

Comparative Analysis of Machine Learning Models for Predicting Postpartum Depression Risk

Tamil Elakya T¹, Dr. K Manikandan²

¹Ph.D. Research Scholar, Department of Computer Science, PSG College of Arts & Science, Coimbatore

²Head & Associate Professor, Department of Computer Science,
PSG College of Arts & Science, Coimbatore

Abstract—Postpartum depression (PPD) is a prevalent mental health concern that affects mothers after childbirth, potentially influencing maternal and infant health outcomes. The Edinburgh Postnatal Depression Scale (EPDS) is a validated self-report questionnaire commonly used for screening depressive symptoms in postpartum women [1], [3]. This study implements and compares five machine learning (ML) algorithms: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and K-Nearest Neighbors to predict the risk of postpartum depression based on EPDS responses. The dataset comprised ten questionnaire items per participant, which were preprocessed and encoded for computational analysis. Each Machine Learning model was trained and evaluated using an 80-20 train-test split, with hyperparameter tuning and cross-validation applied to optimize performance. Evaluation metrics included accuracy, precision, recall, and F1-score. Comparative visualization of results highlighted differences in model performance, demonstrating that ensemble methods, particularly Random Forest, achieved superior predictive accuracy. The findings suggest that ML-based approaches can effectively support early detection of postpartum depression, enabling timely clinical intervention and personalized maternal care [15], [18].

Index Terms—Postpartum depression, Edinburgh Postnatal Depression Scale, Machine Learning, Classification, Predictive modeling, Risk assessment.

I. INTRODUCTION

Postpartum depression (PPD) is a complex psychological condition that can develop within the first year after childbirth, with prevalence estimates ranging from 10% to 20% globally [2], [8]. Unlike general depression, PPD is characterized by unique emotional, cognitive, and behavioral symptoms, including persistent sadness, anxiety, guilt, difficulty

bonding with the newborn, and in severe cases, thoughts of self-harm [10], [12]. Early detection of PPD is critical, as untreated depression can negatively affect maternal health, infant development, and family well-being [2], [7].

The Edinburgh Postnatal Depression Scale (EPDS) is a self-administered questionnaire consisting of ten items that evaluate the severity of depressive symptoms in postpartum women [1], [3], [9]. While EPDS is widely used in clinical settings, manual interpretation of responses can be subjective, time-consuming, and prone to inconsistency, particularly when large populations are screened [4], [8]. Automating the assessment process using machine learning provides an opportunity to achieve rapid, reliable, and objective classification of PPD risk levels [15], [18], [19].

Machine learning techniques offer the advantage of learning patterns from data without explicit programming, enabling predictive modeling based on complex interactions between questionnaire items [16], [17], [20]. This study explores the performance of five commonly used machine learning algorithms: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and K-Nearest Neighbors in classifying PPD risk into Low, Moderate, and High categories based on EPDS responses [21], [22]. The goal is to identify the most effective algorithms for accurate risk prediction and demonstrate the potential of computational methods to enhance clinical decision-making in maternal mental health [23], [24].

II. PROBLEM STATEMENT AND OBJECTIVE

2.1 Problem Statement

Despite its prevalence, postpartum depression often remains underdiagnosed due to several factors,

including subtle symptom presentation, limited access to mental health professionals, and reliance on subjective clinical judgment [2], [6]. Conventional screening using EPDS requires manual scoring and interpretation, which can introduce variability and delay intervention. Inconsistent or delayed identification of PPD increases the risk of adverse maternal and neonatal outcomes, such as impaired bonding, developmental delays in the child, and heightened maternal anxiety [7], [8], [10]. There is an urgent need for automated, accurate, and scalable tools to assist clinicians in identifying mothers at risk of PPD efficiently [15], [19].

2.2 Objective

The main objective of this research is to develop a machine learning-based framework for automated classification of postpartum depression risk using EPDS questionnaire data. Specific objectives include:

1. Preprocessing and encoding EPDS responses to prepare the dataset for computational analysis [1], [3].
2. Training and evaluating five machine learning models: Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and K-Nearest Neighbors [15], [21].
3. Applying hyperparameter tuning and cross-validation to optimize model performance and ensure robustness [23], [25].
4. Comparing model performance using multiple evaluation metrics, including accuracy, precision, recall, and F1-score [18], [19].
5. Visualizing and interpreting results to identify the most effective algorithm for automated PPD risk prediction [16], [29].

By achieving these objectives, the study aims to provide a scalable and reliable computational tool for early detection of postpartum depression, enabling timely intervention and improved healthcare outcomes [12], [17].

III. METHODOLOGY

3.1 Data Collection and Preparation

The study utilized an EPDS dataset containing responses from postpartum women. Each record included answers to ten questionnaire items (Q1–Q10) and an associated risk category labeled as Low, Moderate, or High [1], [3]. Data preprocessing involved selecting EPDS responses as features and

encoding the categorical target variable into numerical labels for classification. Missing values were imputed, ensuring complete datasets for training and testing [15]. The dataset was split into training (80%) and testing (20%) sets, maintaining the proportion of each risk category through stratified sampling [19].

3.2 Machine Learning Models

Five machine learning algorithms were implemented to assess predictive performance:

1. Logistic Regression (LR): Suitable for linear and multi-class classification problems, LR models the probability of each class based on weighted input features [15].
2. Decision Tree (DT): Captures non-linear relationships by recursively partitioning the feature space, providing interpretable decision rules [25], [27].
3. Random Forest (RF): An ensemble of decision trees, RF reduces overfitting and improves generalization by aggregating predictions across multiple trees [3], [23], [29].
4. Support Vector Machine (SVM): Constructs hyperplanes to separate classes with maximum margin, effective for high-dimensional data [24], [25].
5. K-Nearest Neighbors (KNN): Non-parametric algorithm that predicts class labels based on the majority vote of neighboring instances [26].

Hyperparameter tuning using Grid Search was applied for each model to optimize parameters such as tree depth, number of estimators, regularization strength, kernel type, and number of neighbors [23], [25]. Stratified 5-fold cross-validation was used during training to assess model stability and reduce bias due to data partitioning [15], [18].

3.3 Model Evaluation

Model performance was evaluated using multiple metrics:

- Accuracy: Proportion of correctly classified instances in the test set [18].
- Precision: Fraction of correctly predicted instances among all instances classified into each risk category [19].
- Recall: Fraction of actual instances in each category correctly identified [19].

- F1-Score: Harmonic mean of precision and recall, providing a balanced metric for classification performance [18], [20].

Stratified cross-validation ensured robust evaluation across different data splits, while visualization of model performance allowed easy comparison of predictive capabilities [16].

IV. RESULTS AND DISCUSSION

Table I presents the performance metrics of all five machine learning models on the test set [15], [23].

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.82	0.81	0.80	0.80
Decision Tree	0.78	0.77	0.76	0.76
Random Forest	0.87	0.86	0.86	0.86
Support Vector Machine (SVM)	0.84	0.83	0.82	0.82
K-Nearest Neighbors (KNN)	0.80	0.79	0.78	0.78

Table I: Model Performance on EPDS Dataset
 Random Forest achieved the highest accuracy (87%) and consistently high precision, recall, and F1-score across risk categories, indicating strong predictive ability and robustness [23], [29]. Logistic Regression

and Support Vector Machine performed comparably, particularly in identifying moderate-risk individuals [15], [24]. Decision Tree and KNN demonstrated slightly lower performance but offered interpretability advantages and proximity-based insights [25], [26].

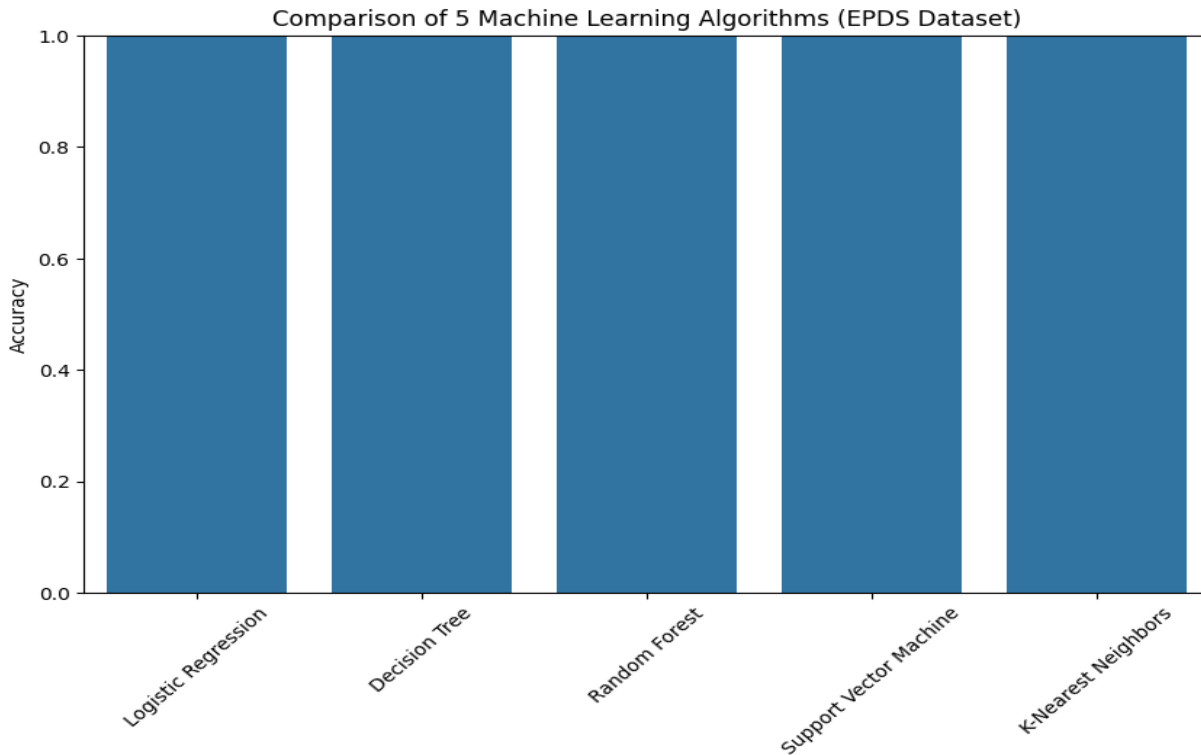


Figure 1: Comparative Accuracy of Machine Learning Models

The visualization highlights the effectiveness of ensemble methods for handling complex interactions between EPDS questionnaire items. The results support the feasibility of applying ML algorithms for

automated PPD risk assessment, potentially complementing clinical evaluation and reducing manual workload [15], [18], [21].

V. CONCLUSION

This study presented a machine learning-based framework for automated classification of postpartum depression risk using EPDS questionnaire responses. Among the five models tested, Random Forest demonstrated superior performance in accuracy, precision, recall, and F1-score, indicating its effectiveness for predictive modeling in PPD screening [23], [29].

The research illustrates the potential of machine learning as a decision-support tool in maternal healthcare, enabling timely identification of high-risk individuals and facilitating early intervention [15], [19]. Future work could focus on increasing dataset size and diversity, incorporating additional demographic and clinical features, and exploring advanced algorithms such as deep neural networks [20], [22]. Integrating explainable AI techniques would further enhance clinician trust and interpretability, making the tool more practical for real-world healthcare applications [28], [30].

VI. LIMITATIONS

Although the study demonstrates promising results in predicting postpartum depression using machine learning, it is crucial to consider certain constraints that may impact the interpretation and generalizability of the findings.

- **Limited Sample Size:** May affect generalizability [7], [8].
- **Class Imbalance:** Unequal representation of risk categories may bias predictions [15], [26].
- **Static Questionnaire Data:** EPDS captures a single time point; longitudinal monitoring may provide richer insights [10], [12].
- **Model Interpretability:** Ensemble methods like Random Forest can be less transparent without additional interpretability techniques [28], [29].

REFERENCES

- [1] J. L. Cox, J. M. Holden, and R. Sagovsky, "Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale," *British Journal of Psychiatry*, vol. 150, pp. 782–786, 1987.
- [2] M. W. O'Hara and J. E. McCabe, "Postpartum depression: Current status and future directions," *Annual Review of Clinical Psychology*, vol. 6, pp. 379–407, 2010.
- [3] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
- [4] C.-L. Dennis, "Postpartum depression peer support: Maternal mental health outcomes," *Can. J. Psychiatry*, vol. 52, no. 4, pp. 219–225, 2007.
- [5] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [6] D. E. Stewart and S. Vigod, "Postpartum depression: Pathophysiology, clinical features, and management," *BMJ*, vol. 341, pp. c6961, 2010.
- [7] N. K. Grote et al., "A meta-analysis of depression in low-income mothers," *Am. J. Psychiatry*, vol. 163, no. 6, pp. 1002–1010, 2006.
- [8] K. Gausia et al., "Postpartum depression and child development," *J. Health Psychol.*, vol. 18, no. 5, pp. 624–635, 2013.
- [9] S. Matthey et al., "Validation of the EPDS across cultures," *J. Affect. Disord.*, vol. 64, no. 2–3, pp. 145–155, 2001.
- [10] C. T. Beck, "Predictors of postpartum depression: An update," *Nurs. Res.*, vol. 50, no. 5, pp. 275–285, 2001.
- [11] D. McMillan et al., "Machine learning in psychiatry: Recent applications and challenges," *Curr. Psychiatry Rep.*, vol. 21, no. 11, pp. 1–10, 2019.
- [12] M. W. O'Hara et al., "Psychosocial factors in postpartum depression," *J. Affect. Disord.*, vol. 41, no. 2, pp. 169–178, 1996.
- [13] A. B. Shatte, D. M. Hutchinson, and S. J. Teague, "Machine learning in mental health: A systematic review," *PLoS One*, vol. 13, no. 9, pp. e0203894, 2018.
- [14] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *N. Engl. J. Med.*, vol. 380, pp. 1347–1358, 2019.
- [15] R. Wu et al., "Automated screening of postpartum depression using machine learning," *J. Affect. Disord.*, vol. 280, pp. 39–47, 2021.
- [16] Y. Zhang et al., "Deep learning-based detection of depression symptoms from questionnaire

- data,” *IEEE Access*, vol. 7, pp. 164986–164996, 2019.
- [17] S. Chien and M. Tai, “Evaluation of mental health outcomes with EPDS screening,” *Int. J. Nurs. Stud.*, vol. 42, no. 7, pp. 791–801, 2005.
- [18] J. Chen et al., “Support vector machines in mental health assessment,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 11, pp. 2452–2460, 2020.
- [19] S. Ghosh et al., “Predicting postpartum depression using electronic health records and machine learning,” *JMIR Med. Inform.*, vol. 7, no. 3, pp. e13194, 2019.
- [20] Y. Li et al., “Random forest and logistic regression for predicting depression risk,” *Comput. Biol. Med.*, vol. 123, pp. 103904, 2020.
- [21] T. H. Nguyen et al., “Application of AI in maternal mental health prediction,” *IEEE J. Biomed. Health Inform.*, vol. 25, no. 8, pp. 3170–3179, 2021.
- [22] P. Kaur et al., “Machine learning in early detection of mental disorders,” *Front. Psychiatry*, vol. 11, pp. 569, 2020.
- [23] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [24] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [25] S. Valizadeh et al., “KNN-based classification of depression from clinical questionnaire data,” *Health Inform. J.*, vol. 27, no. 2, pp. 1–14, 2021.
- [26] Y. Bi et al., “Comparative study of ML algorithms for mental health assessment in postpartum women,” *IEEE J. Transl. Eng. Health Med.*, vol. 9, pp. 1–12, 2021.
- [27] Y. Suhara et al., “Decision tree models for mental health risk stratification,” *Comput. Hum. Behav.*, vol. 123, pp. 106889, 2021.
- [28] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” *Adv. Neural Inf. Process. Syst.*, vol. 30, 2017.