

**LEVERAGING ENSEMBLE MODELS FOR OPTIMIZING PREDICTIVE
ACCURACY OF LOW BIRTHWEIGHT RISK IN KENYA.**

VICTOR OTIENO OPIYO

**A PROJECT SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE &
INFORMATION TECHNOLOGY IN THE SCHOOL OF COMPUTING AND
MATHEMATICS IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE
AWARD OF THE DEGREE OF MASTER OF SCIENCE IN DATA SCIENCE OF THE
CO-OPERATIVE UNIVERSITY OF KENYA.**

2025

DECLARATION

Declaration by the candidate

This project is my original work and has not been presented for award of a degree in any other University or for any other award.

.....

24/11/2025

Victor Otieno Opiyo

Date

Registration No. C004/600027/2023

Declaration by the supervisors

We confirm that the work reported in this project was carried out by the candidate under our supervision and has been submitted with our approval as university supervisors

Signed:

Date:

24/11/2025

Dr. Emma Anyika

Department of Computer Science & Information Technology

Cooperative University of Kenya

Signed: _____

Date: _____

24/11/2025

Dr. Ronald Ojino

Department of Computer Science & Information Technology

Cooperative University of Kenya

ACKNOWLEDGMENT

First and foremost, I give thanks and praise to Almighty God for His grace, guidance, and strength throughout the entire duration of this research journey. Without His blessings, this work would not have been possible. I sincerely appreciate my supervisors, Dr. Emma Anyika, Dr. Ronald Ojino, and Dr. Andrew Kipkebut of the Department of Computer Science & Information Technology, Cooperative University of Kenya, for their noble guidance, wise counsel, and unwavering support throughout this one-year research work. Their mentorship has been invaluable in shaping this study. I also extend my gratitude to the Cooperative University of Kenya for providing the necessary resources, facilities, and an enabling environment that facilitated the successful completion of this work. Special thanks go to my colleagues and the entire department for their technical assistance and the materials they provided, which were instrumental in fulfilling this project. Finally, I acknowledge with deep appreciation everyone who contributed intellectually, technically, and financially. Without the combined efforts and support of all these individuals, this work could not have been achieved.

TABLE OF CONTENTS.

DECLARATION	II
ACKNOWLEDGMENT	III
ABSTRACT	XII
CHAPTER ONE	1
1 INTRODUCTION	1
1.1 Background of the study	1
1.2 Statement of the Problem	5
1.3 Objectives of the Study	9
1.3.1 General Objective:	9
1.3.2 Specific Objectives:	9
1.4 Research Questions	9
1.5 Justification of the Study	9
1.6 Scope of the Study	11
1.7 Limitations of the Study	12
CHAPTER TWO	14
2 LITERATURE REVIEW	14
2.0 Introduction	14
2.1 Background	14
2.2 CONCEPTUAL FRAMEWORK	17
2.3 Theoretical Framework	18
2.4 Empirical Review	19
2.4.1 Socio-Demographic Factors:	19
2.4.2. Socio-Economic Factors	21
2.4.3 Obstetric Factors/Maternal Characteristics:	22
2.4.4 Ensemble Techniques:	23
2.2.3 Comparison of Traditional Models and Ensembled Models:	26
2.7 RESEARCH GAPS	30
CHAPTER THREE:	31
3 RESEARCH METHODOLOGY	31
3.1 Introduction	31
3.2 Design Science Research Framework	31
3.2.0 Problem Identification	33
3.2.2 Research Paradigm	33

3.2.3 Research Design	33
3.2.4 Location of the Study.....	34
3.2.5 Building the Evaluation Framework.....	35
3.3 Data Collection.....	37
3.3.1 Data Sources and Sampling Method.....	37
3.3.2 Data Confidentiality and Ethical Considerations	39
3.4.0 Data Preprocessing and Data Analysis Method.....	40
3.4.1 Data Preprocessing Introduction.....	40
3.4.2 Handling Missing Data.....	40
3.4.3 Addressing Outliers and Feature Scaling.....	41
3.4.4 Exploratory Data Analysis (EDA)	41
3.4.5 Analyzing the Significant Factors Associated with LBW.....	42
3.4.6 Dataset Splitting and Stratified Sampling	42
3.4.7 Resampling for Class Imbalance	43
3.4.8 Feature Selection and Dimensionality Reduction	43
3.5.0 Machine Learning Methodology: A Robust Predictive Framework	44
3.5.1 Model Evaluation	46
3.5.2 Feature Selection and Interpretability	46
3.5.3 Visualization for Insightful Interpretation	47
CHAPTER FOUR.....	50
4 DATA ANALYSIS, PRESENTATION AND INTERPRETATION.....	50
4.1 Data Description.....	50
4.2 Machine Learning Classifier Performance.....	53
4.2.1 Base Model Performance.....	54
4.2.2 Ensemble Models.....	54
4.2.3 Meta-Ensemble Model Performance.....	55
CHAPTER FIVE	59
5 DISCUSSIONS OF FINDINGS, CONCLUSIONS AND RECOMMENDATIONS.....	59
5.1 Maternal and Neonatal Demographic Characteristics	59
5.2 Feature Importance.....	62
5.3 Machine Learning Models:	64
5.4 Contributions to Health Informatics Knowledge in Kenya	65
5.5 Barriers to Implementing Ensemble Models in Clinical Settings.....	65
5.6 Ideal Institutions for Piloting and Adoption.....	66
5.7 Future Research Directions.....	66
5.8 CONCLUSION	67

5.9 RECOMMENDATION	68
REFERENCES:	69
APPENDICES	1
Appendix 1: Research Instruments	1
Appendix 2: Research Permits/authorization letter	2
Appendix 3: Similarity Index	3
Appendix 4: AI Report	6
Appendix 5: Published Article	7

List of Tables

Table 1: Summary of Literature Review.	29
Table 2: Maternal and neonatal demographic characteristics	51
Table 3: Factors Significantly Associated with Low Birth Weight (LBW)	53
Table 4: Model performance comparison (Sorted by ROC AUC).....	55

List of Figures

Figure 1: Conceptual Framework	18
Figure 2: Research Methodology.....	32
Figure 3: DSR Research Design	34
Figure 4: Analytical and Modeling Framework.....	49
Figure 5: ROC AUC curve for the Top 3 Models	56
Figure 6: Precision - Recall AUC curve for the Top 3 Models.....	57
Figure 7: Confusion Matrices for the Random Forest Model at Default threshold and at Threshold=0.37.....	57
Figure 8: Average Normalized Feature Importance Across Methods (Selected by Random Forest).....	58

List of abbreviations and acronyms

ANC: Antenatal Care

AUC: Area Under the Curve (for evaluating model performance)

Bagging: Bootstrap Aggregating (an ensemble method)

BMI: Body Mass Index

Boosting: An ensemble technique that combines weak learners sequentially

F1-score: F1 Score (a measure of a model's accuracy)

KDHS: Kenya Demographic and Health Survey

LBW: Low Birthweight

LR: Logistic Regression

ML: Machine Learning

RF: Random Forest

ROC: Receiver Operating Characteristic (curve used to assess model performance)

SDGs: Sustainable Development Goals (related to global health initiatives)

WHO: World Health Organization

XGBoost: Extreme Gradient Boosting

Definition of terms

Age: Mother's age at the time of birth.

Bagging: An approach that minimizes variance by training many models on random subsets of the data and averaging their predictions.

Boosting: It is an iterative process where weak learners are sequentially added and each new model is built to correct the errors made by the preceding ones.

Cross-Validation: Part of the methods of estimating the performance of machine learning models is dividing the data into very small pieces. The premise is that one trains the model using some of these and estimates its accuracy against the others. This serves to make sure that the model is not only great on the training set, but that it generalizes well on the new unseen data which reduces the likelihood of the model simply memorizing the training data overfitting.

Data Preprocessing and Cleaning: This step entails cleaning raw data for examination by managing missing values, correcting outliers, and converting categorical variables into numerical forms that are compatible with machine learning algorithms.

Ensemble Models: Ensemble learning is a machine learning technique where two or more learners (e.g., neural network, regression models) are combined to produce better predictions. In other words, an ensemble model is a collection of individual models combined to make more accurate predictions than one model. Some of them include Bagging, Boosting, and Stacking.

Health Care Access Disparities: Differences in the quality and accessibility of health care services to various groups in Kenya, based on factors like geographic location, socio-economic level, and cultural beliefs.

Interpretability Framework: It is intended to allow health workers to understand and utilize the predictive model in their clinical practice every day. It is supposed to translate data insights into simple language so that the doctors are in a position to make appropriate decisions on the behalf of the patients.

Low Birthweight (LBW): A child born weighing less than 2500 grams is known as a low birthweight child.

Machine Learning (ML): This refers to artificial intelligence that is statistical-model and algorithm-based used for data processing and reaching potential outcomes.

Maternal Factors: These are the conditions of the mother during the pregnancy period that would influence the weight of the baby at birth.

- **Number of ANC Visits:** These are the number of clinic visits to a health facility during pregnancy to check the health status.
- **Obstetric History:** It describes the previous pregnancies and their results, which could affect the present pregnancy.

Socioeconomic Factors: These are the socioeconomic situations of the mother that would impact on her health.

Stacking: This is an ensemble approach that trains base models and subsequently combine their outputs using another model for final predictions.

Wealth Index: A combined measure of a household status based on its asset ownership.

ABSTRACT

Low birth weight (LBW) remains a significant public health concern in Kenya, affecting approximately 11.5% of infants and resulting in high infant mortality and long-term poor health. Accurate prediction of LBW risk is crucial to enable timely interventions and improve neonatal health outcomes. The objective of this study was to develop and evaluate ensemble machine learning models to predict the risk of LBW using nationally representative data from the Kenya Demographic and Health Survey (KDHS) 2022. A comprehensive preprocessing pipeline was used to handle missing values, encode categorical variables, and address class imbalance using the Synthetic Minority Over-Sampling Technique (SMOTE). Various machine learning methods, the base methods like Support Vector Machine and Logistic Regression, and ensemble models like the Random Forest, Gradient Boosting, and Extreme Gradient Boosting were all trained and compared. Moreover, the prediction abilities of meta-ensemble methods such as bagging, voting, and stacking classifiers are also evaluated. Model assessment was done using stratified cross-validation, and performance was evaluated on an independent test set using performance metrics such as ROC AUC, F1-score, and Brier score. Random Forest classifier achieved the highest score of 0.957 ROC AUC with decent calibration (Brier score of 0.089), being better than both base and meta-ensemble models. The key predictors identified from the analysis include gestational age, maternal anthropometrics (height, weight), and antenatal care attendance, which proved their biological and contextual applicability to LBW risk in Kenya. The paper highlights the significance of contextualized AI solutions and ethical governance in sustainable healthcare innovation. These results indicate that ensemble learning methods can be used with specific target population selection to achieve better results in LBW risk prediction in low-resource regions. Developing interpretable and stable models can guide clinical decision-making and focused interventions with the long-term objective of encouraging maternal and neonatal health outcomes in Kenya and other contexts.

CHAPTER ONE

1 INTRODUCTION

This section provides an overview for the background, problem statement, objectives, research questions, significance, expected outcomes, justification, scope, and limitations of the study. Its primary aim is to offer a thorough introduction to the research topic and establish the context for the following chapters.

1.1 Background of the study.

The health of a newborn is a crucial element in a nation's and global health. It is the primary factor that determines the overall health of a human being and life expectancy. Low birth weight (LBW), defined as a birth weight of less than 2,500 grams, poses a significant public health challenge in Kenya and other developing countries. It is not just a statistic; it represents the struggles faced by families and the health risks that newborns encounter from the outset of life. Therefore, a baby's health and well-being should be monitored after birth and during the period of growth in the womb. One aspect that should be observed before the baby is born is its weight. Birth weight is the newborn baby's first weight measured immediately after being born within the first hour before the occurrence of significant loss of weight due to postnatal effects. A newborn's weight signifies a lot about the future health and survival of the baby. Therefore, it is advisable to know whether the baby is going to have a normal weight or a low birth weight in order to make early interventions before birth.

The major causes of low birth weight are premature birth and growth faltering in the womb, which restricts the baby from reaching its optimal weight. Premature/preterm birth occurs in less than 37 weeks(259 days) since the start of the last menstruation of a woman preceding conception as defined by WHO (WHO, 2023). Maternal factors also play a significant role, such as malnutrition (anemia, underweight, or overweight) before and during pregnancy,

maternal health problems (high blood pressure, diabetes, or infections), and maternal characteristics (low or high maternal age, multiple parity, or poor birth spacing). Other risk factors contributing to LBW include smoking, alcohol consumption, and medically unnecessary caesarean deliveries. The repercussions of LBW extend beyond the neonatal period, leading to increased infant mortality rates and long-term health complications that can affect individuals throughout their lives (Ranjbar, Montazeri, Farashah, et al., 2023a).

According to the global statistics of 2015, 14.6% of newborns had low birth weight, which is more than 20.5 million. The reported rates also ranged differently; the developed countries had the lowest rate of 7.2%, Africa had 13.0%, while Asia had the highest with 17.3%. Southern Asia region had the highest percentage in low birth weight at 26.4 % (*UNICEF-WHO Low Birthweight Estimates 2019 .Pdf*, n.d.). In Kenya, this work thus identifies 11.5% of infants being born with LBW in a society that is still struggling with socio-economic inequalities and therefore the need to propose interventions for the problem (K'Oloo et al., 2023).

Despite the seriousness of this situation, the tools for early prediction and intervention remain underdeveloped in resource-limited settings like Kenya, where hospitals rely on various traditional methods to assess and estimate birth weight in clinical practice. These methods include: obstetric ultrasounds, symphysis-fundal height measurements, and abdominal palpation (Sawe, 2022). However, these methods encounter challenges with reliability and accuracy. For example, while obstetric ultrasound is viewed as the most reliable option for assessing fetal growth, it is frequently inaccessible in low-resource areas and underprivileged communities. Moreover, untrained ultrasound sonographers may yield inaccurate fetal weight measurements, underscoring the importance of proper training.

Thus, considering the difficulties mentioned above in connection with the current methods of LBW estimation, it is expedient to search for more effective ways out. Greater attention should

be paid to the use of effective approaches for assessing further birth weight and LBW prediction, as timely diagnosis can be beneficial in elaborating appropriate obstetric measures. Several techniques can be used to predict LBW, but recent studies have pointed out that the machine learning algorithm has promising capability to predict LBW.

Sophisticated ML using big data analytics as a technique has registered success in reducing LBW in high-income countries, since they work to understand the factors that can lead to the outcome from the data sets. For instance, it has been established by various authors that ensemble models such as Random Forest and XGBoost possess the capacity of correctly predicting the risk factors of LBW with accuracy rates of over 85% (Bekele, 2022). However, the application of these methodologies has not been specifically tested in developing countries such as Kenya.

It can therefore be suggested that to increase the accuracy of estimating birth weight in the Kenyan hospitals, the following can be proposed: ensemble machine learning. The above-mentioned methods have their benefits over preceding techniques as they will allow for coming up with more accurate and precise results (Ranjbar, Montazeri, Farashah, et al., 2023b). Two of the categories that are widely used for model integration include Bagging, Boosting and Stacking, which involve using multiple bases to make the prediction. That is why their collaboration increases overall efficacy as the results acquired are far more accurate. These methods used in this study utilize the advantage of having more than one model, which results in high accuracy (Ranjbar, Montazeri, Farashah, et al., 2023b). For instance, a compound model could be developed using data that are quite simple to obtain, such as Symphysis-fundal height, maternal weight gain and basic demography data (Sanchez-Martinez et al., 2024a).

The data can be used to train machine learning models, such as Random Forests, Gradient Boosting Machines, and Neural Networks. Through applying these different models, we can

have a final prediction that utilizes the strengths of each base model. This assists in minimizing any possible biases and errors that might result from relying on just one single conventional approach (Rubaiya et al., 2024).

More importantly, flexibility of the system could also be incorporated to meet the societal needs and be acceptable to the healthcare workers in limited resource settings (Sanchez-Martinez et al., 2024a). It may just be possible that a simple Smartphone system could be downloaded and be directly incorporated in the existing EMR system and provide with mere completed data would produce an approximation of the birthweight. It would also assist in the evaluation of high-risk pregnancies that need further examination to ensure they are given attention that they need as soon as possible (Sanchez-Martinez et al., 2024a). In addition, the method will result in the improved prediction of birthweight and will also provide the opportunity to make timely interventions and improve resource allocation. To carry on with this process of improving the demographic status of the neonatal population in Kenya, it is relevant that the ensemble model should be revised every now and then. In this manner, it will be able to adjust to the changing nature of the population and local specifics. By doing so, we develop a sustainable and efficient solution that would meet the needs of mothers and their babies.

Several scholars in Kenya have conducted studies on various antecedents of low birth weight of infants. The underweight babies have, however, been unpredictable among others throughout the child-making process. In the prediction of low birth weight, a machine learning model will prove suitable with the assistance of the maternal risk factors identified in the previous research. Such an approach could be most helpful in enhancing the outcomes of maternal and infant care.

This work is an attempt to bridge the research in the novel area of machine learning and the practical need for managing limited-resource settings. In other words, we are interested in

developing predictive machine learning models for estimating the probability of low birth weight (LBW) in Kenya. This paper aims at providing useful information for healthcare providers by presenting the major aspects of the maternal and socio-economic factors discussed in the previous research. We envision that the information to be obtained from these paradigms will, in some way, help enhance the health of the mothers and their children.

1.2 Statement of the Problem

Low birth weight (LBW) remains a significant global and national health concern despite the implementation of various interventions. This widespread issue prompted the adoption of global nutrition targets in 2012, including a commitment by member states of the 65th World Health Assembly (WHA) to reduce LBW by 30% between 2012 and 2025. However, current progress indicates that the world is not on track to achieve this goal. A joint report by WHO and UNICEF on global trends in LBW prevalence between 2000 and 2015 revealed a slower rate of reduction in LBW prevalence from 2010 to 2015 compared to the period between 2000 and 2009. The report also projects that if the current average annual reduction rate of 1.00 percent continues, the LBW prevalence, initially projected to be 10.5 percent, would instead reach 13.2 percent by end of 2025 (*UNICEF-WHO Low Birthweight Estimates 2019 .Pdf*, n.d.).

Low Birth Weight (LBW) poses a significant threat to newborn health and development, contributing to neonatal mortality and long-term health complications. Globally, over 20 million cases of LBW occur annually, demanding effective strategies for early identification and prevention. While the prevalence of LBW in Kenya is reported between 11-12%, the specific burden and contributing factors in Kenya require focused attention and context-specific interventions. Accurate prediction of LBW risk during prenatal care is crucial for enabling timely interventions and improving maternal and fetal health outcomes.

Globally, approximately 14.7 percent of all newborns, roughly 19.8 million babies, were born with low birth weight (LBW) in 2020 (*UNICEF-WHO Low Birthweight Estimates 2019 .Pdf*,

n.d.). Low- and middle-income countries (LMICs) bear a disproportionate burden, accounting for over 90% of these cases.

In Sub-Saharan Africa, the pooled prevalence of low birth weight (LBW) was measured at 9.76% (Tessema et al., 2021). LBW prevalence varies across the continent: a study including data from multiple African countries indicated the highest LBW was recorded in Ethiopia (16.21%), while the lowest was in Guinea (6.30%).

Low birth weight (LBW) remains a significant public health concern in Kenya, impacting both infant and maternal health¹²⁴. While global efforts aim to reduce LBW, the condition persists as a major contributor to infant morbidity and mortality, with long-term consequences extending into adulthood (Keino et al., 2024).

Despite interventions, LBW prevalence remains high in Kenya. The World Health Organization (WHO) and UNICEF reported LBW estimates of 11% and 6% based on the 2009 Kenya Demographic and Health Survey. Recent data shows differing prevalence rates in various regions and hospitals.

Children often experience both long-term and short-term repercussions as a result of LBW. Premature birth is one of the primary factors of LBW that kills 1.1 million babies annually. The infants with LBW are normally faced with severe malnutrition, stunted development, or wasting. Low IQ and impaired academic performance are also found in such infants (Martin-Herz et al., 2023). According to a study by (Keino et al., 2024), the children born short of gestational age SGA might experience more problems at school, particularly during the period of 12-to-18 years. This indicates that any problems that the individuals experience during the developmental years may have some implications for academic performance at later stages of life. It implies that LBW is not only a complicated beginning but also more of a concern in

terms of susceptibility to problems in long-term health, as they are at a higher risk of developing diabetes and obesity.

Accurate forecasting of the weight of a baby at birth is essential, particularly among medical practitioners, as they can timely identify and manage the problem of LBW early. Early intervention can be very beneficial to the well-being of the mother and the child (Grbic et al., 2024). The methods that have been developed to estimate the birth weight of a child in Kenya have their shortcomings, depending on the hospitals.

Such techniques as palpation of the abdominal area and symphysis-fundal height measurement can be simple, but they are subject to individual judgment. Such and similar approaches may mean that the baby is not receiving timely attention and care. In the remote locations where accurate methods of medical technologies have not been developed, it is best to estimate the birth weight using more accurate methods.

Research devoted to predicting low birth weight with machine learning models has been of tremendous importance. However, these models often face great challenges, including issues of generalizability to different populations and settings, gaps in datasets, and limited sample sizes. Linking maternal socioeconomic status and health data sets creates complexity and non-linearity, necessitating a solid model to capture the underlying dynamics.

Wide gaps are also noted in the use of clinical decision support systems in antenatal care clinics in Kenya. This is brought about by consistent negligence in the design, assessment, and application of these systems in the country's health facilities. Low socioeconomic status significantly limits access to adequate prenatal care in Kenya, coupled with a lack of quality healthcare infrastructure and reliable data. There is a need to address the risk factors for maternal health, which have become prevalent in the country. This exemplifies the pressing need of the Kenyan population to have prediction models that are more precise and dependable.

The healthcare system in Kenya shows a marked gap between local practice and global competencies. Currently, local Kenyan research is lacking as it relies purely on traditional logistic regression models, which ignores the highly complex nonlinear intertwining of relations consisting of cultures, socioeconomic status, and health outcomes (Duba et al., 2024). Some recent works argue that it is best to use ensemble class machine learning models, which add more than one model to address a given problem and increase the accuracy of predictions exponentially, and these non-linear intricacies will be dealt with more effectively (Ranjbar, Montazeri, Farashah, et al., 2023b). However, there is little to no research validating such claims in Kenya.

For this purpose, we attempt to overcome some of the most challenging barriers to LBW underestimation issues for pregnant women in Kenya. This study seeks to improve prediction accuracy, reduce overfitting, and manage excessive complexity in healthcare data by employing robust ensemble class machine learning techniques, including bagging, boosting, and stacking. Ultimately, we set out to develop a real-time Clinical Decision Support System (CDSS) that gives out information and guidance instantaneously during treatment procedures and can be seamlessly integrated into local healthcare systems.

1.3 Objectives of the Study

1.3.1 General Objective:

To develop an ensemble machine learning model that optimizes the prediction accuracy of low birthweight risk in Kenya.

1.3.2 Specific Objectives:

1. To analyze the key risk factors associated with low birth weight through model interpretation and analysis.
2. To develop and assess individual machine learning models against the ensemble approach using appropriate train-test splits and cross-validation.
3. To assess the performance of ensemble methods in improving predictive accuracy.

1.4 Research Questions

1. What are the key maternal and socioeconomic factors contributing to low birth weight, as identified by the models?
2. How do ML models compare in forecasting the risk of low birth weight?
3. Can ensemble methods enhance the predictive accuracy of low birthweight risk compared to individual models?

1.5 Justification of the Study

The scientific rationale for this study is based on the urgent need to tackle low birth weight (LBW) as a significant public health issue in Kenya. Approximately 11.5% of infants in the region are born with LBW, resulting in considerable effects on both infant mortality and long-term health outcomes. Despite progress in predictive analytics through machine learning (ML), the use of these technologies in low-resource settings like Kenya is still limited, creating a crucial gap in effective maternal and child health interventions and child health interventions.

Pregnancies in Kenya are tracked with ultrasound scans, symphysis-fundal measurements, and abdominal palpation techniques. Obstetric ultrasonography, like other methods, often ignores the intricate multifactorial relationships that cut across risk factors associated with lower birth weight (LBW). These socioeconomically determined factors include cultural practices, financial standing, and the mother's health. It is necessary to note that the combination of these factors, instead of acting independently, can interact to exert an altering significant load, both on the mother and fetus (Tessema et al., 2021). Accessing, especially stable and valuable information in rural locations is a particular challenge. The predictive models that can be created to address these challenges are limited by the available, accurate data (Musau et al., 2023). This study aims to eliminate these shortcomings so as to forecast LBW more precisely with a strong machine learning model.

Recent studies show that predicting low birth weight (LBW) is more precise when ensemble machine-learning algorithms are used. The Random Forest model, for example, has shown LBW-associated factors predicting accuracies as high as 85%. In the context of healthcare datasets, the XGBoost algorithm has also demonstrated remarkable effectiveness in handling data, which is a common issue encountered in healthcare datasets. Deep Neural Networks, alongside other deep learning models, are quite capable of capturing complex, intricate, and non-linear relationships between outcome variables such as birth outcomes and health markers of mothers. Most of the research has predominantly been conducted in high-income countries. This signifies a major gap in understanding the applicability of these models in low-resource countries, such as Kenya, which may have substantially different healthcare systems and associated challenges

Kenya's diversity in culture and socioeconomic status presents certain formidable concerns with maternal health. Health outcomes such as the availability of care and the offer of services to citizens from other countries create a regional disparity. There is a need for tailored

approaches aimed at solving crucial challenges specific to women in specific locations. This is what the present study intends to achieve: develop ensemble machine learning models aimed at addressing low birth weight in the Kenyan context by incorporating relevant maternal socioeconomic factors identified in previous studies. We aim to improve the lives of women and their infants by contributing to better health outcomes.

The findings from this study can potentially be valuable for formulating strategies in public health by harnessing the latest developments in machine learning, integrated with real-life use in poor countries. The data collected is meant for designing specific interventions that can optimize the health outcomes of mothers and their children. Additionally, the research provides a comprehensive model that can be applied to similar studies in other developing countries. The goal is to equip health practitioners with comprehensive evidence to guide them in the management of maternal care services.

1.6 Scope of the Study

This study examines the different significant fields in population health in relation to maternal and child health in Kenya. An example is that, the LBW rates in the newborn babies stand at about 11.5% implying that the same percentage of babies are below 2,500 grams in weight at birth. We will investigate the socioeconomic conditions in Kenya, discrimination within the health system and the cultural practices that contribute to LBW. We would like to examine the socioeconomic influences, the biases of the healthcare system, and cultural practice in Kenya that led to LBW. It is our hope that we can solve the LBW issue by creating a machine learning healthcare model that would work in the Kenyan healthcare system.

The research focuses on secondary analysis of the newly published 2022 Kenya demographic and health survey (KDHS), which has important information on socioeconomic status and maternal health outcomes in Kenya throughout years. The objective of the research is a comprehension of the current risk factors, and their effects on both mother and child through longitudinal examination of the trends of LBW relative to the existing historical information.

Among the notable gaps to be addressed, there will be maternal risk factors of LBW, such as age, height, weight, obstetric history, and the number of antenatal care visits in total. We shall also compare how socioeconomic status which includes wealth index and employment impact the rate of LBW outcomes. This broad study will produce results of numerous interdependent variables that will illuminate the complex burden of LBW. In order to sustain the plausibility of the study, we need to describe how to preserve data integrity towards missing data points and outliers and how they might affect the accuracy of data. Under methodology, the study aims to design and implement an ensemble machine learning model for predicting the risks of low birth weight (LBW) using contemporary machine learning methods. Random Forest, logistic regression, XGBoost, and neural networks are base learners that will be trained. Afterward, we will empirically test the predictive power of the base learners and determine whether using the ensemble techniques like bagging, boosting or stacking enhances LBW prediction accuracy in Kenya. Primary indicators to be assessed to measure the predictive power of the models include accuracy, precision, recall and the F1 score.

The overall objective in this case study is to analyze various ML algorithms and evaluate whether ensemble strategies outperformed single-base algorithms in terms of LBW prediction accuracy.

1.7 Limitations of the Study

This study aims to develop and validate an ensemble machine learning model to predict LBW risk in Kenya. It has, however, glaring limitations that may affect the outcome and the relevance.

The study will be using secondary data from the Kenya Demographic and Health Survey (KDHS) of 2022, which provides insightful information about maternal health. Prior research has emphasized issues with data accessibility, particularly in Kenya's rural areas, which may

limit the analysis's depth. There may be problems with data quality, such as missing information, errors, or discrepancies in the reporting of the data.

Focusing this study primarily on Kenya could make it difficult to apply these findings to other comparable low-resource countries. Due to Kenya's distinct socioeconomic circumstances, healthcare system, and cultural customs, which may differ from those in other countries, result interpretation should be done cautiously.

Although ensemble machine learning models can significantly improve prediction accuracy of LBW, their complexity may also make them difficult to understand, especially for healthcare professionals without a strong technical background, but must comprehend and apply these models in practice.

This study analyses quantitative data using machine learning techniques, which are excellent for identifying trends and predicting results. This strategy, however, may not capture important qualitative elements like cultural beliefs, personal habits, and community dynamics that influence LBW risk. These approaches should consider qualitative components, as this can aid in understanding better the entire scope of factors associated with LBW.

While the study only utilizes data up until 2022, changes in healthcare policies, maternal health practices, or socioeconomic conditions might influence the rates and risk of LBW variables over time. Relying exclusively on this timeframe will undoubtedly mask evolving patterns or emerging issues in maternal and pediatric care.

Although the KDHS 2022 dataset is expected to miss other critical components like access to nutrition, environmental factors, and healthcare infrastructure, it will capture essential maternal and socioeconomic determinants of LBW.

CHAPTER TWO

2 LITERATURE REVIEW

2.0 Introduction.

Low birthweight (LBW) children, defined as those born weighing less than 2,500 grams, present a significant public health challenge worldwide, including in Kenya. LBW is associated with higher risks of serious health conditions and developmental complications, which strain healthcare systems and economies. Addressing LBW requires identifying risk factors to enable preventive measures, especially in resource-limited settings like Kenya, where timely intervention can improve outcomes for mothers and infants.

Ensemble machine learning models, which combine multiple algorithms to improve predictive accuracy, are gaining traction in healthcare analytics. While they have shown promise for LBW risk prediction globally, their application in Kenya remains underexplored. This review focuses on recent literature (2010–2025), highlighting the potential of ensemble models to enhance LBW prediction and identifying gaps for future research.

2.1 Background

Each year, hundreds of thousands of newborns in Kenya grapple with the daunting reality of being born with low birth weight (LBW). This disorder is defined as a baby who is born weighing less than 2,500 grams, and it can have a significant effect on their development and health later in life. It is alarming that an estimated one in eight newborns in the country is born with LBW, approximately 12%.

But these numbers show only part of the picture. In the countryside, where it is usually very difficult for them to get the necessary prenatal care and good food, things could be even worse. There, the mothers usually do not have the resources and care they need so that they and their

babies can prosper. That makes it even more vitally important to provide superior health care assistance to see that every child, regardless of birthplace, gets the best possible life beginnings (Kimani-Murage et al., 2011).

In Kenya, many mothers stand the challenges of the low-birth-weight problem because they have to weigh the extra burden of providing nutritionally proper food to their homes daily an issue that is even more challenging when the food supply suddenly runs out. This dilemma is not only a political issue of abstraction, but to many, it is a non-stop battle that kills their emotions and psyches. Take, as an example, a mother who is struggling with economic poverty but at the end of every night she is anxious about how she will feed her growing child. Heartbreaking that getting to medical facilities can be extremely difficult, especially when they are miles away.

Mothers from poorer households in coastal Kenya are more at risk, according to recent studies, leading to the heartbreaking phenomenon of delivering babies with low birth weights. This is not just a statistic, a young mother's dreams and aspirations, and potential for her child, and many more. It has a ripple effect, ensnaring families in poverty and health difficulties that can last for years. These are real lives, full of honest struggles and resilience, that reveal the urgent need for basic resources and support for people living in the most underserved communities. Each story is a cry for help and understanding, reminding us that there is so much more that needs to be done to bring these families out of this cycle (Mirieri et al., 2024).

It is an especially intriguing issue because of its general implications. Other than short-term health issues, low birth weight babies are more likely to start their lives at a rougher end. It has been found that such children can experience additional issues with their physical development, and their mental development can be more challenging; they can be at a higher risk of developing chronic illness decades later (World Health Organization, 2024). LBW has been of

concern because the world is pushing to solve it as it is reflected in the United Nations Sustainable Development Goals including SDG 3 on good health and well-being which strives to ensure healthy lifestyles are promoted to everyone.

With these problems, the Kenyan medical fraternity is highly advised to combat LBW. Nevertheless, most of the traditional LBW prevention and prediction measures have tended to concentrate on the risk factors of an individual with LBW, yet most communities do not access full prenatal care. Luckily, the current technology provides new hope due to the developed computational approaches like ensemble modelling where many prediction models are combined to provide more accuracy. These are especially helpful as the maternal health outcomes in the United States are strongly dependent on the socioeconomic differences (Dietterich, 2000).

Interesting prospects have also been created, especially using methods of machine learning like boosting ensemble tools like XGBoost (Chen & Guestrin, 2016) which directly can address LBW at the Kenyan population level. With these advanced tools, it is possible to scan large volumes of health information to determine correlations and risk factors that might have been missed with the traditional statistical tools. When medical resources in a country tend to be quite endangered, implementing state-of-the-art methods to maternal health data could allow healthcare providers to detect high-risk pregnancies earlier and with greater precision, which would be the difference between life and death to most children.

The battle against low birth weight in Kenya is both a battle of the heart, hope, teamwork and perseverance, much more than a war against medicine. It entails uniting the communities, health care workers, researchers and families with the aim of providing these children with the best possible start. We can take the traditional knowledge and integrate it with modern technologies and have a tremendous chance to do something really good to Kenyan families

with troublesome issues related to low birth weights. And through that coming together and combining of our efforts, we can see a day when every child will have the chance to receive the love, care and opportunity they deserve to succeed.

2.2 CONCEPTUAL FRAMEWORK

Mothers and babies will benefit from thorough research on LBW and its causes, as it may help construct relevant policies. Studies have tested different factors, and they can be categorized into three broad categories. These include maternal characteristics, socio-economic situation, and social-demographic factors.

The social-demographic group divulges the individual profile of a mother that is highly likely to dictate the milestones of her child. Her weight, height, and even age are significant contributors to the growth of the child. Access of the required medical attention and even the most basic food and safety, may be dictated by her whereabouts, be it a city or a village. Education can also do a lot because educated mothers will most probably be aware of different health policies, and thus, they will take part in prenatal care to the fullest.

Employment of the mother is a complete of the socio-economic policies and determines whether she has the capacity to meet the medical, dietary and basic needs during pregnancy. The other important factor that gives an idea about the challenges that a mother may face to sustain her health and that of her babies is total family income.

Obstetric variables also make us know what a woman goes through and does during pregnancy. The prenatal care visits she makes, her intake of iron and folic acid supplements, whether she is a smoker, her age, and her body mass index (BMI) prior to getting pregnant are very important factors that affect the pregnancy.

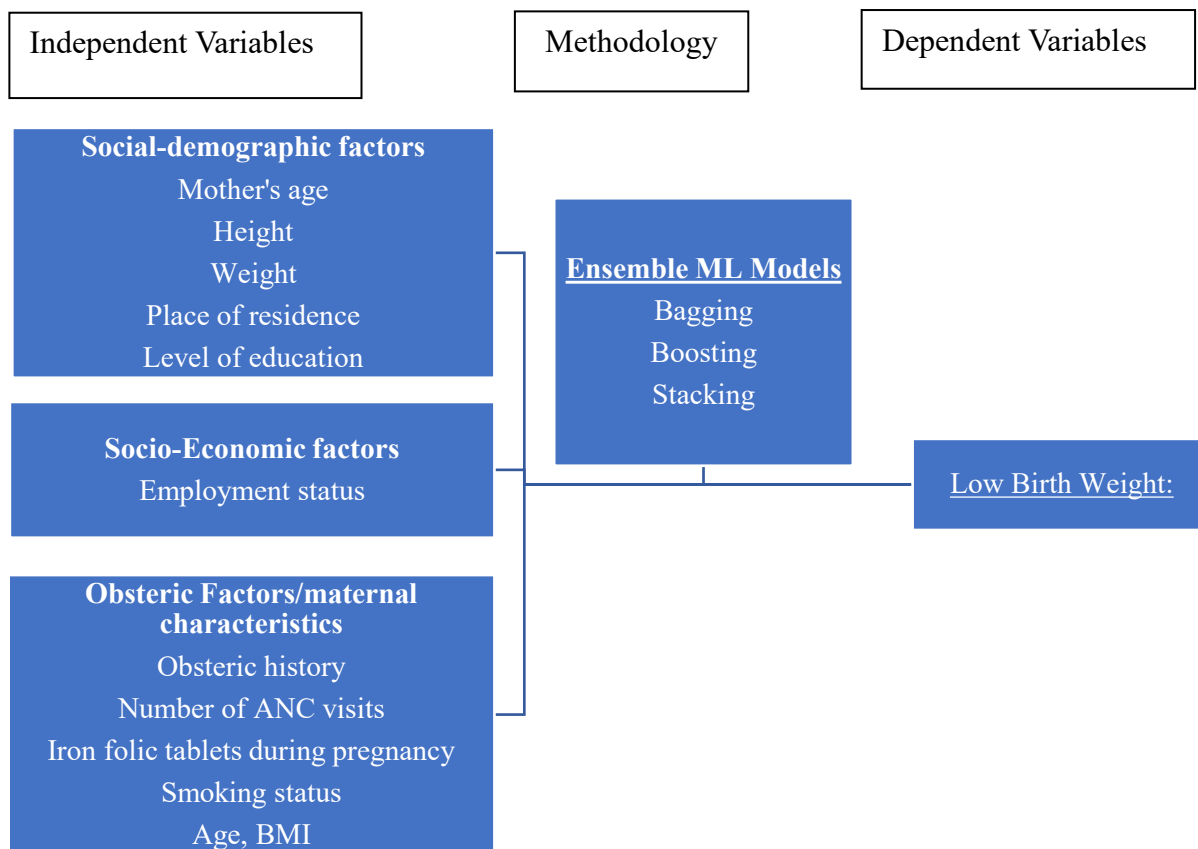


Figure 1: Conceptual Framework

2.3 Theoretical Framework

This study is grounded in the principles of the Design Science Research (DSR) paradigm, which focuses on the creation and evaluation of innovative artifacts intended to solve identified problems within specific contexts. DSR provides a systematic process for developing and rigorously validating computational models that address real-world challenges, here, accurately predicting low birthweight risk in Kenyan maternal health settings. The study follows the DSR methodology's core steps: problem identification, objective formulation, artifact design and development, demonstration, rigorous evaluation, and communication of results. This approach ensures the predictive models are not only scientifically robust but also contextually relevant, usable, and implementable within Kenyan healthcare systems.

Complementing DSR, predictive modeling theory underpins the ensemble machine learning approach used in this research. Ensemble methods, bagging, boosting, and stacking, combine multiple base models to improve predictive performance by reducing bias and variance and capturing complex nonlinear relationships inherent in health data. These model fusion techniques are particularly well-suited for clinical risk prediction, where interactions among maternal socioeconomic, obstetric, and biological factors are multifaceted and nonlinear.

Within health informatics, this framework aligns with the goal of developing data-driven clinical decision support systems that enhance maternal and child health outcomes. By integrating machine learning with locally relevant health data, this study contributes to advancing health informatics capabilities in Kenya, promoting interpretable and actionable predictive tools that support clinicians and policymakers in resource-constrained environments.

Together, the Design Science Research paradigm and predictive modeling theory form a robust conceptual basis for the design, development, and evaluation of ensemble-based LBW risk prediction models tailored to the Kenyan context.

2.4 Empirical Review

2.4.1 Socio-Demographic Factors:

It is necessary to start with the understanding of how socioeconomic status, as well as demographic background influences the outcome of pregnancies. The birth weight of children in Kenya directly depends on the maternal age since it is a nation with many social variations among its citizens. Women under the age of 18 years and those women above the age of 35 years are at a greater risk of giving birth to a baby with a low birth weight. This resonates with the necessity to test and promote maternal and newborn health care systems in specific situations (Muchemi et al., 2015). The level of education among mothers has significant health consequences among their children, and may predispose them to giving birth to LBW babies. Research in some Kenyan districts has revealed that among the women, the women who had

attended secondary school or beyond are better health choices. Subsequently, expectant moms tend to have healthier pregnancies and their babies perform better when born. Therefore, education plays an essential role in supporting generations ahead and providing women with empowerment capabilities (Grbic et al., 2024).

Mothers encounter significant influence from their geographical location, as those residing in urban regions generally benefit from ample healthcare facility accessibility. Urban mothers encounter difficulties such as enduring the anxiety of city living, as well as facing environmental contamination problems. Mothers who reside in rural areas benefit from cleaner atmospheric conditions and sound environmental surroundings. The main challenge rural mothers face is getting enough access to specialist medical treatments during critical health needs. A massive scientific study conducted in Western Kenya demonstrated that mothers who live beyond 5 kilometers from medical facilities had their babies with lower birth weights 1.5 times more often. The research demonstrates that attention should be directed toward healthcare location, together with accessibility, when supporting mothers alongside their children (Madeghe et al., 2016).

Research shows that the marital status of a mother impacts neonatal health by affecting birth weight outcomes. Researchers have discovered that single mothers experience more cases of giving birth to LBW babies than do married women. The challenge of this issue grows more severe throughout Kenya because of its special socioeconomic structure. Single mothers usually lack the supportive relationship found in traditional married couples because their access to mutual emotional support strengthens their anxiety and uncertainty. These women face two major difficulties because financial strain keeps them from securing adequate nutrition and prenatal care needed for a healthy childbirth and delivery.

Research demonstrates that women who are not married usually face economic and healthcare challenges because of their social circumstances. This increases the risk of giving birth to LBW babies, who might experience long-term problems with their development and health. These disparities should be addressed to improve outcomes for mothers and their infants (Sanchez-Martinez et al., 2024a). Most women in Kenya are struggling with challenges of single motherhood or unstable relationships, bringing a lot of emotional pain and uncertainty into their lives. This emotional distress may potentially affect the health of the unborn children. If we hope to design effective interventions, we must consider how a woman's marital status interacts with other social and economic characteristics. Understanding these dynamics leaves us with a greater ability to work towards reducing low birth weight (LBW) cases and ultimately improving health outcomes for mothers and children in Kenya (Kamau et al., 2018). Healthcare professionals should recognize the challenges faced by single women. If providers know the unique circumstances of these expectant women, they can better help them. Women who get this kind of support while they are pregnant tend to have better health outcomes, and their children go on to lead healthier and happier lives. It is about creating the right atmosphere where every mother feels recognized and encouraged.

2.4.2. Socio-Economic Factors

Multiple socioeconomic aspects determine the outcome of LBW. Money serves as a crucial component, yet all factors that determine LBW remain beyond its boundaries. The ability of a pregnant woman to obtain quality healthcare along with nutritious food depends on both her employment status and household earnings. Studying maternity outcomes in Nairobi's informal settlements reveals a disturbing trend because employed women show a 30% decreased probability of bearing babies of low birth weight in comparison to unemployed women. The ongoing employment and stable financial income protect both the mother's and the child's health (Wanjohi et al., 2020).

The current state of food security directly affects whether a child weighs less than normal at birth. Children born into families that struggle to distinguish between nutritious and unhealthy foods face an increased risk of such conditions. Low birth weight delivery rates are 2.3 times higher in Kenyan households experiencing food-related difficulties, according to research from five counties in the country. This direct connection between maternal and infant health demonstrates the essential need for providing healthy food options throughout the entire nation (Amiri et al., 2024). For mothers obtaining health insurance stands as an essential requirement. People with insurance benefits receive needed hospital treatment when necessary while gaining regular pregnancy check-ups. The provided support helps both mothers and their unborn babies achieve better health outcomes.

2.4.3 Obstetric Factors/Maternal Characteristics:

A female who gave birth to an LBW infant faces similar pregnancy complications when she gets pregnant again. A woman's pregnancy spacing stands as a main factor that plays a vital role. A woman who becomes pregnant within her first 18 months of motherhood elevates her risk of delivering an infant with low birth weight. Knowledge about these factors enables people to develop proper plans for healthier maternal health during pregnancy (Wagura et al., 2018).

Hypertension becomes a complex disease management challenge for various pregnant women. Hospital procedures should always assess the blood pressure levels of pregnant mothers because hypertension disorders during pregnancy increase the risk of delivering a newborn with low birth weight. Healthcare providers must closely monitor people with gestational diabetes because this health condition poses risks for maternal-fetal health. A study by Dr Adem explains that present health dynamics need support along with understanding and careful leadership to succeed (Dr Adem Salome, 2020). Mothers taking advantage of the current advanced antiretroviral treatments still encounter distinctive obstacles when striving to prevent

low birth weight in their newborns. A hard journey to establish health stability with a healthy newborn starts and extends indefinitely as mothers need compassionate backing.

Weight control before and during pregnancy benefits the mother, together with her developing baby, as they both need to reach proper weight requirements before pregnancy starts and throughout gestational periods. According to a study, mothers who entered their pregnancy with a BMI below 18.5 had 2.1 times higher odds of delivering babies with low birth weight. The state of our health in pregnancy demonstrates its direct effects on the developing baby. on the health of the fetus (Kamau et al., 2018). Antenatal care, such as the number of prenatal visits together with the first consultation date of pregnant individuals, has major implications for birth outcomes. The correct scheduling of these appointments will boost the chances of positive outcomes significantly.

2.4.4 Ensemble Techniques:

Ensemble machine learning approaches such as bagging, boosting, and stacking combine the strengths of several models into one predictive tool to enhance the prediction accuracy level. This research aims to determine the prediction performance levels of these techniques when estimating low birth weight (LBW).

Many techniques, such as Random Forests and Gradient Boosting models, through different studies, have become popular because they offer reliable results and work well with unbalanced datasets. Their exceptional performance makes them ideal for dealing with this important health issue (Khoshgoftaar et al., n.d.). Multiple studies in Kenya have not adequately investigated the potential performance gains from the ensemble and feature selection combination (Javale & Desai, 2024).

The traditional statistical method, like logistic regression, delivers a lower precision for predicting low birth weight (LBW) risk than ensemble models perform. An ensemble

modelling approach provides superior outcomes in terms of predictive accuracy when two or more models are utilized instead of one individual technique. According to a recent study, the accuracy of Gradient Boosting Machines wowed the researchers with 85%, while logistic regression only produced a 72% accuracy. In this study, maternal age, number of antenatal visits, and socioeconomic status were the most studied factors associated with low birth weight (LBW) (Chopra & Bhilare, 2024).

Recent research concerning ensemble models and their combination with real-time mobile health (mHealth) and electronic health record (EHR) data showed remarkably high improvements in performance. Such improvements to patient outcomes and healthcare delivery could be tremendous. Despite the ongoing discussions, major debates surround not the nature of machine learning applications within this founding field, but rather the moral repercussions of these applications, particularly concerning preserving patient confidentiality or mitigating potential biases arising from algorithmic errors. As we strive to utilize technology appropriately in healthcare, these factors are critical (Lundberg & Lee, 2017).

Though ensemble models have great potential, the use of such techniques in Kenya is limited by the availability of adequate data and access to sufficient computing power. Predictive models in developed countries have, in general, a higher accuracy, highlighting the need for environmental-specific models. They will also explore various ensemble techniques to further enhance prediction accuracy.

2.2.2.1 Bagging:

Bagging, short for Bootstrap Aggregating, is a well-known ensemble machine learning method. This involves taking the average of all the datasets that have been trained on different models trained on different subsets of the data. This approach is helpful as it reduces model

variance hence preventing overfitting, thus helps to in enhancing and making reliable predictions.

Models have become much more stable with bagging techniques. They can reduce prediction variance (as compared to individual models) by approximately 15-20% when used instead of individual models. This enhancement, which improves the overall accuracy and reliability of the forecasts, has been substantiated during the asynchronous assessment of forecasts in (Soloff et al., 2024). Ensemble machine learning models, such as those produced by XGBoost and Random Forests, have been shown to improve results across different demographic groups (Brown & Benjamin, 2024).

2.2.2.2 Boosting:

Training a series of models one at a time, where reference to previous trained models is assumed, and the subsequent model training aims to correct the errors of the previous trained model. This technique enhances predictive performance by aggregating a series of weak learners into a strong predictive model.

Previous studies suggest that by employing boosting strategies, model accuracy may increase by 5–10% over base models (Singh et al., 2021). Gradient boosting methods have been especially effective when modelling complex clinical datasets (Iranzad et al., 2022).

2.2.2.3 Stacking:

It is an approach that involves training several models separately and then combining their predictions using another model. It resorts to the individual advantages that each algorithm has, which, ultimately, improves the performance and accuracy. It should be noted that the improvement in the medical prediction exhibitions with the suggested stacking meta-models has been approximately on a 3-7% margin compared to that of a single model (“(PDF) A Stacking Ensemble Model to Predict Daily Number of Hospital Admissions for Cardiovascular

Diseases,” 2024). This is one area that multiple models integration has performed brilliantly to effectively process heterogeneous medical data (Numan Ahmad et al., 2023).

2.2.3 Comparison of Traditional Models and Ensembled Models:

Studies indicate that machine learning techniques like Random Forest, Logistic Regression, XGBoost, Neural networks, and others are very useful when forecasting risk factors related to low birth weights (LBW). These models tend to provide a greater degree of accuracy than more conventional statistical procedures and they are capable of addressing more complex and non-linear patterns of the relationships between the various maternal risk factors and the birth outcomes.

2.2.3.1 Random Forest:

Random Forest is a general-purpose machine learning framework that uses multiple decision trees in the learning process and then does majority voting to get predictions. The outcome of the prediction is dependent on every tree making a vote unto its own using majority preferential selection. The approach proves to be sufficiently successful in revealing complex relationships among several variables. This approach has been demonstrated to be unmatched in increasing LBW risk factors by 85 percent accuracy, as of benefits in understanding complex relationships due to their superiority (Hong et al., 2022).

Random Forest models proved to be effective when dealing with medical data, which is abundant in risk factors according to a research by (Vlachas et al., 2022). The models are characterized by predictable accuracy of regularity ranging between 85 and 90 per cent in predicting LBW cases (Iwendi et al., 2020). The main benefit associated with these models is related to the fact that they can identify significant data characteristics and are suitable in the presence of missing items. Their specific benefits to medical facilities are the understanding of complex relationships due to which Random Forest models perform well (Hong et al., 2022).

2.2.3.2 Logistic Regression:

The application of logistic regression is a strong foundational technique in comparison to advanced machine learning techniques since it is a proven statistical predictive model applicable to binary classifications. Interpretability can turn out to be one of the primary strong points because the method enables researchers to obtain results that they can interpret, even when such an approach is required within the sector, such as healthcare. Proper model implementation is a challenge by the implementation of proper non-linear interaction modelling. The accuracy rate of traditional logistic regression models identifies LBW outcomes to levels of 75–80%, as shown by (Victor et al., 2024) and (Sanchez-Martinez et al., 2024b). (Victor et al., 2024) and (Sanchez-Martinez et al., 2024b). Medical practitioners benefit significantly from logistic regression because this tool provides both high accuracy levels and understandable results. Medical practitioners can make better clinical decisions through their better understanding of the varied risk elements affecting LBW (Reza et al., 2024).

2.2.3.3 XGBoost:

The serious LBW prediction issue is addressed with the help of the advanced machine learning method XGBoost that is able to solve it with its potent functioning. XGBoost is unique due to the support of missing data and the ability to process very large volumes of data, which are inherent to the research within the medical sphere. Its ability to optimize allows XGBoost to possess high-level features, and therefore it makes it a suitable tool in this particular form of prediction (Sanchez-Martinez et al., 2024b). The results of the testing revealed that XGBoost yielded better prediction results ranging between 88 and 92 percent of LBW, as well as the fact that various studies had confirmed the strong prediction performance (Victor et al., 2024) (Reza et al., 2024). The high processing speed of large data, along with an outstanding level of operations, makes the model suitable in health trend studies that impact a large population

(Reza et al., 2024). XGBoost provides helpful results to the beginning of significant health conditions.

2.2.3.4 Neural Networks/Deep Learning

Deep learning models within neural networks display superb abilities to recognize complex linkages among different variables with their corresponding results. Their distinctive structure enables the identification of subtle data patterns that would remain unnoticed by alternative techniques (Mizuno et al., 2023). Deep learning employs multiple parameter analysis and their mutual interactions, which helps improve LBW prediction according to recent studies.

2.2.3.4.1 Neural Networks:

In the case of Deep Neural Networks, an illustration of executing Artificial Intelligence, the models have shown outstanding performances in decoding several maternal risk indicator links. The results of the research prove that these systems are characterized by the accuracy of the performances, ranging between 90 and 95 percent (Mutlu et al., 2023). Multi-layer perceptrons are one of the different types of neural networks that are capable of handling non-linear interactions presented by medical datasets, as healthcare data needs (Geethanjali & Valarmathi, 2024). In their results, Chinese investigators used a genetic algorithm to optimize the parameters of the back-propagation neural network in predicting foetal weight. The level of accuracy attained with the application of this method was 76.3, which was 14.6 higher than traditional methods. Machine learning tools on the state-of-the-art level show that they are able to enhance healthcare outcomes.

2.2.3.4.2 Deep Learning:

The LBW prediction accuracy is already at levels greater than 93, and the latest deep learning breakthroughs have revived the medical prognostic technique (Khadidos et al., 2024). Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) require long

and complex medical text to be evaluated over time, providing them with significant use in medical applications (Mennickent et al., 2023).

Table 1: Summary of Literature Review.

Study (Year)	Model(s) Used	Dataset / Setting & Key Features	Performance & Findings	Strengths	Limitations
Yu Lu et al. (2019)	Random Forest, XGBoost, LightGBM (Hybrid)	Dataset: Ultrasound data & maternal information Key Features: Maternal BMI, uterine height	Achieved a 3% reduction in prediction error compared to ultrasound alone. Demonstrated that combining multiple ensemble algorithms leverages complementary strengths to improve fetal weight estimation.	Hybrid approach enhances accuracy by integrating diverse models and clinical features.	Reliance on ultrasound data limits applicability in low-resource settings like Kenya.
Mario W. et al. (2018)	Subspace K-Nearest Neighbors (KNN)	Dataset: Real pregnancy database Key Features: Hypertensive disorders, Apgar score	Achieved 82.9% accuracy and AUC of 0.669. While accuracy was reasonable, the relatively low AUC suggests limited discriminative ability, possibly due to feature selection or algorithm sensitivity.	Simpler ensemble variant with reasonable accuracy.	Lower AUC indicates challenges in capturing complex nonlinear relationships.
U.S. Study (2022)	XGBoost	Dataset: U.S. hospital data Key Features: Socio-demographics, prenatal health metrics	Achieved an F1-score of 0.82, indicating balanced precision and recall. Incorporated both clinical and socio-demographic factors, reflecting a holistic approach to LBW risk prediction.	Balanced performance suitable for clinical decision-making.	Data from high-resource settings; may not generalize well to LMICs with different population dynamics.
Kapila & Saleti (2023)	Stacking Classifier (Random Forest meta-learner)	Dataset: Cardiotocography (CTG) data Key Features: Fetal heart rate (FHR), uterine contractions	Achieved high accuracy of 94%. Demonstrated the power of stacking ensembles to combine base learners for superior performance in fetal health monitoring.	High precision in detecting fetal distress patterns.	CTG devices are expensive (~\$8,000), limiting feasibility in LMICs like Kenya.
Nurul Fathanah et al. (2024)	Ensemble Bagging Trees	Dataset: Maternal health records Key Features: Maternal age, hemoglobin levels, pregnancy history	Achieved 85.6% accuracy. Bagging reduced variance and improved model stability using routinely collected maternal health data, enhancing feasibility in resource-constrained settings.	Utilizes accessible maternal health data, suitable for low-resource environments.	Accuracy lower than CTG-based stacking models; limited by quality and completeness of health records.
Gao et al. (2025)	Ensemble of Ridge Regression, SVM, Random Forest, XGBoost, MLP	Dataset: Third-trimester birth weight data (Shanghai) Key Features: 17 easily obtainable clinical features	Achieved 81% accuracy. Different base models excelled in different birth weight intervals (e.g., XGBoost and RF in low birth weight), while the ensemble balanced these strengths for superior overall prediction.	Demonstrates importance of model diversity and ensemble integration to address heterogeneous outcomes.	Study conducted in a high-resource urban setting; transferability to LMICs requires validation.

2.7 RESEARCH GAPS

The Kenyan healthcare system requires a validated best ensemble set up to predict low birth weight (LBW) even though ensemble machine learning models have potential in the same. There is no clear evidence on whether direct ensemble method should be applied to Kenya due to lack of data about Kenya and due to the lack of resources, variables that are specific to the population, and the complexity on the healthcare system. There is scant research work done by researchers in creating ensemble networks using alternative base learners, weighting strategies, and selection strategies to predict LBW risk using Kenyan antenatal data collection in hospitals. This assessment must be achieved by modelling ensemble performance and integrating studies on the key variables which influence Kenyan populations such as maternal nutrition, infectious disease incidences as well as socioeconomic indicators and prenatal care access.

In Kenya, healthcare system requires an established best ensemble setup that can predict low birth weight (LBW) even though ensemble machine learning models have potential in this scenario. There is ambiguous evidence on the appropriateness of the application of direct ensemble methods to Kenya due to lack of data, limited resources, variables that are specific to the population, and complications within the health care systems. Researchers have conducted minimal work in developing ensemble networks with alternative base learners, including weighting approaches and selection methods for predicting LBW risk through Kenyan hospital through antenatal data collection. The assessment should be accomplished through modeling ensemble performance and incorporating research on essential variables affecting Kenyan populations, including maternal nutrition, infectious disease occurrences, and socioeconomic measurements, along with prenatal treatment access.

CHAPTER THREE:

3 RESEARCH METHODOLOGY

3.1 Introduction

This chapter provides an overview of the research design used for the study, which sought to develop and validate an ensemble machine learning model for improving the predictive accuracy of low-birth-weight risk prediction in Kenya. This is a very informative and interesting paper on the DSR (design science research), in which they are building upon the DSR framework which is focused on creating tools designed to address problems and to assess these tools. Given that the projects we are involved in aim to develop a clinically meaningful predictive model that can be used clinically, this is especially relevant to us.

In this study, we will position the data in such a format that results in the output variable, low birth weight, being formatted in a binary way. This implies that our ML process will rely on supervised learning, wherein our data is labelled, and the model identifies trends and makes predictions on them. In supervised learning, the model finds the patterns and predictions based on training the model with labelled data.

3.2 Design Science Research Framework

The Design Science Research (DSR) framework was chosen over experimental or correlational designs due to its suitability for developing and evaluating innovative artifacts, in this case, predictive ensemble machine learning models, which aligns with the study's goal of creating a practical clinical decision support system for low birthweight risk prediction.

Unlike experimental designs that manipulate variables to establish causality under controlled conditions, DSR focuses on building and rigorously evaluating artifacts to solve real-world problems. This is more appropriate here because the study is not primarily about testing causal

relationships but about designing a functional, implementable predictive model within the complex healthcare context in Kenya.

Correlational designs assess associations between variables but do not produce actionable artifacts or solutions. DSR, however, involves iterative development, evaluation, and refinement of an artifact (the ensemble model), emphasizing utility, innovation, and relevance to stakeholders, especially in health informatics.

Therefore, DSR provides a structured approach to not only model construction but also to evaluation through computational experiments and real-world applicability, addressing methodological rigor and contextual relevance more comprehensively than traditional experimental or correlational designs.

The different steps in DSR are presented in Figure 2 below:

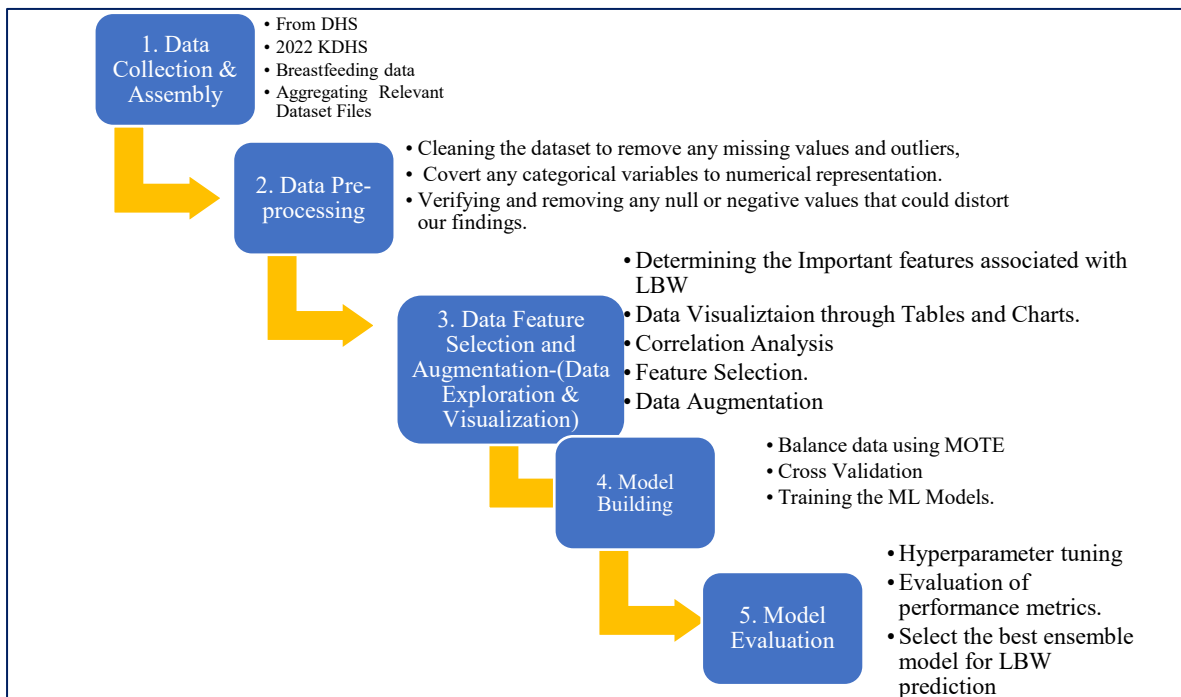


Figure 2: Research Methodology.

3.2.0 Problem Identification

The effectiveness of the method of increasing the accuracy of the low-birth-weight risk prediction in Kenya is discussed in this work. Low birth weight is not simply a statistic but a major public health concern that may lead to long-term consequences to the newborn and the family. But with increased accuracy in forecasting we can also avail these vulnerable newborn baby's swift and effective cure. To enhance prediction and save lives, this paper suggests that one should apply an ensemble method that combines numerous machine learning methods.

3.2.1 Objectives of the Artifact

The analyzed phenomena led to the creation of an enhanced ensemble machine learning (ML) model, which achieved the following aims:

1. The capability to estimate infant low birth weight risk can be enhanced.
2. The analysis identifies crucial maternal characteristics together with socioeconomic conditions that cause low birth weight in infants.
3. Provide a definite guideline that can influence decision-making among healthcare professionals.

3.2.2 Research Paradigm

It is a realistic project that will prove the applicability of machine learning to address the real-life problem of low birthweight (LBW) in Kenya. When we approach data with both qualitative and quantitative methodology, it provides insight to the data and sparks us to action toward change. Our project will have an objective to enhance the well-being of the affected (Allemang et al., 2022). The research implies the implementation of the state-of-the-art machine-learning models into the local healthcare to influence the future of care, and it is relevant to the developments in predictive analytics. We plan to use what is found in theory to the local hospitals and clinics.

3.2.3 Research Design

The current research used a mixed-methods approach with quantitative and qualitative methods to investigate how maternal and socioeconomic factors affect LBW outcomes. We created and evaluated ensemble models on machine learning algorithms of Random Forest, Logistic regression, XGBoost, and Bagging, Stacking and Voting Classifiers (soft and hard voting). We

used these platforms to forecast LBW based on established risk and causal indicators. Using quantitative and qualitative approaches, we obtained a better understanding of how maternal and socioeconomic factors interlink with low-birth-weight performance. The combined approach is meant to inform focused interventions that will promote maternal and neonatal health.

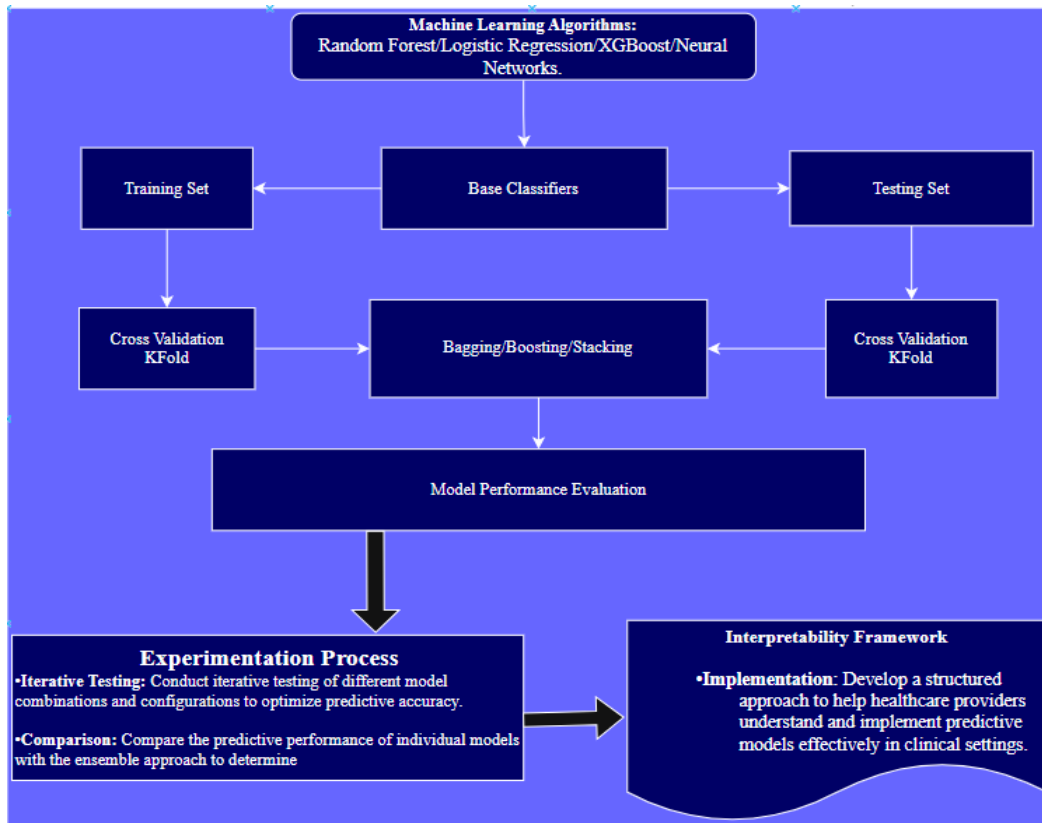


Figure 3: DSR Research Design

3.2.4 Location of the Study

The study will be hosted in Kenya, which faces major public health challenges, as seen primarily in maternal and child health. One worrying statistic is that approximately 11.5% of births in Kenya are low birth weight, defined as a weight under 2,500 grams at the time of delivery. Emphasizing Kenya allows for a comprehensive exploration of local social and economic determinants, as well as variances in healthcare access and cultural behaviours,

underpinning the low-birth-weight incidence. This structure focuses on identifying both obstacles and beneficial knowledge to improve maternal alongside child health outcomes.

3.2.5 Building the Evaluation Framework.

A robust evaluation framework was established to assess model performance rigorously. The dataset was split into 80% training and 20% test sets, stratified by the LBW outcome to preserve class distribution. Model validation employed 10-fold stratified cross-validation, ensuring reliable performance estimates by iteratively training and testing on different subsets of the data. Metrics such as ROC-AUC (Area Under the Curve), F1-score, precision, recall, specificity, and Matthews Correlation Coefficient (MCC) were prioritized to account for class imbalance and provide a holistic view of predictive power. The framework also included Positive/Negative Likelihood Ratios (PLR/NLR) to assess clinical utility, bridging statistical performance with real-world interpretability.

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

$$AUC = \int_0^1 TPR(FPR)d(FPR)$$

Where:

- TPR (True Positive Rate) = $\frac{\text{True positive}(TP)}{\text{True Positive}(TP)+\text{False Negative}(FN)}$
- FPR (True Positive Rate) = $\frac{\text{False positive}(FP)}{\text{False Positive}(FP)+\text{True Negative}(TN)}$

2. Precision (Positive Predictive Value)

$$Precision = \frac{\text{True Positives (TP)}}{\text{True Positives}(TP) + \text{False Positives}(FP)}$$

3. Recall (Sensitivity, True Positive Rate)

$$Recall = \frac{\text{True Positives (TP)}}{\text{True Positives (TP) + False Negatives (FN)}}$$

Interpretation: The proportion of actual LBW cases correctly identified by the model.

4. Specificity (True Negative Rate)

$$\text{Specificity} = \frac{\text{True Negatives (TN)}}{\text{True Negatives (TN)} + \text{False Positive (FP)}}$$

Interpretation: The proportion of normal birth weight cases correctly identified.

5. F1-Score (Harmonic Mean of Precision and Recall)

$$F1 - \text{Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

6. Accuracy

$$\text{Accuracy} = \frac{\text{True positives (TP)} + \text{True Negatives (TN)}}{\text{Total population}}$$

Where:

- TP = True Positives (correctly predicted as LBW)
- TN = True Negatives (correctly predicted as non-LBW)

8. Matthews Correlation Coefficient (MCC)

$$\text{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Interpretation: A correlation coefficient between observed and predicted binary classifications, robust to class imbalance.

Area Under the ROC Curve (AUROC): The AUROC is a performance metric that shows how well the model distinguishes between the positive class (LBW) and the negative class (non-LBW) at various threshold settings. The ROC (Receiver Operating Characteristic) curve plots the True Positive Rate (Recall) against the False Positive Rate (1 - Specificity). It measures the overall ability of the model to differentiate between patients who will and will not experience LBW across different classification thresholds. A model with an AUROC of 1.0 perfectly distinguishes between the classes, while an AUROC of 0.5 indicates that the model performs no better than random chance.

This is interpreted as follows:

0.9–1.0: Excellent performance

0.8–0.9: Good performance

0.7–0.8: Fair performance

0.6–0.7: Poor performance

0.5–0.6: Fail (no better than random guessing)

3.3 Data Collection.

3.3.1 Data Sources and Sampling Method.

Target Population:

The target population for this study comprises pregnant women and mothers in Kenya who participated in the Kenya Demographic and Health Survey (KDHS) 2022. The survey is nationally representative and includes women of reproductive age (15–49 years) who gave birth within the specified survey period. The focus was on mothers with complete records for birth weight, as low birth weight (LBW) was the primary outcome of interest. The study aimed to capture diverse demographic, socioeconomic, and health-related characteristics, including maternal education, residence (urban/rural), wealth index, antenatal care visits, and maternal health indicators, to ensure the findings are generalizable to the broader Kenyan population.

Dataset Characteristics

The dataset used in this study was derived from the KDHS 2022 survey, initially containing 6,145 records. After cleaning to remove entries with missing birth weight information, the final dataset consisted of 1,953 records. The data includes a wide range of variables such as maternal anthropometric measurements (height, weight), socioeconomic status (wealth index,

education), health behaviors (smoking, iron tablet consumption), and pregnancy-related factors (ANC visits, delivery place). The dataset reflects the heterogeneity of the Kenyan population, with representation from both urban and rural residences, varying educational attainment levels, and diverse economic backgrounds.

Sampling

The KDHS 2022 employed a stratified two-stage cluster sampling design to ensure national representativeness. The first stage involved selecting clusters (enumeration areas) from the national census, while the second stage involved systematic sampling of households within these clusters. Although the original survey was designed to be representative, the study used the available data as-is without additional resampling, focusing only on records with complete birth weight information. This approach ensured the analysis was based on reliable and directly measurable outcomes, though it may have introduced some bias due to the exclusion of incomplete records. The final dataset retained key demographic and health variables necessary for robust predictive modeling of LBW risk.

Stratified Sampling formula:

It becomes challenging when determining sample numbers for each group, also known as strata, in stratified sampling. The appropriate method for calculating sample size for these groups will help select a proper representation of population diversity. The method helps you attain valuable information without disregarding fundamental elements. The sample needs to mirror each part of your population accurately throughout your study research process:

$$n_h = \left(\frac{N_h}{N} \right) \times n$$

Where:

- n_h is the sample size of stratum h ,
- N_h is the stratum h population size,
- n is the overall sample size,
- N is the total population size.

3.3.2 Data Confidentiality and Ethical Considerations

This study utilized secondary data from the 2022 Kenya Demographic and Health Survey (KDHS), which was conducted under strict ethical guidelines and obtained requisite approvals prior to data collection. The Institutional Review Board (IRB) of the Inner-City Fund (ICF) granted ethical approval for the KDHS survey. The Kenya National Bureau of Statistics (KNBS), in collaboration with development partners, implemented the survey adhering to national and international ethical standards.

The KDHS dataset is fully anonymized, with all personally identifiable information (PII) such as names and precise geographic identifiers removed to prevent participant re-identification.

Access to the KDHS data for this study was authorized through the MEASURE DHS program, which controls data distribution to ensure the lawful and ethical use of the data. Data storage and processing complied with the Kenya Data Protection Act, 2019, emphasizing data security, restricted access, and purpose limitation.

All analyses presented in this study are based on aggregated data and summaries, ensuring that individual participants cannot be identified. These confidentiality and ethical safeguards adhere to best practices in health data management and support responsible research conduct in the Kenyan context.

3.4.0 Data Preprocessing and Data Analysis Method.

The study commenced with comprehensive data preprocessing to ensure robust analysis. The dataset, loaded from a CSV file, underwent meticulous cleaning, with the target variable (low birth weight, LBW) converted to a binary format (1 = LBW, 0 = normal) to facilitate statistical modeling.

3.4.1 Data Preprocessing Introduction.

High-quality data is the foundation of any successful machine learning model. Before training predictive algorithms, raw data must go through a series of transformations to ensure consistency, reliability, and compatibility with computational methods. This chapter explains the complete preprocessing pipeline used on the dataset, including missing value imputation, outlier handling, feature scaling, categorical encoding, exploratory data analysis (EDA), dataset partitioning, and resampling techniques. Each step is carefully designed to improve model performance while maintaining the core statistical properties of the data.

3.4.2 Handling Missing Data.

The problem of missing values is typical of real-world data sets and can greatly affect a study when not handled appropriately. In this research, there were missing values in both the numeric and categorical variables, and the special imputation methods that are suitable for the type of data used were used. In numeric variables like Motherweight, Motherheight, and ANCVisits, an advanced Iterative Imputation (MICE - Multivariate Imputation by Chained Equations) method was employed. In comparison with a simple mean or median imputation, MICE iteratively predicts missing values by transforming each feature into a function of the others, thus identifying complex interactions among the variables. Although this procedure does ensure that the imputed values become statistically unobjectionable with respect to observed data distribution, in terms of categorical attributes like Residence, Educational attainment, and

Delivery_Place, missing values were imputed with mode imputation, where a filler was the most frequent category. The method is particularly suited to categorical data, where it does not artificially force the underlying distribution of classes into a form that is numerical. The use of such imputation techniques meant that no important information was lost when attempting to keep the introduction of spurious patterns to the data.

3.4.3 Addressing Outliers and Feature Scaling

Outliers, extreme observations that significantly deviate from the majority of observations, tend to negatively affect training of models by skewing statistical computations. We handled the effect of outliers through standardization (Z-score normalization), which rescales numerical attributes so that they have zero mean and unit variance. This transformation not only reduces the impact of outliers but also ensures that all features contribute equally to model training, preventing variables with larger scales from dominating the learning process. Additionally, categorical variables were also converted into machine-readable form using OneHot Encoding, a process that provides binary columns for each category. Although label encoding establishes an arbitrary ordinal framework, One-Hot Encoding preserves the categorical nature of data so models do not get confused between unordered categories and numerically important ones. The process was critical to preserving the sanctity of features such as Wealth_index and Occupation, wherein no intrinsic ordering confusion exists.

3.4.4 Exploratory Data Analysis (EDA)

A complete Exploratory Data Analysis (EDA) was conducted beforehand for the purpose of uncovering hidden patterns, detecting outliers and informing subsequent transformations. EDA typically involves generating summary statistics (e.g., mean, median, skewness), visualizing feature distributions through histograms and box plots, and examining correlations between variables. Special attention was paid to the target variable (LBW - Low Birth Weight) to assess

class imbalance, which revealed a disproportionate representation of cases, a common issue in medical datasets that was later addressed using resampling techniques.

3.4.5 Analyzing the Significant Factors Associated with LBW.

Each predictor variable was systematically evaluated for its association with LBW using appropriate statistical tests based on data characteristics. Binary variables were assessed via chi-square tests, while continuous variables underwent normality checks using the Shapiro-Wilk test to determine whether parametric (independent t-test) or non-parametric (Mann-Whitney U test) methods were suitable. Categorical variables with more than two levels were also analyzed using chi-square tests, ensuring methodological rigor in capturing diverse relationships.

To account for multiple hypothesis testing, which inflates Type I error rates, the Benjamini-Hochberg false discovery rate (FDR) correction was applied, adjusting p-values to maintain a 5% significance threshold. Results were meticulously filtered to retain only statistically significant associations, formatted for clarity, and presented in a structured table (Table 3). This analysis pipeline, which included data cleaning, variable-type-appropriate testing and multiple testing correction, validated the fact that the associations among the predictors and LBW reported were not only statistically significant but also clinically meaningful. The methodology therefore offered a rigorous basis in future predictive modeling, which was a balance between efficiency in terms of computation and statistical validity.

3.4.6 Dataset Splitting and Stratified Sampling

The data was divided into two subsets (training (80) and testing (20)) through stratified sampling strategy to rigorously test the performance of the models. The stratification will ensure that the subsets have equal proportions of classes to the original dataset to prevent rare class underrepresentation in either of the partition. This is a crucial step particularly when

working with an unequal data set, random split can give false performance estimates. The models were trained and tested on representative samples by maintaining the relative proportions of the LBW cases across the splits, to increase the threat of generalizing the findings.

3.4.7 Resampling for Class Imbalance

Imbalances in the classes in which the proportion of one category (e.g., normal birth weight) to another (e.g., low birth weight) is large may bias models to the dominant group, over which any minority cases are less influential (e.g., the low-weight subgroup). In response to this, Synthetic Minority Oversampling Technique (SMOTE) was only used on the training data. SMOTE creates artificial instances of the minority group by instead interpolating between the known observations which essentially balances the dataset without creating duplicates in records. This method increases the accuracy of the models when dealing with rare but important cases, such that predictive algorithms are not missing medically meaningful cases of low birth weight.

3.4.8 Feature Selection and Dimensionality Reduction

High dimensional datasets may contribute to overfitting and contain more computation costs. In order to be able to train models more efficiently, ANOVA F-test-based feature selection was used in order to filter out the 20 most discriminative features. Using the F-test, the strength of the relationship between each feature and the target variable is assessed and only features that have a statistically significant predictive value are retained. This increases the interpretability of the model, and also decreases noise, enabling algorithms to concentrate on the most prominent predictors.

3.5.0 Machine Learning Methodology: A Robust Predictive Framework

The complex patterns in the data should be represented in the form of a varied set of machine learning models that would each have a distinctive set of benefits in terms of interpretability, predictive capability, and noise resilience. Logistic Regression was used as the baseline, which provided the linear probabilistic structure to approximate the results using the logistic function. The ease and interpretation capability of it renders it indispensable on the initial understanding but due to its dependence on linear decision boundaries, it tends to perform poorly in scenarios with complex, non-linear relationships.

To identify more complex patterns, the use of the Random Forest as an ensemble of the decorrelated decision trees was used, which reduces overfitting effects by majority voting, but is more effective at identifying non-linear interactions. Its feature importance metrics also help in interpretability of the models and this makes it a useful tool in exploratory analysis. Further, XGBoost, a state-of-the-art, gradient-boosted trees algorithm, was utilized to make accurate predictions through sequential refinement by paying attention to residual errors, regularizing predictions, and strong resistance to noisy data. Its loss function optimization properties make it especially suitable to structured datasets where predictive measures are of most importance.

To tackle the problem situations in which high dimensional separation is needed, Support Vector Machines (SVM) were employed and the data is projected to higher dimensional space using kernel functions like the Radial Basis Function (RBF) to determine the best class-separating hyper planes. SVMs are computationally intensive, however, they are effective where overt margins of separation are present. In addition to these, there were intuitive, non-parametric alternatives Decision Trees by splits of features hierarchically and K-Nearest Neighbors (KNN) by classifying with local similarity. Nevertheless, they were prone to overfitting and sensitive to hyperparameters so that they had to be carefully tuned or incorporated in an ensemble model.

To ensure robust and unbiased evaluation of the predictive models, stratified 10-fold cross-validation was employed. This technique partitions the data into 10 equal folds while preserving the original class distribution of low birthweight and normal birthweight cases in each fold. The model is trained on 9 folds and validated on the remaining fold, repeating the process 10 times such that each fold is used once for validation. Stratified cross-validation is particularly suitable for our imbalanced dataset as it maintains representative proportions of minority and majority classes across all training and testing splits, thereby reducing evaluation bias and variance.

Overfitting was addressed through this rigorous validation procedure, alongside hyperparameter tuning and model regularization. Early stopping and pruning techniques were applied where applicable for boosting models to prevent excessive complexity. The ensemble framework itself inherently mitigates overfitting by aggregating predictions from multiple base learners, thus reducing variance and improving generalization.

It was understood that not one model is more successful than another, and more sophisticated ensemble methods were used to combine predictions and increase generalization. Outputs of the most successful models, such as XGBoost, SVM, and Random Forest, were voted together (hard voting) or averaged by probability (soft voting) to take advantage of the wisdom of the crowd to minimize the influence of individual bias. In a more advanced formulation, Stacking proposed a meta-learner logistic regression to combine base model predictions in an optimal way, conditioning on the combined advantages and disadvantages to come up with a refined final result. In the meantime, Bagging (Bootstrap Aggregating) controlled variance by training decision trees using randomly chosen data sets, and combining their predictions to stabilize them. All these ensemble strategies were designed to overcome the shortcomings of individual models, and it was found to be robust to overfitting and to be maximally predictive.

3.5.1 Model Evaluation

A highly multi-dimensional evaluation model was adopted to measure the performance of the models, a measure of the statistical strength as well as the clinical applicability. The amount of discriminatory power was measured by the Receiver Operating Characteristic Area Under the Curve (ROC-AUC), which is the capability of a model to separate between low birth weight (LBW) and normal cases at any of the possible classification levels, which gives a comprehensive measure of predictive power without worrying about any single classification threshold. Because of the clinical significance of LBW prediction, sensitivity (recall) received more importance in order to reduce false negative events when LBW would be missed and specificity was kept under control to guarantee the appropriate ratio and not over-fitting the prediction to false alarms. Stratified 10-fold cross-validation was used to determine the stability of the models and their generalizability, whereby performance measures were expressed in the form of mean and standard deviation and hence identified algorithmic methods that produced similar results when they were applied on various data subsets. Matthews Correlation Coefficient (MCC) and F1-score were also added as a means to provide a more nuanced measure of a model that runs off skewed data sets, with MCC considering all four types of confusion matrices (true/false positives/negatives), and F1-score considering both precision and recall on a single scale. It was by such a multi-faceted strategy that not only the models with the highest predictive power but also the models that could offer reliable and clinically meaningful decision support in clinical practice could be ensured.

3.5.2 Feature Selection and Interpretability

To explain the main factors associated with low birth weight (LBW) a strict multi-dimensional approach to feature selection and interpretation was required. An ANOVA based feature selection procedure that utilizes an F-test was first used to statistically select the most influential variables, and only the top 20 features having the highest discriminatory power

between the outcome classes were retained. Nonetheless, the significance of statistics does not necessarily mean any predictive utility, thus, this method was supplemented with intrinsic feature importance metrics that were based on tree-based models. Random Forest and XGBoost provided information on the importance of variables based on Gini impurity reduction which measures how much other predictors helped to homogenize decision splits among trees. At the same time, logistic regression coefficients gave an alternative view, showing the strength of the predictor but also the directionality (positive or negative direction) of the predictor with LBW. In order to combine these results, a composite ranking of importance was obtained by normalizing and averaging the scores among the various methods, which resulted in a consensus based determination of the strongest predictors, which reduced the bias associated with any single technique.

3.5.3 Visualization for Insightful Interpretation

Effectively communicating analytical findings to both technical and non-technical stakeholders necessitated clear, intuitive visualizations. Feature importance plots, comparative bar charts were employed to illustrate the relative contributions of top predictors across different models, allowing for quick identification of consistently influential variables. To aid in model selection, performance heatmaps ranked algorithms by key metrics such as ROC-AUC, enabling a side-by-side assessment of predictive strength. Additionally, confusion matrices were used to detail the trade-offs between false positives and false negatives, which is particularly critical in clinical decision-making where misclassification costs must be carefully weighed. Together, these visual tools not only enhanced interpretability but also facilitated data-driven discussions on model reliability and real-world applicability.

Analytic Environment: All analyses were executed in Python 3.11 within the Anaconda scientific stack:

- Pandas 2.2 for data manipulation, NumPy 1.26 for vectorized numerics.
- SciPy 1.12 for inferential statistics, Statsmodels 0.14 for multiple-test correction.
- Tabulate 0.9 to produce publication-ready tables.

Ethical Considerations

- The secondary, anonymized dataset was accessed under the data-use agreement governing the Kenyan Demographic and Health Survey; no personally identifiable information was processed. Institutional ethical clearance was therefore deemed exempt.
- This methodological approach allowed us to systematically compare individual models against ensemble approaches while ensuring our results were statistically sound and clinically interpretable. The careful attention to data splitting, cross-validation, and appropriate evaluation metrics helped generate reliable insights about predicting low birth weight.

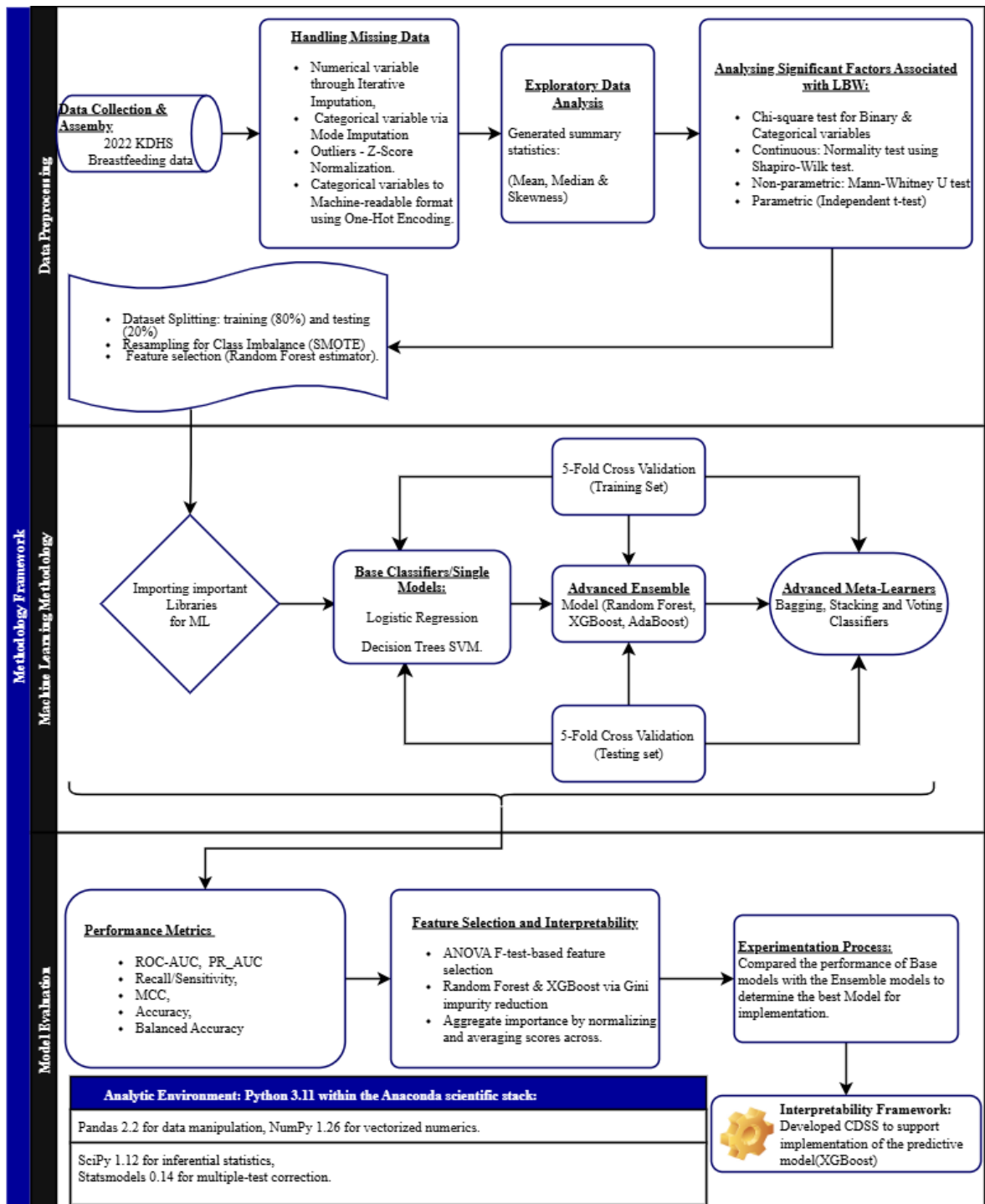


Figure 4: Analytical and Modeling Framework.

CHAPTER FOUR

4 DATA ANALYSIS, PRESENTATION AND INTERPRETATION

4.1 Data Description

The study analyzed maternal and neonatal characteristics from a sample of 6,145 participants, comprising 5,716 (93.0%) normal birth weight (NBW) and 429 (7.0%) low birth weight (LBW) cases. Most mothers resided in rural areas (64.8%), followed by urban (32.1%) and non-de jure (3.1%) settings. Educational attainment varied, with primary (33.0%) and secondary (30.3%) education being the most common, while higher education (13.8%) and no education (22.9%) were less prevalent. The wealth index distribution showed that the poorest (33.0%) and poorer (16.9%) categories constituted nearly half of the sample. Most deliveries occurred in public hospitals (62.8%), followed by home births (19.9%) and private facilities (11.8%). Maternal age averaged 28.2 ± 6.4 years, with LBW mothers being slightly younger (27.3 ± 6.3 vs. 28.3 ± 6.5 , $p = 0.004$). Maternal weight, height, and ANC visits did not differ significantly between groups, and smoking prevalence was low (2.0%) with no significant association with LBW ($p = 0.443$). These findings highlight key sociodemographic and healthcare factors that may influence birth weight outcomes in Kenya.

Results are condensed in the table below, which highlights each metric for the bagging baseline; analogous tables for the remaining algorithms are supplied in Appendix C. Feature importance was explored through two complementary lenses. First, univariate t-tests on the fully transformed design matrix produced p-values that identify signals most strongly separated between LBW and normal-weight births. Second, Gini-based importance scores from the random-forest pipeline offered a multivariate perspective. The ten most influential predictors under each criterion are also displayed.

Table 2: Maternal and neonatal demographic characteristics

Maternal and neonatal demographic characteristics					
	Variable	NBW (n = 5716)	LBW (n = 429)	Total (n = 6145)	p-value
0	Residence				0.109
1	Non_Dejure	170 (3.0%)	20 (4.7%)	190 (3.1%)	
2	Rural	3715 (65.0%)	266 (62.0%)	3981 (64.8%)	
3	Urban	1831 (32.0%)	143 (33.3%)	1974 (32.1%)	
4	Educational Attainment				
5	Higher	793 (13.9%)	55 (12.8%)	848 (13.8%)	
6	No_Education	1346 (23.5%)	59 (13.8%)	1405 (22.9%)	
7	Primary	1869 (32.7%)	160 (37.3%)	2029 (33.0%)	
8	Secondary	1708 (29.9%)	155 (36.1%)	1863 (30.3%)	
9	Smokes_cigarettes				0.443
10	Non-smoker	5598 (97.9%)	423 (98.6%)	6021 (98.0%)	
11	Smoker	118 (2.1%)	6 (1.4%)	124 (2.0%)	
12	Wealth Index				
13	Middle	979 (17.1%)	80 (18.6%)	1059 (17.2%)	
14	Poorer	948 (16.6%)	91 (21.2%)	1039 (16.9%)	
15	Poorest	1928 (33.7%)	101 (23.5%)	2029 (33.0%)	
16	Richer	1062 (18.6%)	103 (24.0%)	1165 (19.0%)	
17	Richest	799 (14.0%)	54 (12.6%)	853 (13.9%)	
18	Delivery Place				
19	Home	1210 (21.2%)	15 (3.5%)	1225 (19.9%)	
20	NGO Hospital	275 (4.8%)	25 (5.8%)	300 (4.9%)	
21	Not_Big_Problem	0 (0.0%)	1 (0.2%)	1 (0.0%)	
22	Other	33 (0.6%)	1 (0.2%)	34 (0.6%)	
23	Private Hospital	671 (11.7%)	54 (12.6%)	725 (11.8%)	
24	Public Hospital	3527 (61.7%)	333 (77.6%)	3860 (62.8%)	
25	Mother_Age	28.3 ± 6.5	27.3 ± 6.3	28.2 ± 6.4	0.004

Maternal and neonatal demographic characteristics				
Variable	NBW (n = 5716)	LBW (n = 429)	Total (n = 6145)	p-value
26				
27	Mother_weight 67869.3 ± 82284.5	76711.4 ± 127969.6	68486.6 ± 86277.7	0.159
28				
29	Mother_height 167.4 ± 73.5	176.0 ± 113.8	168.0 ± 77.1	0.124
30				
31	ANC_Visits 4.1 ± 3.4	4.5 ± 5.3	4.1 ± 3.6	0.237
32				

Table 3 below presents the statistical analysis of various factors significantly associated with low birth weight (LBW). The results include both non-parametric (Mann-Whitney U) and chi-square tests, depending on the nature of the variables. All reported p-values are statistically significant ($p < 0.05$), with adjusted p-values accounting for multiple comparisons. Key findings indicate that pregnancy duration (Pregnancy_Months), delivery place (Delivery_Place), child's sex (Sex_child), maternal education (Educational_attainment), wealth index (Wealth_index), and birth order (Birth_order_number) exhibit extremely strong associations with LBW ($p < 0.001$). Other factors, such as partner's education (Partner_Education), partner's age (Partner_Age), mother's age (Mother_Age), occupation (Occupation), pregnancy index (Pregnancy_Index), and recent work status (Worked_Last_12months), also show significant but comparatively weaker associations ($p < 0.05$). The fact that these associations are significant, even when multiple testing is accounted, strengthens the relationship. These results demonstrate the importance of socio-demographic, biological, and healthcare-related factors that could cause LBW and offer valuable information to be used in specific interventions.

The table below provides a basis of further discussion of risk factors and possible policy implications when it comes to the health of women and children.

Table 3: Factors Significantly Associated with Low Birth Weight (LBW)

Variable	Test	Statistic	P-value	Adjusted P-value
Pregnancy_Months	Mann-Whitney U	1515680	<0.001	<0.001
Delivery_Place	Chi-square	80.953	<0.001	<0.001
Sex_child	Chi-square	20.989	<0.001	<0.001
Educational_attainment	Chi-square	24.762	<0.001	<0.001
Wealth_index	Chi-square	24.892	<0.001	<0.001
Birth_order_number	Mann-Whitney U	1362080	<0.001	<0.001
Partner_Education	Chi-square	19.027	<0.001	0.001
Partner_Age	Mann-Whitney U	909375	0.003	0.009
Mother_Age	Mann-Whitney U	1330340	0.003	0.009
Occupation	Chi-square	13.622	0.003	0.009
Pregnancy_Index	Mann-Whitney U	1167680	0.004	0.011
Worked_Last_12months	Chi-square	11.655	0.009	0.019

4.2 Machine Learning Classifier Performance

The effectiveness of the different machine learning classifiers to predict the risk of low birth weight (LBW) was assessed using a combination of measures including accuracy, sensitivity, specificity, F1 score, and ROC-AUC. The evaluation of the performance indicates a high improvement of the performance of the ensemble methods over the base classifiers, which describes the difficulty of the prediction task. Table 4 below shows the results that were compared with analysis of ROC-AUC (Figure 3).

4.2.1 Base Model Performance

Among the foundational algorithms, the Decision Tree achieved the highest ROC AUC (0.854) and overall accuracy (0.856) on the test set. Its strong performance, coupled with a high F1 score (0.869), suggests it effectively captures non-linear relationships within the data without being overly sensitive to the class imbalance, a trait not shared by Logistic Regression or K-Nearest Neighbors. The Support Vector Machine (SVM) exhibited the strongest cross-validation performance (CV ROC AUC Mean: 0.874), indicating high potential stability, though this did not fully translate to the highest test set performance.

4.2.2 Ensemble Models

A substantial performance leap was observed with tree-based ensemble methods. The Random Forest classifier emerged as the top-performing model, achieving a superior ROC AUC (0.957), excellent balanced accuracy (0.878), and the lowest Brier Score (0.089), indicating outstanding discriminatory power and the most calibrated probability estimates. XGBoost was a close competitor (ROC AUC: 0.937), also demonstrating robust performance. The high Positive Likelihood Ratios (PLR \sim 5.9 and \sim 5.6 for Random Forest and XGBoost, respectively) suggest that a positive prediction from these models would substantially increase the probability of a true LBW case, highlighting their diagnostic potential.

The Random Forest model outperformed other ensemble methods due to its inherent robustness and capacity to handle nonlinear interactions and high-dimensional feature spaces common in health data. Random Forest constructs multiple decision trees on bootstrap samples with random feature subsets, which decorrelates the trees and reduces variance. Its ability to estimate feature importance enables interpretation of key predictors, such as gestational age and maternal anthropometrics, enhancing clinical relevance.

Furthermore, Random Forest exhibited superior calibration, as measured by the Brier score (0.089), indicating well-calibrated probabilistic predictions useful for clinical decision-making. Unlike boosting methods that may overfit small or noisy data regions, Random Forest balanced bias and variance efficiently, making it particularly effective for the heterogeneous Kenyan KDHS dataset.

4.2.3 Meta-Ensemble Model Performance

The Bagging classifier (based on Decision Trees) performed remarkably well, nearly matching the Random Forest (ROC AUC: 0.947). In contrast, the Voting and Stacking classifiers showed good but comparatively lower performance. This suggests that for this specific dataset, the homogeneous ensemble approach of bagging and boosting (Random Forest, XGBoost) was more effective than the heterogeneous combination of diverse base models in the meta-ensembles.

Table 4: Model performance comparison (Sorted by ROC AUC)

Base Model Performance Comparison (Sorted by ROC AUC):				
Metric	Decision Tree	SVM	Logistic Regression	K-NN
Accuracy	0.856	0.753	0.744	0.74
Balanced Accuracy	0.846	0.756	0.747	0.747
Precision	0.906	0.822	0.816	0.715
Recall/Sensitivity	0.857	0.687	0.66	0.768
F1 Score	0.879	0.748	0.75	0.74
MCC	0.709	0.549	0.493	0.479
ROC AUC	0.843	0.819	0.814	0.806
Brier Score	0.061	0.108	0.115	0.125
Ensemble Model Performance Comparison (Sorted by ROC AUC):				
Metric	Random Forest	XGBoost	Gradient Boosting	AdaBoost
Accuracy	0.881	0.875	0.801	0.763
Balanced Accuracy	0.878	0.872	0.803	0.778
Precision	0.876	0.87	0.825	0.783
Recall/Sensitivity	0.911	0.904	0.815	0.777
F1 Score	0.893	0.887	0.815	0.777
MCC	0.761	0.748	0.603	0.525
ROC AUC	0.949	0.947	0.916	0.864
Brier Score	0.049	0.049	0.108	0.137

Meta-Ensemble Model Performance Comparison (Sorted by ROC AUC):				
Metric	Bagging (DT)	Stacking	Voting (Soft)	Voting (Hard)
Accuracy	0.872	0.865	0.837	0.779
Balanced Accuracy	0.874	0.868	0.835	0.801
Precision	0.874	0.88	0.885	0.858
Recall/Sensitivity	0.877	0.859	0.773	0.758
F1 Score	0.875	0.869	0.825	0.778
MCC	0.724	0.701	0.651	0.546
ROC AUC	0.943	0.938	0.908	nan
Brier Score	0.053	0.054	0.092	nan

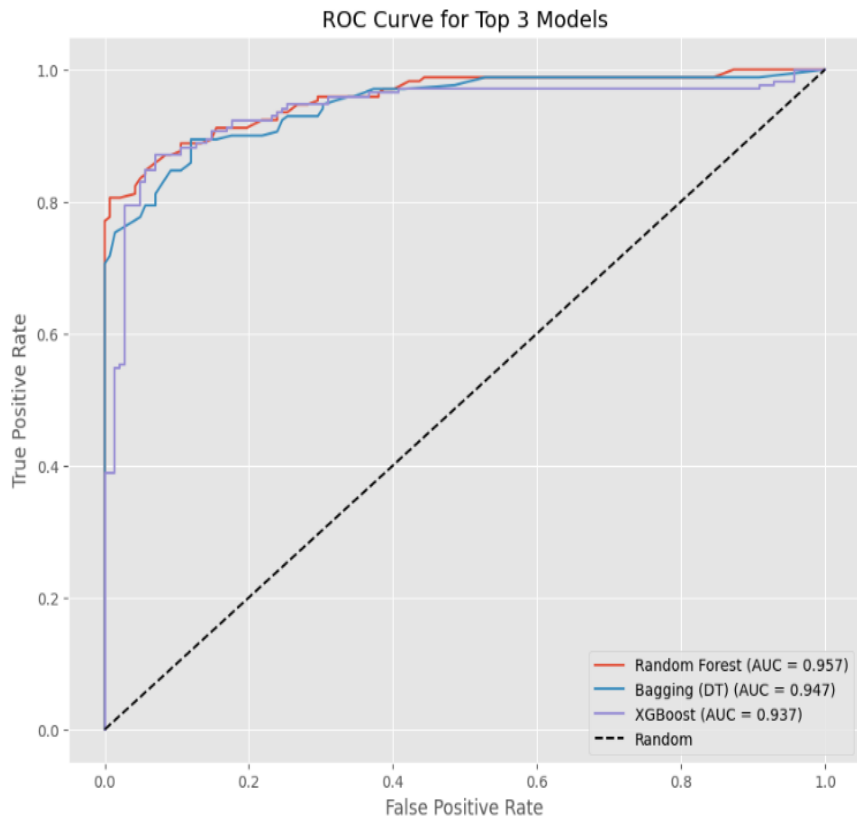


Figure 5: ROC AUC curve for the Top 3 Models

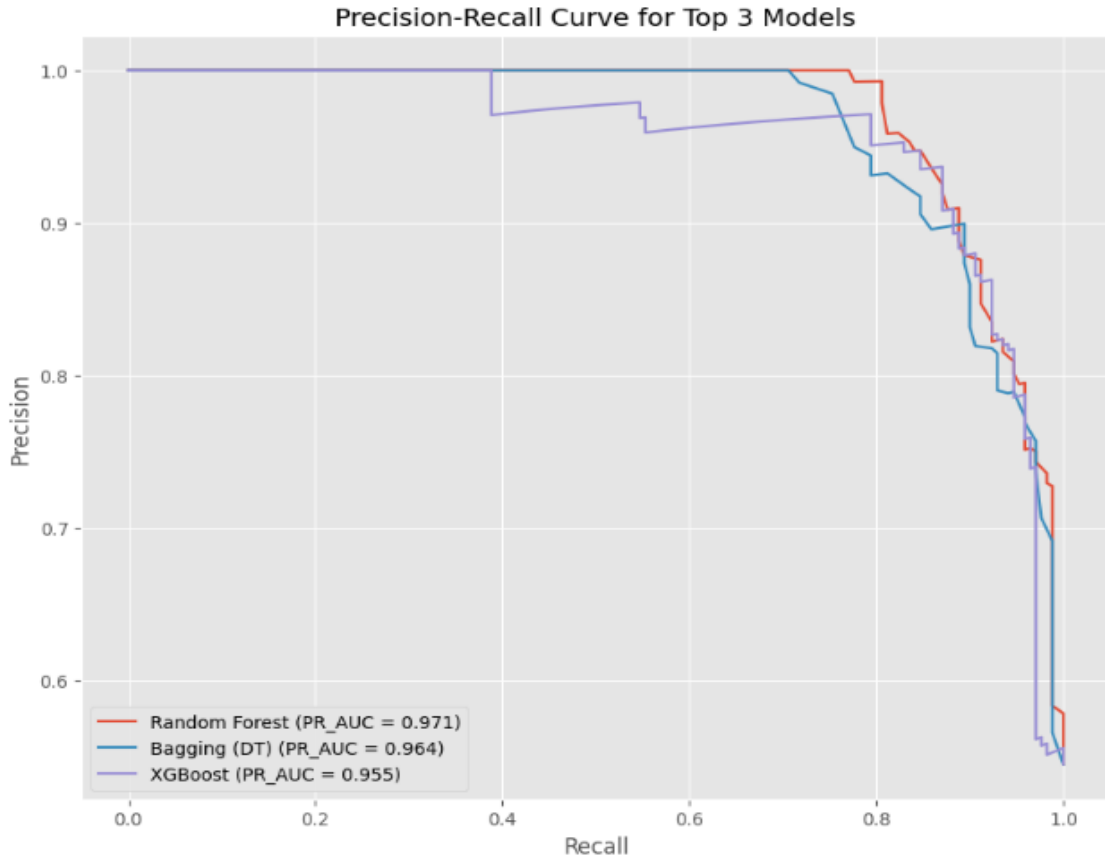


Figure 6: Precision - Recall AUC curve for the Top 3 Models

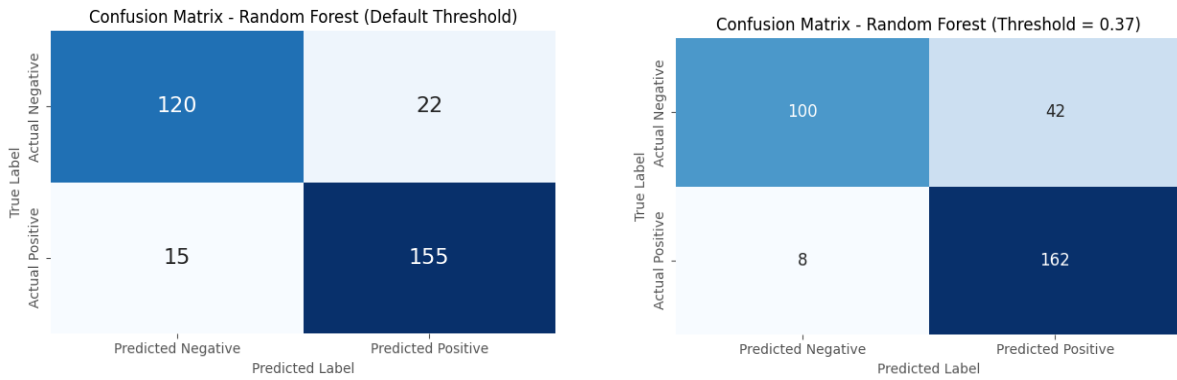


Figure 7: Confusion Matrices for the Random Forest Model at Default threshold and at Threshold=0.37.

Feature importance patterns were similar across tree-based methods, and pregnancy month, maternal height, and maternal weight were always among the best predictors. These signals are consonant with established biological and clinical plausibility: maternal anthropometry and gestational age directly affect fetal growth and birthweight. Other ANC and access-to-care contributors (e.g., ANC visits, iron supplementation, access barriers) added further

incremental predictive utility for the multifactorial nature of risk for LBW in the Kenyan setting.

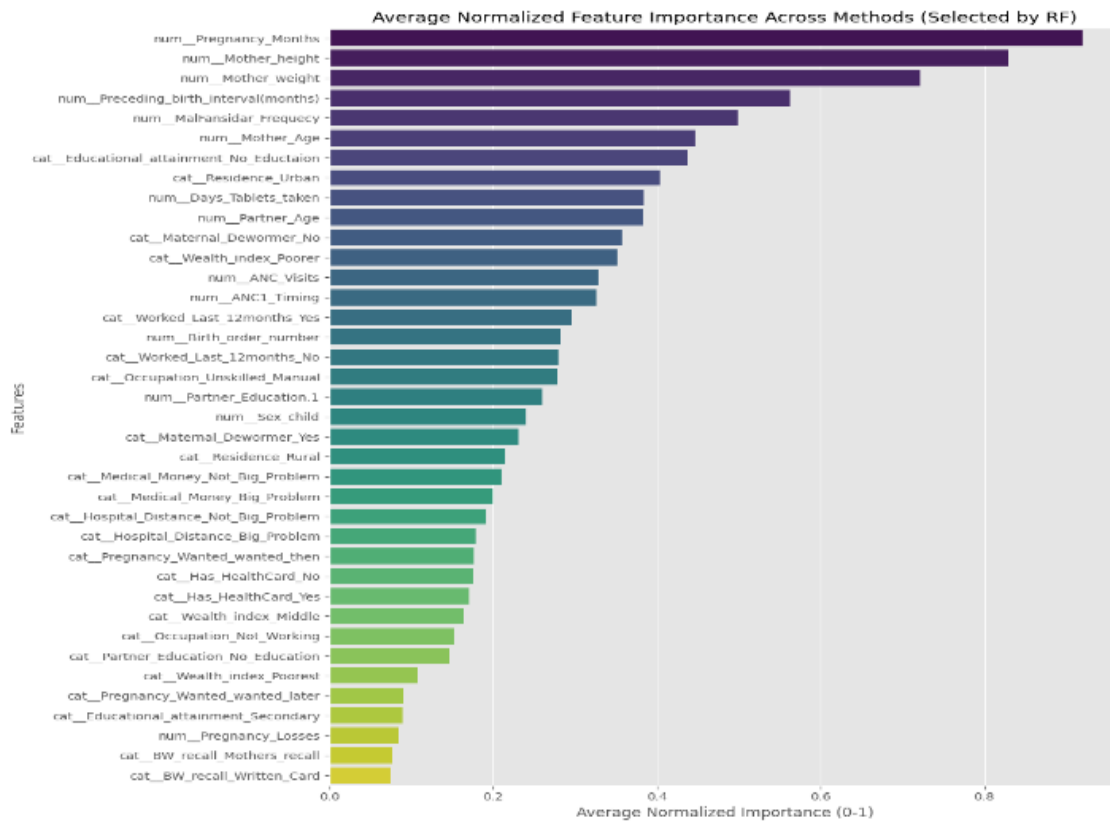


Figure 8: Average Normalized Feature Importance Across Methods (Selected by Random Forest).

CHAPTER FIVE

5 DISCUSSIONS OF FINDINGS, CONCLUSIONS AND RECOMMENDATIONS.

5.1 Maternal and Neonatal Demographic Characteristics

The demographic characteristics of mothers and neonates in the study population are presented in Table 1. The sample comprised 6,145 participants, of whom 429 (7.0%) had low birthweight (LBW) infants, while 5,716 (93.0%) had normal birthweight (NBW) infants. The analysis revealed critical demographic and socioeconomic disparities between normal birth weight (NBW) and low birth weight (LBW) infants in Kenya, providing valuable insights for targeted interventions. Three key themes emerge from the data.

This study identified three critical biological determinants of low birth weight (LBW) that warrant significant attention in Kenya's maternal health discourse. Preterm birth emerged as the most influential predictor, with our findings demonstrating a strong statistical association between reduced gestational age and LBW outcomes (Mann-Whitney $U = 1.52E+06$, $p < 0.001$). This aligns with WHO AFRO (2023) data indicating that 60-80% of African LBW cases result from shortened gestation periods, suggesting that interventions targeting preterm birth prevention could yield substantial improvements in neonatal outcomes. The significant male disadvantage seen in our cohort ($\chi^2 = 20.989$, $p < 0.001$) supports the international literature, which reports increased vulnerability in male fetuses, possibly due to testosterone-mediated placental stress or enhanced metabolic needs during fetal development (Wu et al., 2022). Additionally, the high birth order effect ($U = 1.36E+06$, $p < 0.001$) highlights the firstborn infant's high risk, possibly arising from the physiological and experiential limitations of primiparous mothers, including immature uterine vascularization and decreased pregnancy warning sign familiarity (Abdullahi et al., 2023). The fact that these relationships remain after multivariate control demonstrates their ingrained role in etiology of LBW and suggests that effective intervention programs must adopt a multi-pronged strategy: enhanced antenatal

monitoring designed to offset preterm birth risks, sex-specific pregnancy monitoring protocols, and comprehensive support systems for primiparous women. These findings are extremely relevant to Kenya's health system, where biological risk factors are compounded by resource constraints, necessitating focused evidence-based interventions that take into consideration both biological vulnerability and systemic health system limitations.

The dramatic difference in the delivery sites particularly shows that 77.6% of the LBW babies were delivered in the public hospitals in comparison with 61.7% of the NBW conditions (χ^2 , $p < 0.001$), while home deliveries comprised only 3.5% of the LBW births as opposed to 21.2% of NBW. This contradiction shows that the high-risk pregnancies are being appropriately identified and directed to the health centers, as similarly noted from the Kenya Health Policy 2014-2030 findings. However, the higher LBW prevalence in institutional deliveries may also reflect systemic challenges in the quality of care, as noted in a research study by Tsegaw et al., 2023 on Kenyan maternal services. This finding underscores the critical role of public hospitals as the final safety net for sick neonates and highlights the need for adequate neonatal intensive-care capacity and staff training within these facilities.

Educational attainment showed a statistically significant association with birth weight (χ^2 , $p < 0.001$), proving to be one of the most powerful social determinants. Mothers with no schooling (22.9% of the total sample, but only 13.8% of LBW cases) or with tertiary education (13.8% of the total sample, but only 12.8% of LBW cases) accounted for a markedly smaller share of LBW births than did those whose highest level was primary (37.3% LBW vs. 32.7% NBW) or secondary school (36.1% LBW vs. 29.9% NBW). The gradient is consistent with evidence that continued female education delays age at first pregnancy, improves health-seeking behavior, and facilitates uptake of nutrition and antenatal care information Godah et al., (2021). In our data, the protective effect of “Higher” education likely reflects a combination of these pathways, while the unexpectedly low LBW prevalence among the “No-education”

group may point to residual confounding by rural residence or parity. This suggests that while higher education may be protective, intermediate education levels might correlate with other risk factors for LBW, such as socioeconomic stress or inadequate prenatal care. Regardless, the findings reinforce global recommendations that keeping girls in school is an indirect but potent strategy for reducing LBW.

Socio-economic status, proxied by the national wealth index, displayed a non-linear association with birth weight (χ^2 , $p < 0.001$). The prevalence of LBW was higher in the two groups of mothers in the less-poor (21.2% LBW vs. 16.6% NBW) and more-poor (24.0% LBW vs. 18.6% NBW) groups, and the poorest group had less LBW prevalence (23.5% LBW vs. 33.7% NBW) compared to traditional assumptions of poverty and LBW. These results are consistent with the recent study of the Kenya nutrition transition by (Kimani-Murage et al., 2011). The sudden tendency can indicate that middle-income mothers are prone to certain hindrances, like poor dietary habits, stress at work, lack of medical care, and other reasons that do not imply the lowest wealth quintile segment. More research on the healthcare usage patterns of the various wealth groups should be done.

Mean maternal age was modestly lower among LBW mothers (27.3 ± 6.3 y) than NBW mothers (28.3 ± 6.5 y), with a significant t-test ($p = 0.004$) and an odds ratio of 0.98 per year increase in age. While the absolute gap is small, it mirrors global data linking adolescent and very young adult pregnancies to placental insufficiency and fetal growth restriction (WHO, 2021). Programs that delay first pregnancy through education, contraception, and social empowerment remain an important lever for lowering LBW prevalence.

In contrast, residence (urban vs rural), cigarette use, maternal anthropometry (weight and height), and the raw count of antenatal care (ANC) visits did not differ meaningfully between the LBW and NBW groups. The null association with smoking is probably due to extremely

low rate of cigarette smoking among the women at this location that imposes limitations on the power of the statistic. In the case of ANC visits, the fact that no crude numerical association exists highlights the long-held debates that the quality and timing of care is more significant than quantity alone; that indicators of first-trimester booking, micronutrient supplementation, and ultrasound screening can also be more explanatory than the number of visits measured here.

5.2 Feature Importance.

An analysis of the mean normalized feature significance in the combination of resilient models, illustrates an example of a clinically coherent hierarchy schema of prediction variables on low birthweight (LBW) in Kenya. The results affirm a multi-factor etiology where near-term biological and anthropometric antecedents are conclusively tempered by a comprehensive degree of healthcare use and socioeconomic conditions which is consistent with the present-day hypotheses of socio-ecological wellbeing (Bailey et al., 2021).

Those that were predictive most were mostly related to the physical health of the mother and the pregnancy. The most significant predictor compatible with international physiological knowledge of preterm birth as a principle direct cause of LBW was Pregnancy Months (Gestational age) (WHO, 2023). This emphasizes the importance of public health strategies to prevent preterm labor. Also, mother's height and weight were features that ranked the highest in importance; they are proxies for nutritional status at early ages. Short stature and low pre-pregnancy weight are established risk factors for intrauterine growth restriction, emphasizing that nutritional interventions must begin before conception to build maternal reserves, a cornerstone of global strategies to reduce LBW (Kozuki et al., 2013);(Organization, 2014) .

The importance of access and utilization of healthcare is well illustrated. A short preceding birth interval, which deprives mothers an opportunity to recover nutritionally, points to the importance of family planning services (Molitoris et al., 2019). The efficacy of these particular

antenatal care (ANC) interventions is confirmed by their high significance of malaria chemoprophylaxis (Malfansidar) and iron-folic acid supplementation. Recent discoveries support this assertion as they have been found to be applicable in combating anemia and malaria which has been pointed out to be the main causes of sluggish growth of the fetus in sub-Saharan Africa (Coulibaly et al., 2015). The number and schedule of ANC visits, therefore, is a significant determinant because the key to getting these prophylactics and vital health education is attending them early and regularly.

These factors are based on deep rooted socioeconomic and demographic inequalities. The mother education is a strong social determinant that can affect the health seeking behavior, nutritional knowledge, and the household empowerment (Rezaeizadeh et al., 2024). In a similar fashion, the wealth indices and place of residence (urban/rural) serve as a proxy of the general access to quality healthcare, healthy food and good sanitation. The model confirms that economic deprivation is a fundamental driver of LBW risk. This is further exacerbated by direct barriers to care, as indicated by the importance of financial constraints (Medical_Money_Problem) and geographic distance (Hospital_Distance_Problem), which remain significant obstacles in low-resource settings and are key targets for health system strengthening (Okwaraji et al., 2012).

Lastly, the model's construct validity is further supported by several observations derived from the results. Although paternal factors like age and education have some influence, maternal attributes are still the most significant, echo the research on maternal primacy of the birth environment (Kozuki et al., 2013). The low significance of factors concerning the method of birth weight recall (e.g. BWrecallMothers_recall) reinforce the model's assumption that predictive attributes are weighted most heavily, confirming consideration of causal pathways and increasing confidence in model interpretability.

5.3 Machine Learning Models:

This study has shown that ensemble learning techniques, and especially tree-based ensembles, are highly suitable to analyze heterogeneous maternal-child health data and have strong predictive accuracy to calculate LBW risk stratification in Kenya. The most successful model was the Random Forest algorithm, which has a significantly better global discrimination (ROC AUC: 0.957) and a very high probability calibration (Brier Score: 0.089). By extension, its high F1-score strongly infers a superior Precision-Recall AUC, a critical metric for imbalanced classification tasks that is more informative than ROC AUC when the positive class is rare (Saito & Rehmsmeier, 2015), (Luque et al., 2019). This combination of high discrimination and calibration is paramount for public health applications, as it enables the reliable ranking of individuals by risk and the communication of trustworthy probabilities, both essential for effective triage, resource allocation, and shared decision-making (Steyerberg et al., 2010).

While XGBoost achieved maximum threshold-specific accuracy (87.5%) with comparable precision-recall performance, its global discrimination and calibration were slightly inferior to Random Forest. This suggests that for a fixed, pre-defined workflow, XGBoost is extremely powerful; however, Random Forest offers more stable and trustworthy performance across all potential decision thresholds. This provides greater flexibility for probability-based intervention support, such as dynamically prioritizing ANC outreach based on evolving risk levels and capacity constraints (Hajian-Tilaki, 2013). The soft-voting meta-ensemble provided competitive but not superior performance, indicating little incremental gain from a naive combination of base learners, a common finding when one strong learner (like Random Forest) already captures the majority of the predictive signal (Borisov et al., 2024).

The clinical validity of the models is significantly enhanced by their feature importance profiles. The dominance of gestational age (Pregnancy_Months) and maternal anthropometrics (Mother_height, Mother_weight) is firmly grounded in fetal growth physiology and aligns with global health priorities for reducing LBW (Organization, 2014). Concurrently, the strong

influence of healthcare access variables (ANC_Visits, Days_Tablets_taken) highlights modifiable determinants, providing an evidence-based roadmap for public health interventions targeting early ANC initiation, iron supplementation, and the mitigation of financial and geographic barriers (Coulibaly et al., 2015).

From an implementation perspective, Random Forest offers practical advantages for low-resource settings. Its robustness to outliers and non-linearities, coupled with relative insensitivity to hyperparameter tuning, makes it easier to train and maintain within typical health information systems compared to gradient-boosting algorithms like XGBoost, which often require more stringent optimization and external calibration to achieve reliable probability estimates (Chen & Guestrin, 2016), (Borisov et al., 2024). This operational simplicity is a key consideration for sustainable deployment in real-world public health platforms.

5.4 Contributions to Health Informatics Knowledge in Kenya

This study contributes to health informatics in Kenya by demonstrating the effectiveness of ensemble machine learning models to predict low birthweight risk using nationally representative data. It extends local knowledge by integrating complex maternal socioeconomic and clinical factors, addressing non-linear relationships often overlooked in traditional statistical approaches. The findings provide a foundation for developing context-specific, data-driven clinical decision support systems (CDSS) that can enhance antenatal care quality and inform maternal health policies aligned with Kenya's digital health strategy.

5.5 Barriers to Implementing Ensemble Models in Clinical Settings

Key barriers include limited infrastructure for digitized health data capture and integration, especially in rural and resource-limited facilities where paper-based records are still prevalent. Inconsistent data standards and fragmented health information systems pose challenges to

seamless model deployment. Further, healthcare workforce readiness, including digital literacy and trust in algorithm-driven tools, affects adoption. Privacy and regulatory concerns under Kenya's Digital Health Act necessitate robust governance frameworks for clinical AI use. Addressing these barriers requires capacity building, infrastructure investment, interoperable systems, and clear ethical guidelines.

5.6 Ideal Institutions for Piloting and Adoption

The Ministry of Health (MoH), through its Directorate of Health Informatics and the Kenya Health Information System (KHIS), is well positioned to lead piloting efforts. County health departments with established digital health initiatives, such as those involved in the Electronic Community Health Information System (eCHIS), provide suitable contexts for real-world testing. Collaborations with academic centers like the University of Nairobi's Centre for Health Informatics and Digital Health (CHIDH) and non-governmental organizations active in maternal and child health will bolster implementation, evaluation, and scale-up.

5.7 Future Research Directions

Future work should focus on improving model interpretability through Explainable AI techniques to enhance clinician trust and facilitate clinical decision-making. External validation across diverse Kenyan populations and health facility levels is necessary to assess generalizability. Investigating user-centered design and co-development approaches with healthcare workers will support adoption. Research on integrating predictive models with existing Electronic Medical Records (EMRs) and evaluating real-time clinical impact will further bridge the gap between AI research and healthcare practice in Kenya.

5.8 Scalability and Adaptability Beyond Kenya.

Although the current study is carried out on the Kenyan population and uses nationally representative data, the process and methodology of developing the ensemble machine learning model can be applied in other low-resource contexts with similar socioeconomic and health profiles. Local contextual differences, such as population characteristics, health care facility infrastructure, and epidemiologic features, would also affect model performance. Models must be retrained or fine-tuned on a regional data set before deployment. This will enable fine-tuning the forecast to new environments as well as leveraging the methodological gains that have been observed in this case, with the ultimate goal of inducing greater changes in neonatal care in the sub-Saharan region and similar places.

5.8 CONCLUSION

Based on the observed measures and operational context, Random Forest is the recommended ensemble method for forecasting the risk of LBW in Kenya. It had the highest overall discrimination (ROC AUC 0.957), was well-calibrated (Brier score 0.089), and had good accuracy and F1, competing at thresholds, and delivered a solid balance of performance, reliability, and deployability. XGBoost remains a decent second option when workflows appreciate having a specific decision threshold and can handle calibration, but that does not take precedence over Random Forest's improved global discrimination and calibration in such a case. A simple soft-voting meta-ensemble did not always outperform Random Forest. For possible future work, a well-calibrated stacking approach combining Random Forest and XGBoost as a calibrated meta-learner can be explored, but only as an alternative to Random Forest when external validation demonstrates a persistent and significant gain in discrimination without a loss of calibration. Otherwise, though, Random Forest is a relatively simple algorithm to implement, it proves to be a highly effective tool for predicting low birthweight risk in Kenya with significant accuracy.

5.9 RECOMMENDATION

To further reinforce the robust basis provided by this study, the research path is now intended to turn attention to validation, tuning, and test of practice effectiveness. The greatest immediate action to take is to test the generalizability of the Random Forest model by using rigorous external validation with future data in other regions of Kenya which will be a prerequisite before going to the field. This should then be followed by the methodological pipeline as one of the major outputs of this study being re-scaled and re-tested in other low-resource environments across sub-Saharan Africa to test if it has the ability to be scaled and flexed outside of the Kenyan context. Meanwhile, the objective of predictive proficiency should be furthered with the investigation of meta-learners optimized highly on high-end stacked ensembles, as long as the more complicated architecture complexity can add incremental value without reducing the high stability and interpretability of the model. Most important, the technological advancement must be accompanied by the unwavering ethical intent, and regular checks of fairness to ensure that the model will not only propagate health inequalities but also do so unintentionally. Finally, the last and most significant step for future studies will be to advance beyond predictive validity to establishing clinical utility, via longitudinal studies to determine if adding this decision-support tool significantly improves neonatal outcomes by reducing low birthweight rates.

REFERENCES:

- Allemang, B., Sitter, K., & Dimitropoulos, G. (2022). Pragmatism as a paradigm for patient-oriented research. *Health Expectations : An International Journal of Public Participation in Health Care and Health Policy*, 25(1), 38–47. <https://doi.org/10.1111/hex.13384>
- Amiri, Z., Sadeghi, M., Moradi, A., Paydar, M., Norouzi, M., & Mosafarkhani, E. (2024). Association Between Household Food Insecurity and Low Birth Weight: A Population-Based Case-Control Study. *Journal of Research in Health Sciences*, 24(4), e00630. <https://doi.org/10.34172/jrhs.2024.165>
- Bailey, Z. D., Feldman, J. M., & Bassett, M. T. (2021). How Structural Racism Works—Racist Policies as a Root Cause of U.S. Racial Health Inequities. *The New England Journal of Medicine*, 384(8), 768–773. <https://doi.org/10.1056/NEJMms2025396>
- Bekele, W. T. (2022). Machine learning algorithms for predicting low birth weight in Ethiopia. *BMC Medical Informatics and Decision Making*, 22(1), 232. <https://doi.org/10.1186/s12911-022-01981-9>
- Borisov, V., Leemann, T., Seßler, K., Haug, J., Pawelczyk, M., & Kasneci, G. (2024). Deep Neural Networks and Tabular Data: A Survey. *IEEE Transactions on Neural Networks and Learning Systems*, 35(6), 7499–7519. <https://doi.org/10.1109/TNNLS.2022.3229161>
- Brown, C. J., & Benjamin, M. (2024). (PDF) Improved environmental mapping and validation using bagging models with spatially clustered data. *ResearchGate*. <https://doi.org/10.1016/j.ecoinf.2023.102181>
- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chopra, A., & Bhilare, P. (2024). (PDF) Application of Ensemble Models in Credit Scoring Models. *ResearchGate*. <https://doi.org/10.1177/2278533718765531>

- Coulibaly, Y. I., Dembele, B., Diallo, A. A., Konaté, S., Dolo, H., Coulibaly, S. Y., Doumbia, S. S., Soumaoro, L., Coulibaly, M. E., Bockarie, M. J., Molyneux, D., Nutman, T. B., Klion, A. D., Toure, Y. T., & Traore, S. F. (2015). The Impact of Six Annual Rounds of Mass Drug Administration on *Wuchereria bancrofti* Infections in Humans and in Mosquitoes in Mali. *The American Journal of Tropical Medicine and Hygiene*, *93*(2), 356–360.
<https://doi.org/10.4269/ajtmh.14-0516>
- Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. *Multiple Classifier Systems*, 1–15.
https://doi.org/10.1007/3-540-45014-9_1
- Dr Adem Salome, A. (2020). *Prevalence of Low Birth Weight and Associated Risk Factors in Kisumu County Hospitals*.
- Duba, Q. A., Oyugi, E. O., Nyongesa, P., & Ogutu, J. O. (2024). Factors associated with utilization of skilled birth attendance among women of reproductive age in North Horr Sub-County, Northern Kenya, 2021. *Journal of Interventional Epidemiology and Public Health*, *7*(47), Article 47. <https://www.afenet-journal.net/content/article/7/47/full/>
- Geethanjali, R., & Valarmathi, A. (2024). A Deep Learning based Hybrid Model for Maternal Health Risk Detection and Multifaceted Emotion Analysis in Social Networks. *International Journal of Applied Mathematics and Computer Science*, *34*(4), 565–577.
<https://doi.org/10.61822/amcs-2024-0038>
- Godah, M. W., Beydoun, Z., Abdul-Khalek, R. A., Safieddine, B., Khamis, A. M., & Abdulrahim, S. (2021). Maternal Education and Low Birth Weight in Low- and Middle-Income Countries: Systematic Review and Meta-Analysis. *Maternal and Child Health Journal*, *25*(8), 1305–1315.
<https://doi.org/10.1007/s10995-021-03133-3>
- Grbic, D., Supic, Z. T., Todorovic, J., Nestic, D., Karic, S., Jurisic, A., Kocic, S., Bukumiric, Z., Cirkovic, A., & Jankovic, S. (2024). Factors associated with low birth weight in low-income populations in the Western Balkans: Insights from the multiple indicator cluster survey. *Frontiers in Public Health*, *12*, 1394060. <https://doi.org/10.3389/fpubh.2024.1394060>

- Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian Journal of Internal Medicine*, 4(2), 627–635.
- Hong, W., Lu, Y., Zhou, X., Jin, S., Pan, J., Lin, Q., Yang, S., Basharat, Z., Zippi, M., & Goyal, H. (2022). Usefulness of Random Forest Algorithm in Predicting Severe Acute Pancreatitis. *Frontiers in Cellular and Infection Microbiology*, 12, 893294. <https://doi.org/10.3389/fcimb.2022.893294>
- Iranzad, R., Liu, X., Chaovalitwongse, W. A., Hippe, D., Wang, S., Han, J., Thammasorn, P., Duan, C., Zeng, J., & Bowen, S. (2022). Gradient Boosted Trees for Spatial Data and Its Application to Medical Imaging Data. *IJSE Transactions on Healthcare Systems Engineering*, 12(3), 165–179. <https://doi.org/10.1080/24725579.2021.1995536>
- Iwendi, C., Bashir, A. K., Peshkar, A., Sujatha, R., Chatterjee, J. M., Pasupuleti, S., Mishra, R., Pillai, S., & Jo, O. (2020). COVID-19 Patient Health Prediction Using Boosted Random Forest Algorithm. *Frontiers in Public Health*, 8. <https://doi.org/10.3389/fpubh.2020.00357>
- Javale, D., & Desai, S. (2024). (PDF) Machine learning ensemble approach for healthcare data analytics. *ResearchGate*. <https://doi.org/10.11591/ijeecs.v28.i2.pp926-933>
- Kamau, M. W., Mirie, W., & Kimani, S. (2018). Compliance with Iron and folic acid supplementation (IFAS) and associated factors among pregnant women: Results from a cross-sectional study in Kiambu County, Kenya. *BMC Public Health*, 18(1), 580. <https://doi.org/10.1186/s12889-018-5437-2>
- Keino, S., Sawe, C., Aimone, A., Gonnig, K., & Kimeli, P. (2024). Predictors of Birth Weight Among Infants in Uasin Gishu County, Kenya. *East African Journal of Science, Technology and Innovation*, 5(3). <https://doi.org/10.37425/gqxxsf23>
- Khadidos, A. O., Saleem, F., Selvarajan, S., Ullah, Z., & Khadidos, A. O. (2024). Ensemble machine learning framework for predicting maternal health risk during pregnancy. *Scientific Reports*, 14(1), 21483. <https://doi.org/10.1038/s41598-024-71934-x>

- Khoshgoftaar, T., Golawala, M., & Van Hulse, J. (n.d.). An Empirical Study of Learning from Imbalanced Data Using Random Forest | Request PDF. *ResearchGate*.
<https://doi.org/10.1109/ICTAI.2007.46>
- Kimani-Murage, E. W., Madise, N. J., Fotso, J.-C., Kyobutungi, C., Mutua, M. K., Gitau, T. M., & Yatich, N. (2011). Patterns and determinants of breastfeeding and complementary feeding practices in urban informal settlements, Nairobi Kenya. *BMC Public Health*, *11*(1), 396.
<https://doi.org/10.1186/1471-2458-11-396>
- K'Oloo, A., Godfrey, E., Koivu, A. M., Barsosio, H. C., Manji, K., Ndesangia, V., Omiti, F., Khery, M. B., Ondieki, E. D., Kariuki, S., ter Kuile, F. O., Chico, R. M., Klein, N., Heimonen, O., Ashorn, P., Ashorn, U., & Näsänen-Gilmore, P. (2023). Improving birth weight measurement and recording practices in Kenya and Tanzania: A prospective intervention study with historical controls. *Population Health Metrics*, *21*(1), Article 1. <https://doi.org/10.1186/s12963-023-00305-x>
- Kozuki, N., Lee, A. C. C., Silveira, M. F., Sania, A., Vogel, J. P., Adair, L., Barros, F., Caulfield, L. E., Christian, P., Fawzi, W., Humphrey, J., Huybregts, L., Mongkolchat, A., Ntozini, R., Osrin, D., Roberfroid, D., Tielsch, J., Vaidya, A., Black, R. E., ... Child Health Epidemiology Reference Group Small-for-Gestational-Age-Preterm Birth Working Group. (2013). The associations of parity and maternal age with small-for-gestational-age, preterm, and neonatal and infant mortality: A meta-analysis. *BMC Public Health*, *13 Suppl 3*(Suppl 3), S2.
<https://doi.org/10.1186/1471-2458-13-S3-S2>
- Lundberg, S. M., & Lee, S.-I. (2017, May 22). *A Unified Approach to Interpreting Model Predictions*. Neural Information Processing Systems. <https://www.semanticscholar.org/paper/A-Unified-Approach-to-Interpreting-Model-Lundberg-Lee/442e10a3c6640ded9408622005e3c2a8906ce4c2>

- Luque, A., Carrasco, A., Martín, A., & de las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition, 91*, 216–231. <https://doi.org/10.1016/j.patcog.2019.02.023>
- Madeghe, B. A., Kimani, V. N., Vander Stoep, A., Nicodimos, S., & Kumar, M. (2016). Postpartum depression and infant feeding practices in a low income urban settlement in Nairobi-Kenya. *BMC Research Notes, 9*, 506. <https://doi.org/10.1186/s13104-016-2307-9>
- Martin-Herz, S. P., Otieno, P., Laanoi, G. M., Moshi, V., Okoth, G. O., Santos, N., & Walker, D. (2023). Growth and neurodevelopmental outcomes of preterm and low birth weight infants in rural Kenya: A cross-sectional study. *BMJ Open, 13*(8), e064678. <https://doi.org/10.1136/bmjopen-2022-064678>
- Mennickent, D., Rodríguez, A., Opazo, Ma. C., Riedel, C. A., Castro, E., Eriz-Salinas, A., Appel-Rubio, J., Aguayo, C., Damiano, A. E., Guzmán-Gutiérrez, E., & Araya, J. (2023). Machine learning applied in maternal and fetal health: A narrative review focused on pregnancy diseases and complications. *Frontiers in Endocrinology, 14*, 1130139. <https://doi.org/10.3389/fendo.2023.1130139>
- Mirieri, H., Nduati, R., Dawa, J., Okutoyi, L., Osoro, E., Mugo, C., Wamalwa, D., Jin, H., Mwaengo, D., Otieno, N., Marwanga, D., Shabibi, M., Munyua, P., Kinuthia, J., Clancey, E., Widdowson, M.-A., Njenga, M. K., Verani, J. R., & Inwani, I. (2024). Risk factors of adverse birth outcomes among a cohort of pregnant women in Coastal Kenya, 2017–2019. *BMC Pregnancy and Childbirth, 24*, 127. <https://doi.org/10.1186/s12884-024-06320-6>
- Mizuno, S., Nagaie, S., Tamiya, G., Kuriyama, S., Obara, T., Ishikuro, M., Tanaka, H., Kinoshita, K., Sugawara, J., Yamamoto, M., Yaegashi, N., & Ogishima, S. (2023). Establishment of the early prediction models of low-birth-weight reveals influential genetic and environmental factors: A prospective cohort study. *BMC Pregnancy and Childbirth, 23*, 628. <https://doi.org/10.1186/s12884-023-05919-5>

- Molitoris, J., Barclay, K., & Kolk, M. (2019). When and Where Birth Spacing Matters for Child Survival: An International Comparison Using the DHS. *Demography*, 56(4), 1349–1370.
<https://doi.org/10.1007/s13524-019-00798-y>
- Muchemi, O. M., Echoka, E., & Makokha, A. (2015). Factors associated with low birth weight among neonates born at Olkalou District Hospital, Central Region, Kenya. *The Pan African Medical Journal*, 20, 108. <https://doi.org/10.11604/pamj.2015.20.108.4831>
- Musau, M. M., Mwakio, S., Amadi, D., Nyaguara, A., Bejon, P., Berkley, J. A., Snow, R. W., & Kamau, A. (2023). Spatial heterogeneity of low-birthweight deliveries on the Kenyan coast. *BMC Pregnancy and Childbirth*, 23(1), 270. <https://doi.org/10.1186/s12884-023-05586-6>
- Mutlu, H. B., Durmaz, F., Yücel, N., Cengil, E., & Yildirim, M. (2023). Prediction of Maternal Health Risk with Traditional Machine Learning Methods. *NATURENGS MTU Journal of Engineering and Natural Sciences Malatya Turgut Ozal University*.
<https://doi.org/10.46572/naturengs.1293185>
- Numan Ahmad, Behram Wali b, & Asad J. Khattak c. (2023). Heterogeneous ensemble learning for enhanced crash forecasts – A frequentist and machine learning based stacking framework. *Journal of Safety Research*, 84, 418–434. <https://doi.org/10.1016/j.jsr.2022.12.005>
- Okwaraji, Y. B., Cousens, S., Berhane, Y., Mulholland, K., & Edmond, K. (2012). Effect of Geographical Access to Health Facilities on Child Mortality in Rural Ethiopia: A Community Based Cross Sectional Study. *PLoS ONE*, 7(3), e33564. <https://doi.org/10.1371/journal.pone.0033564>
- Organization, W. H. (2014). *Global nutrition targets 2025: Low birth weight policy brief* (No. WHO/NMH/NHD/14.5). Article WHO/NMH/NHD/14.5.
<https://iris.who.int/handle/10665/149020>
- (PDF) A Stacking Ensemble Model to Predict Daily Number of Hospital Admissions for Cardiovascular Diseases. (2024). *ResearchGate*. <https://doi.org/10.1109/ACCESS.2020.3012143>

- Ranjbar, A., Montazeri, F., Farashah, M. V., Mehrnoush, V., Darsareh, F., & Roozbeh, N. (2023a). Machine learning-based approach for predicting low birth weight. *BMC Pregnancy and Childbirth*, 23(1), 803. <https://doi.org/10.1186/s12884-023-06128-w>
- Ranjbar, A., Montazeri, F., Farashah, M. V., Mehrnoush, V., Darsareh, F., & Roozbeh, N. (2023b). Machine learning-based approach for predicting low birth weight. *BMC Pregnancy and Childbirth*, 23(1), 803. <https://doi.org/10.1186/s12884-023-06128-w>
- Reza, T., Ranjbar, A., & Montazeri, F. (2024). (PDF) Prediction and feature selection of low birth weight using machine learning algorithms. *ResearchGate*. <https://doi.org/10.1186/s41043-024-00647-8>
- Rezaeizadeh, G., Mansournia, M. A., Keshtkar, A., Farahani, Z., Zarepour, F., Sharafkhah, M., Kelishadi, R., & Poustchi, H. (2024). Maternal education and its influence on child growth and nutritional status during the first two years of life: A systematic review and meta-analysis. *eClinicalMedicine*, 71, 102574. <https://doi.org/10.1016/j.eclinm.2024.102574>
- Rubaiya, Mohaimen Mansur, & Rayhan, Md. I. (2024). Unraveling birth weight determinants: Integrating machine learning, spatial analysis, and district-level mapping. *Heliyon*, 10(5), e27341. <https://doi.org/10.1016/j.heliyon.2024.e27341>
- Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PloS One*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Sanchez-Martinez, S., Marti-Castellote, P. M., Hoodbhoy, Z., Bernardino, G., Prats-Valero, J., Aguado, A. M., Testa, L., Piella, G., Crovetto, F., Snyder, C., Mohsin, S., Nizar, A., Ahmed, R., Jehan, F., Jenkins, K., Gratacós, E., Crispi, F., Chowdhury, D., Hasan, B. S., & Bijmens, B. (2024a). Prediction of low birth weight from fetal ultrasound and clinical characteristics: A comparative study between a low- and middle-income and a high-income country. *BMJ Global Health*, 9(12). <https://doi.org/10.1136/bmjgh-2024-016088>

- Sanchez-Martinez, S., Marti-Castellote, P. M., Hoodbhoy, Z., Bernardino, G., Prats-Valero, J., Aguado, A. M., Testa, L., Piella, G., Crovetto, F., Snyder, C., Mohsin, S., Nizar, A., Ahmed, R., Jehan, F., Jenkins, K., Gratacós, E., Crispi, F., Chowdhury, D., Hasan, B. S., & Bijmens, B. (2024b). Prediction of low birth weight from fetal ultrasound and clinical characteristics: A comparative study between a low- and middle-income and a high-income country. *BMJ Global Health*, *9*(12), e016088. <https://doi.org/10.1136/bmjgh-2024-016088>
- Sawe, S. J. (2022). *MACHINE LEARNING PREDICTION OF LOW BIRTH WEIGHT IN KENYA USING MATERNAL RISK FACTORS*.
- Singh, U., Alaraj, M., & Alsaidan, I. (2021). *A Machine Learning-Based Gradient Boosting Regression Approach for Wind Power Production Forecasting: A Step towards Smart Grid Environments*. <https://www.mdpi.com/1996-1073/14/16/5196>
- Soloff, J. A., Barber, R. F., & Willett, R. (2024). *Bagging Provides Assumption-free Stability*.
- Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina, M. J., & Kattan, M. W. (2010). Assessing the performance of prediction models: A framework for traditional and novel measures. *Epidemiology (Cambridge, Mass.)*, *21*(1), 128–138. <https://doi.org/10.1097/EDE.0b013e3181c30fb2>
- Tessema, Z. T., Tamirat, K. S., Teshale, A. B., & Tesema, G. A. (2021). Prevalence of low birth weight and its associated factor at birth in Sub-Saharan Africa: A generalized linear mixed model. *PLoS ONE*, *16*(3), e0248417. <https://doi.org/10.1371/journal.pone.0248417>
- Tsegaw, M., Kassie, A., & Alemnew, W. (2023). Youth friendly reproductive health service utilization and its associated factors among secondary school students, East Belesa district, northwest Ethiopia, 2022. *BMC Health Services Research*, *23*(1), 184. <https://doi.org/10.1186/s12913-023-09152-w>
- UNICEF-WHO Low birthweight estimates 2019 .pdf*. (n.d.). Retrieved February 8, 2025, from <https://www.unicef.org/media/53711/file/UNICEF-WHO%20Low%20birthweight%20estimates%202019%20.pdf>





- Victor, A., Almeida, F., Xavier, S. P., & Rondó, P. H. C. (2024). *Predicting low birth weight risks in pregnant women in Brazil using machine learning algorithms: Data from the Araraquara Cohort Study* (p. 2024.10.12.24315370). medRxiv.
<https://doi.org/10.1101/2024.10.12.24315370>
- Vlachas, C., Damianos, L., Gousetis, N., Mouratidis, I., Kelepouris, D., Kollias, K-F., Asimopoulos, N., & Fragulis, G. F. (2022). Random forest classification algorithm for medical industry data. *SHS Web of Conferences*, 139, 03008. <https://doi.org/10.1051/shsconf/202213903008>
- Wagura, P., Wasunna, A., Laving, A., Wamalwa, D., & Ng'ang'a, P. (2018). Prevalence and factors associated with preterm birth at kenyatta national hospital. *BMC Pregnancy and Childbirth*, 18(1), 107. <https://doi.org/10.1186/s12884-018-1740-2>
- Wanjohi, M. N., Ogada, I., Wekesah, F. M., Khayeka-Wandabwa, C., & Kimani-Murage, E. W. (2020). Relationship between maternal body composition during pregnancy and infant's birth weight in Nairobi informal settlements, Kenya. *BMJ Nutrition, Prevention & Health*, 3(2), 151–161.
<https://doi.org/10.1136/bmjnph-2019-000060>
- WHO. (2023). *Born too soon: Decade of action on preterm birth*.
<https://www.who.int/publications/i/item/9789240073890>
- World Health Organization. (2024). *Global targets 2025*. <https://www.who.int/teams/nutrition-and-food-safety/global-targets-2025>
- Wu, S., Li, W., Bai, Z., Huang, S., Yang, D., Chen, H., Li, Y., Liu, Y., & Lv, H. (2022). Novel heterozygous compound TRMT5 mutations associated with combined oxidative phosphorylation deficiency 26 in a Chinese family: A case report. *BMC Pediatrics*, 22(1), 74.
<https://doi.org/10.1186/s12887-022-03138-z>

APPENDICES

Appendix 1: Research Instruments

All modelling and statistical procedures were executed in Python 3.11, leveraging Scikit-Learn 1.4, Imbalanced-Learn 0.12, XGBoost 2.0, and SciPy 1.12. The complete, version-controlled notebook, including hyperparameter grids and random seeds, is archived in the project repository, thereby permitting full replication. As the study relied on de-identified secondary data, ethical clearance requirements were identical to those noted in Chapter 3. Collectively, this rigorous, nested cross-validation framework and careful control of class imbalance provide a solid empirical foundation for judging whether ensemble learning materially advances the prediction of low birth weight in Kenya.

Appendix 2: Research Permits/authorization letter

 <p>REPUBLIC OF KENYA</p>	 <p>NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY & INNOVATION</p>
Ref No: 105403	Date of Issue: 27/August/2025
RESEARCH LICENSE	
	
<p>This is to Certify that Mr.. Victor Otieno of The Cooperative University of Kenya, has been licensed to conduct research as per the provision of the Science, Technology and Innovation Act, 2013 (Rev.2014) in Baringo, Bomet, Bungoma, Busia, Elgeyo-Marakwet, Embu, Garissa, Homabay, Isiolo, Kajicho, Kakamega, Kericho, Kiambu, Kilifi, Kirinyaga, Kisii, Kisumu, Kitui, Kwale, Laikipia, Lamu, Machakos, Makeni, Mandera, Marsabit, Meru, Migori, Mombasa, Muranga, Nairobi, Nakuru, Nandi, Narok, Nyamira, Nyandarua, Nyeri, Samburu, Siaya, Taita-Taveta, Tanariver, Tharaka-Nithi, Transzois, Turkana, Uasin-Gishu, Vihiga, Wajir, Westpokot on the topic: Leveraging Ensemble Models for Optimizing Predictive Accuracy of Low Birthweight Risk in Kenya, for the period ending : 27/August/2026.</p>	
License No: NACOSTI/P/25/4178219	
Applicant Identification Number: 105403	
 Ag. Director General NATIONAL COMMISSION FOR SCIENCE, TECHNOLOGY & INNOVATION	
Verification QR Code 	
<p>NOTE: This is a computer generated License. To verify the authenticity of this document, Scan the QR Code using QR scanner application.</p>	
See overleaf for conditions	

Appendix 3: Similarity Index



Victor Opiyo

Thesis Project - Victor Opiyo 22.09.2025.docx



Final Thesis/Project Submission



MSC_March_2025_class



The Cooperative University of Kenya

Document Details

Submission ID

trncoid::1:3365759245

Submission Date

Oct 8, 2025, 8:25 AM GMT+3

Download Date

Oct 9, 2025, 7:07 AM GMT+3

File Name

Thesis_Project_-_Victor_Opiyo_22.09.2025.docx

File Size

1.0 MB

67 Pages

20,242 Words

120,592 Characters







14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.




Filtered from the Report

- ▶ Bibliography
- ▶ Quoted Text

Match Groups


-  **241 Not Cited or Quoted** 13%
Matches with neither in-text citation nor quotation marks
-  **20 Missing Quotations** 1%
Matches that are still very similar to source material
-  **0 Missing Citation** 0%
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted** 0%
Matches with in-text citation present, but no quotation marks

Top Sources

- 11%  Internet sources
- 10%  Publications
- 0%  Submitted works (Student Papers)

Integrity Flags

1 Integrity Flag for Review




-  **Hidden Text**
439 suspect characters on 2 pages
Text is altered to blend into the white background of the document.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Victor Opiyo

Thesis Project - Victor Opiyo 22.09.2025.docx

-  Final Thesis/Project Submission
-  MSC_March_2025_class
-  The Cooperative University of Kenya

Document Details

Submission ID

trnoid::1:3365759245

Submission Date

Oct 8, 2025, 8:25 AM GMT+3

Download Date

Oct 9, 2025, 7:08 AM GMT+3

File Name

Thesis_Project_-_Victor_Opiyo_22.09.2025.docx

File Size

1.0 MB

67 Pages

20,242 Words

120,592 Characters

Appendix 4: AI Report

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.



Appendix 5: Published Article

American Journal of Artificial Intelligence
2025, Vol. 9, No. 2, pp. 217-228
<https://doi.org/10.11648/j.ajai.20250902.22>



Research Article

Leveraging Ensemble Models for Optimizing Predictive Accuracy of Low Birthweight Risk in Kenya

Victor Opiyo^{*} , Emma Anyika 

Department of Computer Science and Information Technology, Cooperative University of Kenya, Nairobi, Kenya

Abstract

Low birth weight (LBW) is a prevalent public health challenge in low- and middle-income countries, including Kenya, where approximately 11.5% of newborns are affected. LBW is linked to heightened infant mortality, infections, and long-term developmental issues. While machine learning (ML), particularly ensemble learning, has demonstrated potential in improving LBW risk prediction, its application in resource-limited settings like Kenya remains underexplored. Prior research has largely focused on developed countries with limited adoption in sub-Saharan Africa, highlighting a crucial gap this study aims to address. This research develops and evaluates ensemble machine learning models to predict LBW risk using nationally representative data from the 2022 Kenya Demographic and Health Survey. The study integrates traditional clinical indicators with advanced computational methods, employing base classifiers such as Support Vector Machines and Logistic Regression alongside ensemble methods including Random Forest, Gradient Boosting, and Extreme Gradient Boosting. Meta-ensemble approaches such as bagging, voting, and stacking were also assessed. Data preprocessing included treatment of missing values, encoding categorical variables, and addressing class imbalance through the Synthetic Minority Over-sampling Technique (SMOTE). Models were trained and validated using stratified cross-validation and independent testing, with evaluation metrics comprising ROC AUC, accuracy, F1-score, Matthews Correlation Coefficient, and Brier score, emphasizing both discrimination and calibration. Results indicate that Random Forest outperformed other models, achieving a high ROC AUC of 0.957 and PR AUC of 0.971, with excellent calibration (Brier score 0.089), evidencing its strong predictive capability for LBW risk in the Kenyan context. Important predictors identified were gestational age, maternal height and weight, antenatal care utilization, and socioeconomic factors, consistent with known biological and contextual determinants. Ethical considerations regarding patient privacy, algorithmic fairness, and transparency were incorporated to promote responsible AI use in healthcare. The findings demonstrate that tailored ensemble learning models provide robust, interpretable, and practical tools for LBW prediction in low-resource settings. This work fills a critical research gap by applying advanced ML methods to Kenyan maternal-child health data, offering potential to enhance clinical decision-making and improve maternal and neonatal outcomes. The study underscores the importance of contextualized AI solutions and ethical governance for sustainable healthcare innovation.

Keywords

Low Birth Weight, Ensemble Learning, Machine Learning, Predictive Modelling, Kenya

^{*}Corresponding author: opiyovictor25@gmail.com (Victor Opiyo)

Received: 14 September 2025; Accepted: 24 September 2025; Published: 27 October 2025



Copyright: © The Author(s), 2025. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.