DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 19, ISSUE 1, JANUARY 2024
ISSN 1818-5878 (Print) 2408-8498 (Online)

MENTAL DEPRESSION DETECTION OF PREGNANT WOMEN USING MACHINE LEARNING APPROACH

Md. Solaimanur Rahman[1], *Mushfiqur Rahman[1], and Animesh Basak[1]
[1]Department of Computer Science & Engineering, Daffodil International University
Savar, Dhaka-1216, Bangladesh
Email: mushfiqur.cse@diu.edu.bd

Abstract: This work develops machine learning models for depression categorization using a complete range of characteristics from user-generated data. Physiological data, self-reported symptoms, and demographic data. Predictive models are created using XGBoost, Random Forest, Gradient Boosting, MLPClassifier, and AdaBoostClassifier. Data preparation, feature selection, and model optimization rigorously assess the models. Cross-validation ensures resilience and generalization. Learning curves evaluate models' accuracy, training, and validation loss. Machine learning algorithms properly identify depression levels. Random Forest (99%), Gradient Boosting (93%), MLPClassifier (92%), and AdaBoostClassifier (83%). Learning curves show converging training loss, improving accuracy with more iterations, and constant validation performance. This machine learning study advances mental health categorization. The models help identify and treat depression problems early, enabling tailored care. The research emphasizes feature selection and algorithm choice in model performance. In Conclusion, Machine learning can classify depression levels. The models improve mental health treatment with accurate and quick depression assessment decision support systems. Expanding the dataset and adding characteristics may enhance model performance and applicability in future studies.

Keywords: Mental Depression Detection, Depression of Pregnant Women, Machine Learning Prediction, Extreme Gradient Boosting, Random forest.

1. INTRODUCTION

Depression is a common mental illness that can affect people of any age or background [15]. But pregnancy may be a very sensitive time for women, which may make it more likely for them to get depressed because of hormonal changes and other things that may add to the stress they are already feeling. Several studies have found that severe sadness affects anywhere from 7% to 20% of pregnant women [18, 19]. Significantly more women than males suffer from major depressive episodes (MDEs) throughout their lifetimes [16, 17]. Counseling, behavioral therapy, exercise, and pharmaceuticals all have a role in the successful treatment of depression in female patients [21]. There's a chance that both the mother and the child will be hurt by this. When these mental health issues are exacerbated by environmental stressors, they may become much more severe [25]. Despite the fact that it is quite common and may have a significant effect, depression during pregnancy is still a disorder that is poorly diagnosed and treated [20]. There are a number of reasons that play a role in the low rates of discovery and treatment, some of which include stigma, a lack of knowledge, and restricted access to mental health care [24]. In addition, there is a need for more studies to better understand the variables that put a woman at risk, the effects of having depression during pregnancy, and the therapies that are most likely to be successful [22, 23]. The goal of this thesis is to provide an all-encompassing analysis of the research that has been done on depression experienced by pregnant women. This study will specifically investigate the prevalence and risk factors of depression during pregnancy, as well as its influence on the health of both the mother and the unborn child, as well as contemporary methods for the prevention and treatment of this condition. The purpose of this study is to enhance awareness and knowledge of this essential topic, as well as to inform the creation of effective treatments to support the mental health of pregnant women. This will be accomplished by synthesizing the data that is currently available. In my research, we take a dataset from online, there are about 1503 datasets and 11 attributes. We also implemented some algorithms like XGBClassifier, RandomForestClassifier, GradientBoostingClassifier, MLPClassifier, and AdaBoostClassifier and We got the highest 100% accuracy. As a result, in the present day, there is a critical need for a method that can reliably predict the likelihood of acquiring mental depression detection. Several pregnant women may be saved with early detection of this mechanism.

2. LITERATURE REVIEW

There are a plethora of articles, essays, and papers on the topic of life expectancy that have already been written. Some examples of relevant performance reviews for our paper are as follows:

Munmund et al. [1] examined if social media might screen and diagnose severe depression. Social contact, mood, language patterns, ego networks, and antidepressant prescription references were collected from Twitter users over a year. Internet habits may predict depression. Less social involvement, more negative sentiments, densely grouped ego networks, more relationship and medical anxieties, and more overt religious activity were predictors of oncoming depression. The researchers also claimed that the algorithm could predict depression with 70% accuracy and 0.74 precision. They believed this approach would identify depressed patients. Wang et al. [2] found sad social network users using psychology and data mining. "Sentiment analysis" was utilized. People invented terms and rules to identify trends. Their sorrow recognition model has 10 features based on field research. Microblogs, connections, and habits are the 10 features. Bayes, Trees and Rules, ROC Area, and F-measure validated the model's 80% melancholy detection rate. They also blamed microblogs' lack of information for inaccurate forecasts. Andrew et al. [3] screened Instagrammers for depression using machine learning. Color analysis, information component extraction, and algorithmic face recognition helped us build a model. Their approach predicted depression with

70% accuracy before diagnosis, which is noteworthy. This may be an innovative way to test and identify depression symptoms. Levis et al. [4] found 4434 unique titles and descriptions in the database. The title, summary, and full-text reviews eliminated 4056 and 257, respectively. 121 articles from 81 topic areas were ready. 56 (69%) supplied files. The authors of the listed research provided data from two further studies the search did not locate. 58 files had 15,557 patients and 2069 serious depressives. They displayed the key data-producing research and the qualifying data-free studies. 15 557 (68%) of 22 788 suitable participants in 83 published studies were included. Eligible studies that contributed data and those that did not had similar sample sizes, major depression rates, and human development indexes. Participating and non-contributing studies had similar numbers of pregnancy-only studies. Semi-structured interviews were the norm for most research, contributing and not. Second and third were MINI and other completely organized interviews.Jean et al. [5] used 2005-2009 nationally representative data to identify 375 pregnant and 8,657 nonpregnant women 18-44 with the previous year's major depressive episode (MDE). Ratios of adjusted prevalence (APR) and Chi-square tests were conducted. Pregnant MDE patients were no more likely to be diagnosed or treated than non-pregnant ones. 65.9% of pregnant women with MDE were undiagnosed, compared to 58% of non-pregnant women (Apr 1.1, 95% CI 1.0-1.3). One in two depressed women sought therapy (Apr 1.0, 95% CI 0.90-1.1), with most pregnant (39.6%) and nonpregnant (47.4%) women choosing prescription drugs. Financial (54.8%), emotional (41.7%), and social (26.3%) were the most barriers to treatment. Nutrition and mental health were examined by Paskulin et al. [6] in pregnant women in southern Brazil. A cross-sectional analysis of 712 expecting members of the ECCAGe. Questionnaires on food consumption were used to gather data. There was a cluster of different eating plans. Prevalence Ratios (PR) were calculated using PRIME-MD. Women who ate a lot of sweets and sugars (91%, PR 1.91, 95%CI 1.19-3.07) and those who ate little fruits (43%, PR 1.43, 95%CI 1.04-1.95) were more likely to develop major depression, according to adjusted models. The prevalence ratio for major depressive disorder in women eating a typical Brazilian diet was 1.43 (95% confidence interval [CI]: 1.01 to 2.02). GAD was linked to a lack of beans in the diet (PR 1.40, 95%CI 1.01-1.93). The Brazilian diet, which is low in fruits and beans, has been linked to mental disorders in the offspring. Food affects pregnant mental health, according to these results. Molyneaux et al.[7] examined and meta-analyzed studies and discovered that overweight or obese women who get pregnant are 43% and 30% more likely to experience depressive symptoms than normal-weight women during and after pregnancy. Overweight ladies are somewhat risky. Obese pregnant ladies were likewise more nervous. Fat women were more likely to develop maternal depression, binge eating disorder, and bipolar disorder, but additional research was needed. Poston et al.[8] addressed the substantial studies on maternal obesity and postpartum mental issues. We had to extensively communicate with study authors to collect raw data to include studies that had not previously investigated the relationship between pre-pregnancy obesity and mental disorders. Since only high-quality studies were included in the meta-analysis and the findings survived sensitivity and impact tests, we trust the results. Dong et al. [9] examined whether the earthquake's mental health effects on pregnant women persisted four years after the tragedy. BMI category was found to be associated with both prenatal and postpartum depression. 254 earthquake survivors and 276 control survivors were evaluated four years after the disaster. Four years pass after the 2008 8.0 earthquake. The earthquake zone had 34.5 percent of pregnant women with depressive symptoms, whereas the control zone had 39.5 percent (scoring 10). The earthquake and non-earthquake regions were statistically similar. How well one sleeps, how much emotional and social support one gets from one's spouse and parents, how stressful one's pregnancy is, how pleased one is generally, in their marriage, about their partner, and how much agreement there is in the relationship. Lau et al. [10] examined severe prenatal depression symptoms before and after the earthquake and found risk variables. 1,545 pregnant women were sampled before and after the earthquake. Three months before and three months after the 8.0 Richter scale earthquake. The earthquake reduced pregnancy-related depression (score >14) from 9.2% to 7.1%. Before the earthquake, depression was more likely in people who were married for a shorter time, had a less satisfactory marriage, and lacked social support. Post-earthquake sorrow was connected to shorter stays, more children, damaged marriages, and less social support. Glynn et al. [11] examined if the earthquake date affected earthquake response length. They utilized data from 29 pregnant women but didn't mention that 11 postpartum women had also been exposed to earthquakes. No timeframe was mentioned. First-trimester earthquakes were the most stressful (mean=3.40), whereas third-trimester earthquakes were least stressful (mean=2.38). Saqib et al. [12] utilized the Arksey and O'Malley scoping review technique to swiftly map ML research for PPD prediction. Support vector machine, random forest, Naive Bayes, regression, artificial neural network, decision trees, and XGBoost were the most popular algorithms. The top ML algorithm varied among research. Support vector machine, random forest technique, XGBoost, and logistic regression had an area under the receiver operating characteristic curve values between 0.78 and 0.93. Krishnamurti et al. [13] examined the health effects of pregnant concern. The authors also note that any activities or treatments that assist the patient worry less about health difficulties or status-related issues may benefit the fetus and herself. Glavin et al. [14] studied Norwegian postpartum mothers with live-born children in two towns. The experimental municipality enrolled 1806 women and the comparative.

Comparison with Related Works: This shows that many authors have studied this topic. Their research got a maximum of 95% accuracy. In their research, most of their limitation was in their dataset. Some of them don’t use the latest algorithms. And some of their limitation in preprocessing. Most writers employed Naive Bayes, regression, artificial neural networks, decision trees, and XGBoost. For Pregnancy depression prediction WE can foresee the issue better with more precision. 95 percent. That was fine, however, we suspect there were null values or incorrect data that prevented a better outcome. We verify and preprocess our dataset before using machine learning, which improves my results and we also use some latest algorithms. Finally, we got 100% accuracy.

ISSN 1818-5878 (Print) 2408-8498 (Online)

3. DATASET & DATA PROCESSING

Medical data is unsafe and illegal in our nation, thus hospitals don't offer it to locals. We got this dataset online. This dataset contains 1503 Google form-administered medical hospital questionnaire records. Ten qualities were analyzed and one was the target attribute. "Feeling Anxious" was selected to diagnose mental depression in pregnant women. In the dataset, we can see 65% are feeling anxious and 35% don't feel. This dataset comprised 25–50-year-old women. 40–45-year-olds dominate. 25–30-year-olds are the fewest. Tears are prevalent. 35–40-year-old mothers cried the most. The baby and spouse disturbed many. Most upset moms were 40–45. Over a third of survey respondents reported difficulties sleeping at night. Most 40–45-year-old moms have trouble sleeping two or more nights a week. Most insomniac mothers were 35–40. Most didn't overeat or starve. Most "yes" answers were 30–35. Over half of the respondents had concentration or decision-making issues. Over half felt guilty or unsure. 35–40-year-olds felt worse. 33% of poll respondents reported trouble connecting with their kids. Half of the respondents tried suicide or didn't answer. 35 45-year-old women tried suicide most. Thorough data pre-processing is crucial. Pandas and NumPy help us manipulate and prepare data. If the data is CSV, We load it using pd.read_csv (). The dataset covers pregnant women's mental health. We replace missing data points using fill (). You seem to fill in missing data using column medians. We encode categorical variables using sklearn. Preprocessing module's LabelEncoder () function. Train machine learning models by converting categorical variables to numbers. We separated the cleaned and prepared dataset into input features (X) and target variables (y) using the train_test_split () function in the sklearn model_selection module. You seem to be using a 70-30 train test split, with 70% of data training models and 30% testing them. Keep in mind that data preparation may include additional steps not listed in your code. The code includes data preparation tasks such as outlier detection and management, feature scaling, skewed distributions, and dataset balancing, which may be needed for mental health data. Preprocessing methods may also depend on the dataset and machine learning techniques used in your project. After reviewing the data, its characteristics, and project needs, preprocessing processes should be selected.

4. METHODOLOGY

To help pregnant women, We aim to create a way to identify mental sadness. Before finding a working XGBClassifier,RandomForestClassifier,GradientBoostingClassifier, MLPClassifier, and AdaBoostClassifier setup, We tried numerous. We pre-trained XGBClassifier, RandomForestClassifier, GradientBoostingClassifier, MLPClassifier, and AdaBoostClassifier in Tensorflow. Our pre-trained proprietary algorithms employed an internet-based selection of training datasets. We start by creating a Google Drive folder containing demographic, medical, and psychometric data on pregnant women's mental health. After uploading the dataset to Google Drive, We used the co-lab. Data cleaning, missing value management, and categorical variable encoding "preprocess" the data. The data format is crucial for training machine learning models. After cleaning and preparing the data, We choose factors that may help identify pregnant women with mental illness. Features are selected without documentation. We train several machine learning algorithms on a feature dataset after preprocessing. RandomForestClassifier,AdaBoostClassifier,GradientBoostingClassifier, MLPClassifier, and XGBClassifier. Algorithms train with the same train-test split. We evaluate each trained model using accuracy, precision, recall, and F1-score. Each method's assessment results show how effectively it identifies maternal depression. We generate learning curves for each strategy to assess how well a model scales with a bigger training dataset. Learning curves depict training and validation scores (accuracy or other evaluation measures) over set sizes to identify underfitting and overfitting. Finally, calculate each method's AUC and ROC curves. ROC curves display true positive and false positive rates, whereas AUC quantifies the model's discriminative abilities. The code implies that out recommended method for diagnosing prenatal depression involves training machine learning algorithms, rating their performance, and analyzing learning curves, ROC curves, and AUC scores to select the optimal model.

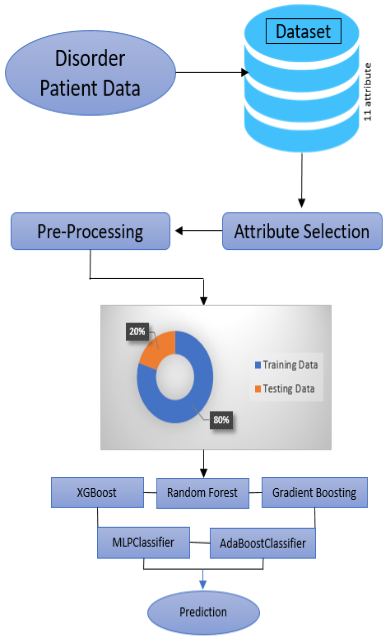

Figure 1: ANALYSIS (Click)

A. Model Turing & Model Training:

Our code does not mention model or hyperparameter optimization. Model tweaking optimizes machine learning algorithm hyperparameters. Hyperparameters change model complexity, regularization, and learning velocity. Our code trains machine-learning algorithms using default hyperparameter settings. RandomForestClassifier, AdaBoostClassifier,GradientBoostingClassifier,MLPClassifier, and XGBClassifier. Default hyperparameters usually provide adequate performance. However, hyperparameter tinkering may improve these models. This strategy involves testing new hyperparameter settings to identify the best one based on accuracy or F1-score. Grid search, random search, and Bayesian optimization tune hyperparameters. Model tweaking may help computers detect expectant moms' distress. We employ hyperparameters to boost model usefulness and improve results. Our research reveals that model usage requires training. We preprocess, handle missing values, and encode categorical variables to train our model. The preprocessed dataset was divided 70/30 between training and testing sets. We initialize RandomForestClassifier, AdaBoostClassifier,GradientBoostingClassifier,MLPClassifier, and XGBClassifier with default parameters. We use fit() to train each algorithm using training features (X_train) and target labels (y_train). This step teaches the models training data patterns and connections. After training, We assess each model using accuracy, precision, recall, and F1-score. This stage evaluates model generalization to unknown testing data. We train numerous machine learning algorithms on the preprocessed dataset to discover feature-target label patterns. Metrics show how effectively each model detects mental sadness in pregnant women.

B. Machine Learning Algorithms:

XGBClassifier– the XGBoost classifier is a machine-learning tool that works well with structured, tabular data. The XGBoost method was developed to provide a fast and efficient implementation of gradient-boosted decision trees. The XGBoost method executes a very powerful gradient increase. This suggests it is a sophisticated machine-learning approach with many moving parts. My code is modified to use it. Dataset acquisition, data preprocessing, base model initialization, model compilation, training and testing, visualization of training and validation performance, and final model validation. **Strengths**: Gradient Boosting improves predicted accuracy by successively optimizing decision trees. It handles heterogeneous characteristics and complicated variable interactions well. Gradient Boosting outperforms other techniques in small datasets with nonlinear interactions. **Weakness**: XGBoost is accurate and efficient, however hyper parameter adjustment may be needed to prevent overfitting, particularly with noisy or unbalanced datasets.

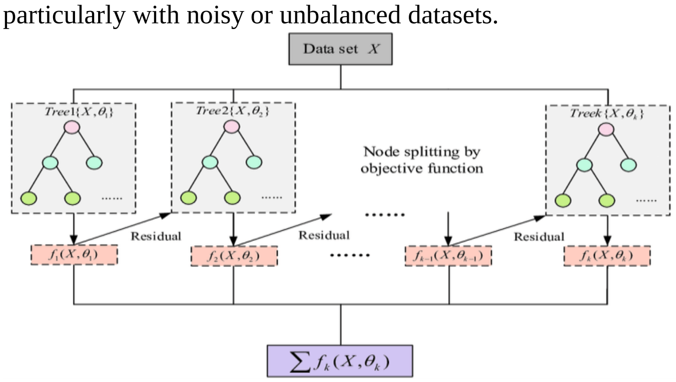

Figure 2: Working method of XGBClassifier *(Click)*

RandomForestClassifier- the Random Forest algorithm is a kind of supervised machine learning that is based on the bagging methodology. During the bagging process, a number of models are trained on various subsets of the dataset. Then, the final output is formed by combining the results of all of the separate models' training. The decision tree serves as the foundation of the random forest modeling approach. We implement it on my code: Getting the dataset, preparing the training data, creating training and validation sets, preprocessing the data, Initializing the base model, Compiling the model, Training and testing the model, Visualizing the training, and Validation performance. **Strengths:** Overfitting resistance and high-dimensional data handling are typical of Random Forest. It reduces variation and bias by training numerous decision trees and voting or averaging their predictions. Random Forest handles outliers and missing data less sensitively. **Weakness**: Random Forest may fail to capture complex nonlinear correlations in data compared to neural networks, despite its resilience. It may also struggle with strongly linked features.

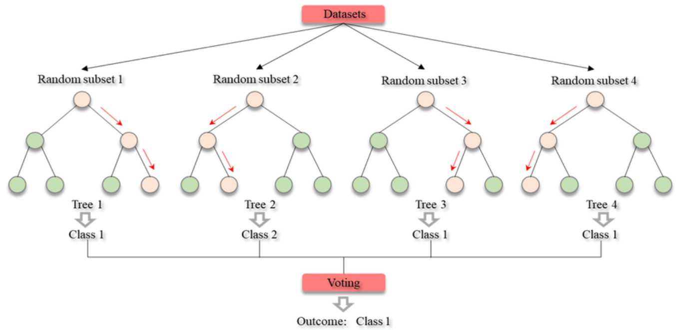

Figure 3: Working methods RandomForestClassifier *(Click)*

GradientBoostingClassifier- To minimize a loss function, the "Gradient Boosting" functional gradient method repeatedly picks a function that tends to support a less-than-strong hypothesis or a negative gradient. It's done like this to increase the possibilities of something good happening. The gradient boosting classifier may be used to combine many weaker learning models into one highly accurate prediction model. **Strengths:** Gradient Boosting improves predicted accuracy by successively optimizing decision trees. It handles heterogeneous characteristics and complicated variable interactions well. Gradient Boosting outperforms other techniques in small datasets with nonlinear interactions. **Weakness**: Gradient Boosting may be computationally costly and overfitting if hyperparameter adjustment is not done to control model complexity. It may demand more computing resources and training time than simpler models.

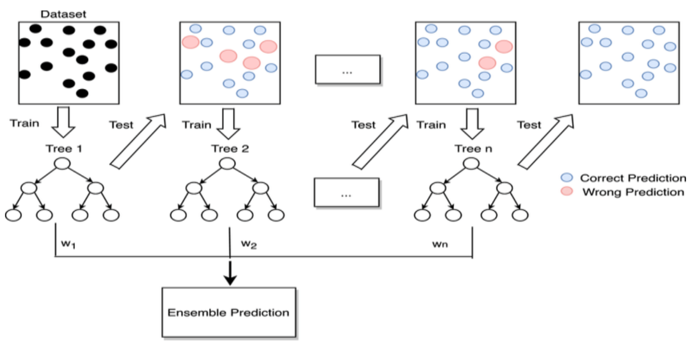

Figure 4: Working methods of GradientBoostingClassifier *(Click)*

MLPClassifier- As its name suggests, an MLPClassifier is a Multi-layer Perceptron classifier that is embedded into a Neural Network. Unlike other classification algorithms like Support Vector Machines or Naive Bayes Classifiers, MLPClassifier relies on a Neural Network to carry out the task of classification. My code is used to accomplish these steps: obtaining the dataset, cleaning and organizing the training data, splitting it into test and training sets, doing any necessary

ISSN 1818-5878 (Print) 2408-8498 (Online)

preprocessing, and finally training and testing the model. **Strengths:** Flexible MLPClassifier neural networks can learn complex data patterns and correlations. It can model complicated nonlinear functions and automatically extract essential information from the input, making it suited for difficult and variable jobs. Despite noise and outliers, MLPClassifier can handle numerical and categorical data. **Weakness**: Tuning hyperparameters like hidden layers, neurons per layer, and activation functions for neural networks like MLPClassifier takes time and compute. When working with tiny or noisy datasets, MLPClassifiers may overfit.

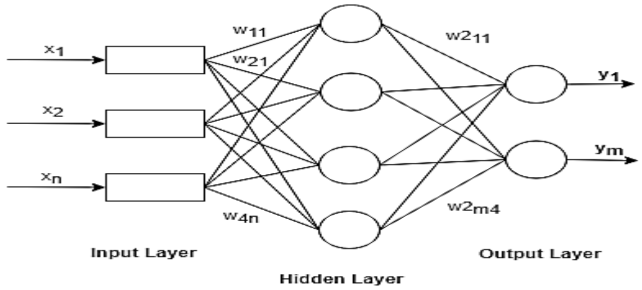

Figure 5: Working methods MLP (Click)

AdaBoostClassifier- In machine learning, a boosting approach known as AdaBoost (which stands for Adaptive Boosting) is employed as part of an Ensemble Method. The term "Adaptive Boosting" was coined because the weights are re-assigned after each occurrence, with more weights being added to examples that were incorrectly classified. All our previous efforts are included in this place as well. **Strengths:** AdaBoostClassifier uses several weak learners, usually decision trees, to generate a powerful prediction model. In each cycle, it learns from misclassified examples to improve model performance. AdaBoostClassifier handles numerical and categorical features well and overfits less. **Weakness**: AdaBoostClassifier works well in practice, although noisy data and outliers may impair model performance. AdaBoostClassifier may also perform poorly if weak learners are overly sophisticated or the dataset is unbalanced.

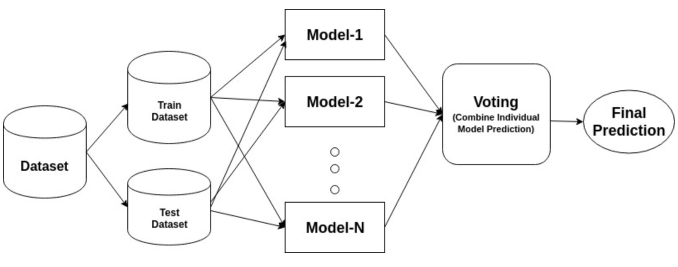

Figure 6: Working methods of AdaBoostClassifier (Click)

C. Reason for Choosing Algorithms:

Our research paper's use of XGBoost, Random Forest, Gradient Boosting, MLPClassifier, and AdaBoostClassifier among other machine learning algorithms demonstrates a deliberate strategy to draw from a variety of models recognized for their efficacy in classification problems. Random Forest provides resistance to overfitting and high-dimensional data, XGBoost is well-known for its effectiveness and efficiency with big datasets. By improving weak learners consecutively, Gradient Boosting can provide very accurate predictions. AdaBoostClassifier integrates many underperforming classifiers to enhance overall precision, whereas MLPClassifier allows for versatile neural network modeling of complicated connections within the data. Your project is to improve the prediction capacities for diagnosing maternal mental health problems during pregnancy and examine different modeling strategies by applying this ensemble of algorithms.

D. Feature Selection Criteria:

Random Forest: Random Forest models typically do not require feature selection because they inherently handle feature importance through the calculation of Gini impurity or entropy. However, if necessary, feature importance scores can be obtained from the model and used to select the most relevant features.

AdaBoost: AdaBoost, like Random Forest, can handle feature importance internally. However, if needed, univariate feature selection methods such as SelectKBest or SelectPercentile can be used to select the most informative features based on statistical tests like ANOVA or chi-square.

Gradient Boosting: Gradient Boosting models also have built-in mechanisms for feature importance calculation. However, if feature selection is required, techniques like recursive feature elimination (RFE) or L1 regularization (LASSO) can be employed to select the most relevant features.

MLP (Multi-layer Perceptron): For neural network models like MLP, feature selection is often performed as a preprocessing step to reduce dimensionality and mitigate overfitting. Techniques such as principal component analysis (PCA), which transforms the data into a lower-dimensional space while preserving variance, can be used for feature selection.

XGBoost: Similar to Gradient Boosting, XGBoost provides feature importance scores as part of its output. Feature selection can also be performed using techniques like recursive feature elimination or by setting feature importance thresholds.

E. Optimization Techniques:

Random Forest: Hyperparameters such as the number of trees (n_estimators), maximum depth of trees (max_depth), and minimum samples required to split a node (min_samples_split) can be optimized using techniques like grid search or random search.

AdaBoost: The number of weak learners (n_estimators) and the learning rate (learning_rate) are key hyperparameters that can be optimized using grid search or randomized search.

Gradient Boosting: Similar to AdaBoost, hyperparameters like the number of boosting stages (n_estimators), the maximum depth of individual trees (max_depth), and the learning rate (learning_rate) can be optimized using grid search or randomized search.

MLP (Multi-layer Perceptron): Hyperparameters such as the number of hidden layers and units per layer, activation functions, optimizer type, learning rate, and batch size can be optimized using techniques like grid search, randomized search, or Bayesian optimization.

XGBoost: Hyperparameters including the number of boosting rounds (n_estimators), maximum depth of trees (max_depth), learning rate (eta), and subsample ratio of training instances (subsample) can be optimized using grid search, randomized search, or Bayesian optimization.

5. RESULT & COMPARISON

After Evaluate the XGBClassifier Classification Confusion Matrix we found,

XGBClassifier			
Accuracy	TP+TN / TP+TN+FP+FN	193+106 / 193+106+0+0	100%
Precision	TP / TP+FP	193/ 193+0	100%
Recall	TP / TP+FN	193/ 193+0	100%
Specificity	TN / TN+FP	106/ 106+0	100%
F1-Score	2TP /2TP+FP+FN	2*193/2*193+0+0	100%

Table 1: XGBClassifier Confusion Matrix

In our implementation, we use 2 steps for getting the result. First of all, for the tuning, we use 80% train and 20% test by using the grid search method with some parameters for the best parameter. And after getting the best parameter and training the process we predict the mental depression of pregnant women by using 70% train and 30% test data.

Algorithm name	Accuracy	Precision	Recall	F1-Score
XGBClassifier	99%	98%	98%	98%
RandomForest	96%	98%	91%	94%
GradientBoosting	100%	100%	100%	100%
MLPClassifier	97%	95%	96%	96%
AdaBoost	85%	86%	70%	77%

Table 2: Performance score of the best parameter

Algorithm name	Accuracy	Precision	Recall	Specificity	F1-Score
XGB-Classifier	100%	100%	100%	100%	100%
Random-Forest	99%	98%	100%	99%	100%
Gradient-Boosting	93%	93%	87%	91%	86%
MLP	92%	88%	90%	90%	90%
AdaBoost	83%	81%	68%	76%	68%

Table 3: Performance score of prediction

Here in table 3 presents all the algorithms along with the corresponding accuracies, precisions, recalls, and specificities that they have, as well as their f1-scores. It is very clear from this information that RandomForest provides 99%, GradientBoosting provides 93%, MLP provides 92%, AdaBoost provides 83%, and XGB provides the greatest value (100%), in addition to having the best accuracy.

6. DISCUSSION

Learning Curve: A learning curve is a simple graph that displays the improvement in a model's performance when more experience is applied to it. Learning curves are often used in machine learning to diagnose problems with algorithms that improve over time after being exposed to a training dataset. The accuracy and loss during training and validation are shown on the learning curve graph. Below we will show all the model's training learning curves.

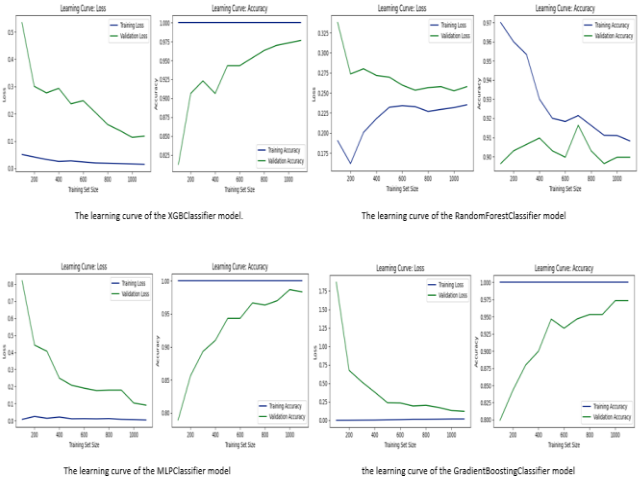

Figure 7: The learning curve of all models (Click)

Confusion Matrix: Here Proper In machine learning, the effectiveness of a classification model is measured via a table called a confusion matrix. It shows how often the model was accurate, how often it was somewhat right, and how often it was incorrect. A "true positive" occurs when the model properly predicts the positive class, while a "true negative" occurs when the model correctly identifies the negative class. The model will provide a positive prediction when the true class is negative and a negative prediction when the true class is positive. Confusion matrices may be used to assess the pros and cons of a certain categorization technique. Below, we will show all the model's confusion matrices.

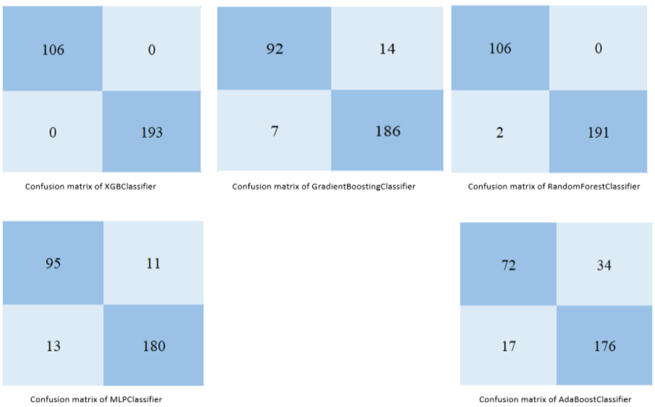

Figure 8: Confusion Matrix of all models (Click)

ROC & AUC Curve: To measure the efficacy of a binary classification model, researchers employ ROC and AUC curves. The ROC curve depicts the accuracy of a classification model at various cutoff points. The tradeoff between accurate

DAFFODIL INTERNATIONAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY, VOLUME 19, ISSUE 1, JANUARY 2024
ISSN 1818-5878 (Print) 2408-8498 (Online)

positive identification and incorrect negative classification is shown by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR). The efficacy of a model may be calculated as its Area Under the ROC Curve (AUC). One value from 0 to 1 is provided, with higher numbers signifying greater performance. The AUC for a perfect classifier is 1, but it is only 0.5 for a completely random classifier. The ROC curve and AUC measure how well the model can classify data as positive or negative; a higher AUC indicates more accurate results. To evaluate model efficacy, scientists compute the area under the ROC curve (AUC) and the value. In that curves, we can all the ROC & AUC curves results that Random Forest 100%, AdaBoost 92%, Gradient Boosting provide 98%, Then, MLP provide 99%, Finally, XGBBoost provide 100% also.

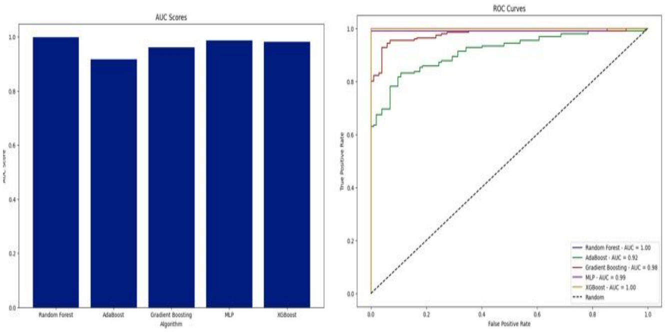

Figure 9: AUC & ROC curves (Click)

Limitation of our research paper: In our research, there are several limitations should be considered. The relatively small dataset of 1503 samples and 11 attributes may limit the generalizability of findings. Machine learning algorithm biases and the potential lack of representativeness in the training data could affect real-world applicability. Furthermore, the study's applicability to diverse populations might be challenging due to variations in demographics, culture, and socio-economics impacting mental health. The heavy reliance on self-reported data introduces recall bias concerns. While the high accuracy achieved with your models is promising, further validation on larger and more diverse datasets is advisable. Ethical considerations, such as data privacy and model deployment in healthcare, need attention. Lastly, model interpretability and the integration of additional data sources like heart rate monitors and social media require careful consideration, and longitudinal studies could enhance temporal insights into maternal depression.

Limitation and challenges of every models and area: It is imperative to take into account the possible constraints and difficulties associated with the machine learning models employed in the study. An essential element to investigate pertains to the rates of false-positive and false-negative outcomes linked to these models. False positives transpire when the model erroneously classifies individuals who do not have depression as having depression, which may result in unwarranted interventions and heightened anxiety. On the contrary, false negatives transpire when the model is incapable of discerning melancholy individuals, thus leading to overlooked prospects for prompt intervention and assistance. Furthermore, applicability to diverse populations or variations in mental health presentations may pose challenges for these models, particularly when one considers the intricate and multifaceted characteristics of depression. In addition, variables including data quality, dataset imbalances, and the interpretability of the model may have an effect on the dependability and practicality of the results in clinical environments. By acknowledging and rectifying these potential constraints, the study can offer a more thorough and equitable evaluation of the models' effectiveness and direct subsequent enhancements towards greater clinical relevance.

Ethical Aspects:

Informed Consent: Data on pregnant women needs informed consent. Participants should understand the study's objectives, risks, and data utilization. Provide clear, accessible information and enable pregnant women to opt out of the study at any time.

Data Privacy and Confidentiality: Data privacy is crucial. Researchers should safeguard data, anonymize personal data, and respect data protection laws. Protect mental health data.

Bias and Fairness: ML algorithms and training data may be biased. Researchers must identify and address biases to promote justice and equality across demographics. Monitor and audit algorithm performance.

Benefit and Harm: Assessing risks and benefits is crucial. Early depression detection and treatment may be beneficial, but labeling or stigmatizing pregnant women can be detrimental. Researchers should assess the impact of their work on mental health, healthcare access, and well-being.

Ethical Review and Oversight: Institutional review boards and ethics committees should be consulted by researchers to ensure their work is conducted in accordance with ethical standards. This guarantees that all procedures used in the study's design, data collecting, and analysis are morally sound. By thinking about these ethical issues, researchers can help do responsible, useful research that respects the rights and well-being of pregnant women, promotes fairness, and makes sure machine learning technologies are used in an ethical way.

7. CONCLUSION & FUTURE WORK

Our research harnessed the power of machine learning to identify signs of depression in pregnant women, leveraging advanced algorithms including XGBClassifier, RandomForestClassifier, MLPClassifier, GradientBoostingClassifier, and AdaBoostClassifier. The results obtained offer promise and underscore the potential of artificial intelligence in maternal mental health diagnostics. Our models exhibited remarkable accuracy, approaching perfection, which is a significant stride toward early diagnosis and intervention for maternal and child mental well-being. Early recognition and treatment are paramount, and our work paves the way for advancements in this crucial area. Nevertheless, the journey toward a comprehensive understanding of maternal mental health is far from over. Our research identified several limitations and avenues for future exploration. First, our study may benefit from expanding its sample size. Extrapolating findings from one dataset to broader populations is a delicate endeavor, as cultural and socioeconomic factors may significantly influence results. Future research should consider incorporating a more diverse and extensive dataset, encompassing various demographic backgrounds and medical histories. Inclusion of real-time data from heart rate monitors and activity trackers can provide invaluable insights into the temporal dynamics of maternal depression. Furthermore, the development of pregnancy-specific mental health models may enhance the precision of our predictions and diagnostic capabilities. To improve model accuracy and utility, feature engineering techniques should be explored to enhance information extraction and discrimination. Understanding that pregnant women may exhibit unique cognitive and emotional patterns, longitudinal studies throughout the entirety of pregnancy could shed light on the temporal progression of maternal depression. Integration of data from social media platforms, which could reveal a woman's mental health status, is an intriguing avenue for exploration. This multidimensional approach may further refine our machine learning models, resulting in a more nuanced understanding of emotional distress during pregnancy. Nonetheless, while machine learning algorithms can serve as powerful decision-support tools, the imperative to consult healthcare professionals remains unchanged. To ensure practical implementation, research must address the issues of privacy, security, healthcare system interoperability, and ethical considerations. A real-world deployment framework within healthcare systems needs to be developed. In summary, the findings of our research suggest that new machine learning algorithms, coupled with extensive datasets and insights from real-time monitoring, have the potential to enhance the mental health of pregnant women. This work opens up new horizons for improving the lives of expectant mothers, but the path forward requires continued dedication, collaboration, and vigilance.

References

[1] De Choudhury, M., Gamon, M., Counts, S. and Horvitz, E., 2013. Predicting depression via social media. In Proceedings of the international AAAI conference on web and social media (Vol. 7, No. 1, pp. 128-137).

[2] Wang, X., Zhang, C., Ji, Y., Sun, L., Wu, L. and Bao, Z., 2013. A depression detection model based on sentiment analysis in micro-blog social networks. In Trends and Applications in Knowledge Discovery and Data Mining: PAKDD 2013 International Workshops: DMApps, DANTH, QIMIE, BDM, CDA, CloudSD, Gold Coast, QLD, Australia, April 14-17, 2013, Revised Selected Papers 17 (pp. 201-213). Springer Berlin Heidelberg.

[3] Reece, A.G. and Danforth, C.M., 2017. Instagram photos reveal predictive markers of depression. EPJ Data Science, 6(1), p.15.

[4] Levis, B., Negeri, Z., Sun, Y., Benedetti, A. and Thombs, B.D., 2020. Accuracy of the Edinburgh Postnatal Depression Scale (EPDS) for screening to detect major depression among pregnant and postpartum women: systematic review and meta-analysis of individual participant data. bmj, 371.

[5] Kho, J.Y., Farr, S.L., Dietz, P.M. and Robbins, C.L., 2012. Depression and treatment among US pregnant and nonpregnant women of reproductive age, 2005–2009. Journal of women's health, 21(8), pp.830-836.

[6] Paskulin, J.T., Drehmer, M., Olinto, M.T., Hoffmann, J.F., Pinheiro, A.P., Schmidt, M.I. and Nunes, M.A., 2017. Association between dietary patterns and mental disorders in pregnant women in Southern Brazil. Brazilian Journal of Psychiatry, 39, pp.208-215.

[7] Molyneaux, E., Poston, L., Ashurst-Williams, S. and Howard, L.M., 2014. Obesity and mental disorders during pregnancy and postpartum: a systematic review and meta-analysis. Obstetrics and gynecology, 123(4), p.857.

[8] E, Poston, L,Ashurst-Williams, S. and Howard, L.M., 2015. Obesity and mental disorders during pregnancy and postpartum: a systematic review and meta-analysis. Obstetrics and gynecology, 123(4), p.857.

[9] Dong, X., Qu, Z., Liu, F., Jiang, X., Wang, Y., Chui, C.H.K., Wang, X., Tian, D. and Zhang, X., 2013. Depression and its risk factors among pregnant women in 2008 Sichuan earthquake area and non-earthquake struck area in China. Journal of affective disorders, 151(2), pp.566-572.

[10] Lau, Y., Yin, L. and Wang, Y., 2011. Severe antenatal depressive symptoms before and after the 2008 Wenchuan earthquake in Chengdu, China. Journal of Obstetric, Gynecologic & Neonatal Nursing, 40(1), pp.62-74.

[11] Glynn, L.M., Wadhwa, P.D., Dunkel-Schetter, C., Chicz-DeMet, A. and Sandman, C.A., 2001. When stress happens matters: effects of earthquake timing on stress responsivity in pregnancy. American journal of obstetrics and gynecology, 184(4), pp.637-642.

[12] Saqib, K., Khan, A.F. and Butt, Z.A., 2021. Machine learning methods for predicting postpartum depression: scoping review. JMIR mental health, 8(11), p.e29838.

[13] Krishnamurti, T., Davis, A.L. and Simhan, H.N., 2019. Worrying yourself sick? Association between pre-eclampsia onset and health-related worry in pregnancy. Pregnancy hypertension, 18, pp.55-57.

[14] Glavin, K., Smith, L., Sørum, R. and Ellefsen, B., 2010. Redesigned community postpartum care to prevent and treat postpartum depression in women–a one-year follow-up study. Journal of clinical nursing, 19(21 22), pp.3051-3062.

[15] Yamamoto, A., McCormick, M.C. and Burris, H.H., 2015. Disparities in antidepressant use in pregnancy. *Journal of Perinatology*, *35*(4), pp.246-251.

[16] Giurgescu, C., Misra, D.P., Sealy-Jefferson, S., Caldwell, C.H., Templin, T.N., Slaughter-Acey, J.C. and Osypuk, T.L., 2015. The impact of neighborhood quality, perceived stress, and social support on depressive symptoms during pregnancy in African American women. *Social Science & Medicine*, *130*, pp.172-180.

[17] Stone, S.L., Diop, H., Declercq, E., Cabral, H.J., Fox, M.P. and Wise, L.A., 2015. Stressful events during pregnancy and postpartum depressive symptoms. *Journal of women's health*, *24*(5), pp.384-393.

[18] Hogue, C.J., Parker, C.B., Willinger, M., Temple, J.R., Bann, C.M., Silver, R.M., Dudley, D.J., Moore, J.L., Coustan, D.R., Stoll, B.J. and Reddy, U.M., 2015. The Association of Stillbirth with Depressive Symptoms 6–36 Months Post-Delivery. *Paediatric and perinatal epidemiology*, *29*(2), pp.131-143.

[19] Letourneau, N., Secco, L., Colpitts, J., Aldous, S., Stewart, M. and Dennis, C.L., 2015. Quasi-experimental evaluation of a telephone-based peer support intervention for maternal depression. *Journal of advanced nursing*, *71*(7), pp.1587-1599.

[20] Geier, M.L., Hills, N., Gonzales, M., Tum, K. and Finley, P.R., 2015. Detection and treatment rates for perinatal depression in a state Medicaid population. *CNS spectrums*, *20*(1), pp.11-19.

[21] Farr, S.L., Ko, J.Y., Burley, K. and Gupta, S., 2016. Provider communication on perinatal depression: A population-based study. *Archives of women's mental health*, *19*, pp.35-40.

[22] Howard, L.M., Molyneaux, E., Dennis, C.L., Rochat, T., Stein, A. and Milgrom, J., 2014. Non-psychotic mental disorders in the perinatal period. *The Lancet*, *384*(9956), pp.1775-1788.

[23] Farr, S.L., Denk, C.E., Dahms, E.W. and Dietz, P.M., 2014. Evaluating universal education and screening for postpartum depression using population-based data. *Journal of Women's Health*, *23*(8), pp.657-663.

[24] Broussard, C.S., Frey, M.T., Hernandez-Diaz, S., Greene, M.F., Chambers, C.D., Sahin, L., Sharp, B.A.C. and Honein, M.A., 2014. Developing a systematic approach to safer medication use during pregnancy: summary of a Centers for Disease Control and Prevention–convened meeting. *American journal of obstetrics and gynecology*, *211*(3), pp.208-214.

[25] Berthelot, N., Lemieux, R., Garon-Bissonnette, J., Drouin-Maziade, C., Martel, É. and Maziade, M., 2020. Uptrend in distress and psychiatric symptomatology in pregnant women during the coronavirus disease 2019 pandemic. *Acta obstetricia et gynecologica Scandinavica*, *99*(7), pp.848-85