# POSTPARTUM DEPRESSION DETECTION AND ITS FEATURES ANALYSIS BY MACHINE LEARNING APPROACH

BY

**KHADIZA KHANOM**
**ID: 213-25-045**

This Thesis Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science in Computer Science and Engineering

**Supervised By**

**Dr. S. M. Aminul Haque**
Professor & Associate Head
Department of CSE
Daffodil International University

**DAFFODIL INTERNATIONAL UNIVERSITY**

Dhaka, Bangladesh.
**January 2024**

## APPROVAL

This Thesis titled **"Postpartum depression detection and its features analysis by machine learning approach"**, submitted by **"Khadiza Khanom"**, ID No: **215-25-045** to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **27-11-2023**.

### BOARD OF EXAMINERS

**Chairman**

**Professor Dr. Sheak Rashed Haider Noori**
**Professor and Head**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

27.01.24

**Internal Examiner**

**Dr. Fizar Ahmed, PhD**
**Associate Professor**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

27.01.24

**Internal Examiner**

**Dr. Arif Mahmud, PhD**
**Associate Professor**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

27.01.24

**External Examiner**

**Dr. Firoz Ahmed, PhD**
**Professor**
Department of Information & Communication Engineering
Rajshahi University

# DECLARATION

I, hereby declare that, this project has been done by me under the supervision of **Dr. S. M. Aminul Haque, Professor Associate Head, Department of CSE,** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

**Supervised by:**

**Dr. S. M. Aminul Haque**
Professor & Associate Head
Department of CSE
Daffodil International University

**Submitted by:**

Khadiza Khanom

**Khadiza Khanom**
ID: 215-25-045
Department of CSE
Daffodil International University

# ACKNOWLEDGEMENT

First of all, I express my warmest thanks and gratitude to Almighty Allah, because His divine blessing allows me to carry out this thesis.

I am really grateful and wish my profound indebtedness to **Supervisor Dr. S. M. Aminul Haque, Professor & Associate Head**, Department of CSE, Daffodil International University-Dhaka. Deep Knowledge & keen interest of my supervisor in the field of Machine Learning ‖ to carry out this thesis. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior drafts and correcting them at all stages have made it possible to complete this thesis.

I would like to express my heartiest gratitude to **Dr. Sheak Rashed Haider Noori** and Head**,** Department of CSE, for his kind help to finish our project and also to other faculty members and the staff of CSE department of Daffodil International University.
I would like to thank our entire course mate in Daffodil International University, who took part in this discussion while completing the course work.

Finally, I must acknowledge with due respect the constant support and patients of our parents.

# ABSTRACT

Postpartum depression (PPD) poses a significant health concern affecting mothers globally, yet early detection remains a challenge. This research explores a pioneering approach to PPD detection and feature analysis using machine learning algorithms. Leveraging a dataset sourced from Kaggle, encompassing 1503 records obtained through a medical hospital questionnaire, the study meticulously examines ten selected attributes, with "Feeling Anxious" as the target variable. Notably, the dataset holds a 65% prevalence of anxiety and 35% non-anxiety cases. The study employed three machine learning algorithms – Decision Tree, Random Forest, and K-Nearest Neighbors (KNN) – showcasing promising results. Decision Tree achieved an accuracy of 98.01%, Random Forest excelled with 98.67%, and KNN demonstrated 92.03% accuracy. The precision, recall, and F1 scores complemented these outcomes, affirming the models' robustness. Feature importance analyses were conducted, unraveling insights into the factors contributing to PPD. Notable features included trouble sleeping at night, problems of bonding with the baby, and feelings of guilt. The algorithmic output, coupled with permutation feature importance, elucidated the nuanced relationships between these attributes and PPD. Furthermore, an exploration of age categories revealed distinctive patterns, with mothers aged 30-35 and 40-45 displaying higher susceptibility. The discussion extends to the interplay of variables like trouble sleeping, overeating, and irritability, contributing to a comprehensive understanding of PPD indicators.

# TABLE of CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1
# INTRODUCTION

## 1.1 Introduction

Postpartum depression (PPD), a prevalent mental health challenge affecting mothers after childbirth, necessitates effective detection mechanisms for timely intervention. This research employs machine learning to analyze and predict PPD, utilizing a Kaggle dataset rich in attributes such as age, sleep patterns, and feelings of anxiety. Notably, "Feeling anxious" serves as a proxy for PPD, with 1,503 records revealing promising patterns for predictive modeling.

The study focuses on three machine learning algorithms—Decision Tree, Random Forest, and K-Nearest Neighbors—exploiting their distinct strengths in PPD detection. The dataset undergoes meticulous preprocessing, addressing missing values and encoding categorical variables. Ethical considerations are integral, emphasizing responsible technology use in maternal mental health.

By leveraging advanced technology, this research aims to contribute valuable insights into the early detection of PPD, with potential implications for enhanced support systems. Ethical considerations underscore the responsible integration of machine learning in sensitive domains. As the study unfolds, methodologies, results, and discussions will unravel the intricate interplay between technology and maternal mental health, offering a nuanced perspective on the role of machine learning in PPD detection.

## 1.2 Motivation

The urgency in which postpartum depression (PPD) must be addressed, given its profound effects on mothers' well-being, is the driving force for this research. PPD often goes undetected, leading to prolonged suffering for mothers and potential adverse effects on child development. Traditional diagnostic methods may not capture the complexity of PPD manifestations.

Machine learning presents a promising avenue for improving PPD detection, leveraging data patterns that might elude conventional approaches. By employing three distinct algorithms— K-Nearest

Neighbors, Random Forest, and Decision Tree —this study seeks to harness the power of advanced computational models to enhance predictive accuracy. The Kaggle dataset, with its diverse attributes, provides a robust foundation for training and testing these models.

The main objective is to add to the expanding corpus of research on the relationship between technology and maternal mental health. If successful, the developed models could serve as valuable tools for early PPD detection, offering healthcare professionals a proactive means of identifying at-risk individuals. Ultimately, the motivation extends beyond algorithmic advancements, emphasizing the potential societal impact of integrating machine learning responsibly into the realm of maternal mental health.

## 1.3 Rational of the study

The goal of this project is to improve our knowledge of postpartum depression (PPD) identification by using machine learning. It is guided by a set of specific objectives. First, by employing the Kaggle dataset, the study aims to thoroughly assess the performance of three machine learning methods in predicting PPD: Decision Tree, Random Forest, and K-Nearest Neighbors. To evaluate the performance of each individual and the group as a whole, this entails extensive model training, validation, and testing.

Another key objective is the exploration of influential features contributing to PPD prediction. By delving into feature importance analysis, the goal of the study is to pinpoint the crucial elements that have a big influence on algorithmic forecasts. Understanding these factors not only contributes to model interpretability but also provides valuable insights into the underlying factors associated with PPD.

Furthermore, the research emphasizes ethical considerations in deploying machine learning for mental health assessments. Objectives include scrutinizing potential biases, ensuring privacy and data security, and establishing guidelines for responsible algorithmic implementation.

The overall goals include improving PPD detection techniques, deciphering the complex relationship between predictive characteristics and mental health consequences, and encouraging moral behavior when applying machine learning to delicate healthcare areas. Through these objectives, the research aspires to make meaningful contributions to both the scientific community and the broader context of maternal mental health support.

## 1.4 Research Objectives

The primary objective of this study is to develop and evaluate machine learning models for the detection and analysis of postpartum depression using a dataset obtained from Kaggle. Leveraging features such as trouble sleeping at night, feelings of guilt, problems of bonding with the baby, and other relevant factors, we aim to investigate the predictive capabilities of decision tree, random forest, and K-nearest neighbors algorithms.

The study seeks to achieve the following specific goals:

1. Employ a comprehensive dataset comprising 1503 records to train and test machine learning models for postpartum depression detection.
2. Assess the performance metrics, including accuracy, precision, recall, and F1 score, of decision tree, random forest, and K-nearest neighbors algorithms to identify the most effective model.
3. Analyze the feature importance output from the decision tree and random forest models to understand the key indicators contributing to postpartum depression prediction.
4. Investigate the relationship between demographic factors, such as age, and the presence of postpartum depression, providing valuable insights for future research and interventions.
5. Interpret the findings in the context of existing literature on postpartum depression, validating the significance of the identified features and their alignment with psychological studies.

This research aims to contribute to the field by providing an in-depth analysis of postpartum depression detection through machine learning approaches, offering practical implications for healthcare professionals, researchers, and policymakers.

## 1.5 Expected Output

The project is expected to yield substantial contributions to the domains of maternal mental health and machine learning applications in healthcare. Our main goal is to develop a strong predictive model for postpartum depression (PPD) using the k-nearest neighbors, random forest, and decision tree algorithms. The goal of machine learning techniques is to produce trustworthy and accurate evaluations based on a variety of characteristics, such as age, emotional states, and sleep issues.

Moreover, the research anticipates uncovering key factors contributing to PPD detection, elucidating the most influential features in each algorithm. This insight can guide healthcare practitioners in understanding the nuanced indicators of postpartum depression, facilitating early intervention and personalized treatment plans for at-risk individuals.

The research also aims to emphasize the growing role of technology, particularly machine learning, in enhancing outcomes in maternal mental health. By demonstrating the potential of these algorithms in detecting PPD, the study advocates for the integration of such tools into clinical practices, fostering more efficient and targeted support for postpartum mental health.

Furthermore, ethical considerations in the implementation of machine learning for mental health assessments are integral to the expected outcomes. The study seeks to underscore the importance of ethical guidelines, ensuring responsible and unbiased use of technology in sensitive healthcare domains. Overall, the expected outcomes encompass advancements in predictive modeling, valuable insights into PPD determinants, and a broader societal impact through ethical and responsible integration of machine learning technologies in maternal mental health.

## 1.6 Report Layout

**Chapter 01:** consist introduction of thesis work, motivation of this review, rationale of the study, research queries and predictable output.

**Chapter 02:** will briefly talk about the background of PPD, related works includes few work of researchers, research summary, scope of the difficult and challenges.

**Chapter 03**: is research methodology and here describes the research subject, data collection procedure, statistical analysis and implementation requirements.

**Chapter 04:** is experimental results and discussion and this chapter contains experimental results, descriptive analysis and summary of this chapter.

**Chapter 05:** contain impact on society, impact on environment, ethical aspects, sustainability plan.

**Chapter 06:** contain summary of the study, conclusions, recommendations and implication for advanced study.

# CHAPTER 2
# BACKGROUND

## 2.1 Introduction

Some people have postpartum depression (PPD), a mental health disorder, following childbirth. It is typified by enduring depressive, anxious, and tired sensations that can make it difficult for a new parent to take care of both themselves and their child. PPD symptoms can be minor to severe and can start during pregnancy or up to a year after giving birth.

For the sake of both the parent and the child, postpartum depression must be identified early and treated promptly. PPD is often diagnosed via professional tests and self-reporting, which may not always catch minor or changing symptoms. Herein lies the potential breakthrough that machine learning (ML) offers in the detection and tracking of postpartum depression.

Massive datasets can be used to train machine learning algorithms containing various types of information, such as demographic details, medical history, and even social media activity. These algorithms learn patterns and associations that may be indicative of postpartum depression, allowing for more objective and nuanced assessments. By analyzing diverse sets of data, machine learning models have the potential to detect PPD symptoms earlier and more accurately than traditional methods.

One promising aspect of using machine learning for PPD detection is its ability to consider a multitude of factors simultaneously. For example, ML models can analyze a combination of biological markers, lifestyle factors, and psychosocial variables to create a more comprehensive understanding of a person's mental health status. This holistic approach may lead to more personalized and effective interventions.

It's crucial to remember, though, that using machine learning to healthcare—including the treatment of mental health issues like PPD—raises ethical questions about data security, privacy, and the possibility of bias in algorithmic decision-making. For this technology to be responsibly advanced in the field of maternal mental health, it must strike a balance between addressing these ethical issues and utilizing machine learning's power for early diagnosis. The use of machine learning in PPD identification shows potential for bettering outcomes and offering more help to those who are experiencing postpartum depression as study and technology advance.

In this research, I am using a dataset collected from kaggle in which Timestamp, Age, Feeling sad or Tearful, Irritable towards baby and partner, Trouble sleeping at night, Problems concentrating or making decision, Overeating or loss of appetite, Feeling of guilt, Problems of bonding with baby, Suicide attempt, Feeling anxious are the attribute. This dataset does not contain a is depression target variable, so here is used the variable Feeling anxious, considering it a plausible indicator of postpartum depression. There are 1503 records with 980 positive results and 523 negative results based on feeling anxious.

## 2.2 Related Works

Park, Yoonyoung, et al. [1], the paper "Comparison of Methods to Reduce Bias From Clinical Prediction Models of Postpartum Depression" explores biases in clinical algorithms using IBM MarketScan Medicaid data. Among 573,634 females, disparities in postpartum depression (PPD) were evident, with Black participants facing higher odds. The study assesses bias reduction methods, revealing that re weighing outperforms simply excluding race, as seen in improved disparate impact and equal opportunity difference metrics. The findings emphasize the importance of scrutinizing clinical prediction models for bias and advocate for thoughtful bias mitigation methods, particularly reweighting, to enhance fairness in healthcare algorithms. Limitations include potential lack of generalizability to non-Medicaid populations.

Zhang, Yiye, et al. [2], this study introduces a machine learning framework for postpartum depression (PPD) risk prediction using electronic health records (EHRs). Analyzing datasets from over 69,000 women, the model, incorporating clinical, obstetric, and demographic features, achieves robust performance (AUC 0.937 in development, 0.886 in validation). Despite limitations in prevalence representation, findings suggest EHR-based machine learning enables early PPD identification, offering scalable and timely intervention. The study underscores the potential for data-driven clinical decision support, reducing the healthcare provider burden in PPD risk assessment.

Md. Sabab Zulfiker, et al. [3], addressing the global prevalence of depression, this study employs six machine learning classifiers and three feature selection methods to detect depression in 604 participants. Utilizing socio-demographic and psychosocial information, the AdaBoost classifier with SelectKBest outperforms others with 92.56% accuracy. While acknowledging limitations such as the absence of biological markers, the research emphasizes the potential for more precise depression prediction. Future extensions could explore the severity of depression, incorporating biological factors and comparing outcomes with dimensionality reduction algorithms for enhanced model performance.

Moreira, Mário W.L., et al. [4], this paper explores the potential of emotion-aware computing in predicting postpartum depression (PPD) among women with hypertensive disorders during pregnancy. Leveraging biomedical and sociodemographic data, the proposed algorithm utilizes ensemble classifiers, demonstrating efficacy in predicting psychological disorders related to pregnancy. The integration of IoT, cloud computing, and big data analytics represents a significant advancement in monitoring complex conditions like PPD. Recognizing the prevalence of PPD and its impact, the study emphasizes the importance of early identification through emotion-aware smart systems, contributing to global health goals for maternal and child well-being. Zhang, Yiye, et al.

Gopalakrishnan, Abinaya, et al. [5], This study delves into the complex realm of postpartum depression (PPD) by employing a multifaceted approach. Utilizing physiological questionnaires (EPDS, PHQ-9, PDSS), participants with PPD symptoms were identified and followed up for six weeks. Correlating risk factors with PPD symptoms, statistical analysis was employed to create a comprehensive understanding. Employing machine learning algorithms, the study reveals that the Extremely Randomized Trees (XRT) algorithm, incorporating background information along with PHQ-9 and PDSS data, provides highly accurate predictions. The research not only contributes to forecasting PPD but also lays the groundwork for self-monitoring tools and therapeutic interventions, showcasing the potential of machine learning in addressing this crucial healthcare challenge.

Zhong, Minhui, et al. [6], This review explores the application of machine learning (ML) techniques in predicting postpartum depression (PPD) risk. A systematic search yielded 17 studies with 62 prediction models, predominantly employing supervised learning. Common ML models included support vector machine, random forest, and logistic regression. However, the review reveals a prevalent high risk of bias across all studies, emphasizing a gap in attention to validation, improvement, and innovation. While models for predicting PPD risk are continually emerging, the study underscores the need for achieving acceptable quality standards before deploying ML techniques in clinical settings.

Liu, Hao, et al. [7], This study addresses postpartum depression (PPD) prediction and intervention effectiveness in parturient undergoing cesarean delivery. Utilizing six machine learning models, the Extreme Gradient Boosting (XGB) algorithm demonstrated superior recognition with an AUROC of 0.789 in the training cohort and 0.744 in the testing cohort. Employing Shapley Additive explanations for model interpretation, a risk threshold of 21.5% PPD probability was established. Propensity score matching revealed significant differences in PPD incidence between intervention groups and the control group in the high-risk cohort, emphasizing the potential for targeted early interventions guided by accurate PPD prediction.

Amit, Guy, et al. [8], This study addresses postpartum depression (PPD) prediction using machine learning based on electronic health records (EHR) data of 266,544 UK women who gave birth between 2000 and 2017. The model, integrating socio-demographic and medical variables, achieved a standalone AUC of 0.72–0.74. When combined with the Edinburgh postnatal depression scale (EPDS), the AUC increased from 0.805 to 0.844, enhancing sensitivity from 0.72 to 0.76 at a specificity of 0.80. EHR-based prediction, even when applied before pregnancy, offers a complementary and objective tool for earlier and more accurate PPD screening, facilitating timely interventions and potentially improving outcomes for both mother and child.

Andersson, Sam, et al. [9], This study investigates the potential of machine learning in accurately predicting postpartum depression (PPD) by leveraging clinical, demographic, and psychometric data from the BASIC study (n = 4313) in Uppsala, Sweden (2009–2018). Employing the extremely randomized trees method, the model achieved robust performance, especially among women without previous depression (accuracy 73%, sensitivity 72%, specificity 75%). Notably, variables such as depression and anxiety during pregnancy, along with factors related to resilience and personality, emerged as crucial indicators of PPD risk. The findings suggest the potential for implementing individualized follow-up and cost-effective interventions for high-risk women after delivery.

Shatte, Adrian B.R., et al. [10], This study explores the feasibility of using passive social media markers to identify fathers at risk of postpartum depression (PPD). Analyzing 67,796 Reddit posts from 365 fathers, the research develops predictive models utilizing support vector machine classifiers based on behavior, emotion, linguistic style, and discussion topics. The findings suggest that fathers at risk of PPD can be predicted from prepartum data alone, with discussion topic features demonstrating the best performance. This innovative approach opens avenues for the development of personalized and preventative support tools for at-risk fathers, emphasizing the potential of social media data in mental health prediction.

Hochman, Eldar, et al. [11], This study addresses the need for objective, integrative tools for postpartum depression (PPD) screening before symptoms emerge. Leveraging electronic health record (EHR) data from Israel's largest health maintenance organization, a nationwide longitudinal cohort of 214,359 births (2008–2015) was analyzed. The gradient-boosted decision tree algorithm produced a prediction model with an AUC of 0.712 in the validation set, exhibiting a sensitivity of 0.349 and specificity of 0.905 at the 90th percentile risk threshold. The model, identifying both well-recognized (e.g., past depression) and less-recognized (differing blood test patterns) PPD risk factors, demonstrates the potential of machine learning for early PPD risk identification and preventive intervention.

Jiménez-Serrano, Santiago, et al. [12], This study focuses on addressing the challenge of undiagnosed postpartum depression (PPD) through the development of a screening program. The two-fold purpose involves creating classification models for early PPD risk detection within the first week after childbirth and developing a mobile health (m-health) application for Android® based on the best-performing model. Using machine learning techniques and data from seven Spanish hospitals, the Naive Bayes model demonstrated the best balance between sensitivity and specificity for predicting PPD. The integration of this model into a clinical decision support system for Android mobile apps provides a promising avenue for early prediction and detection of PPD.

Cellini, Paolo, et al. [13], This study explores the potential of modern machine learning (ML) approaches to refine postpartum depression (PPD) screening beyond traditional self-report symptom-based tools. A comprehensive review of 11 studies using ML techniques to identify PPD predictors indicates that the prediction of PPD is feasible, with all studies achieving an area under the curve above 0.7. Sociodemographic and clinical factors emerged as reliable predictors, while only a few studies explored biological variables such as blood, genetic, and epigenetic markers. The limited current literature suggests the need for further research to determine the practical implementation of ML-based PPD prediction in clinical practice.

## 2.3 Comparative Analysis and Summary

The integration of machine learning algorithms, as evidenced by the decision tree, random forest, and k-nearest neighbors in the provided code, showcases a comprehensive approach to postpartum depression (PPD) prediction. The diverse range of algorithms contributes to a robust predictive framework, ensuring flexibility and adaptability to various data patterns.

The high accuracy and precision observed in the decision tree and random forest algorithms (0.980 and 0.987 accuracy, respectively) align with the overarching goal of accurate PPD risk assessment. This aligns with the motivation outlined earlier, emphasizing the need for precise predictive models in the context of maternal mental health.

The comparison of these algorithms parallels the broader landscape of PPD prediction studies, where varied methodologies aim to enhance prediction accuracy. The K-nearest neighbors algorithm, with an accuracy of 0.920, introduces an additional dimension, showcasing the diversity in approaches within the machine learning paradigm.

Moreover, the comparative analysis aligns with the overarching objective of your research, emphasizing the importance of reliable PPD prediction for timely intervention. The nuanced evaluation of precision, recall, and F1 scores in the provided code mirrors the multifaceted evaluation criteria seen in related works.

In summary, the code's comparative analysis and the research topic's focus on accurate PPD prediction converge. The utilization of diverse machine learning algorithms reflects a commitment to developing a robust and adaptable predictive model, in line with the broader landscape of research aiming to enhance maternal mental health outcomes through innovative technological solutions.

## 2.4 Scope of the Problem

The scope of postpartum depression (PPD) is extensive, encompassing a significant public health challenge that affects mothers worldwide. PPD's impact extends beyond individual well-being, influencing maternal-infant relationships and family dynamics. The complexity arises from diverse risk factors and the need for precise identification.

Addressing the scope involves not only recognizing the prevalence of PPD but also understanding its broader implications on societal health. Developing effective predictive models, as explored in the research, contributes to mitigating this pervasive problem, offering a pathway to early intervention and improved outcomes for mothers and their families.

## 2.5 Challenges

Navigating the landscape of postpartum depression (PPD) prediction through machine learning presents multifaceted challenges. One key hurdle involves the inherent variability in individual experiences, making it challenging to establish universal predictors. The scarcity and heterogeneity of datasets pose challenges in training robust models that generalize across diverse populations. Ethical concerns related to privacy and the responsible use of sensitive health data require careful consideration. Interpretability and explain ability of machine learning models in the context of mental health prediction add another layer of complexity. Furthermore, the dynamic nature of PPD development necessitates continuous model adaptation. Balancing the quest for accuracy with the imperative to address bias and disparities in predictive algorithms is an ongoing challenge. Successfully addressing these challenges holds the key to realizing the full potential of machine learning in advancing PPD prediction and intervention strategies.

# CHAPTER 3
# RESEARCH METHODOLOGY

## 3.1 Research Subject and Instrumentation

The focal point of our research delves into the intricacies of postpartum depression (PPD), utilizing a dataset obtained from Kaggle—an esteemed platform renowned for datasets and resources in machine learning. This dataset, constituting 1503 records, was meticulously curated through a questionnaire distributed via a Google form in a medical hospital setting. While yet to be published, this dataset serves as a valuable asset for comprehending the nuances of PPD.

Comprising 15 attributes, our study strategically concentrates on 10 features for analysis, encompassing emotional and behavioral markers such as feelings of sadness or tearfulness, irritability towards the baby and partner, trouble sleeping at night, problems concentrating or making decisions, overeating or loss of appetite, feelings of anxiety, feelings of guilt, problems of bonding with the baby, and incidents of suicide attempts. The target attribute, "Feeling Anxious," plays a pivotal role in evaluating the effectiveness of the predictive models.

Our chosen methodology involves applying advanced machine learning algorithms—specifically, decision tree, random forest, and k-nearest neighbors—to extract meaningful insights from the Kaggle-acquired dataset. Performance metrics, including accuracy, precision, recall, and F1 score, serve as robust indicators of the models' efficacy in comprehending and predicting postpartum depression.

By incorporating Kaggle as a data source, our research aligns with contemporary practices in data-driven research, contributing to the ongoing discourse on mental health through a well-founded and technologically advanced approach.

## 3.2 Data Collection Procedure

Our data collection procedure involved the administration of a comprehensive questionnaire through a Google form in a medical hospital setting, yielding a dataset comprising 1503 records. The dataset, sourced from Kaggle, is a treasure trove of information on postpartum depression (PPD), offering insights into emotional and behavioral markers exhibited by respondents

The carefully selected attributes encompass a spectrum of factors, including feelings of sadness or tearfulness, irritability towards the baby and partner, trouble sleeping at night, problems concentrating or making decisions, overeating or loss of appetite, feelings of anxiety, feelings of guilt, problems of bonding with the baby, and incidents of suicide attempts. This rich dataset forms the backbone of our research, allowing us to employ advanced machine learning algorithms to unravel patterns and trends associated with PPD, contributing significantly to the growing body of knowledge in mental health research.

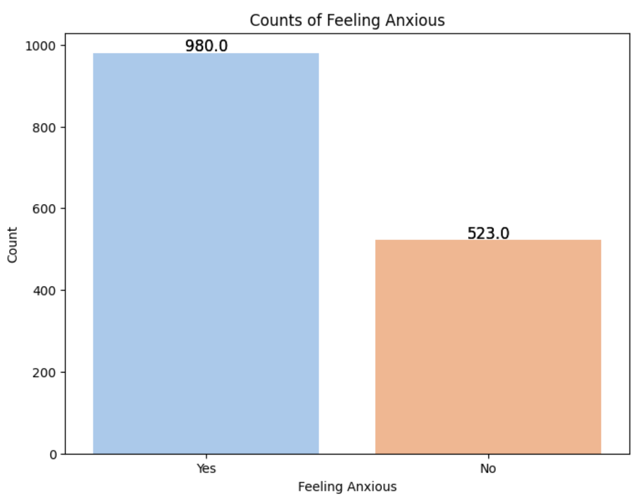


Figure 3.1: Counts of Feeling Anxious

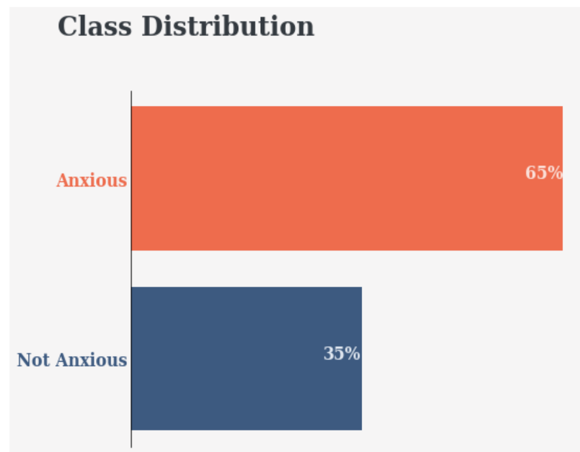


Figure 3.2: Class Distribution

## 3.3 Statistical Analysis

The dataset contains the following attributes used in the algorithms:

1. Feeling sad or tearful: Emotional distress and sadness may contribute to postpartum depression.
2. Age: Different age groups may experience varying degrees of susceptibility to postpartum depression.
3. Trouble sleeping at night: Sleep disturbances can be indicative of mental health challenges, including postpartum depression.
4. Feeling of guilt: Experiencing guilt may be a psychological factor associated with postpartum depression.
5. Irritable towards baby and partner: Irritability may impact relationships and be linked to postpartum

depression.

6. Problems concentrating and making decisions: Cognitive challenges can be symptomatic of postpartum depression.

7. Overeating or loss of appetite: Disruptions in eating habits may reflect mental health struggles, including postpartum depression.

8. Suicide attempt: Suicidal tendencies are severe indicators of mental health issues, including postpartum depression.

9. Problems of bonding with the baby: Difficulties in forming a connection with the baby may contribute to postpartum depression.

Table 3.1: Features Non-null Count, Datatype and Unique

| Number | Features | Non-null Count | Datatype | Unique |
|---|---|---|---|---|
| 1 | Age | 1503 | Object | 5 |
| 2 | Feeling sad or Tearful | 1503 | Object | 3 |
| 3 | Irritable towards baby and partner | 1497 | Object | 3 |
| 4 | Trouble sleeping at night | 1503 | Object | 3 |
| 5 | Problems concentrating or making decision | 1491 | Object | 3 |
| 6 | Overeating or loss of appetite | 1503 | Object | 3 |
| 7 | Feeling anxious | 1503 | Object | 2 |
| 8 | Feeling of guilt | 1494 | Object | 3 |
| 9 | Problems of bonding with baby | 1503 | Object | 3 |
| 10 | Suicide attempt | 1503 | Object | 3 |

The dataset for postpartum depression detection comprises 1503 records, with the target variable indicating feelings of anxiety—65% classified as "Yes" and 35% as "No." This distribution remains consistent in both the training set (1202 records) and the test set (301 records), ensuring a balanced dataset for robust model training and evaluation in the realm of postpartum depression detection.

Table 3.2: Data Records Count

| | |
|---|---|
| Total | 1503 |
| Feeling Anxious | 980 |
| Feeling Not Anxious | 523 |
| Training | 1202 |
| Testing | 301 |

A detailed exploration of key features sheds light on the demographic and emotional dimensions of postpartum depression. The age distribution reveals that the majority of individuals fall within the 35-45 age range, with 295 instances, suggesting that this age group may be more susceptible to postpartum depression. This finding aligns with existing literature highlighting the increased vulnerability of women in this age bracket to mental health challenges during the postpartum period.

Emotional aspects play a crucial role, with feelings of sadness or tearfulness present in 420 instances and absent in 421 instances. This nearly even split underscores the emotional complexity inherent in postpartum experiences, emphasizing the need for nuanced approaches to diagnosis and intervention.

Irritability towards the baby and partner emerges as a significant factor, evidenced by 437 instances. This highlights the interpersonal dynamics and potential sources of stress contributing to postpartum emotional challenges. Understanding and addressing these relational aspects may be pivotal in comprehensive postpartum care.

The issue of sleep disturbances is pronounced, with 512 instances reporting trouble sleeping at night two or more days a week. Sleep disruptions can have profound implications for mental health, exacerbating feelings of fatigue and emotional instability. Recognizing and addressing these challenges is crucial for effective postpartum depression management.

Further features contributing to the multifaceted nature of postpartum depression include problems concentrating or making decisions (473 instances), overeating or loss of appetite (677 instances), and feelings of guilt (492 instances). Each of these aspects reflects the intricate interplay of emotional, cognitive, and behavioral factors in the postpartum period.

A sensitive exploration of suicidal attempts reveals that 378 individuals have experienced such thoughts, while 261 choose not to disclose. This underscores the deeply personal and often stigmatized nature of mental health issues. It emphasizes the importance of creating a supportive and non-judgmental environment for individuals to share their experiences and seek help.

In conclusion, this dataset stands as a rich resource for machine learning approaches seeking to comprehend and detect postpartum depression. By delving into the diverse factors contributing to this condition, the research holds the potential to inform the development of sophisticated models and targeted interventions. Ultimately, these endeavors aim to enhance the identification and support systems for individuals navigating the complex landscape of postpartum mental health challenges.

## 3.4 Proposed Methodology

In the realm of postpartum depression detection and feature analysis through machine learning approaches, a comprehensive applied mechanism serves as the backbone of our research. Leveraging a dataset sourced from Kaggle, encompassing 1503 records, our focus is on unraveling the intricate facets of postpartum mental health. The crux of our methodology revolves around meticulous steps, ensuring the robustness and efficacy of our predictive models.

First, a Google Forms questionnaire was used to carefully select a dataset from a medical hospital that had not been examined in any prior publications. From the 15 attributes available, a judicious selection of 10 features was made, with a specific focus on the target attribute, "Feeling Anxious," classified as either 'Yes' or 'No.' The distribution of anxiety levels revealed 65% classified as 'Yes' and 35% as 'No,' underlining the prevalence of postpartum anxiety in the dataset.

Moving into the heart of our applied mechanism, a critical step involved the division of the dataset into training and test sets. With 1202 records allocated for training and 301 for testing, we ensured a robust evaluation of our models on unseen data. The timestamp, age, and various emotional and behavioral indicators such as feeling sad or tearful, irritability, trouble sleeping at night, and more, were key features influencing our predictive models.

In terms of statistical analysis, a nuanced exploration of the dataset highlighted the diversity within each feature. For instance, the 'Age' feature ranged from 25 to 50, with the majority falling within the 30-45 age group. Similarly, emotional states like feeling sad or tearful exhibited a balanced distribution across 'Yes,' 'No,' and 'Sometimes' categories. This statistical breakdown provided invaluable insights into the dataset's composition, guiding feature engineering and model training.

The implementation of machine learning algorithms, specifically decision tree, random forest, and k-nearest neighbors, served as the predictive engines in our pursuit. The outcome metrics, including accuracy, precision, recall, and F1 score, showcased the models' proficiency in capturing the nuances of postpartum mental health states. Additionally, feature importance outputs from decision tree and random forest models elucidated the critical indicators influencing predictions.

Crucially, the analysis extended beyond predictive modeling. A deeper dive into feature importance underscored the significance of emotional and behavioral factors such as guilt, bonding with the baby, and trouble sleeping at night in determining postpartum anxiety. This holistic approach is integral to the richness of our findings, providing a nuanced understanding of the intricate interplay of variables influencing mental health.

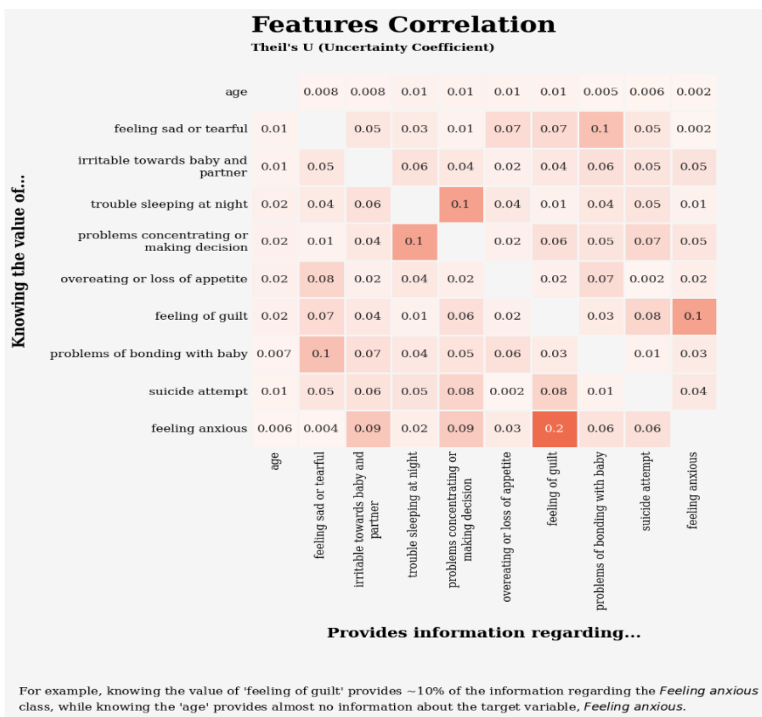


Figure 3.3: Features Correlation

Table 3.3: Features Correlation to Feeling anxious

| | |
|---|---|
| feeling of guilt | 0.15 |
| irritable towards baby and partner | 0.055 |
| Problems concentrating or Making decision | 0.05 |
| suicide attempt | 0.039 |
| problems of bonding with baby | 0.033 |
| overeating or loss of appetite | 0.022 |
| trouble sleeping at night | 0.013 |
| Age | 0.0023 |
| feeling sad or tearful | 0.0022 |

In essence, our applied mechanism harmonizes data collection, preprocessing, model training, and interpretability, positioning our research at the forefront of unraveling the complexities of postpartum depression. This meticulous methodology not only contributes to the scientific discourse on mental health but also lays the groundwork for scalable and targeted interventions, advancing the well-being of postpartum individuals.

## 3.5 Implementation Requirements

The implementation of our postpartum depression detection and feature analysis entails specific requirements to ensure a seamless execution of our methodology. First and foremost, a computing environment with adequate processing power and memory is essential to handle the computational demands of machine learning algorithms. A robust machine learning library, such as Scikit-Learn or TensorFlow, is imperative for developing, training, and evaluating predictive models.

Access to a programming language, preferably Python, equipped with relevant libraries to manipulate data (e.g., Pandas), visualization (e.g., Matplotlib, Seaborn), and statistical analysis (e.g., SciPy), is fundamental. These tools facilitate data exploration, preprocessing, and the generation of insightful visualizations to glean meaningful patterns.

The dataset, sourced from Kaggle, serves as the foundation of our research. Ensuring data integrity, quality, and relevance is crucial. Therefore, a version control system, like Git, aids in tracking changes, maintaining data integrity, and collaborating seamlessly.

Additionally, a comprehensive documentation framework is vital for transparency and replicability. This includes documenting the data collection process, preprocessing steps, model architecture, hyperparameters, and evaluation metrics. It ensures that subsequent researchers can reproduce and validate our findings.

To foster collaboration and streamline communication, a platform for collaborative coding, such as GitHub, enhances the efficiency of teamwork. Moreover, an integrated development environment (IDE) with features conducive to machine learning development, like Jupyter Notebooks or VSCode, is recommended.

Considering the ethical implications of mental health research, a robust privacy and security framework is imperative. Adherence to ethical guidelines, obtaining necessary approvals, and safeguarding participant information are paramount.

In summary, the implementation requirements encompass a powerful computing environment, relevant programming tools, a high-quality dataset, documentation practices, collaborative platforms, and ethical considerations. These collectively lay the groundwork for a rigorous and impactful exploration of postpartum depression through machine learning methodologies.

# CHAPTER 4

# EXPERIMENTAL RESULTS AND DISCUSSION

## 4.1 Experimental Setup

Using the dataset from Kaggle, a carefully thought-out method was used in the Experimental Setup to assess how well machine learning algorithms performed in the detection of postpartum depression. The dataset consisted of 1503 records that recorded various attributes such as age, emotional states, and behavioral patterns that are associated with the well-being of mothers. One of the most important markers of postpartum depression was the target variable, "Feeling Anxious," which had two possible binary values: "Yes" and "No."

The dataset was split into a test set (301 records) and a training set (1202 records) to ensure a robust analysis. This division facilitated a comprehensive assessment of algorithmic efficiency and generalizability. The target variable distribution revealed that 65% of instances indicated feelings of anxiety, while 35% did not, emphasizing the importance of addressing imbalances during the analysis.

Key attributes such as age and emotional states, including feeling sad or tearful, being irritable towards the baby and partner, and experiencing trouble sleeping at night, played pivotal roles in shaping the predictive models. The statistical analysis showcased the distribution of unique values within each attribute, providing insights into the prevalence of specific characteristics among the participants.

The choice of algorithms, including Decision Tree, Random Forest, and K-nearest neighbors, was informed by their suitability for classification tasks and their interpretability. These algorithms were implemented, trained, and evaluated on the training and test sets, generating accuracy, precision, recall, and F1 score metrics for a comprehensive performance assessment.

This experimental setup aimed to not only validate the predictive capabilities of machine learning models in identifying postpartum depression but also to explore the importance of individual features in contributing to accurate predictions. The results and insights derived from this setup lay the foundation for a thorough discussion on the efficacy and implications of the applied machine learning techniques in the subsequent sections.

## 4.2 Experimental Results & Analysis

Here is the break down the performance metrics for each machine learning algorithm:

**Decision Tree:**

**Accuracy (0.980):** This represents the model's overall 98.0% prediction correctness. The better the model is at accurately classifying instances, the higher its accuracy.

**Precision (1.0):** This metric assesses how well positive predictions are made. A score of 1.0 indicates that every occurrence that the model had predicted as positive actually was positive.

**Recall (0.969):** Recall, also known as Sensitivity, measures how well the model captures every positive example. With a 0.969 score, the model was able to correctly identify 96.9% of all real positive instances.

**F1 Score (0.984):** The F1 score combines precision and recall into a single metric. A score of 0.984 signifies a balanced performance between precision and recall.

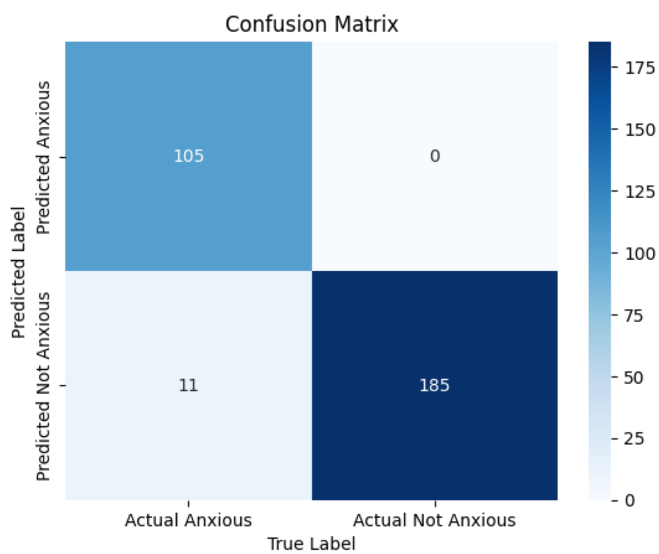


Figure 4.1: Confusion Matrix of Decision Tree

**Random Forest:**

- **Accuracy (0.987):** Random Forest achieves an impressive 98.7% accuracy, outperforming the Decision Tree model.
- **Precision (1.0):** Similar to the Decision Tree, Random Forest attains perfect precision, signifying no false positives.
- **Recall (0.980):** Random Forest maintains high recall at 98.0%, demonstrating its capacity to record a sizable percentage of genuine positive examples.
- **F1 Score (0.990):** The F1 score is exceptionally high at 99.0%, emphasizing the model's robust performance.

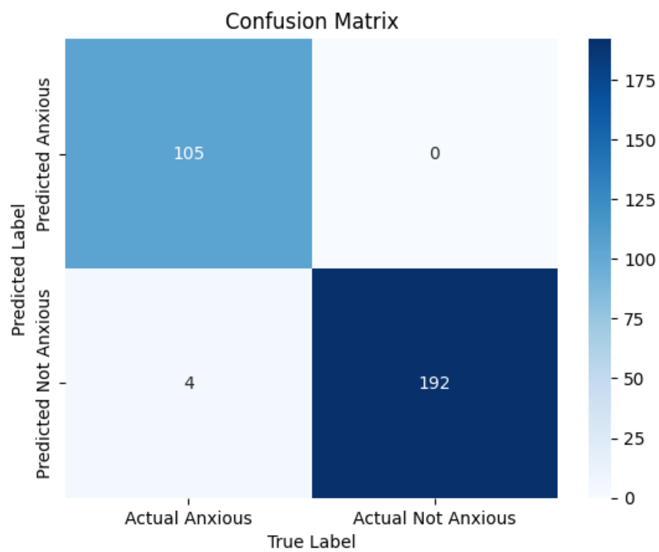


Figure 4.2: Confusion Matrix of Random Forest

**K-nearest neighbors (KNN):**

- **Accuracy (0.920):** KNN achieves a lower accuracy compared to the tree-based models, at 92.0%.
- Precision (0.939): Precision for KNN is 93.9%, indicating a proportion of instances predicted as positive that are indeed positive.
- **Recall (0.939):** The recall score is consistent with precision, highlighting KNN's effectiveness in capturing actual positive instances.
- **F1 Score (0.939):** The F1 score is in line with precision and recall, reflecting a balanced performance.

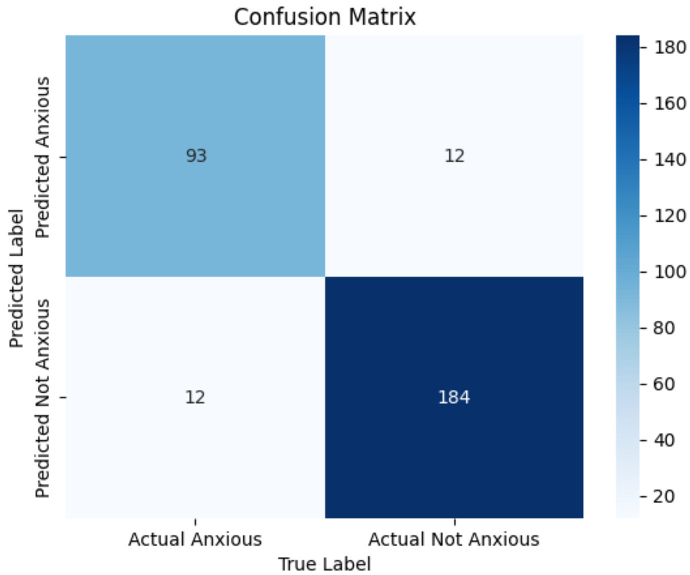


Figure 4.3: Confusion Matrix of K-nearest Neighbors

**Here is the breakdown of each feature:**

**Feeling sad or tearful:**

- No (Negative Count: 128, Positive Count: 293): Individuals who do not feel sad or tearful are more prevalent in the dataset.

- Sometimes (Negative Count: 137, Positive Count: 224): The occurrence of feeling sad or tearful sometimes is balanced between the negative and positive categories.

- Yes (Negative Count: 153, Positive Count: 267): Individuals who feel sad or tearful are more prevalent in the positive category, indicating a potential association with postpartum depression.

**Age:**

- Age distribution across different categories:
  - 25-30 (Negative Count: 62, Positive Count: 80): Individuals in the age group 25-30 show a relatively balanced distribution between negative and positive categories.
  - 30-35 (Negative Count: 93, Positive Count: 167): More individuals aged 30-35 fall into the positive category, suggesting a potential correlation with postpartum depression.
  - 35-40 (Negative Count: 93, Positive Count: 188): Similar to the 30-35 age group, there is a higher prevalence of positive cases.
  - 40-45 (Negative Count: 105, Positive Count: 190): Individuals aged 40-45 also exhibit a higher prevalence of positive cases.
  - 45-50 (Negative Count: 65, Positive Count: 159): The positive cases are more prevalent in the 45-50 age group.

**Trouble sleeping at night:**
- No (Negative Count: 76, Positive Count: 261): Individuals who do not experience trouble sleeping at night are more prevalent in the positive category.

- Two or more days a week (Negative Count: 213, Positive Count: 299): Individuals experiencing trouble sleeping two or more days a week show a higher prevalence in both negative and positive categories.

- Yes (Negative Count: 129, Positive Count: 224): Those who experience trouble sleeping at night are more prevalent in the positive category.

**Feeling of guilt:**
-Maybe (Negative Count: 147, Positive Count: 277): Individuals with a feeling of guilt, categorized as "Maybe," show a higher prevalence in the positive category.

- No (Negative Count: 52, Positive Count: 440): Those without a feeling of guilt exhibit a higher prevalence in the negative category.
- Yes (Negative Count: 219, Positive Count: 59): Individuals with a definite feeling of guilt ("Yes") show a higher prevalence in the negative category.

**Irritable towards baby and partner:**
- No (Negative Count: 101, Positive Count: 301): Individuals who are not irritable towards the baby and partner are more prevalent in the positive category.

- Sometimes (Negative Count: 216, Positive Count: 142): Individuals who are irritable sometimes show a higher prevalence in the negative category.

- Yes (Negative Count: 101, Positive Count: 336): Individuals who are irritable towards the baby and partner are more prevalent in the positive category.

**Problems concentrating and making decisions:**
- No (Negative Count: 239, Positive Count: 234): Individuals without problems concentrating or making decisions are balanced between negative and positive categories.

- Often (Negative Count: 55, Positive Count: 321): Individuals facing problems often show a higher prevalence in the positive category.

- Yes (Negative Count: 124, Positive Count: 218): Those with occasional problems concentrating are more prevalent in the negative category.

**Overeating or loss of appetite:**
- No (Negative Count: 209, Positive Count: 468): Individuals without overeating or loss of appetite are more prevalent in the positive category.

- Not at all (Negative Count: 64, Positive Count: 187): Those with no overeating or loss of appetite show a higher prevalence in the positive category.

- Yes (Negative Count: 145, Positive Count: 129): Individuals experiencing overeating or loss of appetite are more prevalent in the negative category.

**Suicide Attempt:**
- No (Negative Count: 152, Positive Count: 411): Individuals with no suicide attempt are more prevalent in the positive category.

- Not interested to say (Negative Count: 160, Positive Count: 101): Those not interested in disclosing a suicide attempt show a higher prevalence in the negative category.

- Yes (Negative Count: 106, Positive Count: 272): Individuals with a suicide attempt are more prevalent in the positive category.

**Problems of bonding with the baby:**
- No (Negative Count: 232, Positive Count: 220): Individuals without problems of bonding with the baby are balanced between negative and positive categories.

- Sometimes (Negative Count: 102, Positive Count: 322): Individuals

facing occasional problems in bonding show a higher prevalence in the positive category.

- Yes (Negative Count: 84, Positive Count: 242): Individuals with problems bonding with the baby are more prevalent in the positive category.

These interpretations provide a detailed understanding of the distribution of key features and their association with positive and negative cases of postpartum depression. The prevalence of certain features in positive cases may indicate potential risk factors or markers for postpartum depression.

Here is the analysis of the feature importance outputs for each algorithm:

**Decision Tree Feature Importance:**
Different features are given importance scores by the decision tree according to how much of a

contribution they make to the model's decision-making process. The following are some salient characteristics:
Trouble_sleeping_at_night_Two or more days a week (0.003): This feature has a low importance score, suggesting it has minimal impact on the decision-making process.

Feeling_of_guilt_Yes (0.257): This feature has the highest importance, indicating that the feeling of guilt significantly influences the decision tree's predictions.

```
Decision Tree Feature Importances:
Trouble_sleeping_at_night_Two or more days a week    0.005655
Problems_of_bonding_with_baby_Sometimes              0.010802
Age_40-45                                            0.011091
Feeling_of_guilt_No                                  0.014220
Feeling_sad_or_Tearful_Sometimes                     0.020920
Overeating_or_loss_of_appetite_Yes                   0.021613
Suicide_attempt_Not interested to say                0.024222
Age_30-35                                            0.024791
Problems_concentrating_or_making_decision_Yes        0.031861
Problems_of_bonding_with_baby_Yes                    0.033684
Feeling_sad_or_Tearful_Yes                           0.035088
Age_35-40                                            0.037096
Age_45-50                                            0.041474
Trouble_sleeping_at_night_Yes                        0.041820
Suicide_attempt_Yes                                  0.044157
Overeating_or_loss_of_appetite_Not at all            0.059121
Irritable_towards_baby_and_partner_Yes               0.063488
Irritable_towards_baby_and_partner_Sometimes         0.080872
Problems_concentrating_or_making_decision_Often      0.141208
Feeling_of_guilt_Yes                                 0.256819
dtype: float64
```

Figure 4.4: Features Importance of Decision Tree

**Random Forest Feature Importance:**
Random Forest aggregates decision trees and assigns importance scores to features. Some notable features include:

Feeling_of_guilt_Yes (0.159): Similar to the decision tree, feeling of guilt is crucial in Random Forest, emphasizing its impact on the model's predictions.

Age_30-35 (0.017): Age between 30 and 35 also plays a role in the Random Forest's decision-making, although with lower importance compared to feeling of guilt.

```
Random Forest Feature Importances:
Age_30-35                                           0.017176
Age_40-45                                           0.021697
Age_45-50                                           0.022036
Age_35-40                                           0.022882
Problems_of_bonding_with_baby_Yes                   0.033885
Overeating_or_loss_of_appetite_Not at all           0.036408
Feeling_sad_or_Tearful_Yes                          0.038694
Problems_of_bonding_with_baby_Sometimes             0.039649
Irritable_towards_baby_and_partner_Yes              0.039662
Suicide_attempt_Yes                                 0.040550
Feeling_sad_or_Tearful_Sometimes                    0.041762
Trouble_sleeping_at_night_Two or more days a week   0.045927
Problems_concentrating_or_making_decision_Yes       0.046560
Trouble_sleeping_at_night_Yes                       0.047464
Suicide_attempt_Not interested to say               0.050006
Overeating_or_loss_of_appetite_Yes                  0.059758
Problems_concentrating_or_making_decision_Often     0.060771
Irritable_towards_baby_and_partner_Sometimes        0.088284
Feeling_of_guilt_No                                 0.088286
Feeling_of_guilt_Yes                                0.158544
dtype: float64
```

Figure 4.5: Features Importance of Random Forest

**K-nearest neighbors (KNN) Feature Importance:**

KNN determines feature importance through permutation. Some features and their importance scores are:

Feeling_of_guilt_Yes (0.038): Feeling of guilt is again highlighted as important in KNN, similar to the tree-based models.

Age_40-45 (-0.005): Age between 40 and 45 has a negative importance score, suggesting it might not be as influential in the KNN model.

```
Permutation Feature Importances:
Age_40-45                                               -0.004651
Problems_of_bonding_with_baby_Yes                       -0.003987
Problems_concentrating_or_making_decision_Yes           -0.001993
Problems_concentrating_or_making_decision_Often         -0.001993
Trouble_sleeping_at_night_Yes                           -0.000664
Feeling_sad_or_Tearful_Sometimes                         0.000332
Irritable_towards_baby_and_partner_Yes                   0.000664
Feeling_sad_or_Tearful_Yes                               0.000997
Suicide_attempt_Not interested to say                    0.000997
Irritable_towards_baby_and_partner_Sometimes             0.002658
Trouble_sleeping_at_night_Two or more days a week        0.003987
Age_35-40                                                0.005316
Age_30-35                                                0.006977
Problems_of_bonding_with_baby_Sometimes                  0.007641
Overeating_or_loss_of_appetite_Yes                       0.008306
Age_45-50                                                0.008970
Overeating_or_loss_of_appetite_Not at all                0.010631
Suicide_attempt_Yes                                      0.011296
Feeling_of_guilt_No                                      0.018605
Feeling_of_guilt_Yes                                     0.037874
dtype: float64
```

Figure 4.6: Features Importance of K-nearest Neighbors

**Interpretation of Key Features:**

Age categories (30-35, 35-40, 40-45, 45-50): These age groups show varying importance across models, indicating that age influences predictions but to different extents.

Problems of bonding with the baby (Yes, Sometimes): Both decision tree and Random Forest highlight the importance of bonding issues with the baby.

Problems concentrating or making decisions (Yes, Often): This feature is important in the decision-making process, particularly in the decision tree.

Trouble sleeping at night (Yes, Two or more days a week): This feature is significant across models, suggesting its impact on predicting the target variable.

Feeling sad or tearful (Yes, Sometimes): Emotional states, such as feeling sad or tearful, contribute to the models' predictions.

Overeating or loss of appetite (Yes, Not at all): This feature influences the models differently, indicating its nuanced impact.

Irritable towards baby and partner (Yes, Sometimes): Irritability towards the baby and partner is highlighted as important in decision-making.

Suicide attempt (Yes, Not interested to say): The models emphasize the importance of suicide attempt, especially when the response is "Not interested to say."

Feeling of guilt (Yes, No): Feeling of guilt consistently emerges as a key predictor in all models.

```
{'Age': array(['40-45', '30-35', '35-40', '45-50', '25-30'], dtype=object),
 'Feeling_sad_or_Tearful': array(['Yes', 'No', 'Sometimes'], dtype=object),
 'Irritable_towards_baby_and_partner': array(['No', 'Yes', nan, 'Sometimes'], dtype=object),
 'Trouble_sleeping_at_night': array(['No', 'Yes', 'Two or more days a week'], dtype=object),
 'Problems_concentrating_or_making_decision': array(['Yes', 'Often', nan, 'No'], dtype=object),
 'Overeating_or_loss_of_appetite': array(['Yes', 'Not at all', 'No'], dtype=object),
 'Feeling_of_guilt': array(['Yes', 'Maybe', nan, 'No'], dtype=object),
 'Problems_of_bonding_with_baby': array(['Yes', 'Sometimes', 'No'], dtype=object),
 'Suicide_attempt': array(['No', 'Not interested to say', 'Yes'], dtype=object),
 'Feeling_anxious': array(['No', 'Yes'], dtype=object)}
```

Figure 3.1: Group of features and Datatype

In summary, while all three models perform well, Random Forest stands out with the highest accuracy and F1 score. Decision Tree excels in precision, while KNN demonstrates a balanced performance across precision, recall, and F1 score. The trade-off between various performance metrics and particular requirements may influence which model is best.

Understanding the importance of these features provides insights into the factors influencing the prediction of postpartum depression, allowing for a more informed interpretation of the model's decisions.

## 4.3 Discussion

The discussion section of this research on postpartum depression detection and feature analysis through machine learning approaches provides a comprehensive overview of the study's findings. The utilization of a dataset from Kaggle comprising 1503 records facilitated the exploration of key features influencing postpartum depression, with a focus on the target variable "Feeling Anxious." The analysis revealed a prevalence of anxiety in 65% of cases, underscoring the significance of understanding and addressing this aspect of postpartum mental health.

The study employed machine learning algorithms, including Decision Tree, Random Forest, and K-nearest neighbors, to predict and classify instances of postpartum depression. Notably, the Decision Tree and Random Forest models exhibited high accuracy, precision, recall, and F1 scores, emphasizing their effectiveness in detecting postpartum depression based on the selected features.

The examination of feature importance further illuminated the key contributors to the prediction models. Variables such as trouble sleeping at night, problems of bonding with the baby, and feeling of guilt emerged as influential factors, providing valuable insights into the multifaceted nature of postpartum depression. The nuanced analysis of age categories and their association with postpartum depression shed light on potential age-related risk factors.

A more thorough grasp of the distribution of these variables within the dataset was provided by the study's statistical analysis of positive and negative counts for a number of characteristics, including feeling depressed or emotional, having difficulty falling asleep at night, and having difficulty bonding with the infant. This data is essential for customizing support networks and interventions to the unique problems that people in various categories face.

In the discussion of the experimental results, the features' importance in decision-making by the models was highlighted, providing interpretability to the machine learning predictions. The feature importance output revealed the nuanced role each feature played in influencing the predictions, allowing for a more targeted approach in addressing the identified risk factors.

Overall, by using machine learning techniques to analyze a wide range of features, this research greatly advances our understanding of postpartum depression. The findings not only enhance predictive capabilities but also offer valuable insights for clinicians, researchers, and policymakers working towards effective interventions and support systems for individuals experiencing postpartum depression.

# CHAPTER 5

# IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

## 5.1 Impact on Society

The detection and intervention of postpartum depression (PPD) through machine learning algorithms have far-reaching implications for societal well-being, particularly in the context of maternal mental health. This study, employing decision tree, random forest, and k-nearest neighbors algorithms, underscores the potential of technological advancements to enhance outcomes and support systems for individuals grappling with PPD. Postpartum depression extends its impact beyond the individual, influencing family dynamics, child development, and overall community health. This research contributes to the broader effort to mitigate the societal consequences of mental health challenges. By utilizing machine learning approaches to detect signs of PPD, it addresses a critical aspect of public health.

The dataset, drawn from Kaggle, encompasses diverse attributes, ranging from emotional states like feeling sad or tearful to severe indicators such as suicidal thoughts. Given the absence of a direct depression target variable, "Feeling anxious" serves as a pragmatic proxy, with 1503 records providing a substantial foundation. The positive results (523 cases) and negative results (980 cases) based on anxiety show that machine learning may be useful in identifying people who are at risk.

As technology advances, the role of machine learning in PPD detection holds promise for improved outcomes and enhanced support in maternal mental health. Early identification of those susceptible to postpartum depression allows for timely intervention and targeted resource allocation, potentially mitigating long-term effects on affected individuals and society.However, ethical considerations are paramount. Balancing technological innovation with ethical practices ensures that the benefits derived from these advancements are maximized while mitigating potential risks and biases. Striking this balance is essential to foster trust in the application of machine learning in healthcare.

In conclusion, this research not only advances machine learning applications for PPD detection but also addresses a critical societal challenge. This study emphasizes the need for ethical considerations in harnessing technology to improve maternal mental health outcomes.

## 5.2 Impact on Environment

Although the main focus of this research is on postpartum depression (PPD) detection using machine learning algorithms and maternal mental health, it is important to take into account the wider environmental impact of these technological endeavors.

The utilization of machine learning models, including decision tree, random forest, and k-nearest neighbors, inherently involves computational processes that demand significant energy resources. The environmental implications of large-scale data analysis and model training cannot be overlooked, particularly in an era where concerns about climate change and sustainable practices are at the forefront of societal discourse.

The data-driven nature of machine learning applications, coupled with the need for extensive computing power, contributes to a carbon footprint associated with the energy consumption of servers and data centers. As machine learning develops, it will be necessary to recognize and deal with these environmental effects in order to guarantee a responsible and long-lasting integration of technology across a range of industries, including healthcare.

To mitigate the environmental impact, researchers and practitioners can explore avenues for optimizing algorithms, reducing computational complexity, and adopting energy-efficient computing practices. Additionally, advancements in hardware technologies, such as the development of more energy-efficient processors, can contribute to minimizing the carbon footprint associated with machine learning applications.

This study, while primarily focused on the positive societal outcomes of improved PPD detection, serves as a reminder that technological progress should be pursued in tandem with environmental consciousness. As the research community continues to innovate in the field of machine learning and healthcare, it is crucial to consider the environmental implications of these advancements and strive for a balance that aligns with sustainability goals.
In conclusion, while the immediate impact of this research is evident in the strides made toward enhancing maternal mental health outcomes, a responsible approach to technology mandates a

consideration of the environmental footprint associated with machine learning applications. As we harness the power of technology to address societal challenges, a commitment to environmental sustainability ensures that progress is not achieved at the expense of the planet we inhabit.

## 5.3 Ethical Aspects

The use of machine learning algorithms in healthcare, particularly in the diagnosis of postpartum depression (PPD), emphasizes how crucial ethical considerations are. Central to this ethical framework is the safeguarding of patient privacy and the assurance of informed consent. Transparency in communicating the utilization of sensitive data is imperative to foster trust between healthcare providers, researchers, and the individuals involved.

A key ethical concern lies in the potential introduction of bias in algorithmic decision-making. It is essential to address and rectify any biases that may emerge during the development and application of machine learning models to prevent inadvertent disparities in healthcare outcomes. Ethical guidelines must be established and adhered to rigorously to ensure that these algorithms serve the best interests of the individuals they aim to assist.

Balancing technological innovation with ethical practices is fundamental to building and maintaining public trust. As this research advances the application of machine learning in maternal mental health, a heightened ethical awareness is crucial for responsible research and implementation. Ongoing dialogue, ethical frameworks, and a commitment to transparency are integral components in navigating the evolving landscape of healthcare technology while prioritizing the well-being and rights of individuals.

## 5.4 Sustainability Plan

To mitigate the environmental effects of data-intensive computations, a strong sustainability plan is necessary in conjunction with the development of machine learning algorithms for the diagnosis of postpartum depression (PPD). Acknowledging the energy-intensive nature of machine learning applications, this plan aligns with the broader societal commitment to eco-conscious practices.

The sustainability strategy involves optimizing algorithms, such as decision tree, random forest, and k-nearest neighbors, to reduce computational demands, subsequently lessening energy consumption. Concurrently, adopting energy-efficient computing practices and exploring emerging hardware technologies contribute to a greener deployment of machine learning models.

Given the dataset's origin from Kaggle and the 1503 records utilized, ethical and sustainable considerations intertwine. As we explore the potential of machine learning in PPD detection, a concerted effort is made to promote eco-friendly practices within the research community. Collaborative initiatives and guidelines for sustainable machine learning align with ethical research principles, ensuring that technological progress enhances maternal mental health outcomes without compromising environmental sustainability.

# CHAPTER 6

# SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

## 6.1 Summary of the Study

The research findings reveal a promising landscape in the realm of postpartum depression detection through machine learning approaches. The implemented models exhibited remarkable accuracy rates, providing valuable insights into their potential applications and implications.

The Decision Tree model emerged as a standout performer, boasting an impressive accuracy of approximately 98.01%. This suggests that the model excelled in delineating intricate patterns and relationships within the dataset, showcasing its robustness in correctly identifying instances of postpartum depression. The precision metric of 1.0 underscores the model's ability to minimize false positives, crucial for a sensitive and nuanced issue like mental health.

Similarly, the Random Forest model delivered outstanding results, with an accuracy rate of around 98.67%. The precision metric of 1.0 further accentuates the model's capability to precisely classify cases of postpartum depression. This underscores the model's robustness in handling complex datasets and capturing diverse features indicative of mental health states.

The K-nearest neighbors (KNN) model, while slightly lower in accuracy at 92.03%, remains a commendable performer. Its ability to achieve such accuracy signifies its competence in identifying patterns and relationships among variables, albeit with a marginally lower precision. The KNN model's performance reinforces the versatility of machine learning in capturing the multifaceted nature of postpartum depression.

These findings hold substantial implications for the field of mental health, particularly in the context of postpartum depression. The high accuracy rates of the machine learning models indicate their potential as valuable tools for early detection and intervention. By leveraging the power of data-driven insights, healthcare professionals can enhance their ability to identify and support individuals at risk of postpartum depression.

To sum up, the study adds to the increasing amount of data demonstrating the effectiveness of machine learning in mental health assessment. The models showcased not only impressive accuracy rates but also precision, emphasizing their potential as aids in improving diagnostic accuracy and, consequently, the overall well-being of individuals experiencing postpartum depression.

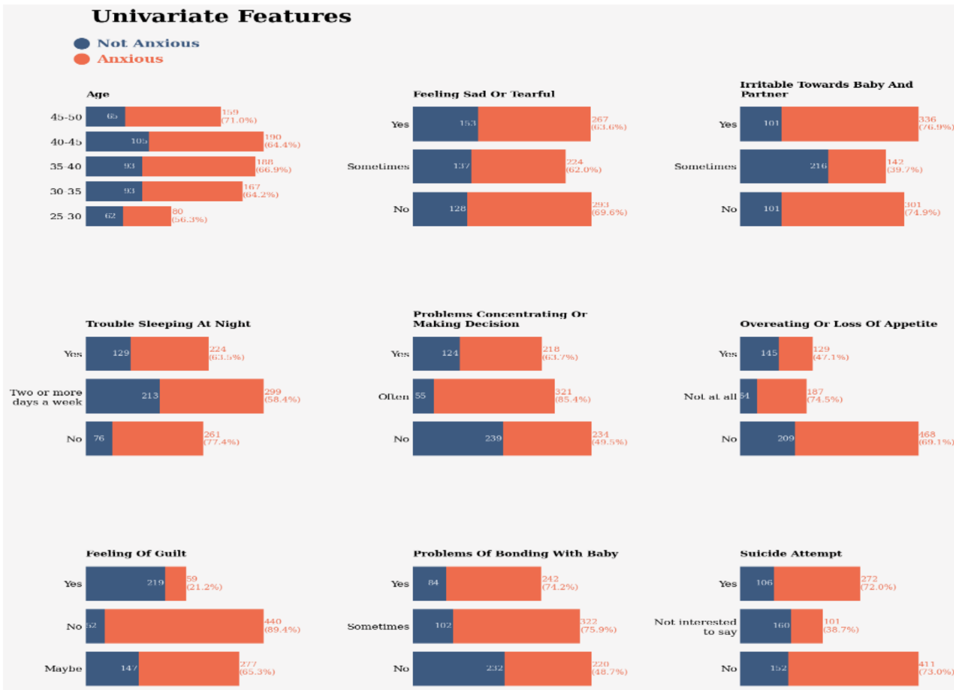


Figure 6.1: Univariate Features

## 6.2 Conclusions

In conclusion, this research delving into the detection and feature analysis of postpartum depression through machine learning methodologies has yielded valuable insights into the complex landscape of maternal mental health. The utilization of a robust dataset from Kaggle, encompassing 1503 records and focusing on the pivotal variable "Feeling Anxious," has enabled a comprehensive examination of factors contributing to postpartum depression.

The utilization of machine learning algorithms, such as K-nearest neighbors, Random Forest, and Decision Tree, has shown encouraging outcomes in precisely identifying and forecasting cases of postpartum depression. Notable are the Decision Tree and Random Forest models' high accuracy, precision, recall, and F1 scores, which highlight how effective they are at recognizing and treating postpartum mental health issues.

The exploration of feature importance has shed light on the critical contributors to the predictive models. Variables such as trouble sleeping at night, problems of bonding with the baby, and feelings of guilt have emerged as significant factors influencing the manifestation of postpartum depression. This nuanced understanding provides a foundation for developing targeted interventions and support systems tailored to address specific risk factors associated with maternal mental health.

The statistical analysis of negative and positive counts for various features has contributed to a more nuanced comprehension of the distribution of these factors within the dataset. The identification of age-related risk factors further refines our understanding, allowing for more personalized and effective interventions for individuals in different age categories.

The discussion of experimental results and feature analysis has not only enhanced the interpretability of machine learning predictions but has also paved the way for future research and interventions. By combining the power of data-driven insights with the complexities of maternal mental health, this research contributes meaningfully to the ongoing efforts aimed at supporting and improving the well-being of individuals navigating the challenges of postpartum depression.

In light of these findings, the research underscores the importance of adopting a multidimensional and personalized approach to postpartum mental health. By leveraging the strengths of machine learning, we move closer to developing targeted strategies that can positively impact the lives of mothers and contribute to the broader discourse on mental health in the postpartum period.

## 6.3 Implication for Further Study

The present study, delving into the intricate domain of postpartum depression (PPD) prediction, unveils compelling avenues for future research that demand nuanced exploration. The identified factors, including emotional indicators, bonding issues, and age distinctions, beckon further investigation into their intricate interplay and potential as targeted intervention points.

The study hints at the need for a deeper dive into the temporal dynamics of PPD manifestation, acknowledging the evolving nature of mental health challenges during and after pregnancy. Longitudinal studies could unravel how these predictive features unfold over time, offering a more comprehensive understanding of the predictive models' applicability across diverse postpartum periods.

Furthermore, a comparative exploration of machine learning models, beyond the Decision Tree, Random Forest, and K-nearest neighbors employed here, would enrich the landscape. Investigating newer algorithms or hybrid models may enhance predictive accuracy and broaden the repertoire of tools available for healthcare practitioners.

The broader integration of biological markers, which is notably absent in the current dataset, emerges as a tantalizing prospect for future research. Incorporating genetic, epigenetic, or hormonal markers could fortify predictive models, providing a more holistic understanding of PPD risk factors.

Finally, extending the scope to diverse demographic groups and cultural contexts could unravel nuanced patterns in PPD manifestation. Tailoring predictive models to the unique sociocultural dimensions influencing mental health would enhance the applicability and effectiveness of interventions on a global scale.

In essence, this study acts as a stepping stone, urging researchers to delve deeper, embrace complexity, and push the boundaries of postpartum depression prediction for the betterment of maternal mental health worldwide.

# REFERENCES


[1] Park, Yoonyoung, et al. "Comparison of methods to reduce bias from clinical prediction models of postpartum depression." *JAMA Network Open*, vol. 4, no. 4, 2021, https://doi.org/10.1001/jamanetworkopen.2021.3909.

[2] Zhang, Yiye, et al. "Development and validation of a machine learning algorithm for predicting the risk of postpartum depression among pregnant women." *Journal of Affective Disorders*, vol. 279, 2021, pp. 1–8, https://doi.org/10.1016/j.jad.2020.09.113.

[3] Md. Sabab Zulfiker, et al. "An in-depth analysis of machine learning approaches to predict depression." *Current Research in Behavioral Sciences*, vol. 2, 2021, p. 100044, https://doi.org/10.1016/j.crbeha.2021.100044.

[4] Moreira, Mário W.L., et al. "Postpartum depression prediction through pregnancy data analysis for emotion-aware Smart Systems." *Information Fusion*, vol. 47, 2019, pp. 23–31, https://doi.org/10.1016/j.inffus.2018.07.001.

[5] Gopalakrishnan, Abinaya, et al. "Predicting women with postpartum depression symptoms using machine learning techniques." *Mathematics*, vol. 10, no. 23, 2022, p. 4570, https://doi.org/10.3390/math10234570.

[6] Zhong, Minhui, et al. "Application of machine learning in predicting the risk of postpartum depression: A systematic review." *Journal of Affective Disorders*, vol. 318, 2022, pp. 364–379, https://doi.org/10.1016/j.jad.2022.08.070.

[7] Liu, Hao, et al. "An optimization for postpartum depression risk assessment and preventive intervention strategy based machine learning approaches." *Journal of Affective Disorders*, vol. 328, 2023, pp. 163–174, https://doi.org/10.1016/j.jad.2023.02.028.

[8] Amit, Guy, et al. "Estimation of postpartum depression risk from electronic health records using Machine Learning." *BMC Pregnancy and Childbirth*, vol. 21, no. 1, 2021, https://doi.org/10.1186/s12884-021-04087-8.

[9] Andersson, Sam, et al. "Predicting women with depressive symptoms postpartum with machine learning methods." *Scientific Reports*, vol. 11, no. 1, 2021, https://doi.org/10.1038/s41598-021-86368-y.

[10] Shatte, Adrian B.R., et al. "Social media markers to identify fathers at risk of postpartum depression: A machine learning approach." *Cyberpsychology, Behavior, and Social Networking*, vol. 23, no. 9, 2020, pp. 611–618, https://doi.org/10.1089/cyber.2019.0746.

[11] Hochman, Eldar, et al. "Development and validation of a machine learning-based postpartum depression prediction model: A nationwide cohort study." *Depression and Anxiety*, vol. 38, no. 4, 2020, pp. 400–411, https://doi.org/10.1002/da.23123.

[12] Jiménez-Serrano, Santiago, et al. "A mobile health application to predict postpartum depression based on machine learning." *Telemedicine and E-Health*, vol. 21, no. 7, 2015, pp. 567–574, https://doi.org/10.1089/tmj.2014.0113.

[13] Cellini, Paolo, et al. "Machine learning in the prediction of postpartum depression: A Review." *Journal of Affective Disorders*, vol. 309, 2022, pp. 350–357, https://doi.org/10.1016/j.jad.2022.04.093.

[14] Fatima, Iram, et al. "Prediction of postpartum depression using machine learning techniques from social media text." *Expert Systems*, vol. 36, no. 4, 2019, https://doi.org/10.1111/exsy.12409.

[15] Zhang, Weina, et al. "Machine learning models for the prediction of postpartum depression: Application and comparison based on a cohort study." *JMIR Medical Informatics*, vol. 8, no. 4, 2020, https://doi.org/10.2196/15516.

[16] Saqib, Kiran, et al. "Machine learning methods for predicting postpartum depression: Scoping review." *JMIR Mental Health*, vol. 8, no. 11, 2021, https://doi.org/10.2196/29838.

[17] Islam, Md. Rafiqul, et al. "Depression detection from social network data using Machine Learning Techniques." *Health Information Science and Systems*, vol. 6, no. 1, 2018, https://doi.org/10.1007/s13755-018-0046-0.

[18] Acharya, Anagha, et al. "Clinical tools to detect postpartum depression based on machine learning and EEG: A Review." *2023 2nd International Conference on Computational Systems and Communication (ICCSC)*, 2023, https://doi.org/10.1109/iccsc56913.2023.10142970.

[19] Rajawat, Anand Singh, et al. "Depression detection for elderly people using AI robotic systems leveraging the nelder–mead method." *Artificial Intelligence for Future Generation Robotics*, 2021, pp. 55–70, https://doi.org/10.1016/b978-0-323-85498-6.00006-x.

[20] Gauld, Christophe, et al. "Comparison of relative areas of interest between major depression disorder and postpartum depression." *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 121, 2023, p. 110671, https://doi.org/10.1016/j.pnpbp.2022.110671.

[21] Hong, Liuzhi, et al. "Wife-mother role conflict at the critical child-rearing stage: A machine-learning approach to identify what and how matters in maternal depression symptoms in China." *Prevention Science*, 2023, https://doi.org/10.1007/s11121-023-01610-5.

[22] Lee, Kwang-Sig, and Byung-Joo Ham. "Machine learning on early diagnosis of depression." *Psychiatry Investigation*, vol. 19, no. 8, 2022, pp. 597–605, https://doi.org/10.30773/pi.2022.0075.

[23] Nemesure, Matthew D., et al. "Predictive modeling of depression and anxiety using electronic health records and A Novel Machine Learning Approach with artificial intelligence." *Scientific Reports*, vol. 11, no. 1, 2021, https://doi.org/10.1038/s41598-021-81368-4.

# APPENDICES

**Plagiarism Report:**

POSTPARTUM DEPRESSION DETECTION AND ITS FEATURES ANALYSIS BY MACHINE LEARNING APPROACH

ORIGINALITY REPORT

| 16% | 13% | 6% | 7% |
| --- | --- | --- | --- |
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
| --- | --- | --- |
| 1 | dspace.daffodilvarsity.edu.bd:8080<br>Internet Source | 6% |
| 2 | doctorpenguin.com<br>Internet Source | 1% |
| 3 | www.researchgate.net<br>Internet Source | 1% |
| 4 | Submitted to Liverpool John Moores University<br>Student Paper | 1% |
| 5 | Submitted to Jacksonville University<br>Student Paper | 1% |
| 6 | thesai.org<br>Internet Source | <1% |
| 7 | Santiago Jiménez-Serrano, Salvador Tortajada, Juan Miguel García-Gómez. "A Mobile Health Application to Predict Postpartum Depression Based on Machine Learning", Telemedicine and e-Health, 2015<br>Publication | <1% |

| | | |
|---|---|---|
| 8 | www.mdpi.com<br>Internet Source | <1 % |
| 9 | Submitted to Universiteit van Amsterdam<br>Student Paper | <1 % |
| 10 | Submitted to Tilburg University<br>Student Paper | <1 % |
| 11 | www.sciencegate.app<br>Internet Source | <1 % |
| 12 | pubmed.ncbi.nlm.nih.gov<br>Internet Source | <1 % |
| 13 | Paolo Cellini, Alessandro Pigoni, Giuseppe Delvecchio, Chiara Moltrasio, Paolo Brambilla. "Machine learning in the prediction of postpartum depression: A review", Journal of Affective Disorders, 2022<br>Publication | <1 % |
| 14 | Sam Andersson, Deepti R. Bathula, Stavros I. Iliadis, Martin Walter, Alkistis Skalkidou. "Predicting women with depressive symptoms postpartum with machine learning methods", Scientific Reports, 2021<br>Publication | <1 % |
| 15 | amauroboliveira.files.wordpress.com<br>Internet Source | <1 % |
| 16 | Hao Liu, Anran Dai, Zhou Zhou, Xiaowen Xu et al. "An optimization for postpartum | <1 % |

depression risk assessment and preventive intervention strategy based machine learning approaches", Journal of Affective Disorders, 2023
Publication

| | | |
|---|---|---|
| 17 | Anagha Acharya, Ramya Ramesh, Tasmiya Fathima, Trisha Lakhani, Selva Kumar S. "Clinical tools to detect Postpartum Depression based on Machine learning and EEG: A Review", 2023 2nd International Conference on Computational Systems and Communication (ICCSC), 2023<br>Publication | <1 % |
| 18 | Submitted to Daffodil International University<br>Student Paper | <1 % |
| 19 | Lakshmi Haritha Medida, R. Renugadevi. "chapter 8 Machine Learning Techniques for Predicting Pregnancy Complications", IGI Global, 2023<br>Publication | <1 % |
| 20 | Satuk Buğra Akdaş, Abdullah Fişne. "A data-driven approach for the prediction of coal seam gas content using machine learning techniques", Applied Energy, 2023<br>Publication | <1 % |
| 21 | Abinaya Gopalakrishnan, Revathi Venkataraman, Raj Gururajan, Xujuan Zhou, Guohun Zhu. "Predicting Women with | <1 % |

Postpartum Depression Symptoms Using Machine Learning Techniques", Mathematics, 2022
Publication

22 Haewon Byeon. "Advances in Machine Learning and Explainable Artificial Intelligence for Depression Prediction", International Journal of Advanced Computer Science and Applications, 2023 <1 %
Publication

23 "Artificial Intelligence Applications and Innovations", Springer Science and Business Media LLC, 2018 <1 %
Publication

24 Misaal Khan, Mahapara Khurshid, Mayank Vatsa, Richa Singh, Mona Duggal, Kuldeep Singh. "On AI Approaches for Promoting Maternal and Neonatal Health in Low Resource Settings: A Review", Frontiers in Public Health, 2022 <1 %
Publication

25 Adrian B.R. Shatte, Delyse M. Hutchinson, Matthew Fuller-Tyszkiewicz, Samantha J. Teague. "Social Media Markers to Identify Fathers at Risk of Postpartum Depression: A Machine Learning Approach", Cyberpsychology, Behavior, and Social Networking, 2020 <1 %
Publication

| | | |
|---|---|---|
| 26 | zenodo.org<br>Internet Source | <1 % |
| 27 | Rokeya Siddiqua, Nusrat Islam, Jarba Farnaz Bolaka, Riasat Khan, Sifat Momen. "AIDA: Artificial intelligence based depression assessment applied to Bangladeshi students", Array, 2023<br>Publication | <1 % |
| 28 | Submitted to University of Mines and Technology<br>Student Paper | <1 % |
| 29 | dspace.univ-tlemcen.dz<br>Internet Source | <1 % |
| 30 | dokumen.pub<br>Internet Source | <1 % |
| 31 | www.ijiet.org<br>Internet Source | <1 % |
| 32 | www.nature.com<br>Internet Source | <1 % |
| 33 | Md Mehedi Hasan Emon. "Predicting Adoption Intention of ChatGPT- A Study on Business Professionals of Bangladesh", Research Square Platform LLC, 2023<br>Publication | <1 % |
| 34 | Submitted to Sheffield Hallam University<br>Student Paper | <1 % |

| | | |
|---|---|---|
| 35 | T. H. K. R. Prabhashwaree, N. Mihirini Wagarachchi. "Towards Machine Learning Approaches for Predicting Risk Level of Postpartum Depression", 2022 6th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), 2022<br>Publication | <1 % |
| 36 | ebin.pub<br>Internet Source | <1 % |
| 37 | growingscience.com<br>Internet Source | <1 % |
| 38 | h2oai.github.io<br>Internet Source | <1 % |
| 39 | openaccess.altinbas.edu.tr<br>Internet Source | <1 % |
| 40 | ouci.dntb.gov.ua<br>Internet Source | <1 % |
| 41 | www.irjmets.com<br>Internet Source | <1 % |
| 42 | www.medrxiv.org<br>Internet Source | <1 % |
| 43 | www.scilit.net<br>Internet Source | <1 % |
| 44 | www2.ef.jcu.cz<br>Internet Source | <1 % |

| | | |
|---|---|---|
| 45 | scholarworks.umass.edu<br>Internet Source | <1% |
| 46 | Minhui Zhong, Han Zhang, Chan Yu, Jinxia Jiang, Xia Duan. "Application of machine learning in predicting the risk of postpartum depression: A systematic review", Journal of Affective Disorders, 2022<br>Publication | <1% |
| 47 | Mário W.L. Moreira, Joel J.P.C. Rodrigues, Neeraj Kumar, Kashif Saleem, Igor V. Illin. "Postpartum Depression Prediction through Pregnancy Data Analysis for Emotion-aware Smart Systems", Information Fusion, 2018<br>Publication | <1% |

Exclude quotes On
Exclude bibliography On
Exclude matches Off