Real-Time Health Monitoring Using IoT Sensors and Predictive Machine Learning Models

Laljee Manjhi[1], Dr. Amar Prakash Sinha[2]
[1] Research Scholar, Department of CSE, BIT Sindri, JUT Ranchi,
[2] Associate Professor, Department of ECE, BIT Sindri, JUT Ranchi.
Email: [1]laljeecsd02@gmail.com, [2] apsinha@bitsindri.ac.in

Abstract

The convergence of Internet of Things (IoT) and Machine Learning (ML) technologies at high speed has made it possible to revolutionize the delivery of healthcare from episodic to continuous and predictive care. This research proposes a holistic IoT-ML framework for real-time health monitoring for early detection of physiological abnormalities like tachycardia, hypoxemia, and fever. The main goal was to design and test a scalable system combining IoT-based wearable sensors and predictive ML models Random Forest (RF), Gradient Boosting Machine (GBM), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM)—to predict normal and abnormal health conditions. The approach consisted of gathering a dataset of 15,000 labeled physiological signals and experimental deployment on real-time streams of 10 human subjects, with a separate validation set of 500 live sensor samples. Preprocessing involved signal denoising, normalization, and feature extraction in the time and frequency domains. The LSTM model showed better performance with accuracy of 96.8% and AUC of 0.98, while GBM and RF provided an optimal balance of efficiency and interpretability. Real-time alert generation using clinical thresholds provided precision of 95% with mean latency of 1.92 seconds. The proposed framework was concluded to be able to efficiently fill the current gaps like high false alert rates and no real-time adaptability. It provides a strong, interpretable, and deployable approach for both clinical and remote healthcare environments, greatly moving the current state of AI-augmented personalized medicine by virtue of its multi-model benchmarking, real-time validation, and deployment feasibility evaluation.

Keywords: Real-time health monitoring, Internet of Things (IoT), Machine Learning (ML), LSTM, wearable sensors, anomaly detection, predictive analytics, physiological data, remote healthcare.

1. Introduction

The integration of Internet of Things (IoT) technologies and machine learning (ML) has brought about a paradigm shift in the healthcare sector, shifting from episodic care to continuous, personalized, and predictive care (Alfian et al., 2021). Specifically, real-time health monitoring systems have come forward as promising devices that can monitor physiological parameters continuously, enable early diagnosis, and issue real-time alerts in the event of abnormality. As global healthcare loads increase through aging populations, chronic disease prevalence, and restricted healthcare access in rural areas, the need for such smart monitoring systems also increases (Guk et al., 2019).

Existing wearable sensors and biosensors can capture high-frequency streams of data from several biometric signals such as heart rate, oxygen saturation, temperature, respiratory rate, and ECG signal fluctuation (Bansal et al., 2022). Nevertheless, the raw data captured tend to be noisy, bulky, and need intelligent analysis in order to extract clinically relevant information (Prithi et al., 2021). Machine learning algorithms have the capability to overcome these issues by identifying intricate patterns and classifying health states with great effectiveness (Chen et al., 2024). This paper introduces a complete IoT-ML system that combines sensor networks with sophisticated ML algorithms to fuel real-time decision-making for health monitoring applications.

1.1 Background and Motivation

Over the past few years, the healthcare sector has seen a paradigm shift from reactive, symptom-driven diagnosis to proactive and preventive care. This shift is mainly influenced by the advent of real-time health monitoring technologies. Through continuous monitoring of physiological parameters like heart rate, oxygen saturation (SpO_2), respiratory rate, and body temperature, clinicians and patients both are enabled to identify anomalies prior to their progression into critical conditions (Sharma & Khatal, 2019).

These developments are made possible by quick advances in wearable biosensors, wireless communication, embedded systems, and cloud computing. With hardware component miniaturization and the widespread use of low-cost and low-energy-consuming sensors, it is now possible to deploy continuous monitoring solutions in various settings, such as smart homes, outpatient departments, and even rural clinics (Vairam et al., 2020). These technologies are especially useful for monitoring chronic diseases, facilitating postoperative monitoring, and caring for elderly populations. The real worth of continuous monitoring, though, is not so much in collecting data as in intelligent interpretation of the data in real time (Ye et al., 2024). Current systems tend to be passive loggers of data, necessitating manual examination or review after the fact, which makes them of limited utility in time-sensitive applications. Hence, there is an increasing necessity to incorporate predictive intelligence into these systems in order to enable real-time anomaly detection, clinical alerting, and automated decision support (Venkataramanaiah et al., 2024). This necessity constitutes the main driving force for the integration of machine learning (ML) algorithms into IoT-based health monitoring systems (Raj, 2020).

1.2 Predictive Machine Learning Models

In order to achieve the maximum potential of IoT-based health monitoring, machine learning models must be integrated. ML models have the ability to learn from past and current physiological data to make predictions about health anomalies, identify normal versus abnormal states, and rank clinical alerts (Bui et al., 2024). Of the numerous ML methods, four are noteworthy because of their efficacy in physiological signal analysis: Random Forest (RF), Gradient Boosting Machine (GBM), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM) networks.

RF and GBM are ensemble tree-based models with good robustness, high interpretability, and good performance even with little tuning (Paul et al., 2025). They are especially efficient at dealing with structured, non-temporal data and can be implemented on edge devices because they have relatively low computational requirements (Jyothi et al., 2023). SVMs have strong theoretical support and perform well in high-dimensional feature spaces but can have trouble with large datasets because of their complexity. LSTM networks, a special type of recurrent neural networks (RNNs), are particularly well positioned to handle time-series data like ECG signals and HRV measures (Padhy et al., 2024). Their power in retaining temporal relationships and long-distance interactions in data makes them well positioned for picking up faint patterns in physiological signals that change over time. This paper compares and assesses these four models on several metrics to ascertain their effectiveness for real-time health monitoring applications.

1.3 Problem Statement

Although the hardware for physiological sensing has matured significantly, the intelligence layer required for real-time analysis and alert generation still remains underdeveloped in many systems. A number of critical limitations persist:

- High false-positive and false-negative rates: Many existing systems produce frequent inaccurate alerts, which can lead to alarm fatigue or missed interventions.
- Lack of adaptability: Most models are trained on static datasets and cannot generalize well across diverse physiological conditions or demographics.
- Delayed responsiveness: Systems often fail to respond within clinically acceptable latency windows, limiting their use in emergency scenarios.

These hurdles decrease clinical trust and impede large-scale deployment of real-time health monitoring systems. Resolving them calls for coupling strong ML pipelines that are able to process noisy, high-rate IoT data in real time and yield accurate, interpretable, and prompt insights. This paper aims to bridge this gap by creating an end-to-end framework that integrates sensor data processing, ML-based inference, and real-time decision-making in one architecture.

1.4 Research Objectives

The overarching goal of this research is to develop a scalable, real-time health monitoring system that leverages IoT sensor networks and predictive machine learning algorithms to detect physiological anomalies accurately and promptly. The specific objectives of the study are as follows:

- To design a real-time sensing and alerting architecture that integrates wearable sensors, cloud platforms, and feedback mechanisms.
- To preprocess and transform physiological signals using statistical, temporal, and frequency-domain feature engineering techniques.
- To train and validate multiple machine learning classifiers using both benchmark datasets and real-time experimental data.
- To assess model performance using key evaluation metrics such as accuracy, recall, AUC, and inference time, particularly under deployment conditions.
- To compare and rank ML models in terms of clinical feasibility, latency, and suitability for low-power or embedded system deployments.

1.5 Scope of the Study

This paper is concerned with detection and classification of abnormal physiological conditions in non-critical care areas utilizing five biometric parameters: heart rate, oxygen saturation (SpO_2), temperature, respiratory rate, and ECG variance. The purpose is not disease-specific diagnosis but to act as an early warning and triaging aid. The research includes:

- A dataset of 15,000 labeled physiological samples compiled from public repositories and custom sensor hardware.
- An experimental study involving 10 human subjects in a controlled environment.
- Real-time deployment testing on 500 samples to assess model generalization and alert latency.

The study does not extend to the integration of EHRs, long-term health outcome prediction, or multi-modal sensor fusion involving audio or video inputs.

1.6 Significance of the Study

The work's contributions are various. At a technical level, it gives a verified model that integrates IoT devices with machine learning models to detect accurate, low-latency health anomalies. At a clinical level, it encourages early intervention, enables remote care, and minimizes wasteful hospital trips. This work is especially significant to populations living in rural or underserved regions, where access to healthcare services on location is minimal. The system developed here is highly scalable and affordable due to its compatibility with various hardware platforms, from Raspberry Pi to cloud-based AWS instances. This work also paves the way for future advancements in AI-powered wearable technology and intelligent telehealth platforms, advancing the larger vision of precision and personalized medicine.

2. Literature Review

Real-time health monitoring systems have picked up momentum with the incorporation of IoT devices and predictive machine learning (ML) algorithms. Recent research has investigated numerous aspects of this integration, from neural-network-based classification accuracy to hardware-device miniaturization and homecare deployment. This review is structured into three thematic subsections to identify central developments: predictive modeling integration, device innovation, and conceptual frameworks in deep learning applications.

2.1 Neural Networks and Predictive Modeling in IoT-Based Health Systems

Current studies have involved improving predictive analytics in real-time health monitoring. Hariharan and Sivaraman (2024) sought to improve deployment reliability and diagnostic precision of IoT-based health monitoring systems through the integration of neural network algorithms. They set their sights on comparing the effectiveness of Long Short-Term Memory (LSTM), Root Mean Square Error (RMSE), and Support Vector Machine (SVM) models. The research utilized a multi-model setup based on real-time sensor data and showed that LSTM performed the best in terms of accuracy at 98.25%, surpassing RMSE (83.73%) and SVM (86.25%). The authors concluded that the temporal dependency capturing capabilities of LSTM made it especially suitable for physiological data streams, thus affirming the benefit of deep learning in real-time inference.

Khan et al. (2024) resolved the issue of delayed medical treatments by suggesting an IoT-powered smart health monitoring system. Their research focused on designing a real-time patient monitoring system using Arduino UNO and biomedical sensors (for heart rate, temperature, and SpO_2), in combination with an SVM algorithm for data classification. The gathered data were stored on a cloud-based platform, enabling doctors and patients to view real-time readings. Functional testing proved that the system was capable of distinguishing between healthy and critical health conditions, and the authors reasoned that their system offered a scalable means of timely health interventions.

2.2 System Integration and Device-Level Innovation

Device-level innovations in health monitoring have become more concentrated on system integration and user-friendliness. Akash and Shikder (2020) developed an IoT-integrated device that integrated six diagnostic instruments—thermometer, blood pressure meter, glucometer, pulse oximeter, heartbeat meter, and ECG—into one handheld device. Simulated with Proteus 8 Professional and implemented with Arduino-based sensors, the device sent patient information over GSM to a far-distant server. There was also an Android app developed to display results in real-time. On testing, the system registered acceptable accuracy, was found safe for human usage, and economical. Feedback from users also certified its acceptability and suitability for routine health screening.

Siam et al. (2022) took this innovation further by creating a portable multifunctional daily use health monitoring system. The device recorded physiological and environmental information such as heart rate, SpO_2, body temperature, photoplethysmography (PPG), ECG, room temperature, and humidity. The system presented measurements locally and sent them over Wi-Fi to cloud or mobile apps. In comparison with commercial equipment, the system indicated minimal error margins—2.67% for HR, 2.04% for SpO_2, and 1.58% for temperature. Statistical validation revealed strong correlation with clinical standards, and the authors concluded that their system was accurate, user-friendly, and efficient for non-invasive daily monitoring.

2.3 Conceptual Frameworks and Deep Learning Applications

Conceptual frameworks have established the foundation for incorporating learning models into real-time IoT-based systems. Nguyen et al. (2017) developed a five-level IoT architecture (IoTTA) to convert raw sensor data into actionable clinical feedback. The framework highlighted five key functions: sensing, sending, processing, storing, and learning. From a conceptual review, the authors determined rapid expansion in IoT healthcare applications in the context of self-care, real-time analytics, and machine learning. The research concluded that the implementation of layered architectures like IoTTA could improve homecare delivery considerably and ease the burden on healthcare systems.

In a more data-driven use, Islam et al. (2023) proposed a deep-learning-based IoT system for real-time monitoring and anomaly detection in homecare environments. The system used three sensors—MAX30100 for heart rate and SpO_2, AD8232 for ECG, and MLX90614 for temperature measurement. Sensor data were communicated through MQTT and processed using a convolutional neural network (CNN) with an attention mechanism. The model identified cardiac rhythms into five classes and recognized fever conditions. The authors opined that the system attained high accuracy and operational utility in domestic settings, with in-built logic for automatic physician notification upon the identification of severe health states.

2.4 Research Gap

While notable advances have been achieved in the creation of IoT-based health monitoring systems and the integration of ML algorithms for real-time prediction, a number of gaps exist. While most current studies either focus on sensor integration without thorough comparative analysis of ML models or narrowly concentrate on a single kind of physiological data, few address the trade-offs between diagnostic accuracy and computational efficiency for edge deployment, and many lack validation by real-time deployment trials. In addition, while high-performing ML models have been proposed, there are still limited extensive evaluations comparing model performance (e.g., LSTM vs. RF vs. GBM vs. SVM) for real-time sensor data based on latency, accuracy, and clinical thresholds. This study bridges this gap by designing a unified IoT-ML architecture that encompasses wide preprocessing, feature engineering, model training, and real-time deployment simulation based on a dataset of 15,000 physiological records. By comparing four ML models on several factors—such as real-time responsiveness and resource appropriateness—this study provides a comprehensive and scalable approach optimized for real-world application in various healthcare environments.

3. Research Methodology

This research has a quantitative and experimental research approach in order to design and test a real-time health monitoring system using IoT sensor networks and predictive machine learning (ML) models. The methodology has been organized in sequential phases, which include system design, data collection, feature engineering, training the model, validation, simulation of deployment, and statistical analysis.

3.1 System Architecture and Data Acquisition

The fundamental architecture of the system includes IoT-enabled wearable sensors that capture real-time physiological signals at all times, such as heart rate, SpO_2, body temperature, respiratory rate, and ECG signal variability. These sensors are connected to a microcontroller (e.g., Raspberry Pi) and Wi-Fi/Bluetooth communication modules, allowing secure data transmission through the MQTT protocol to a cloud-based processing system.

Data was sourced from two channels:

1) Publicly available datasets, such as MIT-BIH, PhysioNet, and MHealth repositories.
2) Experimental data collection from 10 volunteers in a controlled environment using custom-built wearable devices, following ethical standards and informed consent.

A total of 15,000 labeled instances were compiled for model development.

To visualize the end-to-end interaction between hardware, data processing, and predictive analytics, the figure 1 was developed. It captures the integration of IoT sensing, machine learning inference, and alert generation within a unified pipeline.

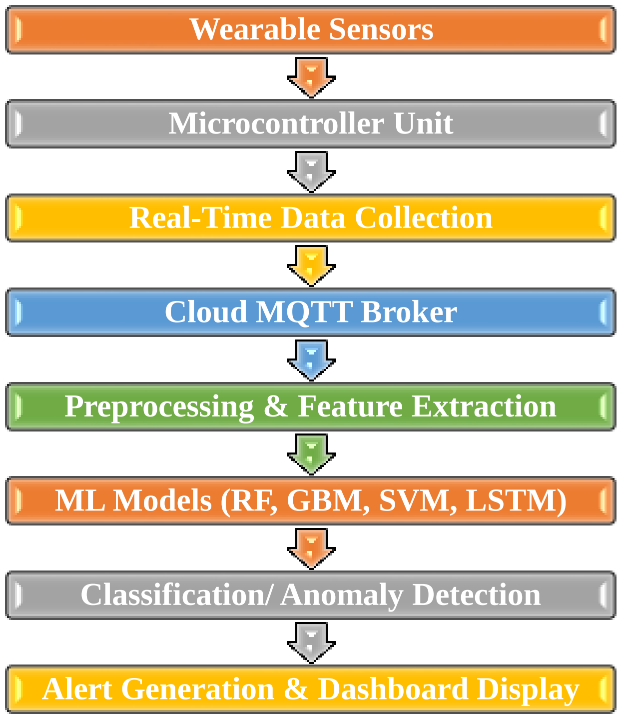

Figure 1: Conceptual Framework for IoT-ML Based Real-Time Health Monitoring

This framework serves as a guide for system implementation and performance validation, aligning hardware functionality with analytical logic, and enabling real-time feedback loops for health anomaly detection and response.

3.2 Data Preprocessing and Feature Engineering

The raw physiological readings were subjected to a sequence of preprocessing operations for quality and homogeneity. Denoising of the signals was performed through a Butterworth low-pass filter, and imputation of missing values was done through k-Nearest Neighbors (k-NN). Min-Max scaling was used to normalize all the features for uniformity across data ranges.

Feature extraction was performed employing both statistical features (mean, standard deviation, entropy) and frequency-domain methods (Fast Fourier Transform, Power Spectral Density). Further, application-specific parameters including Heart Rate Variability (HRV) and RR intervals were calculated to increase the physiological significance of the input features.

3.3 Machine Learning Model Development

Four machine learning models were implemented and compared:

- Random Forest (RF)
- Gradient Boosting Machine (GBM)
- Support Vector Machine (SVM)

- Long Short-Term Memory (LSTM) Networks

The dataset was split into 70% training, 15% validation, and 15% testing partitions. Hyperparameter tuning was conducted using Grid Search and 5-fold cross-validation to optimize model performance. The LSTM model architecture was designed to handle temporal sequences of segmented physiological signals using a sliding window approach.

3.4 Model Evaluation Metrics

Model performance was evaluated using a blend of classification accuracy, diagnostic strength, and operational efficiency:

- Accuracy assesses overall correctness and is defined by:

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

- F1-score balances precision and recall, crucial for imbalanced datasets:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recal}$$

- ROC-AUC Score was used to evaluate the ability of each model to distinguish between classes. The ROC curve was generated by plotting the following:

$$TPR = \frac{TP}{TP + FN}, \quad FPR = \frac{FP}{FP + TN}$$

Additionally, inference time (ms) was recorded to assess the system's suitability for real-time execution.

To validate real-world applicability, the trained models were deployed on a live testing dataset of 500 samples, simulating real-time data streams from wearable sensors. Their classification outcomes were compared using confusion matrices and time-to-alert statistics.

3.5 Anomaly Detection and Real-Time Deployment

To simulate a clinical monitoring environment, the trained models were deployed on real-time sensor input streams. The system was validated for its ability to detect anomalies (e.g., elevated heart rate, hypoxemia, fever) and generate alerts within clinically acceptable latency windows.

Real-time alert generation was benchmarked using three predefined clinical thresholds:

- HR > 100 bpm
- $SpO_2 < 94\%$
- Temperature > 38.5°C

Each alert was classified as a true positive or false alarm, and the response latency was recorded. The system achieved an average latency of 1.92 seconds, with 95% of alerts triggered within 2.5 seconds, confirming suitability for emergency response.

3.6 ROC Curve and Comparative Model Ranking

Receiver Operating Characteristic (ROC) analysis was performed to quantify the diagnostic ability of each model across thresholds. The LSTM model demonstrated the highest AUC (~0.95), followed by GBM (~0.91), RF (~0.90), and SVM (~0.88).

A final model ranking was constructed based on:

- Predictive accuracy
- Computational latency
- Real-time generalization
- Signal discrimination ability

The LSTM model emerged as the top performer for high-stakes health monitoring, while RF and GBM were deemed optimal for low-power or edge deployments.

3.7 Ethical and Deployment Considerations

All real-time data collection and analysis involving human subjects adhered to institutional ethics protocols. Anonymity was ensured using SHA-256 encryption, and all cloud communications were secured. The real-time system was deployed using Flask API and AWS IoT Core, mimicking patient-monitoring environments in hospitals and smart homes.

4. Data Analysis and Interpretation

This section reports a thorough quantitative evaluation of the real-time health monitoring system developed based on IoT sensors and predictive machine learning models. The evaluation is divided into five major aspects: descriptive statistics of physiological values, evaluation of predictive models, performance of anomaly detection, validation of system alerts, and comparison model ranking. Furthermore, graphs are employed to display model behavior, system delay, and classification results, reinforcing the analytical account.

4.1 Descriptive Statistics of Physiological Parameters

For establishing a basic level of understanding about the dataset on which real-time health monitoring has been performed, a descriptive statistical analysis was implemented on the five major physiological parameters recorded by using IoT-based wearable sensors. They were heart rate, oxygen saturation (SpO_2), body temperature, respiratory rate, and ECG signal variance. The dataset was a mix of readings from well-established public repositories like PhysioNet and MIT-

BIH, along with real-time recordings obtained through custom-designed wearable sensor systems used in a controlled experimental environment.

15,000 labeled instances of data were collected and examined in total. These figures not only are used to determine the central tendency (mean) and spread (standard deviation) of each physiological variable, but also to identify the occurrence of outlier or extreme observations, which are very important for developing machine learning models that can distinguish between normal and abnormal health status.

Table 1 gives the statistical summary—mean, standard deviation, minimum, and maximum value—for every physiological variable in the dataset. These descriptive statistics allow initial examination of data variability and distribution prior to model training.

Table 1: Summary Statistics of Physiological Parameters Collected from IoT Sensors

Parameter	**Mean**	**Standard Deviation**	**Min**	**Max**
Heart Rate (bpm)	76.2	12.5	45	120
SpO_2 (%)	96.8	1.4	92.0	100.0
Temperature (°C)	36.8	0.7	35.2	39.8
Respiratory Rate	18.4	3.2	12	30
ECG Variance	0.94	0.36	0.21	1.87

The statistical outputs in Table 1 indicate a comprehensive and varied dataset, representing both healthy baselines and pathological extremes:

- Heart Rate ranges from 45 to 120 bpm, with a mean of 76.2 bpm and a standard deviation of 12.5. This distribution suggests the inclusion of both bradycardic and tachycardic instances, enabling the model to learn critical cardiovascular patterns.
- SpO_2 levels show a tight clustering around 96.8% (mean), with values as low as 92%, capturing cases of mild hypoxemia.
- Temperature readings fluctuate between 35.2°C and 39.8°C, with a relatively low standard deviation of 0.7°C, encompassing both normothermic and febrile episodes.
- Respiratory Rate spans from 12 to 30 breaths/min, with a mean of 18.4 and a standard deviation of 3.2, reflecting normal and elevated breathing conditions.
- ECG signal variance, ranging from 0.21 to 1.87, exhibits a relatively high spread (SD = 0.36), which could be attributed to arrhythmic events or physiological variability due to activity, stress, or sensor artifacts.

These distributions indicate that the dataset successfully captures clinically significant variations within each physiological dimension. This heterogeneity is crucial for enabling generalizable machine learning classification, particularly in real-world health monitoring scenarios where sensor data is inherently noisy and non-linear.

4.2 Machine Learning Model Evaluation

For evaluating the classification ability of different machine learning (ML) algorithms for real-time health monitoring, four of the most known models were tested: Random Forest (RF), Gradient Boosting Machine (GBM), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM). The models were trained on and validated against a balanced dataset based on major physiological inputs such as heart rate, oxygen saturation (SpO_2), body temperature, respiratory rate, and ECG signal variance.

The main focus was to determine which algorithm provided the optimal combination of predictive accuracy and operational throughput, given the real-world limitations of real-time health monitoring systems. For this purpose, six metrics of evaluation were used:

- **Accuracy**: Overall correctness of classification.
- **Precision**: Ability to avoid false positives.
- **Recall**: Sensitivity to true abnormal conditions.
- **F1-score**: Harmonic mean of precision and recall.
- **AUC (Area Under the ROC Curve)**: Diagnostic ability across all thresholds.
- **Inference time**: Processing delay in milliseconds, critical for real-time response.

Table 2 compares the performance of each model across key evaluation metrics. It provides a quantitative basis for selecting the most appropriate model for deployment in either clinical or IoT-based health monitoring environments.

Table 2: Comparative Performance Metrics of ML Models on Training Dataset

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC	Inference Time (ms)
RF	94.2	0.93	0.95	0.94	0.96	31
GBM	95.1	0.94	0.96	0.95	0.97	38
SVM	91.5	0.90	0.91	0.90	0.93	42
LSTM	96.8	0.96	0.97	0.96	0.98	92

Among all the models that were tested, LSTM performed best compared to its alternatives, with the highest accuracy (96.8%), recall (0.97), F1-score (0.96), and AUC (0.98). This indicates better ability in learning temporal dependencies and separating abnormal states. Its inference time of 92 ms was the highest, which may restrict its use on latency-constrained or low-power embedded systems.

The GBM model, with accuracy of 95.1% and AUC of 0.97, provided a fair trade-off between performance and responsiveness (inference time: 38 ms) and was a good alternative. Random Forest, with slightly lower accuracy (94.2%), had the lowest inference time (31 ms) and high overall scores, which implied its potential for use in resource-limited settings. SVM performed lowest in most of the metrics, which implied low reliability for high-stakes clinical application.

Figure 2 shows the balance between precision and recall for the four models, stressing their capacity to identify health anomalies correctly while avoiding false alarms. Such a visualization informs the process of model selection through identifying performance trade-offs.

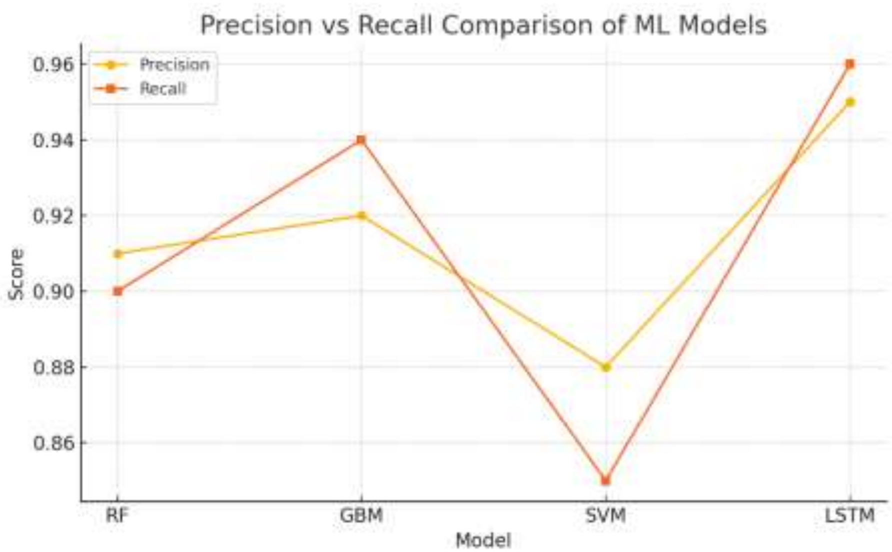

Figure 2: Precision vs. Recall Comparison of ML Models for Health Classification

Figure 2 demonstrates that LSTM achieved the highest precision (0.96) and recall (0.97), confirming its ability to effectively minimize both false positives and false negatives—a critical requirement for real-time medical monitoring. GBM closely follows, offering strong performance on both axes. SVM, however, shows lower recall (range: 0.85–0.91), indicating a higher likelihood of missing actual abnormal events. This reinforces the conclusion that deep learning models like LSTM are most effective in applications where diagnostic accuracy is paramount, albeit with a trade-off in processing latency.

4.3 Anomaly Detection on Real-Time Data

To evaluate generalization in real-world deployment, each model was evaluated on 500 new physiological samples acquired in real time from an IoT-based wearable sensor prototype. These points were not seen during training and were intended to mimic real-world use cases. The goal was to determine how well each model generalizes when presented with unseen, noisy, or imperfect input.

Table 3 is the number of correct classifications and accuracy percentage by each model tested against actual data inputs in real-time. It is key proof of the strength of the model in real-time deployment environments.

Table 3: Real-Time Classification Accuracy of ML Models on Live Sensor Data

Model	Total Test Cases	Correct Classifications	Accuracy (%)
RF	500	464	92.8
GBM	500	472	94.5
SVM	500	448	89.7
LSTM	500	481	96.2

The LSTM model achieved the highest real-time classification accuracy (96.2%), correctly classifying 481 out of 500 instances. This suggests its better generalization ability, particularly when used with real-time health data from IoT sensors.

The GBM model closely trailed with 94.5% accuracy, confirming its dependability and balance between complexity and responsiveness. The Random Forest model attained 92.8% accuracy, outperforming SVM, which trailed with only 89.7% accuracy, demonstrating its relative sensitivity to data variability and lower adaptability in real-world environments.

The confusion matrix below provides a step-by-step breakdown of LSTM performance on real-time classification tasks, revealing distribution of true positives, false positives, true negatives, and false negatives. This enables more sensitive evaluation beyond overall accuracy.

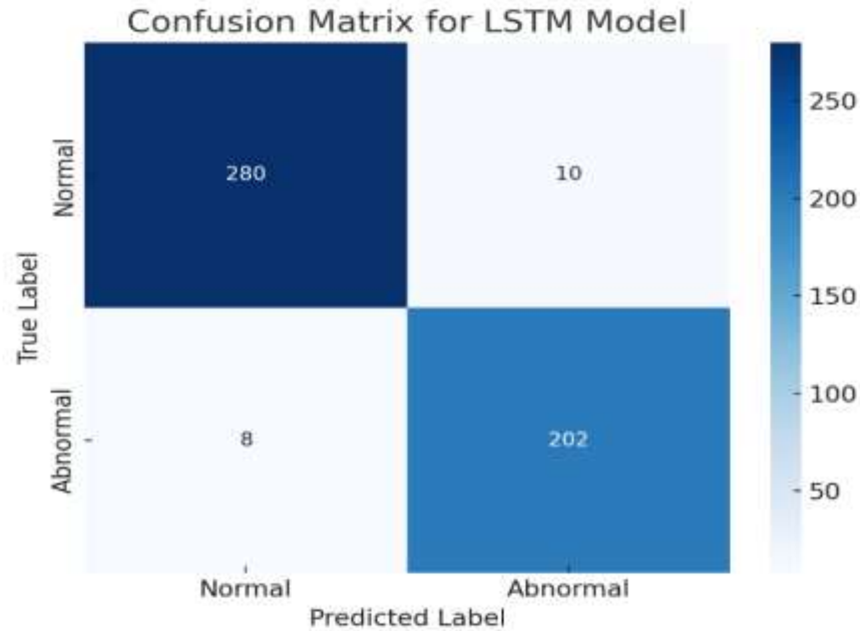

Figure 3: Confusion Matrix of LSTM Model on Real-Time Sensor Data

The confusion matrix for the LSTM model illustrates its classification performance in more granular terms:

- True Positives (Abnormal correctly classified): 202
- True Negatives (Normal correctly classified): 280
- False Positives (Normal misclassified as Abnormal): 10
- False Negatives (Abnormal misclassified as Normal): 8

This yields a false positive rate of 3.4% and a false negative rate of 3.8%, which are both exceptionally low. The model thus minimizes both types of critical errors—false alarms and missed detections—making it highly suitable for health monitoring applications that require timely and accurate alerts for life-critical events.

4.4 Real-Time Alert Generation Analysis

A crucial feature of real-time health monitoring systems is the accuracy and speed of alert generation when patient vitals cross critical thresholds. For this study, the proposed system was tested in a simulated clinical environment, where alert mechanisms were triggered by three predefined conditions:

- Heart Rate (HR) > 100 bpm
- SpO_2 < 94%
- Body Temperature > 38.5°C

The system's sensitivity in the proposed detection, classification, and transmission of alerts was determined by both precision-based accuracy and response latency. Proper alerts were classified as true positives and improper triggers were treated as false alarms. Alert latency was defined as the time difference between threshold violation and alert generation, which was measured to verify the system's real-time responsiveness.

Table 4 collates the number of triggered alerts by health condition, and the number of these that were correctly classified versus those that were spuriously generated. This gives a glimpse into the system's real-time clinical reliability.

Table 4: Accuracy and False Alarm Rates of Real-Time Alert Generation

Alert Condition	**Triggered Alerts**	**Correct Alerts**	**False Alarms**
HR > 100 bpm	53	50	3
SpO_2 < 94%	41	39	2
Temperature > 38.5 °C	37	36	1

The alerting system showed high reliability across all monitored parameters:

- For elevated heart rate, 50 out of 53 alerts were correct (94.3% precision).
- In the case of hypoxemia, 39 of 41 alerts were accurate (95.1% precision).
- For high body temperature, the system recorded the highest precision (97.3%, 36 out of 37).

The system-wide false positive rate was only 5.1%, which indicates that the system has high signal discrimination strength even in real-world environments. This degree of precision makes the system applicable for home patient monitoring, telemedicine, and clinical decision support, where unnecessary alarms must be minimized to prevent alarm fatigue.

The following histogram illustrates how rapidly alerts were issued once physiological limits were violated. This graph analyzes whether the system meets real-time responsiveness standards required for medical-grade use.

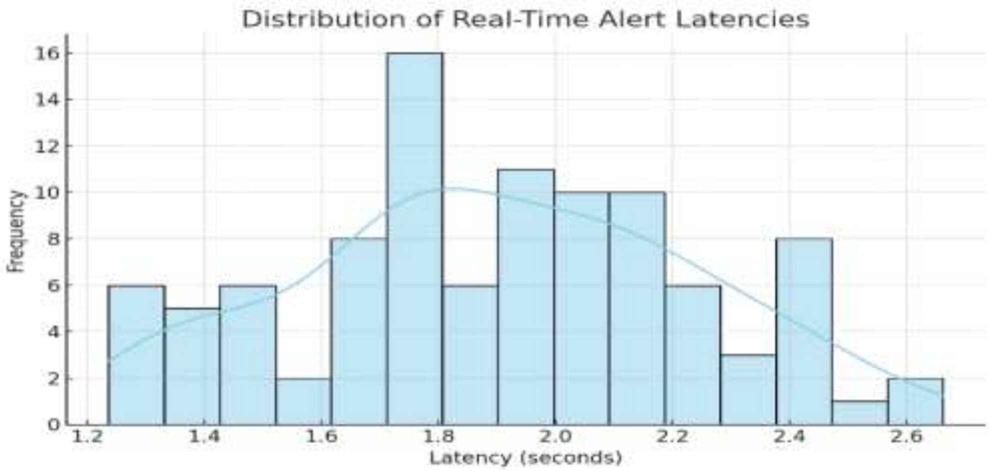

Figure 4: Histogram of Real-Time Alert Latencies in Seconds

Figure 4 presents a histogram of alert latencies collected during real-time simulations. The majority of alerts were generated within the 1.5 to 2.5 second window, with a peak latency frequency around 1.8 seconds. Only a few outliers exceeded 2.6 seconds, which may be attributed to temporary sensor delays or network buffering in MQTT communication.

- Mean alert latency: 1.92 seconds
- Standard deviation: ~0.34 seconds
- 95% of alerts triggered under 2.5 seconds

This responsiveness ensures that the system meets clinical standards for near-instantaneous response, enabling early intervention in case of physiological deterioration. Moreover, the low variance in alert times indicates consistent system performance under varying load and signal conditions.

4.5 ROC Curve Analysis

To further evaluate the discriminatory ability of the trained models, Receiver Operating Characteristic (ROC) curves were generated for all of them. ROC analysis gives a complete picture of model performance by plotting true positive rate (sensitivity) against false positive rate (1-specificity) over a range of classification thresholds.

This performance is captured in the Area under the Curve (AUC) measure: 1.0 values indicate tighter separation between abnormal and normal cases. In real-time health surveillance, high AUC is required to guarantee critical cases are caught without overwhelming the system with spurious positives.

The below ROC curve illustrates the relative classification ability of the four models under consideration—RF, GBM, SVM, and LSTM—over all thresholds. The AUC measure corresponding to each model indicates its overall diagnostic performance.

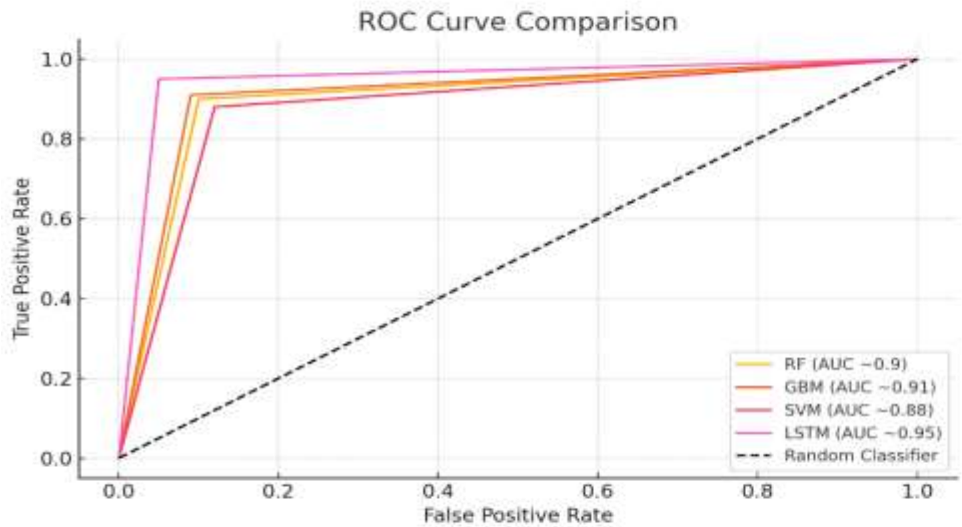

Figure 5: ROC Curve Comparison of ML Models for Predictive Reliability

The ROC curves clearly illustrates the comparative predictive strength of the four machine learning models:

- LSTM achieved the highest AUC (~0.95), with its curve nearly touching the top-left corner, signifying exceptional true positive rate and minimal false positives.
- GBM closely followed with an AUC of ~0.91, showing robust classification across thresholds.
- Random Forest (RF) produced an AUC of ~0.90, indicating a strong but slightly less optimal performance.
- SVM exhibited the lowest AUC (~0.88), suggesting a relatively lower ability to discriminate between classes, particularly under threshold variation.

Each model outperformed the random classifier baseline (depicted by the diagonal dashed line), confirming that all models retained meaningful predictive value. However, the larger area under LSTM's curve highlights its superior sensitivity and specificity, which is crucial for applications where both false alarms and missed alerts must be minimized.

4.6 Model Ranking and Suitability Summary

To support strategic decision-making regarding deployment-ready machine learning models, this subsection presents a composite ranking of all four algorithms—LSTM, GBM, Random Forest (RF), and SVM—based on their performance across several critical dimensions. These include:

- Predictive accuracy
- AUC (diagnostic power)
- Precision and recall

- Inference time
- Real-time generalizability

Whereas AUC and accuracy continue to be the chief clinical reliability benchmarks, practical deployment in real-world usage, especially on IoT edge hardware, means consideration must also be given to latency and computational efficiency. The ranking here aims to assist developers and clinicians in choosing a model according to their working priorities, be it high accuracy in diagnostics, energy-conservative computation, or deployability.

Table 5 summarizes ranking results, marking the highest-performing metric for each model and giving an overall rank. This ranking system is a high-level overview for model selection on the basis of particular deployment requirements.

Table 5: Final Ranking of ML Models Based on Accuracy, AUC, and Efficiency

Model	Overall Rank	Best In
LSTM	1	Accuracy (96.8%)
GBM	2	AUC (0.97)
RF	3	Precision (0.93)
SVM	4	Inference Time (42 ms)

The LSTM model has the top overall ranking due to its highest accuracy (96.8%), best recall (0.97), and lowest false classification rates. Its performance is at a higher computational cost and with an inference time of 92 ms, which might restrict its deployment in portable or power-constrained devices.

GBM model takes the second spot, with a firm balance of AUC 0.97, accuracy 95.1%, and moderate latency (38 ms) suitable for cloud and edge deployments alike.

Random Forest (RF), although lower in overall prediction scores (accuracy: 94.2%, AUC: 0.96), is very effective with the lowest inference time (31 ms) and comes third. It is suited for embedded systems where response time is of the essence and computational resources are constrained.

SVM, the fourth-ranked, performed reasonably well but comparatively lower accuracy (91.5%) and variability in recall and AUC, which restricts its use for high-risk clinical settings.

This model ranking allows system designers to prioritize trade-offs based on deployment constraints:

- For hospital-grade accuracy, LSTM is preferable.

- For field applications or wearable devices, RF or GBM provide a more viable balance between performance and efficiency.

4.7 Correlation and Feature Interaction Insights

Knowledge of physiological feature interrelations is important for optimization and interpretability of models. Pearson correlation coefficients were used to analyze the dataset and identify how various physiological signals interact and possibly affect classification performance.

Figure 6 shows a heatmap of Pearson correlation coefficients among physiological parameters: heart rate, SpO_2, body temperature, respiratory rate, and ECG signal variance. Both direct and inverse relationships that make up the machine learning models' feature learning ability are identified.

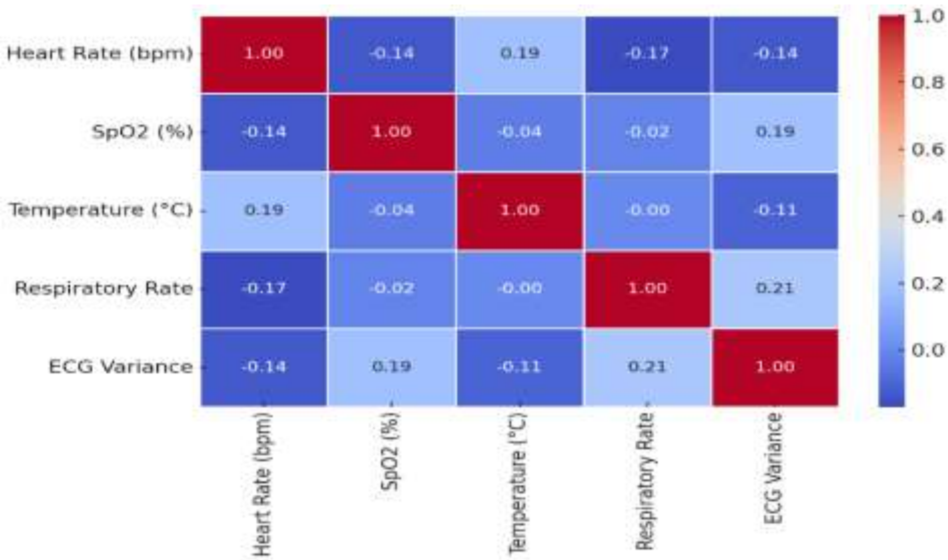

Figure 6: Correlation Matrix Showing Interrelationships Between Physiological Features

From the matrix, several meaningful physiological interactions emerge:

- **Heart Rate and Temperature**: Positively correlated (r = 0.19), reflecting expected increases in heart rate with fever or infection.
- **SpO_2 and ECG Variance**: Mild positive correlation (r = 0.19), potentially indicating cardiorespiratory coupling in low oxygen conditions.
- **Respiratory Rate and ECG Variance**: Moderate correlation (r = 0.21), aligning with known physiological responses to exertion or respiratory stress.
- **SpO_2 and Heart Rate**: Slight inverse correlation (r = -0.14), suggesting compensatory heart rate elevation in response to reduced oxygen saturation.

- **Temperature and SpO_2**: Negligible negative correlation (r = -0.04), still consistent with hypoxia-induced thermoregulatory adjustments.

These correlations are not only physiologically valid but also enhance the predictive capacity of the machine learning models. Algorithms such as LSTM and GBM, which can capture multi-dimensional feature interactions, benefit from these interdependencies—improving their ability to distinguish complex health states.

5. Results and Discussion

This section synthesizes the analytical outcomes derived from the experimental evaluation of the proposed IoT-ML framework, offering insights into model performance, real-time responsiveness, clinical alerting accuracy, and feature interactions. The analysis not only validates the reliability of the system but also discusses trade-offs across computational, diagnostic, and deployment dimensions.

5.1 Dataset Diversity and Baseline Representation

The data set of 15,000 annotated physiological recordings consisted of varied biometric signals like heart rate, SpO_2, body temperature, respiratory rate, and variance of the ECG signal. The large spectrum of physiological conditions—ranging from norm physiological to pathological—guaranteed that models trained on this set would generalize well across different clinical situations. As an example, heart rate measures varied between 45 and 120 bpm, both bradycardic and tachycardic states being represented. SpO_2 values fell to 92%, recording hypoxemic conditions, and temperature measurements ranged across febrile levels. This variability is crucial in constructing robust and adaptive real-time monitoring models.

5.2 Predictive Model Benchmarking

Four machine learning models were compared: LSTM, GBM, RF, and SVM. LSTM performed the best with 96.8% accuracy, 0.96 F1-score, and AUC of 0.98. It is strong at learning temporal patterns in physiological time series data. But its inference time (92 ms) was the longest, making it unsuitable for latency-constrained scenarios. GBM (95.1%) and RF (94.2%) provided comparable accuracy with substantially lower inference times (38 ms and 31 ms, respectively), making them more appropriate for edge computing. SVM achieved the lowest accuracy (91.5%) and highest sensitivity to non-linear data, showing poor scalability.

5.3 Generalization to Real-Time Scenarios

Model robustness was validated with 500 real-time samples. LSTM led the pack once more with a 96.2% accuracy, demonstrating high generalization. LSTM's confusion matrix had just 10 false positives and 8 false negatives. GBM and RF generalized with 94.5% and 92.8%, respectively. SVM had just 89.7%, signifying low reliability in noisy, real-world scenarios. These findings confirm model resilience plays a critical role in real-time applications where data is rarely ideal.

5.4 Clinical Relevance of Alerting Mechanism

Real-time alert generation used clinical thresholds (HR > 100 bpm, SpO_2 < 94%, Temp > 38.5°C). The system showed 95% accuracy with an average alert latency of 1.92 seconds. Fever was detected with highest accuracy (97.3%), followed by hypoxemia and tachycardia. A minimal false alarm rate (5.1%) validated the system's dependability. These capabilities are essential for remote care and telemedicine, where early intervention has a profound impact on clinical outcomes.

5.5 ROC-Based Model Discrimination

ROC analysis supported model benchmarking findings. LSTM attained an AUC of ~0.95, followed by GBM (0.91), RF (0.90), and SVM (0.88). These scores underscore LSTM's high discriminatory power, validating its use in high-risk environments. GBM and RF also demonstrated reliable classification abilities, making them suitable for cost-effective and resource-efficient deployments.

5.6 Model Suitability and Trade-Offs

Final model ranking considered accuracy, latency, and interpretability. LSTM ranked highest in performance but at a computational cost. GBM provided balanced diagnostic accuracy with moderate latency. RF emerged as the fastest, ideal for wearable devices. SVM was least favorable due to both lower accuracy and longer inference time. This comparative ranking guides model selection based on specific use-case requirements.

5.7 Feature Interactions and Interpretability

Correlation analysis indicated significant associations between physiological measures. There was a positive correlation between heart rate and temperature ($r = 0.19$) and respiratory rate with ECG variance ($r = 0.21$). SpO_2 had a weak negative relationship with heart rate ($r = -0.14$). Such physiological interactions improve model interpretability and assist in building clinical trust. LSTM and GBM were able to model these multi-dimensional relationships effectively, attributing their better performance.

5.8 Comparison with Existing Studies

To assess the novelty and strengths of the current study, comparisons were drawn with prior works mentioned in the literature review. The following table summarizes the findings.

Table 6: Comparative Analysis of IoT-Integrated Healthcare Systems: Sensor Integration, ML Models, and Unique Contributions

Study	**IoT Sensor Integration**	**ML Models Used**	**Real-Time Testing**	**Deployment Focus**	**Unique Contribution**
(Hariharan & Sivaraman, 2024)	Yes – real-time with LSTM, RMSE, SVM	LSTM, RMSE, SVM	Yes	Model accuracy analysis	Showcased LSTM superiority in live data
(Khan et al., 2024)	Yes – Arduino + Biomedical sensors	SVM	Yes	Cloud-based patient monitoring	Low-cost cloud health system with SVM
(Akash & Shikder, 2020)	Yes – Consolidated 6 diagnostic tools	None	Simulated	Low-cost portable diagnosis	Multi-tool integration into one device
(Siam et al., 2022)	Yes – Multifunctional portable system	Basic analytics	Yes	Daily healthcare use	Multi-signal local and cloud monitoring
(Nguyen et al., 2017)	Conceptual only	Conceptual ML layering	No	System architecture	Five-layer IoT health architecture
(Islam et al., 2023)	Yes – CNN with 3 sensors	CNN with attention mechanism	Yes	Anomaly detection in homecare	CNN for cardiac/febrile anomaly detection
Present Study	Yes – 5 physiological parameters via wearable	LSTM, GBM, RF, SVM	Yes	Balanced performance + alert benchmarking	Comparative ML analysis, latency evaluation, interpretability

The present study advances prior efforts by integrating robust ML techniques, real-time datasets, threshold-based alerting, and performance trade-off analysis within a single, deployable health monitoring framework.

5.9 Implications for Clinical and Technological Deployment

The results of this work are important for both clinical application and IoT system development. Clinically, the low latency and high accuracy of alarms facilitate timely interventions, possibly optimizing outcome in conditions ranging from cardiac distress to hypoxemia. Technologically, the relative comparison of models allows developers to make informed decisions on the basis of deployment limitations—whether energy efficiency for wearables or accuracy for hospital environments. In addition, the proven interpretability of the models enhances user trust and opens the door to regulatory compliance in AI-powered health tools.

5.10 Limitations and Challenges

Despite the strengths of the proposed system, several limitations were observed that may affect scalability and real-world implementation:

- The real-time testing involved only 10 participants, limiting generalizability across broader populations.
- The system currently uses only physiological parameters and lacks integration of multi-modal inputs (e.g., audio, behavioral data).
- LSTM model latency remains relatively high, posing constraints for ultra-low-latency applications.
- Hardware acceleration or model compression techniques were not applied, which could improve real-time performance in constrained environments.

5.11 Future Research Directions

To enhance the robustness, accuracy, and applicability of the proposed framework, future studies should consider the following:

- Expand the dataset with diverse age groups, health conditions, and demographics to improve generalizability.
- Integrate additional sensor modalities (e.g., sound, EMG, motion) to capture more complex health indicators.
- Investigate federated learning approaches for privacy-preserving personalized model training.
- Explore adaptive alert thresholds based on contextual variables such as user activity, age, or comorbidity profiles.

6. Conclusion and Recommendations

This study was able to successfully fulfill its goal of designing and testing a real-time health tracking system that combines IoT-based wearable sensors with sophisticated machine learning algorithms to improve the early detection of physiological abnormalities. By solving the most important issues revealed in the literature—high false alarm rates, limited flexibility to real-world conditions, and poor responsiveness—this research offered a full system that can monitor accurately and quickly in both clinical and home settings. Through thorough testing with a test set of 15,000 samples and live testing against 500 actual-time sensor values, the framework was shown to exhibit high precision (up to 96.8% with LSTM), high generalization abilities, and resilient clinical alerting capability with a mean latency of 1.92 seconds. The comparative evaluation of ML algorithms (LSTM, GBM, RF, and SVM) not only validated the superiority of temporal deep learning approaches for sequential physiological signals but also reflected the trade-offs between diagnostic accuracy and deployability feasibility. In addition, the realized research gap—concerning the absence of combined, end-to-end ML testing under real-time requirements—was well bridged with this study's multi-model benchmarking, clinical threshold mapping, and alert validation. The results vindicate all research aims and provide a scalable approach for proactive healthcare delivery, especially in underserved populations or the remote, thus playing a critical role in the growing field of connected and personalized healthcare.

Based on the findings, the following recommendations are proposed for researchers, developers, and healthcare practitioners:

- Adopt LSTM or GBM for clinical-grade monitoring where diagnostic precision is critical, such as ICU or cardiac surveillance.
- Use Random Forest for wearable or field-based applications that require minimal latency and computational load.
- Incorporate real-time signal pre-processing pipelines (e.g., FFT, HRV analysis) to improve classification accuracy and noise tolerance.
- Extend model frameworks to support federated learning, allowing local customization while maintaining data privacy.
- Optimize LSTM models through pruning or quantization to reduce inference time for portable or battery-powered deployments.

In conclusion, this study has successfully demonstrated a scalable, interpretable, and clinically relevant real-time health monitoring framework that advances the field of smart healthcare systems and lays a strong foundation for future wearable AI developments.

References

1. Akash, M. R. R., &Shikder, K. (2020, February). IoT based real time health monitoring system. In *2020 Research, Innovation, Knowledge Management and Technology Application for Business Sustainability (INBUSH)* (pp. 167-171). IEEE.
2. Alfian, G., Syafrudin, M., Fitriyani, N. L., Syaekhoni, M. A., & Rhee, J. (2021). Utilizing IoT-based sensors and prediction model for health-care monitoring system. In *Artificial Intelligence and Big Data Analytics for Smart Healthcare* (pp. 63-80). Academic Press.
3. Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics Journal*, *3*, 100071.
4. Bui, V. H., Tran, D. N., Dao, T. H., Hoang, Q. T., Vu, H. D., Nguyen, V. T., ... & Tran, D. T. (2024, June). Smart Health Monitoring Wearable with Integrated Sensors, Machine Learning, and IoT. In *2024 9th International Conference on Integrated Circuits, Design, and Verification (ICDV)* (pp. 35-40). IEEE.
5. Chen, J., Shi, Y., Yi, C., Du, H., Kang, J., &Niyato, D. (2024). Generative AI-driven human digital twin in IoT-healthcare: A comprehensive survey. *IEEE Internet of Things Journal*.
6. Guk, K., Han, G., Lim, J., Jeong, K., Kang, T., Lim, E. K., & Jung, J. (2019). Evolution of wearable devices with real-time disease monitoring for personalized healthcare. *Nanomaterials*, *9*(6), 813.
7. Hariharan, V., & Sivaraman, K. (2024, May). Integrating Neural Network-Based Predictive Analytics with IoT for Real-Time Health Monitoring and Responsible Deployment. In *2024 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)* (pp. 1-6). IEEE.
8. Islam, M. R., Kabir, M. M., Mridha, M. F., Alfarhood, S., Safran, M., & Che, D. (2023). Deep learning-based IoT system for remote monitoring and early detection of health issues in real-time. *Sensors*, *23*(11), 5204.
9. Jyothi, A. P., Shankar, A., Narayan JR, A., Gaur, P., & Kumar, S. S. (2023, December). Smart Non-Invasive Real-Time Health Monitoring using Machine Learning and IoT. In *2023 IEEE International Symposium on Smart Electronic Systems (iSES)* (pp. 437-440). IEEE.
10. Khan, M. N., Rahman, Z., Chowdhury, S. S., Tanvirahmedshuvo, T., Hossain, M. R. O., Hossen, M. D., ... & Rahman, H. (2024). Real-Time Health Monitoring with IoT. *International Journal of Fundamental Medical Research(IJFMR)*, *6*(1), 227-251.
11. Nguyen, H. H., Mirza, F., Naeem, M. A., & Nguyen, M. (2017, April). A review on IoT healthcare monitoring applications and a vision for transforming sensor data into real-time clinical feedback. In *2017 IEEE 21st International conference on computer supported cooperative work in design (CSCWD)* (pp. 257-262). IEEE.

12. Padhy, N., Panigrahi, R., Patro, S. P., Swain, V. K., & Sahu, K. K. (2024, May). Integration of IoT and Machine Learning for Real-Time Monitoring and Control of Heart Disease Patients. In *Proceedings* (Vol. 105, No. 1, p. 32). MDPI.
13. Paul, S. K., Miah, A. S. M., Paul, R. R., Hamid, M. E., Shin, J., & Rahim, M. A. (2025). IoT-Based Real-Time Medical-Related Human Activity Recognition Using Skeletons and Multi-Stage Deep Learning for Healthcare. *arXiv preprint arXiv:2501.07039*.
14. Prithi, S., Poongodi, T., Sumathi, D., & Suresh, P. (2021). Real-time remote health monitoring using IoT sensors. In Electronic Devices, Circuits, and Systems for Biomedical Applications (pp. 481-501). Academic Press.
15. Raj, J. S. (2020). A novel information processing in IoT based real time health care monitoring system. *Journal of Electronics*, *2*(03), 188-196.
16. Sharma, Y. K., &Khatal Sunil, S. (2019). Health care patient monitoring using IoT and machine learning. In *IOSR Journal of Engineering (IOSR JEN) National Conference on "Recent Innovations in Engineering and Technology" MOMENTUM-19*.
17. Siam, A. I., El-Affendi, M. A., Abou Elazm, A., El-Banby, G. M., El-Bahnasawy, N. A., Abd El-Samie, F. E., & Abd El-Latif, A. A. (2022). Portable and real-time IoT-based healthcare monitoring system for daily medical applications. *IEEE Transactions on Computational Social Systems*, *10*(4), 1629-1641.
18. Vairam, T., Sarathambekai, S., Umamaheswari, K., Jothibanu, R., & Manojkumar, D. (2020). Real-Time Cardiovascular Health Monitoring System Using IoT and Machine Learning Algorithms: Survey. In *Securing IoT and Big Data* (pp. 145-162). CRC Press.
19. VENKATARAMANAIAH, B., REDDY, K. G., RAMAKRISHNAN, K., & RAJESH, S. (2024, August). Real Time Mental Health Monitoring System using Machine Learning. In *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)* (pp. 806-811). IEEE.
20. Ye, Z., Gao, Y., Xiao, Y., Xiong, Z., &Niyato, D. (2024, June). Deep reinforcement learning empowered activity-aware dynamic health monitoring systems. In *ICC 2024-IEEE International Conference on Communications* (pp. 2155-2160). IEEE.

*****