iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83884

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

AI-Based Maternal Health Risk Prediction System Using Machine Learning

Namitha M[1], Niriksha G Y[2], R Kanika[3], Sonu N[4], Minugu B[5]
[1, 2, 3, 4]*Department of Electronics and Communication Engineering, P.E.S. College of Engineering, Mandya, Karnataka, India*
[5]*Assistant Professor, Department of Electronics and Communication Engineering, P.E.S. College of Engineering, Mandya, Karnataka, India*

Abstract: ***Maternal mortality continues to be a public health challenge in developing countries, especially in rural and semi-urban areas where regular prenatal monitoring is not available. This paper presents an Artificial Intelligence (AI) based system for prediction of maternal health risk using machine learning techniques which classifies pregnancy risk levels as Low, Moderate or High. The system uses a real-world clinical dataset of 6103 patient records with 8 important physiological parameters: age, body temperature, heart rate, systolic and diastolic blood pressure, BMI and blood glucose levels. The data was trained and evaluated on Random Forest, XGBoost, Gradient Boost and a Voting Ensemble classifier. The XGBoost model performed the best with 99.26% classification accuracy, and 0.99 precision, recall and F1-scores for all three risk categories The system is implemented as a full-stack web application, using React.js for the frontend, FastAPI for the backend, and Supabase as a cloud-based PostgreSQL database, providing real-time risk assessment and accessible healthcare monitoring for pregnant women.***

Keywords: Machine Learning, Maternal Health, Risk Prediction, XGBoost, Random Forest, Healthcare AI, FastAPI, Pregnancy Monitoring

I. INTRODUCTION

Maternal health is the health of women during pregnancy, childbirth and the puerperium. Global health statistics reveal that hundreds of thousands of women die each year from preventable complications during pregnancy and childbirth, most of them in low- and middle-income countries. Maternal mortality remains a big problem in India, especially in rural areas where there is less access to qualified health care personnel and diagnostic facilities.

Early detection of health risks in pregnancy is a major contributor to reducing maternal and infant mortality. Potential complications such as preeclampsia, gestational diabetes and other hypertensive disorders of pregnancy are well-known predictors of clinical parameters such as blood pressure, blood glucose levels, body temperature, heart rate and body mass index (BMI). However, in resource-poor settings, constant monitoring of these parameters is often not possible due to financial constraints, lack of awareness and inadequate healthcare infrastructure.

Recent advances in Artificial Intelligence and Machine Learning enable the development of intelligent decision-support systems to analyze clinical data and to predict potential health risks with high accuracy . These systems can help health professionals make faster data-driven decisions, and help patients monitor their own health remotely.In this paper, we present an AI-Based Maternal Health Risk Prediction System. Pregnant women can enter their daily health vitals through a web-based interface and get an instant AI-based risk classification. The system stratifies risk into Low, Moderate and High tiers and offers appropriate care recommendations. The model is trained on a real clinical dataset and deployed using a modern full-stack architecture, making it scalable and device agnostic.

II. LITERATURE SURVEY

[1]S. N. Khorshid, M. J. Reddy, and P. Gupta, “Machine Learning Based Prediction of Gestational Diabetes Mellitus,”International Journal of Medical Informatics, vol. 112, pp. 1– 8, 2018. This paper presents a machine learning-based approach for early prediction of Gestational Diabetes Mellitus (GDM). Clinical and demographic parameters such as age, BMI, fasting glucose, blood pressure, and insulin levels are used for analysis. Data preprocessing techniques like normalization and missing value handling are applied to improve prediction accuracy. Logistic Regression and Decision Tree algorithms are implemented to classify pregnant women into GDM and non-GDM categories. The results show improved prediction performance compared to conventional diagnosis. However, the study uses static datasets and does not support continuous digital health monitoring.

©IJRASET: All Rights are Reserved

[2]A. Kumar and R. Patel, "Digital Health Monitoring System for Diabetes Management," International Journal of Engineering and Technology, vol. 7, no. 3, pp. 245–250, 2019. This paper discusses the role of digital health platforms in diabetes monitoring using mobile health applications and electronic health records. Patients manually input glucose values, which are stored and analyzed over time. Doctors can remotely monitor patient health and provide recommendations. The system improves accessibility and reduces frequent hospital visits. However, it lacks intelligent prediction mechanisms and does not utilize machine learning algorithms for automated risk assessment.

[3]J. Brown, L. Smith, and K. Taylor, "Application of Machine Learning Algorithms for Blood Glucose Level Prediction,"IEEE Access, vol. 8, pp. 123456–123465, 2020. This paper focuses on predicting blood glucose levels using machine learning algorithms such as Support Vector Machine (SVM), Random Forest, and Artificial Neural Networks. Feature selection techniques are applied to identify significant attributes influencing glucose levels. The results demonstrate that Random Forest provides higher accuracy for nonlinear glucose data. Although the system shows strong predictive performance, it is designed for general diabetes and does not specifically address gestational diabetes or pregnancy-related factors.

[4]P. Ramesh and S. Meena, "IoT and Machine Learning Based Health Monitoring System for Pregnant Women,"International Journal of Advanced Research in Electronics and Communication Engineering, vol. 9, no. 5, pp. 412–417, 2021. This paper proposes an IoT-enabled health monitoring system integrated with machine learning for pregnant women. Sensors are used to measure physiological parameters such as blood glucose, heart rate, and blood pressure. The data is transmitted to a cloud platform for analysis and prediction. The system enables real-time monitoring and early alerts for abnormal glucose levels. However, it requires additional hardware, stable Internet connectivity, and higher power consumption, increasing system cost and complexity.

[5]R. Das and M. Chatterjee, "Data-Driven Decision Support System for Gestational Diabetes Management,"Journal of Healthcare Engineering, vol. 2022, pp. 1–10, 2022. This paper presents a machine learning-based decision support system for managing gestational diabetes. Pregnancy-related health data is analyzed to predict glucose abnormalities and assist doctors in treatment planning. The system improves personalized healthcare and clinical decisionmaking using performance metrics such as accuracy and recall. However, it requires large datasets for training and lacks extensive real-world clinical validation.

[5]M. Alsharif and N. Alotaibi, "Digital Health and Machine Learning Framework for Maternal Health Monitoring,"IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 4, pp. 1765–1774, 2022.

This paper presents a comprehensive digital health framework combined with machine learning techniques for maternal health monitoring, including gestational diabetes management. The framework integrates electronic health records, wearable sensor data, and machine learning models to analyze blood glucose trends and predict potential risks. The system supports early detection, continuous monitoring, and personalized healthcare recommendations. Experimental results show improved prediction accuracy and better patient outcomes. However, the implementation requires secure data handling and integration with hospital information systems, which increases system complexity.

III. PROPOSED METHODOLOGY

A. Dataset

The dataset used in this study was obtained from a real-world clinical maternal health dataset published on Mendeley Data [1] containing 6,103 patient records. Each record contains 8 physiological parameters: Age, Body Temperature, Heart Rate, Systolic Blood Pressure, Diastolic Blood Pressure, BMI, Blood Glucose (HbA1c), Blood Glucose (Fasting). Each record was classified into one of three risk categories (low risk, mid risk and high risk) based on clinical assessment.

B. Data Preprocessing

No missing values were found in the dataset after checking for missing values and inconsistencies. The target variable (risk status) was encoded using label encoding to transform the categorical labels into numerical format suitable for model training. The dataset was split into training (80%) and testing (20%) subsets using stratified sampling to maintain the class distribution between the two datasets.

C. Model Selection and Training

Four machine learning models were trained and compared, as follows:

1) Random Forest Classifier: Ensemble of 200 decision trees with max depth 15 and square root of total features at each split to reduce correlation between trees.

2) XGBoost Classifier: Gradient Boosted Tree Ensemble with 200 estimators, max depth of 8 and learning rate of 0.1, with subsampling and column subsampling of 0.8 to avoid overfitting
3) Gradient Boosting Classifier: Set with 200 estimators, 8 max depth and learning rate 0.1.
4) Voting Ensemble Classifier: Ensemble model that combines the predictions of Random Forest, XGBoost and Gradient Boosting models using soft voting.

All models were trained in Python using the scikit-learn and the XGBoost packages. The model performance was evaluated using accuracy, precision, recall and F1-score on the held-out test set and five-fold cross-validation was performed to verify robustness.

D. System Architecture

The proposed system is based on a four-layer architecture:

1) Frontend Layer: Developed using React.js with Vite as the build tool, it provides screens for user registration, daily vitals input, AI risk result display and historical assessment tracking.
2) Backend Layer: Built on FastAPI, a high-performance Python web framework, provides REST API endpoints for patient registration, vitals submission, risk prediction, and history retrieval.
3) Machine Learning Layer: The trained XGBoost model is serialized using Python's pickle module and loaded by the backend at runtime to make real-time predictions.
4) Data Storage Layer: Supabase, a cloud-based PostgreSQL database, is used to store patient profiles, submitted vitals, and risk assessment records with Row-Level Security enabled to ensure data privacy.

The entire system is deployed with the frontend on Vercel and the backend on Render and can be accessed from any device with internet access.

IV. RESULTS AND DISCUSSION

The four trained models were evaluated on a held-out test set of 1,221 patient records. Table I summarizes the comparative accuracy of each model.

TABLE I MODEL PERFORMANCE COMPARISON

Model	Accuracy
Random Forest	99.18%
XGBoost	99.26%
Gradient Boosting	99.26%
Voting Ensemble	99.26%

The XGBoost classifier was selected as the final deployment model due to its highest accuracy combined with computational efficiency. Table II presents the detailed classification report for the XGBoost model.

TABLE II CLASSIFICATION REPORT (XGBOOST)

Risk Class	Precision	Recall	F1-Score	Support
High Risk	1.00	0.99	0.99	412
Low Risk	1.00	1.00	1.00	400
Mid Risk	0.99	0.99	0.99	409

In five-fold cross validation, the model achieved an average accuracy of 99.23% (ranging from 99.10% to 99.34% for individual folds), indicating a good generalization and no overfitting on the training data.

The feature importance analysis indicated that Fasting Blood Glucose (31.2%), Systolic Blood Pressure (19.8%) and HbA1c (18.7%) were the most important predictors of maternal risk. This is consistent with the clinical understanding that gestational diabetes and hypertensive disorders are the main causes of pregnancy complications.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The whole system was tested in an end-to-end manner with real patient records. Patients with normal vital signs (e.g. blood pressure 136/87 mmHg, glucose 4.4 mmol/L) were correctly identified as low risk and patients with elevated blood pressure (161/100 mmHg) and low hemoglobin were correctly identified as high risk with appropriate care recommendations produced by the system in each case.

V. CONCLUSION AND FUTURE WORK

We present an AI-based maternal health risk prediction system that integrates a high accuracy classifier (XGBoost) with a fully deployed full-stack web application. The system was trained on a real clinical dataset of 6,103 patient records and reached 99.26% classification accuracy with good precision, recall and F1-scores for all risk categories. The deployed system provides real-time accessible risk assessment for pregnant women and is particularly relevant for underserved rural healthcare settings. Future work will involve the integration of IoT-based sensors for automated vital sign acquisition, expanding the dataset through collaboration with local healthcare institutions, developing a dedicated mobile application and including a healthcare provider dashboard for remote patient monitoring. We are also working on multilingual support to improve accessibility for non-English speaking users in rural communities.

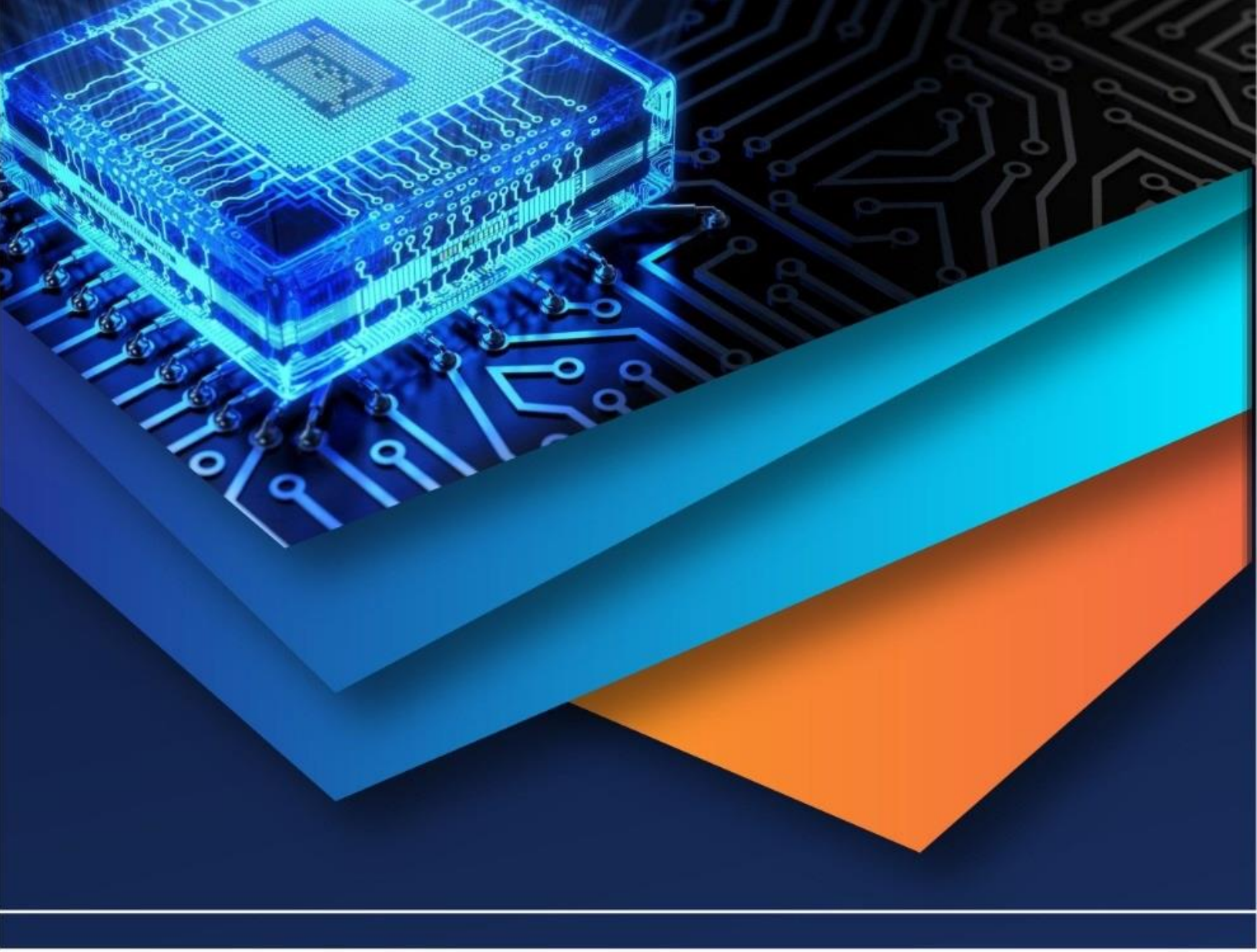
VI. ACKNOWLEDGMENT

The authors would like to thank the Department of Electronics and Communication Engineering, P.E.S. College of Engineering, Mandya, for their support and guidance throughout this project.

REFERENCES

[1] S. N. Khorshid, M. J. Reddy, and P. Gupta, “Machine Learning Based Prediction of Gestational Diabetes Mellitus,”International Journal of Medical Informatics, vol. 112, pp. 1– 8, 2018.
[2] Kumar and R. Patel, “Digital Health Monitoring System for Diabetes Management,” International Journal of Engineering and Technology, vol. 7, no. 3, pp. 245–250, 2019.
[3] J. Brown, L. Smith, and K. Taylor, “Application of Machine Learning Algorithms for Blood Glucose Level Prediction,”IEEE Access, vol. 8, pp. 123456–123465, 2020.
[4] P. Ramesh and S. Meena, “IoT and Machine Learning Based Health Monitoring System for Pregnant Women,”International Journal of Advanced Research in Electronics and Communication Engineering, vol. 9, no. 5, pp. 412–417, 2021.
[5] M. Alsharif and N. Alotaibi, “Digital Health and Machine Learning Framework for Maternal Health Monitoring,”IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 4, pp. 1765–1774, 2022.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)