mathematics

MDPI

Article

Predicting Women with Postpartum Depression Symptoms Using Machine Learning Techniques

Abinaya Gopalakrishnan [1,2,*], Revathi Venkataraman [1], Raj Gururajan [1,2], Xujuan Zhou [2,*] and Guohun Zhu [3]

[1] Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Chennai 603203, India
[2] School of Business, University of Southern Queensland, Springfield, QLD 4300, Australia
[3] School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane, QLD 4072, Australia
* Correspondence: ag1266@srmist.edu.in (A.G.); xujuan.zhou@usq.edu.au (X.Z.)

Abstract: Being pregnant and giving birth are big life stages that occur for women. The physical and mental effects of pregnancy and childbirth, like those of many other fleeting life experiences, have the significant potential to influence a mother's overall health and well-being. They have also been known to trigger Postpartum Depression (PPD) in many cases. PPD can be exhausting for the mother and it may have a negative impact on her capacity to care for herself and her kid if it is not treated. For this reason, in this study, initially, physiological questionnaire Edinburgh Postnatal Depression Scale (EPDS) data were collected from delivered mothers for one week, the score was evaluated by medical experts, and participants with PDD symptoms were identified. As a part of multistage progress, further, follow-up was carried out by collecting the Patient Health Questionnaire-9 (PHQ-9), Postpartum Depression Screening Scale (PDSS) questionnaires for the above-predicted participants until six weeks. As the second step, correlated risk factors with PPD symptoms were identified using statistical analysis. Finally, data were analyzed and used to train and test machine learning algorithms in order to predict postpartum depression from one to six weeks. The extremely Randomized Trees (XRT) algorithm with (Background Information + PHQ-9 + PDSS) data offers the most accurate and efficient prediction. Pregnant women with these features could be identified and treated properly. Moreover, it reduces prolonged complications and remains cost-effective in future clinical models.

Keywords: postpartum depression (PPD); psychometric questionnaire (EPDS, PDSS, PHQ-9); depression analysis; class imbalance problem; classification algorithms

MSC: 37M10

Citation: Gopalakrishnan, A.; Venkataraman, R.; Gururajan, R.; Zhou, X.; Zhu, G. Predicting Women with Postpartum Depression Symptoms Using Machine Learning Techniques. *Mathematics* **2022**, *10*, 4570. http://doi.org/10.3390/math10234570

Academic Editor: Liangxiao Jiang

Received: 12 October 2022
Accepted: 28 November 2022
Published: 2 December 2022

1. Introduction

Postpartum Depression (PPD) is one of the most frequent consequences of pregnancy, affecting 10–15% of women globally, with greater prevalence in underdeveloped countries [1]. PPD accounts for more than 24% of all postpartum fatalities, making it the most prevalent reason for mother perinatal mortality [2]. PPD symptoms have also been linked to poor mother-baby bonding, newborn physical and cognitive development, language development, infant behaviors, and sleep quality [3]. Childbirth exposes women to mental disease significantly, with postpartum depression problems which are the greatest prevalent cause of postpartum hemorrhage after childbirth [4]. Clinical symptoms of postpartum depression such as difficulty falling or staying asleep, prolonged sleeping, fluctuations in mood, loss of appetite, dread of hurting someone, severe worry about the baby, sadness or excessive crying, feelings of guilt and hopelessness, difficulty concentrating and recollecting, absence of desire in interests and daily activities, suicidal thoughts, and persistent

thoughts of suicide were the notable actions of the delivered mother [5]. PPD and its associated risk factors are the primary focus of the vast majority of studies looking at changes in mothers' behaviors following childbirth. Although this research doesn't concentrate on directly detecting PPD, we believe the material to be beneficial from a methodological perspective. According to one of the studies, only less than half of the mothers who admit to being in a depressed state to it out loud [6]. This shows that up to 50% of PPD cases go unrevealed, maybe because the mothers do not seek much attention for themselves [7]. Given the significant problems with detecting PPD that are well documented, it is believed that a prognostic computing methodology could be particularly helpful for before-time identification [8].

Other risk factors for PPD include Postpartum psychosis, life stress, loneliness, economic condition, maternity blues, and the prevailing tendency of the baby [9,10]. These risk factors are reflected in surveys like the Postpartum Depression Predictors Inventory (PPDI) [11], whose results are based on the context of PPD risk variables [9], and social support has been shown to affect new mothers thoughts, feeling, and actions [12]. Social isolation and psychological stress were recognized by Neilson [13] as important forecasters of PPD. However, proxies for a number of these risk markers (such as socioeconomic status) can be tracked through social media posts. Social support can be inferred from a person's level of connectivity and social engagement on social media, for example; reading the mother's posts about the baby will be used to get a sense of the baby's temperament and other notifiable approaches to the mother's depression after delivery.

Various methodologies to predict the depression symptoms [14] was described in that paper. Among those, this Table 1 provides a summary of the existing approaches carried with the single questionnaire such as PHQ-9, Center for Epidemiological Studies Depression (CESD), BDI, and mostly with self-declaration statements. The analysis of the collected datawas carried out with various classification algorithms Logistic Regression (LR), Support Vector Machine (SVM), Principle Component Analysis (PCA), and Radial Basis Function (RBF) kernel. Among those works the major drawbacks are,

1. The usage of a biased dataset leads to a class imbalance problem.
2. Single physiological questionnaire-based assessment of those clinical characteristics may mislead unnecessary fear and treatments in many cases and it is kept either untold or unnoticed.

In order to address the two major drawbacks of the prior research, this study has formed the following objectives:

- To collect non-clinical data (Edinburgh Postnatal Depression Scale (EPDS), Patient Health Questionnaire-9 (PHQ-9), Postpartum Depression Screening Scale (PDSS) questionnaires) using Kobotoolbox and relating the potential risk variables to predict PPD using statistical analysis;
- To construct robust multistage Postpartum Depression detection models using machine learning approaches to automatically predict PPD based on these collected datasets.
- To show how the machine learning techniques might improve PPD in the early diagnosis.

The rest of the paper is organized as follows. Section 2 elaborates on the survey strategy for data collection and is followed by the statistical analysis to predict the association between risk factors and PPD. Section 3 provides a brief of the workflow with machine learning algorithms to validate the predictions of PPD. Section 4 reports the results and analysis. Section 5 discusses the research finding and limitations and future investigation. Conclusions are given in Section 6. A detailed overview of this study is given in Appendix A Table A1.

Table 1. Outcomes of the existing approaches in which single questionnaire used as well as biased dataset leads to poor performance results.

Population	Cases (Conditions, Base Rate BR)	Survey for Mental Illness	Analysis	Outcome
165	PPD: 28, BR: 17	PHQ-9	User activity and social isolation was predicted using Logistic regression	Pesudo-R2 metric was used with performance of 63
209	Depression: 81, BR: 39	CESD	User activity was predicted with SVM	Accuracy of the classitafier was about 69
250	Suicide attempt: 125, BR: 50	Self declaration	User activity was observed	Precision of 70% was observed
476	Depression: 171, BR: 36	BDI + CESD	Social Network is analyzed using PCA, SVM & RBF kernel	Accuracy of 72% was estimated
378	Depression: 105, BR: 28 PTSD: 63, BR: 17	CESD	Time series data were analyzed with RF	AUC was used as metric with depression: 82%, PTSD: 81%
900	Depression: 326, BR: 36	Self declaration	n-grams method was used for feature prediction	AUC was used as one of the metric with 70%.
1957	Depression: 483, BR: 25 PTSD: 370, BR: 19	Self declaration	Age, gender, personality were considered	logistic regression was used AUC was used as metric with depression: 75%, PTSD: 71%
4026	Anxiety: 2013, BR: 50	Self declaration	n-grams, and LIWC methods were used	Precision for identifying Anxiety: 85
9611	Anxiety: 4820, BR: 50	Self declaration	n-grams methods was used for feature prediction	AUC was used as one of the metric with 76%
1749	Depression: 11,866, BR: 54	Self declaration	Activity of user feature was classified with Log linear classifier	AUC was used as metric with depression: 74%, PTSD: 87%
1749	-	Personality	n-grams, LIWC, topics were used as features predicting methods	Correlation was used as metric to analysis

2. Methods

This section provides detailed information regarding the data collection and statistical analysis carried out to explore the associations between risk factors and PPD symptoms. The demographic information, as well as the EPDS, PDSS, and PHQ-9 screening tests responses collected, are used to identify the PPD risk variables in order to predict postpartum depression in a distributed survey. These screenings were performed for one week to six weeks after delivery. During the initial phase of the screening procedure, scores of questionnaires were calculated to identify emerging cases of depression. As per multistage detection, a woman's PHQ-9 and PDSS questionnaires were assessed in the second stage if her EPDS result is affirmative during the initial assessment stage. This was done in order to reduce the number of women who were misdiagnosed with depression analysis based on the single questionnaire assessment method.

2.1. Ethics Declarations

An Institutional Ethical Committee (IEC) at SRM Medical College and Research Center in Chennai, India has authorized the research for collecting the data. In 2022, an online survey was administrated from mid-April through mid-July. Each participant signed a consent form stating that she had read and understood the study's terms and conditions. Everything was done according to the rules and regulations that applied.

2.2. Survey Design

The detailed survey contains the basic demographic as well as EPDS [15], PHQ-9 and Postpartum Depression Screening Scale (PDSS) [16], questionnaire. The women's demographics data (age, nationality, education, family income, and occupation), information about the infant and delivery experience, such as the child's birth date, if the woman was a new primi mother, whether the doctor or nurse screened the mother with postpartum depression at any point after the birth.

2.2.1. Edinburgh Postnatal Depression Scale (EPDS)

The EPDS [15] was generally used to evaluate depression symptoms 6 weeks postpartum. Higher scores indicate lower maternal mood on a 4-point scale. This is the most common tool for assessing postpartum depression and identifying at-risk women [17,18]. Women having EPDS scores greater than 9 were deemed depressed but the depth level of variations in mood is not observed. Predictors of postpartum depression are highly sensitive, specific, and predictive at this point [19–21]. More than half of all new episodes of depression had a history of mild depression, according to the research [22]. Despite the fact that the EPDS is merely a proxy for a clinical diagnosis, this questionnaire focuses on screening rather than treatment. Consequently, screening practitioners can use the EPDS with confidence as a realistic metric.

2.2.2. Patient Health Questionnaire-9 (PHQ-9)

The Patient Health Questionnaire-9 (PHQ-9) [23] is a screening tool for mood disorders that consists of 9 questions. There are four alternative replies for each item, and they correspond to the frequency with which each symptom has happened over the course of the preceding two weeks. The answers are one of the options below: every day, more than half the days, several days, and not at all. In addition to screening for postpartum depression, the PHQ-9 was created to be used in contexts of general healthcare to identify severe depressive disease. It had a significant amount of use and was validated in these different environments [23,24]. For this reason, it is chosen for this multistage analysis.

2.2.3. Postpartum Depression Screening Scale (PDSS)

Postpartum Depression Screening Scale (PDSS) is a self-rating scale that is conceptually based on a collection of empirical data recorded and then assessed by professionals for predicting the range of severity [16]. The PDSS assessment of patients was based on their symptoms in seven different conceptual categories, including suicidal thoughts, loss of self-esteem, guilt and shame, concern and fragile, emotional incontinence, and intellectual disability. Five distinctive signs that new mothers may experience throughout the days and weeks shortly after giving birth make up each area. Such kind of fine tuned segregation is required for the second stage of analysis of this work.

2.3. Subject Selection

Clinicians identified a prospective subject pool based on the labour risk factor and data collection practicality. Every patient diagnosed and admitted with labour pain will be considered in the order they arrived at the hospital, using a sequential subject selection method.

2.3.1. Inclusion Criteria

These individuals were asked for participation consent based on the following criteria:

- Delivered mothers within age group of 19–35 years.
- Subjects able to read and comprehend the study's details, and mentally capable of completing the consent form.
- Subjects such as all Primigravida mothers and Multigravida mothers irrespective of spontaneous or induced delivery.

2.3.2. Exclusion Criteria

These individuals were excluded from participation consent based on the following criteria:

- Multi-fetal Pregnancy Subjects.
- In Vitro Fertilization (IVF) pregnancy Subjects.
- Bad obstetric History Subjects.
- Young and Elderly Primi mother pregnancy subjects.
- Subjects with a high-risk pregnancy (including preeclampsia, gestational diabetes mellitus, chronic disease, intrauterine growth restriction, known fetal anomalies, or chromosomal aberrations)

2.4. Identification of Risk Factors

The databases including EMBASE, PubMed, PsycINFO, CINAHL, and Medline were used in a literature search to discover risky components associated with the evolution of postpartum depression related features for six weeks of delivery. Women were given a postpartum questionnaire that asked about risk factors that had been regularly recorded and were then categorized as follows: pregnancy-related stressors (pregnancy, obstetric), mother adjustment in socio demographic and biological contexts challenges (see Table 2).

To determine potential sample size, the formula published by Harlow and Lisa [25] was used in this study. where N = 104 + m samples were required to assess (m) independent predictors with a Type I error of 0.05 and a Type II error of 0.20. Our study's sample size was generously larger than necessary because we included 27 separate predictors (104 plus 27 equals 132).

2.5. Statistical Methods

The Student's *t*-test and the Mann-Whitney U test were utilized to ascertain the presence or absence of bivariate relationships among the risk factors and PPD. The Student's *t*-test was used to compare two continuous, regularly distributed variables, while the Mann-Whitney U test was used to compare two continuous, non-normally distributed variables. Several groups of potential risk factors for postpartum depression were analyzed using the x2 test. Logistic regression was utilized when simultaneous consideration of several factors was required. Together with the Odds Ratios (OR) were 95% confidence intervals. In this work, the significance level of 5% was used and a two-tailed p value to establish statistical significance. When comparing the risk index distributions of depressed and non-depressed women, the Mann-Whitney U statistic was utilized [26].

Table 2. Identification of Risk variables to determine connection with depression symptomatology and physiological questionnaires after delivery [27].

Questionnaire	Domain	Risky Components	Quantity
Common Factors	Socio demographic	Age	20–24, 25–29, 30–34, >34
		Education	Graduate, School or less
		Ability to manage with income	Always difficult, Sometimes difficult, Not bad, easy
EPDS	Maternal characteristics	offspring number	multiples/singleton
		marital status	married-in a relationship/Single
		distance from the hospital	within 5 km, more than 5 km
		history of anxiety/depression	Yes or No
		prenatal use of antidepressants	Yes or No
	Infant characteristics	birth weight	4 kg: adequate birth weight, 3–3.9 kg Inadequate/Insufficient birth weight, 2.5–2.9 kg Low birth weight
		Birth Gestational Age (weeks)	extremely preterm (<28), very preterm (28–32), moderate to late preterm (32–37)
PHQ-9	Pregnancy	Postpartum depression history	insufficient birth weight,
		Issues with infertility	2500–2999 g: low birth weight
		Planned conception	no definitely not, not exactly at this time, Yes definitely
		Maternal thoughts on pregnancy	Very pleased, very pleased in some respects but not in others
		Paternal thoughts on pregnancy	Very pleased, very pleased in some respects but not in others
		Obstetrical abnormalities	Yes, No for the following complications: abortion or preterm labour threats, Pre-eclampsia, diabetes, a urinary tract infection, severe nausea, or vomiting
	challenges in life	Stress related workplace	No, Yes, all of the time, sometimes, not at all
		Concerned about going back to work	Yes, sometimes, no
PDSS	Obstetric	parents relationship	not close/no relationship, close
		Induction of labour	Yes, No
		Mode of delivery	Yes, No
	Maternal tolerance	Ready to leave the hospital	Yes, No
		Way of feeding babies	almost exclusive breast-feeding, high breast feeding, partial, bottle feeding, token breast feeding
		Regarding the newborn feeding satisfaction	Very unsatisfied, unsatisfied, satisfied, very satisfied

3. Modeling

This modeling section provides a concise workflow strategy such as data pre-processing, data imputation, and attribute selection for data collected to validate the predictions of PDD symptoms with risk factors using machine learning algorithms. The imbalance typically produces biased results that were resolved using the objective function to cut down on the number of false negatives produced and followed by their interpretation, as well as the experimental conclusions that can be drawn with classification algorithms. Finally, various metrics were was used to evaluate how well each categorization model performed.

3.1. Workflow Strategy

In order to create a final classification model, the study was broken down into smaller components and worked on repeatedly, while being conscious of any biases that might have been introduced. Figure 1. provides an overview of the workflow. After that, the raw data were partitioned into datasets for the BI and EPDS questionnaires, in addition to the numerous psychometric questionnaires, such as the (PHQ-9, and PDSS). In order to build multi-stage predictive models and conduct additional studies on each psychological assessment questionnaire separately, this was done in order to ascertain which ones had the best accuracy for the characterization of PPD.

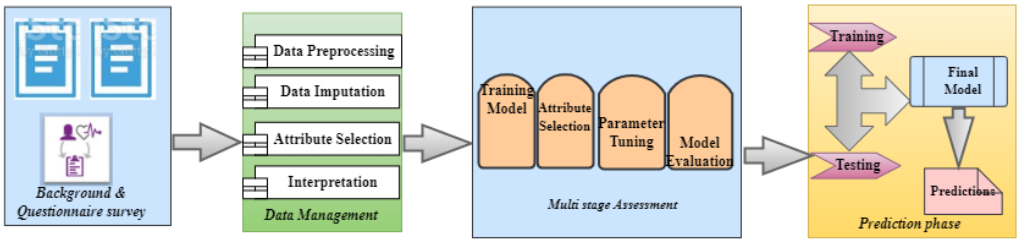


Figure 1. Workflow and analyze strategy of data processing: Data was gathered from the mothers who had given birth. Pregnancy-related variables Background Information (BI) and background data (EPDS, PDSS, PHQ-9) were included in this study. The data were analysed and used to test models or train machine learning algorithms in order to predict postpartum depression within one to six weeks after delivery.

In addition to this, the models constructed were used to determine which psychological assessment questionnaires had the highest accuracy. Further, the Background Information (BI) was linked with the psychometric questionnaires that produced the best levels of accuracy. Predictions were made after the data was compiled (combined dataset). As a result of attribute selection, further models were trained with datasets that were much smaller than the original. Two alternative classification algorithms were trained using the top half and top quarter of Mean Decrease in Impurity (MDI) attributes to ascertain the relative value of each variable in the prediction. Using stratified analyses, participants were also divided into subgroups according to whether or not individuals experienced a history of depressive episodes.

3.2. Data Preprocessing

Preprocessing divided the initial dataset into subsets. For analyzing this research's correctness, Background Information (BI), pregnancy, and EPDS data, as well as psychometric questionnaire data, were kept. Twins and numerous pregnancies were unusual, thus their data were eliminated. Because these populations had a higher risk of PPD, their data were eliminated. Exploratory data analysis was used to confirm the attribute distributions and eliminate uninformative outliers. Psychometric exams and Background Information readings that have supplied information on which women those don't have one week postpartum were omitted. If it had included, the investigation would have used fewer samples. Continuous, nominal, or ordinal variables make up the dataset. The dataset

contains continuous variables with different scales; all variables are standardized to the 0–1 range. To improve machine learning methods, binary numerical representations were utilized to encode nominal and ordinal variables.

3.3. Data Imputation

Missing values can have a significant effect on the performance of machine learning models, a cautious strategy has been chosen to deal with them. First, the samples that had more than fifty percent missing values in the variables that included were removed. After that, the variables (columns, each of which corresponded to a separate variable) that had more than 25% of their data missing were removed as well. In the end, the remaining missing values were computed using an imputation method based on the data that were available. In general Multivariate Imputation by Chained Equations [28] (MICE) approach was used to impute categorical and ordinal features, whereas the Nearest neighbor imputation algorithm [29] was used to impute continuous features. As a result, the total number of pregnancies that were included in the Machine Learning (ML) analyses was 132.

3.4. Attribute Selection

For a machine learning algorithm to be successful, it must be able to generalize and be simply interpreted in addition to being able to make correct predictions. It is beneficial, particularly in the field of medicine, to identify the factors that have a substantial influence on the final result. Gini Index or Mean Decrease in Impurity (MDI) [30] are two methods that can be used to determine the significance of a variable when utilizing Random Forests models. Information that can be utilized to assess the variable's MDI relevance can be obtained by calculating a variable's effectiveness in lowering ambiguity when building decision trees. The variable that is judged to be the most significant is both the most useful and the one that is utilized the most.

3.5. Class Imbalance

Given that a sizable component of the data came from females who reported not having PPD after six weeks after delivery (less than 10% of the individuals comprising PPD symptoms), were included in the population-based sample and clinical scenarios, there is a significant gap between the data classes. Dataset collected was a combination of three questionnaires EPDS, PHQ-9, and PDSS. These were the number of samples which, taken into consideration are shown in Figure 2 for EPDI, PHQ-9, and PDSS respectively among various classes based on score value.

Machine learning classifiers that were taught on datasets with such an imbalance typically produce biased results. A cost in the objective function was included to cut down on the number of false negatives produced by the FNR. A number of studies have shown that using survey data, this strategy is capable of addressing imbalances in relational data sets [31] and accurately forecasting rare diseases [32].

Costs associated with making an incorrect positive or negative prediction are included. To learn to predict Post-Traumatic Stress Disorder (PTSD) using survey data, we use this algorithm. Now obtained a Modified LikeLihood (MLL) by including the new cost function in the likelihood calculation:

$$MLL = \sum_{i} \log \frac{\exp(\psi(\mathbf{x}_i; y_i))}{1 + \exp(\psi(\mathbf{x}_i; y') + \text{cost}(y_i, y'))} \tag{1}$$

The goal function's gradients can be compactly represented as follows:

$$\Delta = I(\hat{y}_i = PPD) - \lambda P(y_i = PPD; \mathbf{x}_i) \tag{2}$$

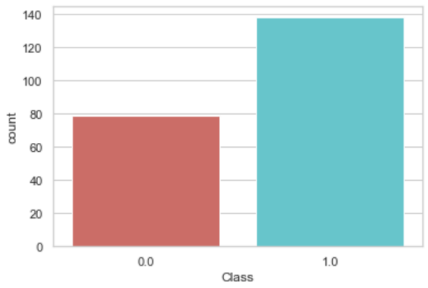
where,

$$\lambda = \frac{e^{c(\hat{y}_i, y=PPD)}}{\sum_{y'} \left[P(y'; \mathbf{x}_i) e^{c(\hat{y}_i, y')} \right]} \tag{3}$$

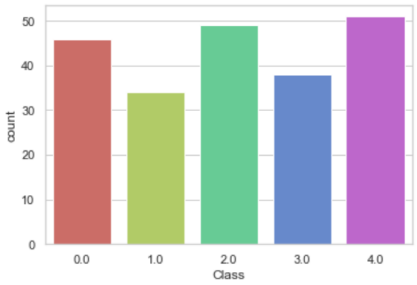
PPD patients (examples) returned the following results:

$$\lambda = \frac{1}{P(y' = PPD; \mathbf{X}_i) + P(y' = \text{Not } PPD; \mathbf{X}_i) \cdot e^{\alpha}} \tag{4}$$

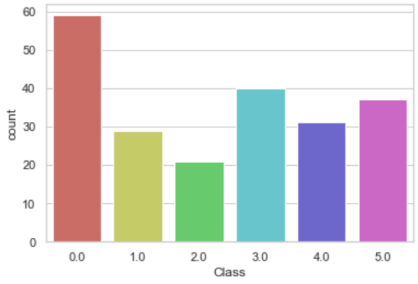
It gets closer to 1 as the gradients gets steeper.$(\Delta \to 1)$ denoting a heavier punishment for incorrectly classified positive examples, $\lambda \to 0$ and the gradients disregard the anticipated probability as, $\mu \to \infty$, which is equivalent to imposing a substantial positive penalty on erroneous negatives. On the other hand, when V $\to -\infty$, gradiants are pushing them closer to 0 $(\Delta \to 0)$, allowing additional flexibility for incorrectly interpreted as negative. By establishing the conditions $\mu > 0$ and $\nu < 0$ the costs of false positive and false negative outcomes into the learning experience, such that the balance between recall and precision may be managed.



(**a**) EPDI questionnaire



(**b**) PHQ-9 questionnaire



(**c**) PDSS questionnaire

Figure 2. Class imbalance representation of collected data items are listed as: number of responses of EPDI, PHQ-9 and PDSS questionnaire shown in (**a**–**c**) respectively.

3.6. Interpretation

Even while gradient-boosting can produce improved results across a number of applications, the interpretation of the trees is a major difficulty. Each consecutive tree learns from the trees that came before it; they "fix" the defects caused by the preceding trees, similar to boosting. The resulting trees can't be understood separately. Because they're all unreliable methods, considering just a few won't effectively describe the model.

Post-pruning removes identical branches by adding their leaf regression numbers. Figure 3 combines two trees. First, two trees were added analytically [33]. And add regression values to each leaf of the first tree, then remove overlapped and redundant branches. Craven's technique explains how neural networks work. Renaming training data allows us to train an infinite number of trees based on the enhanced model [34]. This new

huge tree shows how prior trees made judgments based on training data. First-training data were Boolean. Relabeled data are tree-learned regression values. This research reveals that the created tree is a better PDD vs. non-PDD model. After imputation, the number of balanced samples for EPDI, PHQ-9, and PDSS is displayed in Figure 4.

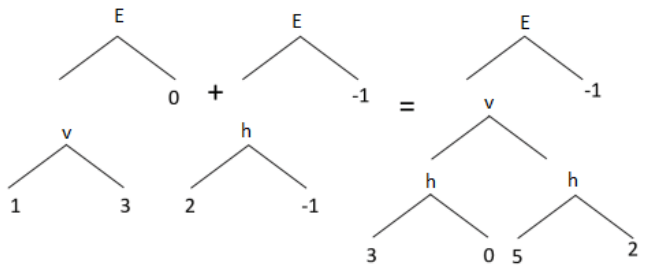
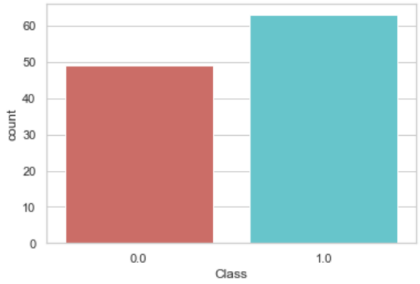
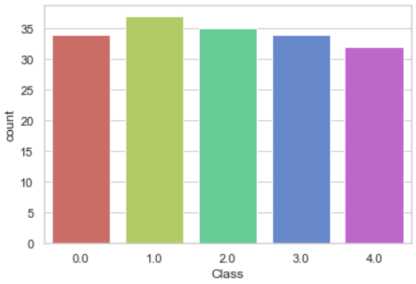


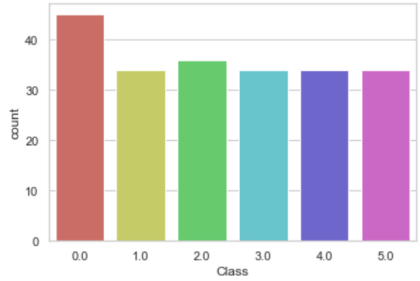
Figure 3. Re-sampling of imbalanced dataset.



(**a**) EPDI questionnaire



(**b**) PHQ-9 questionnaire



(**c**) PDSS questionnaire

Figure 4. Class balance representation of collected data items are listed as: number of responses of EPDI, PHQ-9 and PDSS questionnaire shown in (**a**–**c**) respectively.

3.7. Machine Learning Methods Considered

The primary goal of this study was to predict whether the risk factors associated with the survey questionnaire forecast the prevalence of PPD symptoms. As a result, this study examined a variety of supervised classification algorithms. Nave Bayes, Distributed Random Forests, Extreme Randomized Forest, Ridge Regression, Least Absolute Shrinkage, and Selection Operator (LASSO) Regression, Stacked Ensembles, and Gradient Boosting Machine learning algorithms were deployed in order to offer a full comparison of the techniques.

There are different kinds of linear regression: Ridge Regression and LASSO Regression, a technique for calculating the coefficients of multiple-regression models where the independent features have a high degree of correlation and the later regression method diminution the values of variables into a single cluster point and generates straightforward, skimpy models (i.e., less-parameterized models) respectively. Ensemble learners include

algorithms like Random Forests and Gradient Boosting Machines (GBMs). The Distributed Random Forests (DRF) algorithm selects a selection of attributes and determines the best discriminative thresholds. Instead of creating an ensemble of deep independent trees, GBM specifies a weak and shallow succession of trees, each of which learns from and enhances the previous tree. Lower variance and higher bias by using random thresholds for Extremely Randomized Trees (XRT) rather than the most biased split thresholds.

Like DRF in some respects, but with a larger degree of unpredictability, XRT are also similar. Based on Bayes' Theorem, Naive Bayes (NB) is a probabilistic classifier. According to preliminary data, there is no correlation between the presence of one specific quality and a specific outcome and that feature's existence or absence does not affect the status of any other feature. No matter how interdependent the features are or how necessary other attributes are for an analysis to be successful, NB assumes that each quality contributes independently to the likelihood of the conclusion. By combining the predictions of multiple models, Stacked Ensemble generates a new model When using Stacked Ensembles, the meta-learner learns how to select the best possible mix of base learners. It is necessary for the algorithm to work properly to do this. While bagging and boosting groups together unproductive pupils, learning aims to bring together a diverse group from a variety of backgrounds and experiences [30].

All categorization algorithms were evaluated against the EPDS scores of participants at six weeks postpartum. To calculate this score, a cutoff value of 12 was used as a binary variable, and the above-mentioned Background Information and psychometric data as predictive factors.

3.8. Metrics

A number of different performance indicators were used to assess the accuracy of the model predictions produced by the machine learning classification algorithms. The Confusion Matrix, which served as the foundation for various other measures, was used to evaluate how well each categorization model performed. In addition to the classification accuracy that is most commonly utilized, sensitivity (the rate of true positives) and specificity (the rate of false positives) are also reported. The greater the AUC (which can range from 0 to 1), the better the performance of the classification.

4. Results

The precise description of the investigation results and their analysis were visualized with classification graphs to see how various ML models performed. As well as the assessment's conclusions that were withdrawn Naive Bayes, Distributed Random Forests (DRF), Extreme Randomized Forests (XRT), Ridge Regression, LASSO Regression, Stacked Ensembles models, and Gradient Boosting Machine learning algorithm's performance measurements.

4.1. Descriptive Statistics

Participants who satisfied the inclusion criteria signed an informed consent form, and their stress levels were evaluated in this study. Figure 5 illustrates the process of recruiting participants and collecting their data for the study. 400 people responded to this survey; 138 of the total population are removed because their answers were not complete enough. The remaining responses (a total of 262) served as the basis for the EPDS scale investigation. From this 217 were chosen for further assessments. Approximately 34% of subjects (N = 132) were diagnosed with PPD symptoms. This may be because of various reasons observed such as social background (N = 32), late pregnancy (N = 19), challenges in life (N = 23), and previous medical complications (N = 29).

Table 3 indicates the distribution of study variables among those with depression and those without depression. The features are categorized as sociodemographic, psychopathological, and social support status, as well as prenatal occurrences. All of the study variables' bivariate associations are presented here as their unadjusted odds ratios.

Table 3. Sociodemographic characteristics, Psychopathological status and social support by postpartum outcome. Values are given as %, unless otherwise stated.

Questionnaire	Domain	Risk Factors/ Quantity	Depressed	Non Depressed	Unadjusted Odds Ratio (95% CI)
Common Factors	Socio demographic	**Ability to manage income**			
		Always difficult	3.6%	2.5%	1.5 (0.8–2.9)
		Sometimes difficult	4.4%	3.7%	1.2 (0.6–2.1)
		Not bad	0.7%	0.4%	1.8 (0.4–8.1)
		Easy	6.9%	5.8%	1.2 (0.7–2.0)
		Age			
		20–24	16.4%	12.1%	1.6 (1.1–2.4)
		25–29	32.5%	38.8%	1.0 (reference)
		30–34	34.6%	34.8%	1.2 (0.9–1.6)
		>34	16.4%	14.2%	1.4 (1.0–2.1)
		Education			
		Graduate	94.8%	97.5%	1 (reference)
		School or less	5.2%	2.5%	2.2 (1–2–3.9)
EPDS	Maternal characteristics	**offspring number**			
		multiples	24.3%	20.6%	2.4 (2.4–2.8)
		singleton	75.7%	79.4%	1 (reference)
		Marital status			
		married-in a relationship	15.3%	13.7%	1 (reference)
		single	84.7%	86.3%	0.4 (0.4–1.8)
		History of anxiety/depression			
		Yes	10.6%	12.5%	1.4 (0.8–1.8)
		No	89.4%	87.4	1 (reference)
		Prenatal use of antidepressants			
		Yes	16.2%	23.7%	2.6 (2.6–2.8)
		No	83.7%	76.3%	1 (reference)
	Infant characteristics	**Birth weight**			
		4 kg: adequate birth weight,	15.4%	62.7%	1.8 (1.3–2.4)
		3–3.9 kg Inadequate /Insufficient birth weight	12.3%	2.0%	3.0 (2.5–3.6)
		2.5–2.9 kg Low birth weight	72.3%	35.3%	1 (reference)
		Birth Gestational Age (weeks)			
		extremely preterm (<28)	16.7%	10.2%	4.8 (4.3–5.4)
		very preterm (28–32)	70.4%	65.5%	8.0 (6.5–9.6)
		moderate to late preterm (32–37)	12.9%	24.3%	1 (reference)

Table 3. *Cont.*

Questionnaire	Domain	Risk Factors/ Quantity	Depressed	Non Depressed	Unadjusted Odds Ratio (95% CI)
PHQ-9	Pregnancy	**Postpartum depression history**			
		Yes	83.7%	87.4	1 (reference)
		No	16.2%	12.5%	1.4 (0.8–1.8)
		Issues with infertility			
		Yes	10.6%	8.9%	1.2 (0.8–1.8)
		No	2.5%	1.0%	2.6 (1.2–5.7)
		Planned conception			
		no definitely not	68.1%	62.7%	1.8 (1.3–2.4)
		not exactly at this time	10.6%	2.0%	9.0 (5.5–14.6)
		Yes definitely	21.2%	35.3%	1 (reference)
		Maternal thoughts on pregnancy			
		Very pleased	6.6%	4.6%	1.5 (0.9–2.4)
		very pleased in some respects but not in others	77.7%	83.1%	1 (reference)
		unhappy	13.5%	7.4%	1.9 (14–2.8)
		very unhappy	7.7%	7.4%	0.8 (0.5–1.3)
		Paternal thoughts on pregnancy			
		Very pleased	46.1%	53.1%	1 (reference)
		very pleased in some respects but not in others	37.1%	34.7%	1.2 (0.9–1.6)
		unhappy	6.4%	2.2%	3.6 (2–1–6.2)
		very unhappy	10.4%	10%	0.9 (0.7–1.5)
		Obstetrical abnormalities			
		Yes	83.7%	87.4	1 (reference)
		No	16.2%	12.5%	1.4 (0.8–1.8)
	challenges in life	**Stress-related to workplace**			
		No	3.4%	2.3%	1.3 (0.7–2.6)
		Yes	22.8%	31.3%	0.7 (0.5-0.9)
		all of the time	73.8%	66.4%	1 (reference)
		sometimes	2.1%	1.3%	1.7 (0.7–3.9)
		not at all			
		Concerned about going back to work			
		Yes	10.6%	2.0%	9.0 (5.5–14.6)
		Sometimes	68.1%	62.7%	1.8 (1.3–2.4)
		No	21.2%	35.3%	1 (reference)

Table 3. *Cont.*

Questionnaire	Domain	Risk Factors/ Quantity	Depressed	Non Depressed	Unadjusted Odds Ratio (95% CI)
PDSS	Obstetric	**Parents relationship**			
		not close/no relationship Vs close	83.7%	87.4	1 (reference)
		close	16.2%	12.5%	1.4 (0.8–1.8)
		Induction of labour			
		Yes	99.6%	98.6%	0.3 (0.1–1.9)
		No	0.4%	1.4%	1 (reference)
		Mode of delivery			
		Vaginal	94.8%	97.5%	1 (reference)
		c-section	5.2%	2.5%	2.2 (1–2–3.9)
	Maternal tolerance	**Ready to leave the hospital**			
		Yes	83.7%	87.4	1 (reference)
		No	16.2%	12.5%	1.4 (0.8–1.8)
		way of feeding babies			
		almost exclusive breast-feeding	17.6%	6.0%	3.6 (2.5-5.0)
		high breast-feeding	17.9%	15.3%	1.4 (1.0–2.0)
		partial	64.4%	78.7%	1 (reference)
		bottle-feeding	19.6%	7.3%	3.1 (2–3–4.2)
		token breast-feeding	20.2%	14.0%	1.6 (1.1–2.1)
		Regarding the newborn feeding satisfaction			
		Very unsatisfied	10.6%	2.0%	9.0 (5.5–14.6)
		unsatisfied	68.1%	62.7%	1.8 (1.3–2.4)
		ok	21.2%	35.3%	1 (reference)
		Satisfied	524%	40.2%	1.7 (1.3–2.2)
		Very satisfied	18.6%	18.7%	1.0 (0.7–1.4)

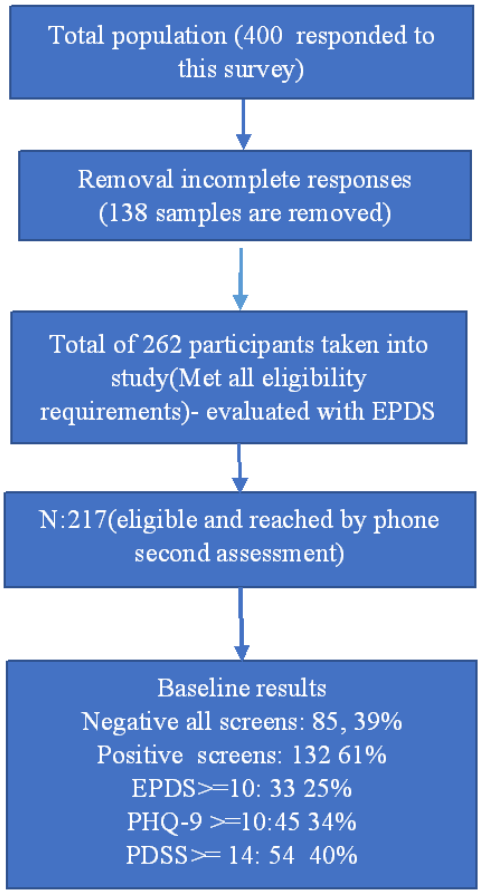


Figure 5. Flow chart illustrates the process of recruiting participants and collecting their data for the study. EPDS, Edinburgh Postnatal Depression Scale; PHQ-9, depression module of Patient Health Questionnaire; PDSS-Short Postpartum Depression Screening Scale.

4.2. Classification Graphs

Data from Background Information and three psychometric questionnaires were integrated into this study to see if machine learning could reliably predict whether female patients would experience depressed symptoms. First, BI data was taken into account in Figure 6 to see how various ML models performed.

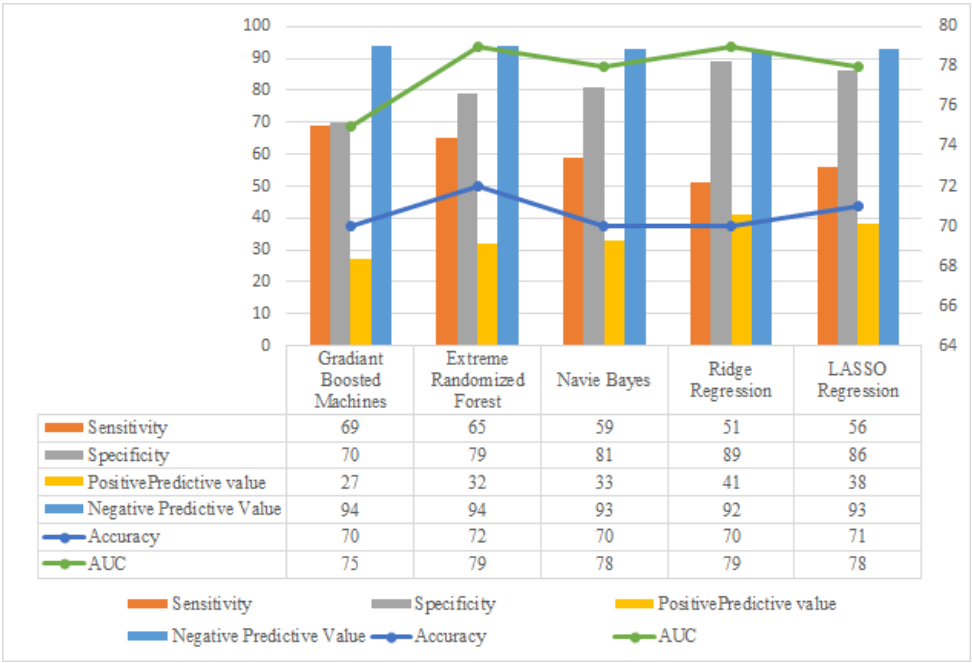
Naive Bayes, Distributed Random Forests (DRF), Extreme Randomized Forests (XRT), Ridge Regression, LASSO Regression, Stacked Ensembles models and Gradient Boosting Machine learning algorithm's performance measurements are presented here. In terms of accuracy, net present value, and area under the curve, XRT was the most accurate at 72% and the most accurate at 79%, respectively. NPV was greater than 92% was noted across the rest of the models. The PPV, specificity, and sensitivity were not all that high, and they varied greatly among models. The sensitivity of the DRF test was the highest, coming in at 84%, whereas the sensitivity of the XRT test was just 65%. However, the DRF test had the lowest specificity and PPV. Ridge Regression and Stacked Ensemble were the ones that demonstrated the highest PPV, coming in at 41%.

Even psychometric considerations were taken into account when evaluating various machine learning models' performance on the pooled dataset (Figure 7). It was found that the accuracy and AUC rates were uniformly high, and tests of performance for the same models showed that the NPV was consistently in excess of 90%. In terms of sensitivity, specificity, and positive predictive value (PPV), there was more heterogeneity among the models. Topping the charts for accuracy (73%) and AUC (81%), XTR also exhibited high sensitivity (72%), specificity (75%), and PPV (33%) as well as the maximum NPV (94%) of all models. Due to the fact that this redressing of the scales is a crucial component of

predictive models that are founded on unbalanced datasets, the subsequent experimental investigation was carried out with XRT alone.

Using the Background Information (BI) dataset and the combined dataset, Figures 8 and 9 compare the XRT model's performance using all variables, using the top 50% of variables, and using the top 25% of variables. There appeared to be a trade-off between the sensitivity of the model and its specificity, both of which were affected by the dataset that was utilized and the percentage of variables that were included (Figure 6). Only the top 25% of the combined dataset was used to obtain the highest levels of sensitivity and specificity, whereas the top 50% of the BI dataset was used to reach the highest levels of sensitivity and specificity. The dataset utilized or the percentage of variables included had no significant effect on any of the other measures (When 25% of the variables were employed, a trend toward reduced PPV was seen).

Figure 8 displays the outcomes of the XRT models' performance following the application of stratification for prior depression. XRT had an area under the curve (AUC) of 81% for all females, a positive predictive value of 33%, a negative predictive value of 94%, a sensitivity of 72%, a specificity of 75%, and an accuracy of 73%. XRT achieved an area under the curve (AUC) of 77%, a positive predictive value of 44%, a negative predictive value of 87%, a sensitivity of 76%, a specificity of 61%, and a balancing accuracy of 69% for women who had experienced depression during pregnancy. There was an area under the curve of 73% for women without a history of depression, and the balanced accuracy was 64%, sensitivity was 52%, specificity was 76%, and the positive predictive value was 13%. (Figure 4). There was not a single questionnaire that obtained an accuracy of greater than seventy percent among the findings that were gleaned from the analysis of the individual forms.



	Gradiant Boosted Machines	Extreme Randomized Forest	Navie Bayes	Ridge Regression	LASSO Regression
Sensitivity	69	65	59	51	56
Specificity	70	79	81	89	86
PositivePredictive value	27	32	33	41	38
Negative Predictive Value	94	94	93	92	93
Accuracy	70	72	70	70	71
AUC	75	79	78	79	78

Figure 6. Analyzing the (n = 132) women with just background, medical, and pregnancy-related factors. Examined models included Extremely Randomized Trees, LASSO Regression, Gradient Boosted Machines, Naive Bayes, and Ridge Regression. Depressive symptoms 6 weeks after delivery were the endpoint, and models were evaluated for accuracy (ACC), sensitivity (SENS), specificity (SPEC), positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC). Bars show performance indicators.

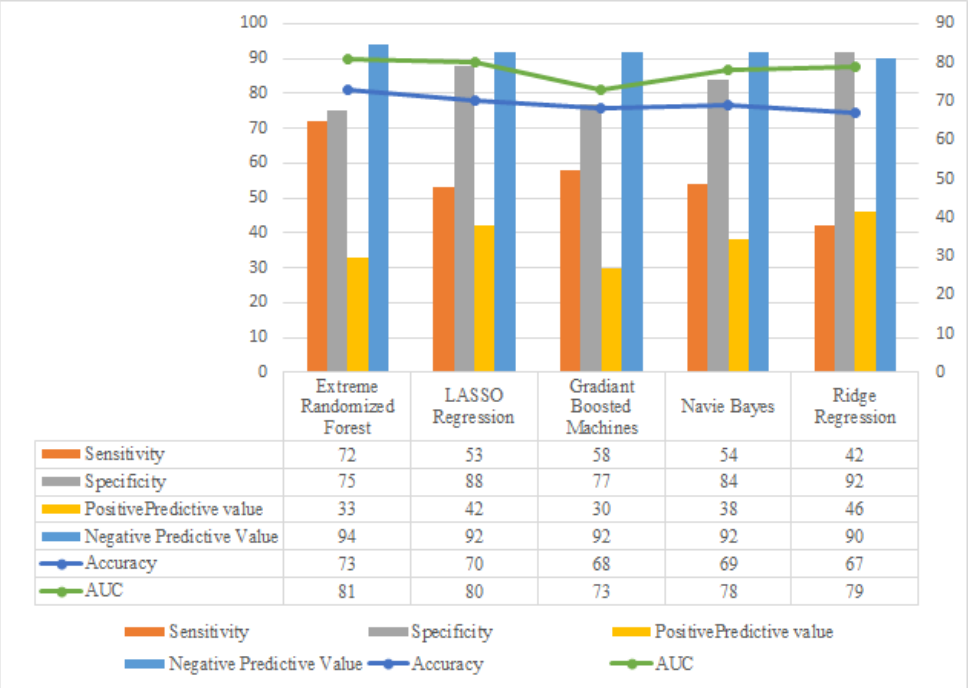


Figure 7. Analyzing the model's overall performance the pooled dataset (n = 132) contained questionnaire responses and background, medical, and pregnancy-related factors. Examined models included Extremely Randomized Trees, LASSO Regression, Gradient Boosted Machines, Naive Bayes, and Ridge Regression. Depressive symptoms 6 weeks after delivery were the endpoint, and models were evaluated for accuracy (ACC), sensitivity (SENS), specificity (SPEC), positive predictive value (PPV), negative predictive value (NPV), and area under the curve (AUC). Bars show performance indicators.

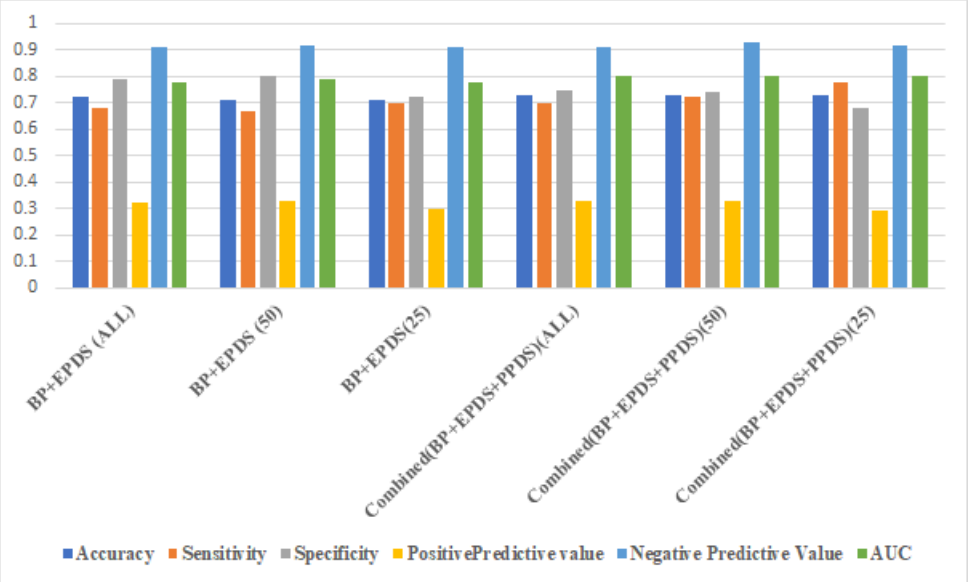


Figure 8. Comparison between datasets that include background information, medical background, and pregnancy factors Background Information (BI) and questionnaire data (BI + EPDS + PDSS). Using XRT, the two datasets' ability to predict postpartum depression was compared. Accuracy, sensitivity, specificity, Predictive Value, Negative Predictive Value, and AUC were calculated for each model.

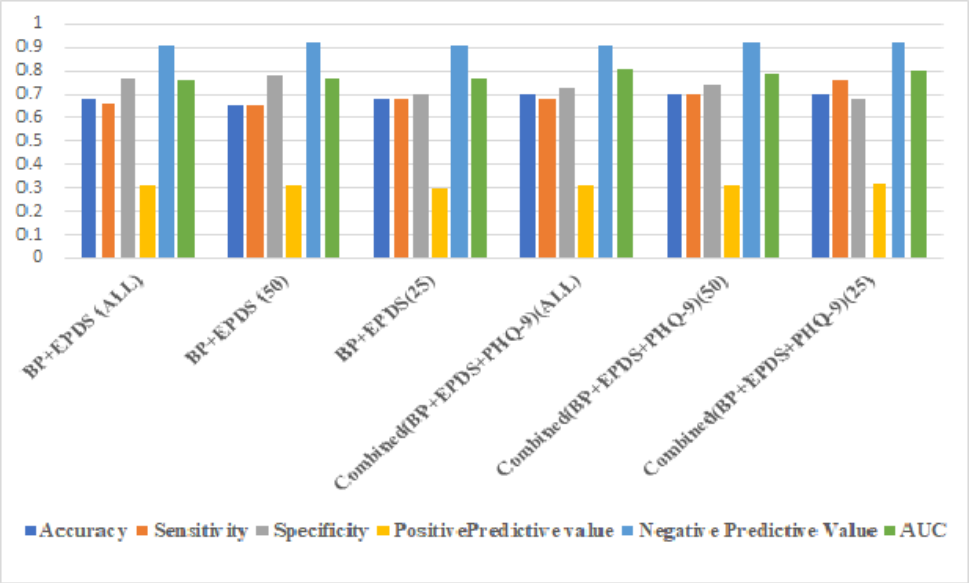


Figure 9. Comparison between datasets that contain background, medical background, and pregnancy factors Background Information (BI) and questionnaire data (BI + EPDS + PHQ-9). The two datasets were compared using the Extremely Randomized Trees (XRT) approach. Each model's accuracy, sensitivity, specificity, Positive predictive value, negative predictive value, and AUC were computed.

5. Discussion

5.1. Research Finding

The goal of this study was to make a multistage predictive model that could find signs of depression between one week and six weeks after giving birth. This would help find high-risk mothers so they could be screened more closely. If the initial step was detected with positive signs, a 6-weeks evaluation was carried out with the help of other questionnaires, and preventive interventions could then be started if necessary. While the study's participants statistical analysis shows those nonclinical data such as socioeconomic and educational backgrounds, biological, life-long stressors, pregnancy-related, obstetric, and maternal adjustment factors had a range of strong association with PPD symptoms. The results arrived in this multistage study data were similar to other postpartum depression prognostic individual models.

- Socioeconomic status was connected with depression symptoms and kept in the predictive model.
- Another risk factor is high blood pressure caused by pregnancy. Several studies suggest that pregnancy-related problems have been associated with postpartum depression similar to this study. A study [35,36] of 1095 women in the United States found that those with serious problems were more likely to have postpartum depression than those without problems.
- Consistent with the current results, previous research of 490 Australian women [35] found obstetric characteristics were related to postpartum depression, including preeclampsia.
- This research includes both elective and emergency cesarean sections. In order to facilitate comparisons, both elective and emergency cesarean sections were included. Those who did not take part in postpartum parenting workshops had a greater chance of developing postpartum depression compared to women who did take part in such classes [37]. Yet, a connection was still shown between both procedures and postpartum depression.

- Partner support influences postpartum depression [9] and is consistent with this study's results. At 1 week postpartum, lack of maternal views of support availability was more indicative in this study of depression.
- With the conclusion that not being ready to leave the hospital was a noticeable risk factor, these data suggest that women need help after giving birth. Astbury et al. say that a mother's lack of trust in the care of her child after she leaves the hospital is a risk factor for postpartum depression [38]. Thus, a readiness-for-discharge assessment should be part of clinical routes for postpartum care.

We also elaborate a variety of machine learning (ML) models to pinpoint freshly delivered mothers who are at risk for postpartum depression (PPD) symptoms were predicted correctly. The accuracy, Negative Predictive Rate (NPV), and AUC of the classification performance of the several ML algorithms under investigation were identical. Differences in specificity, PPV, and sensitivity were the most obvious. As expected, sensitivity and specificity are inversely related. Positive Predictive Rate (PPV) is lower than NPV because PPD prevalence is low. XRT is accurate, with balanced sensitivity and specificity. Self-reported resilience and personality enhance accuracy and AUC. These variables boost XRT's sensitivity but decrease its specificity.

These findings suggest screening new parents as they leave the hospital's delivery ward with machine learning (ML). These approaches can identify a high-risk group, allowing for cost-effective preventative measures, especially by reducing the costs of postpartum depression. These mothers could get more treatment and longer-term follow-ups. Self-reports are employed as screening variables in this study. NPV protects many women who aren't at high risk from postpartum depression. Our categorization algorithms used to target toward high-risk women to improve efficiency and save money. Low-risk women may only be screened during certain times.

Area Under Curve and accuracy remain stable even when models use 100%, 50%, and 25% of all attributes. AUC is stable with 4–8 characteristics. These data support the concept that PPD is influenced by depression or anxiety during pregnancy. In non-depressed patients, only breastfeeding, childhood trauma, birth mode, baby hypoxia, age, and resilience are predictive. This information is crucial for developing screening protocols for women with mental health issues. Variables may need to be changed.The stability of performance measures demonstrates a shorter survey can screen without losing predictability. The lower accuracy in the depressed (during delivery) (n = 132, accuracy = 69%) and never-depressed (n = 85, accuracy = 64%) subgroups may be due to smaller sample sizes and less diversity in the data.

Sensitivity stays the same in depressed women, but it goes down to 52% in women who have never been depressed. This shows how hard it is to find women who are likely to have their first episode of postpartum depression. High NPV implies never-depressed women don't require extra monitoring. In depressive women, the NPV lowers to 86%, showing that postpartum screening may benefit this high-risk population.

Women with a history of depression who are also members of the never before depressed group was also showed in Area Under the Curve (AUC) that was marginally greater than the best prediction models found in most of the previous research (Wang et al. received 79%, while Zhang et al.; received 78%.), ignoring the fact that our accuracy of 73% was markedly smaller than the 84% that was noted by Tortajada et al. [39]. Clinical interviews were also conducted after a lower EPDS cut-off, which may have reduced the risk of wrongly categorizing study participants and controls. But because the study group was so much bigger, it would not have been feasible for us to undertake a medical assessment in our study. Finally, information on related gene mutations was also used in this assessment, in addition to medical and cultural influences that were previously described.

Aside from resiliency, coherence, and personality as a whole, these two traits predict outcomes better than any other single feature. The accuracy for the whole group is 73%, and the area under the curve (AUC) is 81%, which is close to the limit for using these algorithms in clinical settings in the future.

5.2. Limitations and Future Work

A few limitations in predict postpartum depression are:

- Women between the ages of 19 and 32 were investigated in this study. Thus, Women with significant potential of PPD may be missed.
- The study's findings cannot easily be extrapolated to the entire population as a result of these circumstances. One of the causes of selection bias is the omission of women who lack literacy because the surveys require comprehending either the Tamil or the English language. Healthier women are more likely to take part in these types of studies, which is another source of selection bias.
- When necessary, exclusions and imputations were used to fill up the gaps left by missing values. Although it was resolved, the results' class imbalance made algorithm training challenging and with a higher AUC and more predictors [40].

Future studies should assess these predictive factors. In contrast to resilience, coherence, and personality, depression and anxiety during pregnancy are much more likely to lead to postpartum depression. The most precise and effective forecasting algorithm is provided by XRT. The modeling of this approximation's precision will be a fascinating future open problem. Expanding the study to include responses from an ethnicity that match the nationwide levels of various races is a fascinating future study direction.

6. Conclusions

The difficulty in forecasting postpartum depression was considered from outside of a therapy setting by performing an analysis for relating the demographic, behavioral, and socioeconomic data to predicting the PPD symptoms. The research conducted in this area paves the way for the creation of self-observing tools and therapy programs that may be beneficial to women with PPD. In this research, the correlation between the PDD symptoms and risk factors such as sociodemographic, psycho-pathological, social support status, and prenatal occurrences from physiological questionnaires EPDS, PHQ-9, and PDSS were identified using statistical analysis. Further, a multistage evaluation framework was used to predict PDD symptoms and classification algorithms were used to measure the correctness of this framework prediction. This combination of information proved to be useful in predicting PPD using more powerful machine learning techniques in our preliminary research. Furthermore, it was discovered that machine learning methodologies could make a significant contribution to the successful completion of this tough but essential work.

Author Contributions: Conceptualization, A.G.; R.V. and X.Z.; methodology, A.G.; X.Z.; validation, A.G.; R.V.; R.G.; X.Z. and G.Z.; formal analysis, A.G.; investigation, A.G.; R.V.; R.G.; X.Z. and G.Z.; resources, A.G.; R.V.; data curation, A.G.; writing—original draft preparation, A.G.; X.Z. writing—review and editing, A.G.; R.V.; R.G.; X.Z. and G.Z.; supervision, R.V.; R.G.; X.Z.; project administration, A.G.; R.V.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Abinaya Gopalakrishnan, and approved by the Institutional Ethics Committee of SRM Institute of Science and Technology (No: 8376/IEC/2022 and date of approval: 26 May 2022) for studies involving humans.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.Written informed consent was obtained from participating patients who had been identified (included by the patients themselves).

Data Availability Statement: On reasonable request and if data transfer agreements are in place, the corresponding author will provide the datasets created and/or analyzed during the current study available to users.

Acknowledgments: I (Abinaya Gopalakrishnan) would like to acknowledge University of Southern Queensland, SRM Institute of Science Technology and SRM Institute of Science Technology Medical college and Research centre for providing me scholarship and supported with population to collect the data to carried out this research work. I also extend my thanks to the medical experts Arul Saravanan, Department of Psychiatry and Maitrayee Sen, Department of Obstetrics and Gynaecology, SRM Medical College Hospital, Research Centre for supporting me in data collection, analysis and validating the study. Finally, I am willing to thank Premraj, Department of English, Saveetha Engineering College, Thandalam for extensive English revision of our work.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

PPD	PostPartum Depression
PHQ-9	Patient Health Questionnaire-9
EPDS	Edinburgh Postnatal Depression Scale
PDSS	Postpartum Depression Screening Scale
ML	Machine Learning
CESD	Center for Epidemiological Studies Depression
PTSD	Post-Traumatic Stress Disorder
BR	Base Rate
PCA	Principal Component Analysis
SVM	Support Vector Machine
RBF	Radial Basis Function
AUC	Area Under Curve
LIWC	Linguistic Inquiry and Word Count
IEC	Institutional Ethical Committee
BI	Background Information
MDI	Mean Decrease in Impurity
FNR	False Negative Rate
DRF	Distributed Random Forests
XRT	Extremely Randomized Tree
NB	Naive Bayes
LASSO	Least Absolute Shrinkage and Selection Operator
GBM	Gradient Boosting Machine
ROC	Receiver Operating Characteristic
PPV	Positive Predictive Value
NPV	Negative Predictive Value
EHR	Electronic Health Records
CDC	Centers for Disease Control and Prevention
MICE	Multivariate Imputation by Chained Equations
OR	Odds Ratios
IVF	In vitro fertilization
PPDI	Postpartum Depression Predictors Inventory
BDI	Beck Depression Inventory
MLL	Modified LikeLihood

Appendix A. STROBE Statement—Checklist of Items that Should be Included in Reports of Observational Studies

Table A1. STROBE Statement—checklist of items that should be included in reports of observational studies.

	Item No.	Recommendation	Page No.	Relevant Text from Manuscript
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1	
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	1	Line: 1–16
Introduction				
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	1, 2	Line: 17–56
Objectives	3	State specific objectives including any prespecified hypotheses	3	Line: 57–81
Methods				
Study design	4	Present key elements of study design early in the paper	3	Line: 82–132
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	3	Line: 86–132
Participants	6	(a) Cohort study—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up Case-control study—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls Cross-sectional study—Give the eligibility criteria, and the sources and methods of selection of participants	4	Line: 140
		(b) Cohort study—For matched studies, give matching criteria and number of exposed and unexposed Case-control study—For matched studies, give matching criteria and the number of controls per case	-	-
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	-	-
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	4, 5	Line: 163–183
Bias	9	Describe any efforts to address potential sources of bias	-	-
Study size	10	Explain how the study size was arrived at	5	Line: 169–172

Table A1. *Cont.*

	Item No.	Recommendation	Page No.	Relevant Text from Manuscript
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	6	Line : 168
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	6	Line : 173–183
		(b) Describe any methods used to examine subgroups and interactions	6	Line: 173–183
		(c) Explain how missing data were addressed	6, 7	Line: 205–234
		(d) Cohort study—If applicable, explain how loss to follow-up was addressed Case-control study—If applicable, explain how matching of cases and controls was addressed Cross-sectional study—If applicable, describe analytical methods taking account of sampling strategy	8, 9	Line: 236–297
		(e) Describe any sensitivity analyses	11	Line: 333–345
Results				
Participants	13	(a) Report numbers of individuals at each stage of study—e.g.; numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	11	Line: 356–366
		(b) Give reasons for non-participation at each stage	11	Line: 358
		(c) Consider use of a flow diagram	11	Figure 4
Descriptive data	14	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders	11	Line: 356–377
		(b) Indicate number of participants with missing data for each variable of interest	-	-
		(c) Cohort study—Summarise follow-up time (e.g., average and total amount)	-	-
Outcome data	15	Cohort study—Report numbers of outcome events or summary measures over time	13	Table 3
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	-	-
		Cross-sectional study—Report numbers of outcome events or summary measures	-	-

Table A1. *Cont.*

	Item No.	Recommendation	Page No.	Relevant Text from Manuscript
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	12	Line: 369–418
		(b) Report category boundaries when continuous variables were categorized	-	-
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	-	-
Other analyses	17	Report other analyses done—e.g.; analyses of subgroups and interactions, and sensitivity analyses	-	-
Discussion				
Key results	18	Summarise key results with reference to study objectives	17	Line: 420–503
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	18	Line: 509–536
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	19	Line: 539
Generalisability	21	Discuss the generalisability (external validity) of the study results	-	-
Other information				
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	-	-

References

1. Shorey, S.; Chee, C.; Ng, E.; Chan, Y.; San Tam, W.; Chong, Y. Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis. *J. Psychiatr. Res.* **2018**, *104*, 235–248. [CrossRef] [PubMed]
2. Guintivano, J.; Putnam, K.; Sullivan, P.; Meltzer-Brody, S. The international postpartum depression: Action towards causes and treatment (PACT) consortium. *Int. Rev. Psychiatr.* **2019**, *31*, 229–236. [CrossRef] [PubMed]
3. Collins, K.; Onwuegbuzie, A.; Jiao, Q. A mixed methods investigation of mixed methods sampling designs in social and health science research. *J. Mix. Methods Res.* **2007**, *1*, 267–294. [CrossRef]
4. Stocky, A.; Lynch, J. Acute psychiatric disturbance in pregnancy and the puerperium. *Best Pract. Res. Clin. Obstet. Gynaecol.* **2000**, *14*, 73–87. [CrossRef] [PubMed]
5. Patel, M.; Bailey, R.; Jabeen, S.; Ali, S.; Barker, N.; Osiezagha, K. Postpartum depression: A review. *J. Health Care Poor Underserved* **2012**, *23*, 534–542. [CrossRef]
6. Ramsay, R. Postnatal depression. *Lancet* **1993**, *342*, 1358. [CrossRef]
7. MacLennan, A.; Wilson, D.; Taylor, A. The self-reported prevalence of postnatal depression. *Aust. N. Z. J. Obstet. Gynaecol.* **1996**, *36*, 313. [CrossRef]

8. Baagedahl-Strindlund, M.; Börjesson, K. Postnatal depression: A hidden illness. *Acta Psychiatr. Scand.* **1998**, *98*, 272–275. [CrossRef]
9. Beck, C.T. Predictors of postpartum depression: An update. *Nurs. Res.* **2001**, *50*, 275–285. [CrossRef]
10. O'Hara, M. Postpartum depression: What we know. *J. Clin. Psychol.* **2009**, *65*, 1258–1269. [CrossRef]
11. Beck, C. A checklist to identify women at risk for developing postpartum depression. *J. Obstet. Gynecol. Neonatal Nurs.* **1998**, *27*, 39–46. [CrossRef] [PubMed]
12. Fleming, A.; Klein, E.; Corter, C. The effects of a social support group on depression, maternal attitudes and behavior in new mothers. *J. Child Psychol. Psychiatry* **1992**, *33*, 685–698. [CrossRef] [PubMed]
13. Nielsen, D.; Videbech, P.; Hedegaard, M.; Dalby, J.; Secher, N. Postpartum depression: Identification of women at risk. *BJOG Int. J. Obstet. Gynaecol.* **2000**, *107*, 1210–1217. [CrossRef] [PubMed]
14. Gopalakrishnan, A.; Venkataraman, R.; Gururajan, R.; Zhou, X.; Genrich, R.Mobile phone enabled mental health monitoring to enhance diagnosis for severity assessment of behaviours: A review. *PeerJ Comput. Sci.* **2022**, *8*, e1042. [CrossRef] [PubMed]
15. Lee, D.; Yip, A.; Chan, S.; Tsui, M.; Wong, W.; Chung, T. Postdelivery screening for postpartum depression. *Psychosom. Med.* **2003**, *65*, 357–361. [CrossRef] [PubMed]
16. Mancini, F.; Carlson, C.; Albers, L. Use of the Postpartum Depression Screening Scale in a collaborative obstetric practice. *J. Midwifery Womens Health* **2007**, *52*, 429–434. [CrossRef] [PubMed]
17. Cox, J.; Holden, J.; Sagovsky, R. Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *Br. J. Psychiatr.* **1987**, *150*, 782–786. [CrossRef]
18. Teissèdre, F.; Chabrol, H. Detecting women at risk for postnatal depression using the Edinburgh Postnatal Depression Scale at 2 to 3 days postpartum. *Can. J. Psychiatr.* **2004**, *49*, 51–54. [CrossRef]
19. Rogers, C.; Kidokoro, H.; Wallendorf, M.; Inder, T. Identifying mothers of very preterm infants at-risk for postpartum depression and anxiety before discharge. *J. Perinatol.* **2013**, *33*, 171–176. [CrossRef]
20. Lasiuk, G.; Comeau, T.; Newburn-Cook, C. Unexpected: An interpretive description of parental traumas' associated with preterm birth. *BMC Pregnancy Childbirth* 2013, 13, 1–10. [CrossRef]
21. Raines, D. Mothers' stressor as the day of discharge from the NICU approaches. *Adv. Neonatal Care* **2013**, *13*, 181–187. [CrossRef] [PubMed]
22. Latva, R.; Lehtonen, L.; Salmelin, R.; Tamminen, T. Visits by the family to the neonatal intensive care unit. *Acta Paediatr.* **2007**, *96*, 215–220. [CrossRef] [PubMed]
23. Beeghly, M.; Olson, K.; Weinberg, M.; Pierre, S.; Downey, N.; Tronick, E. Prevalence, stability, and socio-demographic correlates of depressive symptoms in Black mothers during the first 18 months postpartum. *Matern. Child Health J.* **2003**, *7*, 157–168. [CrossRef] [PubMed]
24. Spitzer, R.; Kroenke, K.; Williams, J.; Group, P.; Group, P. Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. *JAMA* **1999**, *282*, 1737–1744. [CrossRef] [PubMed]
25. Harlow, L.L. Book review of using multivariate statistics by Barbara G. Tabachnick and Linda S. *Fidell. Struct. Equ. Model.* **2002**, *9*, 621–636. [CrossRef]
26. Campbell G. Advances in statistical methodology for the evaluation of diagnostic and laboratory tests. *Stat. Med.* **1994**, *13*, 499–508. [CrossRef]
27. Dennis, C.-L.E.; Janssen, P.A.; Singer, J. Identifying women at-risk for postpartum depression in the immediate postpartum period. *Acta Psychiatr. Scand.* **2004**, *110*, 338–346. [CrossRef]
28. Azur, M.J.; Stuart, E.A.; Frangakis, C.; Leaf, P.J. Multiple Imputation By Chained Equations: What Is It And How Does It Work. *Int. J. Methods Psychiatr. Res.* **2011**, *20*, 40–49. [CrossRef]
29. Beretta, L.; Santaniello, A. Nearest neighbor imputation algorithms: A critical evaluation. *BMC Med. Inf. Decis. Mak.* **2016**, *16*, 197–208. [CrossRef]
30. Hastie, T.; Tibshirani, R.; Friedman, J.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2009.
31. Yang, S.; Khot, T.; Kersting, K.; Kunapuli, G.; Hauser, K.; Natarajan, S. Learning from imbalanced data in relational domains: A soft margin approach. In Proceedings of the 2014 IEEE International Conference on Data Mining, Shenzhen, China, 14–17 December 2014; pp. 1085–1090.
32. MacLeod, H.; Yang, S.; Oakes, K.; Connelly, K.; Natarajan, S. Identifying rare diseases from behavioural data: A machine learning approach. In Proceedings of the 2016 IEEE First International Conference On Connected Health: Applications, Systems and Engineering Technologies (CHASE), Washington, DC, USA, 27–29 June 2016; pp. 130–139.
33. Craven, M.; Shavlik, J. *Extracting Tree-Structured Representations of Trained Networks, Advances, Neural Information Processing Systems 8*; MIT Press: Cambridge, MA, USA, 1996.
34. Darwiche, A. A differential approach to inference in Bayesian networks. *ACM* **2003**, *50*, 280–305. [CrossRef]
35. Johnstone, S.J.; Boyce, P.M.; Hickey, A.R.; Morris-Yates, A.D.; Harris, M.G. Obstetric risk factors for postnatal depression in urban and rural community samples. *Aust. N. Z. J. Psychiatr.* **2001**, *35*, 69–74. [CrossRef] [PubMed]
36. Burger, J.; Horwitz, S.M.; Forsyth, B.W.C.; Leventhal, J.M.; Leaf, P.J. Psychological sequelae of medical complications during pregnancy. *Pediatrics* **1993**, *91*, 566–571. [CrossRef] [PubMed]

37. Brugha, T.S.; Sharp, H.M.; Cooper, S.-A.; Weisender, C.; Britto, D.; Shinkwin, R.; Sherrif, T.; Kirwan, P.H. The Leicester 500 Project. Social support and the development of postnatal depressive symptoms, a prospective cohort survey. *Psychol. Med.* **1998**, *28*, 63–79. [CrossRef] [PubMed]
38. Astbury, J.; Brown, S.; Lumley, J.; Small, R. Birth events, birth experiences and social differences in postnatal depression. *Aust. J. Public Health* **1994**, *18*, 176–184. [CrossRef]
39. Wang, S.; Pathak, J.; Zhang, Y. Using electronic health records and machine learning to predict postpartum depression. In *MEDINFO 2019: Health and Wellbeing e-Networks for All*; IOS Press: Amsterdam, The Netherlands, 2019; pp. 888–892.
40. Zhang, W.; Liu, H.; Silenzio, V.; Qiu, P.; Gong, W. Machine learning models for the prediction of postpartum depression: Application and comparison based on a cohort study. *JMIR Med. Inf.* **2020**, *8*, e15516. [CrossRef]