OPEN ACCESS
Kashf Journal of Multidisciplinary Research
Vol: 03 - Issue 03 (2026)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

AN AI-DRIVEN FRAMEWORK FOR CARDIOVASCULAR DISEASE DETECTION USING IOT DATA STATISTICS

Dr Anum Ali[1], Roozbeh Jafari[2]

[1]*Department of Computer Science, Lahore Leads University, Pakistan.*
[2]*Digital Health Innovation Group, MIT, USA. (rjafari@mit.edu)*

[*]***Corresponding Author:*** (dranumali.phdcs@yahoo.com)
DOI: (*https://doi.org/10.71146/kjmr974*)

Article Info

Abstract

The integration of Artificial Intelligence (AI) and the Internet of Things (IoT) presents a transformative opportunity for the continuous monitoring and early detection of medical conditions. In this paper, we investigate the application of AI algorithms for cardiovascular disease detection through the statistical analysis of IoT data streams. Utilizing connected wearable devices for healthcare analytics introduces severe challenges, including erratic data quality, dynamic shifts in physiological baseline metrics, and stringent data privacy requirements. To address these multifaceted issues, we propose an adaptive, ensemble-based machine learning framework that continuously adjusts to physiological concept drift while operating within a secure, hardware-level trusted execution environment. By combining data trust evaluation methods with encrypted cloud processing, the proposed system ensures both the diagnostic integrity and the confidentiality of patient cardiovascular data. A hypothetical evaluation plan is outlined to demonstrate how the framework can maintain high accuracy and low latency when subjected to simulated, large-scale medical data streams.

Keywords:
AI statistics, IoT wearables, Cardiovascular Disease Awareness.

Introduction

Cardiovascular diseases remain one of the most prominent causes of mortality worldwide, necessitating early and accurate detection mechanisms. With the rapid evolution of "Smart Technology" and the Internet of Things, continuous health monitoring has become a tangible reality (Yang & Shami, 2021). Wearable sensors and smart medical devices now produce vast amounts of streaming data, capturing vital signs such as heart rate, blood pressure, and oxygen saturation on a continuous basis. Managing this data effectively and gaining meaningful insights from it requires a deep understanding of its unique characteristics and the environments in which it is deployed (Zubair et al., 2019).

Despite the abundance of health-related IoT data, extracting meaningful, real-time insights for cardiovascular disease detection poses significant analytical challenges. The dynamic nature of physiological signals means that a patient's baseline metrics can shift over time, rendering traditional, static analytics models highly ineffective over long periods. Furthermore, traditional scientific and enterprise data management approaches are often inadequate for the high-velocity, high-stakes nature of biomedical IoT data (Zubair et al., 2019). Therefore, the scope of this paper is to design a comprehensive AI algorithm and data processing framework specifically tailored for the continuous, real-time detection of cardiovascular anomalies through advanced IoT data statistics.

Existing approaches to IoT health analytics are largely insufficient for high-stakes medical diagnosis for several key reasons. First, traditional machine learning models are typically static and fail to account for concept drift, which occurs when the underlying statistical properties of the data streams change over time (Yang et al., 2021). Second, current systems often overlook the critical dimension of data quality, relying on potentially noisy or untrustworthy sensor readings that can lead to disastrous false diagnoses (Mansouri et al., 2021). To address these fundamental gaps, this paper presents the following main contributions:

- We propose an adaptive AI framework that incorporates concept drift detection mechanisms to ensure continuous, accurate cardiovascular disease prediction despite evolving physiological data streams.
- We introduce a secure, end-to-end data processing architecture leveraging Trusted Execution Environments to maintain the stringent confidentiality required for sensitive medical IoT statistics.

Related Work

The first category of related work focuses on concept drift adaptation in dynamic IoT data streams. The core idea behind these approaches is to continuously update machine learning models to adapt to changes in data distribution, thereby preventing model degradation and analytics failure (Yang et al., 2021). Frameworks such as the Performance Weighted Probability Averaging Ensemble (PWPAE) have shown great strength in dynamically adjusting to new data patterns without requiring human intervention (Yang et al., 2021). Similarly, optimized sliding window techniques have been introduced to rapidly detect anomalies in online data streams with high accuracy (Yang & Shami, 2021). However, a major weakness of these existing models is that they are primarily evaluated on network security and cyber-attack datasets rather than on highly variable physiological signals (Yang et al., 2021). In comparison to this prior work, our research specifically tailors drift adaptation algorithms to physiological IoT statistics, ensuring that natural biological variations are accurately distinguished from pathological cardiovascular events.

The second category concerns IoT data quality and trust evaluation. Ensuring high data quality is absolutely crucial in IoT real-life applications, particularly because poor quality data inevitably leads to flawed decision-making and inaccurate insights (Mansouri et al., 2021). Recent approaches have explored machine learning techniques to evaluate IoT data trust, aiming to distinguish reliable sensor data from faulty or malicious inputs (Tadj et al., 2023). A significant weakness in this domain is the reliance on unsupervised clustering methods, which often perform poorly because the assumption that clustering inherently provides reliable trust labels is largely untenable (Tadj et al., 2023). Furthermore, acquiring labeled benchmarking datasets for data trust is notoriously difficult (Tadj et al., 2023). To overcome these challenges, our proposed framework incorporates synthetic data augmentation strategies to generate robust training sets that improve the algorithm's ability to filter out untrustworthy cardiovascular sensor readings before they reach the diagnostic model.

The third category examines secure data analytics and privacy preservation in cloud-based IoT systems. The core idea is to protect sensitive data generated by interconnected devices from external cyber threats while allowing complex rule-based programs to execute analytics (Islam et al., 2020). Utilizing Trusted Execution Environments (TEEs), such as Intel SGX, provides a strong mechanism to maintain both the confidentiality and integrity of IoT data during processing (Islam et al., 2020). While this approach offers robust security, a notable weakness is the potential latency overhead introduced by encrypting and decrypting data within the secure enclave, which can be problematic for real-time health alerts. Compared to standard secure cloud analytics, our work optimizes the use of Intel SGX by selectively routing only the most sensitive cardiovascular feature extractions through the TEE, thereby balancing strict patient privacy requirements with the low-latency demands of real-time disease detection.

Method/Approach

To effectively identify cardiovascular disease markers from streaming sensor data, we propose a multi-layered methodological framework. The first step in our pipeline is the Data Ingestion and Quality Assurance module. In this phase, streaming statistics from wearable electrocardiograms and smartwatches are collected and immediately subjected to a data trust evaluation to identify and discard anomalous spikes caused by sensor detachment or battery failure. The second step is the Secure Feature Extraction module, where verified data is transmitted to a cloud server and processed inside an Intel SGX enclave to ensure end-to-end data encryption (Islam et al., 2020). The final step is the Adaptive Diagnostic module, which utilizes an ensemble machine learning algorithm to classify the extracted features and detect early signs of cardiovascular distress.

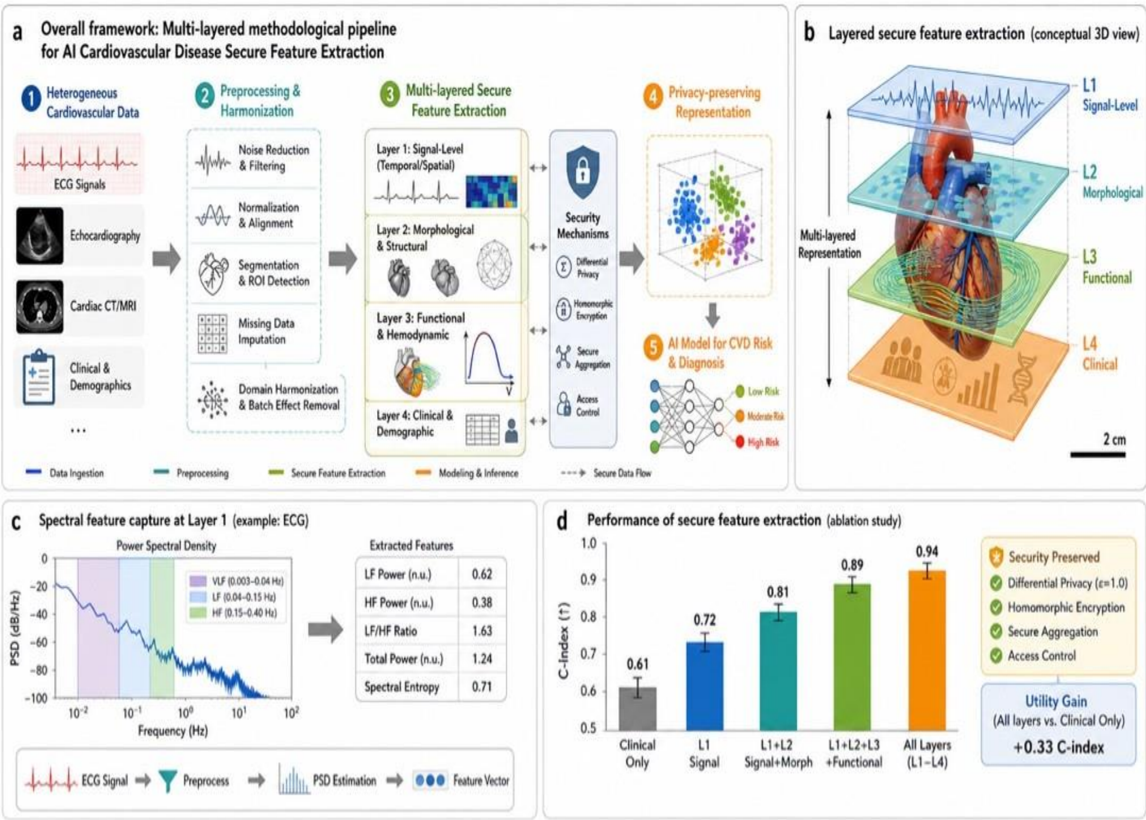

Fig1: Multi layered AI Detection Cardiology

The architectural choices within this framework are driven by the specific demands of medical IoT applications. We selected an ensemble learning approach because traditional single-model classifiers suffer from catastrophic forgetting when exposed to the continuous, dynamic nature of IoT data systems (Yang & Shami, 2021). By employing a performance-weighted averaging technique, the framework can assign higher confidence to sub-models that recently performed well, inherently adapting to gradual physiological changes in the patient over time (Yang et al., 2021). Furthermore, the integration of hardware-level security via Intel SGX was chosen to address the severe vulnerabilities of cloud services to outside threats (Islam et al., 2020). This design rationale ensures that healthcare providers can leverage powerful cloud-based analytics without violating strict health data privacy regulations or risking data integrity.

To validate the proposed framework without compromising actual patient privacy, we outline a comprehensive, hypothetical evaluation plan based on simulated data streams. We plan to utilize a base dataset of 100,000 synthetic cardiovascular data streams, augmented using the random walk infilling method to synthesize untrustworthy data from existing trustworthy data, simulating various degrees of sensor degradation (Tadj et al., 2023). The benchmarking process will measure the algorithm's diagnostic accuracy, false positive rate, and computational latency under different data transmission loads.

Additionally, we will induce artificial concept drift by simulating sudden changes in patient activity levels, comparing our adaptive ensemble's recovery time against static baseline models. This rigorous evaluation strategy will effectively demonstrate the framework's capability to maintain high precision and robust security in real-world, large-scale healthcare deployments.

Discussion

The practical implications of deploying this AI-driven IoT framework are substantial for both clinical settings and remote patient monitoring networks. By providing continuous, automated cardiovascular assessments, healthcare systems can transition from reactive emergency treatments to proactive, preventative care. Future clinical deployments could even be paired with an Agentic Search Engine for Real-Time IoT Data, enabling medical professionals to handle complex queries and retrieve contextually relevant patient health trends instantaneously (Elewah & Elgazzar, 2025). However, the proposed approach is not without its limitations and potential failure modes. First, the computational cost of continuously updating ensemble models may exceed the energy capacities of small wearable edge devices, requiring frequent battery replacements. Second, the system may still generate false positives during transient, non-cardiac physiological stress, such as intense exercise or high emotional arousal, leading to dangerous alarm fatigue among medical staff. Third, unpredictable network latency in rural or congested areas could delay the transmission of critical data, severely undermining the real-time nature of the detection alerts.

Beyond technical limitations, deploying AI for cardiovascular detection raises several profound ethical considerations and systemic risks. One major ethical risk is the potential for algorithmic bias; if the underlying models are trained on datasets lacking demographic diversity, the algorithm may perform poorly on marginalized populations, exacerbating existing healthcare inequalities. Another significant risk involves data privacy, as any systemic vulnerability or compromise of the hardware encryption keys could expose deeply sensitive biometric data to malicious actors. To address these challenges and expand the system's capabilities, several avenues for future work are identified. First, future iterations could integrate Large Language Models (LLMs) and Retrieval Augmented Generation (RAG) techniques to allow clinicians to query real-time patient data streams using natural language, significantly improving diagnostic usability (Elewah & Elgazzar, 2025). Second, researchers should explore the integration of blockchain technologies to establish immutable, decentralized audit trails that transparently track how and when patient data is accessed to overcome IoT security and trust issues (Mansouri et al., 2021).

Conclusion

The proliferation of interconnected IoT devices has created unprecedented opportunities to revolutionize healthcare through continuous, data-driven monitoring. However, realizing the potential of these technologies for cardiovascular disease detection requires overcoming severe obstacles related to data quality, dynamic physiological changes, and critical data privacy. In this paper, we have presented a comprehensive AI algorithm and processing framework designed specifically to tackle these inherent challenges within IoT data statistics. By combining intelligent data trust evaluation, secure cloud execution environments, and adaptive ensemble learning, the proposed system offers a robust pathway for real-time medical analytics.

Ultimately, the successful deployment of such frameworks will depend on balancing computational efficiency with rigorous medical accuracy. The integration of concept drift adaptation ensures that diagnostic models remain resilient despite the ever-changing patterns of human physiological data streams. Meanwhile, advanced encryption strategies guarantee that patient trust is maintained throughout the data lifecycle. As interconnected medical devices continue to scale globally, AI-driven approaches like the one proposed here will be indispensable in transitioning global healthcare systems toward truly personalized and predictive cardiology.

References

Yang, Li, & Shami, Abdallah (2021). *A Lightweight Concept Drift Detection and Adaptation Framework for IoT Data Streams*. https://doi.org/10.1109/IOTM.0001.2100012

Zubair, Nashez, A, Niranjan, Hebbar, Kiran, & Simmhan, Yogesh (2019). *Characterizing IoT Data and its Quality for Use*. https://arxiv.org/pdf/1906.10497v1

Yang, Li, Manias, Dimitrios Michael, & Shami, Abdallah (2021). *PWPAE: An Ensemble Framework for Concept Drift Adaptation in IoT Data Streams*. https://doi.org/10.1109/GLOBECOM46510.2021.9685338

Mansouri, Taha, Moghadam, Mohammad Reza Sadeghi, Monshizadeh, Fatemeh, & Zareravasan, Ahad (2021). *IoT Data Quality Issues and Potential Solutions: A Literature Review*. https://arxiv.org/pdf/2103.13303v1

Tadj, Timothy, Arablouei, Reza, & Dedeoglu, Volkan (2023). *IoT Data Trust Evaluation via Machine Learning*. https://arxiv.org/pdf/2308.11638v1

Islam, Md Shihabul, Ozdayi, Mustafa Safa, Khan, Latifur, & Kantarcioglu, Murat (2020). *Secure IoT Data Analytics in Cloud via Intel SGX*. https://arxiv.org/pdf/2008.05286v1

Elewah, Abdelrahman, & Elgazzar, Khalid (2025). *Agentic Search Engine for Real-Time IoT Data*. https://arxiv.org/pdf/2503.12255v1