iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70118

# Interpretable Machine Leaning Models for Healthcare Decision Support

Ankita Bhide[1], Ms. Rohini Tambe[2], Buddhabhushan Tikte[3], Hemant Gaikwad[4]
*Department of Information Technology Marathwada Mitra Mandal's College of Engineering*

***Abstract: Healthcare decision support systems (DSS) play a crucial role in assisting clinicians with diagnosis and treatment planning. However, the lack of interpretability in machine learning models often leads to trust issues, limiting their adoption in clinical settings. This project presents an interpretable machine learningbased decision support system using the Random Forest algorithm to enhance transparency and accuracy in healthcare predictions. The system preprocesses healthcare data, applies feature selection, and generates interpretable insights through feature importance visualization. It integrates with hospital information systems (HIS) for real-time decision-making, ensuring seamless clinical workflow incorporation. Additionally, the project explores natural language processing (NLP) techniques to analyse text-based medical data, improving contextual understanding and decision accuracy. By balancing interpretability and predictive performance, this system enables clinicians to make informed decisions while enhancing trust in AI-driven recommendations. Model validation is conducted using real-world datasets and evaluated through metrics such as accuracy, F1-score, and AUC-ROC. The proposed approach contributes to the development of reliable, interpretable, and efficient healthcare AI solutions that can significantly improve patient outcomes.***
***Keywords: Interpretable Machine Learning, Healthcare Decision Support, Random Forest, Feature Importance, Natural Language Processing (NLP), Clinical Decision-Making, AI in Healthcare, Model Interpretability, Medical Data Analysis IoT, Real-time health monitoring, Cloud storage, ESP32 microcontroller, Remote healthcare, Data transmission, Vital signs monitoring, Remote patient monitoring.***

## I. INTRODUCTION

The integration of machine learning (ML) in healthcare decision support systems (DSS) has significantly improved clinical decision-making by enhancing diagnostic accuracy and treatment planning. However, many existing ML models lack interpretability, making it challenging for clinicians to trust and adopt these systems in real-world medical applications. This research aims to develop an interpretable ML-based decision support system utilizing the Random Forest algorithm, which provides both high accuracy and explainability through feature importance analysis.

The system processes real-world healthcare data, identifies key predictive features, and generates transparent and justifiable predictions. Additionally, natural language processing (NLP) techniques are incorporated to analyze unstructured medical text, further enhancing the system's capability to provide meaningful insights. The model is validated using real-world clinical datasets and evaluated through metrics such as accuracy, F1-score, and AUC-ROC to ensure reliability.

By balancing predictive performance with interpretability, this research contributes to the development of trustworthy AIdriven healthcare solutions that can be seamlessly integrated into clinical workflows. The proposed system aims to assist healthcare professionals in making informed, data-driven decisions, ultimately improving patient care and clinical outcomes.

## II. LITERATURE REVIEW

Rutika Bhagat and Prof. Pragati Patil (2023) designed a machine learning-based health monitoring system that employs the Decision Tree Classification algorithm.

The system monitors patient health metrics like BMI, age, gender, body temperature, blood pressure, and pulse rate to predict health risks. By integrating a Flask-based web application, the system provides users with real-time health risk predictions and personalized recommendations. Despite achieving an accuracy of 93%, the study identifies the
complexity and cost as key challenges for system adoption

Shalini et al. (2021) developed an IoT-based health monitoring system that utilizes sensors to monitor patient vitals such as heart rate and temperature. Their system successfully reduced the need for manual intervention by automating data collection and transmission to healthcare providers in real-time. However, their study highlighted limitations in system scalability and integration with existing healthcare infrastructure, which are critical for widespread adoption.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 13 Issue V May 2025- Available at www.ijraset.com*

Prasad and Jayaram (2022) introduced a smart health monitoring system using wearable IoT sensors combined with cloud computing for data storage and real-time analysis. Their system achieved higher accuracy in tracking health parameters and allowed healthcare professionals to intervene early by providing instant alerts. However, they encountered challenges in ensuring data privacy and managing cloud dependency.

Kadhim et al. (2020) presented an IoT-based patient monitoring system focusing on wireless sensors for health data collection. Their research emphasized the benefits of remote monitoring, particularly for patients in remote areas. The system allowed for continuous tracking of vital signs, but the study lacked experimental validation and faced challenges with maintaining system reliability during extended use.

Kartikee Uplenchwar et al. (2017) developed an IoT-based health monitoring system using Raspberry Pi and Arduino. It monitored vital health parameters such as ECG, pulse rate, temperature, and body position. While the system enabled monitoring on mobile devices, it lacked live monitoring and data storage capabilities,
which are essential for continuous healthcare solutions

These studies demonstrate the transformative potential of IoT in healthcare, particularly for real-time monitoring and remote care. However, challenges such as system scalability, data privacy, and cloud infrastructure management remain areas for improvement. The first phase of our project builds on these foundations by focusing on the IoT architecture to ensure reliable data transmission and cloud storage, providing a robust platform for real-time healthcare monitoring.

## III. METHODOLOGY

The proposed Interpretable Machine Learning-based Decision Support System (DSS) follows a structured approach to ensure both accuracy and interpretability in healthcare decision-making. The methodology consists of the following key phases:

1) Data Collection & PreprocessingoReal-world healthcare data is collected from clinical records. oPreprocessing includes data cleaning, feature selection, and normalization to enhance model efficiency.
2) Model Developmento Random Forest Algorithm is employed for classification, ensuring interpretability through feature importance analysis.

- Natural Language Processing (NLP) techniques (e.g., tokenization, stemming, and BERT embeddings) are used for analyzing unstructured medical data.

3) Model Training & ValidationoThe system is trained on historical patient records and validated using k-fold crossvalidation to assess reliability.

- Evaluation metrics include accuracy, F1score, and AUC-ROC for performance assessment.

4) System Integration & Deploymento The model is integrated into Hospital Information Systems (HIS) via REST APIs for real-time decision support.

- The system is tested for real-world usability and efficiency in clinical environments.

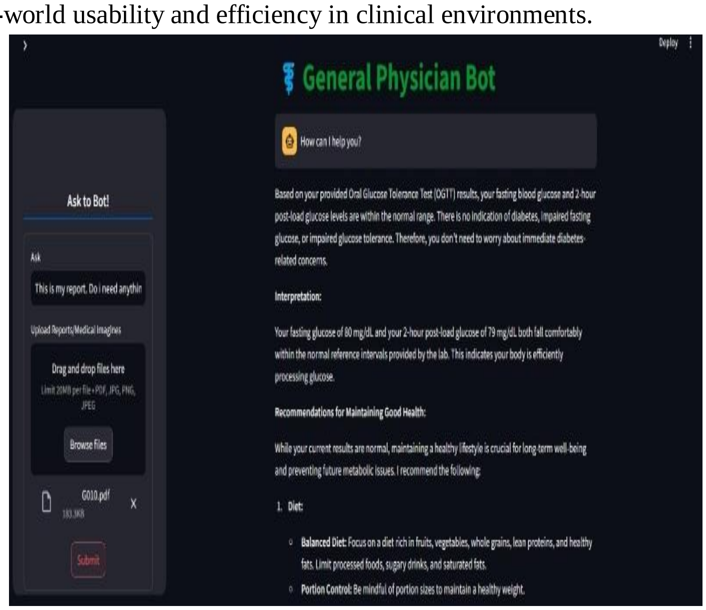



Fig 1: System architecture diagram

TABLE I. USE CASE DIAGRAM

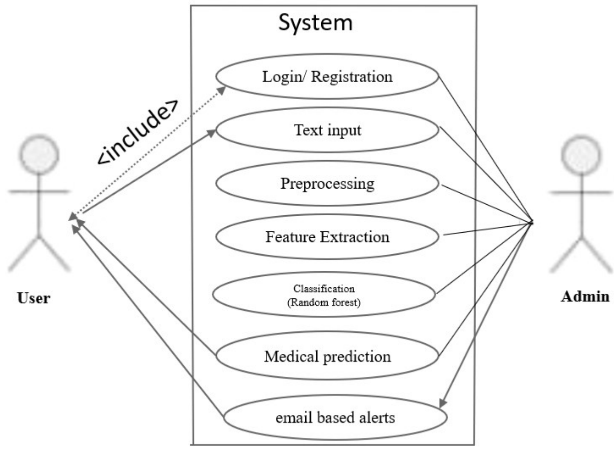
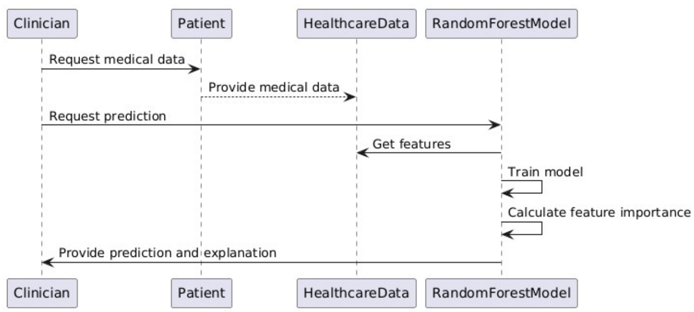



TABLE II. SEQUENCE DIAGRAM




## IV. THE EXISTING SYSTEM AND NEED FOR A NEW SYSTEM

### A. *Existing System*

Current Healthcare Decision Support Systems (DSS) leverage machine learning (ML) and artificial intelligence (AI) to assist clinicians in diagnosis and treatment planning. Many of these systems use deep learning models, support vector machines (SVMs), and neural networks to predict diseases and recommend treatments. However, these models often function as black-box systems, providing accurate predictions without explaining the rationale behind their decisions. The lack of interpretability in such models creates a trust barrier, making it difficult for healthcare professionals to rely on AI-driven recommendations.

Additionally, existing systems face challenges in:

1) Handling Heterogeneous Data – Structured (numerical) and unstructured (textual) medical records require advanced processing.
2) Computational Efficiency – Many AI models are computationally intensive and unsuitable for real-time clinical applications.
3) Integration with Clinical Workflows – Many models lack seamless integration with Hospital Information Systems (HIS), limiting real-world usability.

### B. *Need for a New System*

To overcome these limitations, a transparent and interpretable ML-based DSS is needed. This research proposes a system using the Random Forest algorithm, which balances accuracy with interpretability through feature importance visualization. Additionally, the integration of Natural Language Processing (NLP) allows the model to process unstructured clinical data, improving decision-making capabilities. The system ensures:

1) Enhanced Interpretability – Providing clinicians with clear explanations for AI-driven decisions.
2) Efficient Data Processing – Handling both structured and unstructured data effectively.
3) Real-time Integration – Seamlessly integrating with HIS for clinical usability.

By addressing these challenges, the proposed system aims to build trustworthy, interpretable, and efficient AI-driven healthcare solutions, ultimately improving patient care and clinical decision-making.

Overall, the results demonstrate the successful implementation of an IoT-based health monitoring system capable of real-time data collection and transmission. The system lays a strong foundation for future enhancements, particularly in integrating advanced machine learning algorithms for predictive health monitoring.

## V. PROBLEM DEFINITION

The integration of machine learning (ML) in healthcare decision support systems (DSS) has the potential to improve diagnostic accuracy and patient outcomes. However, a significant challenge in existing ML-based DSS is the lack of interpretability, making it difficult for clinicians to trust and adopt these systems in real-world medical practice. Black-box models provide high accuracy but fail to offer transparent explanations for their predictions, leading to scepticism among healthcare professionals.

Additionally, many DSS struggle with data heterogeneity, where patient records include structured (numerical) and unstructured (textual) data, requiring advanced feature extraction and processing. The inability to effectively handle such data limits the performance and generalizability of current models. Furthermore, clinical environments demand real-time decision support, but existing models often lack the computational efficiency required for seamless integration into hospital information systems (HIS).

To address these challenges, this research proposes an interpretable ML-based DSS using the Random Forest algorithm, which provides both high accuracy and feature importance visualization. The system incorporates natural language processing (NLP) techniques to analyse unstructured medical data and enhance clinical decision-making. By ensuring transparency, efficiency, and real-time integration, this project aims to build a trustworthy AI-driven healthcare solution that empowers clinicians with reliable and explainable insights, ultimately improving patient care.

## VI. CONCLUSION

The proposed Interpretable Machine Learning-based Decision Support System (DSS) addresses the critical challenge of balancing accuracy and interpretability in healthcare decision-making. By utilizing the Random Forest algorithm, the system provides reliable predictions while offering feature importance visualization, enhancing trust among clinicians. The integration of Natural Language Processing (NLP) further improves the analysis of unstructured medical data, ensuring comprehensive decision support.

The model is validated using real-world clinical datasets and evaluated through accuracy, F1-score, and AUC-ROC, demonstrating its effectiveness. Additionally, seamless integration with Hospital Information Systems (HIS) via REST APIs ensures real-time usability.

This research contributes to the development of trustworthy AI-driven healthcare solutions, empowering healthcare professionals with transparent, data-driven insights for improved patient care and clinical outcomes. Future enhancements include optimizing real-time processing and integrating domain-specific NLP models for further accuracy improvements.

In conclusion, the project highlights the transformative potential of IoT in healthcare, demonstrating that innovative technologies can significantly enhance patient monitoring, improve healthcare outcomes, and bridge gaps in access to medical services.

## VII.FUTURE WORK

The successful implementation of the Real-Time E-Healthcare Monitoring System's first phase lays a strong foundation for future enhancements and expansions. The following areas are identified for further development:

### *A. Integration of Additional Sensors*

To enhance the system's capabilities, future work will focus on integrating a wider range of sensors. This may include:

1) ECG Sensors: For continuous monitoring of cardiac activity, providing more comprehensive data for patients with cardiovascular diseases.
2) Glucose Monitors: To support diabetes management by tracking blood sugar levels in real-time.
3) Respiratory Rate Sensors: To monitor breathing patterns, especially for patients with respiratory conditions.

### *B. Advanced Data Analytics*

The integration of machine learning algorithms is crucial for improving predictive analysis capabilities. Future work will include:

1) Predictive Health Monitoring: Developing algorithms that can analyse historical and real-time data to identify trends and predict potential health issues.
2) Anomaly Detection: Implementing advanced techniques to automatically detect deviations in vital signs that could indicate health crises, enabling proactive interventions.

### *C. Enhanced User Interface and Experience*

Improving the user interface of the web application will be essential for better user experience. Future enhancements may include:

*1)* Mobile Application Development: Creating a mobile app for patients and healthcare providers to access health data and receive alerts on-the-go.
*2)* Customizable Dashboards: Allowing users to customize the data they see, improving accessibility and usability.

### *D. Robust Data Security Measures*

Given the sensitivity of health data, ongoing efforts to enhance security will be vital:

*1)* Data Encryption: Implementing stronger encryption protocols for data at rest and in transit to protect against unauthorized access.
*2)* User Authentication: Introducing multi-factor authentication methods to secure user accounts and access to health data.

### *E. Pilot Testing and Evaluation*

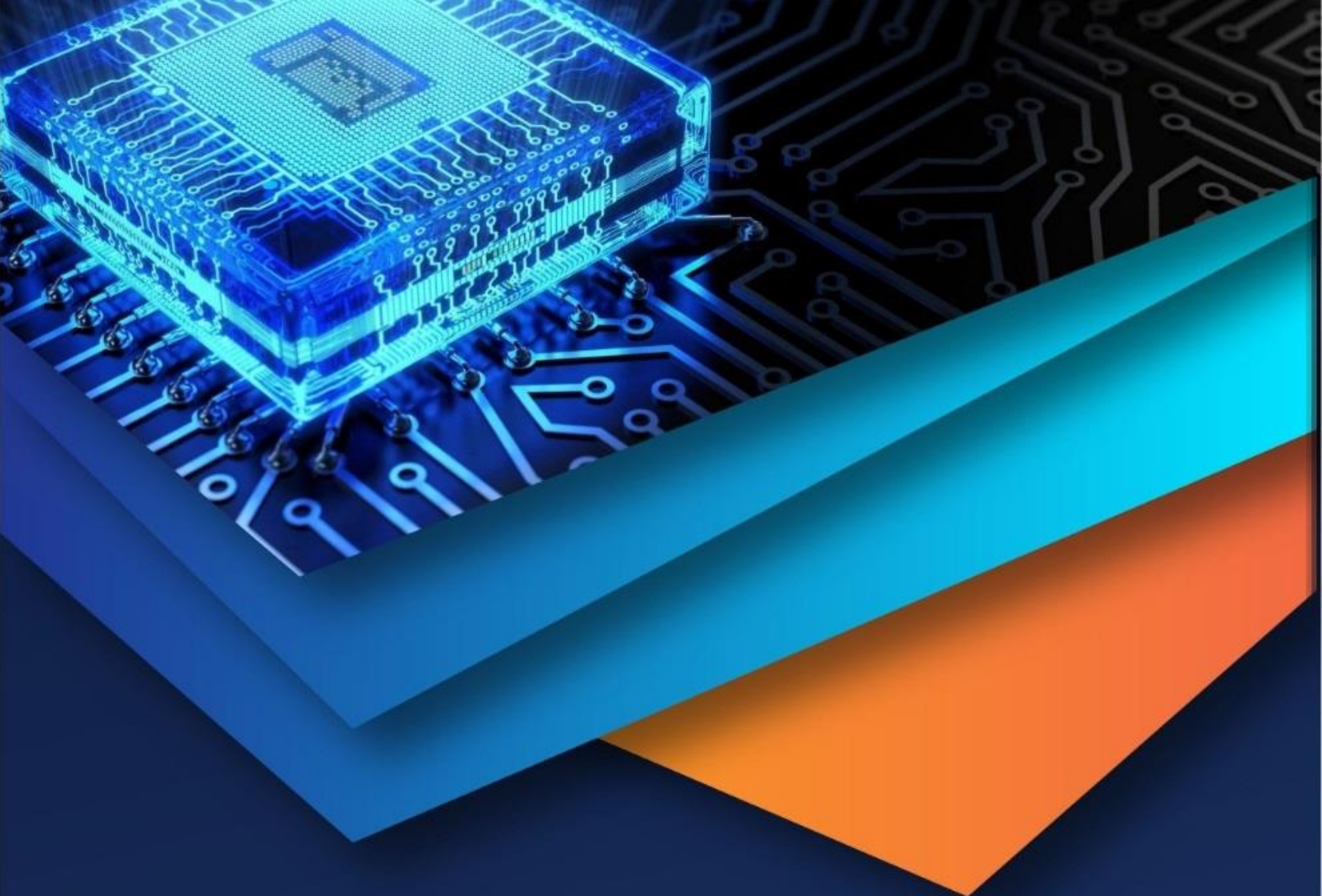
Before a broader rollout, pilot testing in real-world settings will be necessary to evaluate the system's performance and gather feedback from end users. This will involve:

*1)* Collaboration with Healthcare Providers: Engaging healthcare professionals to test the system and provide insights for improvements.
*2)* Feedback Collection: Gathering user feedback to refine the system and ensure it meets the needs of both patients and healthcare providers effectively.

By focusing on these areas, the future phases of the RealTime E-Healthcare Monitoring System aim to enhance its functionality, improve patient outcomes, and contribute significantly to the field of remote health monitoring. This approach will not only broaden the scope of the system but also ensure its adaptability to various healthcare needs and challenges.

## REFERENCES

[1] A New Real Time Clinical Decision Support System Using Machine Learning for Critical Care Unit at Mar 2022 | IRJET
[2] Overview on the Advancements of Support Vector Machine Models in Healthcare Applications at 2024 | MDPI
[3] Bayesian Networks in Healthcare: Distribution by Medical Condition at 2020 | Elsevier
[4] LASSO Regression Modeling on Prediction of Medical Terms among Seafarers' Health Documents at 2022 | MDPI
[5] Classification of Heart Disease Using K-Nearest Neighbor and Genetic Algorithm at 2020 | ScienceDirect
[6] Chinmay Chakraborty; Amit Kishor, "Real-Time Cloud-Based PatientCentric Monitoring Using Computational Health Systems", IEEE Transactions on Computational Social Systems, India, 2022.
[7] Rakhi Bhardwaj; Shiv Narain Gupta; Manish Gupta; Priyesh Tiwari, "IoT based Healthware and Healthcare Monitoring System in India" 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), April 2021. 23
[8] Mahmood. K. Awsaj; Yousif Al Mashhadany; Lamia Chaari Fourati ,"RealTime Healthcare Monitoring and Treatment System Based Microcontroller with IoT" 2023 15th International Conference on Developments in eSystems Engineering (DeSE), India, 2023.
[9] Ahmed Dridi; Salma Sassi; Sami Faiz ,"A Smart IoT platform for personalized healthcare monitoring using semantic technologies" ,2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), 2018.
[10] K. Vijaipriya; C. Priya; Siji Sivanandan; R. Krishnaswamy, "ECG Monitoring System using IoT for Health Care Applications" ,2023 Second International Conference on Augmented Intelligence and Sustainable Systems




10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY