Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Intelligent Health Risk Prediction Framework Using a Hybrid Machine Learning Approach and LLM-Powered Feedback

Pratyusha Chatterjee*[1]**, **Tanima Bhowmik[2]

[1]Computer Science and Engineering, Institute of Engineering & Management, Kolkata, India School of University of Engineering & Management, Kolkata, India.
[2]Computer Science and Engineering, Institute of Engineering & Management, Kolkata, India School of University of Engineering & Management, Kolkata, India
*tanima.bhowmik@iem.edu.in

Abstract

The need for proactive and intelligent healthcare solutions has increased due to the ageing population, chronic disease incidence, and a requirement for real-time patient monitoring, particularly in rural areas. This paper presents an Intelligent Healthcare Monitoring System incorporating fall detection, natural language understanding, prediction of health risk through hybrid machine learning, real-time data gathering, and tracking of emergency events through GPS. An integrated model based on RF and DT is proposed with high precision for predicting and interpretation. RF improves the prediction accuracy, while DT makes decisions transparently. A weighted decision-making strategy is proposed that can be used for handling real-world noise in healthcare data. The model is built around 13 clinically significant features, stored using MQTT and NODE-RED. Preprocessing of the input data with noise removal, elimination of missing values, and data augmentation improves the data quality and generalization of the model. The proposed model is 97.93% accurate in its classification accuracy for the three health risk categories (low, medium, and high) at an 80:20 training-test split. The proposed system uses a chatbot based on Streamlit with a Large Language Model (LLM) to provide conversational, human-readable explanations for health predictions and provides personalized health advice. In the event of a fall or emergency, the system automatically communicates the patient's GPS location to the caregivers or healthcare professionals so that they can respond quickly. It will help elderly care, personal healthcare, and rural healthcare by increasing access to medical care, patient safety, and emergency care using smart technologies.

Keywords: Explainable AI in Healthcare, Fall Detection, Hybrid Machine Learning Model, Health Risk Prediction, IoT in Healthcare, Large Language Model, MQTT Protocol, Real-Time Monitoring, Smart Healthcare System, User Interface.

1.0 Introduction

Intelligent, responsive, and accessible healthcare systems are needed to deal with the growing number of chronic diseases, aging population, and access problems to medical care in remote and underserved areas [1]. The aim of this research project is to create an intelligent healthcare monitoring system that can communicate with patients and doctors to provide timely diagnosis, treatment, and care [2]. The goal of the project is to create and deploy a real-time, intelligent medical monitoring system that can assess a user's health and inform them of health related information [3]. By combining the predictive performance of random forest

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

classification with the ease of decision trees, the model has two strengths: its predictive power and the interpretive power of decision trees. It will detect early warning signs of a patient's health, classify levels of risk as low, medium, and high, and alert the user whether a medical emergency or notifying them immediately of potential problems. The system will also need to be a natural language interface powered by a Large Language Model (LLM) to increase user engagement, a location-based GPS to avoid falls or medical emergencies, and transmitting data securely through IoT protocols such as MQTT. It is also an effort to promote a scalable, modular, and intelligent healthcare infrastructure by providing real-time health information and facilitating a shift from reactive to proactive healthcare, reducing hospital reliance, and improving patient outcomes over time, particularly for the elderly, the chronically ill, and the rural [4]. This research program aims to improve the quality and delivery of future healthcare through innovation and interdisciplinarity [19] [20] .

2.0 Literature Review

Chike Nwibor et al. [5] describe a remote health monitoring device that uses IoT to measure oxygen saturation (SpO2), heart rate, and blood pressure through a single photoplethysmography (PPG) signal and machine learning. Future research is anticipated to focus on user experience and potential therapeutic applications such as monitoring individuals on antihypertensive medications and predicting changes in vital signs.

S. Arun Kumar et al. [6] recommended lowering the mortality rate from lung diseases and premature heart attacks. The aim of this paper is to design and implement an intelligent health monitoring system combining cloud, IoT, and machine learning. Vital signs are collected by sensors, and machine learning predicts 86% of disease types. Wearable sensor technology and telemedicine integration may be utilised in the future to enhance remote health monitoring.

Agrawal et al. [7] demonstrate the use of wireless pressure-sensor insoles to reduce the risk of falls among the elderly. Of the six machine learning models assessed, logistic regression had the highest AUC (0.82), and random forests had the best accuracy (0.82) and specificity (0.88). To enhance fall risk prediction, the subsequent study investigates deep learning methods and novel gait parameters.

Ubaid Ullah et al. [8] examined current developments in healthcare using Quantum Machine Learning (QML). Out of 2,038 articles published between 2018 and 2023, 49 are found to be relevant. Most focus on QML algorithms that utilise photos and patient data from EHRs. The QRF, QKNN, and Qsvm models are known to be used to solve healthcare issues. The difficulty of conducting a thorough search and the lack of precise quantum data are two disadvantages.

Salini et al. [9] stated that pregnancy issues require prompt attention. Manual analysis of cardiotocography (CTG) tests is laborious and imprecise. ML has great potential despite its drawbacks. Care for expectant mothers and fetuses can be enhanced by interdisciplinary teamwork and technological innovation.

Sengupta et al. [10] investigated how ML and wearable technology can be combined to treat strokes. This collection of research, which includes studies in six categories such as activity recognition and clinical score estimation, spans the years 2009 to 2023. By analysing the advantages, disadvantages, and potential applications of wearable sensors and machine learning for post-stroke mobility analysis, the study seeks to guide researchers in this area.

Shanthakumara et al. [11] investigate how to incorporate IoT and ML into health monitoring, with a particular emphasis on wearable devices such as smartwatches. IoT-collected data, including heart rate, blood pressure, and temperature, is used by ML algorithms to examine medical conditions. According to the exhibition examination, the algorithm accuracies are 75% for KNN, 86% for SVM, 83% for Naive Bayes, 74% for

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Decision Tree, and 81% for Logistic Regression. Future developments aim to increase the number of sensors and to develop more complex machine learning algorithms, with a focus on deep learning.

M. Muhammad Arslan et al. [12] examine developments in vital sign monitoring systems that incorporate wireless communication, sensor technology, and machine learning (ML) algorithms. Research indicates that machine learning models such as KNN (92% accuracy), SVM (94% accuracy), and SGD (97% accuracy) are effective at classifying patients' health status. Real-time monitoring is improved by combining window-based applications with robust error-checking protocols; future work will focus on expanding training datasets and improving diagnostic accuracy.

Abdullah Baihan et al. [13] introduce a wearable health-monitoring system for pilgrims performing the Hajj and Umrah that uses a deep reinforcement learning model, the Deep Q Network (DQN), to make context-aware, adaptive alert decisions. To monitor vitals and guarantee safe distancing based on crowd density, the system uses wristbands embedded with sensors (temperature, SpO2, ToF, and heart rate). The system demonstrated improvements in data analysis rate (14.58%), density detection (16.12%), recommendation accuracy (15.79%), and reduced distortion error (11.57%) when evaluated at 5- to 60-minute intervals.

Jie Zhou et al. [14] describe a flexible, encapsulated, wearable, non-invasive blood pressure (BP) monitoring system that uses piezoelectric micromachined ultrasound transducers (PMUTs). This motion-resistant, small (less than 2g) device uses ultrasonic pulse-echo to measure arterial diameter and has a high signal-to-noise ratio (>29 dB). With a mean absolute error of less than 4 mmHg (systolic) and less than 3 mmHg (d(diastolic) and a heart rate error of less than 3 bpm, it computes pulse wave velocity to estimate blood pressure. Its dependability for continuous, real-time blood pressure monitoring is confirmed by a strong correlation (0.951–0.998) with conventional techniques.

Himi et al. [15] present smartwatch-based "MedAi" multi-disease prediction systems. Using machine learning techniques like K-Nearest Neighbour, Random Forest, and Support Vector Machine, it attains exceptional accuracy (99.4%). After that, research will focus on expanding the dataset, creating the smartwatch, adding more diseases, and launching the mobile app.

G. R. Pradyumna et al. [16] examine the development and importance of the Internet of Medical Things (IoMT) in revolutionising healthcare via remote monitoring, telemedicine, and real-time data collection. It highlights how machine learning, security, privacy, interoperability, and ethical issues are integrated. IoMT adoption is hampered by issues such as scalability and data management, despite advances in technology. The study emphasises the need for sustainable, power-efficient IoMT-enabled Smart Healthcare systems to support reliable, safe healthcare device networks and recommends more efficient, lightweight intrusion detection systems.

3.0 Proposed Model

A hybrid machine learning model that combines the strengths of both random forests and decision trees has been built for the proposed intelligent healthcare monitoring system to effectively assess an individual's health based on multiple physiological and environmental factors. When dealing with complex, real-time health data that often contains noise and fluctuations, the hybrid technique was selected to balance interpretability and predictive performance, as shown in Fig. 1. Because it offers an open, rule-based structure for decision-making, the decision tree algorithm is widely recognised for its transparency and usability. However, it could overfit when applied to huge feature spaces or noisy datasets. The combination of random forest, multiple decision trees, and ensemble learning techniques helps reduce variation and increase accuracy and resiliency. By combining both methods, the hybrid model increases the efficiency of the system through the generalisation ability of a random forest and the interpretability of a decision tree. Thirteen parameters are inputted to the model (age, systolic and diastolic blood pressure, body temperature, accelerometer and gyroscope values on

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

each three axes, UV exposure rate, heart rate (bpm), and SpO2 (%) in Fig. 2. Selected these parameters with care due to their clinical relevance and ability to accurately depict the user's physiological state.

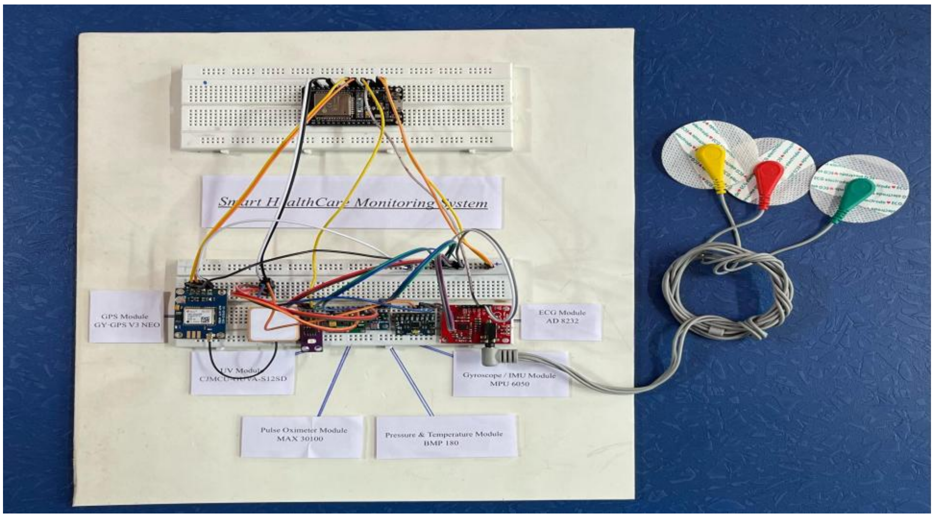

Figure 1. Prototype of Intelligent Health Risk Prediction System

To enhance model generalisation and performance, the data gathered from real-time sensor readings is preprocessed to eliminate noise, normalise outliers, and impute missing values. After preparation, the data is split into training and test sets at an 80:20 ratio. The same dataset is used to train the random forest and decision tree classifiers. To avoid overfitting, the Random Forest model is configured with 100 estimators, and hyperparameters such as the split criterion and depth are optimised. Both models are used to generate predictions for each test sample. When the outputs agree, the shared prediction is maintained. Because of its ensemble nature and better generalisation to unknown data, the random-forest model employs a weighted approach in conflict situations, increasing the confidence of its output. To facilitate a quicker emergency response in the event of a fall or other serious medical condition, the system also includes GPS-based patient location tracking, which records the user's location in real time. The final model creates three categories based on the user's health risk. These classifications are based on predefined thresholds for each feature, derived from medical standards and expert guidelines. Metrics such as accuracy, precision, recall, F1-score, and confusion matrix were used to assess the model's performance. The hybrid model performed better than standalone models with a final prediction accuracy of over 97%. The hybrid model has been serialized using the pickle

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

module of Python, and is integrated in a chatbot from Streamlit. It improves the system's responsiveness and trustworthiness, and can be used for health monitoring in high-risk or remote environments.

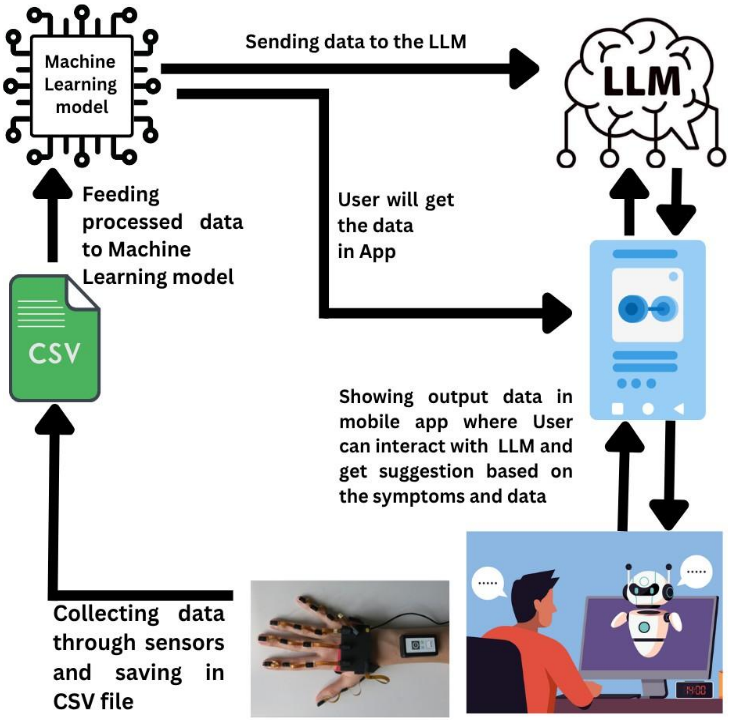

Figure 2. Proposed Intelligent Health Risk Prediction Framework

3.1 Methodology

The Intelligent Healthcare Monitoring System was developed using multiple stages involving real-time data collection, IoT integration, data pre-processing and hybrid machine learning models, alert generation, and deployment using a chatbot interface enhanced by Large Language Model (LLM), shown in Fig. 3. Preventive care is provided by the system with health monitoring, risk classification, and emergency alerts.

Data Gathering in Real-Time from Health Band

A wearable health band is used to gather physiological and environmental health data in real time at the start of the project. Vital signs and parameters included Time in seconds, age, Blood oxygen saturation (%), heart rate (bpm), blood pressure, X, Y, Z Accelerometer readings, X, Y, and Z Gyroscope readings, UV Exposure Level, and Body Temperature in degrees Celsius. These parameters are crucial for assessing a user's physical health as these could indicate illness such as environmental stress, hypoxia, fever, heart failure, over-movement that can lead to falls.

IoT Communication Using MQTT and Node-RED

To provide efficient and lightweight communication between the wearable device and the server, we used the MQTT (Message Queuing Telemetry Transport) protocol via Node-RED. MQTT topics are used to send sensor data to a broker. This data is parsed, logged in CSV and Node-RED subscribes to the topics [17]. The real-time data was analysed using Node-RED dashboards for testing and development. The IoT protocol promises quick data access, reliable data transfer, and the ability to integrate new sensors easily.

Feature Engineering and Data Preprocessing

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

To get it ready for machine learning, the gathered dataset was carefully preprocessed:

- **Noise Handling:** To lessen variability brought on by changes in the environment, raw values were normalised and smoothed.
- **Missing Data:** Incomplete data rows were either discarded or imputed.
- **Synthetic Data Generation:** To simulate real-world sensor inconsistencies, additional samples were generated by adding Gaussian noise.
- **Feature Enhancement:** Using controlled value generation, columns like "Age", "Body Temp (°C)","UV Rate" and "Gyroscope readings" were added or expanded.
- **Random Age Assignment**: Random age values ranging from 18 to 90 were added to replicate a range of age demographics.

The finished dataset, Health_Monitoring_Dataset.csv, was prepared for training and risk labelling.

Classification and Labelling of Health Risks

Based on predetermined medical thresholds, a Health_Risk class was assigned to each data instance:

- **Low Risk (0):** Every parameter falls within the typical range.
- **Mild deviations** (e.g., HR slightly above 90 bpm, SpO\textsubscript{2} 95–96\%) are considered a medium risk (1).
- **Critical abnormalities** (e.g., SpO\textsubscript{2} < 92\%, Temp > 38.5°C, BP > 140/90 mmHg, or excessive gyro/acceleration values) are considered high risk (2).

Each row was classified according to a rule-based function. If several parameters exceeded safe values simultaneously, the second flag was an alert_flag.

Proposed Hybrid Machine Learning Model (Decision Tree and Random Forest)

To accurately classify health states, we combined two tree-based classification models to generate a hybrid ensemble learning model based on both tree-based classification.

- Decision Trees (DTs) are used because of their interpretability, simplicity, and capacity to use if- else rules to model intricate nonlinear relationships.
- To reduce variance, improve accuracy and avoid overfitting, Random Forest (RF) is a set of multiple decision trees trained on different data sets and features.

Model Workflow:

1. By deleting columns such as Time (s), IR Value, Red Value, and Health_Risk, features (X) were extracted.
2. The target label (y) was set to Health_Risk.
3. The train-test ratio was 80:20.
4. We did individual training on both models [18].
5. Weighted approach or majority vote combined their predictions. Random Forest is given more weight as it is more general.
6. We compared our measures of performance (accuracy, precision, recall, F1 score) to those derived from other stand-alone models (e.g., KNN, SVM, Gradient Boosting) [19].

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

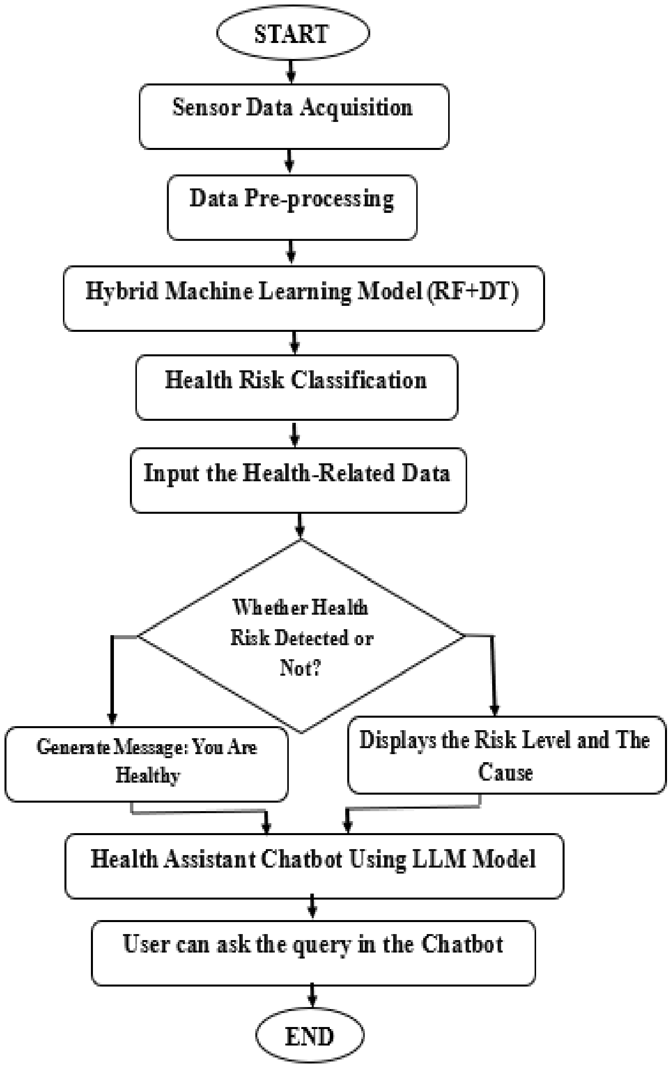

Figure 3. Working Flowchart of Intelligent Health Risk Prediction Framework

7. With over 99% classification accuracy, the hybrid model outperformed the others.
8. Pickle was used to save the finished model so that it could be included in the chatbot.

Intelligent Warning System

One of the primary functions of a sophisticated alert mechanism was to notify me immediately if readings were suspicious:

- Rules for thresholds: Notify patients when a parameter crosses the thresholds of clinical practice (e.g., heart rate 60 or > 100 beats per minute).
- Critical Multi-Condition Alerting: This system will send a critical alert to the user when more than two health conditions have crossed a threshold at the same time. This alerts users of potentially harmful trends that indicate worsening health condition, and of isolated abnormalities..

An Algorithm for Detecting Falls

The logic used to identify falls as a safety risk for senior citizens was based on the following:

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

- Acceleration values on any axis greater than 20 m/s²
- Gyroscope readings that are more than 300°/s A fall is strongly indicated by such a sudden and powerful movement combination. An emergency notification is sent out when the event is flagged by the system.

Online Health Tracking Interface for Chatbots

A chatbot interface built with Streamlit was created to make the system more interactive and easier to use. The following tasks are carried out by the chatbot:

- All vital signs can be entered by the user.
- Predicts health risks by loading the trained hybrid model.
- Shows a positive health assessment message (such as "Healthy", "Mild Risk", or "Critical Risk").
- Enumerates the precise parameters that set off the alarm.
- Alerts users of fall detection.
- Displays recommendations or warnings according to conditions found.

This chatbot empowers users to evaluate their health and take appropriate action using a hybrid model, as shown in Algorithm 1.

Conversational Support Driven by LLM

A Large Language Model (LLM) was incorporated into the system to improve its intelligence and usability to:

- Use natural language to interpret health information and alerts.
- Give conversational advice, such as what to do if you have a fever or low SpO2.
- Respond to user inquiries regarding system usage or health metrics.
- Lower the technical hurdle for users who are not experts.

Algorithm 2 demonstrates that the LLM serves as a virtual health assistant, adding conversational awareness to the analytical backend.

Ngrok Deployment for Remote Access

Without requiring cloud infrastructure, external users can access the local Streamlit app via a public URL to the application's ngrok deployment. This made it possible for:

- User interaction in real time from any device with an internet connection.
- Easy sharing of the monitoring tool with healthcare personnel.
- Demonstration of a prototype without complete cloud hosting.

ALGORITHM 1: Algorithm of Hybrid Model

Let,
$D = \{(p_1, q_1), (p_2, q_2), \ldots, (p_n, q_n)\}$: Dataset of n samples

$p_i \in R^{13}$: Feature vector of 13 health parameters
$q_i \in \{0, 1, 2\}$: Health risk labels (0=LOW, 1= Medium, 2=High)
f_{DT} : Decision Tree Model
f_{RF} : Random Forest Model

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

f_{Hybrid} : Final Ensemble Decision

$LLM()$: Large Language Model Response Generator
$A \subset p$: A Subset of features exceeding safety thresholds
P : Prompt to the LLM
R : Response from the LLM

Step 1: Collected input data $p_i \in R^{13}$ in real time, where:
$p_i = [Age, SpO_2, R, BP, Accel_x, Accel_y, Accel_z, UV, Gyro_x, Gyro_y, Gyro_z, Temp]$

Step 2: Apply preprocessing to D:

- Handle missing data
- Normalize or scale p_i

Step 3: Define function $c(p_i) \rightarrow q_i$ based on medical thresholds: 2, if any critical condition is detected

q_i = 1, if any moderate condition is detected 0, otherwise

Step 4: $D_{train}, D_{test} = TrainTestSplit(D, 80\%\ train, 20\%\ test)$

Step 5: Train both classifiers on D_{train}: $f_{DT} \rightarrow$ Train Decision Tree on D_{train} $f_{RF} \rightarrow$ Train Random Forest on D_{train}

Step 6: For each test instance $\subset D_{test}$: $q_{DT} = f_{DT}(p)$, $q_{RF} = f_{RF}(p)$

Hybrid Prediction Rule:

Final predicted label:

$$f_{Hybrid}(p) = \begin{cases} q_{DT}, & if\ q_{DT} = q_{RF} \\ q_{RF}, & Otherwise\ (RF\ takes\ priority) \end{cases}$$

$q = f_{Hybrid}(x)$

Step 7: Define alert set $A \subset p$:

$A = \{p_j \in p \mid p_j$ exceeds medical thresholds$\}$ Raise:

- alert_flag = 1 if $|A| \geq 1$
- critical_alert = 1 if $|A| > 2$

ALGORITHM 2: Algorithm of Conversational LLM Response

$p_i \in R^{13}$: Input Feature vector
$q \in \{0, 1, 2\}$: Final health risk prediction from the hybrid model
$LLM()$: A pre-trained language model mapping prompts to natural language output
$A \subset p$: Set of alert features where each $p_j \in A$ exceeds the safe threshold
P : A constructed prompt string is given to the LLM

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

R : Final natural language response

Step 1:
Each feature $p_i \in p$ is evaluated against a medical threshold function G_j.
The alert set A is defined as:

$$A = \{p_j \in p \mid p_j > G_{max} \text{ or } p_j > G_{min}\}$$

Where G_{max}, G_{min} are the medically safe bounds for the feature p_j.
Step 2:

Define a formatting function $\emptyset : A \cup \{q\} \rightarrow$ string that converts alerts and risk levels into natural language text. Then:

$$P = \text{The user has the following issues: } + \sum_{p_j \in A} \emptyset(p_j) + \text{"Risk Level"} + \emptyset(q)$$

Where:
Step 3:
$\emptyset(p_j) = p_j = value$, which is outside the safe range.

$$\emptyset(q) = \begin{cases} Low, & q = 0 \\ Medium, & q = 1 \\ High, & q = 2 \end{cases}$$

Given the constructed prompt P, the LLM is a

$$LLM : P \rightarrow R$$

Where $R \in$ NL, the space of natural language explanations.
$R = LLM(P)$

4.0 Results and Discussion

A real-time wearable health monitoring system provided the majority of the dataset used in this investigation. It comprises environmental and physiological factors that are crucial for assessing a user's health. Continuous monitoring with sensor-equipped hardware was used to record these parameters, which were then sent to a central data processing system via the MQTT protocol. Time (s), Age, SpO2 (%), Heart Rate (bpm), Systolic and Diastolic Blood Pressure, Accelerometer readings (X, Y, Z in m/s²), Gyroscope readings (X, Y, Z in °/s), UV Rate, and Body Temperature (°C) are the dataset's primary features as shown in Fig. 4. These attributes were chosen for their clinical relevance in detecting exposures, hazards from falls, respiratory distress, or cardiovascular abnormalities. We classified each record as low, medium, or high risk by thresholds based on the domains. The dataset had sufficient samples to train and assess machine learning models, was balanced, and clean.

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Time (s)	Age	SpO2 (%)	Heart Rat	Systolic Bl	Diastolic B	IR Value	Red Value	Accel X (m	Accel Y (m	Accel Z (m	UV Rate	Gyro X (Â°	Gyro Y (Â°	Gyro Z (Â°	Body Tem	Health_Risk
0	30	98.1988	75.7483	141	116	53812.5	51748.4	-0.3174	0.74534	8.74287	5.02516	-0.508	-4.3489	78.9241	37.6152	1
1	87	96.6846	94.7806	104	101	58730.7	58919	-0.5123	0.55062	9.94163	11.3211	60.4185	75.2918	-37.361	36.1583	2
2	28	99.4414	88.8382	161	68	54522.5	50766.9	0.78226	0.67555	10.0566	8.80959	31.2122	-79.925	-18.982	37.0815	1
3	86	95.3898	63.092	150	72	58543.2	50466.2	-0.8707	-0.1553	9.22535	7.01168	-39.949	9.3552	68.7469	37.693	2
4	68	95.1593	83.8967	110	102	52572	54985.5	-0.4388	0.7113	9.93416	2.05415	88.8169	-48.356	83.6431	36.1622	1
5	55	99.4612	92.215	172	93	56932.4	54077.6	0.32707	-0.4303	9.83944	2.6413	-76.859	20.0699	48.5053	36.5405	1
6	36	98.6421	67.4935	176	75	59484.7	57324.5	-0.8102	0.34297	9.90517	0.48027	14.9584	-94.036	38.5902	37.6362	1
7	64	95.299	71.4206	164	64	50697.4	59467.2	0.5751	-0.5718	9.32143	9.82639	-41.137	15.5067	-14.175	36.1815	1
8	59	97.2829	68.1755	164	60	52411.7	50022.6	0.05373	-0.792	9.27433	7.51849	-73.981	96.1731	-57.988	35.9624	0
9	46	97.1103	68.4421	177	99	54797.6	50683	-0.1253	0.20622	9.19149	8.67421	-56.64	-27.198	-23.921	35.6327	1
10	72	97.9435	61.645	113	81	56782.7	56512.2	-0.7133	0.00234	10.1425	0.82176	-2.5441	-75.607	16.5819	36.3681	1
11	42	96.4813	91.0729	92	107	57666.8	55813.1	-0.5319	0.25984	9.62095	11.9596	17.9994	33.0361	-9.8407	36.4472	2
12	32	97.5893	70.7839	111	105	59564.4	59450.5	0.18918	-0.2086	9.3855	10.3394	53.1915	-15.124	-104.19	35.7688	1
13	73	97.9374	95.2006	142	117	59046.6	58098.8	-0.2904	-0.8641	9.68926	2.28712	73.8882	46.6361	2.66269	37.5987	1
14	43	99.6403	72.2496	91	85	59443.7	58535.2	-0.1646	1.04516	11.1879	1.47458	14.0776	54.7467	54.3731	36.2866	0
15	41	95.4521	66.6105	177	93	50657.2	54078.6	-0.0687	0.52479	9.85319	2.10122	-21.572	51.6909	121.961	36.8728	1
16	74	95.9361	89.8823	119	109	52982.2	54249.2	0.27231	-0.2627	9.69159	3.35263	51.9956	-0.2235	-44.9	36.6341	1
17	58	98.4443	73.5368	127	117	59307.8	55633.5	1.48968	0.04446	9.49514	6.26503	-30.63	91.6728	16.4713	36.9133	0

Figure 4. Dataset of Intelligent Health Risk Prediction Model

Proposed Model

The proposed model combines the features of RF and DT to obtain a hybrid machine learning model. In the hybrid model, interpretation is easily achieved and robustness is assured. In the decision tree, the decision is a path that can be traced. In the random forest, performance is improved by ensemble averaging. A weighted combination of both models is generated by a parameter β. The final prediction is both precise and adaptable. This method significantly improves classification reliability and attains a high degree of accuracy in identifying health risk levels, particularly when dealing with complex or noisy health data.

$$\hat{w} = \beta \cdot \mathrm{RF}_{\mathrm{Pred}} + (1 - \beta) \cdot \mathrm{DT}_{\mathrm{Pred}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

The random forest model is controlled by a weight parameter called β , which affects the outcome prediction.

It's a number between 0 and 1, where

- The decision tree model complete control when β =0.
- The random forest model has full influence when β = 1.

Large Language Model (LLM)

To support the hybrid classifier and enable natural language communication with users, the system incorporates an LLM. The structured input from the hybrid model is converted into conversational, approachable feedback by the LLM. By including a second adjustable parameter α, the system ensures that the final messages are accurate and personalised by combining the LLM's response with pre-existing medical templates. The hybrid predictive engine and the LLM communication lead to an interactive healthcare system that communicates risk and recommendations to the patient, while continuously monitoring key parameters.

$$R_{\mathrm{Final}} = \alpha \cdot L + (1 - \alpha) \cdot T \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

- T : Rule-based template response
- L : Response generated by the Large Language Model (LLM)
- α ∈ [0,1]: Weight parameter controlling the influence of the LLM

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302
Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

- R_{Final} : Final output response shown to the user
-

Table 1. Different Machine Learning Algorithms' Effectiveness in Health Classification

Algorithm	Accuracy	Precision	Recall	F1 Score	Support
KNN	66.95	0.5600	0.4824	0.5009	10000
Decision Tree	87.51	0.6128	0.6455	0.6264	10000
Naïve Bayes	75.21	0.8234	0.5572	0.6122	10000
Random Forest	79.23	0.5858	0.5315	0.5359	10000
Support Vector Machine	62.3	0.5873	0.7060	0.5960	10000
Logistic Regression	69.72	0.6262	0.4968	0.5243	10000
Proposed Model	**97.93**	**0.9510**	**0.9872**	**0.9672**	**10000**

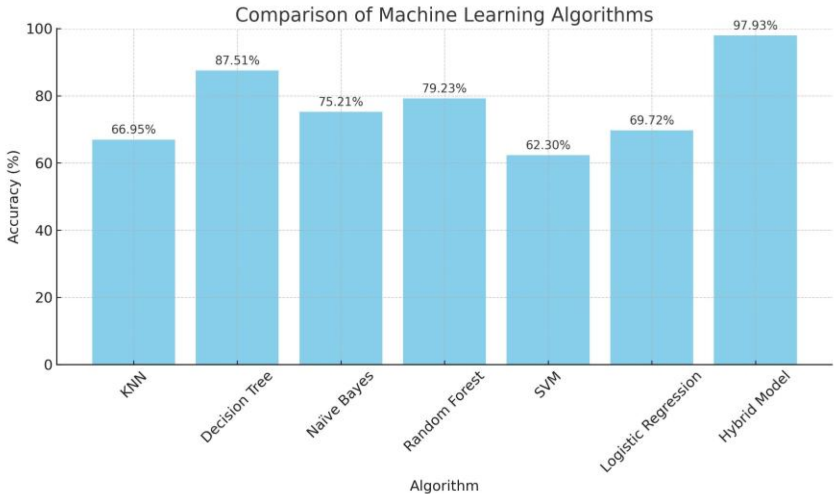

Figure 5. Different Machine Learning Algorithms' Effectiveness in Health Classification

The hybrid health risk classification model, which combines RF and DT classifiers, performs exceptionally well compared to other existing algorithms in predicting health risk levels. As demonstrated in Table 1, the model achieves an impressive 97.93% accuracy, accompanied by outstanding precision, recall, and F1 Scores, as shown in Fig. 5. The combination of RF ensemble averaging and DT transparency ensures a reliable, interpretable model that can handle complex, noisy health data. Because it provides reliable decision paths and high predictive accuracy, this hybrid approach performs incredibly well for health risk classification in real-world applications.

A comparison of training and test accuracy is essential for assessing the performance and generalisation capacity of a hybrid model combining Random Forest and Decision Tree classifiers. Overfitting, a situation in which the model performs well on training data but is unable to generalise to unseen data, may be indicated by a noticeable difference between training and test accuracies, as shown in Fig. 6. The testing accuracy of 97.93% in this instance, however, indicates that the model generalises well and that there is little variation between the training and testing accuracies. The robustness of the model's performance in actual health risk prediction scenarios is confirmed by this high accuracy on the test set.

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

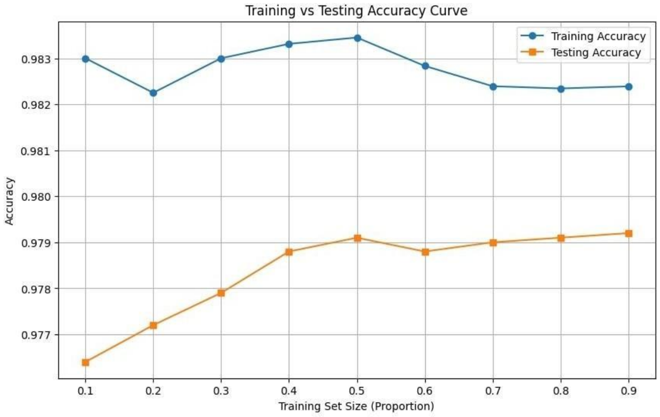

Figure 6. Training vs. Testing Accuracy of Hybrid Model

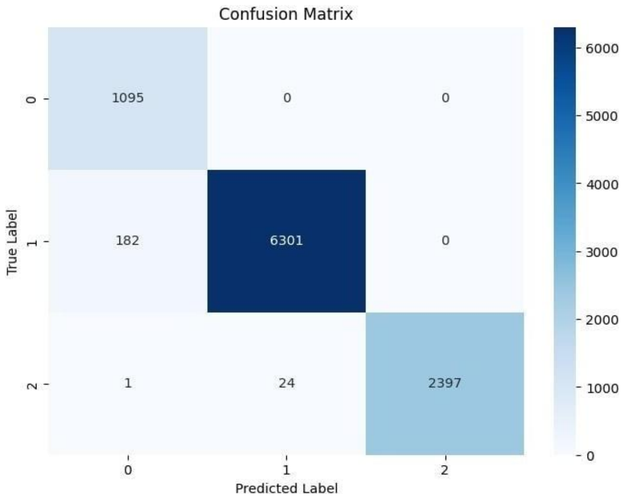

Figure 7. Confusion Matrix of Machine Learning Model

To support the hybrid classifier and enable natural language communication with users, the system incorporates an LLM. The structured input from the hybrid model is converted into conversational, approachable feedback by the LLM. By including a second adjustable parameter α, the system ensures that the final messages are accurate and personalised by combining the LLM's response with pre-existing medical templates. The hybrid predictive engine and LLM-based communication result in an intelligent, interactive healthcare system that effectively communicates risk and recommendations to the user while monitoring key parameters.

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

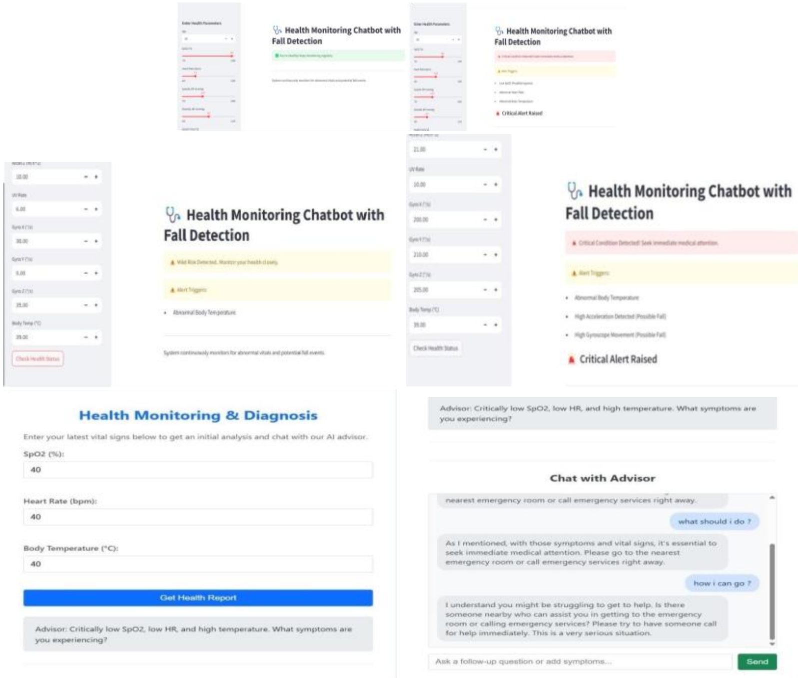

Figure 8. User Interface of Intelligent Health Risk Prediction System

The confusion matrix, which measures the performance of the model across classes, supports this conclusion. Fig. 7 shows the percentages of cases correctly or incorrectly classified at each health risk level (0, 1, 2) in terms of the health risk class (correctly or incorrectly). The model produces accurate predictions for all classes when the confusion matrix has strong diagonal dominance, or high values along the diagonal. Fewer misclassifications are reflected in minimal off-diagonal values, demonstrating the hybrid model's ability to reliably and consistently distinguish among various health risk levels.

The health monitoring system's output interface is made to help users easily comprehend and act upon complex health data. The three health risk levels, low, medium, and high and colour codes for easy identification are displayed. The system identifies which of the parameters, such as SpO2, heart rate, blood pressure, and body temperature, etc., is responsible for the elevated risk for each risk level. Fig. 8 illustrates an example of when the system makes clear that high blood pressure puts the user at risk. The system provides clear instructions by expressing the risk through natural language feedback from a Large Language Model (LLM) as well as the risk level. Based on the identified risk factors, the interface also suggests changes in the user's lifestyle or consultation with a doctor.

5.0 Conclusion

This paper describes the development of an intelligent healthcare monitoring system that can identify falls, assess health in real-time, and engage with users on an intuitive level . The system collects physiological and environmental data. A hybrid ML model utilizing RF and DT classifiers was created to predict a person's health risk. It was able to predict low, medium, and high risk conditions with an accuracy of 97.93%. An LLM was added to the system to provide conversational explanations based on the predicted condition. In emergencies that call medical attention is required, this feature is important. The system is ideal for personal healthcare, senior monitoring, and remote medical support that provides health monitoring, intelligent decision-making, and user-centric communication within one platform . The LLM improves the user's understanding and engagement by translating structured outputs into readable feedback. The system's location-aware capabilities allow the system to determine the user's current location in the event of a fall or unexpected illness. In an emergency requiring prompt medical attention, this feature is essential. The system is ideal for personal healthcare, serial monitoring, and remote medical support since it successfully integrates health monitoring, intelligent decision-making, and user- centric communication into a single platform.

Acknowledgment

This work was supported by the IEM Grant In Aid project and the Centre of Excellence for Advanced Communication & Intelligent Computing.

Authors'Declaration

All authors declare that they have no conflict of interest. There are no financial, professional, or personal relationships that could have influenced the work reported in this paper.

Authors' Contribution Statement

Pratyusha Chatterjee and Tanima Bhowmik wrote the main manuscript text and prepared all figures. All authors reviewed the manuscript.

References

[1] A. Rahman, T. Debnath, D. Kundu, M. S. I. Khan, A. A. Aishi, S. Sazzad, M. Sayduzzaman, and S. S. Band, "Machine learning and deep learning-based approach in smart healthcare: Recent advances, applications, challenges, and opportunities," AIMS Public Health, vol. 11, no. 1, pp. 58–109, 2024.

[2] R. Ferdousi, M. A. Hossain, and A. El Saddik, "Early-stage risk prediction of non-communicable disease using machine learning in health," IEEE Access, vol. 9, pp. 96823–96837, 2021.

[3] P. C, V. Kalpana, S. S. V. Evangelin, A. Pushpalatha, G. Mahendran, and S. Sivarajan, "Development and implementation of an intelligent health monitoring system using IoT and advanced machine learning techniques," Journal of Machine and Computing, vol. 3, no. 4, pp. 456–464, 2023.

[4] D. K. Jain, K. Srinivas, S. V. N. Srinivasu, and R. Manikandan, "Machine learning-based monitoring system with IoT using wearable sensors and preconvoluted fast recurrent neural networks (P-FRNN)," IEEE Sensors Journal, vol. 21, no. 22, pp. 25517–25524, 2021.

[5] C. Nwibor, S. Haxha, M. M. Ali, M. Sakel, A. R. Haxha, K. Saunders, and S. Nabakooza, "Remote health monitoring system for the estimation of blood pressure, heart rate, and blood oxygen saturation level," IEEE Sensors Journal, vol. 23, no. 5, pp. 5401–5411, 2023.

[6] S. A. Kumar, B. Gopinath, A. Kavinraj, and S. Sasikala, "Towards improving patient health monitoring system using machine learning and Internet of Things," in Proc. Int. Conf. Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), 2021, pp. 1–5.

[7] D. K. Agrawal, W. Usaha, and S. Pojprapai, "Fall risk prediction using wireless sensor insoles with machine learning," IEEE Access, vol. 11, pp. 23119–23126, 2023.

[8] U. Ullah and B. Garcia-Zapirain, "Quantum machine learning revolution in healthcare: A systematic review of emerging perspectives and applications," IEEE Access, 2024.

[9] Y. Salini, S. N. Mohanty, J. V. N. Ramesh, M. Yang, and M. M. V. Chalapathi, "Cardiotocography data analysis for fetal health classification using machine learning models," IEEE Access, 2024.

Received: 2026-05-04
Accepted: 2026-07-10
Published Online: 2026-07-20
DOI: 10.70454/JRICST.2026.030302

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

[10] N. Sengupta, A. S. Rao, B. Yan, and M. Palaniswami, "A survey of wearable sensors and machine learning algorithms for automated stroke rehabilitation," IEEE Access, 2024.

[11] V. Yogesh and A. H. Shanthakumara, "Smart real-time health monitoring band using machine learning and IoT," in Proc. Int. Conf. Smart Electronics and Communication (ICOSEC), 2021, pp. 43–46.

[12] M. M. Arslan, X. Yang, Z. Zhang, S. U. Rahman, M. Ullah, and Q. H. Abbasi, "Advancing healthcare monitoring: Integrating machine learning with innovative wearable and wireless systems for comprehensive patient care," IEEE Sensors Journal, vol. 24, no. 18, pp. 29199–29210, Sep. 2024, doi: 10.1109/JSEN.2024.3434409.

[13] A. Baihan, M. Amoon, T. Altameem, and M. Hashem, "Pioneering wearable sensor-driven health-monitoring system for contagious disease prevention with intelligent crowd-counting models," IEEE Sensors Journal, vol. 25, no. 4, pp. 7403–7416, Feb. 2025, doi: 10.1109/JSEN.2024.3524279.

[14] J. Zhou et al., "Continuous monitoring of blood pressure by measuring local pulse wave velocity using wearable micromachined ultrasonic probes," IEEE Transactions on Biomedical Engineering, 2024, doi: 10.1109/TBME.2024.3514878.

[15] S. T. Himi, N. T. Monalisa, M. D. Whaiduzzaman, A. Barros, and M. S. Uddin, "MedAi: A smartwatch-based application framework for the prediction of common diseases using machine learning," IEEE Access, vol. 11, pp. 12342–12359, 2023.

[16] G. R. Pradyumna, B. H. Roopa Hegde, K. B. Bommegowda, T. Jan, and G. R. Naik, "Empowering healthcare with IoMT: Evolution, machine learning integration, security, and interoperability challenges," IEEE Access, 2024.

[17] S. R. Janani, R. Subramanian, S. Karthik, and C. Vimalarani, "Healthcare monitoring using machine learning based data analytics," International Journal of Computers Communications & Control, vol. 18, no. 1, 2023.

[18] N. Mohan, D. Abdelrahman, N. F. Ali, and M. Atef, "An integrated high-gain wide-dynamic range photoplethysmography sensor for cardiac health monitoring," IEEE Sensors Journal, 2024.

[19] T. Bhowmik, D. Mitra, P. Chatterjee, N. Roy, and S. Kundu, "Development of intelligent healthcare monitoring system using machine learning model," in Proc. IEEE 5th India Council Int. Subsections Conf. (INDISCON), Chandigarh, India, 2024, pp. 1–6, doi: 10.1109/INDISCON62179.2024.10744307.

[20] N. S. Tallapaneni, M. Jayanthi, and M. Venkatesan, "IoT-based smart health care system monitoring using machine learning techniques," in Proc. Int. Conf. Systems, Computation, Automation and Networking (ICSCAN), 2021, pp. 1–6.