IJEMD-CSAI, 4 (1) (2025) https://doi.org/ 10.54938/ijemdcsai.2025.04.1.419

International Journal of Emerging Multidiciplinaries:
Computer Science & Artificial Intelligence

Research Paper
Journal Homepage: www.ijemd.com
ISSN (print): 0000-0000

Explainable AI in Maternal Health: Utilizing XGBoost and SHAP Values for Enhanced Risk Prediction and Interpretation

Mohammad Arif [1*]

1.Queen Elizabeth School, Barnet, EN9 4DQ, United Kingdom.

Abstract

This research investigates the integration of Explainable Artificial Intelligence (XAI) in maternal health risk prediction, with a focus on improving the transparency and clinical utility of predictive models. Maternal mortality persists as a global challenge, disproportionately affecting developing nations where healthcare systems often rely on opaque predictive tools trained on limited datasets. To address these gaps, this study analyzes a comprehensive dataset spanning clinical, physiological, and historical health metrics, applying both traditional and advanced machine learning models. By incorporating SHapley Additive exPlanations (SHAP) value analysis, the interpretability of risk predictions was enhanced while maintaining high diagnostic accuracy. The findings indicate that the XGBoost model achieved an impressive accuracy of 96.36%, with body mass index and preexisting diabetes emerging as the most significant risk determinants. Clinical insights from highly-renowned healthcare providers were actively sought during this study to contextualize the model's implications within real-world clinical practice. These insights enable healthcare providers to prioritize high-impact variables when designing interventions, bridging the gap between algorithmic outputs and actionable clinical strategies.

Keywords: Maternal Health, XGBoost, Risk Stratification, Explainable AI, Machine Learning, Predictive Modeling

Email address: mo.tarif1221@gmail.com (Muhammad Arif)

1. INTRODUCTION

Each year, an estimated 295,000 women lose their lives to pregnancy-related complications, a stark reminder of the urgent global health crisis in maternal care [1]. This alarming figure highlights the critical need for better strategies to predict, manage, and mitigate risks during pregnancy. Maternal health complications such as preeclampsia, postpartum hemorrhage, and sepsis are complex and varied, influenced by biological, social, and environmental factors. To address these challenges, a deeper understanding of the conditions that increase vulnerability during pregnancy is essential [2, 3].

The importance of maternal health risk prediction cannot be overstated. Complications like gestational diabetes or preterm birth endanger both mothers and infants, with long-term consequences for families and communities [4]. Risks are amplified by factors, such as older maternal age, pre-existing chronic illnesses (e.g., hypertension), limited healthcare access, and lifestyle habits like smoking. Women in low-income or marginalized communities face even greater disparities due to systemic barriers to quality care [5]. Compounding these issues is the lack of clinical trial data on medication safety during pregnancy, leaving healthcare providers with incomplete guidance [6]. Early detection of risks, paired with timely interventions, is therefore vital to safeguarding maternal and fetal health.

Existing risk assessment tools, such as the Alberta Perinatal Health Program's 39-item checklist and Nigeria's Community Maternal Danger Score (CMDS), have improved how healthcare systems categorize at-risk pregnancies. However, these tools often fall short in terms of accuracy, comprehensiveness, and accessibility [7]. For instance, the Alberta tool identifies high-risk cases but fails to prevent poor outcomes for many in this group [8]. Similarly, while the CMDS outperforms standard evaluations in detecting risks, it fails to provide real-time updates during prenatal visits [9]. Emerging tools like the R4U scorecard reveal that non-medical factors, such as income, education, and social support play a pivotal role in pregnancy outcomes [10]. These gaps underscore the need for more precise, dynamic, and inclusive risk assessment methods to ensure equitable care.

Machine learning (ML) has emerged as a pivotal tool in enhancing maternal health risk assessment. By analyzing vast datasets, encompassing medical histories, demographic trends, and clinical biomarkers, ML algorithms can detect subtle patterns linked to complications like preeclampsia or preterm birth [11]. This enables earlier, more accurate risk predictions, helping healthcare providers tailor interventions to individual needs [12]. As digital health technologies evolve, ML-driven systems are becoming indispensable tools for clinicians, offering real-time decision support to reduce preventable deaths [13]. This study proposes a new ML-based framework designed to improve risk classification using comprehensive, diverse data. By integrating medical and socioeconomic factors, the framework aims to empower healthcare teams with actionable insights, fostering safer pregnancies and healthier outcomes for mothers and newborns.

The main objectives of this study are summarized below:

- This study introduces an inclusive framework for maternal health risk classification using Extreme Gradient Boosting (XGBoost) and comprehensive dataset, comprising clinical, physiological, and historical health information.
- The proposed framework is compared with several other traditional and state-of-the-art ML models K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Trees (DT), Random Forest (RF), Logistic Regression (LR), Light Gradient-Boosting Machine (LightGBM), Categorical Boosting (CatBoost), and Adaptive Boosting (AdaBoost) to show the efficacy of the proposed model. Furthermore,

StratifiedKFold cross-validation is applied to validate its performance complemented by a systematic grid search aimed at hyperparameter optimization.

- The performance of the diverse ML algorithms is rigorously assessed and compared using several metrics, including accuracy, precision, recall, F1, area under the precision-recall curve (AUPRC), and cross-validation scores. Furthermore, Explainable Artificial Intelligence (XAI) methodologies is applied to enhance the transparency and usability of the findings.

The paper is systematically organized to explore the proposed maternal health risk assessment framework. Section 2 presents a literature review on existing methodologies and their limitations. Section 3 outlines the framework, detailing the dataset, preprocessing, exploratory data analysis, and ML model architecture. Section 4 and 5 discusses the results and performance metrics, comparing them with traditional tools, while also addressing the implications for maternal health outcomes. Finally, Section 6 concludes the document, summarizing key insights and suggesting future research directions. This structure ensures clarity and facilitates a smooth transition into the subsequent sections.

2. PERTINENT STUDIES

ML has emerged as a transformative tool for maternal health risk prediction, offering novel ways to address complications such as preeclampsia, gestational diabetes, and perinatal mortality. Recent studies emphasize the effectiveness of tree-based models like LightGBM and XGBoost, which have delivered strong performance across varied clinical scenarios. For instance, LightGBM achieved 85% accuracy in predicting neonatal sepsis by analyzing vital signs and clinical records [14], while XGBoost attained an AUC of 0.91 for preeclampsia risk stratification using biomarkers such as blood pressure and uric acid levels [15]. Naïve Bayes classifiers also proved highly effective, achieving 97% accuracy in diagnosing hypertensive disorders [16], However, these models share a critical limitation: their reliance on imbalanced or geographically restricted datasets, such as retrospective public health data from single countries or homogeneous patient cohorts, which risks biased generalizability to global populations.

Beyond data limitations, methodological inconsistencies hinder progress in this field While gradient boosting models achieved exceptional accuracy (99%) in classifying fetal health [17], few studies incorporated cross-validation (CV) to ensure reliability [14, 18]. Hybrid frameworks like MaternalNET-RF, despite their 95% accuracy in risk categorization [19], often neglect tools to explain decision-making processes. This lack of transparency is widespread; only a minority of studies [15, 20] have adopted interpretability methods like SHapley Additive exPlanations (SHAP) to clarify how factors such as maternal body mass index or blood pressure influence predictions. Without such insights, many algorithms remain “black boxes,” eroding clinical trust and posing ethical dilemmas, particularly in low-resource settings where biased predictions could worsen healthcare inequities.

These technical challenges intersect with significant ethical risks, especially when models are applied to vulnerable populations [21]. For instance, prioritizing sensitivity over explainability as seen in LightGBM’s 72.41% performance in detecting intrahepatic cholestasis [22], may undermine patient autonomy if clinicians cannot validate algorithmic outputs. Furthermore, many studies focus narrowly on technical metrics while overlooking real-world relevance. Ensemble methods for perinatal mortality prediction in Ethiopia [18], which achieved 90% accuracy but relied heavily on region-specific factors like insurance access, underscoring the need for culturally adaptive frameworks. This disconnect highlights a broader tension between algorithmic precision and practical implementation.

Recent efforts to bridge this gap have yielded mixed results like improved accuracy but persistent explainability gaps. Generalized additive models (GAMs), for example, identified novel risk factors like in-vitro fertilization for postpartum hemorrhage [23], but their modest AUROC (0.67) and lack of validation in diverse cohorts limit clinical adoption. Similarly, while CatBoost and AdaBoost have shown promise in benchmarking studies [17, 24], their performance remains untested against comprehensive datasets integrating demographic, psychophysiological, and clinical variables. This narrow focus on isolated conditions, rather than interconnected health factors, impedes the development of universal risk assessment tools.

In summary, the current landscape of maternal health ML research is marked by three unresolved challenges: reliance on fragmented or imbalanced data, inconsistent validation practices, and insufficient explainability to support clinical adoption. Addressing these limitations requires a paradigm shift towards holistic data integration and XAI, as proposed in this study, one that prioritizes robust, transparent frameworks trained on diverse data. By integrating CV, XAI, and systematic benchmarking, future models can transcend narrow applications to deliver equitable, actionable insights for maternal care. The summary of pertinent studies in provided in Table 1.

Table 1: Summary of Pertinent Literature

Reference	Classifiers	Performance	Limitations
Jathanna, R. D. et al. [14]	LR, NB, RF, DT, KNN, SVM, AdaBoost, Extra Trees, Gradient Boosting, Linear Discriminant Analysis, LightGBM	Accuracy: 85.36% (LightGBM)	Imbalanced dataset utilized, No Explainable AI
Kovacheva, V. P. et al. [15]	XGBoost, LR	AUC: 91% (XGBoost)	No cross-validation, only AUC is evaluated
Seeta and Shivali [16]	DT, SVM, RF, NB, XGBoost	Accuracy: 97% (NB)	Imbalanced dataset utilized, No cross-validation, No Explainable AI
Kaliappan, J. et al [17]	DT, RF, KNN, NB, SVC, AdaBoost, Gradient Boosting, Voting classifier, Feed Forward Network	Accuracy: 99% (Gradient Boosting)	Imbalanced dataset utilized
Bogale, D. S. et al. [18]	RF, Gradient Boosting, CatBoost	Accuracy: 90.24 (Gradient Boosting)	Imbalanced dataset utilized, No Explainable AI
Togunwa, T. O. et al. [19]	RF, DT, KNN, LR, NB, SVM, Extra Trees, LightGBM, Gradient Boosting, Linear Discriminant Analysis, Ridge, Quadratic Discriminant Analysis, AdaBoost, Dummy, ANN, MaternalNET-RF	Accuracy: 94.88% (MaternalNET-RF)	Imbalanced dataset utilized, No Explainable AI
Kang, B. S. et al. [20]	LightGBM, XGBoost	AUC: 80.04%, AUPR: 44.02% (XGBoost)	No cross-validation, only AUC and AUPR are evaluated
Zhang, X et al [22]	LR, RF, LightGBM	Accuracy: 72.41% (LightGBM)	No cross-validation, No Explainable AI
Lengerich, B. J. et al. [23]	GAM	AUROC: 67% (GAM)	Imbalanced dataset, only AUROC, Recall, and Precision were evaluated

Assaduzzaman, M. et al. [24]	DT, RF, CatBoost, XGBoost, Gradient Boosting	Accuracy: 90% (RF)	Imbalanced dataset utilized, No cross-validation, No Explainable AI

3. MATERIALS & METHODS

3.1 Data Insights and Operational Framework

The maternal health risk assessment dataset [25] used in this study was sourced from the Mendeley Data repository. This dataset aggregates clinical, physiological, and historical health records of maternal patients to assess pregnancy-related risk factors. It included eleven health metrics, categorized into three domains: metabolic indicators (age, body temperature, blood sugar), cardiovascular parameters (diastolic and systolic blood pressure, heart rate), and psychosocial and medical history markers (mental health, body mass index (BMI), gestational diabetes, preexisting diabetes, and previous complications). All these features were represented as numerical variables. The target variable, designated as Risk Level, was categorical and classified into two classes: high risk and low risk. Thus, totaling twelve variables in the dataset. The dataset encompasses 1,205 instances, of which 727 (60.3%) were classified as low risk and 478 (39.7%) as high risk as depicted in Figure 1. Data preprocessing, predictive modeling, and analytical evaluations were conducted using Python (v3.10.11) utilizing packages such as numpy (1.21.6), pandas (1.3.5), shap (0.40.0), scikit-learn (1.0.2), imbalanced-learn (0.8.1), matplotlib (3.5.2), seaborn (0.11.2).

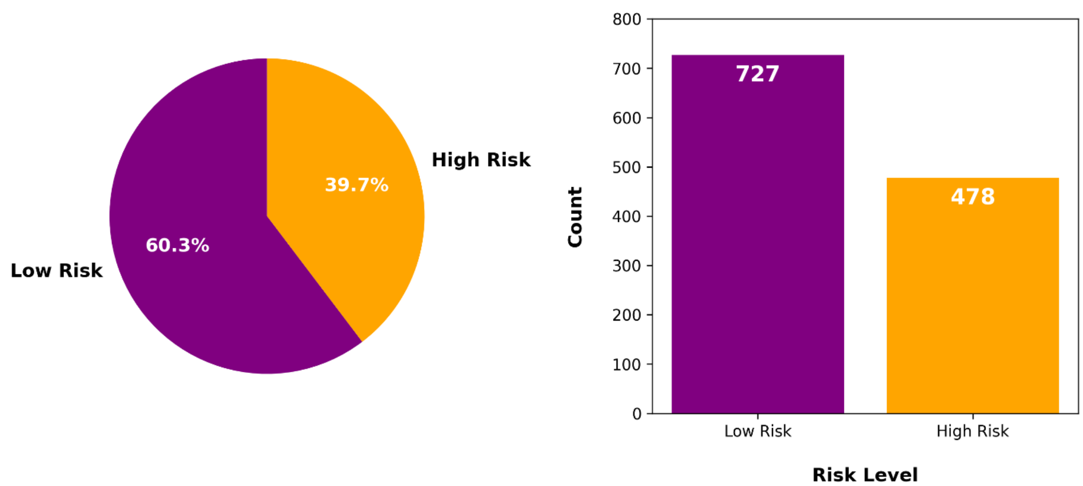


Figure 1: Distribution of maternal risk levels.

3.2 Data Preprocessing and Exploratory Data Analysis

The primary preprocessing steps implemented for the analysis included mode imputation to address missing values in numerical features and label encoding for the categorical target variable, which assigned binary labels (0 = low-risk class, 1 = high-risk class). Initial statistical evaluation identified biologically implausible outliers in the dataset: an age value of 325 and a BMI value of 0, both likely stemming from imputation errors. These two low-risk records were corrected by replacing the outliers with median values derived from the dataset (age: 25, BMI: 23.0) prior to subsequent analysis and modeling. Key statistical descriptors of the processed dataset are summarized in Table 2. To assess variable relationships, correlation analysis was conducted as displayed in Figure 2. The risk level demonstrated the strongest positive associations with preexisting diabetes (r = 0.67), mental health status (r = 0.63), and blood sugar levels (r = 0.59). In contrast, body temperature (r = 0.22) and age (r = 0.18) showed negligible correlations with risk classification.

Table 2: Statistical details of the processed dataset features.

Statistics	Age	Systolic BP	Diastolic BP	Blood Sugar	Body Temp	Body Mass Index	Previous Complications	Preexisting Diabetes	Gestational Diabetes	Mental Health	Heart Rate
Count	1205	1205	1205	1205	1205	1205	1205	1205	1205	1205	1205
Mean	27.48	116.83	77.18	7.5	98.4	23.33	0.18	0.29	0.12	0.33	75.81
Std	9.2	18.68	14.28	3.05	1.09	3.79	0.38	0.45	0.32	0.47	7.23
Min	10	70	40	3	97	15	0	0	0	0	58
25%	21	100	65	6	98	21	0	0	0	0	70
50%	25	120	80	6.9	98	23	0	0	0	0	76
75%	31	130	90	7.9	98	25	0	1	0	1	80
Max	65	200	140	19	103	37	1	1	1	1	92

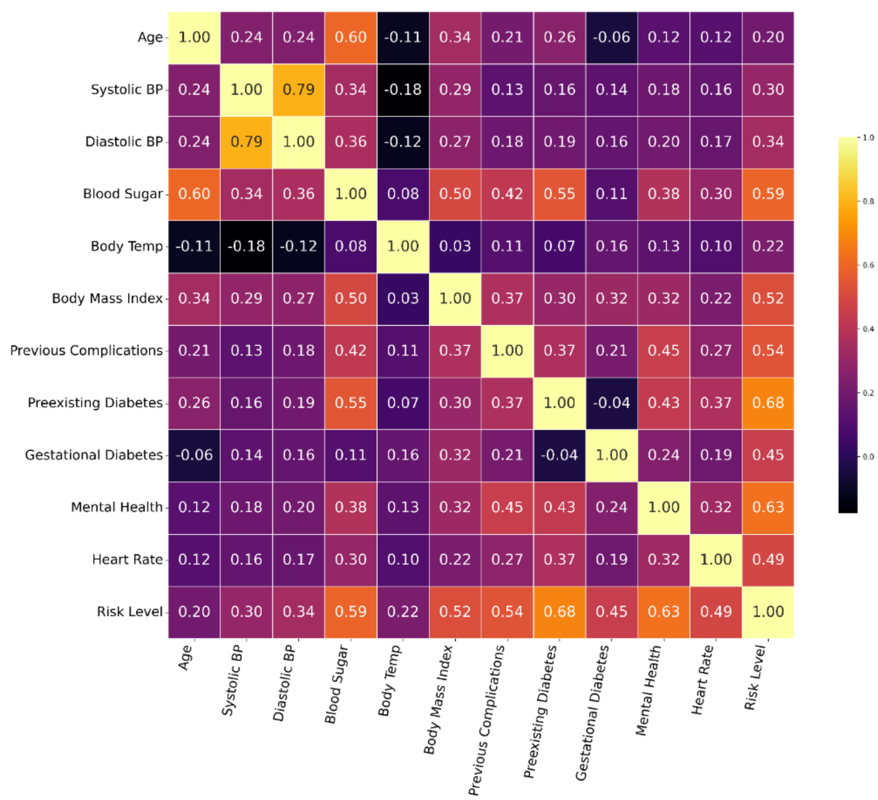
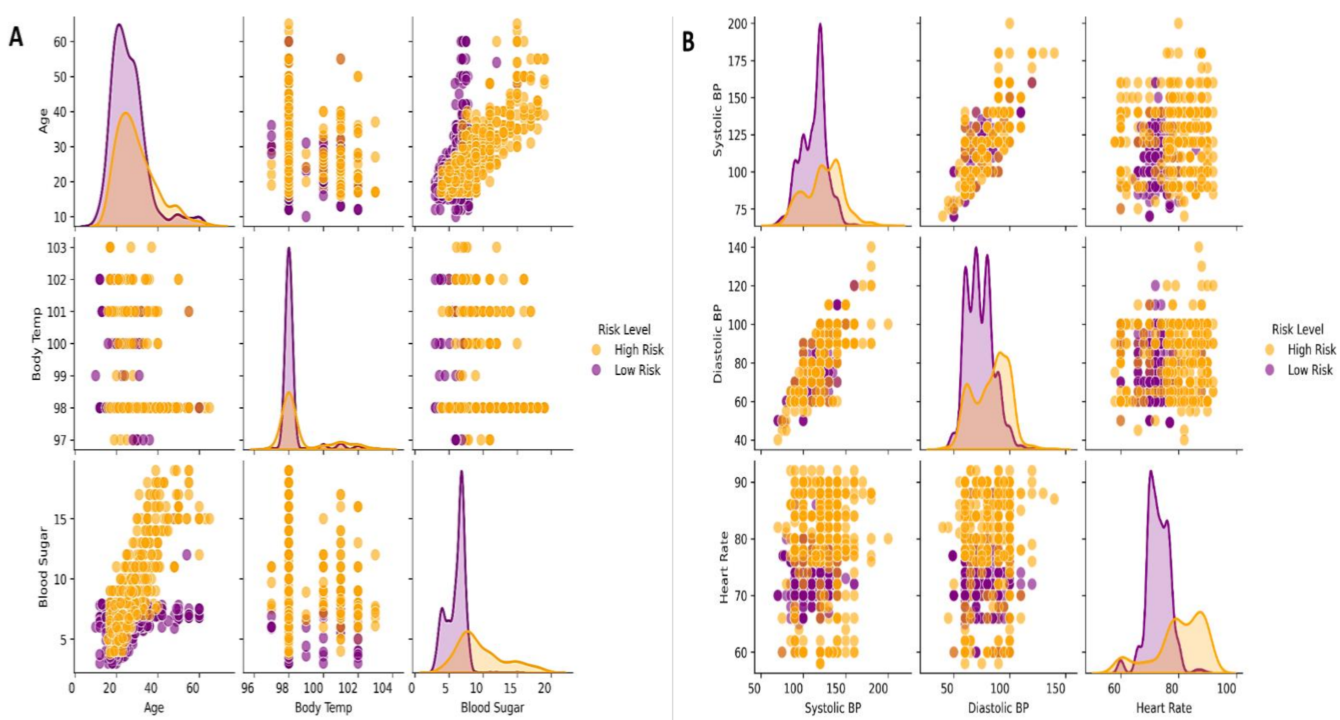


Figure 2: Correlation between risk indicators and maternal risk level.

Subsequent exploratory data analysis (EDA) revealed key associations between clinical risk indicators and maternal health outcomes, as demonstrated in Figure 3A. Maternal patients aged over 40 years with blood sugar levels exceeding 10 mmol/L exhibited higher susceptibility to high-risk classifications, even when body temperature remained within normative ranges (98–100°F). Conversely, individuals with blood sugar levels below 10 mmol/L were predominantly categorized as low-risk across all age groups. Age emerged as a critical determinant of risk stratification, with subjects over 40 years displaying elevated risk density, whereas younger participants (<40 years) demonstrated a higher prevalence of low-risk classifications. Body temperature exhibited no significant correlation with risk levels, as both high- and low-risk cohorts were evenly distributed across observed temperature ranges. These findings underscore blood sugar and advanced maternal age as the most salient predictors of adverse pregnancy-related outcomes. Furthermore, Figure 3B highlights that systolic blood pressure ≥150 mmHg combined with diastolic blood pressure ≥100 mmHg strongly correlated with high-risk designations. In contrast, participants with optimal resting heart rates (65–75 bpm) and normotensive blood pressure levels were predominantly classified as low-risk. Figure 3C illustrates that previous complications and preexisting diabetes were significantly associated with elevated risk profiles. Notably, Figure 3D reinforces BMI as an independent risk factor, with higher BMI values corresponding to increased likelihoods of high-risk categorization

.



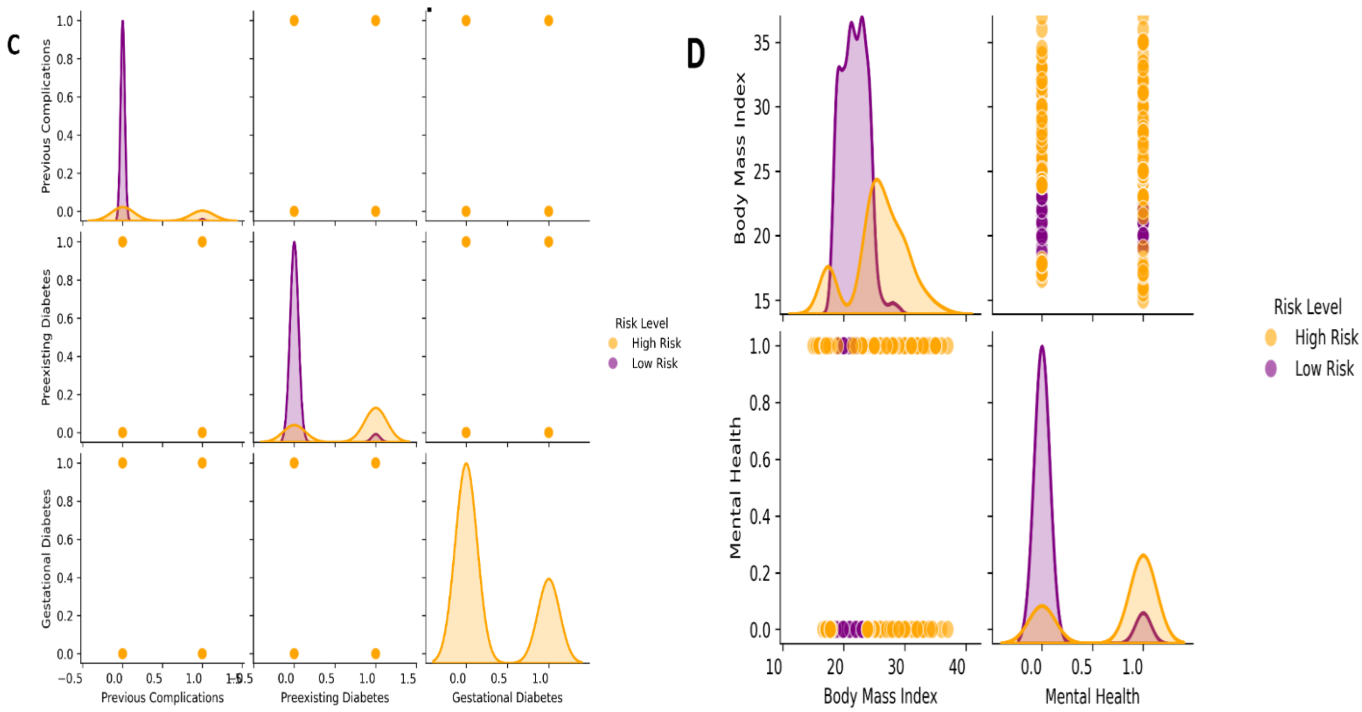


Figure 3: Relationship of risk level with (A) age, blood sugar, and body temperature. (B) systolic BP, diastolic BP, and heart rate. (C) gestational diabetes, preexisting diabetes, and previous complications. (D) body mass index and mental health.

3.3 Machine Learning Model Execution Framework

The dataset was split into two subsets: a training set (75% of the data) and a test set (25%). ML models were trained on the training subset and evaluated on the test subset to assess generalizability. Prior to model development, all features were normalized using Min-Max scaling, which linearly transforms data into a fixed range of 0 to 1. This standardization mitigates biases caused by features with disparate numerical magnitudes, ensuring equitable contributions from all variables during training. The scaling parameters were derived exclusively from the training data, which was both fitted and transformed, while the test data underwent transformation only to prevent data leakage. To identify the optimal classifier, a comparative performance analysis was conducted across models. Model robustness was further strengthened through stratified 10-fold cross-validation, which preserves class distribution in each fold, coupled with a grid search strategy for systematic hyperparameter optimization. An XAI technique, SHAP was utilize to help explain why the model made specific predictions about maternal health risks.

4. RESULTS

Extensive experiments were conducted to evaluate maternal health risk using multiple ML models across varied scenarios. The algorithms' performance was systematically evaluated and compared through metrics including accuracy, precision, recall, F1, AUPRC, and CV scores. Key outcomes of this comparative analysis are summarized in Table 3. Notably, the XGBoost model demonstrated robust predictive performance on the test dataset, achieving an accuracy of 96.36%, with precision and recall values of 96.50% and 96.36%, respectively. The harmonized F1-score of 96.37% underscores the model's balanced capacity to minimize both false positives and false negatives. Notably, the AUPRC reached 99.16%, reflecting exceptional discriminative power in distinguishing between low- and high-risk cases. This performance was achieved with hyperparameters optimized through grid search; *learning_rate: 0.05*, *max_depth: 3*, *n_estimators: 500*, *scale_pos_weight: 1*. The model's consistency is further validated by a CV score of 96.00%, indicating stable generalizability across stratified folds. The findings from the StratifiedKFold-CV of the XGBoost model is provided in Table 4, to showcase position of XGBoost as a highly reliable algorithm for maternal health risk stratification.

Table 3: Comparison of classification matrices of different ML classifiers.

Model	Accuracy	Precision	Recall	F1	AUPRC	Cross-Validation
KNN	0.9769	0.9771	0.9769	0.9769	0.9953	0.9713
SVM	0.9735	0.9735	0.9735	0.9735	0.9962	0.9712
DT	0.9503	0.9512	0.9503	0.9505	0.9578	0.9756
RF	0.9834	0.9841	0.9834	0.9835	0.9982	0.9890
LR	0.9769	0.9769	0.9769	0.9768	0.9963	0.9701
LightGBM	0.9834	0.9841	0.9834	0.9835	0.9998	0.9890
XGBoost	0.9801	0.9806	0.9801	0.9801	0.9988	0.9868
CatBoost	0.9801	0.9806	0.9801	0.9801	0.9971	0.9900
AdaBoost	0.9801	0.9801	0.9801	0.9801	0.9938	0.9890

Table 4: 10-iteration StratifiedKFold cross-validation for the XGBoost classifier.

Fold	Accuracy	Precision	Recall	F1	AUPRC
0	0.989	1.000	0.972	0.986	1.000
1	0.967	1.000	0.917	0.957	0.995
2	0.989	0.973	1.000	0.986	0.998
3	1.000	1.000	1.000	1.000	1.000
4	0.989	1.000	0.972	0.986	1.000
5	0.967	0.946	0.972	0.959	0.995
6	0.978	0.947	1.000	0.973	0.999
7	0.989	1.000	0.972	0.986	1.000
8	1.000	1.000	1.000	1.000	1.000
9	1.000	1.000	1.000	1.000	1.000
Mean	0.987	0.987	0.981	0.983	0.999
Std	0.013	0.023	0.026	0.016	0.002

The confusion matrix, illustrated in Figure 4, provides valuable insights into XGBoost model's performance regarding classification of instances into low risk and high risk. The model demonstrates strong discriminatory performance, achieving 175 true negatives (correctly identified low-risk cases) and 116 true positives (correctly identified high-risk cases). Notably, the low incidence of false negatives (2 cases) underscores the model's sensitivity in detecting high-risk instances, a critical attribute in maternal health contexts where overlooking at-risk patients could lead to adverse outcomes. The 9 false positives indicate a modest rate of over-prediction for high-risk status, which may reflect a cautious approach to risk stratification.

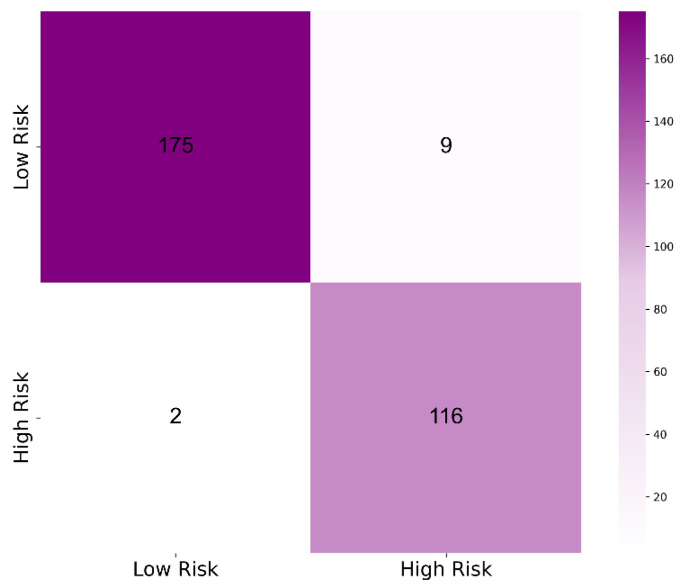
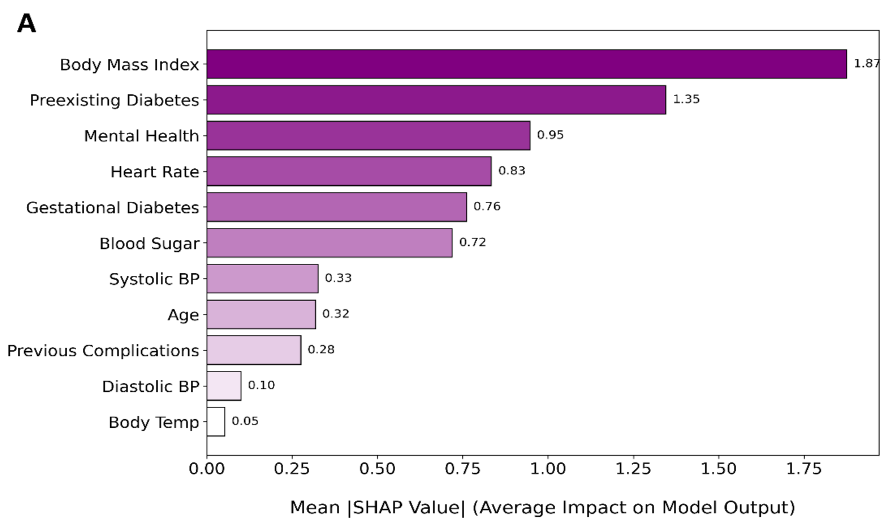
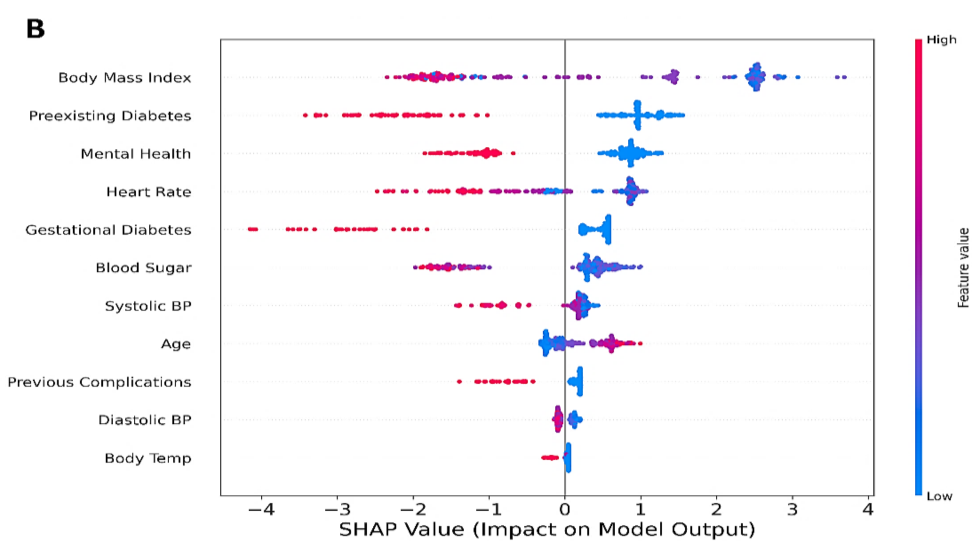


Figure.4: Confusion Matric of XGBoost classifier.

The SHAP analysis of the maternal health risk assessment model as depicted in Figure 5(A-C) revealed critical insights into the factors influencing risk predictions. Figure 5A shows the feature importance plot of XGBoost model. BMI emerged as the most influential feature, with the highest mean absolute SHAP value (1.87), underscoring its pivotal role in maternal health outcomes. This aligns with clinical evidence linking elevated BMI to complications such as gestational diabetes, hypertension, and adverse delivery outcomes. Preexisting diabetes followed closely with SHAP value of 1.35, reflecting its well-documented association with heightened maternal and fetal risks, including macrosomia and preterm birth. Mental health (SHAP: 0.95) and heart rate (SHAP: 0.83) were also significant, suggesting that psychological stress and cardiovascular strain contribute meaningfully to risk stratification. Notably, gestational diabetes (SHAP: 0.76) and blood sugar levels (SHAP: 0.72) reinforced the model's sensitivity to metabolic dysregulation, while systolic blood pressure (SHAP: 0.33) outweighed diastolic measurements (SHAP: 0.10), potentially indicating its stronger correlation with hypertensive disorders like preeclampsia. Age (SHAP: 0.32) and previous complications (SHAP: 0.28) exhibited moderate impacts, possibly reflecting dataset demographics or the model's focus on acute physiological markers over historical or demographic factors. These results validate established clinical risk factors while highlighting underappreciated elements like mental health, advocating for holistic prenatal care frameworks.

Further analysis of risk-specific SHAP summaries, as presented in Figure 5B,C revealed nuanced differences in feature interactions between high- and low-risk cohorts. In high-risk cases, BMI and preexisting diabetes dominated the risk profile, with hierarchical groupings suggesting synergistic effects, for instance, the co-occurrence of elevated BMI and diabetes may amplify cardiovascular or metabolic strain. Mental health and heart rate retained prominence across both cohorts, but their directional impacts diverged: lower feature values (e.g., stable heart rates) were associated with reduced risk in low-risk cases, while deviations correlated with adverse outcomes in high-risk groups. Gestational diabetes and blood sugar exhibited stronger positive SHAP values in high-risk scenarios, emphasizing hyperglycemia's role in exacerbating complications. Conversely, in low-risk profiles, features like age and diastolic blood pressure showed minimal variability, indicating their limited discriminatory power in healthier populations. The model's emphasis on systolic over diastolic blood pressure in high-risk cases aligns with clinical guidelines prioritizing systolic hypertension as a preeclampsia indicator.





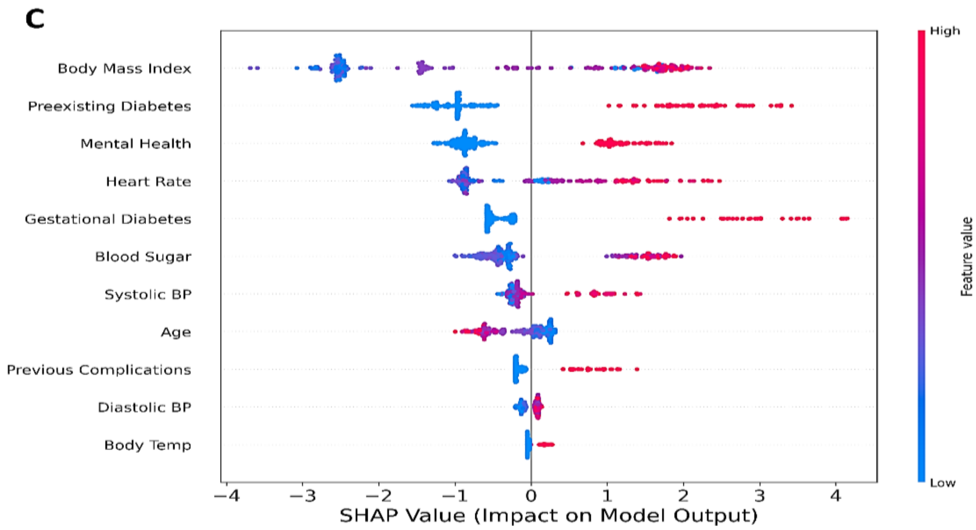


Figure 5: SHAP analysis for XGBoost (A) Feature importance: top contributing features toward prediction, (B) SHAP summary of low risk class, (C) SHAP summary of high risk class.

5. DISCUSSION

The findings of this study underscore the transformative potential of ML models in advancing maternal health risk assessment. This research highlights the applicability of ML algorithms in addressing a critical public health concern i.e. maternal health risk classification. By leveraging the XGBoost model, this study demonstrates how advanced predictive tools can facilitate early detection and intervention, ensuring better health outcomes for mothers and infants [26]. These results are particularly significant in the context of resource-limited settings, where effective risk stratification can guide healthcare providers in prioritizing and delivering targeted care. Importantly, the findings emphasize the role of specific maternal health indicators, such as body mass index, preexisting diabetes, mental health, and blood sugar levels, in assessing risk, offering valuable insights for clinicians and policymakers to refine healthcare strategies and improve maternal outcomes.

In relation to previous studies, the findings of this research corroborate and extend the existing body of knowledge. For example, the importance of body mass index and heart rate as critical predictors of maternal health risks aligns with the conclusions drawn by Noviandy, T. R. *et al.* [27], who also identified similar health metrics as significant, though their proposed LightGBM model reported an accuracy of 84.73%. Similarly, while Mutlu, H. B. *et al.* [28] demonstrated the utility of DT model in classifying maternal health risks with an 89.16% accuracy, this study builds upon their work by employing more advanced algorithms that yield superior performance. Furthermore, the results of the current study align with the work of Ahammad, Md. S. *et al.* [29], who achieved an accuracy of 93.97% using XGBoost for maternal health classification. Table 5 illustrates the performance comparison between the proposed model and existing studies, revealing the superiority of the proposed model. Beyond accuracy, the integration of explainable artificial intelligence (XAI) techniques, such as SHAP values, further distinguishes this study by providing enhanced interpretability. This approach addresses a crucial gap in the field by ensuring that healthcare providers can understand and trust the predictions, a factor that is often lacking in other ML-based studies.

Table 5: Comparison of proposed model with previous studies.

References	Classifier	Accuracy
Noviandy, T. R. et al. [27]	LightGBM	84.73%
Mutlu, H. B. et al. [28]	DT	89.16%
Ahammad, Md. S. et al. [29]	XGBoost	93.97%
Ukrit, M. F. et al. [30]	RF	80%
Maheswari, B. U et al. [31]	XGBoost	89.89%

5.1 Ethical Implication, Deployment and Data Privacy

The ethical implications of deploying ML like XGBoost models in maternal health risk prediction necessitate careful consideration. If the training data lacks diversity, such as underrepresenting certain demographics or regions, the model may yield biased predictions, disproportionately affecting vulnerable populations and exacerbating healthcare disparities. Transparent explainability through techniques like SHAP enhances clinician trust, yet overreliance on algorithmic outputs without human oversight could compromise patient autonomy. Deployment challenges include integrating models into existing healthcare infrastructure, ensuring accessibility in low-resource settings, and training providers to interpret results effectively. Biases may arise from imbalanced datasets or preprocessing choices, such as imputing outliers, which could skew risk assessments for specific groups. Data privacy remains paramount, as health records contain sensitive information; robust anonymization, encryption, and compliance with regulations like HIPAA or GDPR are essential to prevent breaches and protect patient confidentiality. Addressing these concerns holistically ensures equitable, safe, and trustworthy implementation of AI tools in maternal care.

5.2 Implications of Proposed Approach Concerning Traditional Clinical Methods

The proposed ML approach addresses key challenges in traditional clinical methods, which often depend on manual analysis and standardized checklists that may miss subtle patterns in complex patient data. For instance, a gynecologist consulted during this research noted that methods like routine ultrasound or risk scorecards, while foundational, can struggle to consistently detect early signs of complications such as preeclampsia or gestational diabetes. Human fatigue or oversight during lengthy assessments might delay timely interventions. By integrating AI, the model automates data analysis to identify risk factors like shifts in blood pressure or blood sugar trends, that might be overlooked in manual reviews. A maternal health specialist added that AI tools could act as a *"second pair of eyes,"* supporting clinicians in making faster, more informed decisions. However, both professionals stressed that AI should not replace clinical expertise but rather enhance it, ensuring doctors retain final judgment while leveraging data-driven insights to improve care accuracy and efficiency.

6. CONCLUSION AND FUTURE RECOMMENDATION

This study demonstrates the potential of ML in maternal health risk assessment, with the XGBoost model achieving 96.36% accuracy as the top performer in analyzing a comprehensive dataset. The integration of explainable AI techniques, such as SHAP values, revealed BMI and preexisting diabetes as key risk predictors. This research offers significant implications for improving maternal health outcomes, particularly in resource-limited settings, and aligns with global efforts to reduce maternal mortality. Future studies should focus on utilizing larger, more diverse datasets to improve model generalizability and maternal health risk assessment. Exploring hybrid approaches, such as integrating deep learning with ensemble methods, could further enhance predictive accuracy and robustness. Additionally, incorporating longitudinal data and unstructured data (e.g., clinical notes or imaging) alongside structured metrics like body mass index and blood sugar levels may provide a more holistic view of maternal health risk trajectories over time. XAI techniques should be further refined by including domain-specific explanations that align with clinical knowledge, enabling practitioners to understand the underlying factors influencing model predictions.

Declaration of Competing Interests

The authors declares that no conflict of interests is associated with this research work.

Data Availability

The dataset used in this research work is publicly available at data.mendeley.com {Maternal Health Rik Assessment}. The processed data will be shared upon reasonable request.

Funding

The author declare that no funds, grants, or other support were received during the research work.

Author Contribution

M.A. is the sole author of this manuscript and was responsible for all aspects of the research and its presentation. This includes conceptualization, data collection, preprocessing, analysis, interpretation of results, model development, validation, and writing of the manuscript. Additionally, M.A. conducted the literature review, implemented the machine learning models, performed SHAP analysis, and prepared all figures and tables. The author also ensured compliance with ethical guidelines and approved the final version of the manuscript for submission.

References

[1] N. Innab *et al.*, "Automated approach for fetal and maternal health management using light gradient boosting model with SHAP explainable AI," *Frontiers in Public Health,* vol. 12, p. 1462693, 2024.

[2] A. Rahman and M. G. R. Alam, "Explainable AI based Maternal Health Risk Prediction using Machine Learning and Deep Learning," in *2023 IEEE World AI IoT Congress (AIIoT)*, 2023: IEEE, pp. 0013-0018.

[3] M. M. Hosaain, M. A. Kashem, and N. M. Nayan, "Artificial Intelligence-Driven Approach for Predicting Maternal Health Risk Factors," in *2024 9th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, 2024: IEEE, pp. 153-158.

[4] J. Alipour, A. Payandeh, and A. Karimi, "Prevalence of maternal mortality causes based on ICD-MM: a systematic review and meta-analysis," *BMC Pregnancy and Childbirth,* vol. 23, no. 1, p. 821, 2023.

[5] R. Musarandega, M. Nyakura, R. Machekano, R. Pattinson, and S. P. Munjanja, "Causes of maternal mortality in Sub-Saharan Africa: a systematic review of studies published from 2015 to 2020," *Journal of Global Health,* vol. 11, 2021.

[6] F. Di Martino, F. Delmastro, and C. Dolciotti, "Explainable AI for malnutrition risk prediction from m-health and clinical data," *Smart Health,* vol. 30, p. 100429, 2023.

[7] C. W. Oei *et al.*, "Explainable Risk Prediction of Post-Stroke Adverse Mental Outcomes Using Machine Learning Techniques in a Population of 1780 Patients," *Sensors,* vol. 23, no. 18, p. 7946, 2023.

[8] M. Y. Al-Hindi *et al.*, "Association of antenatal risk score with maternal and neonatal mortality and morbidity," *Cureus,* vol. 12, no. 12, 2020.

[9] C. O. Nwokoro *et al.*, "Enhancing Maternal Outcome Prediction Using Explainable Artificial Intelligence for Women of Childbearing Age," *Asian Journal of Science and Applied Technology,* vol. 12, no. 2, pp. 44-55, 2023.

[10] S. Lee *et al.*, "Multiclass Classification by Various Machine Learning Algorithms and Interpretation of the Risk Factors of Pedestrian Accidents Using Explainable AI," *Mathematical Problems in Engineering,* vol. 2023, no. 1, p. 1956865, 2023.

[11] G. Wang, H. Bennamoun, W. H. Kwok, R. Marriott, R. Walker, and J. Kotz, "Codesigning a clinical prediction model for Aboriginal perinatal mental health using Glassbox AI and Aboriginal wisdom and lived experience," in *Health. Innovation. Community: It Starts With Us*: IOS Press, 2024, pp. 196-197.

[12] M. T. Sadiq *et al.*, "Motor imagery EEG signals decoding by multivariate empirical wavelet transform-based framework for robust brain–computer interfaces," *IEEE access,* vol. 7, pp. 171431-171451, 2019.

[13] M. W. Moreira, J. J. Rodrigues, F. H. Carvalho, N. Chilamkurti, J. Al-Muhtadi, and V. Denisov, "Biomedical data analytics in mobile-health environments for high-risk pregnancy outcome prediction," *Journal of Ambient Intelligence and Humanized Computing,* vol. 10, pp. 4121-4134, 2019.

[14] R. D. Jathanna, D. Acharya, L. E. Lewis, and K. Makkithaya, "Early Detection of Late Onset Neonatal Sepsis Using Machine Learning Algorithms," *Engineered Science,* vol. 26, p. 976, 2023.

[15] V. P. Kovacheva, B. W. Eberhard, R. Y. Cohen, M. Maher, R. Saxena, and K. J. Gray, "Preeclampsia prediction using machine learning and polygenic risk scores from clinical and genetic risk factors in early and late pregnancies," *Hypertension,* vol. 81, no. 2, pp. 264-272, 2024.

[16] S. Devi and S. A. Wagle, "Prediction of Hypertensive Disorders in Pregnancy using Machine Learning Algorithms," in *2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2023: IEEE, pp. 1-6.

[17] J. Kaliappan, A. R. Bagepalli, S. Almal, R. Mishra, Y.-C. Hu, and K. Srinivasan, "Impact of Cross-validation on Machine Learning models for early detection of intrauterine fetal demise," *Diagnostics,* vol. 13, no. 10, p. 1692, 2023.

[18] D. S. Bogale, T. M. Abuhay, and B. E. Dejene, "Predicting perinatal mortality based on maternal health status and health insurance service using homogeneous ensemble machine learning methods," *BMC Medical Informatics and Decision Making,* vol. 22, no. 1, p. 341, 2022.

[19] T. O. Togunwa, A. O. Babatunde, and K.-u.-R. Abdullah, "Deep hybrid model for maternal health risk classification in pregnancy: synergy of ANN and random forest," *Frontiers in Artificial Intelligence,* vol. 6, p. 1213436, 2023.

[20] B. S. Kang *et al.*, "Prediction of gestational diabetes mellitus in Asian women using machine learning algorithms," *Scientific Reports,* vol. 13, no. 1, p. 13356, 2023.

[21] M. H. Tilala *et al.*, "Ethical considerations in the use of artificial intelligence and machine learning in health care: a comprehensive review," *Cureus,* vol. 16, no. 6, p. e62443, 2024.

[22] X. Zhang *et al.*, "Prediction of intrahepatic cholestasis of pregnancy in the first 20 weeks of pregnancy," *The Journal of Maternal-Fetal & Neonatal Medicine,* vol. 35, no. 25, pp. 6329-6335, 2022.

[23] B. J. Lengerich, R. Caruana, I. Painter, W. B. Weeks, K. Sitcov, and V. Souter, "Interpretable Machine Learning Predicts Postpartum Hemorrhage with Severe Maternal Morbidity in a Lower Risk Laboring Obstetric Population," *American Journal of Obstetrics & Gynecology MFM,* p. 101391, 2024.

[24] M. Assaduzzaman, A. Al Mamun, and M. Z. Hasan, "Early prediction of maternal health risk factors using machine learning techniques," in *2023 international conference for advancement in technology (ICONAT)*, 2023: IEEE, pp. 1-6.

[25] M. A. Mayen Uddin Mojumdar, Dhiman Sarker, Hasin Arman Shifa, Md. Anisul Haque Sajeeb, Shadikul Bari, Narayan Ranjan Chakraborty, Mohammad Jahangir Alam, "Maternal Health Risk Assessment Dataset," *Mendeley Data,* vol. 1, 2024, doi: 10.17632/p5w98dvbbk.1.

[26] S. A. Mapari *et al.*, "Revolutionizing Maternal Health: The Role of Artificial Intelligence in Enhancing Care and Accessibility," *Cureus,* vol. 16, no. 9, p. e69555, 2024.

[27] T. R. Noviandy, S. I. Nainggolan, R. Raihan, I. Firmansyah, and R. Idroes, "Maternal Health Risk Detection Using Light Gradient Boosting Machine Approach," *Infolitika Journal of Data Science,* vol. 1, no. 2, pp. 48-55, 2023.

[28] H. B. Mutlu, F. Durmaz, N. Yücel, E. Cengil, and M. Yıldırım, "Prediction of maternal health risk with traditional machine learning methods," *Naturengs,* vol. 4, no. 1, pp. 16-23, 2023.

[29] M. S. Ahammad, S. A. Sinthia, M. Hossain, M. M. Ahmed, and M. Ghosh, "Empowering Maternal Health in Bangladesh: Advanced Risk Prediction with Machine Learning," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2024: IEEE, pp. 1-6.

[30] M. F. Ukrit, R. B. Jeyavathana, A. L. Rani, and V. Chandana, "Maternal Health Risk Prediction with Machine Learning Methods," in *2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*, 2024: IEEE, pp. 1-9.

[31] B. U. Maheswari, A. Dixit, and A. K. Karn, "Machine Learning Algorithm for Maternal Health Risk Classification with SMOTE and Explainable AI," in *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, 2024: IEEE, pp. 1-6.