent problems, relating with depressive symptoms and considering it as a secondary predictor. In fig.2(c), it shows that both overeating and loss of appetite were highest in the 45-50 age group, but younger mothers also shown appetite changes. Fig.2(d) indicates that women which are constantly feeling anxious also faced problems with concentration and decision-making, which suggests a strong correlation. And in fig.2(e), the anxiety mostly coincided with bonding problems, with the highest number of anxious women facing significant bonding problems. Lastly, in fig.2(f), it is found that the anxiety strongly correlates with suicidal thoughts, showing the importance of anxiety in predicting postpartum depression.

In figure 3, the Theil's U heatmap, darker green cells shows strong correlations, and lighter cells show weak relations. Important relationships revealed are "feeling sad or tearful," "suicide attempt," "trouble sleeping at night," and "problems bonding with baby." For example, the relation between "feeling anxious" and "trouble sleeping" match with research on anxiety and sleep problems. "Suicide attempt" also shows noticeable relation with other symptoms, suggesting its importance for predictive modelling.

3.4 MODEL BUILDING

This study uses various tree-based machine learning models to predict postpartum depression. The dataset was split into training set of 80% and testing set of 20% ratio to assess each model's performance. A detailed explanation of each model, and their performance metrics is given below.

Decision Tree Classifier: Decision Tree model uses a tree-like structure, in which data is divided into branches, and each branch represents a possible outcome. Random Forest Classifier: Random Forest is an ensemble technique that builds multiple decision trees and combines their outputs. AdaBoost Classifier: AdaBoost is an adaptive boosting method that corrects errors of previous models repetitively, mostly used for improving weak

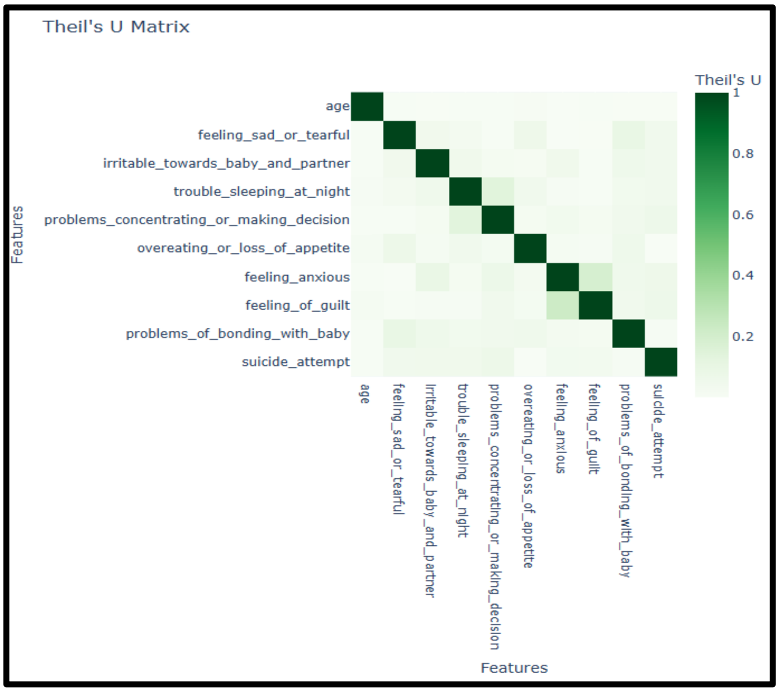


Figure 3: Theil's U Matrix

classifiers. Gradient Boosting Classifier: Gradient Boosting is an ensemble technique that combines weak models to improve them sequentially by minimizing the artifacts of previous models. LightGBM (LGBM) Classifier: LGBM is a gradient boosting framework that handles large datasets by using leaf-wise growth. XGBoost Classifier: XGBoost optimizes both speed and model performance. Extra Trees Classifier: Extra Trees is an ensemble method which works similar to Random Forest, but it builds trees using random splits to reduce variance. CatBoost Classifier: CatBoost, which is also a gradient boosting algorithm, and it is optimized for categorical data. Bagging Classifier: Bagging is an ensemble method which combines multiple versions of the same base learner.

RESULTS

The parameters such as precision, recall, F1 score and accuracy are used to evaluate the results

Precision: Measures the proportion of correctly predicted positive cases out of all cases predicted as positive. It highlights the model's accuracy in identifying true positives.

$$Precision = \frac{True\ Positives\ (TP)}{TP + False\ Positives\ (FP)} \quad \text{...........(1)}$$

Recall (Sensitivity): Assesses the model's ability to identify all actual positive cases. It shows how well the model captures true positives.

$$Recall = \frac{True\ Positives\ (TP)}{TP + False\ Negatives\ (FN)} \quad \text{..........(2)}$$

F1 Score: Provides a harmonic mean of precision and recall, offering a balance between them, especially useful for imbalanced datasets.

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad \text{..........(3)}$$

Accuracy: Represents the proportion of correctly predicted instances (both positive and negative) out of the total instances.

$$Accuracy = 2 \cdot \frac{True\ Positives\ (TP) + True\ Negatives(TN)}{Total\ Instances} \qquad ....(4)$$

Table 2: Results

Model	**Accuracy**	**Precision**	**Recall**	**F1 Score**
Decision Tree	94%	94%	94%	93%
Random Forest	94%	94%	94%	94%
AdaBoost	67%	67%	67%	66%
Gradient Boosting	83%	83%	83%	83%
LGBM	95%	95%	95%	95%
XGBoost	95%	95%	95%	95%
Extra Trees	94%	95%	94%	94%
CatBoost	**96%**	**95%**	**95%**	**95%**
Bagging Classifier	94%	94%	94%	94%

Table 2 shows the results obtained from various machine learning models. CatBoost performed as the best-performing model, scoring the highest test accuracy, precision, recall, and F1 score. Its ordered boosting method and powerful handling of categorical attributes contributed to its excellent performance, which makes it perfect choice for predicting postpartum depression in our dataset.

4.1 MODEL EVALUATION OF CATBOOST:

CLASSIFICATION REPORT:

Table 3: Classification Report scores

	precision	**recall**	**f1-score**	**Support**
Non depressed	95%	98%	96%	136
Mildly depressed	96%	90%	93%	70
Severely depressed	95%	95%	95%	95
Accuracy	95%	95%	95%	95
Micro avg	95%	94%	95%	301
Weighted avg	95%	95%	95%	301

Table 3 shows the classification report of Catboost classifier.

1.2 MODEL PERFORMANCE SUMMARY:

The figure 9 shows the confusion matrix indicates that the model performs good in differentiating between classes, in which Class 0 accomplishing 134 correct predictions and minor errors. Class 1 shows some wrong classifications, in which 3 cases were distinguished as Class 0 and 4 as Class 2. Class 2 carry out well by accomplishing 91 correct predictions and a few wrong classifications as Class 0. This shows that the model is good in differentiating between serious and mild cases, which is a lot important to reduce false diagnosis.

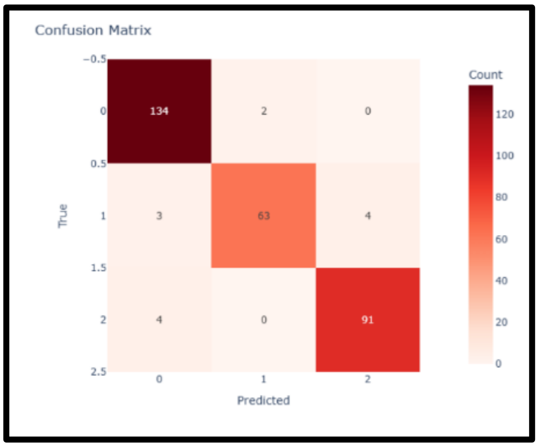


Figure 9: Confusion Matrix

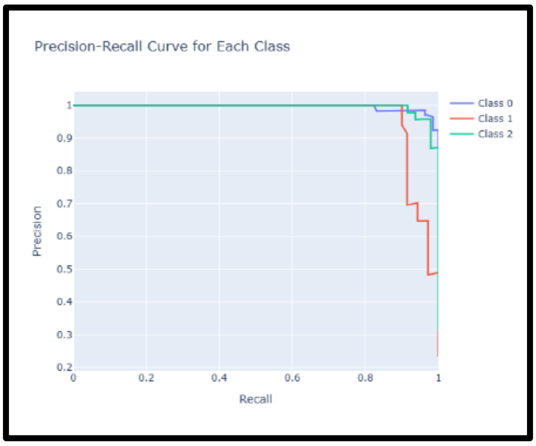


Figure 10: Precision Recall curve

The figure 10 shows the Precision-Recall curves suggests that the model stabilizes precision and recall efficiently, specifically for Class 0 and Class 2, which show high results across different thresholds. Class 1 shows more exchanges between precision and recall, which suggests the model's careful approach to mild cases.

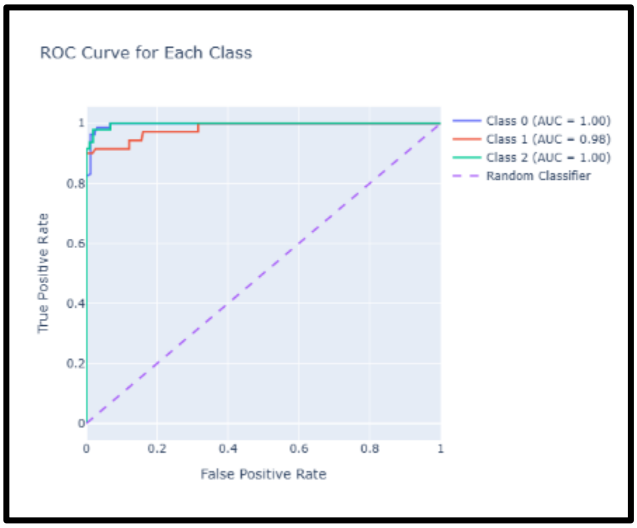


Figure 11: ROC curve

Table 4: Comparison with previous studies

Study	Model(s) Used	Target Classes	Accuracy	Precision	Recall	F1 Score	AUC
[14]	Random Forest	2	79%	83%	73%	77%	88%
[15]	SVM	2	78%	68%	0.69%	68%	78
[16]	Random Forest	2	-	84%	0.79%	81%	81%
[17]	Logistic Regression	2	69%	70%	-	-	82%
[18]	SVM	2	91%	67%	68%	67%	86%
This Study	CatBoost	3	95%	95%	95%	95%	98%

Figure 11 shows the ROC curves shows the model's great performance, where it shows 1.00 AUC of scores for Class 0 and Class 2, and a good AUC of 0.98 for Class 1, indicating its strong capability to distinguish between levels of postpartum depression.

Table 4 shows the comparison of this study with the other existing study carried out. The results shows that the CatBoost classifier outperforms the traditional classifiers. The results shows better performance by Catboost algorithm.

DISCUSSION

The paper demonstrates that machine learning, in general, and the CatBoost algorithm, in particular, can very successfully predict postpartum depression (PPD) from survey data that include psychological and behavioral measurements. CatBoost has shown a maximum accuracy rate of 95.68% which is higher than other machine learning models such as Random Forest and XGBoost concerning accuracy, precision, recall, and F1 score. The model's capabilities in interpreting categorical data along with its ability to deal with class imbalance issues were indispensable to the accurate classification of PPD severity, pointing it out as a trustworthy diagnostic means. Unlike most of the traditional ways of diagnosis that mainly involve self-reporting and clinical observation, this technique is data-driven and consequently more consistent, timely, and hence, will likely improve the outcomes of postpartum women.

However, challenges still remain. The many facets of the survey-based approach to collecting data for this study may fail to reflect the real-time physiological and behavioral indicators of PPD, which could be more apt to be found in wearables or digital biomarkers. Also, by combining CatBoost with deep learning models, specifically for overlapping symptoms, the result might be more accurate. Hence, future research should emphasize drawing valid conclusions using this technique in heterogeneous populations and adopting multimodal data sources to, in consequence, better the predictive power, ending, in more personalized mental health care.

CONCLUSION

This study shows the transformative power of machine learning in the process of dealing with postpartum depression (PPD), a very important but often overlooked mental health issue. The CatBoost model proved its accuracy by reaching a very high result of 95.68%, thus confirming its suitability for this particular application. The new methods employ dynamic real-time monitoring and patient-specific diagnoses rather than being over reliant on subjective self-reporting and clinical assessments. These are the innovations that empower immediate interventions; thus, the symptoms are prevented from the overgrowth and the whole family is in better health prospects. Besides that, the study illustrates the robustness of the CatBoost model in dealing with the problem of imbalanced data, which is quite common in the medical research realm. This property makes it possible to make dependable forecasts even in such situations, thus, it is a valuable instrument that can be utilized for health care analytics. This research seeks to demonstrate ways in which machine learning can act as a bridge in the traditional methods to mental health, thereby making the solutions more effective and, most importantly, accessible to all.

FUTURE WORK

The future of prostate prediction diagnosis should be more about collecting data from the portable devices, like the fitness trackers and smartwatches, than concentrating solely on the analytical capability of the models. For example, using such technologies, we can record and capture the required data from the patient's body to the relevant database, and then it can be shown on the screen of the device. Furthermore, such research can be broadened to out of the mainstream groups thus addressing the issues related to health care accessibility, as well as making it possible to have a more general model that can be applied to different people's cultural and demographic backgrounds. Additionally, the utilization of CatBoost and a deep learning framework provides another way of improving regression neural networks to better capture the nonlinear characteristics of data. These models view this as a bridge for intercommunication between the domains to propel them to the point where they can exert control over their data and extract by the use of the other method the full arrays of elementary patterns. In doing so, these potential projects scientists can explore a variety of methods to producing high-quality, cost-effective, and accessible psychologically guided modes for healthcare.

REFERENCES

[1] Khooharo, Y., Majeed, T., Das, C., Majeed, N., Majeed, N., & Choudhry, A. M. (2010). Associated risk factors for postpartum depression presenting at a teaching hospital. *Annals of King Edward Medical University*, *16*(2), 87-87.

[2] Levis, B., Negeri, Z., Sun, Y., Benedetti, A., & Thombs, B. D. (2020). Accuracy of the Edinburgh Postnatal Depression Scale (EPDS) for screening to detect major depression among pregnant and postpartum women: systematic review and meta-analysis of individual participant data. *bmj*, *371*.

[3] Al-Antari, M. A. (2023). Artificial intelligence for medical diagnostics—existing and future aI technology!. *Diagnostics*, *13*(4), 688.

[4] Bhagwat, S., & Mukherji, P. (2020, August). Temporal Feature Extraction for Improving Myoelectric based Recognition of Prosthetic Hand. In *2020 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)* (pp. 67-71). IEEE.

[5] Mahbub, M. A., Nishat, M. M., Ahmed, D. A., & Mamun, M. A. A. Anxiety Detection Among Engineering Students: The Impact of Perpetual Connectivity to Mobile Devices. *Available at SSRN 4621068*.

[6] Xin, L. K. (2021, March). Prediction of depression among women using random oversampling and random forest. In *2021 International Conference of Women in Data Science at Taif University (WiDSTaif)* (pp. 1-5). IEEE.

[7] Katarya, R., & Maan, S. (2020, December). Predicting mental health disorders using machine learning for employees in technical and non-technical companies. In *2020 IEEE International Conference on Advances and Developments in Electrical and Electronics Engineering (ICADEE)* (pp. 1-5). IEEE.

[8] Usman, M., Haris, S., & Fong, A. C. M. (2020, December). Prediction of depression using machine learning techniques: A review of existing literature. In *2020 IEEE 2nd International Workshop on System Biology and Biomedical Systems (SBBS)* (pp. 1-3). IEEE.

[9] Aragon, M. E., Lopez-Monroy, A. P., González-Gurrola, L. C., & Montes-y-Gómez, M. (2021). Detecting mental disorders in social media through emotional patterns-the case of anorexia and depression. *IEEE transactions on affective computing*, *14*(1), 211-222.

[10] Nasim, S., Al-Shamayleh, A. S., Thalji, N., Raza, A., Abualigah, L., Alzahrani, A. I., ... & Abd Elminaam, D. S. (2024). Novel Meta Learning Approach for Detecting Postpartum Depression Disorder Using Questionnaire Data. *IEEE Access*.

[11] Skaik, R. S., & Inkpen, D. (2022). Predicting depression in Canada by automatic filling of Beck's depression inventory questionnaire. *IEEE Access*, *10*, 102033-102047.

[12] Mukherjee, S., Rintamaki, L., Shucard, J. L., Wei, Z., Carlasare, L. E., & Sinsky, C. A. (2022). A statistical learning approach to evaluate factors associated with post-traumatic stress symptoms in physicians: Insights from the COVID-19 pandemic. *IEEE Access*, *10*, 114434-114454.

[13] Bhagwat, S., & Mukherji, P. (2020). Electromyogram (EMG) based fingers movement recognition using sparse filtering of wavelet packet coefficients. *Sādhanā*, *45*, 1-11.

[14] BABU, P. B., JAYALAKSHMI, V., HARSHA, P. N. V. S., KARTHIK, P., & SADHANALA, V. A. (2024). AN EFFCIET FORCASTING MENTAL HEALTH CONDITION USING MACHINE LEARNING. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, *15*(1), 146-150.

[15] Shan, X. (2024). College Students' Mental Health Prediction Model Based on Time Series Analysis. *Journal of Electrical Systems*, *20*(7s), 2952-2963.

[16] Shin, D., Lee, K. J., Adeluwa, T., & Hur, J. (2020). Machine learning-based predictive modeling of postpartum depression. *Journal of clinical medicine*, *9*(9), 2899.

[17] Zhang, W., Liu, H., Silenzio, V. M. B., Qiu, P., & Gong, W. (2020). Machine learning models for the prediction of postpartum depression: application and comparison based on a cohort study. *JMIR medical informatics*, *8*(4), e15516.

[18] Hurwitz, E., Butzin-Dozier, Z., Master, H., T O'Neil, S., Walden, A., Holko, M., ... & Haendel, M. A. (2024). Harnessing Consumer Wearable Digital Biomarkers for Individualized Recognition of Postpartum Depression Using the All of Us Research Program Data Set: Cross-Sectional Study. *JMIR mHealth and uHealth*, *12*(1), e54622.

[19] Saqib, K., Khan, A. F., & Butt, Z. A. (2021). Machine learning methods for predicting postpartum depression: scoping review. *JMIR mental health*, *8*(11), e29838.

[20] Miller, L. J. (2002). Postpartum depression. *Jama*, *287*(6), 762-765.

[21] Mosaraf, M. P. (2023). *PostPartum Depression [Data set]. Kaggle*.