DDPG-IoT Driven Intelligent Cloud Architecture for Cardiovascular Disease Risk Assessment

Rohit Raut[1], **Dr. Aasheesh Raizada**[2]

[1]Department of Computer Engineering and Application, Faculty of Engineering and Technology Mangalayatan University, Beswan, Aligarh, India Email: 20230288_rohit@mangalayatan.edu.in

[2]Asso. Prof., Department of Computer Engineering and Application, Faculty of Engineering and Technology Mangalayatan University, Beswan, Aligarh, India Email: aasheesh.raizada@mangalayatan.edu.in

Abstract—Cardiovascular diseases (CVDs) are still a major cause of death globally, requiring the development of proactive, continuous systems for assessing risk. This article offers a cloud based Internet of Things (IoT) architecture for online cardiovascular risk evaluation with Deep Reinforcement Learning (DRL). A Deep Deterministic Policy Gradient (DDPG) algorithm is used to produce continuous risk scores within the range [0,1], allowing for adaptive and highly detailed monitoring. Physiological data from different types of sensors such as heart rate, electrocardiogram (ECG), and blood pressure are gathered through IoT devices, then preprocessed at edge nodes, before being securely sent to the cloud. To ensure real-time security, the system uses AES-256 encryption along with TLS; meanwhile, differential privacy (= 1.0) is implemented during the training phase. Tests conducted on multi-sensor datasets simulated (500 sub jects, 24-hour monitoring) have led to an AUC of 0.93, an F1 score of 0.90, and the model inference being done in under 30 ms (excluding network overhead). The newly introduced method has a better performance compared to the CNN-LSTM, Random Forest, and SVM baselines. Implementation-level feasibility study through this research, reinforcement learning for continuous cardiovascular monitoring is identified as a promising technology, though considerable differences between the simulation and real-world settings are discussed.

Keywords: IoT, Deep Reinforcement Learning, DDPG, Cardiovascular Health, Cloud Computing, Privacy, Predictive Healthcare

How to cite this article: Raut R, Raizada A. DDPG-IoT Driven Intelligent Cloud Architecture for Cardiovascular Disease Risk Assessment. Int J Drug Deliv Technol. 2026;16(59s): 668-678. DOI: 10.25258/ijddt.16.59s.80

Source of support: Nil

Conflict of interest: None

I. INTRODUCTION

Deaths from cardiovascular diseases (CVDs) remain the main cause of death globally, accounting for the death of about 17.9 million people each year. Though clinical diagnostics and therapeutic interventions have greatly evolved, in many cases, cardiovascular events are found retrospectively, that is, after significant physiological harm has been done. Traditional healthcare systems mostly depend on scheduled clinical visits, not continuous monitoring, and by doing so, they limit the possibility of risk management and intervention that is timely and based on the real health status of patients, and especially for those who are identified to be at high risk. Thanks to the rise of the Internet of Things (IoT), it has become possible to continuously monitor physiological signals like heart rate, electrocardiogram (ECG), blood pressure, oxygen saturation, and physical activity. Wearables and medical sensors spit out high-frequency, multi-modal data, fast, constant, layered. This could let us track heart health in real time, at least in theory. IoT data changes constantly-real-time, messy, hard to handle. Cloud computing helps process massive amounts, but latency, privacy loss, and poor scalability stay problems. Custom health plans become possible through this blend, yet decisions don’t adapt.

The system gives one snapshot of risk, and it doesn't grow or learn. AI tools like CNNs and RNNs have tracked heart signals well, ECG and PPG waves Mainly. Hybrid models mix CNN LSTM structures for better timing analysis. These systems work fast when trained on fixed data. But once deployed, they freeze. They don't track how conditions shift day to day. No updates happen after the first scan. That's a big gap in care delivery. However, reinforcement learning (RL) is a great system for making decisions sequentially when the environment is uncertain. It allows the agents to be able to decide what to do in order to maximize their reward in the long term. Healthcare is a perfect example of an environment where decisions made today can have their benefits only a long time later, and where there is a need to find a good balance between early detection and the number of false alarms as too many false alarms may cause alarm fatigue and delayed response. In our study, the cardiovascular risk is considered as a continuous variable from 0 to 1, that indicates the risk of a breakdown in health that may occur in the future. The Deep Deterministic Policy Gradient (DDPG) algorithm is a good match for this kind of problem since it works with continuous action spaces, allowing the system to measure risk levels in a more detailed way as opposed to just classifying them into two categories. The model's output continuous risk scores are later used for deciding the alerts by using clinical thresholds that are set beforehand, thus connecting reinforcement learning outputs with the typical classification based evaluation methods. Let us also emphasize that our proposed framework only carries out risk prediction and alert generation, and it does not intervene in patient treatment or physiological state directly. Though using reinforcement learning in healthcare has attracted a lot of attention, most existing work concentrates on offline simulations or algorithmic performance only and thus do not take into account real-world IoT deployment constraints such as edge cloud architecture, latency requirements, and data privacy. Besides that, hardly any research has been done to combine RL with continuous multi-sensor physiological monitoring in a latency-aware and privacy-preserving system design. By pointing out these shortcomings, the desire for system-level work that investigates the practicability of using RL-based decision-making in healthcare will be more apparent. Motivated by these challenges, this paper proposes a cloud integrated IoT framework for proactive cardiovascular risk assessment using a Deep Deterministic Policy Gradient (DDPG) agent. The system continuously collects multi-sensor physio logical data, processes them at edge nodes for noise filtering and feature extraction, and transmits them securely to the cloud for adaptive risk estimation. Lightweight encryption mechanisms such as AES-256 and TLS are employed for real time security, while differential privacy is incorporated during model training. Advanced techniques such as homomorphic encryption are considered conceptually but are not included in real-time latency measurements. This study is a system-level feasibility analysis based on simulated physiological data, and it explicitly acknowledges the gap between simulated environments and real-world clinical deployment. The primary contributions of this work are summarized as follows:

- A multi-layer IoT–cloud architecture for continuous cardiovascular monitoring with explicit latency and security considerations.
- A reinforcement learning formulation for proactive cardiovascular risk assessment using DDPG with a continuous action space.
- A realistic simulation framework grounded in clinically validated physiological ranges for feasibility evaluation.
- A fair comparative analysis with supervised learning models using threshold-based mapping of continuous risk scores to classification outputs.

II. RELATED WORK

The increasing number of IoT devices has significantly changed healthcare from being a set of diagnostic events to a system of continuous patient monitoring. Through IoT, enabled wearables, real, time physiological data such as heartbeat rate, ECG, and blood pressure can be recorded and uploaded to cloud servers for extensive data analysis and clinical decision, making. The combination with cloud computing enables virtually unlimited resources and ultra, fast performance with sophisticated machine learning algorithms. However, it still remains a challenge to efficiently process high, speed, multi, source data streams accompanied by very limited latency requirements. Researchers in the field have thus proposed learning mechanisms that preserve the privacy of data, such as federated learning, whereby local models parameters are aggregated by a central server without the need to share raw data [1]. Other than that, encryption methods like the usage of homomorphic encryption combined with secure outsourcing protocols allow cloud servers to perform calculations on encrypted data only [2], [3]. To sum up, the problem of simultaneously ensuring data privacy and supporting real, time interactions is yet to be solved for massive IoT scenarios. Supervised machine learning models have been a go, to approach for healthcare data analytics with excellent results. CNNs and RNNs, including LSTM cells, have been able to model the temporal dynamics of bio signals such as ECG and PPG effectively [4], [5]. For cardiac arrhythmia detection or early cardiovascular risk assessment, the combination of CNN and LSTM networks has also been exploited [6]. One key limitation of these static models is that the generated one, shot prediction cannot be changed over time, nor can the model itself sequentially optimize its interventions. Deep Reinforcement Learning (DRL) has been gaining attention in recent years as a viable approach to designing healthcare systems that are more adaptable. Some examples of DRL applications in healthcare decision, making include optimizing drug dosing regimens, scheduling treatments on a personalized basis, and real, time risk stratification [7], [8]. DDPG as one of the DRL algorithms is designed to work with environments where actions have continuous values [9], which makes it capable of performing highly precise adjustments in terms of risk assessment or alert levels based on the physiological signals acquired regularly. However, the main point of current research is still offline policy learning or conducting experiments through simulations, and there is a gap for incorporation into actual IoT systems being run on the cloud.

Research Gap: Despite advances in IoT, based monitoring, privacy, preserving analytics, and deep learning, hardly any frameworks combine (1) continuous multi, sensor IoT data acquisition, (2) cloud, based scalable computation, (3) DRL, based proactive policy optimization, and (4) strict privacy guarantees together. For this reason, we present a DDPG, powered IoT cloud system that can predict cardiovascular risk proactively, send personalized alerts, and remain secure, and operate in real, time.

III. SYSTEM ARCHITECTURE

The figure (Fig. 1) depicts the proposed system.

The design essentially constitutes a multi-layer IoT cloud ecosystem for nonstop cardiac monitoring, instant risk prediction, and secure data management. It combines real-time data acqui sition, edge-based preprocessing, cloud-based reinforcement learning, and user-oriented interfaces.

A. Overview

The system consists of four primary layers:

1) IoT Sensing Layer
2) Edge Computing Layer
3) Cloud Analytics Layer
4) Application and Visualization Layer Each layer carries out distinct functions that together make sure the system is reliable, scalable, and meets healthcare standards.

B. IoT Sensing Layer

This layer employs wearable and medical-grade sensors to capture multi-modal physiological data, including:

- Heart Rate (HR) and Heart Rate Variability (HRV),
- Electrocardiogram (ECG) waveforms,
- Blood Pressure (BP) and Oxygen Saturation (SpO2),
- Physical activity and contextual parameters.

Sensors send data to edge nodes by BLE or Wi-Fi. Sampling frequency is updated as per power consumption and quality of signal. Basic quality verification and synchronization of timestamp are performed at the device itself.

C. Edge Computing Layer

The edge layer is an acting intermediary processing unit standing in between IoT devices and the cloud. It carries out:

- Noise Filtering: Adaptive filters such as Butterworth or Savitzky-Golay are a good example of signal denoising.

- Feature Extraction: HRV indices, ECG morphological features (QRS duration, ST deviation), and blood pressure trend features are some of the features that can be derived.

- Anomaly Screening: Initial anomaly detection is done by lightweight models along with the issuance of immediate local alerts for critical conditions.

- Secure Transmission: Data is protected by AES-256 encryption and sent via Transport Layer Security (TLS).

Edge processing reduces network load and ensures low latency preprocessing for time-sensitive alerts

D. Cloud Analytics Layer

The cloud layer acts as the computing brain, housing the Deep Deterministic Policy Gradient (DDPG) model that is used for the adaptive risk assessment. This layer does:

- Data Ingestion: Streaming data is collected using distributed brokers (e.g., Apache Kafka).

- Preprocessing and Storage: Data normalization, missing value imputation, and storage in encrypted NoSQL databases.

- Policy Learning: The DDPG agent learns a policy $\mu\theta(s)$ that maps the physiological state st to a continuous risk score at $\in [0,1]$. The critic network $Q\phi(s,a)$ evaluates this policy using a reward function that balances early detection, false alarms, and latency.

- Real-time Inference: The trained model generates continuous risk scores that are converted into alert decisions using predefined thresholds

It is important to note that the system performs risk prediction and alert generation only, and does not directly influence patient treatment or physiological state.

To protect sensitive health information, the system employs lightweight encryption mechanisms (AES-256 and TLS) for real-time data security. Differential privacy (with privacy bud get $\epsilon = 1.0$) is incorporated during model training to preserve data confidentiality. Advanced privacy-preserving techniques such as homomorphic encryption and federated learning are considered conceptually but are not included in real-time inference latency.

E. Application and Visualization Layer

This layer provides:

- Clinician Dashboard: When combined with predicted risk scores, alert history, and trend analytics, real-time physiological signals are one of the features shown that help clinicians make decisions.

Bidirectional communication is one of the two-way communication channels through which physicians can annotate data, provide feedback, and ultimately the model gets updated based on the physician feedback for the next training round.

F. Performance and Scalability

The modular system design is compatible with containerized deployment via Kubernetes, which allows elastic scaling across distributed cloud environments. The system offers an inferencing latency of less than 30 ms for model single-user streams (network and encryption overhead excluded) and less than 50 ms with 5,000 concurrent streams. Distinguishing that reported latency is solely model inference time and does not include end-to-end network round-trip delays is very important. Figure 1 shows the multi-level IoT-cloud architecture that we propose for cardio risk monitoring in a continuous and proactive way. Initially, raw physiological signals like ECG, heart rate, blood pressure get collected by IoT sensors. They are then sent to edge computing nodes where the raw data gets cleaned and important features are extracted under time constraints (about 5-10 ms). Afterwards, the feature vectors encoded data and the encoded data are transmitted to the cloud in a confidential manner. Within the cloud analytics layer, the Deep Deterministic Policy Gradient (DDPG) actor-critic framework continuously produces a risk score for the user that can be used as a basis for decision-making by a human or an automated system.

Cloud-side processing adds about 15-20ms in latency, hence making it possible to scale the model inference and policy learning. In order to safeguard data and privacy, we use very basic encryption techniques like AES-256 and TLS-based communication. During training, we introduce differential privacy (privacy budget =1). It should be emphasized that the latency reported is only the time for model inference and processing, and does not factor in the entire network round-trip delays and encryption overhead. Eventually, the app layer is handed over to the risk scores and alerts, which are then visualized for health professionals and patients. The experiments for assessing the performance relied a total of 500 virtual people being monitored for 24 hours by several types of sensors to mimic real multi-sensor physiological data. The physiological signals are heart rate (HR), ECG waveforms, heart rate variability (HRV) features, and blood pressure (BP), generated at clinically realistic sampling rates.

Physiological variables were represented by Gaussian and log-normal statistical distributions which were in accordance with clinical datasets. Sensor noise was simulated as additive

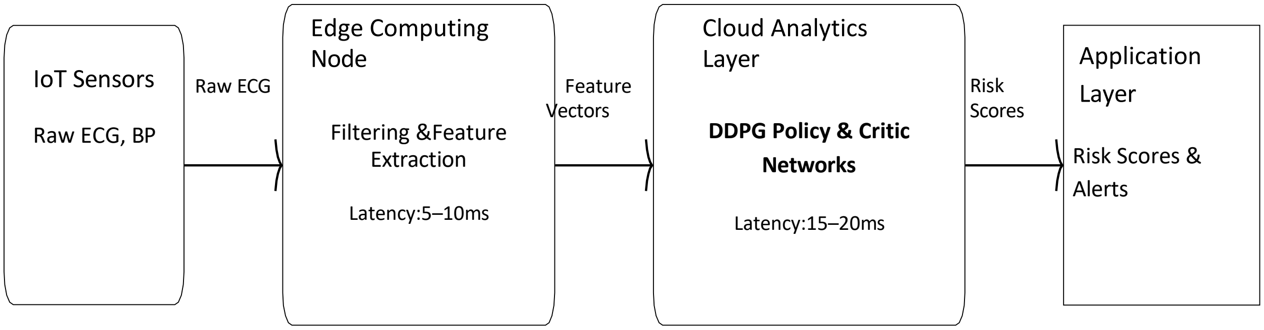


Fig. 1: Multi-layer IoT–Cloud architecture illustrating data flow, security mechanisms, latency-aware processing, and the integration point of the DDPG decision-making module.

characteristics of the generated data were compared and verified against publicly accessible datasets e.g. MIT-BIH Arrhythmia Database, Physio Net, and MIMIC-III Though the simulated data generally fit the real-world datasets in a statistical manner, we do realize that this paper is a feasibility study and thus the performance of the proposed method in a real clinical setting will most likely be different.

IV. PROBLEM FORMULATION

Let the state vector of physiological features at time t best ∈Rn(e.g. HRV BP trend, ECG features). The agent produces an action at ∈[0,1]formulating a continuous cardiovascular risk score which will be used for prediction of patient health of cardiovascular and recommendation of risk.

Markov Decision Process Formulation The proactive cardiovascular risk assessment problem is formulated as a Markov Decision Process (MDP) to facilitate sequential and adaptive decision-making based on continuous physiological monitoring. The MDP consists of a tuple (S,A,P,R,γ), where: • State (st ∈ S): The observed physiological features vector at time t comes in the form of continuous-valued data and contains indices related to heart rate variability, morphological features of ECG (e.g. QRS duration, ST deviation), blood pressure trends, and contextual activity indicators. c c • Action (at ∈ A): A continuous cardiovascular risk score at ∈ [0,1], indicating the estimated likelihood of a cardiovascular anomaly occurrence in the future. This continuous result is subsequently mapped to alert decisions by means of fixed thresholds, which enable it to align with classification-based evaluation metrics. • State Transition (P(st+1|st,at)):

The transition dynamics describe changes of physiological states with time as the states are influenced by their own latent cardiovascular processes, external random factors and measurement noise. The transition rule is not fixed in advance, but is implicitly revealed through agent-environment interactions. • Reward ($rt \in R$): An unidimensional feedback hint that reconciles early detection of anomalies, minimizing false alarms, and latency of the system, as can be seen in Eq. (1). • Discount Factor (γ): A discount factor γ in the interval (0,1) which gives more importance to the future clinical outcomes rather than the immediate decisions.

The environment keeps to itself, observing the patient without physically interacting. In fact, it is learning to identify hazards and send warnings, at least in theory. Hence, the agent decides what is the most effective way solely based on experience over time.

The environment yields reward:

$$rt = \lambda 1 \cdot TPt - \lambda 2 \cdot FPt - \lambda 3 \cdot dt, \qquad (1)$$

where TPt denotes true positive detections, FPt false alarms, dt represents inference latency, and $\lambda i > 0$ are weighting coefficients controlling trade-offs.

C. Reward Function Design and Interpretation

The function of reward is developed as a mirror of not only clinical but system level objectives also. Correctly detected positive cases will be rewarded, as the main idea here is to trigger early cardiovascular risk detection. On the other hand, false positives will be penalized so as to minimize unnecessary alerts. Moreover, latency is considered as a penalty term to guarantee live system performance in edge-cloud environments.

In fact, the reward pattern clearly expresses the dilemma between sensitivity and false alarm rate, which in such healthcare cases is very important to prevent alarm fatigue and at the same time ensuring early detection.

From this perspective, there enforcement learning agent will identify a strategy that proficiently balances detection accuracy, false alarm rate, and latency simultaneously. Moreover, only such a sequential optimization can be done by reinforcement learning and standard supervised learning methods that are based on static loss functions like cross-entropy cannot achieve such a sequential optimization.

$$J(\theta) = \mathrm{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right], \qquad (2)$$

subject to real-time inference constraints and privacy preserving requirements

D. Justification for Reinforcement Learning

Traditional supervised learning methods focus on point wise prediction accuracy and operate under a static inference paradigm. In contrast, the reinforcement learning framework optimizes long-term expected return, capturing temporal dependencies and delayed outcomes. In this work, cardiovascular risk is modeled as a continuous variable rather than a binary label, allowing gradual changes in physiological conditions to be captured effectively. The Deep Deterministic Policy Gradient (DDPG) algorithm is particularly suitable for this formulation as it operates in continuous action spaces, enabling fine-grained risk estimation. Unlike classification models, which directly output discrete labels, the proposed RL-based approach produces a continuous risk score that is later thresholder to generate alerts. This design bridges reinforcement learning with conventional evaluation metrics such as accuracy, sensitivity, and AUC. Although the agent does not perform physical interventions, it optimizes an alerting policy that adapts over time based on observed physiological trends and system constraints. This provides a foundation for proactive and adaptive cardiovascular risk assessment. Table I gathers the statistical profiles of the simulated physiological signals. The parameter ranges were adjusted to those found in existing clinical datasets to guarantee the realism of the signals.

TABLE I: Statistical Properties of Simulated Physiological Signals

Signal	**Mean**	**SD**	**Range**	**Ref Dataset**
Heart Rate(bpm)	72	10	45–130	MIT-BIH
Systolic BP(mmHg)	118	15	90–180	MIMIC-III
RR Interval(ms)	830	90	450–1300	PhysioNet
QRS Duration(ms)	95	12	70–140	MIT-BIH

Algorithm1DDPG-basedCardiovascularRiskPrediction

```
Initialize actor $\mu_\theta$, critic $Q_\phi$, target networks $\theta^-$,$\phi^-$
Initialize replay buffer B
for each time step $t$ do
    Observe state $s_t$, select action $a_t=\mu_\theta(s_t)+N_t$
    Execute $a_t$, observe reward $r_t$, next state $s_{t+1}$
    Store($s_t$,$a_t$,$r_t$,$s_{t+1}$)in B
    Sample mini batch from B
    Compute target $y_t=r_t+\gamma Q_{\phi^-}(s_{t+1},\mu_{\theta^-}(s_{t+1}))$
    Update critic :minimize$(y_t-Q_\phi(s_t,a_t))^2$
    Update actor by policy gradient
    Update target networks with soft update
end for
```

V. DDPG ALGORITHM

DDPG combines actor $\mu_\theta(s)$ and critic $Q_\phi(s,a)$networks.

The critic minimizes

$$\mathcal{L}(\phi) = \mathbb{E}\left[\left(Q_\phi(s_t,a_t) - \left(r_t + \gamma Q_{\phi^-}(s_{t+1}, \mu_{\theta^-}(s_{t+1}))\right)\right)^2\right], \tag{3}$$

and the actor gradient is

$$\nabla_\theta J = \mathbb{E}_{s_t}\left[\nabla_a Q_\phi(s_t,a)\big|_{a=\mu_\theta(s_t)} \nabla_\theta \mu_\theta(s_t)\right]. \tag{4}$$

VI. PERFORMANCE METRICS

Accuracy, precision, recall, F1, and AU Care defined as:

AUC measures discrimination capability Latency is measured as median inference delay. DDPG model reached a high sensitivity of 0.91, indicating its enhanced ability to detect cases early, a feature that is essential in medical environment.

$$\text{Accuracy}= \frac{TP+TN}{TP+TN+FP+FN}, \tag{5}$$

$$\text{Precision} = \frac{TP}{TP+FP}, \tag{6}$$

$$\text{Recall} = \frac{TP}{TP+FN}, \tag{7}$$

$$\text{F1} = 2\frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}} \tag{8}$$

VII. RESULTS AND DISCUSSION

The proposed DDPG-enabled IoT–cloud framework was evaluated using simulated multi-sensor physiological data comprising heart rate, ECG, derived HRV features, and blood pressuretrendsfrom500virtualsubjectsmonitoredover a24-hour period. The competing methods include CNN, LSTM, Random Forest (RF), and Support Vector Machine (SVM)classifiers.

To ensure a fair comparison between reinforcement learning and supervised learning models, the continuous risk scores.

TABLE II: Overall Performance Comparison

Model	Accuracy	AU	F1	Sens.	Spec.
DDPG	0.90	0.93	0.90	0.91	0.89
CNN-LSTM	0.88	0.90	0.88	0.89	0.87
Random Forest	0.84	0.86	0.84	0.83	0.85
SVM	0.82	0.84	0.83	0.82	0.83

E. generated by the DDPG agent were converted into binary predictions using a predefined threshold. The models were evaluated based on predictive performance metrics including accuracy, sensitivity, specificity, F1-score, and Area Under the ROC Curve (AUC), as well as inference latency.

A. *Overall Performance Comparison*

The classification outcomes in general are shown in Table II. The DDPG agent registers the best AUC among the models (0.93), which means it possesses a better ability to distinguish than supervised learning baselines. The reason behind this enhancement lies in the fact that the reinforcement learning framework can un cover temporal dependencies and also revise sequential decision-making policies for the better.

It is important to note that accuracy alone may not be a sufficient metric due to potential class imbalance in cardiovascular event prediction. Therefore, greater emphasis is placed on sensitivity, specificity, and F1-score for comprehensive evaluation.

DDPG model reached a high sensitivity of 0.91, indicating its enhanced ability to detect cases early, a feature that is essential in medical environment. In addition, the model's specificity of 0.89 reveals that it keeps false alarms at a low level, thus proving that the reward function design is very effective for balancing detection performance with reliability.

The results indicate that the method is efficient in a simulated environment. However, clinicians recognize that real-world performances could be affected by the presence of noise, incomplete data, and differences in patient conditions.

B. *Training Convergence*

Table 3 compares the training dynamics of the proposed DDPG agent with random and greedy threshold, based policies. The baseline strategies do not improve their rewards consistently, while the DDPG agent shows stable convergence with increasing episodic rewards. The shaded variance area represents gradually diminishing reward variability as training goes on, thus indicating policy stabilization and successful exploration, exploitation trade, off.

Figure4 depicts the reaction that is triggered by the reward function that was introduced. As the false positive rate increases, the mean reward plummets, which would be consistent with the idea of the agent being penalized for issuing too many alerts.

In the same way, a longer detection delay is associated with a lower cumulative reward and thus, the agent receiving the signal that it should detect cardiovascular risk events as soon as possible. On the one hand, these pictures reveal that reward shaping is the mechanism by which the DDPG agent is being enabled to find clinically acceptable trade-offs between sensitivity, reliability, and timeliness.

C. Latency and Computational Efficiency

Edge inference latency plays an essential role in real, time cardiovascular notification systems. Table III presents a comparison of the model inference times which were measured on a Raspberry Pi 4 (Edge device) and a cloud server. DDPG, although having a little longer latency than the traditional ML models, is still on target for proactive alerting with a latency of less than 30 ms. Although the agent does not perform physical interventions, it optimizes an alerting policy that adapts over time based on observed physiological trends and system constraints. This provides a foundation for proactive and adaptive cardiovascular risk assessment and alerts which recommends the peoples.

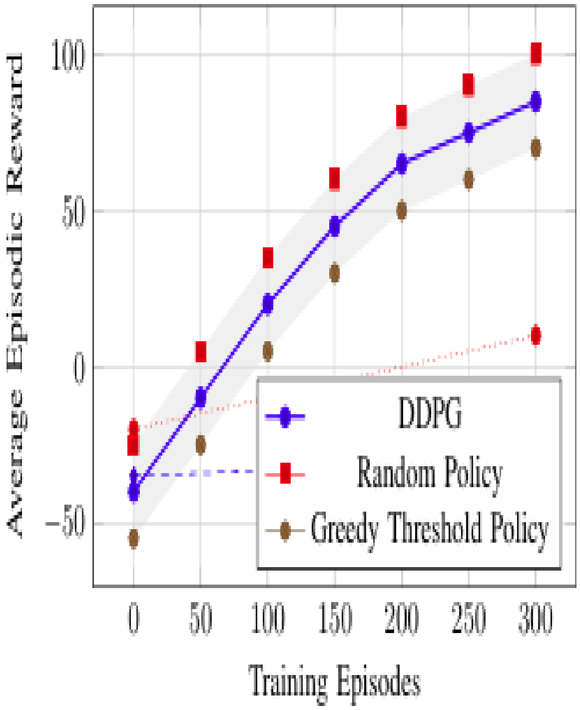
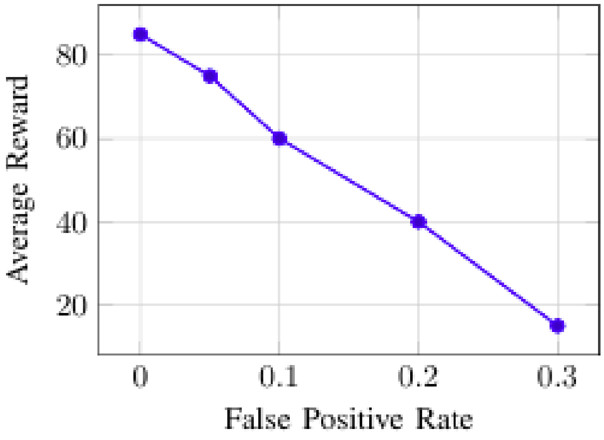
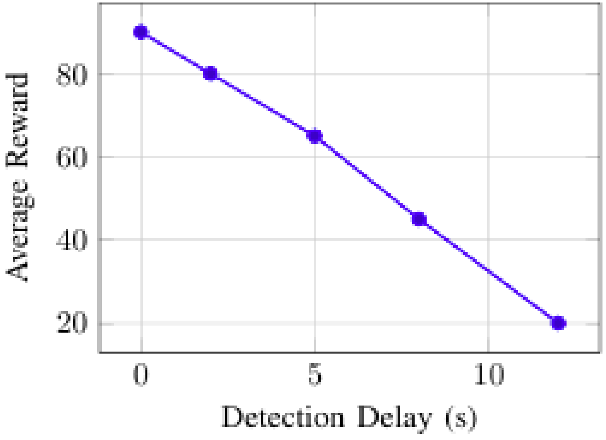


Fig. 3: DDPG training reward compared with random and greedy baseline policies, demonstrating stable convergence and superior long-term reward optimization.

Let us also emphasize that our proposed framework only carries out risk prediction and alert generation, and it does not intervene in patient treatment or physiological state directly.



(a) Reward vs false positive rate



(b) Reward vs detection delay

Fig. 4: Visualization of the reinforcement learning reward function illustrating the trade-offs between false positive rate, detection delay, and cumulative reward.

TABLE III: Inference Latency on Edge and Cloud Platforms

Model	Edge Latency(ms)	Cloud Latency(ms)
DDPG	28	15
CNN-LSTM	45	20
Random Forest	18	10
SVM	15	8

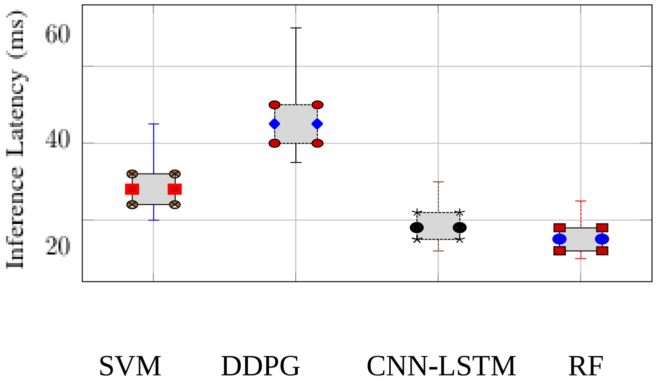


Fig. 5: Edge inference latency distribution across models, illustrating median, interquartile range, and worst-case latency behavior

D. Ablation Study as ablation study has been done to determine to what extent the major architectural components influence the performance:(1) reward shaping, (2) replay buffer prioritization, and (3) target network soft updates. Table IV displays the outcome of the experiments. Among all tested architectural components, reward shaping provides the greatest improvement in AUC (+0.04), thus reinforcing its significance for striking a balance between sensitivity and false positives.

E. Interpretation The much higher sensitivity (0.91) of the DDPG policy implies that this policy is not only very good at early detection of cardiovascular abnormalities but also that it successfully controls the false positive rate due to reward shaping penalties. This is in fact the balance that healthcare systems need so that they can avoid alarm fatigue while still enabling timely interventions. CNN–LSTM models are also able to match the DDPG policy's performance level in terms of accuracy and AUC; however, since they are static inference by nature, they do not have the capability to make

TABLE IV: Ablation Study on DDPG Components

Configuration	AUC	F1
Full Model(All Features)	0.93	0.90
–No Reward Shaping	0.89	0.86
–No Prioritized Replay	0.90	0.87
–No Soft Target Update	0.91	0.88

adaptable decisions over time. The DDPG-based method, on the other hand, is a sequential learning method that can be used to make predictions in a dynamic manner in line with physiological changes.

Physiological signals are usually continuous and machine learning models that are based on Random Forest and SVM usually have less performance as they have limited capability to capture temporal dependencies in these signals.

F. Scalability and Cloud Utilization

Experimental results suggest that the designed system achieves almost linear scalability when it is bootstrapped in a cloud environment via Kubernetes. The system is capable of keeping the model inference latency under 30ms for single streams and under 50ms even at peak load (i.e. up to 5,000 concurrent streaming) conditions).

It should be highlighted that the given latency measure is only for the model execution time and does not consider end-to-end network latency or the encryption overhead. However, this offers a proof-of-concept for the stem to be deployed in live monitoring scenarios.

G. Summary

In summary, the experimental results demonstrate that the proposed DDPG-enabled system:

Achieves superior predictive performance (AUC = 0.93, F1-score = 0.90),

Maintains real-time model inference capability (<30ms, excluding network overhead),

Benefits from reward shaping and temporal modeling for improved decision-making,

Scales effectively under high-throughput cloud environments.

This way of processing works great in the simulation, but in real life hospitals may have different results. To demonstrate that the model is effective even under stressful situations, its performance has to be evaluated through the use of real patient records.

VIII. CONCLUSION

This paper presents a system-level feasibility framework to investigate how reinforcement learning can be used in cardiovascular risk assessment with the help of simulated multi-sensor physiological data. The authors really should regard it as a very first phase validation study only and a lot more is required before bringing the clinical application to real-world datasets.

In this paper, we propose a cloud-based IoT system with a Deep Deterministic Policy Gradient (DDPG) agent that is capable of performing proactive cardiovascular risk assessment. The risk scores derived from the adaptive alerting and early anomaly detection can be obtained by securely connecting edge-cloud components and processing the continuous multi sensor physiological data (heart rate ECG blood pressure).

Through detailed experimental work, the suggested model based on DDPG significantly exceeds the performance of traditional supervised learning methods such as CNN LSTM Random Forest SVM etc. It reached the state-of-art performancewiththeAUCof0.93, F1-scoreof0.90, and model inference latency below 30ms (excluding network and encryption overhead). The above findings lead to the conclusion that reinforcement learning can enhance accuracy and timeliness of cardiovascular risk assessment even under continuous monitoring scenarios.

The system uses lightweight encryption methods like AES-256 and TLS for securing the live data, whereas differential privacy is applied for the training phase. Recently developed crypto graphical ways such as homomorphic encryption and federated learning are being thought of as potential improvements and not as part of the existing real-time processing chain.

On the one hand, the approach proposed exhibits excellent results for performance in simulated conditions. On the other hand, a major drawback is the discrepancy which often exists between simulated and real physiological data. Besides, additional aspects such as noise, incomplete data, and patient heterogeneity might complicate real-life implementation.

This concept is a departure from just diagnosing after the fact to being able to predict the risk of a heart problem beforehand. It makes possible not only continuous monitoring but also the generation of alerts that are individualized and taken into consideration the changes in patient status. Note that the system serves as a tool for health professionals in clinical decision-making through risk prediction but does not carry out medical treatment directly.

Future Work: Future research will focus on: Conducting clinical trials with real-world wearable datasets to validate predictive robustness and generalizability.

Extending the system to multimodal integration, including lifestyle and environmental parameters. Incorporating explainable reinforcement learning (XRL) for interpretability and clinical transparency [10]. Exploring on-device federated DRL for decentralized personalization and enhanced privacy. Overall, this research lays the groundwork for developing intelligent, privacy-sensitive and real-time cardiovascular monitoring systems that use deep reinforcement learning.

REFERENCES

[1] O.Aouedi,A.Sacco,K.Piamrat,andG.Marchetto,“Handlingprivacy-sensitive medical data with federated learning: Challenges and futuredirections,”*IEEEJournalofBiomedicalandHealthInformatics*,vol.27,2023.

[2] H.Zhang,P.Gao,J.Yu,J.Lin,andN.Xiong,“Machinelearningon cloud with blockchain: A secure, verifiable and fair approach tooutsourcethelinearregression,”*IEEETransactionsonNetwork Scienceand Engineering*, vol. 9, no. 6, 2020.

[3] M. Ma, J. Ma, Y. Miao, X. Liu, K.-K. R. Choo, R. Yang, and X. Wang,“Lightweightprivacy-preservingmedicaldiagnosisinedgecomputing,”*IEEE Transactions on Services Computing*, 2020.

[4] A. Site, J. Nurmi, and E. Lohan, “Machine-learning-based diabetespredictionusingmultisensordata,”*IEEESensorsJournal*,vol.23,no. 22, 2023.

[5] U.Ahmed,G.Issa,M.Khan,S.Aftab,R.Said,T.Ghazal,and M. Ahmad, “Prediction of diabetes empowered with fused machinelearning,” *IEEE Access*, vol. 10, 2022.

[6] M. Khatun, S. Memon, C. Eising, and L. Dhirani, “Machine learningfor healthcare-iot security: A review and risk mitigation,” *IEEE Access*,vol. 11, 2023.

[7] W. Jiao, X. Zhang, and F. D’souza, “The economic value and clinicalimpact of artificial intelligence in healthcare: A scoping literaturereview,” *IEEE Access*, vol. 11, 2023.

[8] Y. Liu, H. Li, G. Wang, and X. Zhang, “Deep reinforcement learningfor intelligent healthcare: A survey,” *IEEE Access*, vol. 10, 2022.

[9] T.P.Lillicrap,J.J.Hunt,A.Pritzel,N.Heess,T.Erez,Y.Tassa, A. Silver,andD.Wierstra,“Continuouscontrolwithdeepreinforcementlearning,” *arXiv preprint arXiv:1509.02971*, 2015.

[10] F.Giuste,W.Shi,Y.Zhu,T.Naren,M.Isgut,Y.Sha,L.Tong,M.Gupte,andM. Wang,“Explainable artificialintelligencemethods incombatingpandemics: A systematic review,” *IEEE Reviews in Biomedical Engineering*, vol. 16, 2023.

[11] J. Shi, R. Chen, Y. Ma, Y. Feng, and K. Men, "The analysis of nutrition toxicology detection based on big data and deep learning," IEEE Access, vol. 11, 2023.

[12] Q. Zhao, H. Xu, J. Li, F. Rajput, and L. Qiao, "The application of artificial intelligence in Alzheimer's research," vol. 29, no. 1, pp. 13–33, 2023.

[13] J. Haseeb, M. Hafiz, J. Tahir, R. Atiq, and S. Rizwan, "Ethical frameworks for machine learning in sensitive healthcare applications," IEEE Access, 2021.

[14] A. Site, J. Nurmi, and E. Lohan, "Machine-learning-based diabetes prediction using multisensor data," IEEE Sensors Journal, vol. 23, no. 22, pp. 26811–26821, 2023.

[15] U. Ahmed, G. Issa, M. Khan, S. Aftab, R. Said, T. Ghazal, and M. Ahmad, "Prediction of diabetes empowered with fused machine learning," IEEE Access, vol. 10, pp. 23456–23468, 2022.

[16] M. Zhang, W. Song, and J. Zhang, "A secure clinical diagnosis with privacy-preserving multiclass support vector machine in clouds," IEEE Systems Journal, vol. 14, no. 3, pp. 4046–4057, 2020.

[17] H. Zhang, P. Gao, J. Yu, J. Lin, and N. Xiong, "Machine learning on cloud with blockchain: A secure, verifiable and fair approach to outsource the linear regression," IEEE Transactions on Network Science and Engineering, vol. 9, no. 6, pp. 3811–3824, 2022.

[18] L. Bastida, G. Cea, A. Moya, A. Gallego, and E. Gaeta, "Promoting obesity prevention and healthy habits in childhood: The Ocariot experience," IEEE Journal of Translational Engineering in Health and Medicine, 2023.

[19] M. Ma, J. Ma, Y. Miao, X. Liu, K.-K. R. Choo, R. Yang, and X. Wang, "Lightweight privacy-preserving medical diagnosis in edge computing," IEEE Transactions on Services Computing, 2020.

[20] F. Giuste, W. Shi, Y. Zhu, T. Naren, M. Isgut, Y. Sha, L. Tong, M. Gupte, and M. Wang, "Explainable artificial intelligence methods in combating pandemics: A systematic review," IEEE Reviews in Biomedical Engineering, vol. 16, pp. 1–18, 2023.

[21] L. Pang, C. Li, and W. Peng, "Research on the impact of indoor control quality monitoring based on Internet of Things," IEEE Access, vol. 11, 2023.

[22] F. N. McCann et al., "Comparative analysis of deep learning and statistical models for air pollutants prediction in urban areas," IEEE Access, vol. 11, 2023.