Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Perinatal Mortality Prediction and Risk Factor Identification Using Machine Learning on Recent Sub-Saharan African DHS Data Affiliations

**Tadele Chekol Maru**

tadelle2007@gmail.com

University of Gondar

**Andualem Enyew**

University of Gondar

**Makda Fekadie Tewelgne**

University of Gondar

**Eliyas Addisu Taye**

University of Gondar

**Agerie Mengistie Zeleke**

Debark University

**Belayneh Jejaw Abate**

University of Gondar Comprehensive Specialized Hospital

**Deresse Abebe Gebrehana**

University of Gondar

**Azanaw Amare Muche**

University of Gondar

**Research Article**

**Keywords:** Perinatal mortality, machine learning, maternal education, parity, Sub-Saharan Africa, Demographic and Health Survey

**Posted Date:** March 30th, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9092921/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** No competing interests reported.

## Abstract

# Background

Perinatal mortality stillbirths after 28 weeks of gestation and early neonatal deaths within seven days of birth persists as a major public health challenge in Sub-Saharan Africa. Despite global declines in neonatal mortality, progress remains uneven in low-resource settings. Limited use of advanced analytics, especially machine learning, has hindered predictive accuracy and insights into complex risk factors.

# Methods

This cross-sectional study analyzed pooled Demographic and Health Survey data from 19 Sub-Saharan African countries (2016–2023), including a weighted sample of 802,470 women aged 15–49. Seven supervised machine learning algorithms (Random Forest, Decision Tree, Logistic Regression Voting Classifier, XGBoost, Naïve Bayes, Stacking, LightGBM) predicted perinatal mortality. Synthetic Minority Over-sampling Technique (SMOTE) addressed class imbalance. Performance was assessed via accuracy, precision, recall, F1-score, and AUC-ROC; SHAP analysis interpreted feature importance.

# Results

Perinatal mortality prevalence was 5.2%, highest in Côte d'Ivoire (9.1%), Tanzania (7.5%), and Ghana (7.0%), lowest in Gabon (2.5%), Zambia (3.4%), and Madagascar (3.6%). Random Forest excelled with 81.7% accuracy, 80% precision, 84% recall, 82% F1-score, and 0.91 AUC-ROC. SHAP analysis highlighted maternal age, education, and parity (number of living children) as top predictors. Higher maternal education offered strong protection; first and high-order births elevated risk.

# Conclusion

Random Forest effectively predicted perinatal mortality. Low maternal education was the primary risk factor, with parity showing a U-shaped pattern. Prioritizing female education, antenatal care quality, and high-risk parity groups could substantially reduce regional perinatal deaths.

## Background

Perinatal mortality includes death within seven days of birth and any fatality that occurs at or after 28 weeks of pregnancy stillbirth(1). In many developing nations, these untimely deaths are a public health issue that has a significant impact on families, society, and the economy(2).

In the world, there are more than 5 million perinatal deaths per year, and many of them are avoidable (3).Ten and a half million children under five pass away every year, with four million of those deaths

occurring in the first four weeks of life. Stillbirths total 3.3 million more.

In 2019, the World Health Organization (WHO) reported that out of 2.6 million newborns worldwide, more than 8,200 died within their first day. Of the 133 million newborns born each year, 2.8 million die in the first week of life or at birth, and the majority of these deaths take place in developing countries(4).

In comparison to high-income countries, the risk of newborn death is nine times higher, while the risk of stillbirth is 7.6 times higher in low-income countries(5).

Furthermore, with 27 deaths per 1000 live births in 2017, sub-Saharan Africa had the highest neonatal mortality rate among the regions designated by the Sustainable Development Goals. Southern Asia, Oceania, Latin America, and the Caribbean had rates of 26, 21, and 10 deaths per live birth, respectively(6–8).

In East Africa, perinatal mortality remains a significant issue, with rates varying between countries and subregions. Eastern Africa has the highest pooled perinatal mortality rate at 34.5 per 1000 births (9). Globally, the newborn mortality rate decreased from 37 deaths per 1000 live births in 1990 to 19 in 2016, while the average stillbirth rate per 1000 total births decreased from 24.7 in 2000 to 18.4 in(10, 11).The international community encourages low- and middle-income nations to keep up their commitment to lowering maternal and infant mortality rates by expanding access to maternal, neonatal, and perinatal health services, with the Millennium Development Goals deadlines rapidly approaching(12).The region was committed to achieving the Sustainable Development goals, and it has been working to lower newborn mortality to less than 12 per 1000 live births by 2030(13). Nevertheless, the reduction of neonatal, infant, and under-five mortalities cannot be realized without a substantial reduction of perinatal mortality(14).

Various risk factors for neonatal mortality have been found in numerous studies, such as low birth weight, preterm, prior history of perinatal deaths, insufficient iron supplementation, and inadequate tetanus toxoid vaccination(15). furthermore, Factors influencing perinatal outcomes include parity(16, 17), maternal age(18), maternal education(19, 20), attendance of antenatal care visits (ANC)(21), presence of a skilled birth attendant(22), previous history of perinatal mortality(23), low family income, birth interval(24), the total number of children under five(23), place of delivery(25), and residency(18). Although the alarmingly high incidence of perinatal mortality in East Africa, our review of the literature reveals significant gaps in the methodology of the determinants of prenatal health and death in resource-limited settings. A significant barrier to the global attempts to instantly reduce perinatal and under-five child mortality is the absence of the application of machine learning. To overcome these problems, machine learning algorithms perform better than conventional approaches, allowing for quicker and more accurate results and learning provides new tools to address issues that traditional statistical methods are not well adapted. Therefore, this study aimed to address the method gap in predicting perinatal deaths in East Africa by developing the best model using machine learning algorithms.

# Methods

# Design

This investigation adopted a cross-sectional analytical approach, drawing on data from the Demographic and Health Surveys (26) implemented across 19 sub-Saharan African nations. As a standardized, nationally representative survey program, the DHS systematically collects health and demographic information in low- and middle-income countries. The weighted analytical sample consisted of 802,470 women of reproductive age (15–49 years). Data preprocessing and subsequent analyses were performed using Python (version 3.12). To predict perinatal mortality, a range of supervised machine learning algorithms—including Random Forest, Decision Tree, Logistic Regression-based voting classifier, Extreme Gradient Boosting (XGBoost), Naïve Bayes, stacking ensemble, and LightGBM—were employed. LightGBM was prioritized for its computational efficiency in processing large-scale, high-dimensional datasets (13). The predictive performance of the models was evaluated using standard classification metrics: accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC) (9).

## Study period and setting

The study used data from recent surveys conducted in 19 Sub–Sharan African countries between 2016 and 2023 G.C. This investigation compiles individual-level data from multiple countries to examine women who experienced perinatal loss. The pooled dataset, weighted to ensure representativeness, comprises 802,470 cases. All data utilized in this research are publicly accessible through the Demographic and Health Surveys (26) program repository at https://dhsprogram.com/data/available-datasets.cfm.

## Study Area

This investigation examines perinatal outcomes across Sub-Saharan Africa's extensive geographic expanse, targeting 19 nations: Burundi, Côte d'Ivoire, Ethiopia, Gabon, Ghana, Gambia, Kenya, Liberia, Lesotho, Madagascar, Malawi, Mozambique, Rwanda, Sierra Leone, Senegal, Tanzania, Uganda, and Zambia. Secondary data analysis was conducted using Demographic and Health Surveys (26) as the primary data source. Among Sub-Saharan African countries, 19 maintain complete and accessible DHS information; the remaining four possess either outdated datasets or restricted access (including Eritrea). Consequently, this study incorporated the most recent standardized DHS data from 19 countries meeting all quality and accessibility criteria.

## Target Population

The source population comprised all women of reproductive age in sub-Saharan Africa who had given birth within the five years preceding the respective Demographic and Health Surveys. The study population was restricted to women residing in the selected enumeration areas across the nineteen included countries, who provided complete information on delivery method during the survey period.

## Eligibility Criteria

Women aged 15–49 with ongoing pregnancies of 28 + weeks (seven or more months) at the time of the survey, within the five-year period preceding the survey, were included. Women with pregnancy losses before seven months (e.g., abortions) or live births occurring outside the five-year window preceding the survey were excluded.

# Sample Size and Sampling Procedure

# Sample Size Determination

A total weighted sample of 802,470 women of reproductive age was drawn using the Demographic and Health Surveys' two-stage stratified cluster sampling design. In the first stage, enumeration areas were selected with probability proportional to size. In the second stage, a fixed number of households were systematically selected from each enumeration area. This sampling approach ensured national representativeness and permitted disaggregation of estimates by region, urban-rural residence, and socioeconomic characteristics (15).

## Sampling Procedure

This study utilized data from the Demographic and Health Surveys (26), which employ a stratified multi-stage cluster sampling design. Enumeration areas were initially selected from census regions (strata) using stratified random sampling, followed by the systematic random selection of households within each enumeration area. Data were collected via standardized questionnaires covering various aspects of reproductive health, including the mode of delivery. The 19 sub-Saharan African countries included in this analysis were selected based on the completeness and recency of their DHS data (16) as shown in Fig. 1 and Table 1

Table 1
Countries' years of survey and sample size (annex

| Countries | Years of survey | Unweighted sample | Weighted sample |
|---|---|---|---|
| Burundi | 2016-17 | 48,745 | 45,419 |
| Ethiopia | 2016 | 90,838 | 41,392 |
| Kenya | 2022 | 39,979 | 77,381 |
| Madagascar | 2021 | 41,392 | 47,720 |
| Malawi | 2015-16 | 23,232 | 68,074 |
| Mozambique | 2022-23 | 34,663 | 33,900 |
| Rwanda | 2019-20 | 31,665 | 30,820 |
| Tanzania | 2022 | 77,381 | 40,394 |
| Uganda | 2016 | 24,765 | 57,906 |
| Zambia | 2018 | 10,445 | 38,446 |
| Burkina Fasso | 2021 | 47,720 | 48,745 |
| Cote d'Ivoire | 2021 | 68,074 | 39,979 |
| Gabon | 2019 | 33,900 | 23,232 |
| Ghana | 2022 | 30,820 | 34,663 |
| Gambia | 2019 | 40,543 | 31,665 |
| Liberia | 2019 | 21,562 | 24,765 |
| Lesotho | 2023 | 40,394 | 10,445 |
| Sierra Leone | 2019 | 57,906 | 40,543 |
| Senegal | 2019 | 38,446 | 21,562 |
| Total | | 1,279,054 | 802,470 |

Source demographic and health survey dataset survey 2016–2023

# Variables and Operational Definitions

## Study Variable

The primary outcome variable in this study was perinatal mortality, defined as the composite of stillbirths specifically, fetal deaths occurring after seven completed months of gestation—and early neonatal deaths, referring to the death of live-born infants within the first seven days following birth(27).

Perinatal death was operationalized as a binary variable, coded as “yes” for cases in which death occurred during the perinatal period and “no” for all other outcomes.

A total of 25 predictor variables were included in the machine learning models, selected on the basis of prior evidence from the literature(28, 29), which are available in DHS dataset. These encompassed maternal and paternal sociodemographic characteristics, including age (in years), area of residence (rural or urban), educational status (highest level of education attained (no, primary, secondary, or higher), marital status (single, married, or widowed/divorced), and occupation (unemployed, employed), sex of the house hold leader( male, female), sex of child(male, female) place of delivery(home, Health institution) and clinical characteristics antenatal care visit status, Postnatal care visits status, numbers of living children, current family planning status, numbers of tetanus toxoid vaccine injection taken during pregnancy, and mode of and delivery.

# Data Sources and Tools

The analysis drew on data from The DHS Program database (https://dhsprogram.com), with the primary source being the birth record components of Demographic and Health Surveys implemented in ten sub-Saharan African nations. As these surveys utilize standardized questionnaires, they facilitate valid and reliable cross-country comparisons.(26).

# Data and Management Preprocessing

Data processing and analysis were carried out using STATA version 17 and Python. Following initial data management in STATA, the datasets were converted to. DTA format and imported into Python. Within this environment, data manipulation was facilitated by Pandas and NumPy, and visualizations were generated using Matplotlib and Seaborn. The model development phase relied on the Scikit-learn and random forest libraries. The preprocessing strategy included mean imputation for missing data, the removal of outliers, and the application of the Synthetic Minority Over-sampling Technique (SMOTE) to address class imbalance where necessary(30).

# Data Analysis Framework

The analytical approach adhered to stage machine learning workflow delineated by Guo (2018). This framework encompasses data cleaning, feature selection, model selection, training, validation, hyperparameter tuning, testing, and interpretation(31).

# Feature Engineering

As part of the data preprocessing pipeline, categorical variables were transformed via one-hot encoding. Subsequently, feature normalization was applied to standardize all variables to a common scale, thereby facilitating unbiased comparisons across different algorithms.

# Machine learning models

This schematic depicts the sequential analytical process, beginning with data acquisition from the Demographic and Health Surveys (2016–2023) and proceeding through data preprocessing, model development using seven classifiers, performance assessment, optimal model identification via hyperparameter tuning, and concluding with validation on a 20% test sample shown in Fig. 2.

# Model Evaluation

The performance of the predictive models was evaluated using both train-test split and cross-validation techniques. Model performance on the test set was assessed using standard classification metrics, including accuracy, precision, recall, F-measure, and the receiver operating characteristic (ROC) curve. To provide a detailed account of classification outcomes, a confusion matrix was generated for each model. This N × N matrix (where N represents the number of classes) tabulates the counts of correct and incorrect predictions, thereby illustrating the model's ability to correctly identify cases of perinatal mortality versus instances of misclassification.

Model performance was assessed using metrics derived from the confusion matrix, which tabulates the counts of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions. In the context of this study, true positives refer to cases of perinatal mortality that were correctly classified by the model. True negatives denote non-mortality cases that were correctly identified as such. False positives represent instances where the model incorrectly predicted perinatal mortality for a non-case, while false negatives indicate actual mortality cases that were misclassified as negative (32).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

Model performance was evaluated using several standard classification metrics derived from the confusion matrix.

### Recall (Sensitivity)

Recall, also referred to as sensitivity or the true positive rate, measures the proportion of actual positive cases that were correctly identified by the model. It quantifies the model's ability to detect positive events and was calculated as follows (33)

$$\text{Recall}(\text{Sensitivity}, \text{truepositiverate}) = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

### Precision (Positive Predictive Value)

Precision, or positive predictive value, assesses the accuracy of positive predictions by measuring the proportion of predicted positive cases that were truly positive. It indicates how many of the cases identified as high-risk for perinatal mortality were correctly classified (34). The formula used was

$$\text{Precision}\,(\text{Positivepredictivevalue}) = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

### F1-Score

The F1-score represents the harmonic mean of precision and recall, providing a balanced measure that accounts for both false positives and false negatives. It is particularly useful when dealing with imbalanced classes, as a higher F1-score indicates better model performance (35). It was computed as

$$\mathrm{F1score} = \frac{2*\mathrm{Precision}*\mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

### Area Under the Receiver Operating Characteristic Curve (AUC-ROC)

The Receiver Operating Characteristic (ROC) curve illustrates the trade-off between sensitivity (true positive rate) and specificity (true negative rate) across different classification thresholds. The area under this curve (AUC) provides a single measure of the model's ability to discriminate between positive and negative classes. AUC values range from 0 to 1, with higher values indicating better discriminatory power. In this study, model efficiency was assessed using the AUC, where values closer to 1.0 signify superior predictive performance (36).

## Model Development and Evaluation

The dataset was randomly partitioned into training (80%) and testing (20%) subsets.

To enhance model robustness and generalizability, 10-fold cross-validation was applied during the training phase. Model performance was subsequently assessed on the test set using standard classification metrics, including accuracy, precision, recall, F1-score, and the area under the receiver operating characteristic curve (AUC)(32).

## Model Interpretability using SHAP

To improve the interpretability of the model outputs, shapley Additive exPlanations (SHAP) values were calculated. This game-theoretic framework quantifies the contribution of each feature to individual predictions, thereby enhancing transparency and illuminating the underlying decision-making processes of the machine learning models(33).

## Ethical Considerations

Permission to access and utilize the data was sought from the Demographic and Health Surveys Program. This was accomplished by submitting the study title and objectives through their online request platform, which resulted in the issuance of an official authentication letter authorizing data access(34).

## Results

A total of 802,470 weighted participants were included in the analysis to predict perinatal mortality and identify its determinants in sub-Saharan Africa (SSA). The largest contributors to the sample were

Burundi (90,838; 11.3%), Kenya (77,381; 9.6%), Malawi (68,074; 8.3%), and Uganda (57,906; 7.2%). Conversely, the smallest sample sizes were observed in Lesotho (10,445; 1.3%), Senegal (21,562; 2.6%), Gabon (23,232; 2.9%), and Liberia (24,765; 3.1%). With respect to maternal age, women aged 35–49 years constituted the largest proportion (57.2%). The majority of participants resided in rural areas (72.2%), while 27.8% lived in urban settings. Regarding educational attainment, 305,982 participants (40.4%) had no formal education, while 148,064 (37.4%) had completed primary education, and only 21,961 (2.9%) had attained higher education. Furthermore, approximately 362,263 participants (47.9%) were from households classified as having poor wealth status. A detailed summary of participant characteristics is presented in 2 and Table 3.

Table 2
Frequency distribution of weighted samples from selected sub-Saharan countries, 2016–2023 (N = 802470)

| Country | Pregnancy loss and early neonatal death | | | |
|---|---|---|---|---|
| | Perinatal death | | | |
| Yes | No | Mortality % | Total | |
| BurkinaFasso | 3,037 | 45,708 | 6.2 | 48,745 |
| Burundi | 4,342 | 86,496 | 4.8 | 90,838 |
| Cote d'Iovore | 3,645 | 36,334 | 9.1 | 39,979 |
| Ethiopia | 2,085 | 39,307 | 5.0 | 41,392 |
| Gabon | 581 | 22,651 | 2.5 | 23,232 |
| Ghana | 2,429 | 32,234 | 7.0 | 34,663 |
| Gambia | 1,701 | 29,964 | 5.4 | 31,665 |
| Kenya | 4,739 | 72,642 | 6.1 | 77,381 |
| Liberia | 1,332 | 23,433 | 5.4 | 24,765 |
| Lesotho | 658 | 9,787 | 6.3 | 10,445 |
| Madagascar | 1,711 | 46,009 | 3.6 | 47,720 |
| Malawi | 2,568 | 65,506 | 3.8 | 68,074 |
| Mozambique | 2,133 | 31,767 | 6.3 | 33,900 |
| Rwanda | 1,191 | 29,629 | 3.9 | 30,820 |
| Sierra Leone | 1,650 | 38,893 | 4.1 | 40,543 |
| Senegal | 1,054 | 20,508 | 4.9 | 21,562 |
| Tanzania | 3,048 | 37,346 | 7.5 | 40,394 |
| Uganda | 2,576 | 55,330 | 4.4 | 57,906 |
| Zambia | 1,296 | 37,150 | 3.4 | 38,446 |
| Total | 41,776 | 760,694 | 5.2 | 802,470 |

Table 3
Sociodemographic characteristics of the study participants in the Sub-Saharan Africa demographic and health survey, 2016 – 203

| Variables | Category | Frequency | Percentage |
|---|---|---|---|
| Age | 15–24 | 72,032 | 9.5 |
| | 25–34 | 252,258 | 33.3 |
| | 35–49 \| | 432,761 | 57.2 |
| Residence | Urban | 210,847 | 27.9 |
| | Rural | 546,204 | 72.1 |
| Respondent education status | No education | 330,850 | 41.1 |
| | primary education | 314,183 | 39.2 |
| | secondary education | 135,179 | 16.9 |
| | higher education | 22,258 | 2.8 |
| Husband educational status | No education<br>primary education | 347,939 | 41 |
| | secondary education | 129,404 | 80 |
| | higher education | 55,180 | 7.29 |
| | Male | 559,719 | 73.93 |
| Sex of the house hold | Female | 197,332 | 26.07 |
| | Married | 631,903 | 83.47 |
| Marital status | Single | 125,148 | 16.53 |
| | poor | 362,263 | 47.85 |
| Wealth status | Middle | 152,129 | 20.09 |
| | Rich | 242,659 | 32.05 |

# Maternal and prenatal-related characteristics

The study comprised a weighted sample of 802,470 women. Regarding antenatal care (ANC) utilization, the vast majority of participants (703,190; 92.9%) attended at least four ANC visits, while only 6,841 women (0.9%) had no ANC contact during their pregnancy. With respect to postnatal care (PNC), the proportion of women who received a postnatal check-up within 24 hours of delivery was 3,055 (0.4%). In contrast, 741,202 women (92.4%) received their first postnatal care after seven days or later as shown in Table 4.

Table 4
Maternal health and fertility-related factors *in* the Sub-Saharan Africa demographic and health survey, 2016 – 203

| Variables | Category | Frequency | Percentage |
|---|---|---|---|
| ANC visit | no visit | 6,841 | 0.90 |
| | 1–4 | 703,190 | 92.89 |
| | 4 and above | 47,020 | 6.21 |
| Place of delivery | home delivery | 38,368 | 5.07 |
| | institutional delivery | 718,683 | 94.93 |
| delivery by caesarean section | No | 745,587 | 98.49 |
| | Yes | 11,464 | 1.51 |
| sex of child | male | 383,743 | 50.69 |
| | female | 373,308 | 49.31 |
| multiple birth | single birth | 733,761 | 96.92 |
| | 1st multiple \| | 11,579 | 1.53 |
| | 2nd multiple | 11,577 | 1.53 |
| | 3rd multiple \| | 121 | 0.02 |
| number of living children | zero | 2,780 | 0.37 |
| | one-Two | 145,052 | 19.16 |
| | Three and above | 609,219 | 80.47 |
| Postnatal care | Within 24 hrs. | 3,055 | 0.40 |
| | 25–48 hrs. | 51 | 0.01 |
| | 49–72 hrs. | 5,067 | 0.67 |
| | 73hrs-7days | 7,676 | 1.01 |
| | After 7days | 741,202 | 97.91 |
| pregnancy, given or bought iron tablets/syrup | No | 20,683 | 2.73 |
| | Yes | 736,082 | 97.23 |
| | Unknown | 286 | 0.04 |
| tetanus injections before birth | Less than 2 | 63,418 | 8.38 |
| | greater than 2 | 693,633 | 91.62 |

| Variables | Category | Frequency | Percentage |
|---|---|---|---|
| ANC visit | no visit | 6,841 | 0.90 |
| | 1–4 | 703,190 | 92.89 |
| anemia level | severe | 4,369 | 0.58 |
| | Moderate | 56,971 | 7.53 |
| | Mild | 77,600 | 10.25 |
| | Not Anemic | 618,111 | 81.65 |

# Machine learning Analysis

# Balancing data set

Descriptive statistics revealed that the prevalence of prenatal mortality among births was 5.2%, indicating a class imbalance where approximately 95% of observations fell into the majority class. To address this and support the development of reliable predictive models, the dataset was balanced by adjusting the class distribution to 50 instances per category, resulting in a symmetric distribution. The selected machine learning models were then trained on this balanced dataset using an 80:20 train-test split and 10-fold cross-validation, as illustrated in Fig. 3.

# SMOTE oversampling technique

To address the class imbalance inherent in the dataset, the Synthetic Minority Over-sampling Technique (SMOTE) was implemented. This method augments the minority class by generating synthetic instances to achieve parity with the majority class, thereby producing a more balanced class distribution, as depicted in Fig. 3

# Model evaluation and ROC analysis

The confusion matrix provides a comprehensive breakdown of the model's predictions, detailing the counts for true positives, true negatives, false positives, and false negatives.

False Positives (FP): 26,523 – These are cases where the model incorrectly predicted perinatal mortality for babies who did not actually experience it.

False Negatives (FN): 20,529 – These are cases where the model failed to detect perinatal mortality, predicting "no mortality" for babies who did experience it.

True Negatives (TN): 102,126 – These are cases where the model correctly predicted that no mortality occurred.

True Positives (TP): 108,120 – These are cases where the model correctly identified the presence of perinatal mortality.

The high number of true positives indicates that the model is effective at detecting perinatal mortality when it occurs. Overall, the model achieved an accuracy of 81.7**%**, meaning it correctly classified the outcome in more than 8 out of every 10 cases. While this reflects solid performance overall, it is important to interpret accuracy alongside the class distribution shown in the confusion matrix as shown in Fig. 4.

**Comparisons of model performance metrics**

**figure 4: confusion matrix**

The performance of seven machine learning models was evaluated using Accuracy, Precision, Recall, F1 Score, and ROC/AUC on both imbalanced and balanced datasets. Balancing the data led to improved performance across most models. Random Forest demonstrated consistent results across all metrics, achieving an AUC of 91.9%, an F1 score of 82%, accuracy of 81.7%, precision of 80%, and recall of 84%. In contrast, Naïve Bayes showed comparatively lower performance, with an AUC of 59.4%, F1 score of 65.5%, accuracy of 54%, and precision of 52.4%. These findings highlight Random Forest's superior robustness in handling class imbalance and capturing complex patterns. Moreover, data balancing significantly enhanced predictive reliability for perinatal mortality risk modeling, as illustrated in Fig. 5.

# Feature ranking

Understanding the reasoning behind machine learning predictions is essential for effectively applying these models in practice. By using the Random Forest algorithm to assess predictor importance for nutritional status, we can identify which factors most strongly influence prenatal mortality. Features with longer bars at the top of the importance ranking have the strongest associations with the outcome. The most influential predictors include the participnats age, , educational status level, and the number of living children key insights that offer actionable guidance for healthcare providers. Interestingly, factors such as the number of antenatal care visits and the use of contraceptive family planning methods ranked low in importance. This unexpected finding may be due to data quality issues, such as missing values or measurement errors in self-reported responses. Nonetheless, clinicians should continue to prioritize these established risk factors during patient assessments as shown in Fig. 6.

**Figure. 6 Feature Importance from Random Forest Model for perinatal mortality Prediction Among gave birth mothers demographic and health survey, 2016–2023**

The Receiver Operating Characteristic (ROC) curve was used to evaluate the performance of the support vector machine (random forest classifier applied to a balanced dataset. Each curve represents the true positive rate against the false positive rate for a specific class, along with its area under the curve (AUC) score. The model demonstrated consistently high performance across all perinatal mortality see Figure. 7.

# SHAP Summary Plot

The SHAP dependence plots show that red and pink colors correspond to higher values of the interacting variable, while blue indicates lower values. The analysis identified maternal education, parity, and maternal age as key predictors of perinatal mortality. Higher levels of maternal education and greater household wealth were associated with a lower predicted risk, demonstrating their protective effect. In contrast, an increase in the number of living children substantially raised the predicted probability of perinatal mortality, reflecting the detrimental impact of high parity. Maternal age contributed modestly and positively to the risk, which is consistent with the greater obstetric complications observed at advanced ages. Moreover, the interaction patterns indicate that these associations are influenced by the socioeconomic context.

SHAP analysis was employed to assess the contribution of individual predictors to perinatal mortality. Maternal education emerged as the most influential factor, exhibiting the widest SHAP value range and thus the strongest impact on the model output. Higher educational attainment was associated with a reduced predicted risk of perinatal mortality. The number of living children ranked as the second most important predictor, with increasing parity markedly elevating the predicted risk. Maternal age showed a smaller yet consistent positive effect on perinatal mortality. Furthermore, interaction patterns revealed that the protective effect of maternal education was more pronounced among women from wealthier households. Overall, the SHAP summary plot demonstrated that maternal education was the leading predictor of perinatal mortality, followed by parity and maternal age, as illustrated in the Fig. 8.

# Discussion

Seven machine-learning models were developed to predict perinatal mortality in Sub-Saharan Africa using Demographic and Health Survey data from 2016–2023. Côte d'Ivoire (9.1%), Tanzania (7.5%), Ghana (7.0%), Lesotho (6.3%), and Mozambique (6.3%) recorded the highest proportions of perinatal deaths, whereas Gabon (2.5%), Sierra Leone (2.8%), Zambia (4.0%), and Madagascar (3.6%) had the lowest. The Random Forest model outperformed the other algorithms, with an accuracy of 81.7%, precision of 80%, recall of 84%, F1-score of 82%, and an AUC-ROC of 0.91. Class imbalance was handled using the Synthetic Minority Over-sampling Technique (SMOTE).

The primary objective of this study was to determine the optimal machine learning model for predicting perinatal mortality, identify the most influential predictors, and extract decision rules that describe the underlying relationships in the dataset across selected Sub-Saharan African countries. SHAP results from the Random Forest model demonstrated that maternal age, maternal educational status, number of living children, multiple gestion were among the most significant predictors variables. Women with higher levels of education showed substantially lower SHAP values, indicating a protective effect against perinatal mortality.

The findings indicated that low maternal educational attainment was associated with an increased risk of perinatal mortality. This result is in line with earlier studies from Indonesia (35), the United States, Canada, and India (36), and the WHO Global Survey conducted in 24 countries across Africa, Asia, and Latin America between 2004 and 2005, which also documented a higher burden of perinatal deaths among women with limited education (37).

This may be attributed to the fact that low maternal education is closely linked to reduced use of antenatal care services. In many settings, the availability of ANC does not guarantee its utilization, as sociocultural and economic constraints such as low female social status, limited access to formal education, and poor decision-making power indirectly hinder women from seeking appropriate care. Lower levels of education may be a proxy for poor health-seeking behavior, suboptimal lifestyle choices. As study as shown in Indonesia (35), India(36).

Educated mothers are more likely to practice optimal neonatal care behaviors, particularly clean cord care and early initiation of breastfeeding, which are essential for improving newborn survival. Enhances the uptake of antenatal care services, increases the likelihood of institutional delivery, promotes appropriate newborn care practices, and improves timely healthcare-seeking behavior. Across different outcomes, the results provide strong and consistent support for the research question, indicating that maternal education is a key determinant of increased utilization of maternal and newborn health services. The education of women is a critical determinant of maternal and child health, primarily through its effect on the utilization of essential health services. Core maternal healthcare indicators reflecting this utilization include tetanus toxoid immunization, consumption of iron and folic acid supplements, antenatal care visits, skilled delivery care, and institutional deliver as the study supported in Bangladesh(37), Pakistan(38), Nepal(39), India(40).Nigeria(41)and Ethiopia(42)

Low maternal education consistently increases perinatal, neonatal, and under-five mortality across countries, largely through preventable pathways, supporting investments in girls' education and improved maternal–newborn care.

According to this study large number of living children indicated that attainment was associated increased risk of perinatal mortality, particularly in low-resource contexts. Importantly, the risk pattern is often U-shaped, meaning that both first births and very high-order births are linked to greater mortality compared with intermediate parity levels. this finding was supported by research conducted in(43), (44), (45), (46)and (47)

One of the possible reasons could be delays in reaching health facilities among high-parity women may be explained by a combination of structural and individual-level factors. The cumulative burden of repeated childbirth, extensive domestic responsibilities, financial hardship, low educational attainment, restricted decision-making autonomy, and reduced perception of pregnancy risk may collectively hinder timely access to appropriate care(48). Evidence from the literature suggests that factors such as financial capacity, women's empowerment, physical distance to healthcare facilities, and level of knowledge play significant roles in shaping maternal healthcare utilization(49–51).

Repeated childbearing may lead to maternal depletion syndrome; whereby cumulative nutritional deficiencies develop over successive pregnancies. High parity is also associated with a higher likelihood of obstetric complications, such as pregnancy-induced hypertension and hemorrhage, which can adversely affect perinatal outcomes(52, 53).

The risk of perinatal and neonatal mortality follows a U-shaped pattern, with both first births and very high-order births carrying greater risk than moderate parity. This pattern is particularly evident in resource-constrained settings where access to quality maternal and newborn care is limited. Implementing targeted surveillance for high-risk pregnancies and strengthening family planning programs to prevent early and high-order births are key measures to reduce perinatal mortality.

# Strength and limitations

# Strength

This study offered several notable strengths. It leveraged a large, multi-country dataset from the Demographic and Health Surveys, which enhanced the generalizability of the findings across Sub-Saharan Africa. The application of advanced machine learning algorithms enabled more accurate prediction of perinatal mortality and uncovered complex patterns that traditional statistical methods might have overlooked. To improve interpretability, the study employed SHAP analysis, which clarified how individual predictors contributed to the model's outputs. Additionally, the researchers conducted validation procedures to mitigate the limitations often associated with the "black-box" nature of machine learning models.

# Limitations of the study

This study had several limitations. First, the DHS data relied on self-reported information from respondents, which may introduce recall bias, particularly when participants were asked to recall events that occurred years earlier. Second, the cross-sectional design limited the ability to establish causal relationships, as it could only identify associations rather than determine the direction of effects. Third, although the findings provide valuable insights for Sub-Saharan Africa, they may not be generalizable to other regions with distinct cultural practices or healthcare systems. Lastly, to enhance cross-country comparability, the analysis excluded country-specific variables, which may have reduced the representativeness of the data and omitted certain contextual factors that could have enriched the investigation.

## Conclusion and recommendation

### Conclusion

In this research applied seven distinct machine-learning algorithms to Demographic and Health Survey datasets (2016–2023) from selected Sub-Saharan African nations to predict perinatal mortality. The prevalence of perinatal deaths varied significantly across countries, with the highest rates observed in

Côte d'Ivoire (9.1%), Tanzania (7.5%), Ghana (7.0%), Lesotho (6.3%), and Mozambique (6.3%). Conversely, the lowest figures were recorded in Gabon (2.5%), Sierra Leone (2.8%), Madagascar (3.6%), and Zambia (4.0%). The Random Forest algorithm emerged as the most effective model, delivering the best performance metrics: 81.7% accuracy, 80% precision, 84% recall, an 82% F1-score, and an AUC-ROC of 0.91. To ensure the model's reliability despite imbalanced data, the Synthetic Minority Over-sampling Technique was employed.

The study aimed to pinpoint the most effective predictive model, uncover the key drivers of perinatal mortality, and elucidate the complex relationships within the data. Through SHAP analysis of the top-performing Random Forest model, several critical factors were identified. These included the mother's age, her total number of living children, the educational levels of both the mother and her husband, and whether she received a tetanus toxoid vaccination during pregnancy. The analysis revealed a clear pattern: women with advanced education exhibited significantly lower SHAP values, reinforcing the idea that education serves as a strong protective factor.

The evidence consistently pointed to a mother's lack of formal education as a major contributor to higher perinatal mortality risks. This connection likely stems from the broader challenges faced by less-educated women, including socioeconomic barriers that limit their access to and use of antenatal care. Even when services are available, factors like limited autonomy and social standing can prevent women from seeking timely care. Education appears to empower women to make healthier choices, such as attending regular check-ups, adopting proper newborn care practices like immediate breastfeeding and hygienic cord management, and opting for skilled birth attendance in a health facility. This empowerment also translates into greater use of vital maternal health services, including necessary vaccinations and nutritional supplements.

Furthermore, the study highlighted that the number of children a woman has previously given birth to also plays a significant role. The relationship between parity and perinatal death was found to be U-shaped; both first-time mothers and those with many previous children faced a heightened risk, especially in settings with limited resources.

Ultimately, these findings strongly affirm that improving educational opportunities for girls is a foundational strategy for enhancing maternal and newborn health. The evidence suggests that the benefits of maternal education are multifaceted, directly influencing health-seeking behavior and the utilization of essential healthcare services. Therefore, sustained efforts to promote female education, coupled with strengthened maternal and neonatal care systems, are vital interventions for lowering perinatal mortality rates across Sub-Saharan Africa.

## Recommendation

### Promote girls' Education and Female Empowerment

Since maternal education strongly protects against perinatal death, governments should:

Invest in girls' access to secondary and higher education, remove school fees, offer scholarships, and tackle cultural barriers that disrupt girls' schooling, provide literacy programs for women who missed formal education, incorporate maternal and child health topics into school lessons and women's group activities

Improve Antenatal Care Quality and Reach

Given the role of tetanus vaccination and ANC services in reducing perinatal risks:

Guarantee universal access to quality ANC, especially for rural and marginalized women

Deliver comprehensive care packages including immunization, supplements, and routine monitoring

Use community outreach to connect with pregnant women not attending ANC

**Focus on Parity-Based Risk Groups**

Design tailored care for first-time mothers, with extra education and support throughout pregnancy and delivery expand family planning services to enable healthy birth spacing and limit high-order pregnancies in resource-limited areas offer intensified monitoring and skilled delivery access for mothers with many previous children and educate women of reproductive age on the dangers of both very low and very high parity

# Abbreviations

**ANC** Antenatal Care
**AUC** Area Under the Curve
**AUC-ROC** Area Under the Receiver Operating Characteristic Curve
**CSV** Comma-Separated Values
**DHS** Demographic and Health Surveys
**DTA** Data (STATA data file format)
**FN** False Negative
**FP** False Positive
**GBM** Gradient Boosting Machine
**IRB** Institutional Review Board
**LightGBM** Light Gradient Boosting Machine
**ML** Machine Learning
**NN** Neural Network
**PNC** Postnatal Care
**ROC** Receiver Operating Characteristic
**SDG** Sustainable Development Goals
**SHAP** Shapley Additive exPlanations
**SMOTE** Synthetic Minority Over-sampling Technique

**SSA** - Sub-Saharan Africa
**STATA** Statistics and Data (software)
**TN** True Negative
**TP** True Positive
**WHO** World Health Organization
**XGBoost** xtreme Gradient Boosting

# Declarations

### Disclosure statement

The authors declared that there was no conflict of interest in this work.

### Ethics approval and consent to participate

This study was based on secondary data analysis, and we obtained permission from the MEASURE DHS program to download and use the data for our research purposes. As a result, ethical approval and participant consent are not required for this study. The dataset is publicly available in the MEASURE DHS program's official database, with no personal identifiers.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

# Author details

1Department of Health Informatics, Institute of Public Health, College of Medicine and Health Sciences, University of Gondar, Gondar, Ethiopia. 2 Department of Health Informatics, School of Public Health, College of Medicine and Health Sciences

# Funding

This particular study did not receive funds from any funding agencies.

# Author Contribution

1. Tadele Chekol Maru: Proposal development, data analysis, writing the document, thesis write up and principal investigator. 2. Andualem Enyew Gedefaw: Document review. 3.Makda Fekadie Tewelgne,

document reviewer. 4.Eliays Addisu Taye, data extraction. 5. Agere mengistie Zeleke, data extraction. 6. Belayeneh Jejaw Abate: data extraction. 7.Deresse Abebe Gebrehana, proposal reviewer. 8. Azanaw Amare Muche: Document review.


# Acknowledgements

The authors would like to thank the MEASURE DHS program for providing the data for further analysis.


# Data Availability

The dataset used for this study is publicly available at the MEASURE DHS program website https://www.dhsprogram.com/data.

## References


1. Jain V, Chatterjee JM. Machine learning with health care perspective. Cham: Springer. 2020:1–415.
2. Organization WH. Youth-centered digital health interventions: a framework for planning, developing and implementing solutions with and for young people. World Health Organization; 2020.
3. Meraji M, Jafari M. Can International Classification of Disease Perinatal Mortality (ICD-PM) be a solution to overcome neglected tragedy? A scoping reviews. Health Sci Rep. 2024;7(11):e70134.
4. Rasaily R, Saxena N, Pandey S, Garg BS, Swain S, Iyengar SD et al. Effect of home-based newborn care on neonatal and infant mortality: a cluster randomised trial in India. BMJ global health. 2020;5(9).
5. Hug L, Mishra A, Lee S, You D, Moran A, Strong KL et al. A neglected tragedy the global burden of stillbirths: report of the UN inter-agency group for child mortality estimation, 2020. 2020.
6. Madaj B, Smith H, Mathai M, Roos N, Van Den Broek N. Developing global indicators for quality of maternal and newborn care: a feasibility assessment. Bull World Health Organ. 2017;95(6):445.
7. Housseine N, Snieder A, Binsillim M, Meguid T, Browne JL, Rijken MJ. The application of WHO ICD-PM: Feasibility for the classification of timing and causes of perinatal deaths in a busy birth centre in a low-income country. PLoS ONE. 2021;16(1):e0245196.
8. Ouyang F, Zhang J, Betrán AP, Yang Z, Souza JP, Merialdi M. Recurrence of adverse perinatal outcomes in developing countries. Bull World Health Organ. 2013;91:357–67.
9. Akombi BJ, Renzaho AM. Perinatal mortality in sub-Saharan Africa: a meta-analysis of demographic and health surveys. Annals global health. 2019;85(1):106.
10. De Bernis L, Kinney MV, Stones W, ten Hoope-Bender P, Vivio D, Leisher SH, et al. Stillbirths: ending preventable deaths by 2030. lancet. 2016;387(10019):703–16.
11. Manu A, Ewerling F, Barros AJ, Victora CG. Association between availability of children's book and the literacy-numeracy skills of children aged 36 to 59 months: secondary analysis of the UNICEF

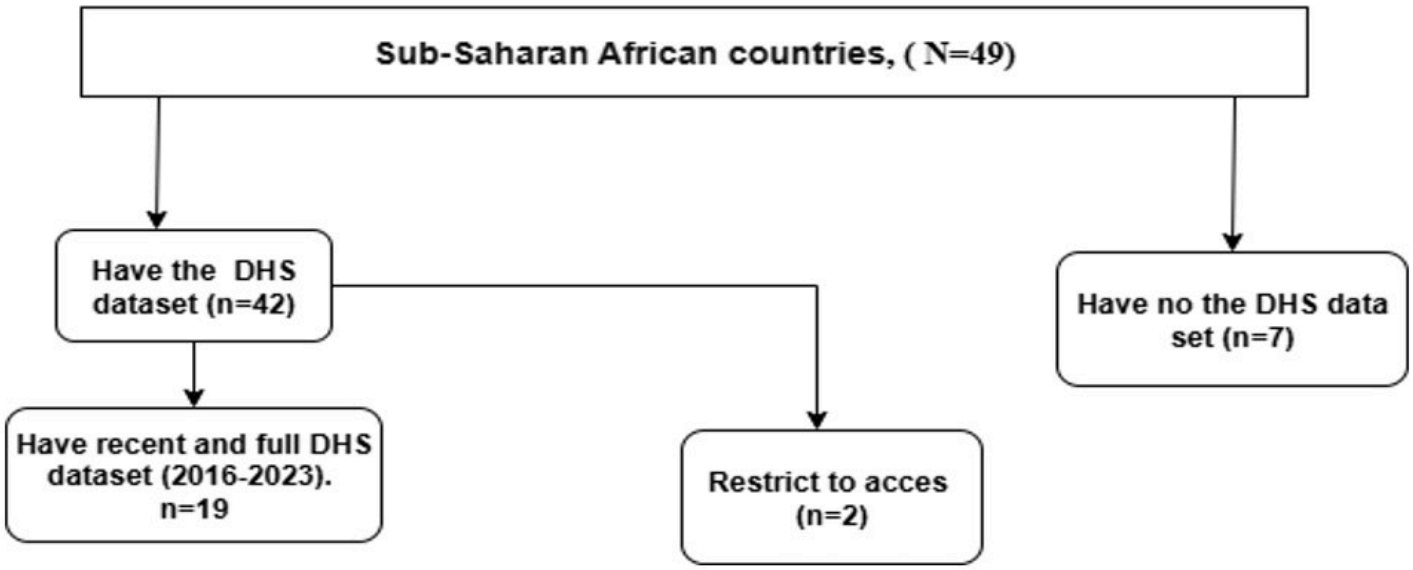
Multiple-Indicator Cluster Surveys covering 35 countries. J global health. 2018;9(1):010403.

12. Comfort AB, Peterson LA, Hatt LE. Effect of health insurance on the use and provision of maternal health services and maternal and neonatal health outcomes: a systematic review. J Health Popul Nutr. 2013;31(4 Suppl 2):S81.
13. Ghimire PR, Agho KE, Renzaho AM, Nisha MK, Dibley M, Raynes-Greenow C. Factors associated with perinatal mortality in Nepal: evidence from Nepal demographic and health survey 2001–2016. BMC Pregnancy Childbirth. 2019;19(1):88.
14. Hassan ZA, Ahmed MJ. Factors associated with immunisation coverage of children aged 12–24 months in Erbil/Iraq 2017–2018. Int J Psychosoc Rehabil. 2020;24(08):12222–35.
15. Goba GK, Tsegay H, Gebregergs GB, Mitiku M, Kim KA, Alemayehu M. A facility-based study of factors associated with perinatal mortality in Tigray, northern Ethiopia. Int J Gynecol Obstet. 2018;141(1):113–9.
16. Arach AAO, Tumwine JK, Nakasujja N, Ndeezi G, Kiguli J, Mukunya D, et al. Perinatal death in Northern Uganda: incidence and risk factors in a community-based prospective cohort study. Global Health Action. 2021;14(1):1859823.
17. Seyoum E, Bekele A, Tsegaye AT, Birhanu S. Magnitude and determinants of adverse perinatal outcomes in Tefera Hailu Memorial Hospital, Sekota Town, northern Ethiopia. Global Pediatr health. 2021;8:2333794X211015524.
18. Al-Ghamdi S, Alfauri TM, Alharbi MA, Alsaihati MM, Alshaykh MM, Alharbi AA et al. Current self-medication practices in the Kingdom of Saudi Arabia: an observational study. Pan Afr Med J. 2020;37(1).
19. Martins EF, Rezende EM, Almeida MCM, Lana FCF. Perinatal mortality and socio-spatial inequalities. Rev Latinoam Enferm. 2013;21(05):1062–70.
20. Usynina AA, Grjibovski AM, Krettek A, Odland JØ, Kudryavtsev AV, Anda EE. Risk factors for perinatal mortality in Murmansk County, Russia: a registry-based study. Global health action. 2017;10(1):1270536.
21. Geda A, Shemsu S, Debalke R. Determinants of perinatal mortality in public hospitals of Iluu Abbaa Boor Oromia Region, South West Ethiopia, 2019: unmatched case–control study. Res Rep Neonatology. 2021:57–66.
22. Debelew GT. Magnitude and determinants of perinatal mortality in Southwest Ethiopia. J Pregnancy. 2020;2020(1):6859157.
23. Hossain MB, Kanti Mistry S, Mohsin M, Rahaman Khan MH. Trends and determinants of perinatal mortality in Bangladesh. PLoS ONE. 2019;14(8):e0221503.
24. Jena BH, Biks GA, Gelaye KA, Gete YK. Magnitude and trend of perinatal mortality and its relationship with inter-pregnancy interval in Ethiopia: a systematic review and meta-analysis. BMC Pregnancy Childbirth. 2020;20(1):432.
25. Roro EM, Sisay MM, Sibley LM. Determinants of perinatal mortality among cohorts of pregnant women in three districts of North Showa zone, Oromia Region, Ethiopia: Community based nested

case control study. BMC Public Health. 2018;18(1):888.
26. Program DPTD. [Available from: https://dhsprogram.com.
27. Duke T, AlBuhairan FS, Agarwal K, Arora NK, Arulkumaran S, Bhutta ZA, et al. World Health Organization and knowledge translation in maternal, newborn, child and adolescent health and nutrition. Arch Dis Child. 2022;107(7):644–9.
28. Hug L, Alexander M, You D, Alkema L. National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. Lancet Global Health. 2019;7(6):e710–20.
29. Mitao M, Philemon R, Obure J, Mmbaga BT, Msuya S, Mahande MJ. Risk factors and adverse perinatal outcome associated with low birth weight in Northern Tanzania: a registry-based retrospective cohort study. Asian Pac J Reprod. 2016;5(1):75–9.
30. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321–57.
31. learning Avitm. [Available from: https://r2d3.us/visual-intro-to-machine-learning/
32. documentation. S-l. [Available from: https://scikit-learn.org/stable/
33. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Adv Neural Inf Process Syst. 2017;30.
34. Croft TN, Marshall AM, Allen CK, Arnold F, Assaf S, Balian S. Guide to DHS statistics. Rockville: ICF. 2018;645:292–303.
35. Tosepu R. Distribution Patterns of Maternal and Child Health in West Sulawesi Province, Indonesia: A Descriptive Analysis from 2017 to 2020. J Health Sci Pharm. 2025;2(2):265–71.
36. Chopra I, Juneja SK, Sharma S. Effect of maternal education on antenatal care utilization, maternal and perinatal outcome in a tertiary care hospital. Int J Reprod Contracept Obstet Gynecol. 2018;8(1):247.
37. Shahjahan M, Ahmed MR, Rahman MM, Afroz A. Factors affecting newborn care practices in Bangladesh. Paediatr Perinat Epidemiol. 2012;26(1):13–8.
38. Khan Z, Soomro GY, Soomro S, Hafeez S. Mother's education and utilisation of health care services in Pakistan [with comments]. Pakistan Dev Rev. 1994;33(4):1155–66.
39. Thasineku OC. Association of Mother's Education with Maternity ond Newborn Care Proctices in Nepol. 2022.
40. Govindasamy P, Ramesh B. Maternal education and the utilization of maternal and child health services in India. 1997.
41. Umar AS. Does female education explain the disparity in the use of antenatal and natal services in Nigeria? Evidence from demographic and health survey data. Afr Health Sci. 2017;17(2):391.
42. Gedefaw Abeje Fekadu GAF, Fentie Ambaw Getahun FAG, Seblewongiel Ayenalem Kidanie SAK. Facility delivery and postnatal care services use among mothers who attended four or more antenatal care visits in Ethiopia: further analysis of the 2016 Demographic and Health Survey. 2018.

43. Andrejevic A, Cvetkovic S, Vitosevic Z, Andrejevic L, Relic G. Multiparity, perinatal morbidity and mortality. Clin Experimental Obstet Gynecol. 2011;38(1):71.
44. Kozuki N, Lee AC, Silveira MF, Sania A, Vogel JP, Adair L, et al. The associations of parity and maternal age with small-for-gestational-age, preterm, and neonatal and infant mortality: a meta-analysis. BMC Public Health. 2013;13(Suppl 3):S2.
45. Siahaan A, Ariawan I. The effect of parity on neonatal mortality in indonesia. Jurnal Ilmu Kesehatan Masyarakat. 2021;12(3):250–62.
46. Straneo M, Beňová L, van den Akker T, Abeid MS, Ayebare E, Dossou J-P, et al. Mixed vulnerabilities: the biological risk of high parity is aggravated by emergency referral in Benin, Malawi, Tanzania and Uganda. Int J Equity Health. 2025;24(1):19.
47. Dasa TT, Okunlola MA, Dessie Y. Effect of grand multiparity on the adverse birth outcome: A hospital-based prospective cohort study in Sidama Region, Ethiopia. Int J women's health. 2022:363–72.
48. Straneo M, Beňová L, van den Akker T, Pembe AB, Smekens T, Hanson C. No increase in use of hospitals for childbirth in Tanzania over 25 years: Accumulation of inequity among poor, rural, high parity women. PLOS global public health. 2022;2(8):e0000345.
49. Murray SF, Pearson SC. Maternity referral systems in developing countries: current knowledge and future research needs. Soc Sci Med. 2006;62(9):2205–15.
50. Schmitz MM, Serbanescu F, Arnott GE, Dynes M, Chaote P, Msuya AA et al. Referral transit time between sending and first-line receiving health facilities: a geographical analysis in Tanzania. BMJ Global Health. 2019;4(Suppl 5).
51. Vidler M, Kinshella M-LW, Sevene E, Lewis G, von Dadelszen P, Bhutta Z. Transitioning from the Three Delays to a focus on continuity of care: a qualitative analysis of maternal deaths in rural Pakistan and Mozambique. BMC Pregnancy Childbirth. 2023;23(1):748.
52. Lawn JE, Ohuma EO, Bradley E, Idueta LS, Hazel E, Okwaraji YB, et al. Small babies, big risks: global estimates of prevalence and mortality for vulnerable newborns to accelerate change and improve counting. Lancet. 2023;401(10389):1707–19.
53. Sequeira Dmello B, John TW, Housseine N, Meyrowitsch DW, van Roosmalen J, van den Akker T, et al. Incidence and determinants of perinatal mortality in five urban hospitals in Dar es Salaam, Tanzania: a cohort study with an embedded case–control analysis. BMC Pregnancy Childbirth. 2024;24(1):62.

# Figures




## Figure 1

Sampling procedure

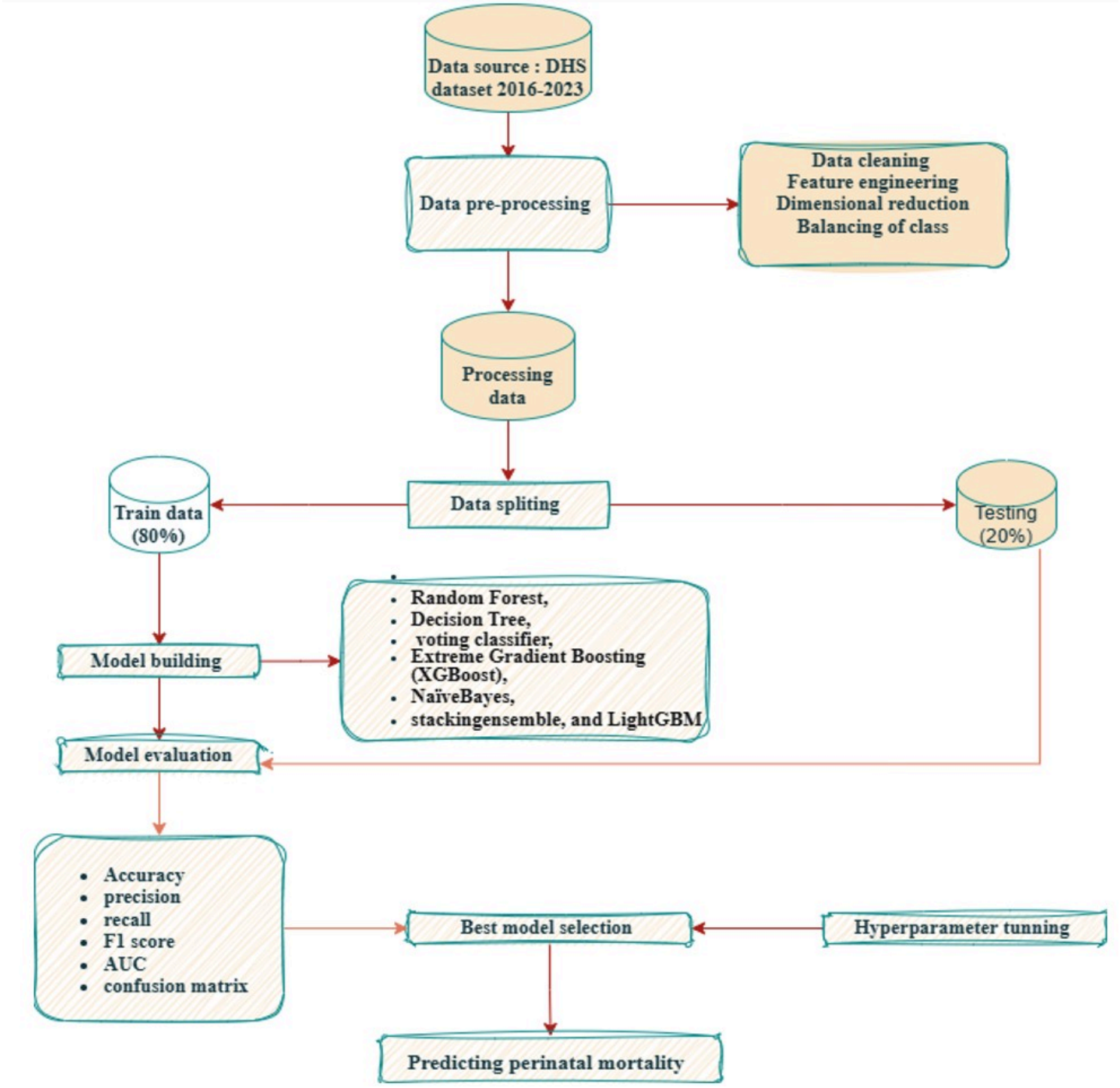



Figure 2

Methodological framework and machine learning workflow.

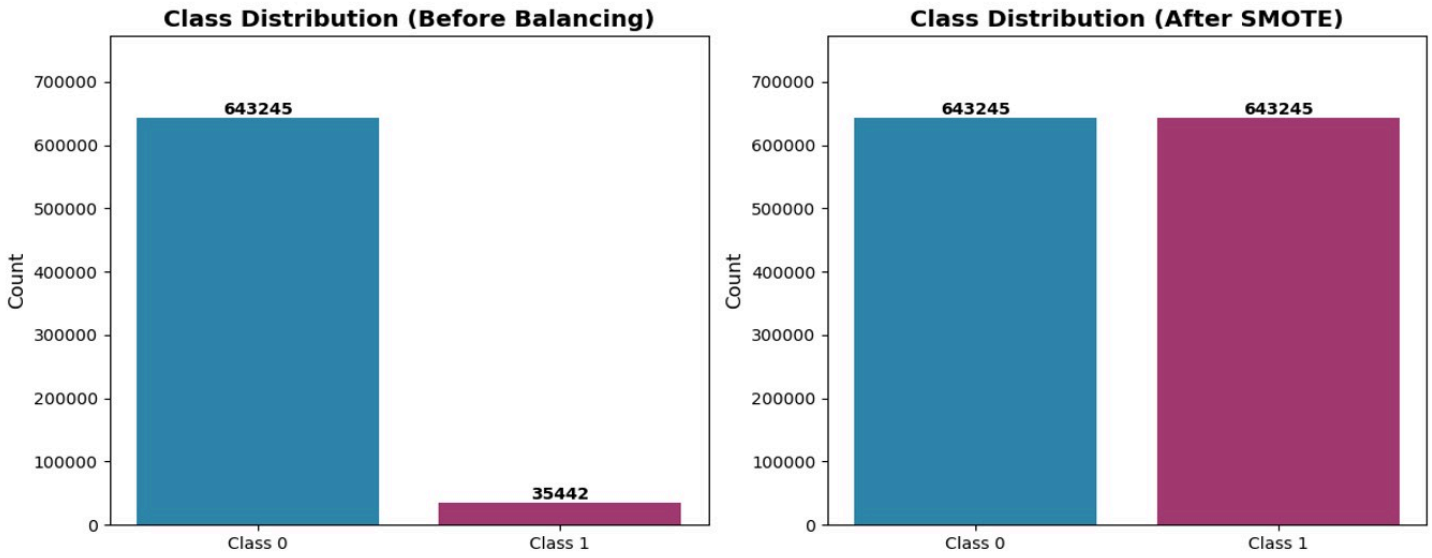
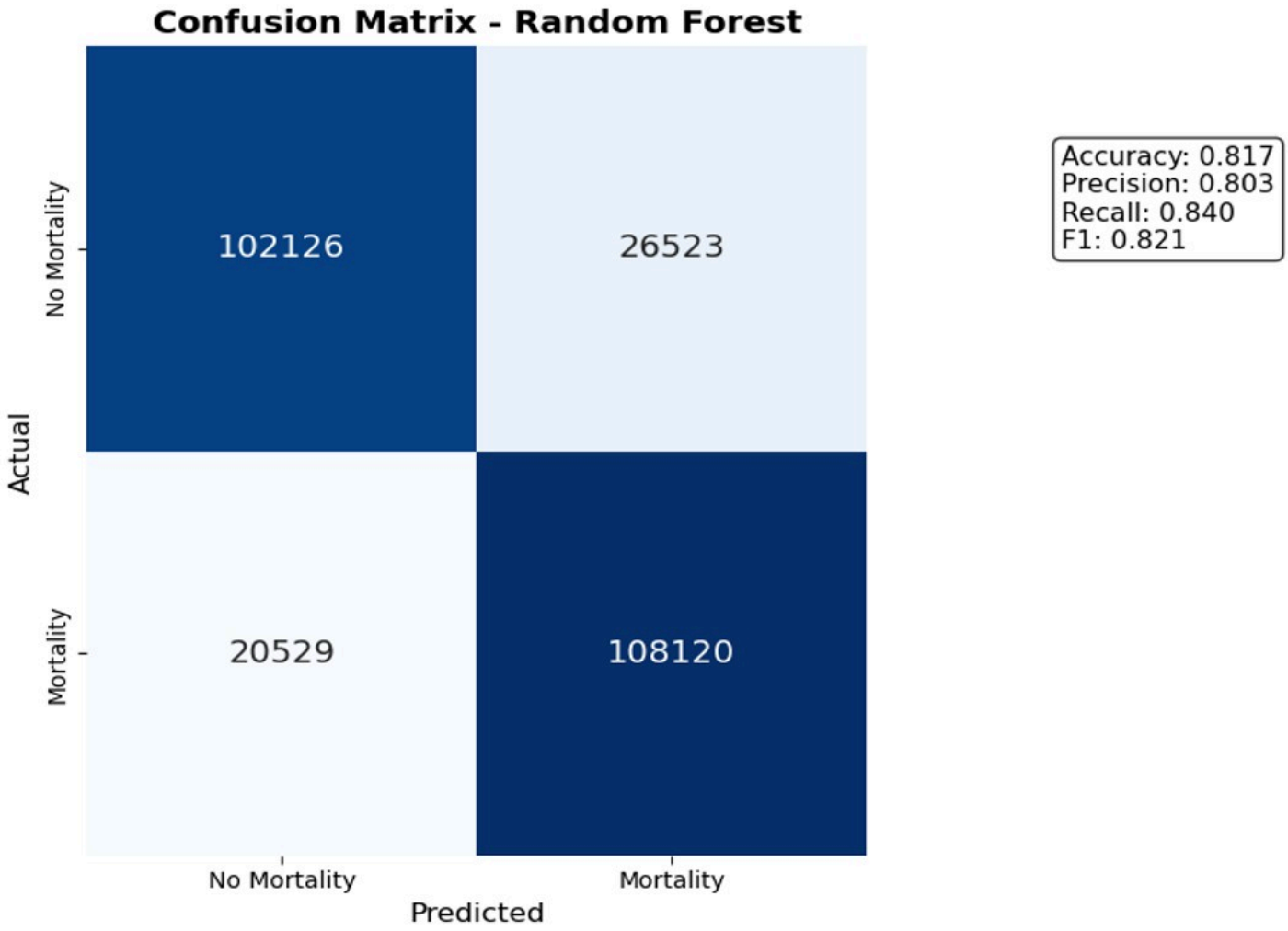


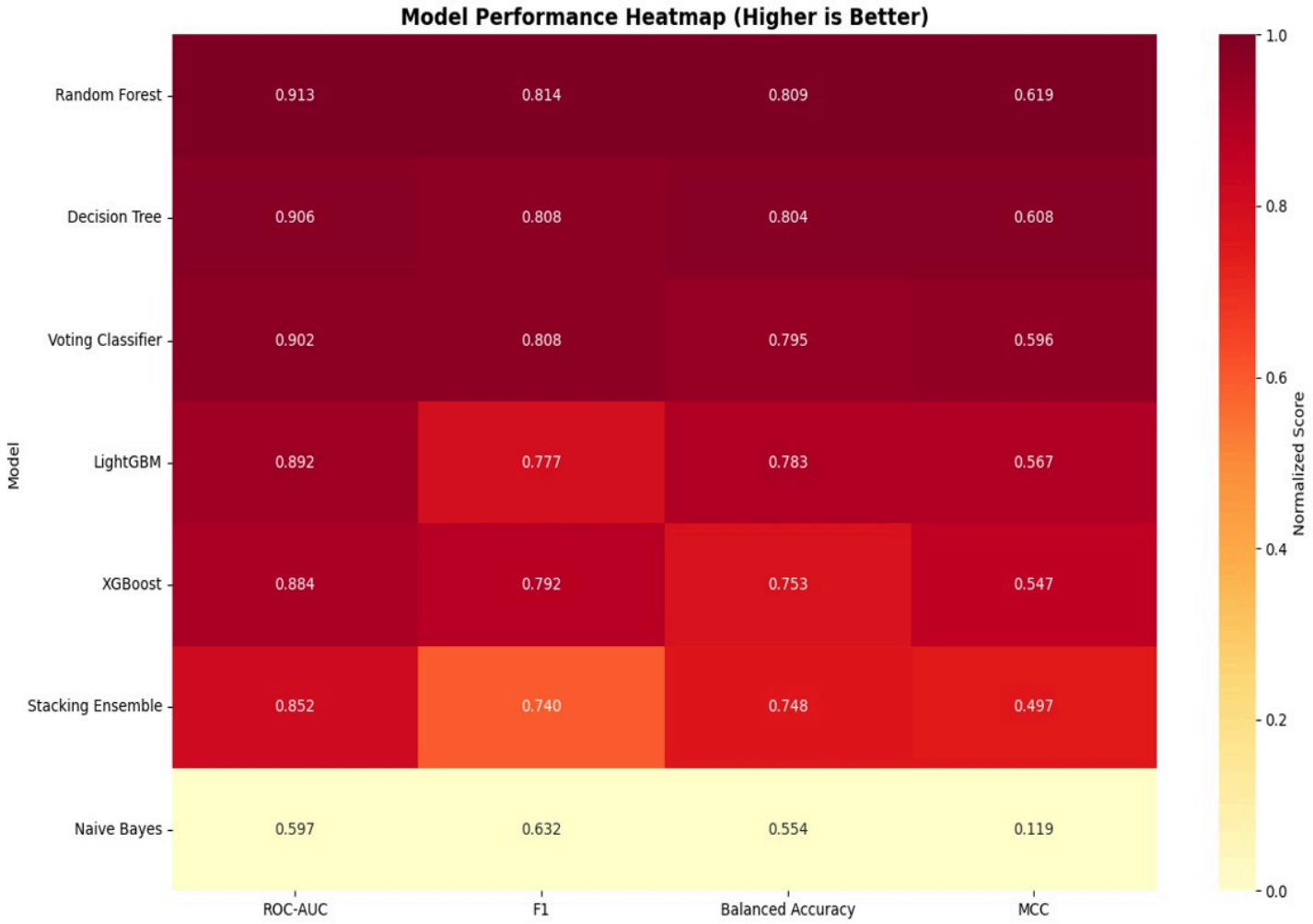

Figure 3

Class distribution of perinatal mortality Demographic and health survey in SSA, 2016-2023



## Figure 4

## confusion matrix




## Figure 5

## model performance metrics

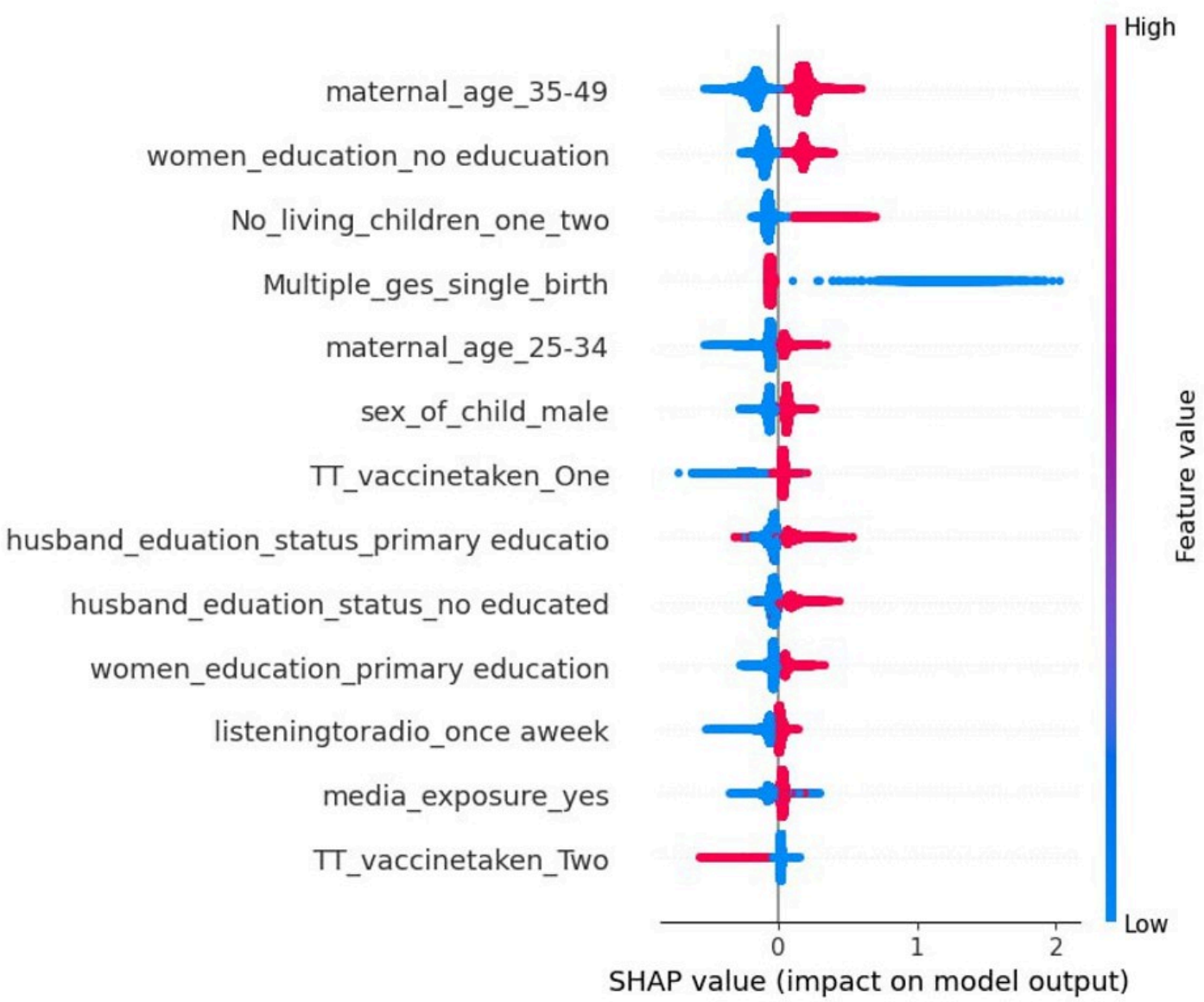
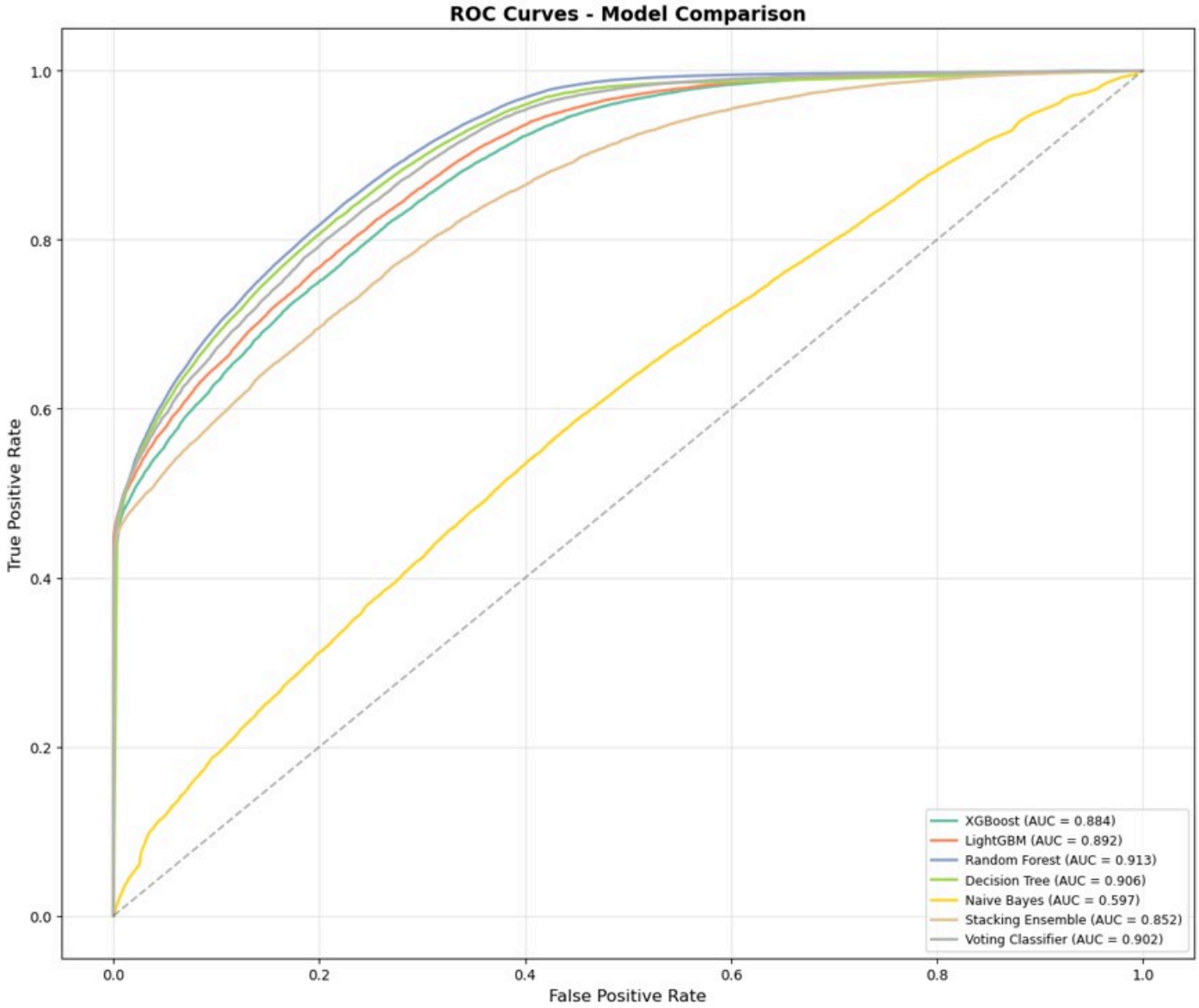



**Figure 6**

**Feature Importance from Random Forest Model for perinatal mortality**




## Figure 7

## ROC Curve for Random Forest Model

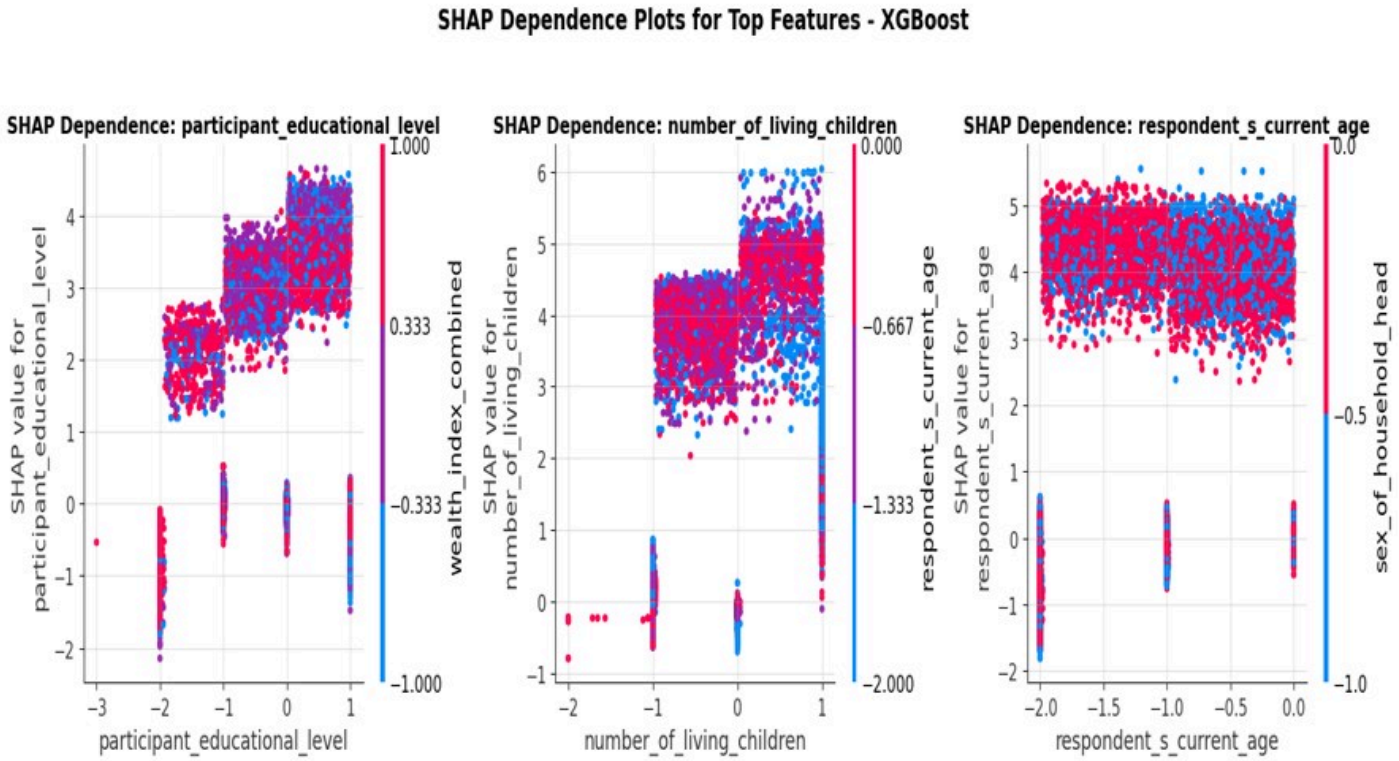



Figure 8

SHAP Summary Plot