Machine Learning-Based Diagnosis of Postpartum Depression Methods and Insights

Neesha Nayee [1,*], Saurabh Shah [2], Hetal Bhaidasna [3]

[1]Parul Institute of Engineering and Technology, Parul University, Gujarat, India
[2] Parul Institute of Engineering and Technology, Parul University, Gujarat, India
[3]Parul Institute of Engineering and Technology Diploma-Studies, Parul University, Gujarat, India
neeshakumari.dholakiya31948@paruluniversity.ac.in

Abstract A serious mental health condition that affects new moms, postpartum depression (PPD) frequently has a negative impact on both the mother and the child. Conventional screening techniques, including the Edinburgh Postnatal Depression Scale (EPDS), limit early intervention because they mainly concentrate on detecting symptoms after they have started. In this study, the scope for further development of ML methods in early detection of PPD is investigated based on the analysis of a Kaggle dataset with 1,503 entries. To classify the risk of PPD following these influential factors, ML algorithms were employed, including Support Vector Machine (SVM), XGBoost, Random Forest, Decision Tree, and RBFNN. Such as Support Vector Machine (SVM), XGBoost, Random Forest, Decision Tree, and K-Nearest Neighbors (KNN). On the other hand, XGBoost and Random Forest tree-based models, when used in conjunction with Recursive Feature Elimination (RFE), obtained the highest accuracy (96.99%), and these two are the best performing models of them all. through accuracy, precision, recall and F1-score based as performance evaluation. These results demonstrate that ML-based techniques may facilitate early detection of PPD, which may lead to early interventions. Future work will focus on enhancing model generalizability for clinical utility, incorporating multimodal data, as well as enlarging datasets.

Keywords Postpartum Depression, machine learning, SVM, RF, XGBoost, Decision Tree, Kaggle Dataset.

1 Introduction

Postpartum depression (PPD) is a serious mental condition and it is prevalent during the first year after delivery. It happens among 1 in 7 moms in the U.S. and the global incidence is 13 to 19 percent that is subject to sociodemographic conditions. PPD is a disease in which the mother has enduring emotional, behavioral and cognitive impairments with some of the symptoms including feeling of hopelessness, ir-

J. Saxena et al. (eds.), *Proceedings of the International Conference on Sustainable Micro-Nano Materials & Innovative Technology (ICSUMMIT 2026)*, Atlantis Highlights in Engineering 44,
https://doi.org/10.2991/978-94-6239-727-9_19

ritability, anxiety and sleeping and eating disorder. It may also have long-term consequences of mother-infant interaction, child development and family systems when left untreated, [1][2][3].

The specific etiologies of PPD are complex and comprise the hormone changes, specifically the sudden decrease of estrogen and progesterone levels after childbirth, and the psychosocial conditions, including the low socioeconomic status and the previous mental illness, and the absence of support. A new literature also points to the implication of biomarkers, including neurostaroid changes, dysregulation of the immune system, and genetic susceptibility to PPD as causative factors to avoid adverse effects. The traditional screening scales like the Edinburgh Postnatal depression scale (EPDS) have been known to focus on the secondary prevention by controlling the symptoms.

Predictive modeling based on machine learning (ML) and artificial intelligence (AI) has created the means of identifying people at risk in the context of primary prevention.[1][2][3]. Postpartum depression (PPD) is a severe mental condition, which is common in new mothers and is characterized by persistent moods of sadness and anxiety and a dysfunctional capacity to bond with the child. The rate is about 10-15 percent of women around the world yet 18-25 percent in the developing nations. Not only does PPD affect the maternal health negatively, it also affects child growth which may have long lasting effects in the long run[4][5][6].

Recently, machine learning (ML) algorithms have been created as potentially effective strategies to identify and forecast PPD. These models are highly suitable at modeling complex nonlinear relationships and can work with many forms of data, including text, audio and demographic data. One generalization is that, state of art architectures (hybrid CNN-LM models) use both the temporal and spatial properties of the data to provide a superior prediction. Similarly, ensemble ML models like Random Forest, SVM and Neural Networks are applied to pool the predictions together and obtained result is more robust and reliable[4][5][6].

Mental health diagnostics can be used with ML to early diagnose and intervene in time without compromising the- mental health of mothers and children, as it happened in this scenario. This research direction indicates a promising area of developing data-driven and scalable interventions to be effective in treating PPD[4].

2 Review Methodology

This review applies the structured narrative approach to offer a dedicated and critical evaluation of accessible literature on the prediction of postpartum depression (PPD) through machine learning (ML) and deep learning (DL). Recent methodological advances that are published within the years 2020-2025 are pointed out in the review. There was a concentrated review of literature in the recognized academic databases of Scopus, IEEE Xplore, PubMed, ScienceDirect and MDPI. The keywords used in the search included postpartum depression, postnatal depression, machine learning, deep learning, PPD prediction, and artificial intelligence in maternal

health. Keeping this in mind, 11 best representative publications (journal articles/ conference papers) were reviewed in detail to objectives relevance. Review: (i) ML/DL model, using questionnaire, clinical, behavioral or multimodality data to predict, classify, or estimate the risk of PPD; (ii) results of reported performance of experiments, such as accuracy, precision, recall, F1 score and/or area under curve (AUC); (iii) not reported.

The following information from each included study was extracted and is summarized in table 2 method, dataset characteristics, evaluation measures, findings, and limitations, reported by the authors. A cross comparison analysis was then carried out among these to find similar modelling methods, performance trends, datasets dependant results and open issues. This approach for reviewing is conducive for a balanced discussion of the effectiveness and interpretability of the methodologies and the feasibility of the application in a clinical setting, rather than suggesting an any means of exhaustive coverage of the current literature.

3 Review Process Overview

The adopted review process was based on a flow and is conceptually described to ensure transparency and clarity when reporting study inclusion and exclusion.

3.1 Database Search

A targeted literature search of the scholarly databases Scopus, IEEE Xplore, PubMed, ScienceDirect, and MDPI was performed.

Keywords associated with postpartum depression and machine learning such as "PPD prediction", "machine learning", "deep learning", "artificial intelligence in maternal health" etc.

3.2 Selection Criteria

Studies were eligible if they were published from 2020 to 2025 and utilizing ML or DL for the prediction or diagnosis of postpartum depression. We excluded any articles not related to postpartum or maternal depression or that did not provide empirical evaluation.

3.3 Selecting 11 Studies

A total of 11 relevant studies were included for full review after taking into account the inclusion criteria. In this studies we consider peer-reviewed journal articles and conference papers covering popular ML and DL methods.

3.4 Methods, Datasets, and Results Extraction

From each study, information was extracted in a structured way on algorithms, datasets, type of features, evaluation metrics, and achieved performances.

3.5 Comparative Analysis

A comparison between the results of the retrieved data was made in order to recognize methodological tendencies, strengths, weaknesses, and voids in the research. "Predictive performance, dependency on the dataset, degree of interpretability, as well as the potential for clinical deployment were rigorously considered."

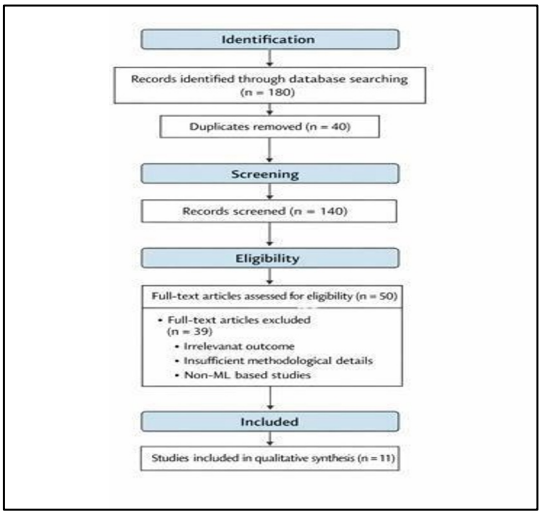

Fig. 1. PRISMA flow diagram illustrating the study selection process of the systematic review

4 Literature Review

Table 1. Literature Review

Reference	Methodologies Used	Strength	Limitation	Year
[1]	MDKR (meta-learning with DT, KNN, RF+ MLP)	Ensemble learning boosts performance; validated-with statistical tests	High complexity; resource-heavy	2024
[2]	RF, SVM, KNN, DT	Three-wave follow-up; strong temporal analysis	High attrition rate; limited generalizability	2024
[3]	Gradient boosting models (CatBoost, LightGBM); emotional and behavioral survey data	Effective handling of categorical features; identification of key risk indicators	Self-reported data; limited generalization; no severity-level classification	2024

[4]	Hybrid CNN-based transfer learning + Improved Bi-LSTM with attention; text and audio data	Multimodal analysis; high accuracy, precision, recall; improved risk factor identification	High computational cost; privacy concerns; model complexity	2024
[5]	Chi-square feature selection KNN, SVM, RF, LR, DT, XGB, ANN, stacking	Multi-layered ensemble improves robustness	Dataset limited to 1503 samples; potential bias from questionnaire data	2024
[6]	LSTM–CNN Deep learning model; comparison with Logistic Regression; sentiment analysis	Captures temporal and spatial patterns; early symptom detection; better Accuracy than traditional ML	Small dataset; reliance on social media data; limited medical interpretability	2024
[7]	SVM, ANN, LR	High accuracy; Robust model validation	Questionnaire-only data; limited feature diversity	2021
[8]	ML (RF, DT, XGB) Survival Analysis (Cox, Kaplan-Meier)	High accuracy, interpretability, combined approaches.	Small dataset, single population, limited biological factors, possible overfitting	2024
[9]	GCNs, RF, SHAP analysis	High interpretability; symptom-level insights	Dataset limited to 1503 samples; potential bias from questionnaire data	2025
[10]	RF, AdaBoost, DT, LR	Large dataset (2570); strong feature analysis	Limited social variables; self-reported data may introduce bias	2025
[11]	Decision Tree (depth-constrained)	Simple, interpretable model usable in low-resource settings	Based on self-reported EPDS scores; lacks clinical diagnosis	2025

5 Different Methods used in PPD

5.1 Logistic Regression

Logistic regression was often employed to calculate the odds of PPD based on demographic, psychosocial, and clinical predictors to determine risk factors of PPD. In cross-sectional and cohort research, logistic regression models have been used suc-

cessfully to estimate the strength of the associations between exposures such as prenatal depression and social support and PPD outcomes. Recent predictive modelling efforts demonstrated that logistic regression achieves competitive performance with more complex approaches for early prediction of PPD risk. Logistic regression has also demonstrated good classification performance for PPD when combined with social determinants of health [12].

5.2 Decision Tree

A decision tree (DT) is a non-parametric supervised learning method that can be used to do both regression and classification. The objective is to derive decision rules that are simple and which are determined based on the data characteristics in order to predict a target variable value. A model is approximated to foretell the worth of a target variable. One can imagine a Tree as something that is an approximation of a piecewise constant. Easily understandable and interpretable. One can imagine trees. Data is not preprocessed much. Other techniques are usually data normalized with the creation of dummy variables and the elimination of missing values. There are tree and algorithm combinations that deal with missing values. The larger the number of data points to define the tree, the larger the cost of the tree evaluation. can deal with categorical and numerical data. Nevertheless, currently the scikit-learn version does not accept categorical variables. Other techniques are often modified to the analysis of data sets that comprise of only single type of variable. To deal with multiple-output problems. uses the white box model. see algorithms. Something true in a model can be described by the following: the condition can be easily described using the logic of booleans. In comparison, the results of a black-box model may be more difficult to interpret. Statistical tests can be used to test a model. By doing so, one can take into account the model reliability, which works well regardless of the fact that the real model that the data were obtained actually do not correspond to the given assumptions to a small extent[13][14].

5.3 Random Forest

A popular supervised machine learning algorithm to address regression and classification tasks is the random forests. Random forests are applied in supervised machine learning, where there is a goal variable that is labelled. Regression (numerical target variable) and classification (categorical target variable) issues can be solved using random forests. Random forests being an ensemble method, combine predictions of different models. Each of the smaller models is a decision tree in the random forest ensemble. A random forest classification builds several decision trees based on multiple random sampling of the data and features. All decision trees will provide different classification of the data. A forecast on each of the decision trees is made and the most popular decision is then chosen to make predictions[15].

5.4 Support Vector Machine

Support Vector Machine (SVM) is one of the most popular supervised learning programs in both classification and regression problems. However, its major application is in blockchain machine learning: classification problems. In order to make it easy to classify new data points, the SVM is aimed at identifying an optimal line or a decision boundary that has the capacity to separate space n-dimensionally. We describe this ideal border of decision as hyper plane. SVM selects the outliers and vectors to construct the hyperplane. The algorithm is called a Support Vector Machine because such extreme cases are called support vectors[16].

5.5 KNN

KNN is a simple concept of supervised learning and it can be used to predict by utilizing the distance between two data points. All these properties of KNN include it in the best baseline machine learning algorithm. It has been effectively used in both classification and regression tasks and its simplicity and efficiency is a benefit that gives it a preference. It is followed by identifying the k nearest neighbors of the input data point with the help of distances (In the KNN algorithm, k is a parameter that shows how many neighbors we want to explore). To classify, the method predicts the label of the majority of the K neighbours to make a prediction of the data point. The KNN algorithm can be used to solve both regression and classification problems[17].

5.6 XGBoost

XGBoost is a faster and performance based implementation of the gradient boosted decision trees which falls under the gradient boosting framework, a type of ensemble learning. Approaches based on regularization of model generalization, decision tree being a base learner. XGBoost is most notorious in its efficiency in com-putation (e.g. missing values are supported, split finding is split sparsity-conscious, feature importance, parallelizing), and it is a machine learning algorithm, and therefore lies within the category of ensemble learning. It is popular in the case of supervised learning tasks such as classification and regression. XGBoost uses a sequence of separate predictive models, often decision trees, combined together, to form a predictive model. The algorithm works in the way that weak learners are joined in the ensemble in sequential order where each learner focuses on correcting the errors made by the others. XGBoost is an ensemble learning method. The results of a single machine learning model might not necessarily be sufficient. The ensemble learning is a methodological way of combining the prediction capability of multiple learners. The final product is one model which is an amalgamation of numerous models. The ensemble may consist of the models based on the same learning algorithm or different learning algorithms, which are sometimes called basic learners. Two such widely used ensemble learning methods are boosting and bagging[18].

5.7 Graph Convolutional Networks

Graph Convolutional Networks (GCNs) have shown potential for predicting postpartum depression by treating symptoms or patient similarities as graph-structured data. Contrary to conventional machine learning models, GCNs utilize the relational dependencies of the nodes, thus effectively aggregating information of the neighbours to enhance the performance of the prediction. Recent research has shown that symptom-based GCN models are able to capture complex inter-feature relationships that are pertinent to postpartum depression risk assessment [19]. Similar good performance on more general depression-detection task also validate the usefulness of GCNs for maternal mental-health study [20]

5.8 SHAP

Shapley Additive Explanations (SHAP) has been more frequently applied in research on postpartum depression (PPD) to increase the interpretability of ML models. SHAP has the advantage of global and local interpretability in that it explains the impact of each feature on a prediction using game theory. For PPD prediction, SHAP has been also used with ensemble methods such as XGBoost to detect the most influential risk factors, such as prenatal depression, anxiety, quality of sleep, and social support. Greater transparency and clinical applicability of SHAP fosters trust and enables informed decision-making in applications for Maternal Mental Health [21].

Table 2. Comparison method used in PPD

Algorithm	Type	Key Strengths	Limitations	Use Case in PPD
Logistic Regression (LR)[12]	Linear, Statistical	Simple, interpretable, good for baseline models	Limited with non-linear or complex data	Early risk assessment using demographic & clinical data
Decision Tree (DT)[13][14]	Non-linear, Supervised	Easy to interpret, visual representation	Prone to Overfitting with small datasets	Screening with simple clinical & psychosocial factors
Random Forest (RF)[15]	Ensemble, Supervised	High accuracy, Robust to noise and overfitting	Less interpretable, Requires larger data	Comprehensive risk prediction using multimodal data
Support Vector Machine (SVM)[16]	Supervised, Classifier	Effective for high-dimensional data, good with small datasets	Computationally expensive, less interpretable	Predicting PPD using clinical + behavioral features
K-Nearest	Instance based	Simple to implement, non-parametric	Sensitive to	Suitable for small

Neighbors (KNN)[17]	Learning		noisy/imbalanced data, slow with large datasets	datasets with clear clusters
Gradient Boosting Machines (GBM/XGBoost/LightGBM)[18]	Ensemble, Boosting	High predictive performance, handles complex patterns	Prone to overfitting if not tuned properly	Large datasets with mixed data types (clinical + social)

6 Discussion

The results of this study, along with those of recent articles, indicate that machine learning algorithms can provide higher feasibilities for early identification of postpartum depression (PPD). Several of the reviewed articles indicate that ensemble-based models perform better than traditional classifiers in prediction of PPD as they are able to capture non-linear relationships among demographic, psychological and behavioral risk factors [1],[2],[5].

The results of current experiments are consistent with previous work that tree-based ensemble methods are effective. Random forest and gradient boosting algorithms have been demonstrated to perform well with various combinations of features and datasets [2],[3],[5]. In particular, reports employing CatBoost and stacking-based ensembles demonstrated superior classification performance by considering implicit interactions between features as well as categorical features [3], [5]. These results are in line with the high accuracy obtained in the present study using ensemble learning along with feature selection.

It was claimed that feature selection played an important role in enhancing the accuracy of the model in PPD prediction. Several contributions point out that the dataset based on questionnaires are heavily redundant, or correlated features could harm the performance of the classifier [5], [7]. Present application of RFE is in line with the previous results suggesting that by minimizing feature space while maintaining discriminativeness, one can achieve higher accuracy and robustness.

Deep learning and hybrid models have demonstrated capabilities of learning temporal and multimodal features for postpartum depression. Attention and CNN–LSTM based approaches achieved better predictive results on text, audio, or time-series data [4], [6]. Nevertheless, these models usually need to be trained with large size of data and they consume more computational resources and are opaque, which restrict their use in clinical decision-making [4], [6].

Interpretability is still a significant challenge in the application of ML models for maternal mental health. Although complex ensemble and deep learning models perform well in terms of predictive accuracy, clinicians need systems that are transparent and explainable to have trust in automated screening tools. Logistic Regression and depth-bounded decision trees demonstrate that interpretable models

can have predictive performance comparable with non-interpretable models when evaluated on the clinical usability context [10], [11]. Our findings reinforce the focus of the current study: the trade-off between predictive performance and interpretability.

Another major issue that emerges from multiple researches is the dependency of the dataset. A large part of the literature that was reviewed is based on publicly available questionnaire-based dataset, most notably the Kaggle dataset with 1,503 records [1], [5], [7]. While such datasets make benchmarking and reproducibility easier, there is risk that repeated use may also cause datasets bias and hinder generalizability. This highlights the need to test predictive algorithms developed in such large-scale cohort studies [10], [11] on more extensive, heterogeneous, and clinically ascertained samples.

The recent developments in graph-based learning also imply new possibilities for PPD prediction. Graph Convolutional Networks have been applied to symptom-level interactions to capture relational dependencies that traditional classifiers fail to represent model [9].

Although these methods are promising, issues concerning interpretability and availability of data need to be solved before they can be integrated in clinical practice. Taken together, the discussion synthesizes ensemble learning strategies (when enabled by suitable feature selection) as the current best candidate for a predictive solution for postpartum depression. However, to facilitate the development of scalable and ethically sound solutions for machine learning based PPD screening systems, future work must focus on dataset heterogeneity xAI and clinical validation in real-world settings.

7 Conclusion and Future Work

Machine learning models, in particular XGBoost and Random Forest, have shown good results in detecting postpartum depression. Advanced detection and treatment of maternal PPD can be significantly improved by using AI based models in mother healthcare, which in turn can reduce maternal and infant burden. To render these models suitable for therapeutic use, research efforts in the future need to be focused on scalability, real-world implementation, and integration of multimodal data. Using a Kaggle dataset, the paper investigated the application of machine learning models to postpartum depression (PPD) prediction. It is the pioneering research that will prove machine learning as a promising tool to detect PPD in its initial stages, which might enable medical personnel to take proactive steps to prevent it. The excellent accuracy values achieved by Random Forest and XGBoost suggest that these models can be adopted in clinical practice to support the decision-making of mental health professionals. A dataset of 1,503 records was used in this work. For better generalization of the model, future work shall involve larger and diverse datasets, e.g., real clinical data. Feature selection may be further extended to include biomarkers, genetic data, and social factors to increase prediction accuracy and model robustness.

Disclosure of Interest: The authors have no competing interests to declare that are relevant to the content of this article.

References:

1. Nasim, Shazia, et al. "Novel meta learning approach for detecting postpartum depression disorder using questionnaire data." IEEE Access 12 (2024): 101247-101259.
2. Wang, Yu, et al. "Trajectory on postpartum depression of Chinese women and the risk prediction models: A machine-learning based three-wave follow-up research." *Journal of Affective Disorders* 365 (2024): 185-192.
3. Gupta, Vinayak, et al. "Predictive Algorithms for Early Postpartum Depression Detection: CatBoost vs. LightGBM." *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*. IEEE, 2024.
4. Lilhore, Umesh Kumar, et al. "Prevalence and risk factors analysis of postpartum depression at early stage using hybrid deep learning model." *Scientific reports* 14.1 (2024): 4533.
5. Sharma, Yash, Vansh Jain, and Sandhya Tarwani. "Ensemble machine learning model for predicting postpartum depression disorder." 2024 IEEE Region 10 Symposium (TENSYMP). IEEE, 2024.
6. Srivatsav, P., and S. Nanthini. "Detecting early markers of post-partum depression in new mothers: an efficient LSTM-CNN approach compared to logistic regression." *2024 5th International Conference on Innovative Trends in Information Technology (ICITIIT)*. IEEE, 2024.
7. Zulfiker, Md Sabab, et al. "An in-depth analysis of machine learning approaches to predict depression." *Current research in behavioral sciences* 2 (2021): 100044.
8. Juliet, Sujitha, and J. Anitha. "Machine Learning and Survival Analysis Models for Postpartum Depression: A Comprehensive Risk Factor Analysis." *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*. Vol. 1. IEEE, 2024.
9. Mathew, Jessica Sarah, and Gobi Ramasamy. "Graph Convolutional Networks for Predicting Postpartum Depression: A Symptom-Based Analysis." *2025 International Conference on Emerging Technologies in Computing and Communication (ETCC)*. IEEE, 2025.
10. Kim, Hyunkyoung. "Predictive analysis of postpartum depression using machine learning." Healthcare. Vol. 13. No. 8. 2025.
11. Matsumura, Kenta, et al. "Decision tree learning for predicting chronic postpartum depression in the Japan Environment and Children's Study." Journal of Affective Disorders 369 (2025): 643-652.
12. Zhu, Xiaotong, et al. "Predicting Postpartum Depression Risk Using Social Determinants of Health." Studies in Health Technology and Informatics 329 (2025): 851-855.
13. R. Zhang et al., "An interpretable machine learning model for predicting postpartum depression," JMIR Medical Informatics, vol. 13, e58649, 2025.
14. Breiman, Leo, et al. Classification and regression trees. Chapman and Hall/CRC, 2017.
15. L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
16. R. Zhang et al., "An interpretable machine learning model for predicting postpartum depression," JMIR Medical Informatics, vol. 13, e58649, 2025.
17. T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 21–27, Jan. 1967.
18. X. Huang et al., "Postpartum depression risk prediction using explainable machine learning algorithms," Frontiers in Medicine, vol. 12, Art. no. 1565374, 2025.

19. M. Ramasamy, R. K. Patel, and S. Verma, "Graph convolutional networks for predicting postpartum depression: A symptom-based analysis," in Proc. Int. Conf. Artificial Intelligence and Healthcare, 2024, pp. 112–118.
20. J. Song, W. Lin, and H. Zhang, "Graph neural networks for depression detection using symptom and patient similarity graphs," IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 5, pp. 2143–2153, May 2023.
21. R. Zhang et al., "An interpretable machine learning model for predicting postpartum depression," JMIR Medical Informatics, vol. 13, e58649, 2025.