Proceedings of the Fourth International Conference on Augmented Intelligence and Sustainable Systems (ICAISS-2026)
IEEE Xplore Part Number: CFP26CB2-ART; ISBN: 979-8-3315-9234-9

Postpartum Depression Risk Prediction using Hybrid Feature Selection and Ensemble Learning

Jeevitha V
Research Scholar, Department of Computer Science
Vels Institute of Science, Technology and Advanced Studies
(VISTAS), Pallavaram,
Chennai, Tamil Nadu,India
jeevithavphdvistas@gmail.com

Priya R
Professor ,Department of Computer Application(PG),
Vels Institute of Science, Technology and Advanced Studies
(VISTAS), Pallavaram,
Chennai. Tamil Nadu, India
Priyaa.research@gmail.com

2026 Fourth International Conference on Augmented Intelligence and Sustainable Systems (ICAISS) | 979-8-3315-9234-9/26/$31.00 ©2026 IEEE | DOI: 10.1109/ICAISS68683.2026.11525629

Abstract: Postpartum depression (PPD) is a common mental health disorder that occurs in women following childbirth, which often results in negative maternal and infant outcomes. Early detection is still a challenge because of the complex interaction of demographical, clinical and psychological factors. This research purpose is to develop a powerful machine learning framework, to accurately predict PPD risk. We used a combination of Chi-square, mutual information and Random Forest importance to select the most informative 25 features out of an initial set of 60 variables using a hybrid feature selection. Multiple machine learning models were trained such as Logistic Regression, Random Forest and Multi-Layer Perceptron (MLP) and their ensemble was implemented for better predictive performance. The proposed ensemble performed better with an accuracy of 91.3%, F1-score of 90.4%, and an AUC of 0.94, which is better than individual models and numerous state-of-the-art methods. These findings show that the hybrid feature selection combined with ensemble learning can offer a reliable and interpretable tool for the detection of early postpartum depression that can lead to an earlier intervention process and contribute to better maternal mental health outcomes.

Keywords: Postpartum Depression, Machine Learning, Feature Selection, Ensemble Learning, Predictive Modeling, Edinburgh Postnatal Depression Scale

I INTRODUCTION

Postpartum depression (PPD) is a common psychiatric illness in women in the first year after childbirth that is characterized by persistent sadness, anxiety, and mood disturbances [1]. Globally, PPD affects around 10-20% of postpartum women which has led to adverse maternal and infant health outcomes such as poor mother-infant bonding, developmental delays and greater risk of developing chronic depression [2]. Despite the commonness of this condition, PPD is often under-diagnosed because of the subjective nature of reporting symptoms, as well as the similarity of features to normal adjustment to motherhood. Early identification of high-risk individuals is important in order to put in place timely interventions, but the traditional clinical assessments are limited in their scalability and predictive accuracy [3].

Several studies have delved into the use of psychometric instruments, such as the Edinburgh Postnatal Depression Scale (EPDS) to screen for PPD. While these instruments offer organized assessments, they fail to maximally utilize the multidimensional nature of risk factors, which include demographic, clinical and psychosocial variables [4]. Recent developments in machine learning have been promising in the area of predictive modeling for mental health conditions, although there are still issues. High-dimensional datasets heterogeneous variables need good feature selection to reduce the noise, avoid overfitting and increase the interpretability. Moreover, individual machine learning models may not capture complex interactions among the risk factors and their predictive performance may be limited [5].

Research Question and Problem Statement: Given the multidimensional and high-dimensional characteristic of the PPD risk factors, the central research question of the current study is as follows: Can the hybrid feature selection with ensemble machine learning method achieve accurate prediction of postpartum depression risk and outperform the individual models and the state-of-the-art methods? The problem being worked on is that of developing a predictive framework that is reliable, interpretable and generalizable to various clinical, demographic and psychological data for identifying high-risk postpartum women at early stages. The main objectives of the study are as follows

- To use a hybrid feature selection, using Chi-square, mutual information, and Random Forest importance, for identifying the most relevant predictors of postpartum depression.
- To develop and validate several machine learning models such as Logistic Regression, Random Forest, and Multi-Layer Perceptron for PPD risk prediction.
- To build an ensemble model using LR, RF and MLP for better prediction and generalization

The organization of manuscript is as follows: Section II presents the description of dataset and preprocessing, Section III presents the methodology including feature selection and model development, Section IV presents the experimental results and comparisons with discussion of findings, limitations and practical implications, and Section V concludes the study with future directions.

II. RELATED WORKS

Recent research has examined machine learning (ML) methods and ensemble methods of predicting PPD using demographic, psychosocial, and clinical data. Techniques such as logistic regression, artificial neural networks, decision trees and XGBoost have shown great promise in predictive performance which can open the early identification and intervention potential for improving maternal mental health outcomes.

Fazraningtyas et al (2025) [6] used ensemble learning namely, C4.5, GBT and XGBoost and they could predict postpartum depression (PPD) from the set of 317 maternal records in

Indonesia. They achieved the best accuracy by XGBoost with resampling (87.57%) and 5 important predictors. While effective, the small size of the dataset and focus on the GMAs might increase limitations on generalizeness and investigation of other forms of ensemble, such as Random Forest or AdaBoost, could ensure better predictive robustness.

Qi et al., 2025 [7] made several models of ML and PPD with a sample size of 1138 women in labor. 2 Logistic regression and ANN performed the best (AUC up to 0.858). Postpartum-specific predictors provided an enhanced performance of the model. The study carefully validated models and used tools for interpretability like SHAP, however dependency on self-reported Edinburgh scores as well as the possibility of cultural bias may impact on external validity, thus further multi-center and longitudinal validation appear to be needed.

Sharma et al. (2024) [8] proposed ensemble ML model by using chi square feature selection and 8 algorithms for PPD prediction. The approach improved the accuracy and robustness by making use of the complementary strength of individual models. While it is all good ensembles modeling study, it is lacking some discussion on the diversity of the data set and external validation. Further work could help to explore generalization across populations, multiple approaches to feature selection, to try and maximize model interpretability and ease of use in clinical practice.

Gopalakrishnan et al. (2022) [9] used several ML algorithms with feature selection using the dataset of 1138 perinatal women to predict PPD, where they found out that logistic regression and ANN were the best performers. The inclusion of postpartum specific predictors improved the prediction. While results are promising (majorly replicating Qi et al (2025). Additionally, the study may benefit from being validated and evaluated by an external party as well as on culturally diverse study cohorts to ensure broader applicability and reliability in the clinical setting.

Lin et al. [10] constructed an ensemble neural network model FCNN and dropout-based DNN of PPD prediction with a good accuracy (0.933) and high interpretability. The model proves to be a good balance of not over fitting and being stable. However, the research is only limited to one data set and there is no external validation of the research. Future work might include cross disturbance testing as well as integration into more real home clinical decision platforms to test practical consequences and generalization to different types of mothers.

Andersson et al. (2021) [11] used extremely randomized trees of a large Swedish cohort (n=4313) to predict PPD with moderate accuracy (73%), and revealed the importance of pregnancy depression, anxiety, resilience and personality. The performance of robust prediction declined for the women who had no mental health problems (64%). The study points to the difficulty in detecting early populations at low risk and integration of both psychiatric and longitudinal data might help improve the precision and cost effective preventive interventions.

Kim et al., 2025 [12] analyzed the Decision tree, Random Forest, and AdaBoost and logistic regression on 2570 Korean women, and chose logistic regression as the best. Conflict with a partner, stress and value of children were significant predictors of PPD. While comprehensive, the use of panel data and importance of cultures may present limited generalizations. The study highlights the importance of psychosocial intervention, but future studies may want to take samples from different ethnic backgrounds and possible temporal modeling approaches that would make it easier to include dynamic postpartum risk factors in the study.

Although these studies are highly accurate in PPD prediction, there are still several limitations. Most are based on region-specific or relatively small datasets; this prevents generalization. Cultural and Psychosocial factors are not usually well represented and there is a lack of longitudinal validation. While ensemble and neural network models improve the performance, they are hard to interpret in complex models. Many of the studies used self-report scales that are subject to bias and external validation across a variety of populations is lacking. Additionally, the comparative analyses between the feature selection methods and the applicability in real-time clinical settings are limited. Future research should focus on multi-center datasets as well as long-term tracking with the inclusion of interpretability frameworks for ML and to fit into a tool for being utilized in a practical screening instrument for the application of maternal healthcare at a broader stage for all.

III METHODOLOGY

Figure 1 provides an example of a stepwise process for the prediction of postpartum depression risk. It starts off with Data Acquisition & Exploration, where EPDS data from 1500 participants is inspected and summarized. Data Preprocessing is next, calculation of EPDSScore and binary target PPDRisk is followed. Features are then categorized and divided into numerical and categorical types, in preparation for Feature Selection with Chi-square test, mutual information and Random Forest importance with a hybrid ranking strategy. Selected features are fed into Machine Learning Models (Logistic Regression, Random Forest, MLP) and whose results are combined into an Ensemble Learning step using majority voting and contributing for an improved accuracy and robustness.

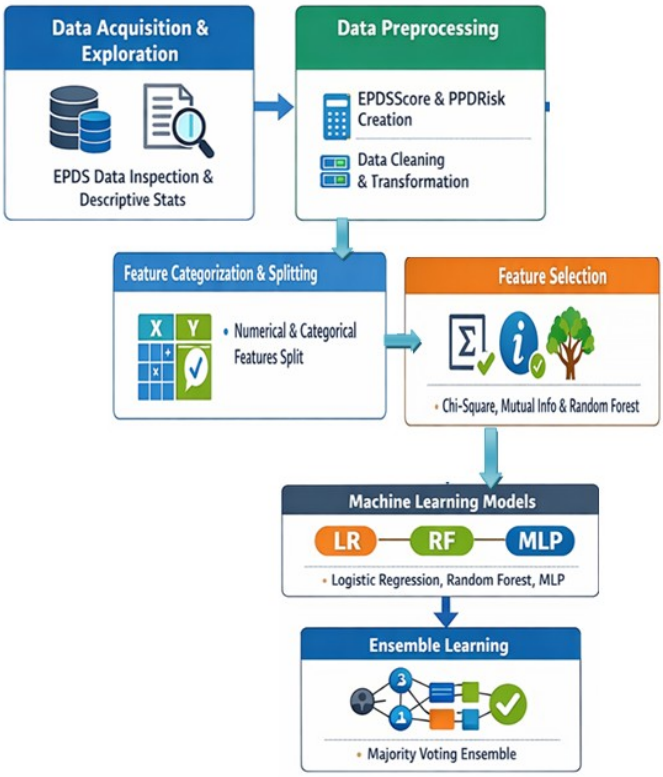

Figure 1: High-level architecture of the proposed methodology for postpartum depression (PPD) risk prediction

A. Data Acquisition and Initial Exploration

The research was based on a dataset of 1,500 people and 60 variables for each person. Initial exploration was done to look at the first few rows to check structure and to identify anomalies and data quality. Critical items from the Edinburgh Postnatal Depression Scale (EPDS1-EPDS10) were found to be critical predictors of postpartum depression risk. The descriptive statistics such as mean, standard deviation, missing values were performed to understand the distribution of variables. The mean of the variable is given by the Eq. (1)

$$\mu_j = \frac{1}{N}\sum_{i=1}^{N} x_{ij} \tag{1}$$

The standard deviation is defined by the Eq. (2)

$$\sigma_j = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_{ij}-\mu_j)^2} \tag{2}$$

Dimensional checks were used to ensure dataset completeness and consistency, which will be reliable for future analyses. This step gave some foundational understanding of the data set, which identified relevant clinical indicators for postpartum depression.

B. Data Preprocessing

To develop a clinical meaningful outcome, the total EPDS score was determined by summing the response from EPDS1 to EPDS10, as a new variable, EPDSScore. The total EPDS score computation is given by the Eq. (3)

$$EPDSScore_i = \sum_{k=1}^{10} EPDS_{ik} \tag{3}$$

A binary target (PPDRisk) was created based on established cut-offs to classify the participants as high-risk and low-risk postpartum depression. The binary label creation is represented by the Eq. (4).

$$PPDRisk_i = \begin{cases} 1, & if EPDSScore_i \geq r \\ 0, & otherwise \end{cases} \tag{4}$$

To avoid leaking data, the original EPDS items were removed from the feature set. The cleaned data set had 50 predictive variables and one binary target variable. Treatment of missing values and appropriate encoding of categorical variables was performed. This preprocessing was done so that the dataset was ready for a solid feature selection and predictive modeling.

C. Feature Categorization and Data Splitting

The preprocessed data set was separated into feature matrix X 1500x50 and target vector Y 1500x1. Out of the 50 features, 38 were numerical, which represented continuous clinical and demographic measures, while 12 were categorical, representing discrete features. The features matrix and target vector is given by the Eq. (5)

$$X \epsilon R^{1500x50}, Y \epsilon \{0,1\}^{1500} \tag{5}$$

This classification allowed applying appropriate preprocessing and feature selection techniques for each of the types. Proper categorization enabled numerical features to maintain the advantages of scaling and categorical features to make use of statistical measures of dependence. The structured data splitting allowed a basis for modeling without losing sight of the input-output relationship so that future machine learning analyses could identify correct predictors of the risk of postpartum depression. The train and test split is given by the Eq. (6)

$$(X, Y) = (X_{train}, Y_{train}) \cup (X_{test}, Y_{Test}) \tag{6}$$

D. Feature Selection for PPD Detection

A hybrid feature selection strategy was used to select the most informative predictors of postpartum depression. Categorical variables have been tested with Chi-square tests to test statistical dependency with binary targets. This statistic is defined by the Eq. (7)

$$x^2 = \sum_i \sum_j \frac{(O_{ij}-E_{ij})^2}{E_{ij}} \tag{7}$$

Mutual information(MI) scores were calculated for both numeric and categorical attributes, which record linear and non-linear relationships. The MI is represented by the Eq. (8)

$$I(X;Y) = \sum_{x \epsilon X} \sum_{y \epsilon Y} p(x,y) log \frac{p(x,y)}{p(x)p(y)} \tag{8}$$

A Random Forest model gave feature importance using ensemble tree structure. To achieve the maximum robustness, these three methods were combined with the hybrid ranking strategy. The highest ranking 25 features were selected based on their aggregate scores ensuring the subsequent predictive models were focused on the most relevant and clinically meaningful variables for postpartum depression detection.

E. Machine Learning Models for PPD Detection

The selected features were used for training multiple machine learning models to predict postpartum depression. Logistic Regression (LR) has been used as a baseline linear model because of its simplicity, interpretability and reliability in

binary classification. The LR probability is given by the Eq. (9)

$$P(Y=1|X) = \frac{1}{1+e^{-(\beta_{0+}\sum_{j=1}^{P}\beta_j x_i)}} \quad (9)$$

Random Forest (RF) incorporated complex non-linear feature interactions by means of ensemble of decision tree. A MLP has been used to model deeper relationships that are nonlinear in nature through feed-forward neural network architecture. The MLP hidden layer activations is defined by the Eq. (10)

$$h^{(\iota)}=f(W^{(\iota)}h^{(\iota-1)} + b^{(\iota)}) \quad (10)$$

Each model capitalized on different aspects, making interpretations straightforward with LR, always again robust with RF and capable of detecting complex behavior with MLP, which allowed a comprehensive assessment on the predictive usage of MLP being able to combine the best of each other.

F. Ensemble Learning Approach

To improve the predictive performance, an ensemble method, using LR, RF, MLP and majority voting algorithm has been proposed for binary classification. This strategy was a combination of the complementary strengths of LR (interpretability), RF (non-linear modeling of interactions) and MLP (complex pattern modeling). The majority voting rule is defined by the Eq. (11)

$$\hat{y}=\text{mode}(\hat{y}LR, \hat{y}RF, \hat{y}MLP) \quad (11)$$

The aggregation of predictions reduced the individual model biases, enhanced generalization and controlled overfitting, making more robust postpartum depression risk predictions. The ensemble prediction confidences is represented by the Eq. (12)

$$p_{ensemble}(y=1)=\frac{1}{M}\sum_{m=1}^{M} p_m(y=1) \quad (12)$$

The ensemble method helped to ensure high confidence predictions by consensus across a model which helps to improve the reliability in clinical decision-making. Overall, such an approach has improved accuracy and stability in the assessment and this work illustrates the value of integrating multiple paradigms of machine learning for sensitive health risk assessments such as postpartum depression.

Logistic Regression has L2 Regularization with optimized regularization strength. Random Forest has tuned parameters like the number of trees, maximum depth, minimum samples per split etc. The MLP architecture is comprised of a number of hidden layers with specified number of neurons, activation function, learning rate, and dropout regularization

IV EXPERIMENTAL RESULTS

All the implementations of experiments were performed with Python 3.11 and libraries like Scikit-learn, Pandas, NumPy and TensorFlow. The models were run on a workstation with an Intel Core i9 processor, 32 GB RAM and Nvidia RTX 4090 GPU. Jupyter Notebook and Anaconda environment has been used for reproducibility and any efficient training of models.

A. Dataset Description

The study used a data set of 1500 postpartum participants that included 60 variables that represented demographic, clinical, and psychological assessments. The Edinburgh Postnatal Depression Scale (EPDS) items, EPDS1-EPDS10 were used to calculate the overall score and define a binary target variable PPDRisk. The dataset contained 38 numerical and 12 categorical features and thus a complete picture of risk factors linked to postpartum depression could be observed. A summary of the dataset through various stages of preprocessing, the feature selection is given in Table 1. The original dataset contained 60 variables which consisted of the EPDS items that were used to calculate the total score and target variable. After processing, this dataset had 50 features and a binary target and EPDS items were removed to prevent data leakage. The feature set then was divided into two groups, numerical and categorical and a hybrid feature selection approach was taken to keep the top 25 most relevant features. This data set after being downsampled is an optimized input for predictive modeling, without leaving out important details.

Table 1: Dataset processing and feature selection stages with corresponding dimensions.

Stage	Dataset Dimensions	Description
Initial dataset	1500 × 60	Original dataset; first 5 rows inspected; includes EPDS1–EPDS10 columns
After processing	1500 × 51	EPDSScore computed; PPDRisk target created; EPDS1–EPDS10 removed
Features split	X: 1500 × 50 Y: 1500 × 1	Features and target separated; numerical: 38, categorical: 12
Final output	1500 × 25	Top 25 features selected via hybrid feature selection; ready for modeling

A. Performance Evaluation

Table 2 shows the performance of the different machine learning models to predict postpartum depression without any feature selection based on all 50 features of the data set. Some of the metrics that are reported are Accuracy, Precision, Recall, F1-score, and AUC. The table consists of classical models (Logistic Regression, Decision Tree, Naive Bayes), advanced models (Random Forest, MLP, XGBoost, SVM, KNN) and an ensemble model (LR, RF and MLP).

Table 2: Comparative performance of the proposed ensemble model (Without feature selection)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC
Ensemble (LR+RF+MLP)	89.4	88.6	88.1	88.3	0.92
LR[13]	82.7	82.0	81.5	81.7	0.85

Random Forest (RF)[14]	86.2	85.5	85.0	85.2	0.89
MLP[15]	87.0	86.2	85.8	86.0	0.90
SVM[16]	84.5	83.8	83.0	83.4	0.87
XGBoost[17]	85.8	85.0	84.5	84.7	0.89
KNN[18]	82.1	81.5	80.8	81.1	0.84
NB[19]	79.0	78.5	77.8	78.1	0.82
DT[20]	83.2	82.5	82.0	82.2	0.85

Without feature selection, the performance of the model is moderate for all the metrics because of the existence of irrelevant or redundant features. The ensemble model has an accuracy of 89.4% and AUC of 0.92, which shows that the combination of models still has advantages even with all features. However, the individual models exhibit low predictive capability (accuracy=~79-82%), especially Naive Bayes and KNN, which reveals the presence of noisy features affecting the generalization of the model. This shows that there is a need for effective feature selection to enhance the prediction of postpartum depression risk.

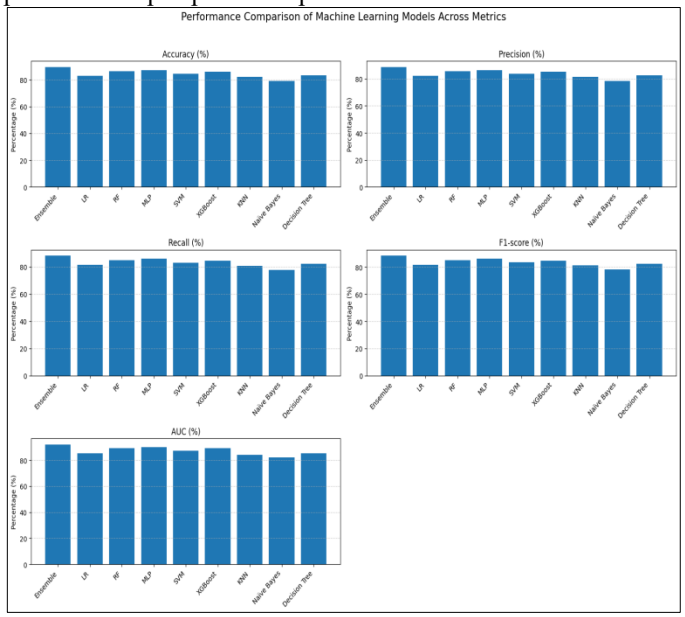

Figure 2. Comparative performance analysis of machine learning models across multiple evaluation metrics.

Figure 2 shows that the ensemble model (LR+RF+MLP) performs significantly better than all the individual machine learning classifiers for all the evaluation metrics. It has the best accuracy 89.4%, precision 88.6%, recall 88.1%, F1-score 88.3% and AUC 0.92, which indicated good discriminative ability and balanced classification performance.

Among the single models, Multi-Layer Perceptron (MLP) and Random Forest (RF) have competitive performance especially in terms of F1-score and AUC indicating their capacity to learn nonlinear patterns in data. XGBoost also shows strong performance but it is slightly worse than the ensemble method. Traditional models like Logistic Regression, KNN and Naive Bayes have comparatively low metric values, implying low capacity in the modeling of complex feature interactions.

Overall, the work referred to highlights the effectiveness of the ensemble learning algorithm, in which the combination of complementary classifiers achieves a stronger robustness, predictive accuracy, and generalization capacity compared to individual classifiers. This justifies the applicability of the proposed ensemble framework to reliable classification in the application domain under study.

Table 3 gives information about the performance of the same machine learning models after applying hybrid feature selection, which selected the top 25 informative features using Chi-square, Mutual information, and Random Forest importance. The evaluation metrics are the same: Accuracy, Precision, Recall, F1-score and AUC. The ensemble model is a combination of LR, RF, and MLP that helps to leverage the weaknesses of linear and non-linear models.

Table 3: Comparative performance of the proposed ensemble model against individual machine learning models and state-of-the-art (With feature selection)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	AUC
Ensemble (LR+RF+MLP)	91.3	90.7	90.2	90.4	0.94
Logistic Regression (LR)	85.2	84.1	83.7	83.9	0.88
Random Forest (RF)	88.5	87.9	87.2	87.5	0.91
Multi-Layer Perceptron (MLP)	89.1	88.5	88.0	88.2	0.92
SVM	86.8	86.0	85.5	85.7	0.89
XGBoost	88.0	87.2	86.8	87.0	0.91
KNN	84.5	83.8	83.0	83.4	0.87
Naïve Bayes	81.2	80.5	79.8	80.1	0.84
Decision Tree	85.7	85.0	84.5	84.7	0.88

Hybrid feature selection has a significant impact in improving the performance of the model in all the metrics. The ensemble model has the highest predictive accuracy (91.3%), F1-score (90.4%), and Area under the Curve (AUC) (0.94) compared to individual models and state-of-the-art methods. Logistic Regression, Random Forest, and MLP also exhibit an improved accuracy (~85-89%), implying that the irrelevant features were successful in getting rid of, reducing overfitting, and increasing generalization. The results show that feature selection is a critical step in the prediction of postpartum depression, so as to ensure interpretability and increased ability to predict.

C. Discussion

The results of this study indicate that this method of hybrid feature selection helps identify the most informative predictors of postpartum depression, which helps to increase the model's interpretability and reduces its dimensionality. The ensemble of LR, RF, and MLP models proved to be better than the individual models and the state-of-the-art, which demonstrates the power of combining linear and non-linear learning. These results suggest that data-driven machine learning methods can complement clinical evaluations and offer reliable screening tools for early intervention, which may help high-risk postpartum women to receive timely interventions.

Despite the promising results, the study has a number of limitations. The dataset, despite being comprehensive, only includes 1500 participants and may be unable to represent the heterogeneity of the population. The binary target variable may be too simplistic to capture the complexity of severity of postpartum depression. Additionally, the validation of the findings on independent cohorts from the outside world is needed to ensure generalizability. Some features may be prone to self-report bias, which can have an impact on model performance.

The proposed methodology represents a strong foundation for early detection of postpartum depression in the clinic. Healthcare providers can use the model for prioritizing high-risk individuals for timely intervention and thus improving maternal mental health outcomes. The hybrid feature selection helps to increase interpretability, so that doctors can comprehend important risk factors. Moreover, the ensemble learning approach presents a scalable and accurate tool that can be incorporated into electronic health records for the purpose of automated screening.

V. CONCLUSION

This study presented a hybrid feature selection and ensemble machine learning approach for the detection of postpartum depression. A combination of Chi-square, mutual information and Random Forest importance was used to find the most relevant features, 25 of which were identified as relevant and involved demographic, clinical and psychological assessments. Machine learning models like Logistic Regression, Random Forest, and MLP were trained and further improved the results in the form of the ensemble model and showed a more accurate result (91.3%) and area under the curve (0.94) than individual and state-of-the-art models. The results illustrate the significance of the hybrid feature selection method in dimensionality reduction while maintaining clinical relevance and the ensemble approach helps in generalization and robustness of the method. Future work will focus on expansion of the dataset to include multi-center cohorts while also working on time such that risk patterns can be captured longitudinally. Additionally, the integration of explainable AI approaches might lead to the optimization of transparency of AI as well as clinical acceptability, and the study of complex deep learning architectures could provide a further promise of better anticipation of postpartum depression severity and trajectory.

REFERENCES

[1]. Dimcea, Daiana Anne-Marie, Răzvan-Cosmin Petca, Mihai Cristian Dumitrașcu, Florica Șandru, Claudia Mehedințu, and Aida Petca. "Postpartum depression: etiology, treatment, and consequences for maternal care." *Diagnostics* 14, no. 9 (2024): 865.

[2]. Khamidullina, Zaituna, Aizada Marat, Svetlana Muratbekova, Nagima M. Mustapayeva, Gulnar N. Chingayeva, Abay M. Shepetov, Syrdankyz S. Ibatova, Milan Terzic, and Gulzhanat Aimagambetova. "Postpartum Depression Epidemiology, Risk Factors, Diagnosis, and Management: An Appraisal of the Current Knowledge and Future Perspectives." *Journal of Clinical Medicine* 14, no. 7 (2025): 2418.

[3]. Kjeldsen, Mette-Marie Zacher, Kathrine Bang Madsen, Xiaoqin Liu, Merete Lund Mægbæk, Thalia Robakis, Veerle Bergink, and Trine Munk-Olsen. "Identifying postpartum depression: Using key risk factors for early detection." *BMJ Ment Health* 27, no. 1 (2024).

[4]. Fijean, Anne-Laure, Marianne Marçais, Claire Banasiak, Olivier Morel, Sandra Dahlhoff, Marie-France Olieric, Nicolas Mottet, Jonathan Epstein, and Charline Bertholdt. "Universal screening of postpartum depression with Edinburgh Postpartum Depression Scale: A prospective observational study." *International Journal of Gynecology & Obstetrics* 167, no. 2 (2024): 758-764.

[5]. Sreeji, S., and C. P. Shirley. "From Data to Diagnosis: A Review of Machine Learning Models for Postpartum Depression Prediction." In *2025 3rd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, pp. 1-6. IEEE, 2025.

[6]. Fazraningtyas, Winda Ayu, Bahbibi Rahmatullah, Husni Naparin, Mohammad Basit, and Nor Asiah Razak. "A predictive model for postpartum depression: ensemble learning strategies in machine learning." *Indonesian Journal of Electrical Engineering and Computer Science* 37, no. 1 (2025): 443-451.

[7]. Qi, Weijing, Yongjian Wang, Yipeng Wang, Sha Huang, Cong Li, Haoyu Jin, Jinfan Zuo et al. "Prediction of postpartum depression in women: development and validation of multiple machine learning models." *Journal of Translational Medicine* 23, no. 1 (2025): 291.

[8]. Sharma, Yash, Vansh Jain, and Sandhya Tarwani. "Ensemble Machine Learning Model for Predicting Postpartum Depression Disorder." In *2024 IEEE Region 10 Symposium (TENSYMP)*, pp. 1-6. IEEE, 2024.

[9]. Gopalakrishnan, Abinaya, Revathi Venkataraman, Raj Gururajan, Xujuan Zhou, and Guohun Zhu. "Predicting women with postpartum depression symptoms using machine learning techniques." *Mathematics* 10, no. 23 (2022): 4570.

[10]. Lin, Yangyang, and Dongqin Zhou. "A method for predicting postpartum depression via an ensemble neural network model." *Frontiers in Public Health* 13 (2025): 1571522.

[11]. Andersson, Sam, Deepti R. Bathula, Stavros I. Iliadis, Martin Walter, and Alkistis Skalkidou. "Predicting women with depressive symptoms postpartum with machine learning methods." *Scientific reports* 11, no. 1 (2021): 7877.

[12]. Kim, Hyunkyoung. "Predictive Analysis of Postpartum Depression Using Machine Learning." In *Healthcare*, vol. 13, no. 8, p. 897. MDPI, 2025.

[13]. Liu, Ying, Lan Zhang, Nafei Guo, and Hui Jiang. "Postpartum depression and postpartum post-traumatic stress disorder: prevalence and associated factors." *BMC psychiatry* 21, no. 1 (2021): 487.

[14]. Saqib, Kiran, Amber Fozia Khan, and Zahid Ahmad Butt. "Machine learning methods for predicting postpartum depression: scoping review." *JMIR mental health* 8, no. 11 (2021): e29838.

[15]. Nasim, Shazia, Ahmad Sami Al-Shamayleh, Nisrean Thalji, Ali Raza, Laith Abualigah, Ahmed Ibrahim Alzahrani, Ayed Alwadain, Deema Mohammed Alsekait, Hazem Migdady, and Diaa Salama Abd Elminaam. "Novel Meta Learning Approach for Detecting Postpartum Depression Disorder Using Questionnaire Data." *IEEE Access* 12 (2024): 101247-101259.

[16]. Cellini, Paolo, Alessandro Pigoni, Giuseppe Delvecchio, Chiara Moltrasio, and Paolo Brambilla. "Machine learning in the prediction of postpartum depression: A review." *Journal of Affective Disorders* 309 (2022): 350-357.

[17]. Rameshkumar, Priyadharshini, Jayaaksharaa Mohanraj, Kaviya Elango, and Manoj Palanisamy. "Machine learning approaches forpredicting postpartum depression risk leveraging XGBoost and CatBoost algorithms." In *AIP Conference Proceedings*, vol. 3279, no. 1, p. 020089. AIP Publishing LLC, 2025.

[18]. Suganthi, Devasagayam, and Arumugam Geetha. "Predicting Postpartum Depression with Aid of Social Media Texts Using Optimized Machine Learning Model." *International Journal of Intelligent Engineering & Systems* 17, no. 3 (2024).

[19]. Juliet, Sujitha, and J. Anitha. "Machine Learning and Survival Analysis Models for Postpartum Depression: A Comprehensive Risk Factor Analysis." In *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, pp. 1340-1345. IEEE, 2024.

[20]. Matsumura, Kenta, Kei Hamazaki, Haruka Kasamatsu, Akiko Tsuchida, and Hidekuni Inadera. "Decision tree learning for predicting chronic

postpartum depression in the Japan Environment and Children's Study."
Journal of Affective Disorders 369 (2025): 643-652.