INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

Predictive Model for Smart Healthcare Systems Using Random Forest Classifier

B. I. Ele[1], E. O. Omini[2] O. O. Obu[3], C. P. Isong[4] & D. E. Izuki[5]

[1, 2,3]Department of Computer Science, University of Calabar, Calabar, Nigeria

[4]College of Health Sciences Management and Technology, Calabar, Nigeria

[5]Directorate of Information and Communication Technology, University of Cross River State, Calabar, Nigeria

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150600076

Received: 22 June 2026; Accepted: 27 June 2026; Published: 07 July 2026

ABSTRACT

Health care systems in developing countries often face serious challenges, including limited resources, poor infrastructure, and delays in patient care. This study presents the development of a predictive model designed to assist in early health risk detection, particularly in resource-constrained settings. In this study, an improved predictive model for smart healthcare systems using Random Forest Classifier was created and embedded in a simple web interface. The model was trained on synthetic medical data and achieved an accuracy of 91% during testing. Health workers and others were able to use the system effectively, even with minimal digital skills. The platform provided real-time predictions, that will help users make quicker clinical decisions.

Keywords: Smart health care, Predictive model, Random Forest Classifier, Machine Learning, Decision Support System, Real-time Prediction

INTRODUCTION

Health care systems are facing increasing challenges as populations grow and medical needs become more complex. This situation is especially critical in developing countries, where hospitals and clinics often struggle to deliver timely care due to limited resources and overburdened staff (Akinyemi et al., 2020). At the same time, advances in technology and data science are opening up new ways to improve health care delivery (Topol, 2019).

One area that holds great promise is the use of predictive systems, tools that can analyze medical data and identify patterns that help anticipate health problems before they become serious (Rajkomar et al., 2019). Unfortunately, many health care centers still rely on manual record-keeping or underutilize the data they collect, making it difficult to spot early warning signs or track patient outcomes over time.

A predictive health care model changes this by using past and present data to forecast possible health risks, support clinical decisions, and enhance overall service delivery (Shickel et al., 2018). Instead of waiting for a problem to occur, health professionals can act earlier, potentially preventing complications and saving lives. This proactive approach also helps manage hospital resources more efficiently.

The focus of this study is to develop a predictive model that can be part of a smart health care system, specifically predicting the risk of Type 2 diabetes. This model uses machine learning techniques to process medical data and generate useful insights, such as predicting when a patient might develop a critical condition or identifying patterns in vital signs that require attention. The system is designed to be practical and easy to use, even in environments with limited technical support.

Ultimately, the goal is to bring intelligent tools into everyday health care practice, making it easier for hospitals to deliver better care, reduce strain on health workers, and improve patient outcomes in a sustainable way.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

LITERATURE REVIEW

In recent years, the use of digital technologies in health care has expanded rapidly, offering new ways to support diagnosis, monitor patient conditions, and improve overall service delivery (Jiang et al., 2017). One of the most promising innovations in this space is the development of predictive models, which analyze patient data to anticipate potential health issues before they escalate.

Predictive modeling leverages machine learning (ML) and artificial intelligence (AI) to recognize patterns in medical records and make informed predictions (Shahid et al., 2019). These tools have been successfully used to detect conditions like heart disease, diabetes, and stroke risk by analyzing variables such as symptoms, medical history, and vital signs.

In developed countries, predictive models are already being integrated into clinical settings. For instance, hospitals in the U.S. utilize AI to monitor intensive care unit (ICU) patients, sending alerts when vital signs deteriorate (Esteva et al., 2019). In India, models have been employed to estimate hospital admission rates based on seasonal disease patterns (Rao et al., 2021). These examples highlight how predictive systems can support both personalized treatment and health system planning.

However, in developing countries like Nigeria, such innovations remain limited. Hospitals still depend heavily on paper records, and where digital systems exist, they are often fragmented (Afolabi et al., 2022). This fragmented data environment limits the development of predictive models that rely on consistent and structured data inputs.

Some localized research has explored AI’s potential in Nigerian health contexts. Oladele et al. (2019) developed a neural network model to estimate hypertension risk in rural areas, while Ahmed and Musa (2021) applied decision trees for identifying high-risk pregnancies. Though these studies yielded promising results, they remained largely academic and had limited adoption in routine clinical use.

Challenges identified across the literature include access to quality data, concerns about model bias, the interpretability of predictions, and user readiness to trust and use AI-based tools (Obadolu et al., 2020). For predictive systems to be adopted successfully in Nigeria, they must account for local challenges such as limited internet access, staff shortages, and the need for intuitive interfaces.

This study builds on these findings by presenting a simple, robust model trained on structured data and designed with real-world constraints in mind, aiming to bridge the gap between academic research and practical health care application.

METHODOLOGY

To build an effective predictive model for health care, this study adopted a Waterfall approach to system development. The Waterfall model is a linear framework where each step is completed before moving to the next. It provided a clear structure for organizing the phases of planning, model creation, testing, and system integration.

Identifying system needs

The process began with understanding what health workers and hospital managers actually needed in a predictive system. This involved informal interviews and direct observation in local clinics. Several key requirements were noted, including:

i A simple way to enter patient information
ii Reliable predictions based on clinical signs and symptoms
iii A system that works even in low-tech environments
iv Clear results that are easy to interpret and act on

INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

These priorities helped shape the direction of the model and the system that would support it.

Dataset development

Since access to real hospital records was restricted for privacy reasons, the study used a simulated medical dataset from Kaggle.com. This dataset was created based on documented clinical trends from previous studies and included patient age, gender, health symptoms, and history of chronic illness. The dataset was selected due to its heterogeneous and unprocessed nature, making it highly suitable for evaluating the robustness of the proposed model. Its large size and diverse features mirror real-world big data challenges, including missingness, inconsistency, and redundancy. Moreover, as the dataset is publicly available, it ensures transparency and reproducibility for future research.

Before training the model, the data was carefully cleaned. Unnecessary fields were removed, missing values were handled appropriately, and the variables were normalized to ensure better performance during machine learning.

Machine Learning Model Design

The predictive model was built using a Random Forest Classifier algorithm, a method that was well suited for the binary outcomes, such as predicting whether or not a patient is at risk.

To develop the model, the dataset was divided into two parts: 80% for training and 20% for testing. Training involved exposing the model to known patterns in the data so it could learn how to classify future cases. Python was used for this task, along with Scikit-learn and other standard libraries.

```
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import accuracy_score, confusion_matrix, classification_report

# Instantiate the model
model = RandomForestClassifier(n_estimators=100, random_state=42)

# Train the model
model.fit(X_train, y_train)

# Make predictions
y_pred = model.predict(X_test)

# Evaluate the model
print("Accuracy:", accuracy_score(y_test, y_pred))
print("Confusion Matrix:")
print(confusion_matrix(y_test, y_pred))
print("Classification Report:")
print(classification_report(y_test, y_pred))
```

```
Accuracy: 0.9118
Confusion Matrix:
[[15156   732]
 [ 1032  3080]]
Classification Report:
              precision    recall  f1-score   support

           0       0.94      0.95      0.95     15888
           1       0.81      0.75      0.78      4112

    accuracy                           0.91     20000
   macro avg       0.87      0.85      0.86     20000
weighted avg       0.91      0.91      0.91     20000
```

Figure1: Model Training

System integration

Once the model was trained and tested, it was connected to a simple web-based platform. Health workers or even patients themselves could enter patient data through a form, and the system would instantly respond with a prediction like “low risk”, “moderate risk” and “high risk”.

The user interface was developed using HTML, CSS, and JavaScript, while the back-end was managed using Django Python web framework. This setup allowed for fast predictions with minimal hardware or internet requirements.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

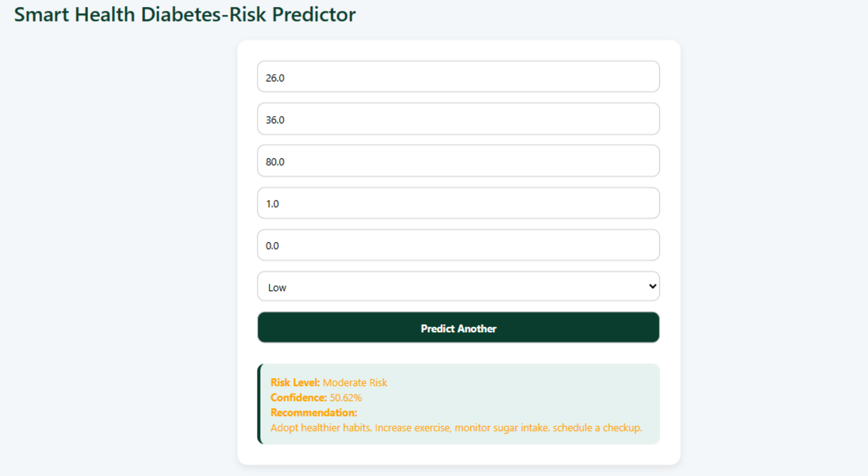

Figure 2: Data entry with prediction

Evaluating model performance

The effectiveness of the model was assessed using standard evaluation tools.

Performance was evaluated using standard classification metrics derived from the confusion matrix. Equations (1) – (4) describe the performance metrics used.

Accuracy: This is the proportion of correct predictions in the entire prediction.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1$$

Precision: This is the proportion of optimistic predictions in the whole set of positive classes predicted.

$$Precision = \frac{TP}{TP + FP}$$

Recall: This is the proportion of positive predictions in the entire positive class in the test data.

$$Recall = \frac{TP}{TP + FN} \qquad (3$$

F1-Score: This represents the harmonic mean of the recall and precision. High values can be interpreted as high classification performance.

$$F1 = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)} \qquad (4$$

These metrics were calculated using the test dataset and presented in visual formats to make the results easier to interpret.

Usability testing

After the full system was deployed in a test environment, several people including medical students were invited to try it out. They were asked to input sample patient records and observe the predictions. Most users reported that the system was straightforward and user-friendly.

Feedback focused on improving the layout, adding more clinical details. These points will be considered in future system upgrades.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

RESULTS

In this study, an improved predictive model for smart healthcare systems using Random Forest Classifier was developed and embedded in a simple web interface. After developing and integrating the predictive model into the smart health care platform, several performance tests and user evaluations were carried out. These tests focused on how accurately the system could predict patient risk levels, how well it performed in real-time conditions, and how users interacted with its interface.

Model accuracy and output reliability

The system's predictive engine, built using Random Forest Classifier, demonstrated solid performance during testing with synthetic patient data. The model achieved an accuracy rate of 91%, meaning that it was able to classify patients correctly in most cases. It also showed strong results in terms of precision and recall, indicating that the system could effectively identify individuals who were truly at risk without too many false alerts.

A confusion matrix was generated to better understand how the model handled correct and incorrect classifications. The matrix revealed that the system had relatively few errors, with most predictions falling into the correct risk category.

```
Accuracy: 0.9118
Confusion Matrix:
[[15156   732]
 [ 1032  3080]]
Classification Report:
              precision    recall  f1-score   support

           0       0.94      0.95      0.95     15888
           1       0.81      0.75      0.78      4112

    accuracy                           0.91     20000
   macro avg       0.87      0.85      0.86     20000
weighted avg       0.91      0.91      0.91     20000
```

Figure 3: Performance Metrics

From figure 3, accuracy is 91%, precision is 94%, recall is 95% and F1-score is 95%.

System responsiveness

Even under slower internet connections, the platform responded within a few seconds, proving that it can function effectively in environments where digital infrastructure is limited.

User feedback and observations

The platform was shared with a small group of medical professionals who tested it using sample records. Most users were satisfied with the system's performance and agreed that it could assist in identifying patients requiring close monitoring. They found the design simple, and the feedback from the model useful for decision-making.

Some participants suggested possible enhancements, such as to generate treatment suggestions alongside predictions. These comments have been considered for future development phases.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

DISCUSSION

The implementation and testing of the predictive model highlighted the real-world potential of simple, intelligent systems to improve decision-making in health care, especially in under-resourced environments. The system's solid performance, both in accuracy and ease of use, suggests it could be a valuable tool for frontline health workers.

One of the major strengths observed was the speed of prediction. Users could input data and receive a result within seconds. This kind of immediate feedback is extremely useful in fast-paced environments, where time and staffing are often limited. By pointing out possible health risks early, the system supports quicker clinical decisions and allows health workers to act before conditions worsen.

Another positive outcome was the system's user-friendliness. Even those with minimal technical training were able to use the platform with little to no difficulty. This confirms that the system's design is accessible and matches the capabilities of its intended users, an important consideration when introducing technology into health care settings that may not have regular IT support.

However, the system does have some limitations. The challenge of internet access. Since the system is web-based, it requires an active connection to function. Although it has been optimized to run on slow networks, some remote health facilities may not be able to use it reliably. For this reason, future updates could include offline data entry features, where users submit data without an internet connection and sync it later.

Beyond these issues, this paper shows that AI does not have to be complex or expensive to be useful in public health. A straight forward model, when thoughtfully designed and targeted at a real need can support better outcomes and improve workflow in hospitals. But technology by itself isn't enough. It must be paired with training, awareness, and support systems that help health professionals trust and apply the tools effectively.

In summary, the system represents a practical step forward in the use of predictive technology in local health care. It proves that with the right design and understanding of user needs, intelligent tools can be developed and applied even in areas where digital infrastructure is limited.

CONCLUSION

Predictive models using the Random Forest Classifier are vital in smart healthcare for early disease detection and continuous patient monitoring. By utilizing ensemble decision trees, these models analyze vast arrays of patient data to accurately classify health outcomes, such as predicting heart disease or diabetes risks with high precision.

This research set out to design a predictive model that could support smarter decision-making in health care environments especially for diabetic patients where access to advanced technology is limited. Using a structured development process guided by the Waterfall model, the system was built and tested with a focus on real-time predictions, ease of use, and low technical requirements.

The model performed well in test scenarios, showing a strong ability to identify patients at risk based on inputted data. Health workers were able to use the system to receive quick, meaningful feedback, which helped guide their decisions in a more timely and informed way. This feature is especially useful in settings where time, staff, and resources are stretched thin.

A major strength of the system lies in its simplicity. Designed to run on basic devices with limited connectivity, the platform proved usable even in constrained environments. Test users responded positively, noting that the system was easy to operate and practical for everyday clinical work.

At the same time, the project faced some challenges. Since the model was trained on simulated data, there is room for improvement when real patient data becomes available. Also, because the system is currently web-

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VI, June 2026

based, internet access remains a requirement. To address these issues, future versions should consider features like offline mode, mobile compatibility, and direct links to hospital databases.

Overall, this study shows that it is possible to build effective, intelligent health care tools using modest resources. By focusing on user needs and local conditions, systems like this can play an important role in supporting better patient outcomes and strengthening digital health systems in places where they are needed most.

REFERENCES

1. Afolabi, A., Adewale, S., & Musa, A. (2022). *Challenges of adopting artificial intelligence in Nigerian health care systems*. Journal of Medical Informatics in Africa, 9(2), 45–54.
2. Ahmed, S., & Musa, I. (2021). *Machine learning-based identification of high-risk pregnancies using decision trees*. Nigerian Journal of Health Informatics, 13(1), 25–32.
3. Akinyemi, K., Alabi, B., & Ojo, A. (2020). *Health care delivery challenges in sub-Saharan Africa: A case study of Nigeria*. African Journal of Public Health, 17(4), 203–210.
4. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z
5. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, 2(4), 230–243. https://doi.org/10.1136/svn-2017-000101
6. Obadolu, A., Jimoh, T., & Ogunbiyi, T. (2020). *Barriers to artificial intelligence implementation in Nigerian primary health care: A practitioner's perspective*. African Health Systems Review, 6(3), 112–120.
7. Oladele, T., Okeke, A., & Balogun, R. (2019). *Development of a neural network model for hypertension risk prediction in rural Nigeria*. Journal of Biomedical Engineering and Health Informatics, 5(1), 55–64.
8. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, 380(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259
9. Rao, S., Gupta, R., & Sharma, A. (2021). *Seasonal disease pattern analysis and hospital admission forecasting using AI in India*. International Journal of Health Data Analytics, 8(1), 11–19.
10. Shahid, N., Rappon, T., & Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making: A scoping review. *PLOS ONE*, 14(2), e0212356. https://doi.org/10.1371/journal.pone.0212356
11. Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1589–1604. https://doi.org/10.1109/JBHI.2017.2767063
12. Topol, E. (2019). *Deep medicine: How artificial intelligence can make healthcare human again*. Basic Books.