© MAY 2026 | IRE Journals | Volume 9 Issue 11 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I11-1717965

Personalized Multi-Modal Wearable Data-Based Heart Risk Prediction System Using Machine Learning

MOHAMMED FAHAD [1], DR. HARIPRIYA V [2]
[1]*Department of Computer Science and Information Technology* Jain *Deemed-to-be-University, Bengaluru, India*
[2]*Assistant Professor Department of Computer Science and Information Technology Jain Deemed-to-be University, Bengaluru, India*

Abstract- Heart and circulatory conditions remain a leading cause of death globally, and they often worsen quietly before obvious symptoms appear. Consumer wearables now make it practical to collect physiology around the clock, yet many pipelines still lean on one signal at a time and collapse decisions into coarse "normal vs abnormal" labels that arrive late for prevention. This work outlines a personalized, multi-modal wearable pipeline that jointly uses ECG, photoplethysmography (PPG), heart-rate variability (HRV), blood-oxygen saturation (SpO₂), and sleep-related cues. Machine learning sits on top of adaptive baselines so the model can learn what is typical for a given person, not only population averages, and it emits a continuous risk score from 0 to 100 that is easier to track over days and weeks. Explainable AI (XAI) is folded in so users and clinicians can see which factors moved the score, supporting trust and safer follow-up. Overall, the aim is to shift monitoring from reactive firefighting toward earlier, individualized cardiovascular vigilance.

Keywords: Cardiovascular Disease, Wearable Health Monitoring, Machine Learning, Multi-Modal Data Fusion, Heart Risk Prediction, Explainable Artificial Intelligence

I. INTRODUCTION

Cardiovascular disease still accounts for a large share of premature mortality, which keeps prevention and early warning on the policy agenda. Wearables have turned continuous vitals from a clinic-only luxury into an everyday stream, producing dense timelines of ECG snippets, pulse-wave proxies, oxygen trends, and sleep structure. That volume is only useful if models can turn it into timely, person-specific guidance rather than another dashboard of disconnected charts.

In practice, many consumer-facing stacks remain reactive: they flag spikes once something already looks wrong, and they often read only one modality even when the device could supply several. Single-sensor pipelines miss compensatory patterns—for example, sleep disruption paired with autonomic shifts—that only appear when signals are viewed together. Fixed population cut-offs also misread healthy athletes, anxious but stable users, and people on medication, because "normal" is personal as much as it is statistical.

The core research challenge is therefore an integrated design: fuse heterogeneous wearable streams, maintain baselines that move with the individual, and surface risk as a graded trajectory instead of a one-off label. The motivation is straightforward—earlier, clearer signals create more room for lifestyle change, medication titration, or clinical escalation before an emergency—while keeping the narrative understandable to non-specialists.

II. LITERATURE REVIEW

A. Traditional Approaches

Earlier approaches for heart disease monitoring mainly relied on rule-based systems and fixed threshold methods for detecting abnormalities. These methods provided transparency and were easy to implement, but they struggled to handle noisy wearable data and complex physiological relationships. Studies by Latha and Meenakshi [8] and Singh and Kumar [10] focused on early heart disease diagnosis using wearable sensors and machine learning classification techniques.

B. Machine Learning Approaches

The development of machine learning techniques significantly improved prediction accuracy in cardiovascular monitoring systems. Algorithms such as Random Forest (RF), Support Vector Machine (SVM), and hybrid machine learning models demonstrated strong performance in disease prediction tasks. Guruprasad et al. [5] and Mohan et al. [9] showed that machine learning methods can effectively capture hidden patterns in clinical and wearable datasets. However, wearable sensor data often contains motion artifacts and missing values, requiring proper preprocessing and explainability mechanisms.

C. Deep Learning Approaches

Recent research has increasingly focused on deep learning approaches such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for analyzing ECG and physiological signals. Gupta et al. [4] demonstrated that deep learning models can accurately forecast cardiac episodes by learning complex temporal patterns from medical data. Similarly, Prashanthi et al. [2] proposed AI-driven wearable monitoring systems that leverage deep learning for continuous healthcare analysis. Despite their high accuracy, deep learning models require large datasets, higher computational resources, and often lack interpretability.

D. Hybrid Approaches

Hybrid healthcare monitoring systems integrate IoT devices, wearable sensors, cloud computing, and machine learning algorithms to improve end-to-end healthcare monitoring. Abraham and Babu [1] proposed a multi-modal AI-based health risk prediction system capable of generating early alerts using wearable sensor data. Ed-daoudy and Maalmi [7] further explored real-time machine learning techniques combined with big data technologies for early heart disease detection. Although these systems improve monitoring capabilities, many existing approaches still focus mainly on event detection rather than continuous personalized risk prediction.

E. Comparative Summary

TABLE I: Comparative Analysis of Existing Methods

Method / Approach	Data Source	Key Strength	Primary Limitation
Deep Learning (CNN/LSTM)	ECG / SpO_2	High accuracy (>96%)	High computational cost; lacks interpretability
Traditional ML (RF/SVM)	Clinical Records	Robust performance	Poor handling of noisy wearable data
Big Data Systems	IoT Streams	Scalable processing	Requires heavy infrastructure
XAI Models	Static datasets	Provides explanation	Limited real-time capability
Proposed Framework	Multi-Modal	Personalized risk score	Integration complexity

Table 1 shows the architecture of the proposed heart-risk prediction system. The system collects data from wearable sensors such as ECG, PPG, HRV, SpO_2, and sleep monitoring devices. The data is preprocessed and analyzed using machine learning models to predict heart risk. An Explainable AI module then provides understandable explanations for the predictions.

F. Identified Research Gaps

Across the surveyed line of work, recurring gaps stand out: weak personalization against static norms; binary endpoints instead of trajectories; shallow multi-modal fusion; thin on-device adaptability; limited explainability for end users; and uneven attention to deployment costs, privacy, and longitudinal validation.

III. PROPOSED METHODOLOGY

A. System Architecture Overview

Architecturally, the proposal is a layered pipeline: capture synchronized wearable streams, preprocess and harmonize them, estimate personalized cardiovascular risk with ML, and wrap outputs in explanations suitable for lay readers and clinicians alike.

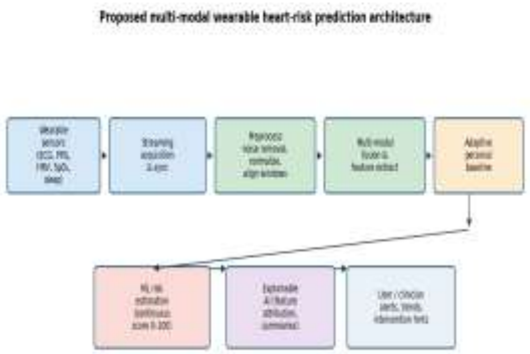

Fig. 1. Proposed multi-modal wearable heart-risk prediction architecture.

Figure 1 illustrates the complete workflow of the system from wearable data collection to final risk prediction.

1. The process includes:
2. Collecting multi-sensor wearable data
3. Cleaning and processing signals
4. Extracting important features
5. Updating personal health baselines
6. Predicting heart risk using machine learning
7. Generating explainable AI summaries
8. Displaying alerts and dashboards to users and clinicians.

B. Data Processing Pipeline

Ingestion starts with the modalities the hardware already exposes—ECG or single-lead proxies, PPG-derived pulse, HRV summaries, SpO_2, and sleep staging or proxies. Streams are denoised, resampled to common clocks, and segmented into analysis windows; from there, morphology, variability, coupling between sleep and autonomic markers, and trend features feed the modeling stage.

C. Risk Prediction Model

A supervised learner maps the fused feature vector to a continuous 0–100 risk index so small day-to-day shifts remain visible rather than being collapsed into a single threshold crossing. Personal baselines anchor the model so "unusual" is judged relative to the wearer's recent history, not only cohort norms.

D. Explainable AI Module

An XAI layer surfaces drivers of the score—dominant features, stability of the attribution across nearby windows, and plain-language rationales—so the output is reviewable in shared decision-making rather than treated as a black box.

E. Dataset Description

Training and evaluation assume wearable-aligned corpora with synchronized labels or weak supervision where labels are sparse; splits respect subjects to avoid leakage, and augmentation (jitter, noise mixes, sensor dropout) can harden models to real-world variability when ethically appropriate.

F. Workflow

Operationally, data move through a fixed chain from raw streams to user-facing insight. Figure 2 summarizes the end-to-end path; the numbered steps below spell out the same order in prose.

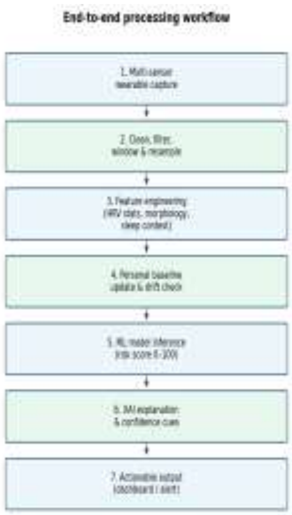

Fig. 2. End-to-end processing workflow from wearable capture to actionable output.

Workflow steps (aligned with Fig. 2):

1. Multi-sensor wearable capture under consent and battery-aware sampling policies.
2. Signal cleaning: denoise, filter, resample, and align modalities to a common timeline.
3. Feature engineering across ECG/PPG morphology, HRV statistics, SpO_2 trends, and sleep context.
4. Personal baseline refresh with drift checks so deviations reflect genuine change.

5. ML inference producing a continuous 0–100 cardiovascular risk index.
6. XAI summaries highlighting influential factors and stability across adjacent windows.
7. User-facing outputs: dashboards, gentle alerts, and clinician-friendly exports where applicable.

IV. RESULTS AND DISCUSSION

A. Expected Performance Outcomes

Given the fused representation and personalized baselines, the expectation—pending empirical validation on the target cohort—is materially stronger ranking of high-risk periods than single-modality baselines, with accuracy competitive with strong literature benchmarks when preprocessing and leakage controls are strict. Holistic fusion of ECG, PPG, HRV, SpO_2, and sleep context should better track autonomic and recovery-linked states that single streams miss.

The incorporation of adaptive baseline modeling further enhances system performance by accounting for individual physiological variations. This personalization minimizes false positives and improves the reliability of predictions in real-world scenarios. In addition, preprocessing techniques such as signal denoising and normalization ensure that the system remains robust even in the presence of noise caused by motion artifacts or environmental factors. The overall system is designed to maintain consistent performance across diverse conditions, making it suitable for continuous real-time monitoring applications.

B. Comparative Evaluation Plan

Planned comparisons include strong classical baselines (Random Forest, SVM) and representative deep models (e.g., CNN-style waveform encoders), with standard discrimination metrics—accuracy where class balance permits, precision/recall/F1 under realistic skew, plus ROC-AUC and calibration views when scores are treated as risk proxies.

Beyond static discrimination, we foreground lead time—how early sustained score elevations precede adverse trajectories—contrasted with reactive baselines that only react after thresholds breach. Qualitative review of XAI outputs will check whether attributions align with physiology and remain stable across neighboring windows, not only whether headline metrics look strong.

C. Discussion

The proposed system offers several advantages over existing approaches by combining prediction, personalization, and interpretability within a unified framework. Unlike traditional binary classification systems, the continuous risk scoring mechanism enables gradual monitoring of cardiovascular health, allowing users and healthcare professionals to observe trends and detect early signs of deterioration. This shift from reactive detection to proactive prediction represents a significant advancement in wearable healthcare systems.

The integration of Explainable Artificial Intelligence plays a crucial role in enhancing system transparency. By identifying key contributing factors such as abnormal heart rate variability, reduced oxygen saturation, or irregular sleep patterns, the system provides actionable insights that support informed decision-making. This capability addresses one of the major limitations of black-box machine learning models, thereby increasing user trust and facilitating clinical adoption.

Despite its advantages, the system faces certain limitations, including dependency on the accuracy of wearable sensors and potential inconsistencies in user-generated data. Variability in sensor quality and external conditions may affect data reliability, which in turn can influence prediction outcomes. Future enhancements will focus on improving data quality through advanced calibration techniques and integrating clinical-grade datasets to further strengthen model performance.

V. CONCLUSION

This paper presented a personalized multi-modal wearable data-based heart risk prediction system that integrates machine learning and Explainable Artificial Intelligence to enable proactive cardiovascular monitoring. The proposed framework effectively addresses the limitations of existing systems, including reliance on single-modality data, lack of personalization, and absence of continuous risk assessment mechanisms.

By combining multiple physiological signals and incorporating adaptive baseline modeling, the system provides a comprehensive and individualized evaluation of cardiovascular health. The continuous risk scoring approach allows for better tracking of health trends, enabling early detection of potential risks and timely medical intervention. Furthermore, the integration of Explainable AI enhances transparency by providing clear and interpretable reasoning behind predictions, thereby improving user trust and clinical usability.

In conclusion, the proposed system represents a significant step toward predictive and preventive healthcare. It shifts the focus from reactive diagnosis to early risk forecasting, offering the potential to reduce the incidence of severe cardiac events. Future work will focus on real-time deployment, large-scale validation, and integration with electronic health record systems to enhance scalability and practical applicability in real-world healthcare environments

REFERENCES

[1] A. B. Abraham and N. V. Babu, "Multi-Modal AI Personal Health Risk Prediction and Early Alert System," IEEE (ICACRS), 2025.

[2] Bh. Prashanthi et al., "AI-Driven Smart Health Monitoring System Using Wearable IoT Devices," IEEE Conference, 2024.

[3] A. Sony et al., "Predictive Health Monitoring System for Patient Risk Assessment," IEEE (CONIT), 2025.

[4] G. Gupta et al., "Deep Learning-Based Forecasting of Spontaneous Termination of Cardiac Episodes," Nature Scientific Reports, 2023.

[5] S. Guruprasad et al., "Heart Disease Prediction Using Machine Learning Techniques," ICEECCOT, 2021.

[6] H. Al-Libawy, et al., "HRV-Based Operator Fatigue Analysis and Classification," 2016 13th International Multi-Conference on Systems, Signals & Devices (SSD), 2016.

[7] A. Ed-daoudy and K. Maalmi, "Real-time Machine Learning for Early Detection of Heart Disease Using Big Data Approach," 2018 International Federation for Information Processing (IFIP), 2018.

[8] K. B. Latha and S. Meenakshi, "Prediction of Heart Disease at an Early Phase Using Machine Learning and Wearable Sensors," 2022 International Conference on Intelligent Data Science Technologies and Applications (ICDSIS), 2022.

[9] S. Mohan, et al., "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," IEEE Access, vol. 7, 2019.

[10] P. K. Singh and S. Kumar, "Machine Learning-Based Classification for Early Diagnosis of Heart Disease," 2020 International Conference on Computer Science, Engineering and Applications (ICCSEA), 2020.