IJO -INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND ENGINEERING
(ISSN: 2814-1881) Mojirade.A. AWODUN *
https://ijojournals.com/ Volume 09 || Issue 03 || March, 2026 ||
"Advancing Maternal and Child Health in Nigeria through Support Vector Machines"

Advancing Maternal and Child Health in Nigeria through Support Vector Machines

[1]Treasure.O. ADEFEHINTI, [2]Ayorinde.O. IDOWU

[3]Mojirade.A. AWODUN

[1,2,3]Department of Computing and Information Science, Bamidele Olumilua University of Education, Science and Technology, Ikere-Ekiti, Ekiti State, Nigeria

Abstract

The alarming rate of maternal and neonates' morbidity and mortality rates in the country prompted a genuine commitment to enhance the health and well-being of mothers and children. This research aims to address the critical health challenges that causes maternal and neonates mortality using SVM to improve health outcomes, optimize healthcare efficiency and address existing disparity. SVM, a supervised machine learning algorithm, is employed to predict risks during pregnancy and childbirth by analyzing health indicators such as blood pressure, age, and heart rate so as to develop a predictive model to identify high-risk cases early and support healthcare professionals in implementing timely interventions.

Keywords: Maternal and Child Health, Machine Learning, Support Vector Machines, Predictive Analytics, Healthcare Optimization, Nigeria

Introduction

Maternal and neonate mortality rates in Nigeria are among the highest in the world especially in rural communities. Most rural areas face challenges in access quality health care service due to inequality in distribution of medical facilities and manpower leads to an increase rate in maternal and child mortality. Maternal and child health care in Nigeria is severely underfunded, with only 7% of the annual budget allocated to the health sector since the early 1990s (Adeyanju, *et al.*, (2017). This chronic under-funding impacts the quality and availability of maternal healthcare services, contributing to high maternal mortality rates. With limited resources, health facilities often struggle to provide essential services, which leads to disparities in care across different regions.

Monitoring and evaluation systems for maternal and child health care in Nigeria is critically underdeveloped. Effective health information systems are necessary for tracking health outcomes and evaluating the quality of delivered. Without robust monitoring frameworks, it becomes challenging to identify areas needing improvement or to implement effective interventions that could enhance maternal and child health care delivery (World Health Organization 2024).

A significant challenge faced in maternal and child health care decision making is the low autonomy women many women experience. Approximately only 40% of women have the ability to experience independent decision regarding their health (Obasohan, *et al.*,2019). A Support Vector Machine (SVM) is a supervised machine learning (ML) method capable of learning from data and making decisions. SVM is one of the popular supervised learning algorithm which is used for classification and regression problems. The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n-dimensional space into classes so that we can easily put the new data point in the correct category in the future. This best decision boundary is called a hyperplane.

Support Vector Machines (SVM) is pivotal tool in addressing various challenges within maternal and child health care, especially regarding health risk prediction, monitoring systems, and decision-making processes. By providing accurate classifications and predictive analytic, SVM enhance the ability of healthcare providers to identify high-risk pregnancies, improve monitoring of maternal health, and contribute to informed decision-making. Predicting risk levels for pregnant women can significantly minimize the likelihood of adverse outcome of pregnancies and child birth by analyzing numerous factors such as blood pressure and glucose levels, SVM can accurately classify the risk associated with pregnancies enabling health workers to identify high- risk cases early (Aman, *et al.*, 2024).

Women and children face unique health challenges that require specialized attention. Maternal and child health issues are often multifaceted, involving complex interactions among biological, social, and economic factors. SVM can analyze large datasets to identify patterns and risk factors associated with adverse health outcomes, enabling healthcare providers to develop targeted interventions.

In the past, clinical decision making are depended significantly on healthcare providers and traditional models which often lack the predictive accuracy required for effective maternal and child health interventions. Conventional approach like logistic regression have struggled to account for the complexities and interactions in health data leading to insufficiently accurate predictions of maternal risks and outcomes. Furthermore, past models have demonstrated issues with class imbalance, wherein the majority of cases are healthy, resulting in a failure to accurately identify and respond to critical conditions such as low birth weight. (Malndu, *et al.*, 2023) In recent years, advances in Machine Learning has created new opportunities for developing reliable models for health care. Support Vector Machine has demonstrated remarkable performance as a leading machine learning tool for maternal and child health care, outperforming several other algorithms, their ability to handle complex datasets, provide high accuracy and adapt to unique challenges health care data makes SVM valuable in improving health care outcomes for mothers and children. The robustness of SVM not only enhances predictive analytic but also supports healthcare professionals in making informed clinical decisions.

Studies have shown that SVM consistently achieve higher F1 scores and accuracy levels in maternal and child health applications than alternatives such as Decision Trees (DT) or Random Forests (RF). This superior performance is particularly beneficial in critical areas such as predicting maternal complications, assessing fetal health risks, and determining appropriate care interventions. The integration of SVM into clinical decision-making processes aids health workers in identifying at-risk populations, allocating resources effectively, and offering timely interventions that are crucial for improving outcomes. (Halomoan, *et al.*, 2024) To demonstrate the benefits of Support Vector Machines (SVM) in this field, this research is to develop an SVM-based predictive and decision making model for maternal and child health complications and evaluate its performance against other machine learning techniques.

Related Works

Support Vector Machines (SVM) are a widely utilized machine learning technique in healthcare due to their effectiveness in classification and predictive tasks.

SVM plays a crucial role in disease diagnosis, particularly in the classification of complex medical data. For instance, SVM have been successfully implemented in the diagnosis of conditions such as pneumonia, where different kernel functions have been analyzed for their effectiveness in classifying lung sounds. Studies indicate that the Radial Basis Function kernel (RBF) outperforms others, achieving an accuracy of 97.8% in diagnosing pneumonia, which is pivotal given the traditional reliance on less accessible diagnostic methods like chest X-ray. (Amose, *et al*, 2022)

Furthermore, another significant application of SVM is in predicting chronic kidney disease (CKD) risk. A cloud-based healthcare service paradigm employing boosted SVM has been developed to predict CKD, enhancing prediction accuracy while addressing resource maximization in cloud environments. (Aswini, *et al*,. 2021)

This exemplifies how SVM can facilitate timely interventions in chronic diseases, potentially improving patient outcomes through early detection. SVM also prove beneficial in scenarios with imbalanced datasets, which is common in healthcare. In predicting apnea of prematurity in neonates, SVM were combined with techniques like the Synthesized Minority Oversampling Technique (SMOTE) to balance data classes, achieving satisfactory prediction accuracy. This application highlights SVM adaptability in preprocessing data for better model performance in critical areas such as neonatal care.

Additionally, Advancements in genomics have also seen the application of SVM in stratifying patients based on specific biomarkers, essential for targeted therapies in oncology. SVM are well-adapted to handle high-dimensional genomics data, assisting in identifying patient subgroups likely to benefit from particular treatments. Their capability to effectively classify linear and non-linear data makes them instrumental in generating insights from complex datasets. (Lanza, *et al.* 2020)

SVM are recognized for their robust performance in risk prediction tasks across various domains, including healthcare. In cardiovascular risk prediction, for instance, SVM achieved an accuracy of 75.01% in a study, indicating their effectiveness in handling complex, high-dimensional health data. (Yarasuri, *et al.* 2022)

In contrast, other models like logistic regression and decision trees often yield varying results. Logistic regression can achieve around 81.00% accuracy, while decision trees provide slightly better performance at 82.30%. However, advanced ensemble methods, such as XGBoost and LightGBM, significantly outperform SVM, achieving peak accuracies of 92.72% and 90.60% respectively. This suggests that while SVM are capable, they may not always rank at the top compared to more contemporary ensemble methods in risk prediction scenarios.

SVM demonstrate exceptional efficacy in handling small datasets, primarily due to their reliance on a small subset of data points known as support vectors. This feature mitigates the risk of over fitting, which often plagues models like decision trees that may learn noise from the training data. For example, in applications involving medical diagnoses where data is often limited, SVM can still provide reliable predictions, with studies indicating that they maintain robust performance even with small sample sizes. Their capacity to efficiently navigate hyperplane in high-dimensional space allows SVM to generalize well from limited data, unlike models that require large datasets for effective training.

In light of this, algorithms such as K-Nearest Neighbors (KNN) may struggle with small datasets, as their performance heavily relies on the local structure of the data and may lead to over fitting or misclassification. Random Forests, while generally effective, can introduce unnecessary complexity with small datasets due to the numerous decision trees they build, which may not always contribute to improved performance.

Ensemble methods like Random Forests and Gradient Boosting have shown improved performance over SVM, particularly in scenarios with abundant data. Random Forests, for instance, achieve higher accuracies (90.76% in one study), leveraging multiple decision trees to enhance predictive accuracy. However, SVM still retain an advantage in simplicity and interpretability, which can be crucial in clinical settings where model transparency is essential.

Additionally, on small datasets, the advantage of SVM in avoiding over fitting can make them more reliable than ensemble methods that may require careful tuning of hyper parameters to prevent capturing noise rather than signal. This contrasts with methods like LightGBM and XGBoost, which may excel in ensemble strategies but necessitate larger datasets for optimal performance without risking over fitting.

In comparing Support Vector Machines with other machine learning models for risk prediction and small dataset analysis, it becomes clear that SVM are particularly well-suited for scenarios where data is limited and interpretability is vital. While ensemble methods generally outperform SVM in accuracy when sufficient data is available, the robustness of SVM against overfitting and their ability to capitalize on a small number of support vectors make them a powerful tool in healthcare applications and beyond. Understanding these trade-offs is critical for practitioners seeking the best model for specific data scenarios and predictive tasks.

In addressing under-five mortality, SVM can analyze demographic health survey data to identify critical risk factors leading to high mortality rates among children.(Samuel, et al. 2024) By profiling these risk factors, interventions can be tailored to target the most vulnerable groups effectively. SVM can also be employed in neonatal care to predict adverse birth outcomes and complications in neonates, such as apnea of prematurity. Accurate predictions can facilitate timely medical interventions, ensuring better survival rates and healthcare quality for newborn. (Ngusie, *et al.* 2024). By leveraging SVM to analyze health service delivery data, stakeholders can optimize the allocation of limited healthcare resources. This includes ensuring that maternal and child health programs are directed to the regions and populations most in need, thus maximizing the impact of health initiative.

The application of Support Vector Machines in maternal and child health in Nigeria has the potential to revolutionize healthcare delivery. Through their ability to predict health outcomes, identify at-risk populations, and optimize healthcare resource allocation, SVM can significantly enhance the health and well-being of mothers and children. The expected benefits include improved diagnostic accuracy, early risk identification, informed policy-making, personalized management, and enhanced access to care, all of which are critical to improving maternal and child health outcomes in Nigeria.

Materials and Methods

This study focuses on SVM as the sole machine learning algorithm due to its strong theoretical foundation, versatility, and ability to handle complex decision boundaries in classification tasks. This section discusses the research design, data collection, preprocessing steps, model selection, evaluation metrics, and the justification for choosing SVM for this multi-class classification task with three classes: low risk, mid risk, and high risk.

Mathematical Intuition behind Support Vector Machine

Hyperplane Equation:

For binary classification, the hyperplane that separates the classes is defined as: $w \cdot x+b = 0$w \cdot x + b = 0$w \cdot x+b=0$ (1)

Where:

- w: Weight vector (determines the orientation of the hyperplane).
- x: Feature vector (input data: Age, SystolicBP, BS, BodyTemp, HeartRate).

- b: Bias term (shifts the hyperplane).

Decision Rule:
The classification decision is based on the sign of the decision function: f(x) =w·x+bf(x) = w \cdot x + bf(x)=w·x+b (2)

- If f(x) > 0f(x) > 0f(x) >0: Classify as one class (e.g., Low Risk).
- If f(x) ≤ 0f(x) \leq 0f(x)≤0: Classify as the other class (e.g., Mid Risk or High Risk).

Multi-Class Classification:
Since this is a multi-class classification problem (Low Risk, Mid Risk, High Risk), SVM employs the **one-vs-rest strategy**:

- A separate hyperplane is trained for each class to distinguish it from all other classes.
- The class with the highest confidence score is selected as the final prediction.

Objective Function:
The objective of SVM is to maximize the margin (distance between the hyperplane and the nearest data points) while minimizing classification errors. For linearly separable data, the optimization problem is:
min12‖w‖2\min \frac{1}{2} \| w \|^2min21‖w‖2
Subject to:
yi(w·xi+b) ≥1,∀iy_i (w \cdotx_i + b) \geq 1, \quad \foralliyi (w·xi+b)≥1,∀ I
Where: yiy_iyi is the label of the ithi^{th}ith sample.

Soft Margin for Non-Separable Data:
To handle overlapping classes, slack variables (ξi\xi_iξi) are introduced: min12‖ w‖ 2+C∑I = 1nξi\min \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{n} \xi_imin21‖ w‖ 2+Ci=1∑nξi
Subject to:
yi (w· xi+b) ≥ 1−ξi, ξi ≥ 0y_i (w \cdotx_i + b)\geq 1 - \xi_i, \quad \xi_I \geq 0yi(w· xi+b) ≥1−ξi, ξi ≥ 0

- CCC: Regularization parameter that balances maximizing the margin and minimizing misclassifications.
- ξi\xi_iξi: Slack variable penalizing misclassified points.

Kernel Trick for Non-Linear Separation:
For non-linearly separable data, the input features are mapped to a higher-dimensional space using a kernel function. The Radial Basis Function (RBF) kernel is commonly used.
K(xi,xj) = \exp(-\gamma \| x_i - x_j \|^2) (3)
K(xi,xj) = exp(−γ‖ xi−xj‖ 2) (4)
where:

- γ\gammaγ: Determines the influence of a single training example.
- x_jxi,xj: Input feature vectors.

Infusing Input Data into the Equation

Input Data (Features): Suppose a patient has the following measurements:

- Age = 30 years
- SystolicBP = 140 mmHg
- BS = 7.5 mmol/L
- BodyTemp = 98.6°F
- HeartRate = 85 bpm

The corresponding feature vector is:
x = [30,140,7.5,98.6,85]

Weight Vector and Bias Term:
Assume the trained SVM model produces the following weight vector and bias term for the hyperplane separating **Low Risk** from the other classes:

- W = [0.02,0.1,−0.5,0.3,0.05]
- b = −0.4b

Substitute into the Decision Function:
The decision function is:

f(x) = w· x+bf(x) = w \cdot x + bf(x)=w· x+b (6)

Substituting the values:

f (x) = (0.02· 30)+(0.1· 140)+(−0.5· 7.5)+(0.3· 98.6)+(0.05· 85)−0.4 f(x) = (0.02 \cdot 30) + (0.1 \cdot 140) + (-0.5 \cdot 7.5) + (0.3 \cdot 98.6) + (0.05 \cdot 85) - 0.4 f(x) = (0.02· 30)+(0.1· 140)+(−0.5· 7.5)+(0.3· 98.6)+(0.05· 85)−0.4

Calculate:

f (x) = 0.6+14−3.75+29.58+4.25−0.4 = 44.28 f(x) = 0.6 + 14 - 3.75 + 29.58 + 4.25 - 0.4 = 44.28 f(x) = 0.6+14−3.75+29.58+4.25−0.4 = 44.28

Since f(x)>0f(x) > 0f(x)>0, the model classifies this input as **Low Risk**.

Repeat for Other Hyperplanes:
The above process is repeated for hyperplanes separating Mid Risk and High Risk from the rest. The class with the highest confidence score (i.e., largest f(x)f(x)f(x)) is chosen as the final prediction.

Example Workflow

Input Features:
Age = 30, SystolicBP = 140, BS = 7.5, BodyTemp = 98.6, HeartRate = 85

Output:
After evaluating the decision function for all classes, the model predicts the RiskLevel as **Low Risk**.

Research Design

The research adopts a quantitative research design that leverages machine learning to predict health risks based on numerical indicators. The design includes several phases:

Data Collection: The dataset is collected from a publicly available Kaggle dataset on maternal health.

Data Preprocessing: Necessary cleaning and transformations are applied to the dataset such as Cleaning, feature selection, scaling, and addressing class imbalance.

Model Training and Evaluation: SVM was trained and optimized to classify maternal health risks into low risk, mid risk, and high risk.

Deployment: The trained SVM model was deployed for real-time risk prediction using Streamlit.

System Architecture

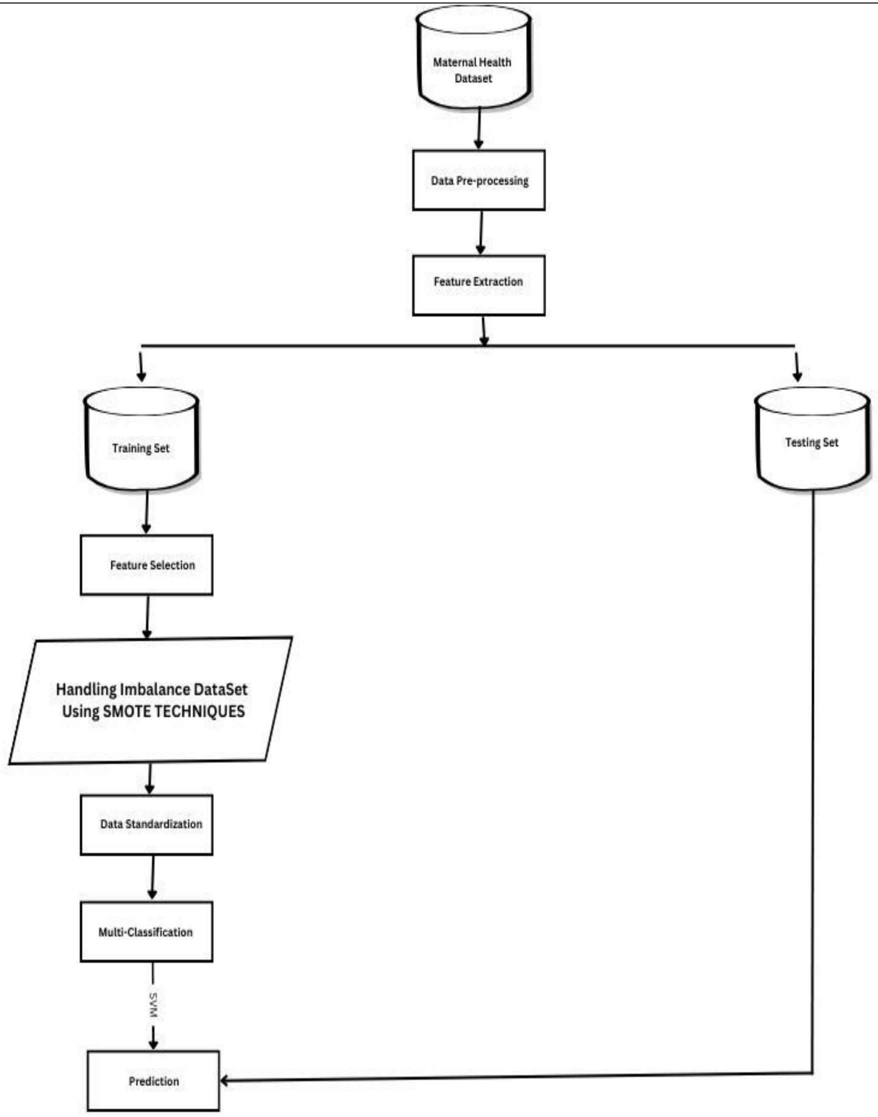


Figure 1: System Workflow Architecture

Data Collection and Description

The dataset utilized in this study is a public Kaggle Maternal Health Risk Data Set. It contains indicators commonly used in predicting maternal health risks during pregnancy. The features in the dataset include:

Age: The mother's age (years).
SystolicBP: Systolic blood pressure (mmHg).
DiastolicBP: Diastolic blood pressure (mmHg).
BodyTemp: Body temperature (°F).
HeartRate: Pulse rate (beats per minute).
Risk Level: The target variable representing maternal health risk, categorized as:
Low Risk (0)
Mid Risk (1)
High Risk (2)

The dataset is publicly available on Kaggle and was preloaded as a CSV file, making it suitable for machine learning model training and testing.

Data Preprocessing

Data Cleaning

The initial dataset was reviewed to identify any issues such as missing data, irrelevant features, or inconsistencies.

Missing Values: The dataset was checked for missing values, which were addressed appropriately by imputation or removal of affected rows/columns.

Dropping Irrelevant Features: While DiastolicBP was included in the initial dataset, it was dropped after analyzing its correlation with SystolicBP. Both features were highly correlated, which could lead to multicollinearity.

Feature Encoding

The target variable Risk Level, which contains categorical values (low risk, mid risk, high risk), was converted into numeric format using label encoding:

Low Risk: 0
Mid Risk: 1
High Risk: 2

This transformation allowed the target variable to be used in machine learning models.

Handling Class Imbalance

The Risk-Level variable exhibited significant class imbalance, with a higher number of low risk cases compared to mid risk and high risk cases. This imbalance could bias the machine learning models toward predicting the majority class. To address this, SMOTE (Synthetic Minority Over-sampling Technique) was applied to the training data to oversample the minority classes, ensuring that all classes were equally represented in the training process.

Feature Scaling

Since the dataset contains features with varying scales, Standard-Scaler was used to standardize all feature values to have zero mean and unit variance. This step ensures that no single feature dominates the learning process, especially for models like SVM, which are sensitive to the scale of input features.

Support Vector Machine (SVM)

SVM was selected due to its efficacy in classification tasks, especially in high-dimensional spaces. SVM is particularly useful when dealing with complex decision boundaries, which is often the case in healthcare-related datasets.

Kernel: The RBF kernel was used, as it is effective for non-linear classification problems.

Hyperparameter Tuning: GridSearchCV was employed to optimize the following hyperparameters:

C: Regularization parameter to control overfitting.
Gamma: Kernel coefficient that controls the influence of individual training examples.
Kernel: Fixed as RBF for non-linear decision boundary creation

Hyper-parameter Optimization

The performance of SVM depends on key hyperparameters such as:

C: Controls the trade-off between maximizing the margin and minimizing classification error.
Gamma: Determines the influence of individual data points on the decision boundary.

These hyperparameters were optimized using GridSearchCV with 5-fold cross-validation.

Model Training and Validation

The dataset was split into training (80%) and testing (20%) subsets using train_test_split fromscikit-learn. The training data was then balanced using SMOTE to ensure equal representation of each class. Each model was trained on the preprocessed training data and evaluated on the test data. GridSearchCV was used for hyperparameter optimization with 5-fold cross-validation to ensure robust model validation.

Results and Discussion

This section presents the implementation, analysis, and evaluation of the Support Vector Machine (SVM) model used to predict maternal health risks. The model was trained and tested using a datasets of maternal health indicators, and its performance was evaluated based on metrics such as accuracy, precision, recall, and F1-score. The chapter also includes the deployment of the model for real-time predictions using Streamlit.

Data Cleaning

The datasets originally contained multiple features related to maternal health indicators, including systolic and diastolic blood pressure (BP). During feature correlation analysis using a heat map, it was observed that Diastolic BP exhibited a high correlation with Systolic BP. Since Systolic BP is more clinically significant in identifying risk levels for hypertension-related complications during pregnancy, Diastolic BP was dropped to avoid redundancy and multicollinearity issues.

This decision aligns with established medical knowledge, as systolic BP is often prioritized when screening for conditions such as pre-eclampsia and gestational hypertension, both of which are critical risk factors for maternal mortality.

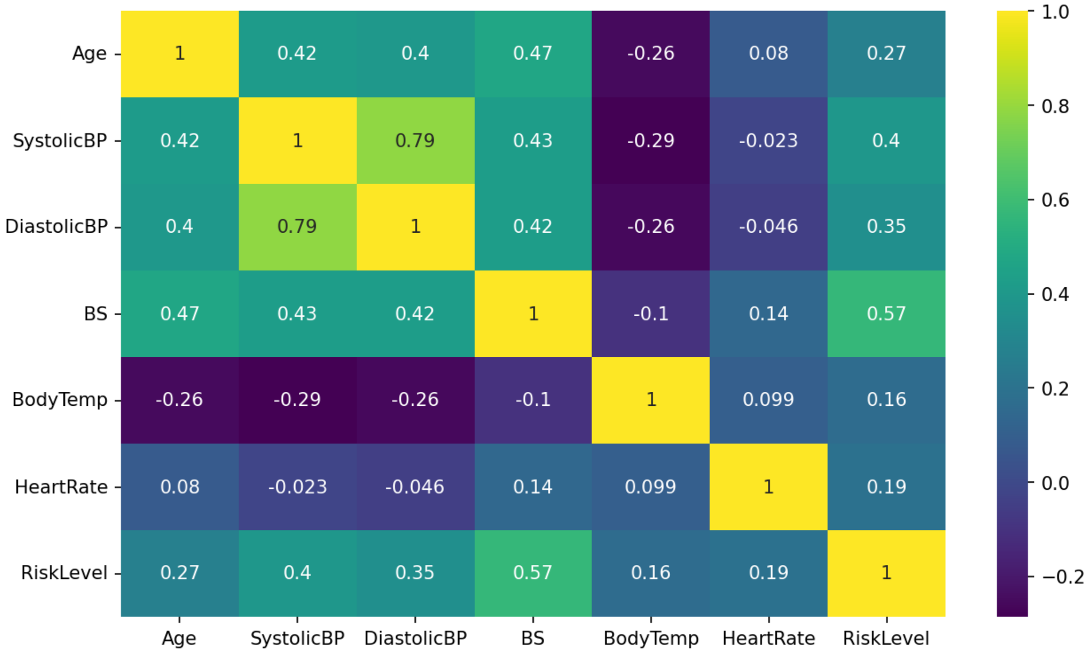


Fig 2:*The above heat map denote that Systolic BP is highly correlated to Diastolic BP*

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is the initial step in data analysis where you examine and summarize the main characteristics of your data to gain insights and understand its structure, it calculates descriptive statistics like mean, median, standard deviation, and quartile for numerical features using df.describe() function. This gives you a sense of the central tendency and spread of your data.

Table 1. Features of data studied

Statistic	Age	Systolic BP	BS	Body Temp	Heart Rate
Count	1014.0000	1014.0000	1014.0000	1014.0000	1014.0000
Mean	29.8718	113.1982	8.7260	98.6651	74.3018
Std	13.4744	18.4039	3.2935	1.3714	8.0887
Min	10.0000	70.0000	6.0000	98.0000	7.0000
25%	19.0000	100.0000	6.9000	98.0000	70.0000
median (50%)	26.0000	120.0000	7.5000	98.0000	76.0000
75%	39.0000	120.0000	8.0000	98.0000	80.0000
Max	70.0000	160.0000	19.0000	103.0000	90.0000

Feature Engineering

Other continuous variables like maternal age, body temperature, and pulse rate were retained for analysis. The target variable, Risk Level, originally contained categorical labels (low risk, mid risk, and high risk). These were encoded numerically as 0, 1, and 2, respectively, to facilitate machine learning.

Addressing Class Imbalance

Class distribution analysis revealed significant imbalances in the Risk Level variable, with low risk instances dominating the datasets. This imbalance was addressed using Synthetic Minority Over-sampling Technique (SMOTE), which generated StandardScaler was applied to standardize all features by removing the mean and scaling them to unit variance:

z = x−μσz = \frac{x - \mu}{\sigma}z=σx−μ

Where:

- xxx: Original feature value.
- μ\muμ: Mean of the feature.
- σ\sigmaσ: Standard deviation of the feature.

Results

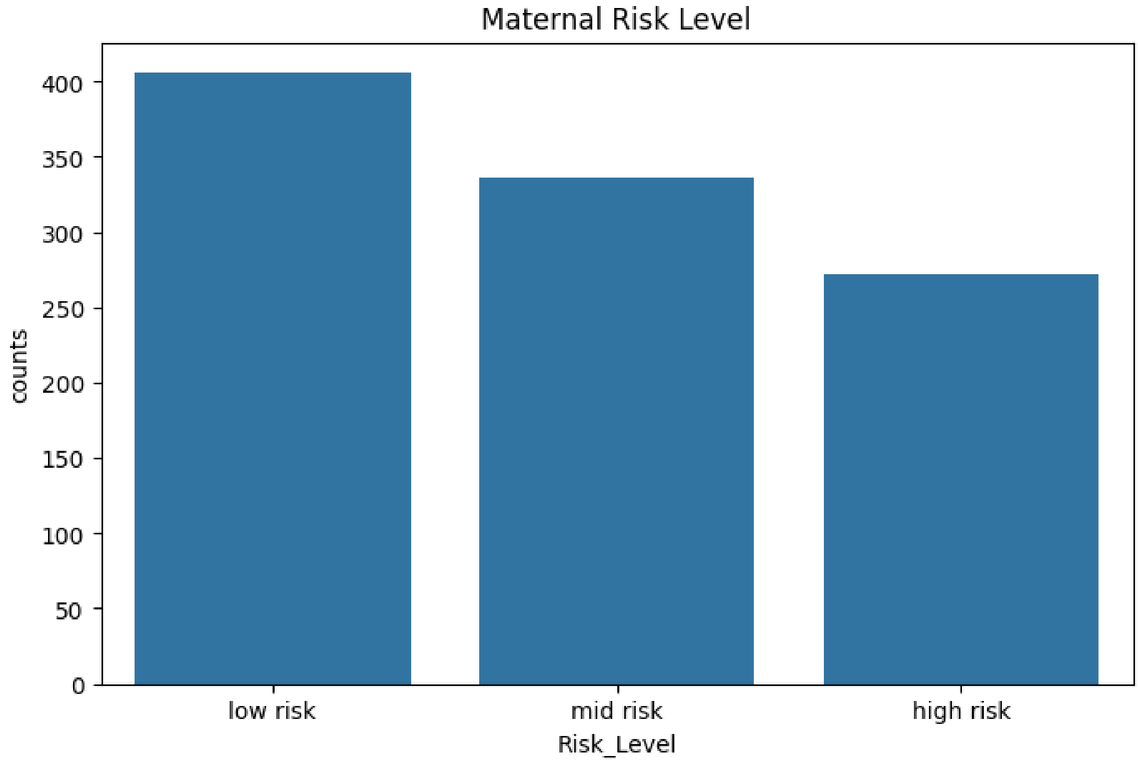


Fig 3:Pregnancy Risk Level

Handling Imbalanced Datasets Using SMOTE

The target variable, Risk Level, exhibited a significant imbalance, with most records categorized as Low Risk and fewer as Mid Risk or High Risk. An imbalanced datasets can bias the model toward the majority class, reducing its ability to correctly predict minority classes. To address this, the **Synthetic Minority Over-sampling Technique (SMOTE)** was applied:

- **How SMOTE Works**: SMOTE generates synthetic samples for minority classes by interpolating between existing samples. For example, for a given feature vector xxx in the minority class, SMOTE selects one of its nearest neighbors xneighborx_{\text{neighbor}}xneighbor and creates a synthetic sample: xsynthetic = x+λ· (xneighbor−x)x_{\text{synthetic}} = x + \lambda \cdot (x_{\text{neighbor}} - x) xsynthetic =x+λ· (xneighbor−x). Where: λ\lambdaλ is a random number between 0 and 1.
- **Outcome**:
  The training datasets became balanced, ensuring that the SVM model was trained equally on Low Risk, Mid Risk, and High Risk samples.

Data standardization

Data standardization is a preprocessing technique used in machine learning and data analysis to transform numerical data into a standard format. This ensures that features have a similar range of values and helps improve the performance of algorithms that are sensitive to the scale of features

Model Implementation

Support Vector Machine (SVM)

The SVM model used the RBF kernel, which is suitable for non-linear decision boundaries often encountered in healthcare datasets. Hyper parameters such as C (regularization strength) and gamma (kernel coefficient) were optimized using GridSearchCV, yielding the following results:
Optimal Parameters: C = 10, gamma = 0.1, kernel = 'rbf'
The model was trained and evaluated, achieving the following metrics:
Training Accuracy: 0.92%
Test Accuracy: 0.73%

Evaluation Metrics

The evaluation metrics used for both algorithms included accuracy, precision, recall, F1-score, and confusion matrices.

SYM Classification Report:

	precision	recall	f1-score	support
0	**0.85**	**0.66**	**0.75**	**80**
1	**0.74**	**0.80**	**0.77**	**76**
2	**0.69**	**0.87**	**0.77**	**47**

accuracy 0.76 203
macro avg 0.76 0.78 0.76 203
weighted avg 0.78 0.76 0.76 203

Confusion Matrix Comparison

SVM: Displayed misclassifications between midrisk and high risk categories.

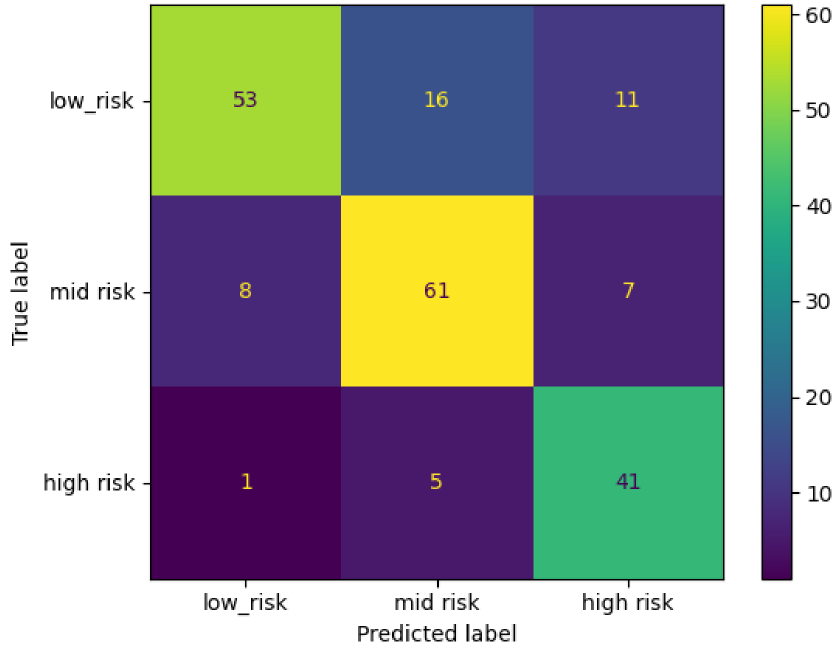


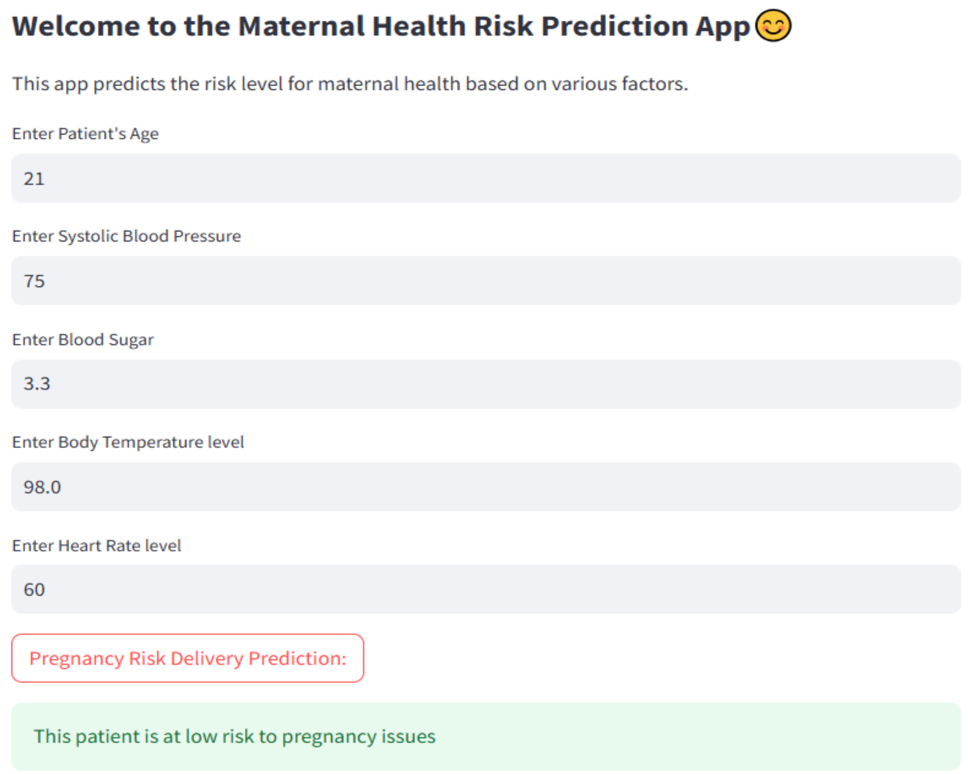
Fig 4:Misclassification between mid-risk and high risk categories

Deployment Platform:

The application was hosted on a cloud platform (Streamlit) to make it accessible to healthcare professionals via a web browser.

Example Workflow in the Application

Input: Age = 35, SystolicBP = 120, Blood Sugar= 7.5, BodyTemp = 98.0, HeartRate = 76.
Output: Predicted Risk Level: Low Risk



Welcome to the Maternal Health Risk Prediction App 😊

This app predicts the risk level for maternal health based on various factors.

Enter Patient's Age

21

Enter Systolic Blood Pressure

75

Enter Blood Sugar

3.3

Enter Body Temperature level

98.0

Enter Heart Rate level

60

This patient is at low risk to pregnancy issues

Figure 5: Model implementation

Discussion

This research explores the use of **Support Vector Machines (SVM)** to address maternal and child health challenges in Nigeria, where high maternal and child mortality rates persist. SVM, a supervised machine learning algorithm, is employed to predict risks during pregnancy and childbirth by analyzing health indicators such as blood pressure, age, and heart rate. The study aims to develop a predictive model to identify high-risk cases early and support healthcare professionals in implementing timely interventions.

The methodology involves data collection, preprocessing, model development, and evaluation. Key steps include data cleaning, feature engineering, and handling class imbalances using techniques like SMOTE. The model achieved reasonable accuracy in predicting maternal health risks and was deployed using Streamlit for real-time predictions.

Conclusion

The findings highlight the potential of machine learning techniques, especially SVM, in improving maternal and child healthcare in resource-constrained environments like Nigeria. The model demonstrated effective performance in classifying health risks into low, mid, and high categories, providing a scalable and cost-effective solution for healthcare providers. Its deployment ensures real-time decision-making capabilities that can significantly reduce mortality rates by enabling early detection and targeted interventions.

Despite its success, the study acknowledges limitations such as reliance on a single dataset, exclusion of socioeconomic variables, and potential biases introduced during synthetic oversampling. These factors may limit the model's generalizability to other populations.

List of Abbreviations

Support Vector Machine (SVM),
Synthetic Minority Over-sampling Technique (SMOTE)
Exploratory Data Analysis (EDA)

References

Ajegbile, M. L. (2023). Closing the gap in maternal health access and quality through targeted investments in low-resource settings. *Journal of Global Health Reports*.

Akeju, D., Okusanya, B., Okunade, K., Ajepe, A., Allsop, M., &Ebenso, B. (2022). Sustainability of the effects and impacts of using digital technology to extend maternal health services to rural and hard-to-reach populations: Experience from Southwest Nigeria. *Frontiers in Global Women's Health*.

Amose, J., Manimegalai, P., Scholar, R., &Narmatha, C. (2022). Comparative performance analysis of kernel functions in support vector machines in the diagnosis of pneumonia using lung sounds. *2022 2nd International Conference on Computing and Information Technology (ICCIT).*

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*.

Aswini, J., Yamini, B., Jatothu, R., Nayaki, K. S., &Nalini, M. (2021). An efficient cloud-based healthcare services paradigm for chronic kidney disease prediction application using boosted support vector machine. *Concurrency and Computation: Practice and Experience*.

Azuogu, B., Akamike, I., Okedo-Alex, I., Adeke, A., Agu, A., Akpa, C., Obiechina, N., Akpa, W., Nwali, D., & Anyigor, C. (2020). Knowledge, attitude, and perceived partner and socio-cultural support for family planning among women of reproductive age in a rural community in Ebonyi State, Nigeria.

Badillo, S., Banfai, B., Birzele, F., Davydov, I. I., Hutchinson, L., Kam-Thong, T., & Zhang, J. D. (2020). An introduction to machine learning. *Clinical Pharmacology & Therapeutics*, *107*(4), 871–885.

Bukar, L. (2019). The National Health Insurance Scheme: An alternative source of healthcare funding in Nigeria—A case study of Borno State. *Texila International Journal of Management*.

Dahab, R., &Sakellariou, D. (2020). Barriers to accessing maternal care in low-income countries in Africa: A systematic review. *International Journal of Environmental Research and Public Health*.

Dixon, D., Sattar, H., Moros, N., Kesireddy, S. R., Ahsan, H., Lakkimsetti, M., Fatima, M., Doshi, D., Sadhu, K., &Junaid Hassan, M. (2024). Unveiling the influence of AI predictive analytics on patient outcomes: A comprehensive narrative review.

Ekpenyong, B., Agbaje, E., & Omotoye, J. (2019). Healthcare delivery in Nigeria: Addressing maternal and child mortality. *African Health Sciences, 19*(4), 3024–3032.

Feldner-Busztin, D., Nisantzis, P. F., Edmunds, S. J., Boza, G., Racimo, F., Gopalakrishnan, S., Limborg, M., Lahti, L., & Polavieja, G. G. de. (2023). Dealing with dimensionality:

The application of machine learning to multi-omics data. *Bioinformatics*.

ForeSee Medical. (2024). *Predictive analytics in healthcare: Explore benefits & applications*.

Habehh, H., &Gohel, S. (2021). Machine learning in healthcare. *Current Genomics, 22*(4), 291.

Mennickent, D., Rodríguez, A., Opazo, M. C., Riedel, C. A., Castro, E., &Eriz-Salinas, A. (2022). What is machine learning: Definition, types, applications and examples. *Potentia Analytics*.

Harahap, T. H., Mansouri, S., Abdullah, O. S., Uinarni, H., Askar, S., Jabbar, T. L., Alawadi, A. H., & Hassan, A. Y. (2024). An artificial intelligence approach to predict infants' health status at birth. *International Journal of Medical Informatics*.

IBM. (2024). What is support vector machine? *IBM*.

Iqbal, F., Satti, M. I., Irshad, A., & Shah, M. A. (2023). Predictive analytics in smart healthcare for child mortality prediction using a machine learning approach. *Open Life Sciences*.

Lanza, B., & Parashar, D. (2020). Do support vector machines play a role in stratifying patient population based on cancer biomarkers? *Archives of Proteomics and Bioinformatics*.

Mlandu, C., Matsena-Zingoni, Z., & Musenge, E. (2023). Predicting the dropout from the maternal, newborn, and child healthcare continuum in three East African Community countries: Application of machine learning models. *BMC Medical Informatics and Decision Making*.

Ngusie, H., Mengiste, S., Zemariam, A., & Molla, B. (2024). Predicting adverse birth outcomes among childbearing women in Sub-Saharan Africa: Employing innovative machine learning techniques. *BMC Public Health*.

Nigeria Health Watch. (2024). *An unseen grief: Maternal mortality's impact on infants and children*. *Medium*.

Obasohan, P. E., Gana, P., Mustapha, M. A., Umar, A. E., Makada, A., & Obasohan, D. N. (2019). Decision-making autonomy and maternal healthcare utilization among Nigerian women. *International Journal of Maternal and Child Health and AIDS (IJMA)*.

Ogu, U. U., et al. (2023). Demand and supply analysis for maternal and child health services at the primary healthcare level in Nigeria. *BMC Health Services Research*.

Ope, B. W. (2020). Reducing maternal mortality in Nigeria: Addressing maternal health services' perception and experience. *Journal of Global Health Reports*.

Olonade, O., Olawande, T. I., & Alabi, O. J. (2019). Maternal mortality and maternal healthcare in Nigeria: Implications for socioeconomic development.

Petrova, B. (2024). Predictive analytics in healthcare | Reveal. *Reveal Embedded Analytics*.

Samuel, O., Zewotir, T., & North, D. (2024). Application of machine learning methods for predicting under-five mortality: Analysis of Nigerian demographic health survey 2018 dataset. *BMC Medical Informatics and Decision Making*.

Sarker, I. H. (2021). Machine learning: Algorithms, real-world applications, and research directions. *SN Computer Science, 2*(3), 160.

Mao, W., & Watkins, D. (2023). Effects of public financing of essential maternal and child health interventions across wealth quintiles in Nigeria: An extended cost-effectiveness analysis.

Yarasuri, V. K., Reddy, D. S., Muneesh, P. S., Kaushik, R. V. S., Vardhan, T. N., & Nisha, K. L. (2022). Developing machine learning models for cardiovascular disease prediction. *2022 2nd Asian Conference on Innovation in Technology*.