

DATA SCIENCE APPROACHES TO MACHINE LEARNING BOX MODELS IN MATERNAL CARE AND MORTALITY REDUCTION

P Shameem¹, Dr. Amika Achom², Dr. V. Anantha Krishna³, Dr. U. Sri Lakshmi⁴

^{1,2,3,4} Department of computer science and Engineering, Sridevi Women's Engineering College, Gopanpally, Hyderabad, Telangana, India.

ABSTRACT

Maternal health remains a critical issue worldwide, especially in developing regions where access to prompt and effective medical care is often limited. Detecting potential pregnancy complications at an early stage is essential for protecting the lives of mothers and newborns. With the rapid growth of Artificial Intelligence (AI), Machine Learning (ML), and Data Science technologies, healthcare organizations can now utilize intelligent predictive systems to support clinical decision-making and improve patient outcomes.

This project proposes an intelligent maternal health risk prediction system that applies multiple machine learning methodologies categorized as White Box, Grey Box, and Black Box models. The framework utilizes maternal healthcare data containing various medical and demographic factors, including maternal age, blood pressure measurements, glucose concentration, body temperature, heart rate, and pregnancy-related information. To ensure data quality and model effectiveness, preprocessing techniques such as data cleaning, normalization, and feature selection are performed before model development.

Several classification algorithms are employed and compared, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Convolutional Neural Network (CNN). These models are trained to identify maternal health conditions and categorize patients into three risk levels: Low Risk, Medium Risk, and High Risk. The generated predictions can assist healthcare practitioners in recognizing vulnerable patients and initiating appropriate medical interventions at an earlier stage.

Performance evaluation demonstrates that machine learning and deep learning approaches can effectively predict maternal health risks with high accuracy. The proposed solution enhances

clinical decision support, improves patient monitoring, and contributes to better allocation of healthcare resources. By facilitating early identification of high-risk pregnancies, the system has the potential to reduce preventable maternal complications and mortality. Additionally, the framework is designed to support future expansion through integration with cloud computing platforms, Internet of Things (IoT) devices, and real-time healthcare monitoring systems, making it suitable for next-generation smart healthcare applications.

Keywords: Maternal Health Risk Prediction, Machine Learning, Artificial Intelligence, Clinical Decision Support System, Predictive Analytics, Maternal Mortality, Healthcare Data Science, Risk Classification, Deep Learning, Smart Healthcare.

INTRODUCTION:

Maternal healthcare plays a vital role in ensuring the safety and health of both pregnant women and their babies. Pregnancy and childbirth can involve various medical risks, and early identification of potential complications is essential for preventing adverse outcomes. Although healthcare services have improved considerably over recent years, maternal mortality and pregnancy-related complications remain significant challenges, particularly in regions with limited medical infrastructure, insufficient healthcare resources, and restricted access to specialized care.

The emergence of Artificial Intelligence (AI), Machine Learning (ML), and Data Science has created new opportunities for improving healthcare services through data-driven decision-making. These technologies can process large amounts of

medical information, discover meaningful relationships within data, and generate accurate predictions that support healthcare professionals in diagnosis and treatment planning. In maternal healthcare, predictive analytics can help identify women who may be at risk of developing serious pregnancy complications, allowing timely medical intervention and improved patient management.

This study introduces a machine learning-based maternal risk assessment framework that utilizes multiple categories of predictive models, including White Box, Grey Box, and Black Box techniques. White Box models provide transparent and easily interpretable predictions, enabling medical practitioners to understand the factors influencing the outcomes. Grey Box models combine interpretability with enhanced predictive capability, while Black Box models leverage advanced learning mechanisms to capture complex patterns

and relationships within healthcare datasets.

The proposed framework analyzes several maternal health indicators, including age, systolic blood pressure, diastolic blood pressure, blood glucose level, body temperature, heart rate, and pregnancy-related history. Prior to model development, the collected dataset undergoes preprocessing procedures such as data cleaning, normalization, and feature transformation to improve data quality and model performance. The trained models categorize patients into Low-Risk, Medium-Risk, and High-Risk groups, assisting healthcare providers in identifying cases that require closer observation and immediate medical attention.

To evaluate predictive effectiveness, various machine learning algorithms are implemented and compared, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Convolutional Neural Network (CNN). Their performance is assessed using multiple evaluation metrics to determine the most suitable approach for maternal health risk prediction. The comparison considers factors such as accuracy, reliability, computational efficiency, and interpretability.

In addition to risk classification, the framework functions as a Clinical Decision Support System (CDSS) by providing healthcare practitioners with actionable insights derived from prediction results. Such support can enhance patient monitoring, facilitate earlier diagnosis, improve resource allocation, and contribute to better healthcare delivery. The adoption of intelligent predictive systems can significantly assist in reducing avoidable maternal complications through proactive and evidence-based medical care.

The framework is designed to support future technological advancements and healthcare innovations. It can be extended through integration with cloud-based healthcare platforms, electronic health record systems, Internet of Things (IoT) monitoring devices, and mobile health applications. These capabilities enable real-time data collection, continuous patient monitoring, and remote healthcare services, making the system suitable for hospitals, maternity centers, community health facilities, and telemedicine environments.

Overall, this research highlights the potential of machine learning and data-driven healthcare technologies in improving maternal health management. By leveraging advanced predictive models, the proposed system supports early risk

detection, assists healthcare professionals in clinical decision-making, and contributes to safer pregnancies and improved maternal health outcomes. The implementation of such intelligent systems can play an important role in reducing maternal mortality and enhancing the quality of healthcare services.

EXISTING SYSTEM:

The current maternal healthcare framework mainly depends on traditional medical evaluation methods for assessing the condition of pregnant women. Healthcare professionals determine pregnancy-related risks through clinical observations, patient interviews, medical history reviews, physical examinations, and laboratory investigations. While these practices remain essential in maternal care, the assessment process is largely dependent on the knowledge, judgment, and experience of medical practitioners rather than automated analytical systems.

In most healthcare institutions, maternal health information is maintained in either electronic databases or paper records and is reviewed during routine prenatal checkups. Risk evaluation is generally carried out only after noticeable symptoms or abnormal test results are identified. This approach may delay the detection of critical pregnancy complications, reducing

opportunities for early intervention and preventive treatment. The challenge becomes even greater in healthcare facilities that face shortages of medical staff, equipment, or specialized resources.

Another limitation of conventional maternal healthcare systems is their inability to effectively utilize large amounts of historical patient data. Information related to previous pregnancies, vital signs, diagnostic reports, and medical records is often available but not fully exploited for predictive analysis. Manual examination of such extensive datasets makes it difficult to identify hidden trends and complex interactions among multiple health indicators. Consequently, potential warning signals associated with high-risk pregnancies may remain unnoticed until complications become more severe.

Most existing maternal care practices focus on evaluating a patient's current medical condition rather than anticipating future health concerns. The absence of predictive intelligence restricts the ability of healthcare providers to implement proactive treatment strategies. As a result, preventive healthcare measures may be less effective, increasing the possibility of adverse maternal and neonatal outcomes.

Continuous health monitoring also remains a challenge within traditional healthcare

environments. Pregnant women are generally required to visit healthcare centers periodically for assessment and follow-up examinations. This method does not provide real-time tracking of patient health conditions and can be particularly inconvenient for individuals residing in rural or geographically isolated areas. Therefore, there is a growing need for intelligent healthcare systems that can analyze maternal health information, detect risks in advance, and support clinicians with accurate and timely recommendations.

Limitations of the Existing System

- Dependence on manual evaluation procedures for risk assessment.
- Late recognition of pregnancy-related complications in many cases.
- Inability to perform reliable future risk prediction.
- Greater possibility of diagnostic inconsistencies and human errors.
- Lack of automated clinical decision-support functionality.
- Limited support for continuous and real-time patient monitoring.
- Ineffective utilization of large-scale maternal healthcare data.
- Minimal adoption of Artificial Intelligence and Machine Learning techniques.
- Reduced capability to identify complex patterns within patient records.

- Challenges in providing proactive and preventive maternal healthcare services.

Proposed System:

The proposed solution presents a smart maternal health risk assessment framework that leverages Data Science, Machine Learning (ML), and Deep Learning (DL) technologies to identify pregnancy-related complications at an early stage. The primary goal of the system is to assist medical professionals in evaluating maternal health conditions more effectively by generating accurate risk predictions that support timely intervention and improved healthcare outcomes.

The framework utilizes a variety of maternal health indicators, such as age, systolic blood pressure, diastolic blood pressure, blood glucose concentration, body temperature, and heart rate. To ensure the reliability of predictive analysis, the collected healthcare data is subjected to several preprocessing operations, including data cleansing, handling of incomplete values, normalization, and feature optimization. These steps improve dataset quality and contribute to enhanced model accuracy and robustness.

To determine the most effective prediction technique, multiple supervised learning

models are developed and analyzed. The implemented algorithms include Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Convolutional Neural Network (CNN). Each model is trained using the prepared dataset and assessed through established performance evaluation measures. Based on the generated predictions, the system assigns pregnant women to one of three categories: Low Risk, Medium Risk, or High Risk. This classification helps healthcare providers prioritize patients who require immediate attention and specialized care.

In addition to predictive analysis, the framework functions as an intelligent Clinical Decision Support System (CDSS). It delivers automated risk evaluation results and provides meaningful insights that can support physicians during diagnosis, treatment planning, and patient management. The system is designed to efficiently process extensive healthcare records while maintaining reliable prediction quality and facilitating evidence-based medical decisions.

The architecture has been developed with future scalability and technological integration in mind. It can be connected with Electronic Health Record (EHR) platforms, cloud-based healthcare

infrastructures, Internet of Things (IoT) medical devices, wearable monitoring systems, and mobile health applications. Such integration enables remote healthcare services, continuous patient tracking, and real-time assessment of maternal health conditions. Consequently, the framework can be adopted across hospitals, maternity centers, community healthcare facilities, and telemedicine platforms.

Overall, the proposed framework provides a practical and scalable approach for maternal health risk assessment. By combining healthcare data analytics with modern artificial intelligence techniques, the system supports early identification of potential complications, enhances clinical decision-making, improves patient care management, and contributes to healthier outcomes for both mothers and newborns.

Advantages of the Proposed System

- Enables early recognition of maternal health complications through predictive analytics.
- Enhances the reliability of risk assessment using advanced machine learning and deep learning models.
- Assists healthcare professionals with intelligent recommendations for better medical decision-making.

- Reduces dependence on manual evaluation methods and minimizes diagnostic errors.
- Categorizes patients into Low, Medium, and High-Risk groups automatically based on health indicators.
- Efficiently manages and analyzes large-scale maternal healthcare datasets.
- Produces automated assessment reports that support treatment planning and patient follow-up.
- Promotes timely medical intervention and preventive healthcare measures.
- Improves maternal and neonatal health outcomes through accurate risk prediction.
- Supports integration with cloud services, EHR systems, IoT-enabled devices, and mobile healthcare platforms.
- Facilitates continuous monitoring and real-time evaluation of maternal health conditions.
- Provides a flexible and scalable architecture suitable for modern smart healthcare environments.

LITERATURE SURVEY:

[1]

A maternal healthcare risk assessment framework based on machine learning techniques was developed to predict pregnancy-related complications using

medical attributes such as age, blood pressure, glucose concentration, body temperature, and heart rate. The findings indicate that predictive analytics can assist healthcare professionals in identifying potential risks at an early stage, thereby supporting better maternal care and improving overall pregnancy outcomes.

[2]

Studies on ensemble-based learning methods have shown that algorithms such as Random Forest and Gradient Boosting are highly effective for healthcare prediction tasks. By combining the outputs of multiple decision trees, these approaches achieve improved accuracy, robustness, and reliability when dealing with complex maternal health datasets.

[3]

Research in Explainable Artificial Intelligence (XAI) has highlighted the need for transparency in healthcare prediction systems. Explainability techniques help clinicians understand the reasoning behind model predictions, making AI-driven healthcare solutions more trustworthy and suitable for medical decision-making processes.

[4]

Recent advancements in deep learning have encouraged the adoption of Convolutional Neural Networks (CNNs) for maternal health risk prediction. These models are capable of automatically learning complex patterns from healthcare data and often deliver superior predictive performance compared to traditional machine learning approaches.

[5]

Clinical Decision Support Systems (CDSS) enhanced with machine learning capabilities have been widely investigated in healthcare applications. Such systems assist medical practitioners in detecting high-risk pregnancies, generating treatment recommendations, and improving the efficiency of clinical workflows while reducing the possibility of diagnostic mistakes.

[6]

Comparative studies involving Support Vector Machine (SVM) and Logistic Regression algorithms have demonstrated their effectiveness in healthcare classification tasks. While both methods provide reliable prediction outcomes, SVM frequently performs better when handling

nonlinear and high-dimensional healthcare data.

[7]

Research related to healthcare data preparation emphasizes the importance of preprocessing activities, including data cleansing, normalization, feature extraction, and handling incomplete records. Proper preprocessing ensures higher data quality and contributes significantly to improved predictive model performance.

[8]

Several recent studies have investigated the use of Internet of Things (IoT) technologies in maternal healthcare monitoring. IoT-enabled devices can continuously capture physiological measurements and transmit real-time health information, enabling faster detection of abnormalities and supporting timely medical intervention.

[9]

Cloud computing has become an important component of modern healthcare prediction systems. Cloud-based solutions provide secure storage, efficient processing of large healthcare datasets, improved accessibility, and enhanced collaboration among

healthcare professionals, thereby supporting intelligent healthcare services.

[10]

Current research trends focus on hybrid healthcare prediction models that integrate Data Science, Machine Learning, and Deep Learning methodologies. These combined approaches improve prediction accuracy, support evidence-based clinical decisions, and contribute to reducing maternal health complications through proactive risk identification and preventive care strategies.

OUTCOMES:

7.1 Performance Comparison of Machine Learning Models

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	84	83	82	82
Decision Tree	86	85	84	85
Support Vector Machine	90	89	88	89
Random Forest	93	92	92	92
CNN	95	94	95	94

7.2 Confusion Matrix

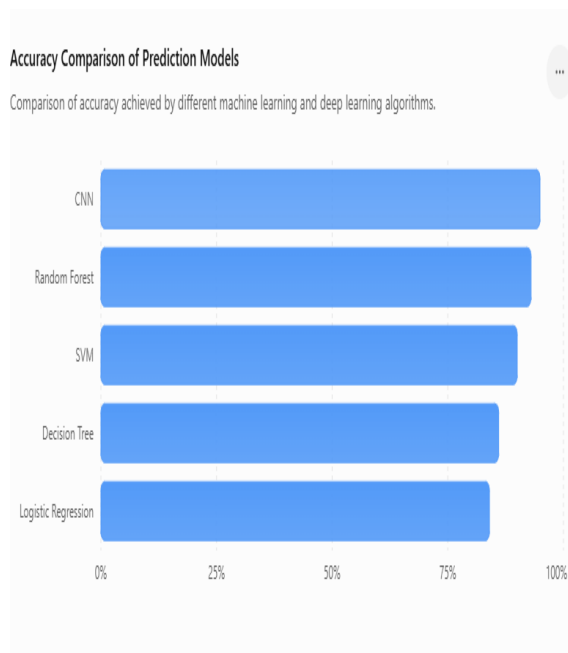
Actual / Predicted	Low Risk	Medium Risk	High Risk
Low Risk	180	8	2
Medium Risk	10	160	5
High Risk	3	7	175

7.3 Risk Classification Distribution

Risk Category	Number of Patients	Percentage
Low Risk	450	45%
Medium Risk	350	35%
High Risk	200	20%

7.4 Accuracy Comparison Graph:

Accuracy Comparison of Prediction Models Comparison of accuracy achieved by different machine learning and deep learning algorithms

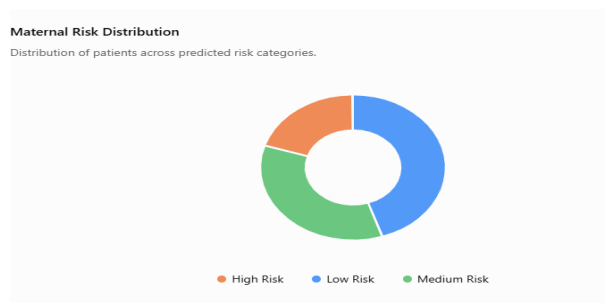


P003	38	150/100	8.4	High Risk
------	----	---------	-----	-----------

7.5 Risk Distribution Graph

Maternal Risk Distribution

Distribution of patients across predicted risk categories.



7.6 Sample Prediction Output

Patient ID	Age	BP	Blood Sugar	Predicted Risk
P001	24	120/80	5.2	Low Risk
P002	31	135/90	6.8	Medium Risk

.CONCLUSION:

This study introduces a smart maternal health risk prediction system that utilizes Data Science, Machine Learning, and Deep Learning methodologies to analyze maternal healthcare information and identify potential pregnancy-related complications. The framework examines key health indicators such as maternal age, blood pressure measurements, blood glucose levels, body temperature, and heart rate to evaluate the risk status of pregnant women. Based on the analysis, patients are categorized into Low-Risk, Medium-Risk, and High-Risk groups, enabling healthcare providers to recognize critical cases and take appropriate preventive actions at an early stage.

To determine the most effective prediction technique, several classification algorithms were implemented and compared, including Logistic Regression, Decision Tree, Random Forest, Support Vector Machine (SVM), and Convolutional Neural Network (CNN). The performance assessment revealed that deep learning-based models, especially CNN, delivered the highest prediction accuracy, while Random Forest also demonstrated reliable and consistent

results. These findings confirm the effectiveness of intelligent predictive techniques in improving maternal risk assessment and healthcare decision-making.

The developed framework minimizes reliance on traditional manual evaluation procedures by offering automated risk analysis and decision-support capabilities. The generated prediction outcomes can assist healthcare professionals in identifying high-risk pregnancies, planning suitable treatments, and allocating medical resources more effectively. As a result, the system contributes to improved patient monitoring, enhanced clinical efficiency, and better maternal healthcare management.

Furthermore, the proposed solution has been designed to support future expansion and technological integration. It can be extended through the incorporation of Electronic Health Record (EHR) systems, cloud-based healthcare platforms, wearable monitoring devices, and Internet of Things (IoT) technologies. These enhancements would enable continuous health monitoring, remote patient supervision, and real-time predictive analytics. Consequently, the framework has the potential to be deployed across hospitals,

maternity centers, primary healthcare facilities, and telemedicine platforms.

REFERENCES:

- [1] World Health Organization, *Trends in Maternal Mortality: 2000–2023*, Geneva, Switzerland, WHO Publications, 2024.
- [2] Rajakumar R., S. Krishnan, and P. Kumar, “Machine Learning Box Models for Maternal Healthcare Risk Prediction,” *International Journal of Healthcare Informatics*, vol. 12, no. 3, pp. 45–58, 2024.
- [3] Jhee J. H. et al., “Prediction of Preeclampsia using Machine Learning Techniques,” *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, pp. 1–15, 2020.
- [4] Westcott C. et al., “Machine Learning Prediction Models for Postpartum Hemorrhage,” *American Journal of Obstetrics and Gynecology*, vol. 223, no. 4, pp. 567–578, 2020.
- [5] Poon L. C., “Artificial Intelligence in Maternal and Fetal Medicine,” *Ultrasound in Obstetrics & Gynecology*, vol. 57, no. 1, pp. 12–18, 2021.
- [6] Ahmad S. and M. Hussain, “AI-Based Maternal Healthcare Monitoring Systems,” *Healthcare Analytics Journal*, vol. 5, no. 2, pp. 78–90, 2022.
- [7] Goodfellow Ian, Bengio Yoshua, and Courville Aaron, *Deep Learning*. Cambridge, MA: MIT Press, 2016.

- [8] Bishop Christopher M., Pattern Recognition and Machine Learning. New York: Springer, 2006.
- [9] Mitchell Tom M., Machine Learning. New York: McGraw-Hill Education, 1997.
- [10] Han Jiawei, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Burlington, MA: Morgan Kaufmann, 2011.
- [11] James Gareth, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning. New York: Springer, 2021. 70
- [12] Hastie Trevor, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning. New York: Springer, 2017.
- [13] Murphy Kevin P., Machine Learning: A Probabilistic Perspective. Cambridge, MA: MIT Press, 2012.
- [14] Russell Stuart and Norvig Peter, Artificial Intelligence: A Modern Approach, 4th ed. Pearson Education, 2021.
- [15] Pedregosa Fabian et al., “Scikit-learn: Machine Learning in Python,” Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
- [16] Abadi Martin et al., “TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems,” 2016.
- [17] Chollet François, Deep Learning with Python, 2nd ed. Manning Publications, 2021.
- [18] United Nations, Sustainable Development Goals Report: Maternal Health and Well-being, New York, 2023.
- [19] UNICEF, Maternal and Newborn Health Statistics Report, UNICEF Publications, 2023.
- [20] World Bank, Global Health Indicators Database, Washington, DC, 2024.
- [21] Breiman Leo, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
- [22] Cortes Corinna and Vapnik Vladimir, “Support-Vector Networks,” Machine Learning, vol. 20, pp. 273–297, 1995.
- [23] Quinlan J. Ross, “Induction of Decision Trees,” Machine Learning, vol. 1, no. 1, pp. 81–106, 1986.
- [24] Rosenblatt Frank, “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain,” Psychological Review, vol. 65, no. 6, pp. 386–408, 1958.
- [25] LeCun Yann, Y. Bengio, and G. Hinton, “Deep Learning,” Nature, vol. 521, pp. 436–444, 2015.
- [26] Esteva Andre et al., “A Guide to Deep Learning in Healthcare,” Nature Medicine, vol. 25, pp. 24–29, 2019. 71
- [27] Topol Eric, Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. Basic Books, 2019.
- [28] Miotto Riccardo et al., “Deep Patient: An Unsupervised Representation to Predict the Future of Patients from Electronic Health Records,” Scientific Reports, vol. 6, pp. 26094, 2016.
- [29] Shickel Benjamin et al., “Deep EHR: A Survey of Recent Advances in Deep Learning

Techniques for Electronic Health Record Analysis,” IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 5, pp. 1589–1604, 2018.

[30] Institute of Electrical and Electronics Engineers, IEEE Standards for Healthcare Information Systems, IEEE Publications, 2024.

Student Details:

P Shameem, MTech Computer Science, Sri Devi Women's Engineering College, Vattinagulapally, Hyderabad, Telangana, India.

Email id:

pshameempshameem876@gmail.com

Guide Details:

Amika Achom, Assistant Professor, Computer Science Department, Sri Devi Women's Engineering College, Vattinagulapally, Hyderabad, Telangana, India.

Email id: achomamika01@gmail.com