I. J. Education and Management Engineering, **2026, 3, 46-61**
Published Online on June 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijeme.2026.03.04

An Intelligent, Bilingual Pregnancy Health Monitoring System

Isah Omeiza Rabiu*
Federal University of Technology / Department of Computer Engineering, Minna, 234, Nigeria
E-mail: isah.rabiu@futminna.edu.ng
ORCID iD: https://orcid.org/0000-0002-2784-8891
*Corresponding author

Bitrus Judah Tanko
Federal University of Technology / Department of Computer Engineering, Minna, 234, Nigeria
E-mail: judah.m1801151@st.futminna.edu.ng
ORCID iD: https://orcid.org/0009-0000-2139-713X

Nuhu Bello Kontagora
Federal University of Technology / Department of Computer Engineering, Minna, 234, Nigeria
E-mail: nuhubk@futminna.edu.ng
ORCID iD: https://orcid.org/0000-0003-0130-2468

Received: July 18, 2025; Revised: November 14, 2025; Accepted: May 18, 2026; Published: June 8, 2026

Abstract: This research implements an intelligent, bilingual pregnancy health monitoring system for expectant mothers. A significant problem commonly experienced by expectant mothers in rural areas in Nigeria is the unavailability of a decent antenatal system and a shortage of experienced medical personnel and equipment. The proposed system comprises IoT sensors, including Electrocardiogram (ECG), body temperature, and heart rate sensors, connected to an ESP32 microcontroller for data acquisition and transmission. A predictive system built using Random Forest and Support Vector Machine (SVM) classifiers categorises pregnancy risk into low, medium, and high. A Flask-based web application for real-time data visualization and diagnosis was developed to display the collected data and visually represent the risk level diagnosis. The performances of the predictive models, Random Forest and Support Vector Machine (SVM), were evaluated using accuracy, precision, recall, and F1-score. Random Forest achieved an accuracy surpassing SVMs by a margin of 5.28%. Random Forest and SVM precision were then compared and there was an improvement of 6.49%.

In addition, Random Forest had a higher recall than SVM by 6.58%, and also had a performance increase of 6.49% on F1-score as compared to SVM. The comparative analysis shows that the Random Forest model works better than SVM in all the main measures. In this project, the Random Forest model was better than the SVM because it uses ensemble learning to manage the non-linear relationship, imbalance data and noise better to achieve superior accuracy, recall, and the F1 Scores. It was also more reliable in categorizing risks in pregnancy, as it was interpretable, which was also strong and guaranteed the timely and suitable intervention of health care.

Index Terms: Bilingual System, Random Forest, Support Vector Machine (SVM), Expectant Mothers, Pregnancy Health

1. Introduction

The life of a Woman passes through numerous phases and pregnancy is one of the most significant moments in her life, that is why mother, and child should both be well taken care of. The survey by [1] indicated that thousands of women lose their life every day during pregnancy and during labour such issues include still birth/pre-term/miscarriage etc. which can be prevented early in the stages. The maternal mortality rate is also unacceptable, 287 000 women died during and after childbirth and pregnancy in 2020 alone. In 2020, nearly 95 percent of all maternal deaths are found in low and lower middle-income countries, and primary preventable in nature [1].

Women need prompt and proper assistance and access to quality care by a medical care provider at the appropriate time during pregnancy. The reason why many pregnant women are not receiving proper care during pregnancy is due to

This work is open access and licensed under the Creative Commons CC BY 4.0 License.

such reasons as lack of education, late coming, lack of information, poverty, insufficient skilled health professionals and poor health care services [2]. In the rural and poor regions, the mother to death ratio is very high since the regions are too far to access immediate medical care. Monitoring of these crucial health indicators in the course of pregnancy may contribute to identification, diagnosis, and even prevention of a great number of risks inherent to pregnancy [3].

Missing timely antenatal care increases the risk of various complications during pregnancy, such as premature labor, low birth weight, and health issues for the mother. Without regular check-ups, it's harder to detect and treat potential problems, which further endangers both the mother's and the baby's health [4,5].

The solution to these problems is that new concepts are required to make maternal healthcare useful, particularly in the villages with hard conditions. Digital health technologies represent a means of sealing these gaps, thankfully. It can be seen that this e-antenatal device demonstrates a new practice where technology is employed to monitor health, real-time data analysis and, consequently, access to the essential healthcare information and services is fast [6]. The smart functionalities such as classification of risk levels enables pregnant women in rural areas to be empowered to take an active role in properly pregnancy through this project. However, the development and implementation of targeted interventions to address the unique needs of pregnant women in rural areas are yet to be developed and implemented [6]. The main objective of this research, therefore, should be to develop a smart e-antenatal device that is specifically targeted at the rural setting of healthcare. The project will improve the state of maternal health and reduce maternal health risks in rural areas.

Although some of the beneficial achievements have been realized in maternal health care, there are still numerous pregnant women in parts of the world that are rural who are exposed to very severe difficulties in obtaining proper antenatal care and support. Although the problem of pregnancy outcomes may be alleviated by carrying out antenatal care visits, most of the problems can be exacerbated by socioeconomic constraints, geographical factors and absence of healthcare facilities [7]. Some of the pregnancy complications encompassed low birth weight, preterm births, maternal morbidity and absence of proper and due antenatal care some of the practices like irregular monitoring complicate the diagnosis and treatment of maternal health complications that can endanger the well-being of both the mother and the child [8].

Nigeria is not an exception since it does not have maternal healthcare services in rural communities and past interventions are not efficient in responding to the needs of pregnant women in rural societies. The conventional practices of antenatal care may not always be possible, workable or effective in the rural setting, thus exerting a considerable contribution to the healthcare disparities, hence we certainly require some new solutions, which may make use of the digital health technologies in enhancing the maternal healthcare services that the mothers living in the rural setting are receiving [9].

A high-level e-antenatal device is a convenient means of addressing these issues and monitoring, analyzing the data in real-time, and offering valuable healthcare information and services. Nevertheless, the development and use of such a device will be achieved through a keen insight into the different factors that influence maternal healthcare in rural communities. This involves the technological, infrastructural and socio-economic issues that must be resolved to ensure that it is successful and can be broadly applied [9].

2. Related Works

Pregnancy healthcare (PH) is very significant in ensuring that the mother and the baby are healthy throughout the pregnancy. Frequent check-ups allow clinicians to monitor the progress of the pregnancy, address the risks that may emerge and provide expectant mothers with the necessary guidance. In this way, PH allows reducing the risks of pregnancy and childbirth, including preeclampsia, gestational diabetes, preterm birth, etc. [10]. The recommendation of the World Health Organization has now been adjusted to make the pregnant women undertake at least eight check-ups during pregnancy. The visits are used to ensure that the mother and the baby remain healthy as the doctors can easily identify and control any problems that occur in the pregnancy at an early stage [11]. PH is also important in promoting healthy behaviors, administering valuable vaccinations and employment of nutritional assistance, which are significant in enhancing the outcomes of pregnancies [12]. Regular and good-quality antenatal care is not always easy to find in the rural areas, thus the necessity to resort to innovative methods that would help to bridge the gap in healthcare. Through technology, such as mobile health apps and monitoring gadgets, the healthcare providers will be capable of delivering a more holistic healthcare in the remote communities [13].

2.1. Internet of Things (IoT) in Healthcare

Internet of Things (IoT) has indeed imprinted itself in the healthcare industrial by linking instruments that constantly monitor and process real time medical statistics. Numerous researches indicate the way IoT is applied in systems to check the health of pregnant women: [14] constructed an IoT system that was used to measure maternal and fetal health and track indicators such as heart rates and blood pressure. A One-Dimensional Convolutional Neural Network (1-D CNN) is a type of deep learning model they used to sort the data and assist in diagnosing emergencies, and this system had a 92.59 diagnostic accuracy. Although it proved effective, the system had the problem of managing complex data and its reliance was very high on the availability of a stable internet connection.

[15] created an IoT platform capable of transmitting the health data of the mother, over GSM, to a cloud platform. Such an arrangement allows healthcare professionals to remotely track patients and receive notifications in a short period of time, which contributes to the enhancement of maternal care. Nevertheless, the system was not without its predicaments, such as ensuring that the technology was effective in the rural regions and its vulnerability to an uninterrupted internet connection.

2.2. Wearable and Sensor-Based Monitoring

The internet of things proposed by [16] is an IoT system that can monitor vital parameters such as blood pressure and heart rate, as well as the kicks made by the fetus using acceleration sensors. The system should have provided to provide low cost continuous monitoring. Nevertheless, it had certain drawbacks: it relied on technology to transmit information too much and the conventional ultrasound could not be used to track the information continuously.

The sensor network that [17] has developed is aimed at tracking the behavior of pregnant women as well as their vital signs with the aim of transmitting early warnings that can help manage risks. Nevertheless, so doing, they encountered difficulties in making sure that the sensor data was trustworthy and overcoming the complexity of a microservice-based structure.

2.3. Integrated Sensor Solutions

In [18], the authors created a system with sensors such as the Digital Humidity and Temperature (DHT11) to monitor temperature and humidity. The readings are displayed on an LCD screen and also relayed to the cloud as well as to the sensors which in most cases have low error rates, but the research did not ignore the fact that there are problems with the overall accuracy of the system and also added that it works best when the user is already familiar with the operational process of the system.

A high-risk pregnancy monitoring system was designed by [19] to monitor ECG, fetal heart rate and uterine contractions with the use of sensors. The system was scalable and affordable thus suited well to different regions. Nevertheless, the question of maintaining sensitive data confidentiality in the course of transmission remains.

2.4. Machine Learning in Maternal Health Monitoring

The functioning of the IoT systems related to the monitoring of maternal health has been positively affected by machine learning that allowed predicting potential issues and diagnosing the emergencies promptly:

[20] introduced a clever fetal monitoring device using the Healthcare IoT technology to constantly measure the wellbeing of the pregnant women and the unborn infant to improve the prenatal care by means of collecting real-time data.

A Gated Recurrent Unit (GRU) was presented in [21] to identify emergencies. Their findings indicated that this model was more effective than other approaches, such as Random Forest and SVM, in terms of accuracy and sensitivity in identifying problems.

2.5. Risk Management through AI

[22] introduced algorithms to process data on maternal health and presented graphs to display the risks of healthcare providers in various ways. The system was made to be economical whereas it was feared that information, which would be considered vital, would be omitted due to omission of some sensors.

Wearable AI-equipped belts were used in [23] to categorize fetal heart rates and identify risks and provide a superior alternative to traditional ones, but with generalizability and data security limitations.

2.6. Biosensors for Vital Monitoring

[24] installed sensors that detect the heart rate, temperature and pressure and were attached to an Arduino. In case of any abnormal readings, then the system would send out a buzzer as a reminder to the users, and although the system was very effective in forecasting health risks, when it had difficulties accessing the internet in rural locations, it got into issues.

The sensors were used by [25] to monitor some of the key maternal health indicators such as blood pressure and oxygenation. The data was transferred to a mobile application through IoT hence the information was easily accessible. The system, however, despite all these benefits, encountered problems of maintaining low costs.

[26] created a smart IoT platform to monitor high-risk pregnancies that incorporates various sensors to monitor the baby and the mother. They employed sophisticated algorithms to ensure that the system is able to forecast health related problems accurately which will enable the doctors to detect problems at an early stage and offer them better treatment.

3. Proposed Approach

3.1. Dataset Preparation

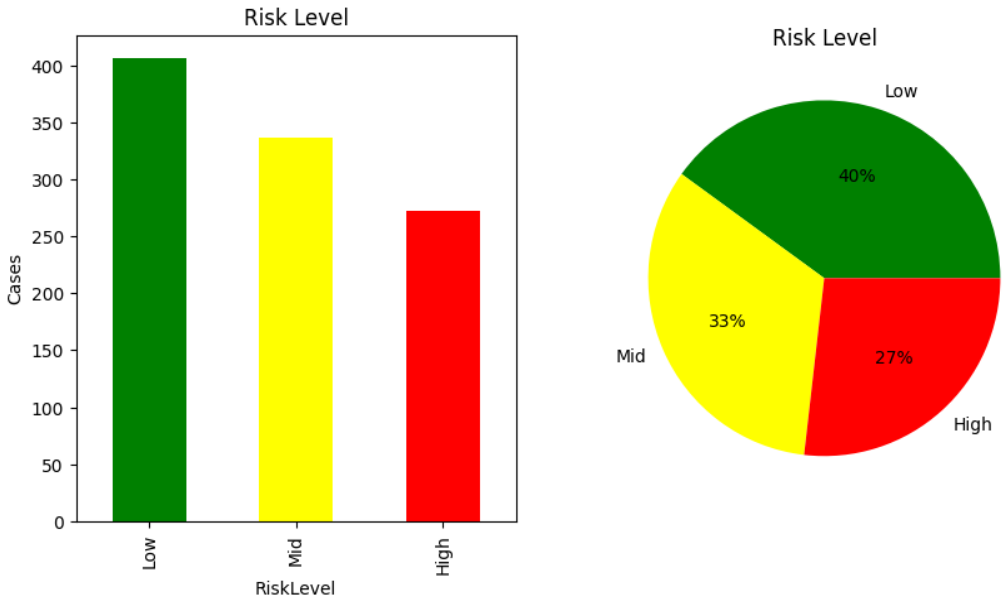
The method employed in my code to eliminate the outlier is a conditional filtering and selective removal method, which entails the identification and elimination of data according to the probable difficulty of the model to cope with the cases. This method plays an important role in data analysis, and this method is usually applied to remove any false or extreme values in the data to enhance the quality and accuracy of the data [27]. This is a manual outlier filter using index-based filtering. It is a straightforward and plain method particularly when the outliers are certain and known. This approach is used in contrast to more sophisticated methods, such as z-scores or IQR (Interquartile Range) filtering, in which data points that are known to be invalid or unrealistic are explicitly removed, in particular, based on the domain knowledge (o.k., a heart rate of 7 is implausible). One outlier heart rate value 7BPM was found and removed out of the training data (since heart rates lower than 40 are typically classified as abnormally low and can be due to an error in measurement or unusual physiological states) [27]. This outlier suggests a trend break data point, which is far out of the usual trend. Having such an outlier, however, increases the potential of an error in data input [28]. When the outliers were taken out, the remaining total dataset that was left was 1013 observations. The dataset employed in this study is the Maternal Health Risk dataset publicly available (University of California, Irvine (UCI) machine learning Repository) (https://archive.ics.uci.edu/dataset/863/maternal+health+risk) which includes physiological data of patients in rural hospitals and community clinics in Bangladesh.

3.2. Classification of Risk Levels

Risk Categories: The model is designed to classify pregnancy into one of three risk levels:

i. Low-Risk (or class 0): Indicates a healthy pregnancy with minimal risk of complications.
ii. Mid-Risk (or class 1): Suggests a moderate risk where regular monitoring is required.
iii. High-Risk (or class 2): Flags a pregnancy with a high likelihood of complications, necessitating immediate medical attention.

The data show a significant distribution of risk levels as in Fig.1. Most of the cases are categorized under the Low Risk category (406 cases) which means normal health factors and increased chances of avoiding complications during pregnancy. There are a significant number of cases that are considered as the Mid Risk (336 cases), indicating that there is a moderate risk level and further monitoring is necessary. In the meantime, there are 272 cases of the High-Risk category, which indicate a high degree of risk that requires special attention and even treatment.



(a) (b)

Fig. 1. Risk Level Distribution.

The classifier model is trained to analyze real-time information of various sensors such as heart rate and body temperature and classify risk as low, medium, or high. There is a fixed blood pressure value (120 mmHg systolic) which is temporarily used instead of a live blood pressure sensor because hardware was not available at the time of development. This simplification has been recognized to be a limitation: replacing a dynamic physiological parameter with a fixed value lowers the clinical sensitivity of the blood pressure component of the model and likely leads the model to undervalue the risk conditions associated with blood-pressure such as pre-eclampsia. The improvement will be made by incorporating real blood pressure sensor readings in the future to enhance clinical validity. Because the

research involves the publicly available and anonymised dataset and the authors do not directly collect data regarding human subjects, the machine learning component did not require the formal ethical approval and the personal informed consent. The authors, however, accept that until the proposed system can be fully deployed in clinical practice, a full range of ethical clearance, informed consent processes and maintenance of the relevant data protection regulations will have to be adhered to and implemented.

3.3. Mapping Risk Levels to Integer Values

The pregnancy risk levels are classified into three categories: **low risk**, **mid risk**, and **high risk**. To facilitate machine learning algorithms and other data processing tasks, categorical variables such as these need to be converted into numerical form.

To achieve this, the following **mapping** was applied:

i. **Low Risk**: Mapped to 0
ii. **Mid Risk**: Mapped to 1
iii. **High Risk**: Mapped to 2

This mapping ensures that the level of risks is converted into numbers to enable machine learning models to operate with it. We did this by using the panda's library of Python and its map () method. Then to check again that all was converted correctly we ran the df.info () function. This operation provides a summary of the DataFrame, with the type of data and the number of non-empty values per column. This was one of the steps that ensured that the RiskLevel column was transformed into integers and was now ready to be used in the predictive models.

3.4. Feature Extraction

Patient body temperature is usually recorded in Fahrenheit according to standard medical practice in some areas. However, to meet our regional preference and local healthcare standards, it is necessary to convert the body temperature data from Fahrenheit to Celsius, which is more widely used in the Nigeria. The conversion from Fahrenheit to Celsius is based on the following well-known formula shown in equation 1.

$$C = 5/9\ x(F - 32) \tag{1}$$

where,
C represents the temperature in Celsius.
F represents the temperature in Fahrenheit.

This formula accounts for the fact that 32 ℉ is the freezing point of water in Fahrenheit, which corresponds to 0 ℃ in Celsius, and it adjusts for the different scales of measurement. In our system, we used Python's pandas library in our data processing pipeline to apply this conversion to the "BodyTemp" column, which originally stores temperature readings in Fahrenheit .

3.5. Splitting the Data into Training and Testing Sets

For this project, the dataset is splitted into 80:20 ratio for training and testing, respectively.

i. Training Set: This contains 80% of the data and is used to train the machine learning model.
ii. Testing Set: This contains 20% of the data and is used to evaluate the model's performance, ensuring it works well on new, unseen data.

3.6. Model Training

Fig. 2 shows a flowchart of the training process, and this section explains each step taken to build and fine-tune my machine learning model for classifying patient risk levels (low, mid, high) based on physiological data.

It started with gathering and preparing the data, then, an initial analysis was carried out to understand its structure and quality. The data was cleaned, transformed, and key features were selected to ensure the dataset was ready for modeling [29]. The reason behind choosing the random forest and SVM is their established capabilities in the medical classification tasks: random forest is a highly competent noise-resistant, class-imbalance, and feature-importance ranking model, and the SVM was the choice of a strong linear-kernel baseline that can compete with random forest on small to medium-sized datasets. In the case of Random Forest, the number of estimators was established to 100 and the maximum depth of the trees were established through 5-fold cross-validation grid search in the range of 5, 10, 20, None. In the case of SVM, tuning of regularisation parameter C and the type of kernel (Radial Basis Function (RBF), linear) was also done. First, a simple baseline model was set up to serve as a reference, and then, applied the above optimisation techniques to improve its performance. After training, the model was evaluated and its parameters saved for future use. This structured approach not only helped boost the model's accuracy and reliability but also provided valuable insights into the key factors influencing patient risk levels.

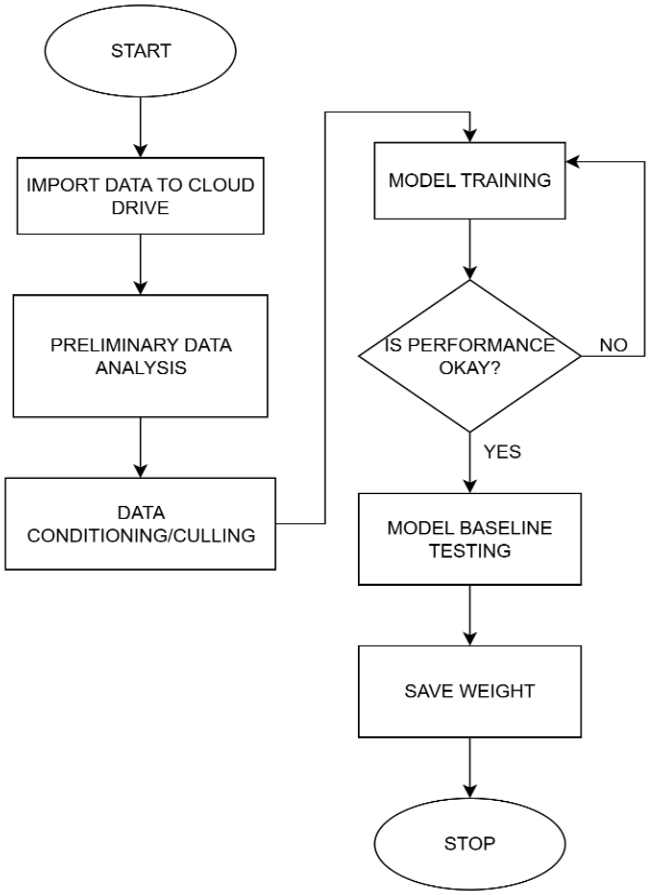


Fig. 2. Model Training Flowchart.

3.7. The System Flow Diagram

Fig.3 is a data flow and integration of the intelligent e-antenatal monitoring system which is specifically created to understand vital signs of the mother such as heart rate, body temperature and blood pressure and forecast possible health threats. The ESP32 microcontroller is the main component in the system, as it serves as the central processing unit to coordinate communication, between hardware and software and predictive model.

The sensors act as the body temperature sensor, a heart rate sensor, and an ECG sensor that has been modified to measure blood pressure, which means that the IoT sensors are the start of the process. These sensors will continuously receive real-time health data about the mother and transmit it to be processed to the ESP32 microcontroller. ESP32 will serve as a communication hub, controlling the flow and directing the data to the necessary elements. After receiving the raw data in ESP32, it forwards the raw data to a risk classification model in a random forest of machine learning. The machine learning model proceeds to provide the likelihood as to whether the mother stands a chance of any risk that may be considered as low risk, mid risk or high risk based on the input parameters. After processing it, it sends it back to the ESP32 microcontroller, and then to the web application that is created with Flask.

The web application built using flask is the user interface of the system. It shows the sensor information (body temperature, heart rate, blood pressure) and the predicted level of the risk by the model in a transparent and understandable manner. The system incorporates text-to-speech (TTS) capability to promote interaction and access to the system by users. Prediction of risk level is pre-translated into Hausa and play button is provided on the web application. Upon pressing, this button enables the user to listen to the audible Hausa version of diagnosis so that the information is easily perceived by the user, especially those who are more conversant with the sound of the Hausa language. In addition, the system relies on cloud storage to ensure the security of data and simplify its analysis. The ESP32 transmits sensor data and risk forecasting to the cloud, so that all the data is safely stored to be used in the future or further analysis. This option of recovery helps in the trend monitoring, increases reliability and long-term health surveillance. Simply put, the system is an extremely integrated platform where hardware and software can co-exist comfortably with each other. The sensors gather the health data in real time and are fed through the ESP32 and forwarded, the Random Forest model predicts the risk levels and displays them as a web application which is simple to read and is written in Flask and has Hausa TTS feedback. This data is saved on cloud storage and made accessible and secure. The given system can be viewed as an effective and straightforward way of tracking the health of pregnant mothers by integrating IoT sensors, machine learning, and a web platform.

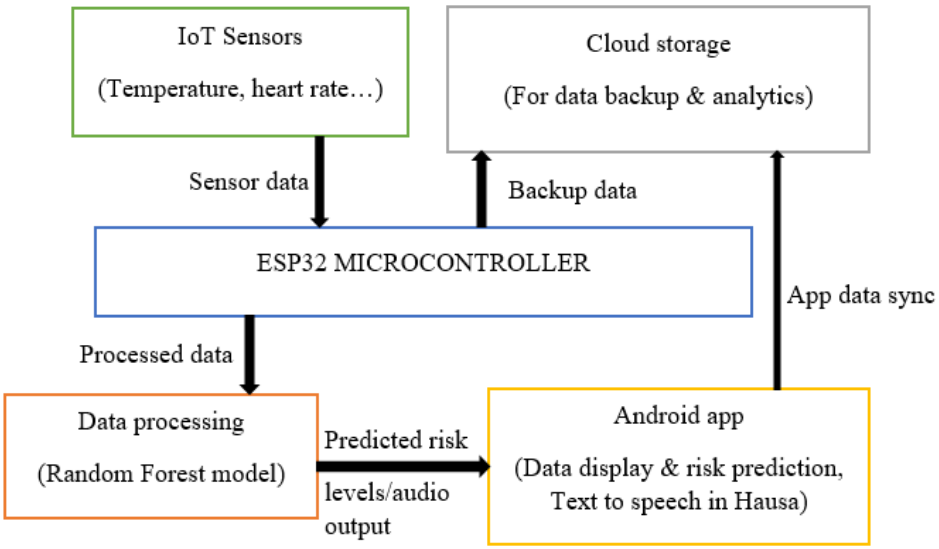


Fig. 3. Block Diagram of the Units that make up the E-Antenatal Device.

3.8. Circuit Calculations

List of Components and Specifications

a. ECG Sensor (e.g., AD8232):
i. Voltage: 3.3V
ii. Current: 170 µA (typical)
b. DHT22 Temperature Sensor:
i. Voltage: 3.0V to 5.5V
ii. Current: 1mA (standby), 1.5mA (active)
c. MAX30102 Heart Rate Sensor:
i. Voltage: 1.8V (core), 3.3V (I/O)
ii. Current: 600 µA (typical during sampling)
d. ESP32 Microcontroller:
i. Voltage: 3.3V
ii. Current: 160mA (average), up to 500mA in Wi-Fi transmission
e. Power Supply:
i. Typically 5V (regulated), with on-board 3.3V for sensors.

Power Consumption Calculations

For each component, power consumption PPP is calculated using the formula:

$$P = V \times I \tag{2}$$

ECG Sensor (AD8232)
a. Voltage: 3.3V
b. Current: 170 µA = 0.00017A

$$P_{ECG} = 3.3V \times 0.00017A = 0.000561W = 0.561mW$$

DHT22 Temperature Sensor
a. Voltage: 3.3V (assuming it's powered at 3.3V)
b. Current: 1.5mA = 0.0015A (active mode)

$$P_{DS18B20} = 3.3V \times 0.0015A = 0.00495W = 4.95mW$$

MAX30102 Heart Rate Sensor
a. Voltage: 3.3V
b. Current: 600 µA = 0.0006A

$$P_{MAX30102} = 3.3V \times 0.0006A = 0.00198W = 1.98mW$$

ESP32 Microcontroller
a. Voltage: 3.3V
b. Current: 160mA = 0.16A (average during operation)

$$P_{ESP32} = 3.3V \times 0.16A = 0.528W$$

Total Power Consumption

Next, sum the power consumption of all components:

$$P_{total} = P_{ECG} + P_{DS18B20} + P_{MAX30102} + P_{ISD1820} + P_{Arduino} + P_{ESP32} \quad (3)$$
$$P_{total} = 0.561mW + 1.98mW + 650mW + 528mW$$
$$P_{total} = 1180.541mW \approx 1.44W$$

3.9. Circuit Design and Power Consumption Calculations

The total power consumption of the system was calculated based on the individual power requirements of each component. The total estimated power consumption considering all components (including the ISD1820 and Arduino Uno referenced in Equation 3) was found to be approximately 1.44W. However, Table 1 presents only the four implemented hardware components (ECG sensor, DHT22, MAX30102, and ESP32), whose combined power is 0.54W. The selected 5V power supply, rated for at least 500mA, provides adequate capacity to power all components. These are depicted in Table 1.

Table 1. Power Consumption of Hardware Components.

Component	Voltage (V)	Current (A)	Power (W)
ECG Sensor (AD8232)	3.3	0.00017	0.000561
DHT22 Temperature Sensor	3.3	0.0015	0.00495
MAX30102 Heart Rate Sensor	3.3	0.0006	0.00198
ESP32 Microcontroller	3.3	0.16	0.528
Total			0.54W

This ensures that the system operates reliably within the power limits and that all components receive the appropriate voltage and current

3.10. Natural Language Processing

Considering the flaws of the Natural Language Processing (NLP) implementation, another strategy was developed to achieve the aims of the project without being too complicated in terms of functionality and accessibility. The solution included the following steps:

1. Pre-Prepared Translations: The diagnostic key words, low-risk, mid-risk and high risks, were translated into Hausa manually with the assistance of the linguistic specialists or credible sources.
2. Text-to-Speech (TTS) Integration: The Hausa translations that are already prepared were translated to audio using Google Text-to-Speech (TTS) technology. This was to make sure that the machine was able to provide auditory responses to Hausa speakers.
3. Application Integration: TTS-generated audio files were obtained and incorporated in the software application. Whether a diagnostic result is created (e.g., low-risk), the appropriate Hausa audio file is played, which allows the user to comprehend the result in their native language.

4. Results and Discussion

The study examined the ability of the Random Forest and the SVM models to classify pregnancies in various risk categories. It also covered the integration of the system with the IoT sensors and the web app explaining that the system gathers data in real time and provides timely and helpful feedbacks.

4.1. Comparative Study between the Random Forest and SVM Classifiers.

The training of the Random Forest and SVM models in this project was to classify pregnancies into the low-risk, mid-risk, or high-risk categories. In this section, the performance of each of them is compared and discussed in detail based on the metrics; accuracy, precision, recall and F1-score.

Random Forest model stood by 84 percent and SVM model at 79 percent. This means that the Random Forest model is more capable of classifying the risk of pregnancy in all the categories. Although both models have been quite accurate, the difference of 5 percent suggests that the model of the Random Forest is a bit more valid. This is likely because it has a number of decision trees which could better handle many types of data.

4.2. Precision, Recall, and F1-Score

A closer look at the individual class metrics highlights key differences between the two models. Table 2 and Figure 4 provide a comparison of precision, recall, and F1-score for each class.

Table 2. Comparison of Classification Metrics for Random Forest and SVM.

Class	Metric	Random Forest	SVM
0 (Low-risk)	Precision	0.85	0.84
	Recall	0.79	0.70
	F1-Score	0.82	0.77
1 (Mid-risk)	Precision	0.78	0.72
	Recall	0.79	0.78
	F1-Score	0.79	0.75
2 (High-risk)	Precision	0.88	0.83
	Recall	0.96	0.95
	F1-Score	0.92	0.88
Overall	Accuracy	0.84	0.79
Macro avg.	Precision	0.84	0.80
	Recall	0.85	0.81
	F1-Score	0.84	0.80

The two models are generally high in precision but in most instances, Random Forest is slightly better than SVM. In class 0 (low-risk), they are almost the same 0.85 in Random Forest and 0.84 in SVM, which demonstrates that they are similarly effective in finding true positives. But in class 1 (mid-risk) and class 2 (high-risk), Random Forest is ahead, having precision scores of 0.78 and 0.88 as opposed to SVM having a precision score of 0.72 and 0.83.

Recall is used to determine the ability of the models to discover true positive cases in each of the classes. Random Forest model is significantly stronger than SVM in the identification of low-risk pregnancies with recall of 0.79 other than SVM which has a recall of 0.70. This implies that SVM is misclassified more on low-risk cases. In the case of mid-risk pregnancies (class 1), the two models have a similar result of 0.79 and 0.78 under Random Forest and SVM respectively. Both models are particularly effective when it comes to high-risk pregnancies (class 2), though the first is a bit more successful with 0.96 than SVM with 0.95.

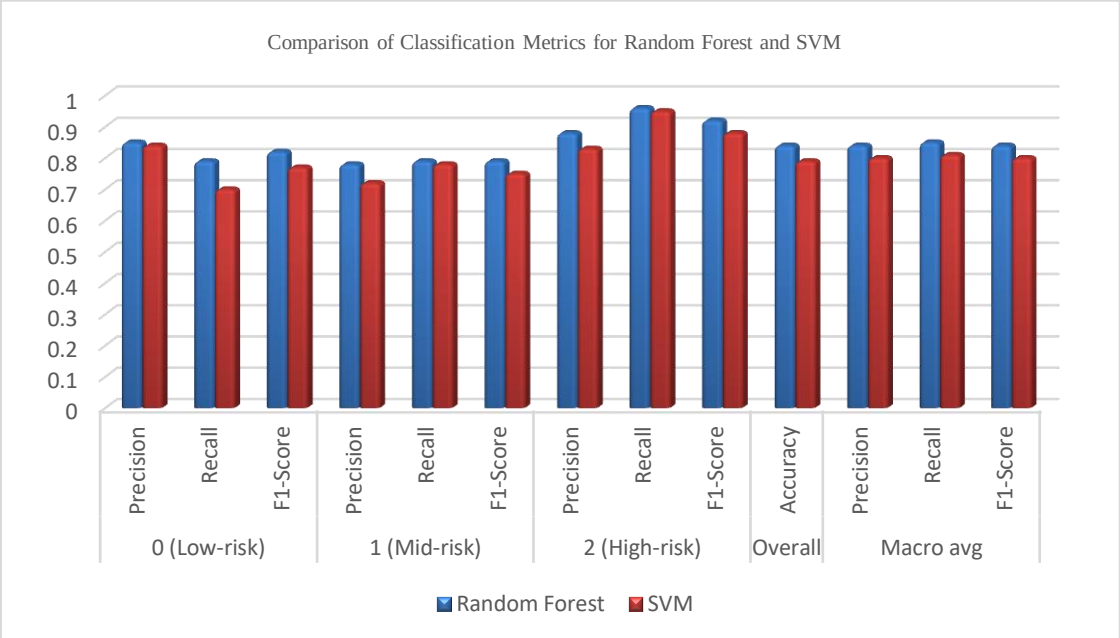


Fig. 4. Comparison Of Classification Metrics For Random Forest And SVM.

F1-score that balances between accuracy and recall indicates that the implementation of the Random Forest always yields a better overall rating when compared to SVM. Random Forest has an F1-score of 0.92 with high-risk pregnancies (class 2) which is better than SVM with F1-score of 0.88, and hence it is the better option in such a critical

healthcare category. Both the models show similar performance in mid-risk cases (class 1), with the F1-scores being 0.79 in Random Forest and 0.75 in SVM.

The confusion matrices for both models further illustrate their differences in performance (Tables 3 and 4).

Table 3. Confusion Matrix of Random Forest.

	Predicted Class 0	**Predicted Class 1**	**Predicted Class 2**
Actual Class 0	64	15	2
Actual Class 1	9	53	5
Actual Class 2	2	0	53

Table 4. Confusion Matrix of SVM.

	Predicted Class 0	**Predicted Class 1**	**Predicted Class 2**
Actual Class 0	57	19	5
Actual Class 1	9	52	6
Actual Class 2	2	1	52

Analysis of Class 0 (Low-risk)

Random Forest classifier has achieved 64 correct low-risk prediction cases out of 81 compared to 57 correct low-risk prediction cases of SVM model.

SVM had a higher number of false classifications of more low-risk pregnancies as mid-risk with 19 cases wrongly classified compared to 15 by the Random Forest. This proves that the Random Forest is more useful at differentiating between low-risk pregnancies and mid-risk pregnancies.

Analysis of Class 1 (Mid-risk)

Both models were similar regarding mid-risk cases. Random Forest had an accuracy of 53/67 and SVM had an accuracy of 52/67.

The models have comparable numbers of mid-risk cases that are mistaken to be low risk, 9 cases per model. Random Forest however falsely categorized 5 cases as high-risk whereas SVM falsely categorized 6 cases.

Analysis of Class 2 (High-risk)

The high-risk group performed the best in both models. Random Forest was able to correctly classify 53/55 cases whereas SVM was able to correctly classify 52/55 cases.

This category had a very few instances of misclassification by the two models (Random Forest and SVM), which increased their efficacy in identifying high-risk pregnancies. From a clinical perspective, the most consequential error is misclassifying a high-risk pregnancy as low or mid-risk (a false negative), as this could delay critical medical intervention and place the mother or foetus at serious risk. For Random Forest, only 2 out of 55 high-risk cases were misclassified (false negative rate: 3.6%), compared to 3 for SVM (false negative rate: 5.5%). While these rates are encouraging, clinical deployment would require further validation to confirm that such false-negative rates are acceptable within the target healthcare setting.

Table 5. Percentage Improvement of Random Forest Model over SVM.

Metric	**Random Forest (%)**	**SVM (%)**	**Percentage Improvement (%)**
Accuracy	83.7	79.5	5.28
Precision	85.3	80.1	6.49
Recall	84.2	79.0	6.58
F1-Score	83.8	78.7	6.49

The predictive models, randomly forest and support vector machine (SVM) were evaluated based on four important measures which included accuracy, precision, recall, and F1-score as indicated in Table 5. The findings reveal that, Random Forest was better in every aspect, as discussed below.

1. Accuracy

The Random Forest model was more accurate at 83.7 than SVM, 79.5% with a difference of 5.28. This implies that Random Forest will be more precise in determining the various risk levels in maternal health.

2. Precision

The precision of SVM was 80.1% and that of Random Forest was 85.3% which is an improvement of 6.49. It implies that, the cases that are high- risk are more accurately predicted by the Random Forest with less probability of false positives.

3. Recall

Random Forest also had a recall of 84.2% which was 6.58 higher than that of SVM, 79.0%. This implies that random forest has a higher capacity to know true positive cases, and this is essential in identifying high-risk pregnancies.

4. F1-Score

The F1-score of SVM was 78.7% and the 83.8% of Random Forest represented a difference of 6.49. Because the F1-score is a weighted average between recall and precision, it can be affirmed that Random Forest tends to be more useful in terms of classifying the level of risk with accuracy.

The results show that the most appropriate algorithm to apply in this application is the Random Forest. It is effective on complex patterns, it is insensitive to noise and is an ensemble learning algorithm which reduces overfitting. This is why it is particularly suitable to the study of maternal health statistics. On the whole, the Random Forest is a better predictor of this project than SVM because it is superior in all the essential measures. This greater precision will result in better classifying of maternal health risks to enhance the performance of the smart e-antenatal system. It should be noted that the reported results are based on a single stratified 80:20 train-test split, which is a recognised limitation. Future work will incorporate k-fold cross-validation (k = 10) and bootstrapped confidence intervals to provide statistically robust estimates of generalisation performance. Additionally, given the moderate class imbalance in the dataset (406 low-risk, 336 mid-risk, 272 high-risk), complementary metrics such as macro-averaged ROC-AUC and balanced accuracy would further strengthen the evaluation and are recommended for future iterations of this work.

4.3. Sensor Data Processing

The system employed a set of sensors such as temperature sensors, ECG and heart rate sensors to collect vital maternal health information. The sensors were in turn connected wirelessly to the ESP32 microcontroller that relayed the data to the cloud itself which then processes the data, classifies it and transmits the ensuing information to the mobile application via Wi-Fi as depicted in Fig.5.

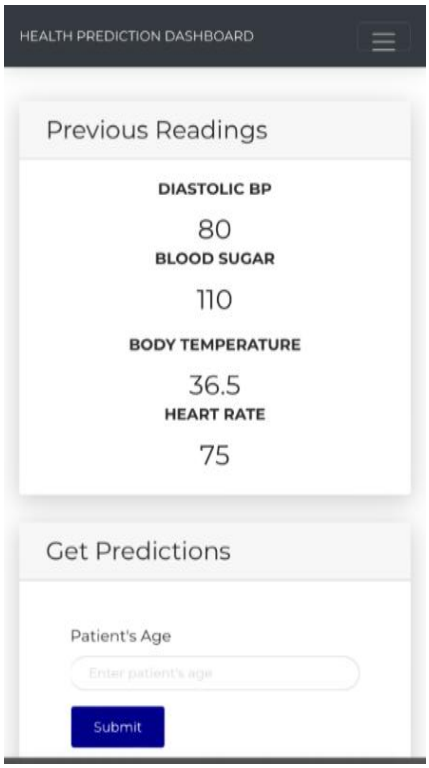


Fig. 5. Health Prediction Dashboard (Home Page) of the Flask Application.

The sensor data is displayed in real time using the Flask application which gives patients and healthcare providers the immediate feedback as it is presented below in Fig.6. The visualization of the risk classification is also provided in the application, where the low, mid, and high-risk pregnancies are reflected using different colour codes.

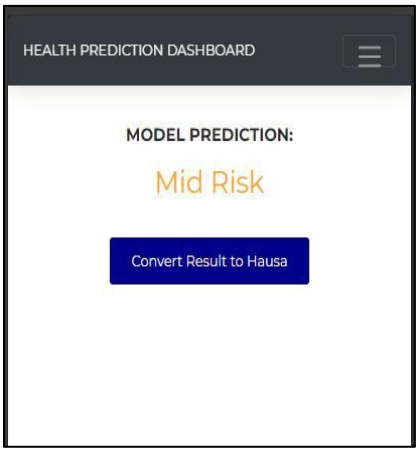


Fig. 6. Model Risk Prediction Page of the Flask Application.

4.4. Adapted Approach Outcomes

In order to overcome the problems with the translation, a less complicated approach has been chosen: the diagnostic risk levels (low-risk, mid-risk, high-risk) are translated with the help of the ready-made Hausa translations, and the Text-to-Speech (TTS) tool offered by Google is used. Some of the major enhancements included in this approach included:

- Precision and Consistency: The fact that the translations were ready-made meant that the system presented good and credible audio feedback in Hausa, which ruled out chances of misinterpretation.
- Better Accessibility: Hausa-speaking patients would have no problem interpreting the diagnostic test results since the data was presented orally. This made the device far easier to use by non-English speakers.
- Smooth Implementation: The audio files of TTS were effectively integrated into the system and the right Hausa translation was immediately provided once the risk level was identified.

4.5. Quantitative Results

Although the original NLP model was not able to provide quantitative measures of accuracy because of its underperformance, the alternative solution was able to achieve:

- 100% Consistency in Translations: The translations were already prepared and there could be no variability and mistakes.
- Real-Time Feedback: The implementation of the TTS provided feedback results in milliseconds of diagnosis, which guaranteed no delays in audio feedback.

4.6. User Testing and Feedback

The modified solution was tested by a group of Hausa speaking participants through user testing. Feedback obtained consisted of:

Ease of Understanding: All respondents said that the auditory feedback was understandable and comprehensible.

System Reliability: The participants have pointed out that the audio playback system was also reliable, and the system did not exhibit any errors in pronunciation or any functionality.

4.7. Impact on Overall System Design

The results show that while the NLP model didn’t perform as expected, the alternative approach still met the project’s key objectives. This highlights the importance of adaptability and problem-solving when facing unexpected challenges. As shown in Fig.7, the device successfully shares diagnostic results in Hausa, ensuring it stays easy to use and accessible for patients.

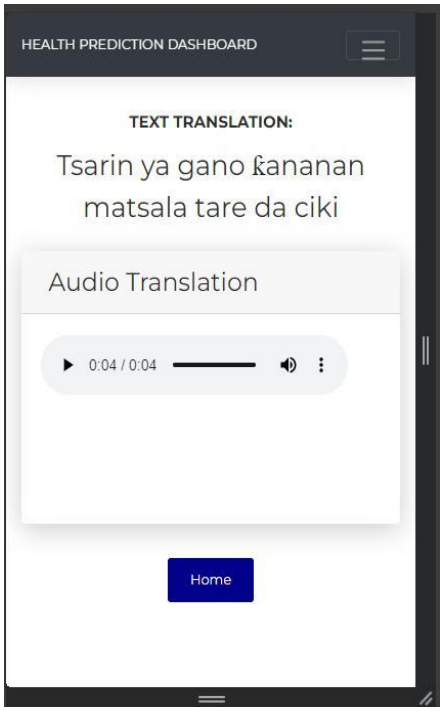


Fig. 7. Translation and Text-To-Speech Page of the Flask Application.

5. Conclusion

This paper presented the findings of the predictive system implemented with the Random Forest and SVM models to identify levels of pregnancy risk. In addition to connecting IoT sensors for collecting data and a Flask-based web application for real-time monitoring, the system incorporated Hausa-language audio feedback via Google Text-to-Speech (TTS) using pre-prepared static translations. It should be noted that an NLP-based dynamic translation component was initially planned but could not be realised due to technical limitations; the TTS-with-static-translations approach is therefore the implemented bilingual contribution of this work. The results show that the Random Forest model outperformed SVM across accuracy, precision, recall, and F1-score, and was therefore adopted as the final model. The resulting system demonstrates potential suitability for rural healthcare environments by providing real-time risk feedback in the Hausa language; however, further clinical validation, pilot deployment, and user-centred evaluation involving healthcare professionals are required before the system can be considered ready for real-world deployment. Future work will address the blood pressure sensor integration, external dataset validation, and a formal clinical pilot study.

All the Declarations and Statements

Author Contributions Statement

Isah Omeiza Rabiu – Conceptualization, Methodology, and Supervision: Proposed research ideas, Constructed the overall framework, and supervised project execution, Drafted the initial manuscript, Review and Editing.
Bitrus Judah Tanko – Data Curation and Software Implementation: Handled data acquisition, dataset preprocessing, and implementing the research model.
Nuhu Bello Kontagora – Model Training, Validation, and Performance Evaluation: Led the model training process, validated results using standard metrics, and benchmarked performance against existing methods.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement

The authors declare no conflicts of interest.

Funding Declaration

This research receives no funding.

Data Availability Statement
This study analyzed publicly available datasets. The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Ethical Declarations
There are no ethical declarations because this research does not involve human subjects and/or animals.

Acknowledgments
The authors are grateful to the management of the Federal University of Technology, Minna, for providing the enabling environment for the execution of the study.

Declaration of Generative AI in Scholarly Writing
During the preparation of this manuscript/study, the authors used Grammarly for English grammar checking and validation. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Abbreviations
The following abbreviations are used in this manuscript:

1-D CNN - One-Dimensional Convolutional Neural Network
AD8232 - ECG Sensor Module
AI - Artificial Intelligence
CNN - Convolutional Neural Network
DHT22 - Digital Humidity and Temperature Sensor
ECG - Electrocardiogram
e-antenatal - Electronic Antenatal
ESP32 - ESP32 Microcontroller
Flask - Flask (Python Web Framework)
GRU - Gated Recurrent Unit
GSM - Global System for Mobile Communications
IoT - Internet of Things
MAX30102 - Heart Rate / Pulse Oximeter Sensor
mHealth - Mobile Health
ML - Machine Learning
NLP - Natural Language Processing
PH - Pregnancy Healthcare
RBF - Radial Basis Function
SVM - Support Vector Machine
TTS - Text-to-Speech
UCI - University of California, Irvine
WHO - World Health Organization

References

[1] World Health Organization, *WHO recommendations on antenatal care for a positive pregnancy experience*, Geneva, Switzerland: World Health Organization, 2016.
[2] R. D. Aliyu, E. I. Gloria, O. Adebowale, and I. Rebecca, "mHealth for self-management in pregnancy: Perceptions of women in low-resource settings," *Procedia Comput. Sci.*, vol. 181, pp. 738–745, 2021, doi: 10.1016/j.procs.2021.01.226.
[3] P. Vinothiyalakshmi, V. Pallavi, N. Rajganesh, and V. Adityavignesh, V. "Internet of Things (IoT)-Based Smart Maternity Healthcare Services", In *Intelligent Systems and Sustainable Computational Models* (pp. 266-274), 2024. Auerbach Publications.
[4] S. Ashraf, S. P. Khattak, and M. T. Iqbal, "Design and implementation of an open-source and Internet-of-Things-based health monitoring system", *Journal of Low Power Electronics and Applications*, vol. *13*, no. 4, 57, 2023.
[5] A. Bansal, V. A. Athavale, K. Kaur, and A. Garg, "IoT enabled prenatal health monitoring system for pregnant women," in *Proc. 3rd Int. Conf. ICT Digit. Smart Sustain. Dev. (ICIDSSD)*, May 2023, p. 289.
[6] B. Dhanwanth, R. Saravanakumar, T. Tamilselvi, and K. Revathi, "A smart remote monitoring system for prenatal care in rural areas," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, no. 3, pp. 30–36, 2023.

[7] S. G. Tenaw, D. Tsega, B. T. Zewudie et al., "Completion of the maternal continuum of care and its association with antenatal care attendance during previous pregnancy among women in rural areas of the Gurage Zone, Southwest Ethiopia: A community-based cross-sectional study," BMJ Open, vol. 12, p. e066536, 2022. doi: 10.1136/bmjopen-2022-066536.
[8] W. C. Yang, S. Sabwa, A. D. Mebratie, B. Amboko, I. Mugenya, S. Kim, ... and C. Arsenault, "Quality of antenatal care and perinatal outcomes: evidence from a cohort study in Ethiopia, Kenya, South Africa, and India", *medRxiv*, 2025-04, 2025.
[9] H. Rehan, "Bridging the digital divide: A socio-technical framework for AI-enabled rural healthcare access in developing economies," *Eur. Vantage J. Artif. Intell.*, vol. 2, no. 1, pp. 19–27, 2025.
[10] M. T. Sangy, M. Duaso, C. Feeley, and S. Walker, S. "Barriers and facilitators to the implementation of midwife-led care for childbearing women in low-and middle-income countries: A mixed-methods systematic review", *Midwifery*, vol. *122*, 103696, 2023.
[11] A. A. Abuosi, E. A. Anaba, A. A. Daniels, A. A. A. Baku, and J. Akazili, "Determinants of early antenatal care visits among women of reproductive age in Ghana: evidence from the recent Maternal Health Survey", *BMC pregnancy and childbirth*, vol. *24*, no. 1, 309, 2024.
[12] F. S. Tanberika, T. B. Sansuwito, and H. C. Hassan, "Knowledge in enhancing antenatal care compliance and improving maternal and neonatal health outcomes", *Journal of Angiotherapy*, vol. *8*, no. 12, 1-7, 2024.
[13] M. Nazari, H. Sadr, H. Emami, and R. Rabiei, "A telemonitoring system for high-risk pregnancies: A novel approach to enhancing prenatal care and reducing maternal and fetal complications," in Proc. 2025 11th Int. Conf. Web Res. (ICWR), Apr. 2025, pp. 431–437.
[14] R. Ettiyan, and V. Geetha, V. "Iod-Nets–An IoT based intelligent health care monitoring system for ambulatory pregnant mothers and fetuses", *Measurement: Sensors*, *27*, 100781, 2023.
[15] X. Li, Y. Lu, X. Fu, and Y. Qi, "Building the Internet of Things platform for smart maternal healthcare services with wearable devices and cloud computing," Future Gener. Comput. Syst., vol. 118, pp. 282–296, 2021, doi: 10.1016/j.future.2021.01.016.
[16] B. Priyanka, V. M. Kalaivanan, R. A. Pavish, and M. Kanageshwaran, "IoT based pregnancy women health monitoring system for prenatal care," in *2021 7th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Mar. 2021, pp. 1264–1269, doi: 10.1109/ICACCS51430.2021.9441677.
[17] R. Ivanov, S. Yordanov, and D. Dinev, "Internet of Things–based pregnancy tracking and monitoring service," in *2022 Int. Conf. Autom. Informatics (ICAI)*, Oct. 2022, pp. 298–302.
[18] H. Hashim, S. F. B. Salihudin, and P. S. M. Saad, "Development of IoT based healthcare monitoring system," in *2022 IEEE Int. Conf. Power Eng. Appl. (ICPEA)*, Mar. 2022, pp. 1–5.
[19] A. Velusamy, J. Akilandeswari, and M. Priya, "IoT-enabled intelligent maternal intensive care: A research study," in *2023 Int. Conf. Self Sustain. Artif. Intell. Syst. (ICSSAS)*, Oct. 2023, pp. 1377–1382.
[20] P. A. Mathina, S. Saravanadevi, and M. Pavithra, "Smart fetes monitoring system using healthcare IoT for pregnant women," in *Proc. 2024 7th Int. Conf. Circuit Power Comput. Technol. (ICCPCT)*, vol. 1, Aug. 2024, pp. 607–612.
[21] J. Zhang, M. Lau, and Z. Zhu, "Hybrid CNN-GRU model for exercise classification using IMU time-series data," *J. Mach. Intell. Data Sci. (JMIDS)*, vol. 5, no. 1, pp. 54–64, 2024.
[22] A. Siddika and M. Sultana, "Maternal health risk analysis by using exploratory data analysis and machine learning algorithms," *Int. J. Eng. Res. Comput. Sci. Eng.*, vol. 10, pp. 146–151, 2023.
[23] A. Rattanasak, T. Jumphoo, W. Pathonsuwan, K. Kokkhunthod, K. Orkweha, K. Phapatanaburi, P. Tongdee, P. Nimkuntod, M. Uthansakul, and P. Uthansakul, "An IoT-enabled wearable device for fetal movement detection using accelerometer and gyroscope sensors," *Sensors*, vol. 25, no. 5, Art. no. 1552, Mar. 2025, doi: 10.3390/s25051552.
[24] S. Veena, and D. J. Aravindhar, "Remote monitoring system for the detection of prenatal risk in a pregnant woman", *Wireless Personal Communications*, Vol. *119, no.* 2, 1051-1064, 2021.
[25] M. Nazari, H. Sadr, H. Emami, and R. Rabiei, "A telemonitoring system for high-risk pregnancies: a novel approach to enhancing prenatal care and reducing maternal and fetal complications", In *2025 11th international conference on web research (ICWR)* (pp. 431-437). IEEE.
[26] S. Addanke and R. Anandan, "Secure IoT based smart system for monitoring health care for ambulatory and fetal," *J. Auton. Intell.*, vol. 7, no. 3, 2024.
[27] B. Bala and S. Behal, S. "A brief survey of data preprocessing in machine learning and deep learning techniques", In *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 1755-1762), 2024. IEEE.
[28] R. Foorthuis, "On the nature and types of anomalies: A review of deviations in data," *Int. J. Data Sci. Anal.*, vol. 12, no. 4, pp. 297–331, 2021, doi: 10.1007/s41060-021-00259-1.
[29] C. Bulut, and E. Arslan, E. "Comparison of the impact of dimensionality reduction and data splitting on classification performance in credit risk assessment", *Artificial Intelligence Review*, vol. *57, no.* 9, 2024.

Authors' Profiles



Isah Omeiza Rabiu earned his Bachelor of Engineering (B.Eng.) Degree in Electrical and Computer Engineering from the Federal University of Technology, Minna, Nigeria. He later pursued advanced studies at Ahmadu Bello University, Zaria, Nigeria where he obtained his M.Sc. Degree and Ph.D. Degree in Computer Engineering. Over the years, he has developed a strong research portfolio, with interests spanning intelligent systems, telemedicine, and biomedical imaging. His work focuses on leveraging emerging technologies to address complex real-world challenges, particularly in healthcare and automation, with the aim of enhancing efficiency, accessibility, and innovation in both local and global contexts.



Bitrus Judah Tanko holds a B.Eng. in Computer Engineering from the Federal University of Technology, Minna, Niger State, Nigeria. His research interests span data science, artificial intelligence, embedded systems, and public health.



Nuhu Bello Kontagora is a Lecturer in the Department of Computer Engineering at the Federal University of Technology Minna, Nigeria. He obtained his M. Tech in Computer Science and Engineering at the Ladoke Akintola University of Technology Ogbomoso, Nigeria. And had his B.Eng in Electrical & Computer Engineering from the Federal University of Technology Minna, Nigeria in the year 2010. He is currently a doctoral student at the Ahmadu Bello University, Zaria, Nigeria. His research interests include Artificial and Computational Intelligence, localization in sensor networks, Computer/network security, Internet of Things (IoT) and Software Defined Networking.

How to cite this paper:

Isah Omeiza Rabiu, Bitrus Judah Tanko, Nuhu Bello Kontagora, "An Intelligent, Bilingual Pregnancy Health Monitoring System", International Journal of Education and Management Engineering (IJEME), Vol.16, No.3, pp. 46-61, 2026. DOI:10.5815/ijeme.2026.03.04