FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 8 No. 6, December, 2024, pp 55 - 62
DOI: https://doi.org/10.33003/fjs-2024-0806-2731

MACHINE LEARNING MODELS FOR CLASSIFICATION AND PREDICTION OF PREECLAMPSIA IN KADUNA, NIGERIA

[1]Agozie Eneh, [1]Deborah U. Ebem, *[1,2]Eli. A. Jiya, [3]Emmanuel Etuh and [4]Bulus Raskeb

[1]Department of Computer Science, University of Nigeria, Nsukka, Nigeria.
[2]Department of Computer Science, Federal University, Dustin-ma, Nigeria.
[3]Department of Mathematics/Statistics and Computer Science, Kwararafa University, Wukari
[4]Transport Technology Centre, Nigeria Institute of Technology Zaria, Nigeria

*Corresponding authors' email: jiyaeli@fudutsinma.edu.ng

ABSTRACT

Preeclampsia is a significant complication in pregnancy characterized by high blood pressure and damage to organs, posing serious risks to both the mother and fetus. Early prediction and management are crucial for improving outcomes. In Nigeria, where healthcare resources are often limited, and prenatal care access can be uneven, advanced predictive models can enhance early detection and intervention. This paper developed machine learning-based classifiers and predictors of preeclampsia. The data used was collected from general hospital Hunkuyi, Kaduna state, Nigeria. The models were based on Adaboost, Support Vector Machine (SVM), and Naïve Bayes (NB) algorithms. In both classification and prediction of preeclampsia among the population studied, SVM has 0 error in MAE, RMSE, RAE and ERSE, with accuracy level of 100%. Adaboost and NB had accuracy levels of 98% and 85%, which are very good. This paper recommends the use of these models for prediction of onset of preeclampsia among pregnant women in Hunkuyi and Kaduna state. Since the data used to develop the models represent impartially the various set of people within the state, it can be used for all women. we believe the models can assist the health personnel to predict onset of preeclampsia and help proper planning and intervention. It will also reduce maternal and child mortality that could result for preëclampsia.

Keywords: Preeclampsia Prediction, Preeclampsia classification, Machine learning, Preeclampsia in Nigeria

INTRODUCTION

Annually, large and significant number of women die due to Preeclampsia and related pregnancy induced problems (Schindler & Schindler, 2018). Some papers estimated the mortality rate due to this disorder to be about 100, 000 for women and 500,000 for babies globally. It is estimated that about 10 -15% of pregnant women globally come down with Preeclampsia and therefore consider as a leading factor in maternal and child mortality (Maric et al., 2020). In developing countries, 1.8–16.7% of events are reported, compared to 0.4% in industrialized nations (Mou et al, 2021). In Nigeria, the prevalence of preeclampsia ranges from 2% to 16% (Ugwu et al, 2022). A similar study in university of Ibadan found 7.2% among the population studied (Suleman et al., 2022). The enormity of this disorder is more common in developing countries where health facilities are not sufficiently equipped to manage health complexities. A number factors influence increase the challenges of early identification of the condition, these include difficulty in identifying the root cause of the disorder and multiple pathogenic phenotypes of the condition (Maric et al., 2020). Therefore, finding a good machine learning predictor of the condition can increase early detections of women with risk of the disease and monitoring. In some studies which examined the factors which may influence onset of preeclampsia, Abubakar et al., (2009) and Ola et al, (2023) highted a number of factors like ethnic group, family history of diabetes, age, poor education, lack of antenatal care, hypertension, heart disease, or renal disease and could be caused by regular bad diet like alcohol (Weissgerber and Mudd, 2015)

Preeclampsia is a pregnancy-induced disorder characterized by the onset of hypertension and/or significant proteinuria after the 20th week of gestation (Lu et al., 2021). Medically, the specific cause of Preeclampsia is not known, however, some studies suggested placental dysfunction as a major contributory factor for onset of Preeclampsia. The dysfunction of placenta can further lead systemic inflammation and damage to vital organs of the body (Melchiorre et al., 2020). Preeclampsia is associated with various clinical manifestations, including high blood pressure, proteinuria, oedema, organ dysfunction (such as liver or kidney abnormalities), and in severe cases, the development of seizures or eclampsia (Ngwenya et al., 2021). Preeclampsia can be in mild, moderate or severe form. Mild Preeclampsia symptoms include heartburn, nausea, high blood pressure, insufficient urination, and obesity. In severe cases, preeclampsia symptoms may manifest in form of difficulty in breathing, renal failure, oedema, impaired vision, and eye irritation and may progress to eclampsia.

The diagnosis of Preeclampsia is primarily based on the presence of hypertension and proteinuria (Portelli & Baron, 2018). Other signs, symptoms, and laboratory findings may also be considered in the diagnosis and management of the condition. The management of Preeclampsia involves close monitoring of maternal and unborn child, blood pressure control, prevention of seizures (in severe cases), and timely delivery of the baby (Enaruna & Sodje, 2015). Depending on the severity and gestational age, various treatment options, including antihypertensive medications and corticosteroids to accelerate fetal lung maturity, may be considered. Like the works of Schindler and Schindler, (2018).noted, preeclampsia patient have higher probability of suffering cardiovascular health problems later in life than those who never had it.

Preeclampsia cases was reported as a leading cause of maternal mortality in Nigeria (Okunade et al., 2014, Rana et al, 2019) and account for about 40% of maternal death in northern Nigeria (Abubakar et al., 2009). The Abubakar et al., (2009) found that ethnic group within the region are predispose to severe form of preeclampsia could progress to eclampsia; this was further corroborated Vanderjagt et al.,

(2004) which found it to be responsible for (24.2%) of women in Gombe. In two related studies in Kaduna state and northwest Nigeria, Attahir et al., (2010), Mohammed et al., (2023), and Akaba et al., (2021) noted prevalence of preeclampsia in different part of the region. Therefore, this paper developed machine learning based prediction models of Preeclampsia in northern Nigeria for early detection and classification of women with high risk of Preeclampsia. The model can be a useful aid to health professional in treatment planning and reduce the mortality rate of women and unborn children. Machine learning has become a potent tool for predictive analytics in several industries, including healthcare, in recent years. Using machine learning algorithms, it is possible to create prediction models for disease diagnosis, prognosis, and risk assessment by revealing hidden patterns and relationships in large, complicated datasets. Machine learning has the potential to enhance the early prediction and identification of women who are at high risk for developing Preeclampsia.

The works of Venkatesh et al., (2021) use random forest and gradient-boosted trees and logistic regression to develop Preeclampsia prediction model. The model used features that include maternal age, parity, chronic hypertension, gestational hypertension, Preeclampsia, infant birthweight, and maximum postpartum diastolic blood pressure. They dd a retrospective study using dataset of over 10000 delivery records. The result of their work showed that logistic regression performed best with highest accuracy. A model for early identification of Preeclampsia was developed by Maric et al., (2020). The specific maternal features like age, race of the woman, the age group, diabetes mellitus, height, history of Preeclampsia weight, and blood pleasure were used as parameters to 2 algorithms: The gradient boosting algorithm and the elastic net. The major finding was that parameters like high blood pleasure, the parity, history of Preeclampsia were strong predictors of Preeclampsia. Also include in the list is the chronic hypertension, and diabetes mellitus. The work emphasized these variables as essential in a model that can detect early onset of Preeclampsia.

One notable study by Smith et al. (2017) used a dataset of clinical and demographic information from pregnant women was to predict the onset of preeclampsia using machine learning techniques. They used methods for feature selection to determine the most relevant predictors and evaluated the effectiveness of various algorithms such as logistic regression, support vector machines, and random forests. The study yielded promising findings, with machine learning models predicting preeclampsia with good accuracy and area under the receiver operating characteristic curve (AUC). The research of Chaemsaithong et al., (2019) developed a model for prediction of Preeclampsia in first trimester. Model used regression analysis and 2 years dataset of women who attended antenatal clinic in 7 countries in Asian countries. The model was claimed to perform almost equal with the standard procedural formular used to calculate risk of Preeclampsia by health agencies in some those nations. Tarca et al.,(2021) formulated preeclampsia prediction model based on maternal risk factors, the biophysical and the age of the pregnancy. Their study was aimed was to study the influence of age of gestation and hypertension on the accuracy of Preeclampsia prediction models. The work multivariable Posson regression models which was able to predict preterm, term and post term Preeclampsia with area under cover of 0.7

In another investigation by Johnson et al. (2019), the authors explored the use of physiological data, such as blood pressure, heart rate, and uterine contractions, collected from pregnant women for early detection of preeclampsia. They applied deep learning algorithms, specifically recurrent neural networks (RNN), to capture temporal patterns in the physiological signals. The RNN-based model showed superior predictive performance, outperforming traditional machine learning algorithms, and providing valuable insights into the dynamic changes preceding preeclampsia onset. Furthermore, a study by Chen et al. (2020) focused on incorporating novel features, including circulating microRNA expression profiles, into the prediction of preeclampsia using machine learning. They utilized a deep learning model called a stacked denoising autoencoder to extract relevant features from microRNA data and combined it with clinical variables. The integrated model achieved improved accuracy and AUC-ROC compared to models using clinical variables alone, highlighting the potential of incorporating molecular data in predictive models for preeclampsia.

MATERIALS AND METHODS

Several steps were taking during the model development. The methodology adopted follow several the steps as stated in the following subsections.

Data Acquisition and Labelling

The data used in developing the Preeclampsia classifier and prediction model was collected from General Hospital Hunkuyi. The health facility is located in Kudan Local Government Area of Kaduna state. The hospital is a secondary health facility which handles medium health complexities and server hospital of referral for primary health care facilities within the LGA.

The study involved mothers attending antenatal care clinic, whose gestation period has exceeded 20 weeks. This was to ensure that the gestational periods were ripe enough to detect both chronic and gestational hypertension. The data used covers a period of 2018- 2023. A total of 208 data was used for model development and training.

During the process of data acquisition, the team worked with several midwives and nurse in the facility to arrive at various parameters and variables that are used to diagnosed Preeclampsia. The basic features of the data are shown in Table 1. While other features are familiar, order of marriage represent the number of husbands or marriage the woman has had. For example, if a woman has been divorced twice and currently married, the order in marriage is 3.

Table 1: Dataset Description

S/No	Parameter	Description	Data Type
1	Age	Age of the Patient	Integer
2	Parity	Parity	Integer
3	OrdMar	Order in Marriage	Integer
4	GesAge	Gestational Age (Weeks)	Integer
5	Prot	Presence of Proteinuria	Nominal
6	Gluc	Glucosuria	Nominal
7	SysPres	BP Systolic Value (mmHg)	Integer

8	DiasPres	BP Diastolic Value (mmHg)	Integer
9	PHH	Previous History of Hypertension	Nominal
10	PHP	Previous History of Preeclampsia	Nominal
11	Odm	Oedema	Nominal
12	Weight	Patients weight (kg)	Numerical

Data Preprocessing and feature selection

The model make use of all the features of the data collected. The decision of the usage of them is based on previous works (Chen et al., 2021; Dathan-stumpf et al., 2020; E.V. et al., 2009; Myatt, 2020; Okunade et al., 2014; Wright et al., 2019) which has listed these features as important predictors of Preeclampsia.

The nominal data like OrdMar, Prot, Gluc, SysPres, DiasPres, PHP, PHH and Ddm columns were left in their original forms, but the class column was label with "PE" for presence of preeclampsia while "NO-PE" was used to represent absence of Preeclampsia.

Machine learning Algorithms

Three algorithms: Adaptive Boosting (AdaBoost), Support Vector Machine (SVM) and Naïve Bayes (NB) were used to model Preeclampsia. SVM and Adaboost are two of a kind algorithms that have thieved in classification task. Both uses optimization functions that minimize the error of misclassification during the training process (Li et al., 2022; Mathanker et al., 2011).

Adaptive Boosting is a powerful algorithm used for classification, especially in binary classification. The algorithm develops multiple weak learners which are continually improved during the training, this process produces strong learners at the end of the training. AdaBoost assigns weights to each of the weak base classifiers uniformly, the algorithm then adjusts the weights of the misclassify data samples in other to improve weak classifiers and produce strong classifier (Wang & Sun, 2021). The algorithm of Adaboost (Hastie, Tibshirani, & Friedman, 2017; Li et al., 2008) is presented in algorithm 1.

Support Vector Machine (SVM) is a supervised learning algorithm used for classification and regression tasks. Its primary goal is to find the best decision boundary (or hyperplane) that separates data points of different classes with the maximum margin.

Algorithm 1: Adaboost algorithm

1. Initialize the observation weights $w_i = 1/N$, $i = 1, 2, \ldots, N$.
2. For $m = 1$ to M:
   (a) Fit a classifier $G_m(x)$ to the training data using weights w_i.
   (b) Compute
   $$\mathrm{err}_m = \frac{\sum_{i=1}^{N} w_i I(y_i \neq G_m(x_i))}{\sum_{i=1}^{N} w_i}.$$
   (c) Compute $\alpha_m = \log((1 - \mathrm{err}_m)/\mathrm{err}_m)$.
   (d) Set $w_i \leftarrow w_i \cdot \exp[\alpha_m \cdot I(y_i \neq G_m(x_i))]$, $i = 1, 2, \ldots, N$.
3. Output $G(x) = \mathrm{sign}\left[\sum_{m=1}^{M} \alpha_m G_m(x)\right]$.

SVMs is robustness when dealing with problems with high-dimensional spaces.(Aworka et al., 2022). The algorithm uses the features of the dataset to predict or classify Preeclampsia according to the hyperplane. SVM, the hyperplane is used to separate data points of different classes. The objective of an SVM is to find the hyperplane that maximizes the margin between the classes. This helps in achieving the best generalization on unseen data.

$$f(x) = W^r X + b \quad (1)$$

Where w is the weight vector normal to the hyperplane and b is the bias term. The margin is the distance between the hyperplane and the nearest data points from either class (support vectors). The function of the nonlinear boundaries generated by the introduction of the kernel is given below as

$$f(x) = \sum_{i=1}^{n} k(X, X_i) + b \quad (2)$$

Naive Bayes (NB) is a family of probabilistic algorithms based on Bayes' Theorem, used primarily for classification tasks. The "naive" part comes from the assumption that all features (or attributes) in the dataset are independent given the class label, which simplifies the computation but might not always hold in real-world data. NB makes use of a function $f(x)$ which maps input value x into output class. As a classifier, NB forecast the predict the probability that an item belongs to a certain class (Salmi & Rustam, 2019). It uses the bayes theorem based on the following equation. the probability of Y given X can be expressed as seen in the equation below

$$P(Y|X) = \frac{P(X|Y)\,P(Y)}{P(X)}. \quad (3)$$

Modelling of Preeclampsia and evaluation

The dataset was split into 70% for training and 30% for testing. The model will be evaluated using the class-based error of precision, Recall, F-Measure and Receiver Operating Characteristic (ROC) area. ROC area measures the usefulness of a model. The other measures for the model will concentrate on the errors of prediction form the model. This metrics include Root Mean Square Error (RMSE), Root-Relative Square Error (RRSE) and Mean Absolute Error (MAE) and accuracy.

$$Precision = \frac{True_{\,positive}}{True_{\,positive} + False_{positive}} \quad (4)$$

$$Accuracy = \frac{True_{positive} + True_{negative}}{True_{positive} + True_{negative} + False_{positive} + False_{negative}} \quad (5)$$

$$Recall = \frac{True_{positive}}{(True_{positive} + False_{negative})} \quad (6)$$

$$F - Measure = \frac{2 \cdot (Precision \cdot Recall)}{Precision + Recall} \quad (7)$$

$$MAE = {}^{1}/_{n} \sum_{j=1}^{n} |y_j - \hat{y}_j| \quad ()$$

$$RMSE = \sqrt{{}^{1}/_{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \quad (9)$$

$$sMSE = \frac{1}{m} \sum_{i=1}^{m} |y_i - \hat{y}_i| \quad (10)$$

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2 \quad (11)$$

RESULTS AND DISCUSSION

The first part of the result covers the statistical analysis of the data and the correlation of the features of the data used in developing the models. Table 2 and figures 1 to 5 show some statistical characteristics of the data.

Table 2: Statistics of some features

	Minimum	Maximum	Mean	StdDeviation
Age	17	40	26.726	6.955
Parity	0	10	5.159	3.119
Order_Mar	1	2	1.048	0.214
Gest_age	22	39	33.091	5.18
Systolic BP	120	230	162.212	26.41
Diastolic BP	90	150	103.121	15.788
Weight	47	83	65.514	9.322

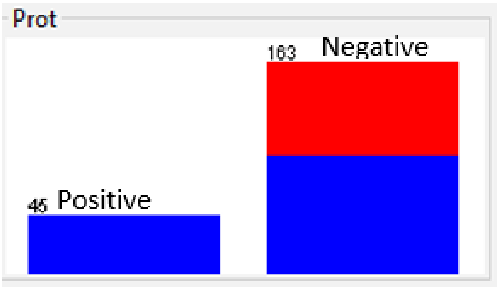

Figure 1: Summary of Proteinuria

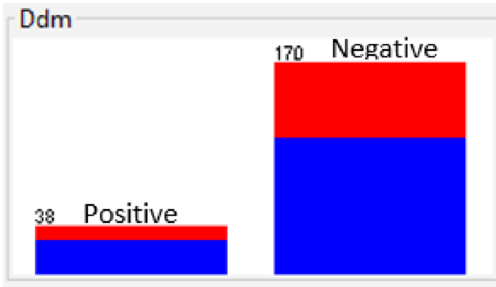

Figure 2: Summary of Oedema

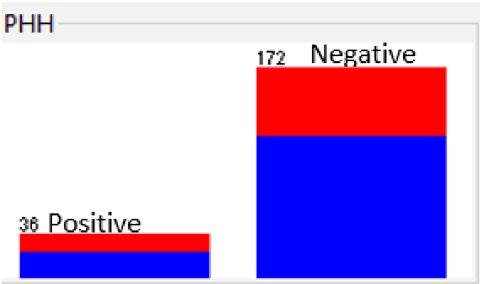

Figure 3: Previous history of hypertension

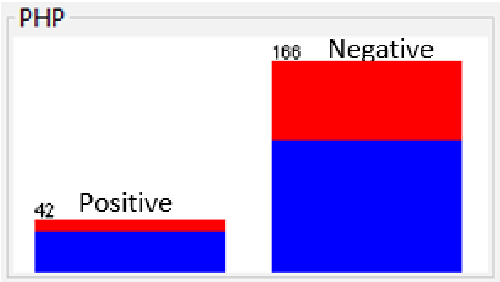

Figure 4: Previous history of PH

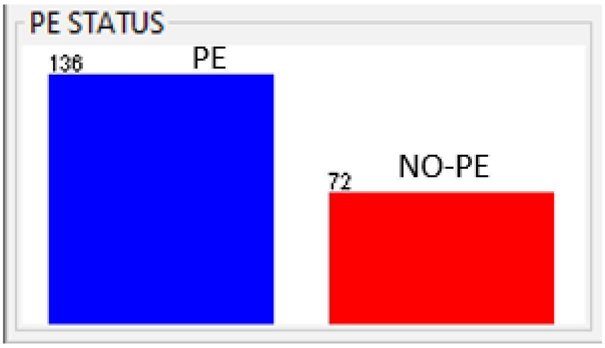

Figure 5: Status of preeclampsia

name	GesAge	Weight	PE-STATUS	SysPres	DiasPres	OrdMar	Age	Parity
GesAge	1							
Weight	0	1						
PE-STATUS	-0.18	0.34	1					
SysPres	0.03	0.42	0.65	1				
DiasPres	0.06	0.38	0.6	0.9	1			
OrdMar	0.07	0.11	0.16	0.15	0.21	1		
Age	-0.14	0.3	0.27	0.48	0.55	0.35	1	
Parity	-0.14	0.2	-0.04	0.19	0.23	0.28	0.73	1

Figure 6: correlation coefficient of some feature of Preeclampsia model

Some of the features with numerical values were evaluated for their ability to correectly predict the Preeclampsia. The correlation of various features of the model shows that systolic and diastolic values of blood pressure, age of the patient, order of marriage and weight of the patient all have positive correlation with the onset of Preeclampsia. The other variable like parity and gestational age has negative correlation, this is same with findings of Attahir et al.,(2010). The chart in figure 6 show a small positive (0.16) correlation between preeclampsia and order of marriage. This result is the same with the finding of Attahir et al.,(2010) which corelated and number of marriages. The implication of this small but positive corelation of order of marriage means that as women continue to change their sexual partners, there is higher chances of developing Preeclampsia for pregnancy with new partners.

The mean age of the women that came down with Preeclampsia is 17 years (table 1), which is closely related to the findings in the works of Abubakar et al., (2009) that mean age of Fulani women who had Preeclampsia in Borno was 18.9 years.

Classification and prediction performance evaluation

The result of performance evaluation of preeclampsia classification and prediction models developed using Adaboost, SVM and NB is presented in the Tables 2 and 3 and figures 7 and 8.

Table 2: Classification performance evaluation

	TP	FP	Precision	Recall	F-measure	ROC Area
Adaboost	0.98	0.01	0.98	0.97	0.98	1.00
SVM	1.00	0.00	1.00	1.00	1.00	1.00
NB	0.8	0.092	0.894	0.855	0.857	0.883

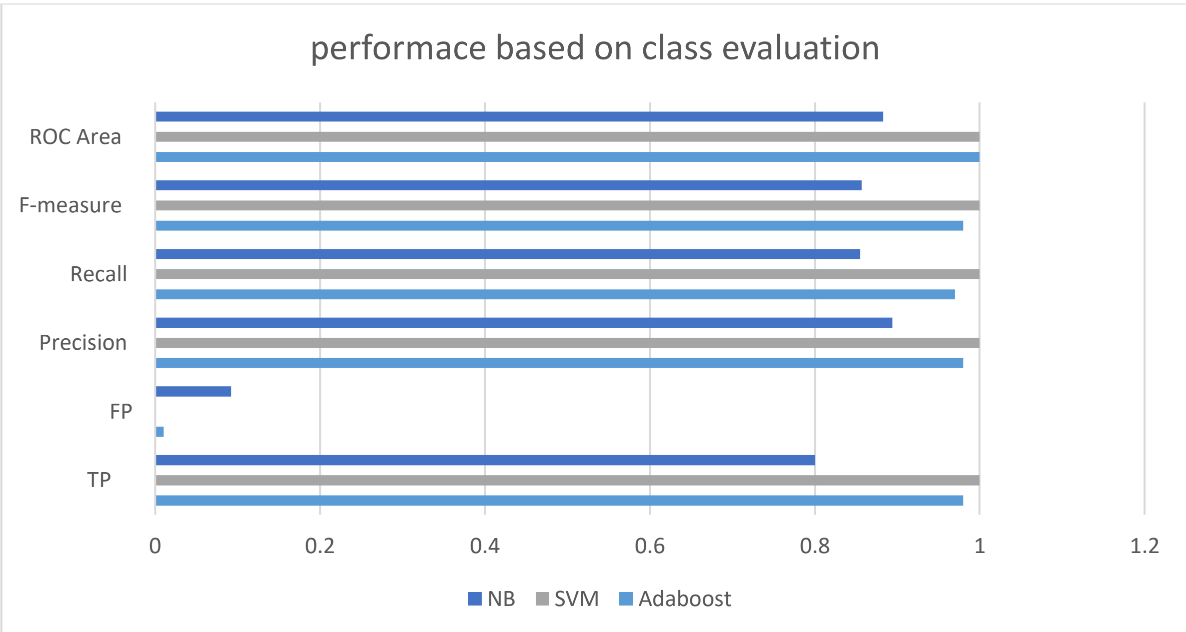

Figure 7: Classification performance chart for various models

The metrics used in measuring the performance of the models on the classification of Preeclampsia gave information on the accuracy and useability of the models developed in this paper. The SVM model demonstrated highest accuracy and performance in all the metrics. It was able to classify Preeclampsia with TP, Precision, Recall, F-measure, and ROC Area all equal to 1 while the FP is 0. This represents 100% accuracy level. The Adaboost model has TP, Precision, Recall, F-measure, and ROC values of 0.98, 0.01, 0.98, 0.97, 0.98, and 1.00 respectively. The result of the NB closely follows the Adaboost with TP of 0.8, FP of 0.092, Precision of 0.894, Recall of 0.855, F-measure of 0.857, and ROC of 0.883

The result of the models' performance show that SVM and Adaboost were able to classify Preeclampsia at 100% while NB was able to classify Preeclampsia at 88%. These results demonstrated that SVM and Adaboost are best in classifying Preeclampsia; however, based on these metrics, SVM is still superior to Adaboost.

The models developed in this work were also evaluated based on their ability to predict feature accuracy of Preeclampsia. The result of their performance is shown in table 3 and figure 8.

Table 3: Prediction performance evaluation

	MAE	RMSE	RAE	RRSE	Accuracy
Adaboost	0.05	0.15	12.97%	30.67%	98.38
SVM	0.00	0.00	0.00%	0.00%	100
NB	0.24	0.39	52.49%	81.49%	85.48

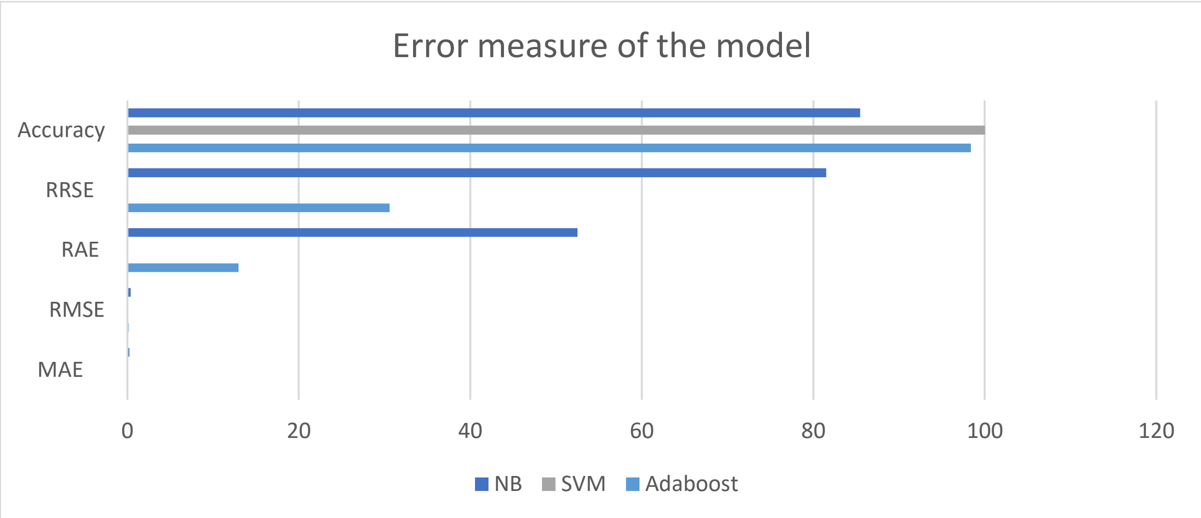

Figure 8: prediction performance chart for various models

The Adaboost, SVM, and NB models were each evaluated using MAE, RMSE, RAE, RRSE and Accuracy. From the data in figure 8 and Table 3, SVM demonstrated superior performance with MAE, RMSE, RAE, and RRSE values of 0 and accuracy values equal 100. On the other hand, the Adaboost followed closely with MAE, RMSE, RAE, RRSE, and accuracy values of 0.05, 0.15, 12.97%, 30.67%, and 98.38 respectively. NB model had the least performance values of with MAE, RMSE, RAE, RRSE, accuracy values of 0.24, 0.39, 52.49%, 81.49%, 85.48 respectively.

Discussion of Result

The developed models were each evaluated based on their performance on classification and prediction of Preeclampsia among pregnant women in who had their antenatal in Hunkuyi general hospital. From the figures 2, SVM outperformed Adaboost and NB in the classification of Preeclampsia. SVM achieved TP rate of 1 while Adaboost has TP of 0.98 and NB got TP of 0.8. This result show that SVM is accurate and precise in classifying Preeclampsia more than other models. Also, the result of precision and recall of all the models show that SVM was leading with values of 1 while Adaboost is second and NB third; however, the values indicate that they are strong models for classifying preeclampsia with high values of precision. The values of ROC, which represents the usefulness of the models reveal both SVM and Adaboost models have equal values of 100% while NB model has 85.48%. This suggests that SVM and Adaboost exhibited superior ability in accurately classifying Preeclampsia compared to NB.

A critical analysis of performance of the three models in predicting preeclampsia (table 3 and figure 8) further reveals that SVM demonstrated exceptional results. When tested in the prediction of preeclampsia, it had 0 error in MAE, RMSE, RAE, and RRSE. This result suggests a perfect model for prediction of preëclampsia with 100% accuracy and reliability. Furthermore, Adaboost error were also low with 0.05 and 0.15 for MAE and RMSE and accuracy level of 98%. This performance level is within exceptionally good prediction level since the error values are incredibly low. When compared with other models, higher levels of errors were observed in NB model when used to predict preeclampsia but with a very accuracy level of 85%.

A over-all interpretation of these performances suggests that SVM and Adaboost achieved excellent score of 100% in predicting preëclampsia, while NB got 85% accuracy. All the 3 models showed excellent level of performances in both classification and prediction of Preeclampsia, but SVM has the overall highest performance. These results suggest SVM, Adaboost and NB are good predictor of preeclampsia incidence.

When compared with the works of Maric et al., (2020), which predicted onset of preëclampsia with ROC of 0.79; the models developed in this paper are better with higher ROC of 1 for

SVM and Adaboost while that of NB has 0.88. This suggest these models are more useful than that of their work.

CONCLUSION

This paper developed machine learning-based classifiers and predictor of preeclampsia. The data used was collected from general hospital Hunkuyi, Kaduna state, Nigeria. The models were based on Adaboost, Support Vector Machine (SVM), and Naïve Bayes (NB) algorithms. The result of the corelation showed that parity and gestational age has negative correlation while other variables like blood pressure and order in marriage has positive. This shows that as women continue to change their sexual partners, there is higher chances of developing Preeclampsia for pregnancy with new partners. This result has grave implication for many women and should be a further research topic.

In both classification and prediction of preeclampsia among the population studied, SVM has 0 error in MAE, RMSE, RAE and ERSE, with accuracy level of 100%. Adaboost and NB had accuracy levels of 98% and 85%, which are very good. This paper recommends the use of these models for prediction of onset of preeclampsia among pregnant women in Hunkuyi and Kaduna state. Since the data used to develop the models represent impartially the various set of people within the state, it can be used for all women. we believe the models can assist the health personnel to predict onset of preeclampsia and help proper planning and intervention. It will also reduce maternal and child mortality that could result for preëclampsia

REFERENCES

Abubakar, A., Abdullahi, R. A., Jibril, H. Z., Dauda, M. H., & Poopola, M. A. (2009). Maternal Ethnicity and Severity of Pre-Eclampsia in Northern Nigeria. *Asian Journal of Medical Sciences 1(3):*, *1*(3), 104–107.

Akaba, G. O., Anyang, U. I., & Ekele, B. A. (2021). Prevalence and materno-fetal outcomes of preeclampsia / eclampsia amongst pregnant women at a teaching hospital in north- central Nigeria : a retrospective cross- sectional study. *Clinical Hypertension*, 1–10. https://doi.org/https://doi.org/10.1186/s40885-021-00178-y

Attahir, A., Dikko, A. A. U., Sufiyan, M. B., Salihu, A., & Rabiu, A. M. (2010). Association between Maternal Socio – economic Status , Polygamy and Risk of Pre – eclampsia in Rural Areas of Northern Nigeria. *Journal of Family and Reproductive Health*, *4*(1), 47–52.

Aworka, R., Cedric, L. S., Adoni, W. Y. H., Zoueu, J. T., Mutombo, F. K., Kimpolo, C. L. M., Nahhal, T., & Krichen, M. (2022). Agricultural decision system based on advanced machine learning models for yield prediction: Case of East African countries. *Smart Agricultural Technology*, *2*(December 2021), 100048. https://doi.org/10.1016/j.atech.2022.100048

Chaemsaithong, P., Pooh, R. K., Zheng, M., Ma, R., Tokunaka, M., Shaw, S. W., Seshadri, S., Choolani, M., Wataganara, T., Yeo, G. S. H., Wright, A., & Leung, W. C. (2019). Prospective evaluation of screening performance of first trimester prediction models for preterm preeclampsia in Asian population. *American Journal of Obstetrics and Gynecology*. https://doi.org/10.1016/j.ajog.2019.09.041

Chen, Y., Lin, T., Cheng, P., Chan, K., Huang, H., & Shaw, S. W. (2021). Taiwanese Journal of Obstetrics & Gynecology Taiwanese new direction in prediction of early pregnancy preeclampsia. *Taiwanese Journal of Obstetrics & Gynecology*, *60*(1), 66–69. https://doi.org/10.1016/j.tjog.2020.11.009

Chen, Y., Schisterman, E. F., Srinivasan, S., & Buck Louis, G. M. (2020). Deep learning in the prediction of preeclampsia: A novel approach. *Scientific Reports*, 10(1), 1-10.

Dathan-stumpf, A., Czarnowsky, V., Hein, V., Andraczek, T., & Stepan, H. (2020). Real-world data on the clinical use of angiogenic factors in pregnancies with placental dysfunction. *The American Journal of Obstetrics & Gynecology*, *December 2017*, 1–13. https://doi.org/10.1016/j.ajog.2020.10.028

E.V., K., C., A., & M., C. (2009). Hypertensive disorders and severe obstetric morbidity in the United States. *Obstet Gynecol Ogy*, *113*, 299–306.

Enaruna, N. O., & Sodje, J. D. (2015). Early-Onset Severe Preeclampsia in Nigerian Women : Determining a Balance between Maternal Wellbeing and Fetal Survival in a Resource-Limited Setting. *British Journal of Medicine & Medical Research*, *10*(3), 1–9. https://doi.org/10.9734/BJMMR/2015/19217

Hastie, T., Tibshirani, R., & Friedman, J. (2017). The elements of statistical learning: data mining, inference, and prediction, PP 342-348.

Johnson, A. C., Mendoza, M., Caplice, S., & Brennecke, S. P. (2019). Recurrent neural networks in the prediction of preeclampsia onset. Scientific Reports, 9(1), 1-9.

Li, J., Li, G., Hai, C., & Guo, M. (2022). Transformer Fault Diagnosis Based on Multi-Class AdaBoost Algorithm. *IEEE Access*, *10*, 1522–1532. https://doi.org/10.1109/ACCESS.2021.3135467

Li, X., Wang, L., & Sung, E. (2008). AdaBoost with SVM-based component classifiers. *Engineering Applications of Artificial Intelligence*, *21*(5), 785–795. https://doi.org/10.1016/j.engappai.2007.07.001

Lu, M. Y., Blanchard, C. T., Oglesby, K. R., Page, M. R., Lazenby, A. J., Elizabeth B. Ausbeck1, G. D. C., Rogers, R. D. D. M., Bushman, E. T., Kaplan, E. R., Ruzic, M. F., Mahalingam, M., Dunk, S. A., Champion, M. L., Casey, B. M., Tita, A., Kim, D.-J., Szychowski, J. M., & Subramaniam, A. (2021). A prediction model for postpartum readmission for hypertension or preeclampsia. *American Journal of Obstetrics & Gynecology*, *February*, 41–42. https://doi.org/10.1016/j.ajog.2020.12.057

Maric, I., Tsur, A., Aghaeepour, N., Montanari, A., Stevenson, D. K., Shaw, G. M., & Winn, V. D. (2020). Early prediction of preeclampsia via machine learning. *Am J Obstet Gynecol*, *May*, 1–17. https://doi.org/10.1016/j.ajogmf.2020.100100

Mathanker, S. K., Weckler, P. R., Bowser, T. J., Wang, N., & Maness, N. O. (2011). AdaBoost classifiers for pecan defect classification. *Computers and Electronics in Agriculture*, *77*(1), 60–68. https://doi.org/10.1016/j.compag.2011.03.008

Melchiorre, K., Giorgione, V., & Thilaganathan, B. (2020). The placenta and preeclampsia : villain or victim ? *The American Journal of Obstetrics & Gynecology*, 1–9. https://doi.org/10.1016/j.ajog.2020.10.024

Mohammed, C., Adze, J., Bature, S., Mohammed-Durosinlorun, A., Taingson, C., Abubakar, A., Airede, L. ., Onwuhafua, P., & Musa, J. (2023). Severe Preeclampsia and Eclampsia: A 3-Year Review in Barau Dikko Teaching Hospital Kaduna, Northwest Nigeria. *J Med*

Mou AD, Barman Z, Hasan M, Miah R, Hafsa JM, Das Trisha A, Ali N (2021): Prevalence of preeclampsia and the associated risk factors among pregnant women in Bangladesh. Sci Rep. 2021, 11:21339.10.1038/s41598-021- 00839-w

Myatt, L. (2020). Expert Review The prediction of preeclampsia : the way forward. *The American Journal of Obstetrics & Gynecology*. https://doi.org/10.1016/j.ajog.2020.10.047

Ngwenya, S., Jones, B., Mwembe, D., Nare, H., Alexander, E., & Heazell, P. (2021). Pregnancy Hypertension : An International Journal of Women ' s Cardiovascular Health Development and validation of risk prediction models for adverse maternal and neonatal outcomes in severe preeclampsia in a low-resource setting , Mpilo Central Hospital . *Pregnancy Hypertension: An International Journal of Women's Cardiovascular Health*, *23*(November 2020), 18–26. https://doi.org/10.1016/j.preghy.2020.10.011

Okunade, K. S., Gbadegesin, A., & Akinsola, O. J. (2014). Risk Factors for Pre-eclampsia in Multiparous Women in Lagos , Nigeria. *International Journal of Biomedical Research*, *9633*, 3–6. https://doi.org/10.7439/ijbr

Ola, D., & Suliburska, J. (2023). Risk factors of preeclampsia in Nigeria and Poland. *J Obstet Gynecol Investig*, *May*, 7–14. https://doi.org/https://doi.org/10.5114/jogi.2023.130145

Portelli, M., & Baron, B. (2018). Clinical Presentation of Preeclampsia and the Diagnostic Value of Proteins and Their Methylation Products as Biomarkers in Pregnant Women with Preeclampsia and Their Newborns. *Journal of Pregnancy*, *2018*(Table 1).

Rana S, Lemoine E, Granger JP, Karumanchi SA (2019). Preeclampsia: pathophysiology, challenges, and perspectives. Circ Res., 124:1094-112. 10.1161/CIRCRESAHA.118.313276

Salmi, N., & Rustam, Z. (2019). Naïve Bayes Classifier Models for Predicting the Colon Cancer. *IOP Conf. Series: Materials Science and Engineering*, *546*. https://doi.org/10.1088/1757-899X/546/5/052068

Schindler, A. E., & Schindler, A. E. (2018). New data about preeclampsia : some possibilities of prevention. *Gynecological Endocrinology*, *34*(8), 636–637. https://doi.org/10.1080/09513590.2018.1441401

Smith, G. C., Seaman, S. R., Wood, A. M., Royston, P., & White, I. R. (2017). Correcting for optimistic prediction in small data sets. American Journal of Epidemiology, 186(6), 777-784.

Suleman, B. ., Olayemi, O. ., Adesina, O. ., & Bankole, A. . (2022). Combination Of Spot Urine Protein-Creatinine Ratio and Uterine Artery Doppler in Predicting Preeclampsia in Ibadan : A Prospective Cohort Study. *African Journal of FetoMaternal Medicine*, *1*(1), 31–37.

Tarca, A. L., Mph, A. T., Dmedsci, R. R., Jung, E., Paredes, C., Bhatti, G., Ghita, C., Chaiworapongsa, T., Than, N. G., & Mph, C. H. (2021). maternal characteristics and biophysical and. *The American Journal of Obstetrics & Gynecology*. https://doi.org/10.1016/j.ajog.2021.01.020

Ugwu AO, Owie E, Oluwole AA, Soibi-Harry AP, Garba SR, Okunade KS, Makwe CC, Omisakin SI, Ani-Ugwu NK, Ojiefoh LG, Okafor IT, Egba A, Olatunji RA, Yusuf-Awesu SA (2022). Maternal and perinatal outcomes of preeclampsia at a tertiary hospital in lagos, Nigeria. Int J Med Health Dev [serial online] 2022 ;27:197-200. Available from: https://www.ijmhdev.com/text.asp?2022/27/2/197/339034

Venkatesh, K. K., Beckham, J., Friedman, A. M., Boggess, K., Jelovsek, E., & Stamilio, D. M. (2021). A prediction model for postpartum readmission for hypertension or preeclampsia. *The American Journal of Obstetrics & Gynecology*, *224*(2), S42–S43. https://doi.org/10.1016/j.ajog.2020.12.058

VANDERJAGT, D. J., PATEL, R. J., EL-NAFATY, A. U., MELAH, G. S., CROSSEY, M. J., & GLEW, R. H. (2004). High-density lipoprotein and homocysteine levels correlate inversely in preeclamptic women in northern Nigeria. *Acta Obstet Gynecol Scand*, *83*, 536–542. https://doi.org/10.1111/j.1600-0412.2004.00513.x

Wang, W., & Sun, D. (2021). The improved AdaBoost algorithms for imbalanced data classification. *Information Sciences*, *563*, 358-374.

Weissgerber TL, Mudd LM (2015): Preeclampsia and diabetes. Curr Diab Rep. 2015, 15:9. 10.1007/s11892-015-0579-4

Wright, D., Wright, A., & Nicolaides, K. H. (2019). THE COMPETING RISK APPROACH FOR PREDICTION OF PREECLAMPSIA. *American Journal of Obstetrics and Gynecology*. https://doi.org/10.1016/j.ajog.2019.11.1247