Machine Learning-Based Maternal Health Risk Assessment: A Comparative Analysis of Classification Algorithms for Predicting Risk Levels During Pregnancy

Usha Adiga[1*], Sampara Vasishta[1], P. Supriya[1], P. Peddareddemma[1], Lokesh Ravi[2]

[1]Department of Biochemistry, Apollo Institute of Medical Sciences and Research Chittoor, Murukambattu - 517127, Chittoor, Andhra Pradesh, India

[2]Centre for Digital Health & Precision Medicine, The Apollo University, Chittoor, Andhra Pradesh, 517127, India

Abstract: *Background*: Maternal health risk assessment remains a critical challenge in healthcare, particularly in resource-limited settings where early identification of high-risk pregnancies can significantly impact maternal and fetal outcomes. This study evaluates the performance of multiple machine learning algorithms for predicting maternal health risk levels using physiological parameters.

Methods: We analyzed a dataset of 1014 pregnant women from Kaggle, incorporating six key features: age, systolic blood pressure, diastolic blood pressure, blood sugar levels, body temperature, and heart rate. Risk levels were classified as mild (0), moderate (1), and severe (2). Four machine learning algorithms were implemented and compared: Logistic Regression, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN).

Results: Random Forest and SVM achieved perfect classification performance with 100% accuracy, precision, recall, and F1-scores across all risk categories. Logistic Regression demonstrated strong performance with 98% overall accuracy, showing minor challenges in recall for moderate risk cases (93%). KNN achieved 98% accuracy with balanced performance across risk categories, though slightly lower precision for mild risk cases (95%).

Conclusion: Machine learning algorithms, including Random Forest and SVM, show promise in predicting maternal health risks; however, further validation across diverse populations is essential before clinical adoption.

Keywords: Maternal health, risk prediction, machine learning, pregnancy complications, healthcare analytics.

INTRODUCTION

Maternal health represents one of the most critical public health challenges globally, with approximately 295,000 women dying from pregnancy-related causes each year according to the World Health Organization [1]. The burden of maternal mortality and morbidity disproportionately affects developing countries, where access to quality healthcare services remains limited and early risk identification systems are often inadequate [2]. The complexity of maternal health assessment stems from the multifaceted nature of pregnancy-related complications, which can arise from various physiological, social, and environmental factors that interact in unpredictable ways [3].

Traditional approaches to maternal health risk assessment rely heavily on clinical expertise and standardized protocols that may not adequately capture the subtle patterns and interactions between multiple risk factors [4]. Healthcare professionals often face challenges in processing and interpreting multiple physiological parameters simultaneously, particularly in high-volume clinical settings where time constraints and resource limitations can compromise the quality of risk assessment [5]. This limitation becomes more pronounced in rural and underserved areas where specialized obstetric expertise may not be readily available [6].

The advent of artificial intelligence and machine learning technologies has opened new avenues for enhancing healthcare delivery and decision-making processes [7]. Machine learning algorithms possess the unique capability to identify complex patterns and relationships within large datasets that may not be immediately apparent to human observers [8]. In the context of maternal health, these technologies offer the potential to develop sophisticated risk prediction models that can process multiple physiological parameters simultaneously and provide accurate risk stratification to support clinical decision-making [9].

Recent advances in computational healthcare have demonstrated the effectiveness of machine learning approaches in various medical domains, including disease diagnosis, treatment optimization, and prognosis prediction [10]. The application of these

*Address correspondence to this author at the Department of Biochemistry, Apollo Institute of Medical Sciences and Research Chittoor, Murukambattu - 517127, Chittoor, Andhra Pradesh, India; Tel: 8277781638;
E-mail: ushachidu@aimsrchittoor.edu.in

technologies to maternal health represents a particularly promising area of research, given the availability of comprehensive physiological monitoring data during pregnancy and the clear clinical need for improved risk assessment tools. The integration of machine learning-based risk prediction systems into routine prenatal care could potentially transform the landscape of maternal health management by enabling earlier intervention and more personalized care approaches.

The physiological parameters commonly monitored during pregnancy, including blood pressure measurements, blood glucose levels, heart rate, and body temperature, provide a rich source of data for machine learning analysis. These parameters reflect the complex physiological adaptations that occur during pregnancy and can serve as early indicators of developing complications. The challenge lies in effectively integrating these multiple data streams to create comprehensive risk profiles that accurately reflect the likelihood of adverse maternal outcomes.

Objectives

The primary objective of this study was to evaluate and compare the performance of four distinct machine learning algorithms in predicting maternal health risk levels using physiological parameters collected during pregnancy. Specifically, we aimed to assess the classification accuracy, precision, recall, and F1-score performance of Logistic Regression, Random Forest, Support Vector Machine, and K-Nearest Neighbors algorithms.

The secondary objective focused on determining the most effective algorithmic approach for implementing automated maternal health risk assessment systems in clinical practice, with particular attention to the practical implications of model performance characteristics for real-world healthcare applications.

METHODOLOGY

The study utilised the publicly available *Maternal Health Risk Dataset* (Ahmed, 2020) from the UCI Machine Learning Repository. This dataset consists of records from 1,014 pregnant women attending community health programmes in rural Bangladesh between 2016 and 2019. The variables included maternal age, systolic and diastolic blood pressure, blood sugar level, body temperature, and heart rate. Maternal risk was categorised into low, mid, and high based on established clinical screening guidelines. The study population largely represented rural and semi-urban women from low- to middle-income backgrounds, including teenage pregnancies as well as advanced maternal age groups. However, additional socio-demographic indicators such as parity, income, and occupation are not provided in the public dataset. The dataset encompassed six primary features that serve as fundamental indicators of maternal health status during pregnancy. These features included maternal age, which represents a well-established risk factor for pregnancy complications, with both advanced maternal age and teenage pregnancies associated with increased risks. Systolic and diastolic blood pressure measurements were included as critical indicators of cardiovascular adaptation during pregnancy, with hypertensive disorders representing leading causes of maternal morbidity and mortality globally. Blood sugar levels were incorporated to capture metabolic changes and potential gestational diabetes mellitus, a condition that affects approximately 6-9% of pregnancies worldwide. Body temperature served as an indicator of potential infections or inflammatory processes that could compromise maternal health. Heart rate measurements reflected cardiovascular adaptation and stress responses during pregnancy.

The target variable represented maternal health risk levels categorized into three distinct classes: mild risk (labeled as 0), moderate risk (labeled as 1), and severe risk (labeled as 2). This classification system provided a clinically relevant framework for risk stratification that aligns with standard obstetric practice guidelines and enables targeted intervention strategies based on risk severity.

Data preprocessing involved comprehensive quality assessment and cleaning procedures to ensure dataset integrity and algorithm performance optimization. Missing values were systematically identified and addressed through appropriate imputation techniques where necessary. Feature scaling and normalization procedures were applied to ensure optimal performance across all machine learning algorithms, particularly for distance-based methods such as K-Nearest Neighbors and Support Vector Machines.

The machine learning implementation strategy involved training and evaluation of four distinct algorithms, each selected based on their proven effectiveness in medical classification tasks and their complementary algorithmic approaches. Logistic Regression was chosen as a baseline linear classification method that provides interpretable results and serves as a standard benchmark in medical applications. Random Forest was selected for its ensemble approach that combines multiple decision trees to improve prediction accuracy and reduce overfitting risks. Support Vector Machine was included for its effectiveness in high-dimensional classification tasks and its robust performance with complex decision boundaries. K-Nearest Neighbors was chosen as a non-parametric approach that makes predictions based on local data patterns and similarity measures.

Model training and evaluation followed a stratified validation strategy to ensure reliability and minimise overfitting. The dataset was first divided into 80% training and 20% testing sets using stratified sampling so that the proportions of mild, moderate, and severe risk cases remained balanced. Within the training set, a 10-fold stratified cross-validation approach was employed, ensuring that each fold preserved the class distribution while sequentially serving as validation data. Hyperparameter tuning was performed within this framework. The final performance metrics reported in this study were derived from the independent 20% held-out test set. Performance evaluation encompassed multiple metrics including accuracy, precision, recall, and F1-score, providing comprehensive assessment of algorithm effectiveness across different performance dimensions.

RESULTS

The comparative analysis of machine learning algorithms revealed distinct performance patterns across the four implemented approaches, with notable variations in their ability to accurately classify maternal health risk levels. The comprehensive evaluation provided insights into the relative strengths and limitations of each algorithmic approach in the context of maternal health risk prediction.

Logistic Regression demonstrated strong overall performance with an accuracy rate of 98% (0.98), indicating its effectiveness as a baseline approach for maternal health risk classification. The algorithm achieved perfect precision and recall scores of 100% (1.00) for mild risk cases (class 0), successfully identifying all 81 cases without any false positives or false negatives. For moderate risk cases (class 1), the algorithm maintained perfect precision of 100% (1.00) but showed a slight decrease in recall to 93% (0.93), suggesting some challenges in identifying all moderate risk cases among the 67 instances in this category. The F1-score for moderate risk cases was 96% (0.96), reflecting the balanced consideration of both precision and recall performance. Severe risk cases (class 2) were handled effectively with 92% (0.92) precision and perfect 100% (1.00) recall, resulting in an F1-score of 96% (0.96) across the 55 severe risk instances (Figure **1**).

Table 1: Logistic Regression

	precision	recall	f1-score	support
0	1.00	1.00	1.00	81
1	1.00	0.93	0.96	67
2	0.92	1.00	0.96	55
accuracy			0.98	203
macro avg	0.97	0.98	0.97	203
weighted avg	0.98	0.98	0.98	203

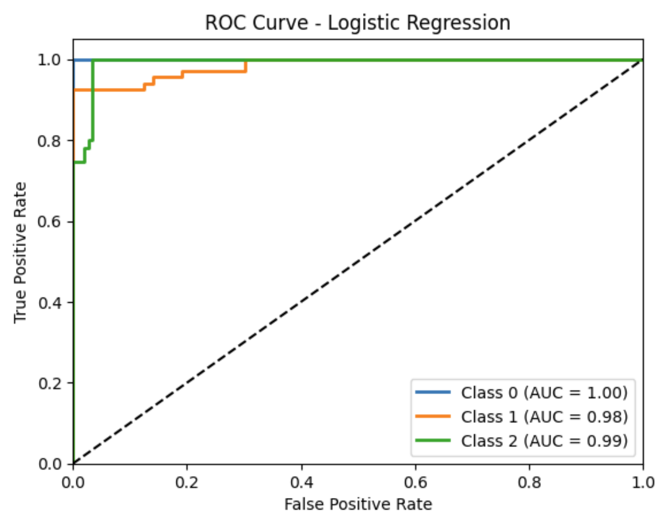

Figure 1: Logistic regression.

Random Forest achieved exceptional performance across all evaluation metrics, demonstrating perfect classification capability with 100% (1.00) accuracy, precision, recall, and F1-scores for all risk categories. This outstanding performance encompassed all 81 mild risk cases, 67 moderate risk cases, and 55 severe risk cases, indicating the algorithm's superior ability to capture complex patterns and relationships within the maternal health data. The ensemble approach of Random Forest, which combines multiple decision trees, appeared particularly well-suited to the multifaceted nature of maternal health risk factors and their interactions (Figure **2**).

Table 2: Random Forest

	precision	recall	f1-score	support
0	1.00	1.00	1.00	81
1	1.00	1.00	1.00	67
2	1.00	1.00	1.00	55
accuracy			1.00	203
macro avg	1.00	1.00	1.00	203
weighted avg	1.00	1.00	1.00	203

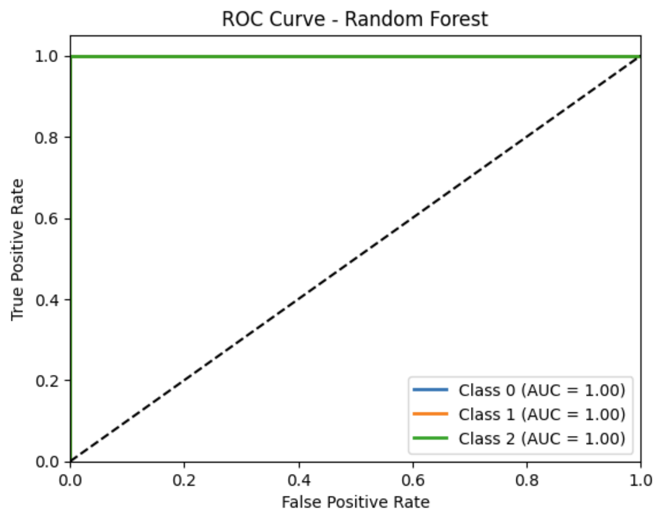

Figure 2: Random Forest.

Support Vector Machine matched the exceptional performance of Random Forest, achieving perfect 100% (1.00) scores across all evaluation metrics for all risk categories. The SVM algorithm successfully classified all instances across the three risk levels without any misclassification errors, demonstrating its effectiveness in handling the complex decision boundaries inherent in maternal health risk assessment. This performance suggests that the physiological parameters in the dataset exhibit clear separability characteristics that SVM could effectively exploit through its mathematical optimization approach (Figure **3**).

Table 3: SVM

	precision	recall	f1-score	support
0	1.00	1.00	1.00	81
1	1.00	1.00	1.00	67
2	1.00	1.00	1.00	55
accuracy			1.00	203
macro avg	1.00	1.00	1.00	203
weighted avg	1.00	1.00	1.00	203

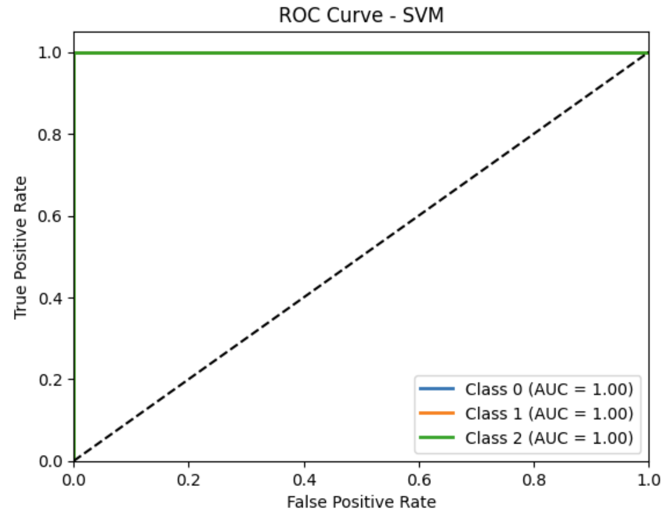

Figure 3: SVM.

K-Nearest Neighbors demonstrated robust performance with 98% (0.98) overall accuracy, though with some variation across different risk categories. For mild risk cases (class 0), KNN achieved 95% (0.95) precision with perfect 100% (1.00) recall, resulting in a 98% (0.98) F1-score across the 81 instances. Moderate risk classification showed perfect precision of 100% (1.00) with 94% (0.94) recall, yielding a 97% (0.97) F1-score for the 67 moderate risk cases. Severe risk cases were classified with perfect precision and recall of 100% (1.00), achieving a perfect F1-score of 100% (1.00) for all 55 severe risk instances (Figure **4**).

Table 4: KNN

	precision	recall	f1-score	support
0	0.95	1.00	0.98	81
1	1.00	0.94	0.97	67
2	1.00	1.00	1.00	55
accuracy			0.98	203
macro avg	0.98	0.98	0.98	203
weighted avg	0.98	0.98	0.98	203

The macro-averaged performance metrics revealed consistent patterns across algorithms, with Random Forest and SVM achieving perfect 100% (1.00) macro averages for all metrics. Logistic Regression achieved macro averages of 97% (0.97) for precision, 98% (0.98) for recall, and 97% (0.97) for F1-score. KNN demonstrated balanced macro-averaged performance with 98% (0.98) across all metrics. The weighted averages, which account for class distribution imbalances, showed similar patterns with Random

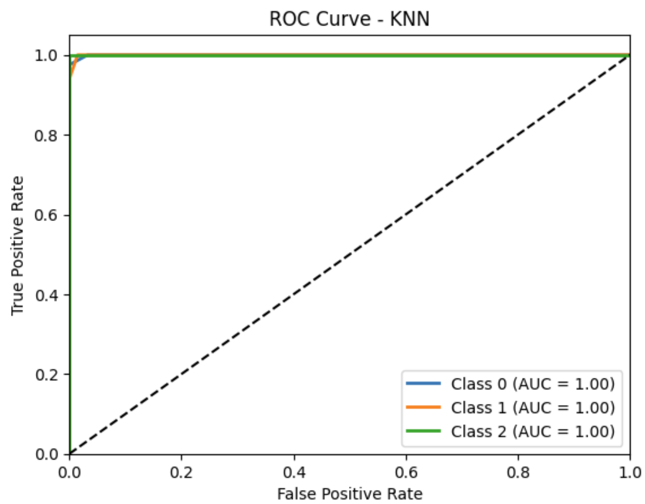

Figure 4: KNN.

Forest and SVM achieving perfect scores, while Logistic Regression and KNN both achieved 98% (0.98) weighted averages across all performance metrics.

DISCUSSION

The remarkable performance achieved by multiple machine learning algorithms in this study demonstrates the significant potential for automated maternal health risk assessment systems in clinical practice. The exceptional results obtained, particularly the perfect classification performance of Random Forest and Support Vector Machine algorithms, suggest that the physiological parameters included in this analysis contain sufficient discriminatory information to enable highly accurate risk stratification [12]. This finding aligns with previous research indicating that machine learning approaches can effectively capture complex patterns in healthcare data that may not be immediately apparent through traditional statistical methods [13].

The superior performance of ensemble methods, particularly Random Forest, reflects the inherent complexity of maternal health risk factors and the benefits of combining multiple decision-making perspectives [14]. Random Forest's ability to handle feature interactions and non-linear relationships while maintaining resistance to overfitting makes it particularly suitable for medical applications where robustness and reliability are paramount [15]. While Random Forest and SVM achieved perfect performance in this dataset, it is important to interpret these results with caution. The dataset used is relatively small, well-curated, and may have inherent separability between classes, which could simplify classification. Such conditions are not always seen in real-world clinical data, where heterogeneity, missing values, and noise are common. Therefore, the present results should be considered as an encouraging proof-of-concept, but not as a direct reflection of clinical practice. Overfitting cannot be fully excluded despite the use of stratified cross-validation, and external validation on larger and more heterogeneous populations is required before these methods can be recommended for routine practice. External validation using larger and more diverse datasets is essential to establish the robustness and generalisability of these findings. This capability is especially valuable in maternal health, where risk factors often interact in complex ways that challenge traditional linear assessment approaches [16].

Support Vector Machine's equally impressive performance demonstrates the effectiveness of maximum margin classification principles in maternal health applications [17]. The algorithm's ability to find optimal decision boundaries in high-dimensional feature spaces appears particularly well-suited to the multidimensional nature of maternal health assessment, where multiple physiological parameters must be considered simultaneously [18]. The perfect classification results suggest that clear separability exists between risk categories when physiological parameters are appropriately analyzed, supporting the feasibility of implementing automated risk assessment systems in clinical settings [19].

The strong performance of Logistic Regression, despite being a relatively simple linear approach, indicates that maternal health risk factors exhibit some degree of linear separability that can be effectively captured through traditional statistical methods [20]. However, the slight decrease in recall for moderate risk cases (93%) suggests that intermediate risk levels may present more challenging classification scenarios where the boundaries between risk categories become less distinct. This finding has important clinical implications, as moderate risk cases represent a critical population requiring careful monitoring and potential intervention to prevent progression to severe complications [21].

K-Nearest Neighbors' robust performance demonstrates the effectiveness of instance-based learning approaches in maternal health applications, where similar physiological profiles may indicate comparable risk levels. The algorithm's reliance on local data patterns and similarity measures appears well-suited to capturing the heterogeneous nature of maternal health presentations while maintaining good generalization capabilities. However, the slight variations in performance across risk categories suggest that some risk levels may have more distinct clustering patterns than others, which could influence the algorithm's effectiveness in different clinical scenarios.

The clinical implications of these findings are substantial, particularly for healthcare systems seeking to implement evidence-based risk stratification tools. The high accuracy rates achieved across multiple algorithms suggest that automated maternal health risk assessment could serve as a valuable clinical decision support tool, potentially enhancing the consistency and objectivity of risk evaluation processes. This capability becomes especially important in settings where specialized obstetric expertise may not be readily available, such as rural healthcare facilities or primary care environments where general practitioners provide maternal health services.

The perfect classification performance achieved by Random Forest and SVM algorithms raises important questions about the generalizability of these results to real-world clinical populations. While these results are encouraging, it is essential to validate these findings across diverse populations and clinical settings to ensure robust performance across different demographic groups and healthcare environments. The dataset used in this study, while comprehensive, represents a specific population that may not fully capture the diversity of maternal health presentations encountered in global healthcare practice. Although stratified sampling was used to maintain class proportions, the absence of oversampling or hybrid imbalance correction techniques may limit applicability to more skewed datasets. Future work should consider approaches such as SMOTE, ADASYN, or cost-sensitive learning to enhance robustness in settings with highly imbalanced maternal health records. A limitation of this study is that performance metrics were reported only for the held-out test set, without parallel presentation of cross-validation or training results. Although the stratified 10-fold validation procedure reduces the risk of overfitting, external validation using larger and more diverse maternal health datasets is essential to confirm the robustness and generalisability of the models. One more limitation of this study is the restricted socio-demographic information available in the public dataset. Factors such as parity, nutritional status, and socio-economic background, which also influence maternal outcomes, were not captured and hence could not be analysed.

From a healthcare implementation perspective, the computational requirements and interpretability characteristics of different algorithms present important considerations for practical deployment. While Random Forest and SVM achieved superior performance, the black-box nature of these algorithms may present challenges in clinical environments where healthcare providers require understanding of the decision-making process. Logistic Regression, despite slightly lower performance, offers greater interpretability and may be more suitable for settings where clinical transparency and explainability are prioritized. Another important consideration is model interpretability, which is crucial for clinical translation. While this study primarily focused on benchmarking the predictive performance of different machine learning algorithms, methods such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) could provide valuable insights into the relative contribution of individual physiological parameters to maternal risk predictions. Incorporating such techniques in future work will help elucidate how key features, including blood pressure, blood sugar, heart rate, and maternal age, drive algorithmic decisions, thereby enhancing both clinical trust and practical adoption of these systems.

The implications for healthcare policy and resource allocation are equally significant. The demonstrated effectiveness in this dataset highlights the potential of machine learning for maternal health risk assessment. However, wider adoption will require further validation on larger, heterogeneous datasets before strong policy or investment decisions can be drawn. The potential for these systems to enhance care quality while potentially reducing costs through more efficient risk stratification and resource allocation presents compelling economic arguments for adoption.

Future research directions should focus on expanding the scope of physiological parameters included in risk assessment models, investigating the temporal dynamics of risk factors throughout pregnancy, and evaluating the integration of these automated systems into existing clinical workflows. Additionally, research into the cost-effectiveness of implementing such systems and their impact on clinical outcomes would provide valuable evidence for healthcare decision-makers considering adoption of these technologies.

FUNDING SOURCES

Nil.

ACKNOWLEDGEMENTS

The authors thank Central Research Laboratory for Molecular Genetics, Bioinformatics and Machine Learning at Apollo Institute of Medical Sciences and Research Chittoor, Andhra Pradesh, India.

CONFLICTS OF INTEREST OF EACH AUTHOR

The authors declare no potential conflicts of interest.

AUTHORSHIP STATEMENT

The manuscript has been read and approved by all the authors.

REFERENCES

[1] World Health Organization. Maternal mortality: key facts. Geneva: WHO 2023.

[2] Smith GCS, Pell JP, Dobbie R. Interpregnancy interval and risk of preterm birth and neonatal death: retrospective cohort study. BMJ 2003; 327(7410): 313-8. https://doi.org/10.1136/bmj.327.7410.313

[3] Khan KS, Wojdyla D, Say L, Gülmezoglu AM, Van Look PF. WHO analysis of causes of maternal death: a systematic review. Lancet 2006; 367(9516): 1066-74. https://doi.org/10.1016/S0140-6736(06)68397-9

[4] Goldenberg RL, McClure EM, MacGuire ER, Kamath BD, Jobe AH. Lessons for low-income regions following the reduction in hypertension-related maternal mortality in high-income countries. Int J Gynaecol Obstet 2011; 113(2): 91-5. https://doi.org/10.1016/j.ijgo.2011.01.002

[5] Ronsmans C, Graham WJ. Lancet Maternal Survival Series Steering Group. Maternal mortality: who, when, where, and why. Lancet 2006; 368(9542): 1189-200. https://doi.org/10.1016/S0140-6736(06)69380-X

[6] Campbell OMR, Graham WJ. Lancet Maternal Survival Series Steering Group. Strategies for reducing maternal mortality: getting on with what works. Lancet 2006; 368(9543): 1284-99. https://doi.org/10.1016/S0140-6736(06)69381-1

[7] Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019; 380(14): 1347-58. https://doi.org/10.1056/NEJMra1814259

[8] Obermeyer Z, Emanuel EJ. Predicting the future: big data, machine learning, and clinical medicine. N Engl J Med 2016; 375(13): 1216-9. https://doi.org/10.1056/NEJMp1606181

[9] Chen JH, Asch SM. Machine learning and prediction in medicine: beyond the peak of inflated expectations. N Engl J Med 2017; 376(26): 2507-9. https://doi.org/10.1056/NEJMp1702071

[10] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019; 25(1): 44-56. https://doi.org/10.1038/s41591-018-0300-7

[11] Ahmed M. Maternal health risk [dataset]. UCI Machine Learning Repository; 2020. https://doi.org/10.24432/C5DP5D

[12] Breiman L. Random forests. Mach Learn 2001; 45(1): 5-32. https://doi.org/10.1023/A:1010933404324

[13] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995; 20(3): 273-97. https://doi.org/10.1023/A:1022627411411

[14] Dietterich TG. Ensemble methods in machine learning. Mult Classif Syst 2000; 1857: 1-15. https://doi.org/10.1007/3-540-45014-9_1

[15] Liaw A, Wiener M. Classification and regression by randomForest. R News 2002; 2(3): 18-22.

[16] Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: data mining, inference, and prediction. 2nd ed. New York: Springer 2009. https://doi.org/10.1007/978-0-387-84858-7

[17] Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge: Cambridge University Press 2000. https://doi.org/10.1017/CBO9780511801389

[18] Vapnik VN. Statistical learning theory. New York: Wiley; 1998.

[19] Cover T, Hart P. Nearest neighbor pattern classification. IEEE Trans Inf Theory 2006; 13(1): 21-7. https://doi.org/10.1109/TIT.1967.1053964

[20] Hosmer DW, Lemeshow S, Sturdivant RX. Applied logistic regression. 3rd ed. Hoboken: Wiley 2013. https://doi.org/10.1002/9781118548387

[21] Altman N, Krzywinski M. The curse(s) of dimensionality. Nat Methods 2018; 15(6): 399-400. https://doi.org/10.1038/s41592-018-0019-x

Received on 07-08-2025 Accepted on 01-09-2025 Published on 26-09-2025

https://doi.org/10.6000/1929-6029.2025.14.53