Received: 30 May 2025 | Revised: 3 December 2025 | Accepted: 28 January 2026 | Published online: 12 February 2026

RESEARCH ARTICLE

Artificial Intelligence and Applications
2026, Vol. 00(00) 1–17
DOI: 10.47852/bonviewAIA62026306

BON VIEW PUBLISHING

Deep Reinforcement Learning-Enabled IoT Framework for Real-Time and Personalized Diabetes Diagnosis Using Wearable Sensors

M. Alamelu[1], Meera Alphy[2], Finney Daniel Shadrach[3,*] and Jayaraj Velusamy[4]

[1]*Department of Information Technology, Kumaraguru College of Technology Coimbatore, India*

[2]*Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology, India*

[3]*Department of Electronics and Communication Engineering, KPR Institute of Engineering and Technology, India*

[4]*Electronics and Communication Engineering, Nehru Institute of Engineering and Technology, India*

Abstract: The increasing number of diabetes patients worldwide reveals an immediate necessity for technical diagnostic solutions that offer smart operations and full data protection capabilities. This research presents FL-Hybrid, which unites the CNN-LSTM-Attention network structure together with Dueling Deep Q-Network (DQN) to perform individual diabetes screening using wearable sensor information through federated learning. The PhysioNet BIG IDEAs dataset with CGM and physiological sensor signals feeds into the model that combines spatiotemporal transformers with context-driven decision functions. The system protects user privacy through its deployment on edge devices, combined with federated learning training that eliminates the need to send sensitive information to centralized servers. The CNN-LSTM backbone finds deep physiological patterns, and the attention mechanism focuses on crucial diagnostic indicators such as post-meal glucose spikes and sleep-related heart rate variability decreases. The Dueling DQN agent develops optimum diagnostic actions through integration of clinical accuracy with efficiency and prediction reliability measures. Experimental tests conducted on multiple models show that FL-Hybrid outperforms all other systems by reaching a 97.85% accuracy rate with 97.31% precision, 97.96% recall, and 98.41% AUC-ROC. These results exceed those of CNNs and RNNs as well as centralized DRL models. The model displays 95.34% accuracy when noise occurs while also demonstrating exceptional user adaptation ability at a 95.62% adaptation rate. The proposed system provides organizations with a scalable, privacy-focused solution for continuous diabetes monitoring, which represents an advanced mobile health technology.

Keywords: diabetes diagnosis, attention mechanism, Dueling DQN, personalized health care

1. Introduction

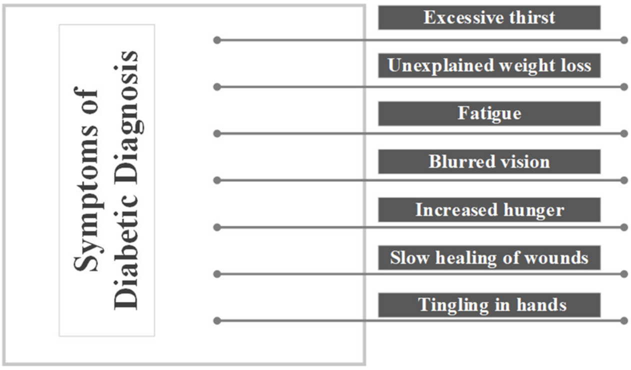
Diabetes is a chronic metabolic disorder characterized by high blood glucose levels resulting from the body's inability to produce or effectively use insulin. With its rising global prevalence, diabetes has become a significant public health concern, contributing to serious complications such as cardiovascular disease, kidney failure, nerve damage, and vision loss [1, 2]. Timely and accurate diagnosis plays a crucial role in preventing or managing these complications. The diagnosis of diabetes remains most securely performed using three standard tests: fasting blood glucose (FBG), oral glucose tolerance test (OGTT), and HbA1c tests. These testing approaches need periodic checking, yet they fail to monitor instant blood sugar changes primarily during times of early diabetes or pre-diabetes [3, 4]. The development of artificial intelligence (AI) and machine learning (ML) and deep learning (DL) models produced advanced diagnostic capabilities for diabetes. Such tools use massive medical data for detecting weak patterns while advising on early signs of diabetes. The implementation of Internet of Things (IoT) devices featuring continuous glucose monitors and biosensors enables simultaneous real-time noninvasive observation [5]. Healthcare professionals still have to address several obstacles to fully realize AI benefits because of data privacy risks and requirements for diverse large datasets and the need to explain AI models and their incorporation within existing healthcare infrastructure. The development of efficient diagnostic tools needs to resolve existing challenges to create tools that help early detection and deliver improved lifestyle quality for diabetic and pre-diabetic patients [6, 7].

Multiple diagnostic hurdles persist, which make testing and managing diabetes effectively more challenging despite recent advances in diagnostic technologies. The three commonly used diabetic diagnostic tests, including FBG, OGTT, and HbA1c, cannot detect diabetes in its earliest stages or silent cases, thus resulting in late medical care [8, 9]. Glucose tests provide static measurements instead of continuous monitoring because this specific monitoring method is essential for correct disease identification and tracking. The price of real-time continuous glucose monitoring devices remains high as they need calibration and do not provide universal accessibility, especially in low-resource situations. The field of diabetic diagnosis uses deep learning as a robust developing technique to detect patterns in medical data which standard diagnostic approaches cannot identify. CNNs and similar deep learning techniques achieve high evaluation results when detecting diabetic retinopathy in retinal fundus images while also determining glucose levels through analysis of physiological signals [10, 11]. The automated feature extraction capability of these models

*Corresponding author: Finney Daniel Shadrach, Department of Electronics and Communication Engineering, KPR Institute of Engineering and Technology, India. Email: finneydaniel@kpriet.ac.in

makes early detection of diabetes along with its complications possible by removing the necessity of manual input. Deep learning systems that work with big data can develop self-learning capabilities which enable precise fast medical diagnosis while requiring minimal human involvement [12]. Figure 1 shows the common symptoms of diabetic diagnosis.

Figure 1
Common symptoms of a diabetic diagnosis



Diabetic diagnosis of deep learning has major limitations of requiring large-sized, correctly labeled datasets, challenging to acquire because of privacy concerns and varying data formats [13, 14]. There could be bias in algorithms because of limited data variety, which can impact the predictions of algorithms on various populations. Moreover, deep learning models are computationally intensive, an aspect that raises the cost of deployment and makes them difficult to apply in resource-limited healthcare facilities. These technical, logistical, and fairness issues need to be addressed as soon as possible to guarantee ethical and effective application [15–18]. In this study, we propose a new Federated Learning Hybrid (FL-Hybrid) model that performs both individualized diagnosis and privacy-preserving diagnosis of diabetes through time-series data collected by wearable sensors. The proposed system unites the temporal sequence processing from CNN-LSTM networks with attention mechanism interpretation methods together with Dueling Deep Q-Networks (DQN) for real-time dynamic diagnostic action optimization. Through federated learning, the model can conduct distributed training on various users' devices without accumulating their confidential health information at one location for privacy protection including HIPAA and GDPR standards.

1.1. Main contribution of the work

1) The network involves CNN-LSTM-Attention and Dueling DQN to provide the customized diagnosis of diabetes.
2) Federated learning allows updating the world model and ensures the privacy of local data.
3) Attention mechanisms would prioritize such vital time points as glucose sudden elevations and HRV variations.
4) Reinforcement feedbacks enable the system to change according to user and clinical reactions.
5) An early risk detection and proactive health decision system gets underpinned by a reward-based system.

The remainder of this paper is structured as follows. Section 2 provides an overview of related studies focusing on diabetes prediction using wearable sensor data, deep learning techniques, and federated learning approaches. It highlights existing gaps in personalization, privacy, and decision adaptability. Section 3 details the proposed methodology, including the CNN-LSTM-Attention encoder, reinforcement learning policy design, and federated edge deployment strategy. Section 4 presents the experimental results and a comprehensive discussion on performance metrics, robustness, and system efficiency. Finally, the "Conclusion and Future Work" section summarizes key contributions and outlines directions for expanding the model's capabilities across broader health domains.

2. Related Works

High-level IoT systems allow smooth connection of gadgets to assess medical conditions. Web-connected sensing in the management of diabetes provides a precise, real-time measurement of blood glucose levels, aiding in self-management and avoiding overreliance by patients. Since diabetes is a chronic ailment, monitoring systems are required, which are invasive-free and effective [19]. An IoT-driven healthcare system was proposed that would integrate big data analytics, machine learning, and cloud technology that would help in processing information about diabetes and improving patient care outcomes and autonomy.

A decision tree-based diabetes prediction algorithm (DPA) using IoT-enabled sensor data in early detection of diabetes was proposed. Configuring 15,000 records, including 11,250 training records and 3,750 testing records, DPA performed remarkably well with an accuracy of 92.60% precision, 89.17% specificity, 90.02% accuracy, and an error rate of 9.98% [20]. These parameters outperform the usual models when it is time-sensitive to detect diabetes. With machine learning, the system is self-recognizable, and the patterns in the health data are not coded manually, enhancing automation in the healthcare system as well as diagnostic efficiency.

Machine learning also changes health care since it minimizes human touch and increases the accuracy of diagnosis. In the study, the authors point out feedforward neural networks (FNNs) concerning diabetes prediction and also mention the issues concerning navigating within the so-called complex multimodal data landscape that impacts the accuracy [21]. The prediction is enhanced utilizing the Pima Indians Diabetes Dataset using optimization techniques such as Whale Optimization Algorithm (WOA) and Particle Swarm Optimization (PSO). Evaluation of performance is done by values of best fit, means, medians, and standard deviations that help to enhance the choice of decisions during the diabetes diagnosis.

The diabetic retinopathy (DR) originates as a dire complication of long-lasting diabetes and impairs vision, becoming one of the leading causes of blindness worldwide. The typical classification of DR was based on the early lesions, such as exudates or aneurysms, with disregard of severe signs [22]. The focus is to suggest a deep learning solution, which is to perform feature extraction of comprehensive lesions based on APSO-powered GoogleNet and ResNet. Classifiers are trained on the hybrid model features, and this is done by using two or three of the following: random forest, SVM, decision tree, and linear regression as proper DR diagnosing tools.

The International Diabetes Federation defines that, out of the population of diabetic people, 33% of them have diabetic retinopathy as a significant cause of blindness and loss of visual capabilities. Automated diagnostic tools are becoming important because the number of diabetes cases will amount to 700 million by the year 2045. The proposed method is the Robust Fuzzy Local Information K-Means Clustering algorithm for robust learning in classifying retinal images [23]. This achieves improved precision by allowing localized data in every cluster. It provides better results when compared with the Modified Fuzzy C-Means clustering with better image segmentation and higher reliability in the diagnosis.

Federated learning (FL) has rapidly evolved in health care, enabling privacy-preserving collaborative model training across distributed institutions [24]. A systematic review highlights FL's role in tackling siloed datasets, class imbalance, and data shifts via distributed optimization, with applications in disease diagnosis like COVID-19 detection from X-rays, sepsis prediction, and ECG arrhythmia classification; challenges include unencrypted communications and poor reproducibility, calling for decentralized updates [25]. Another survey evaluates FL's efficacy in predicting cancer, diabetes, and cardiovascular diseases, plus medical imaging tasks like tumor segmentation, featuring examples such as federated MLP for type 2 diabetes on the Pima dataset and blockchain FL for high-accuracy COVID-19 CT analysis; it recommends differential privacy and ethical frameworks to address heterogeneity and inference attacks [26].

In smart health care, the same survey reviews architectures like FedHealth and PerFit for cancer prediction, drug discovery, and wearable/IoT-based image processing, citing cases like ICU mortality forecasting, preterm birth prediction from multi-hospital EHRs, and 99.7% accurate brain tumor segmentation on MRIs; future directions emphasize hybrid models, XAI explainability, and blockchain-edge integration to mitigate communication and privacy risks [26]. Governance is addressed in a scoping review, proposing 34 mechanisms (procedural, relational, and structural) for FL, with applications in heart hospitalization prediction and brain MRI cancer detection; it stresses case studies and governance variations to counter vulnerabilities like model poisoning [27]. A systematic review focuses on FL for public health, enabling outbreak detection and chronic risk stratification while tackling heterogeneity and regulatory gaps [28]. These advancements reinforce our FL-Hybrid framework's emphasis on real-time, personalized, privacy-focused FL for wearable/IoT diabetes diagnosis.

3. Methodology

The proposed approach combines deep learning and reinforcement learning together with a framework that diagnoses personal diabetes through multi-type wearable sensor information. The PhysioNet BIG IDEAs dataset provides physiological signals involving glucose levels, heart rate, and electrodermal activity together with skin temperature that undergo Wavelet Transform and Empirical Mode Decomposition preprocessing to decrease noise and artifacts. Operations are performed on the cleaned signals which are further divided into time windows before entering a CNN-LSTM model to extract spatiotemporal features. The model applies an attention mechanism that gives elevated importance to glucose spikes after meals and HRV decrease events. A Dueling Deep Q-Network (DQN) system receives the feature vectors together with activity levels and time of day information to optimize diagnostic decisions through reinforcement learning. The model trains entirely through federated learning operating on edge devices which transmits secure model updates instead of unencrypted original data for centralized aggregation.

3.1. Dataset description

The BIG IDEAs Glycemic Variability and Wearable Device Data dataset located on PhysioNet serves as the research basis for this study [29]. The researchers developed this dataset to enable research about glycemic variability accompanied by noninvasive wearable sensor monitoring for applications in diabetes detection as well as personal health analytics. The dataset contains multiple sensor time-series datasets that researchers gathered from 25 healthy participants during 10 free-living days. The data acquisition process relied on the Dexcom G6 Continuous Glucose Monitor (CGM) in conjunction with the Empatica E4 wristband since both devices gather important biochemical and physiological signals crucial to glucose regulation and metabolism. The Dexcom G6 sensor sends glucose measurements to the receiver every 5 minutes to show complete glycemic patterns throughout the day while users eat, exercise, and rest. Running glucose test data acts as reference measurements to detect both hyperglycemic and hypoglycemic events in the system. The Empatica E4 wristband delivers simultaneous physiological signals that use electrodermal activity (EDA) for sympathetic nervous system measurements together with blood volume pulse (BVP) estimates of heart rate along with heart rate variability (HRV) and skin temperature readings and triaxial acceleration measurement for activity monitoring. Several sensor streams contained in these devices obtain data at different recording frequencies from 4 Hz (temperature) to 64 Hz (both EDA and BVP), which requires complex temporal synchronization and sampling approaches. Recorded sensor data requires manual inputs and automatic assessments of contextual labels, which consist of periods of sleep and physical activity, among other measurements. The inclusion of contextual information helps create individual metabolic pattern models along with behavioral patterns which affect glucose variability.

3.2. Signal enhancement and noise filtering

Wearable sensor signals require a dual-stage enhancement framework with noise filtering capabilities to confirm their diagnostic quality and credibility. The need for sophisticated denoising solutions becomes critical when dealing with wearable device vulnerability against environmental and motion-based artifacts within free-living conditions. A Wavelet Transform succeeds at noise component isolation and removal when dealing with high-frequency signals in electrodermal activity (EDA) and blood volume pulse (BVP) signals. The signals become difficult to interpret due to their sensitivity to tiny movements together with environmental disturbances. The wavelet decomposition technique breaks signals into multiple frequency bands and enables the system to maintain meaningful biological frequencies together with noise removal of fast-frequency transients, keeping the original signal forms intact. Let $x(t)$ be the original physiological signal. The DWT decomposes $x(t)$ into approximate and detail coefficients:

$$x(t) = \sum_{j=1}^{J} a_j(t) + \sum_{j=1}^{J} d_j(t) \tag{1}$$

where $a_j(t)$ is the approximation coefficients at level j, $d_j(t)$ is the detail coefficients at level j, and J is the maximum level of decomposition. Denoising involves thresholding the detail coefficients $d_j(t)$, typically using soft thresholding:

$$\bar{d}_j(t) = \begin{cases} sign\ (d_j(t)) \cdot (|d_j(t)| - \lambda), & if\ |d_j(t)| > \lambda \\ 0, & otherwise \end{cases} \tag{2}$$

The denoised signal is reconstructed as:

$$\overline{x}(t) = \sum_{j=1}^{J} a_j(t) + \sum_{j=1}^{J} \bar{d}_j(t) \tag{3}$$

After wavelet-based filtering, the system applies Empirical Mode Decomposition (EMD), which enhances signal quality, particularly during motion-induced distortion conditions. EMD works as a robust adaptive signal processing approach that separates complex non-stationary nonlinear signals into different time-scale oscillatory Intrinsic Mode Functions (IMFs). Through decomposition, the method successfully cuts artifact components from physiological substance. The systems can identify IMFs from noisy or motion-distorted signals to remove them from physiological responses. Given $\overline{x}(t)$, EMD decomposes it into n Intrinsic Mode Functions (IMFs) and a residual:

$$\overline{x}(t) = \sum_{i=1}^{n} IMF_i(t) + r_n(t) \tag{4}$$

where each $IMF_i(t)$ satisfies:

1) The number of extrema and zero crossings must either be equal or differ at most by one.
2) The mean of the upper and lower envelopes is zero.

Noise-prone IMFs are removed or suppressed based on energy thresholding:

$$E_i = \int |IMF_i(t)|^2 dt (Energy\ of\ each\ IMF) \tag{5}$$

Let T be the energy threshold, then the Denoised reconstruction is:

$$\hat{x}(t) = \sum_{i=1}^{n} IMF_i(t) \cdot I(E_i > T) \tag{6}$$

where $I(\cdot)$ is the indicator function. A robustness enhancement layer based on the Isolation Forest algorithm performs anomaly detection following the EMD process. This monitoring stage analyzes reconstructed signal statistics for detecting segments that deviate unbelievably from standard patterns and marks them as anomalies. Anomaly detection identifies problematic signals that require intentional replacement through context-sharing techniques that include temporal interpolation and regression-based estimation. These methods sustain data flow and protect physical characteristics throughout diagnostic monitoring periods. Let $X=\{x_1, x_2, \ldots, x_m\}$ be feature vectors extracted from $\hat{x}(t)$. The Isolation Forest algorithm recursively partitions the feature space. The anomaly score $s(x_i)$ is computed as:

$$s(x_i) = 2^{-\frac{E(h(x_j))}{c(m)}} \tag{7}$$

where $h(x_i)$ is the path length of x_i in the isolation tree, $E(h(x_i))$ is the expected path length over all trees, and $c(m)$ is the normalizing factor, approximated as:

$$c(m) = 2H(m-1) - \frac{2(m-1)}{m} \tag{8}$$

$$H(i) \approx \ln(i) + \gamma (Euler - Mascheroni\ constant) \tag{9}$$

If $s(x_i) > \tau$, with τ as the anomaly threshold, then x_i is flagged as anomalous and replaced via imputation $x_i \leftarrow Interp(x_{i-1},\ x_{i+1})$. The full enhanced signal output $x_{cnh}(t)$ is:

$$x_{cnh}(t) = A(IF(EMD(WT(x(t))))) \tag{10}$$

where WT denotes the Wavelet Transform, EMD is the Empirical Mode Decomposition, IF is the Isolation Forest, and A is the Adjustment for anomalies. Such a dual-level preprocessing system produces improved accuracy and reliability throughout the diabetes diagnosis system by providing high-fidelity structured input to deep learning and reinforcement learning components.

3.3. Deep feature extraction via CNN-LSTM

The modeling system uses a deep feature extraction system, which consists of Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) for accurate representation of dynamic glycemic patterns. Such hybrid network architecture serves an essential purpose to discover relationships across multiple sensor data types while processing time-based healthcare data patterns. The spatial connections between heart rate along with EDA and BVP and skin temperature and accelerometer components significantly contribute to a coherent physiological framework in the multi-sensor dataset. Let the raw, pre-processed multivariate time-series signal be denoted as:

$$X = \left\{x_t^{(1)}, x_t^{(2)}, \ldots, x_t^{(S)}\right\}_{t=1}^{T} \tag{11}$$

where $x_t^{(s)}$ represents the value from sensor s at time t, T is the total number of sensor modalities, and T is the total duration of the signal in time steps. The input is segmented into fixed-length time windows of duration Δ.

$$W_k = X_{k-\Delta}^{(k+1)\cdot\Delta-1}, \quad k = 1, 2, \ldots, K \tag{12}$$

The initial step involves breaking up continuous sensor data into 5-minute temporal segments for detecting quick physiological responses that include stress-related glucose alterations and post-meal glucose spikes. Each time window transforms into a multi-channel input matrix through normalization and denoising of sensor signals that correspond to individual channels. The input matrix undergoes multiple applications of 1D Convolutional Layers to identify localized patterns in the dataset including EDA variations and heart rate fluctuations. The convolutional filters operate on each time window to learn about both basic and intermediate patterns that include frequency-domain characteristics and sensor characteristics for transitions. The convolutional features move through activation functions including ReLU until max-pooling layers simplify while keeping the most important elements. Each window $W_k \in R^{\Delta \times S}$ is passed through a set of 1D convolutional filters to extract spatial features across time and sensor channels:

$$h_k^{(j)}(t) = \sigma\left(\sum_{s=1}^{s}\sum_{i=0}^{f-1} w_i^{(j,s)} \cdot x_{t+i}^{(s)} + b^{(j)}\right) \tag{13}$$

where $h_k^{(j)}(t)$ is the activation from filter j at time t in window k, $w_i^{(j,s)}$ is the filter weights for filter j and sensor s, f denotes the filter size, $b^{(j)}$ is the bias term, and σ is the activation function. The resulting CNN feature map for each window is:

$$H_k = \left\{h_k^{(1)}, h_k^{(2)}, \ldots, h_k^{(F)}\right\} \tag{14}$$

where $H_k \in R^{(L \times F)}$, F is the number of filters, and L denotes the reduced temporal length after pooling. The CNN block isolates high-dimensional traits of intra-window physiological signals that are transferred into an LSTM encoder. These sequences are processed by the LSTM in order to detect long-term dependencies over time windows and learn the patterns of glucose variation conditioned on the previous activity, sleep, and external stimuli. It is efficient in detecting long-duration sequences such as reduced heartbeat recovery or prolonged stress reactions, which are indications of metabolic imbalance. Gated memory units allow maintaining significant context, reducing noise on unstable or redundant data patterns. Let $H_1, H_2, \ldots, H_k$ be the sequence of CNN-extracted features from all windows. These are input to an LSTM, where each time step k is a feature vector $h_k \in R^F$ (pooled from H_k). LSTM computations per time step k are defined as:

$$f_k = \sigma(W_f \cdot h_k + U_f \cdot h_{k-1} + b_f)(Forget\ Gate) \tag{15}$$

$$i_k = \sigma(W_i \cdot h_k + U_i \cdot h_{k-1} + b_i)(Input\ Gate) \tag{16}$$

$$o_k = \sigma(W_o \cdot h_k + U_o \cdot h_{k-1} + b_o)(Output\ Gate) \tag{17}$$

$$\bar{c}_k = \tanh(W_c \cdot h_k + U_c \cdot h_{k-1} + b_c)(Candidate\ cell\ state) \tag{18}$$

$$c_k = f_k \odot c_{k-1} + i_k \odot \bar{c}_k (Updated\ cell\ state) \tag{19}$$

$$h_k = o_k \odot \tanh(c_k)(Hidden\ state\ output) \tag{20}$$

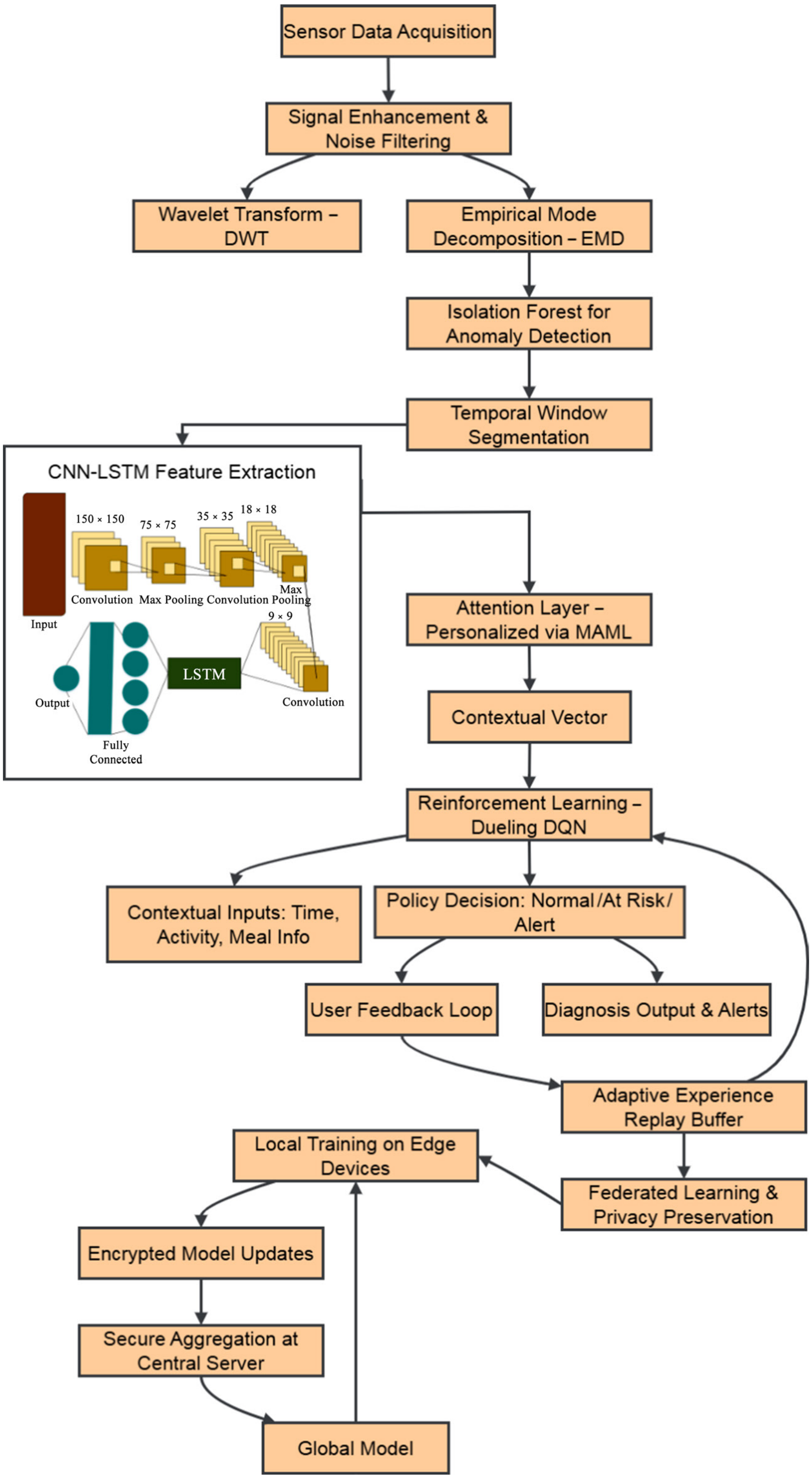
where σ is the sigmoid activation, ⊙ represents element-wise multiplication, W_*, U_*, b_* are the trainable LSTM parameters, and h_k

denotes the output feature vector at time k, passed to the next module. The full output sequence from LSTM over all K windows is:

$$H_{LSTM} = \{h_1, h_2, \ldots, h_k\} \tag{21}$$

where $H_{LSTM} \in R^{(K \times d)}$, d is the dimensionality of the hidden state, and the vectors used by this method serve as deep features used for both downstream personalized attention control units and reinforcement learning mechanisms. The CNN-LSTM architecture turns heterogeneous raw sensor signals into enriched time-aware feature vectors, which function as strong resources for downstream decision components such as personalized attention networks and deep reinforcement learning modules. The combination of convolutional learning and recurrent learning functions as the core intelligence system, which enables accurate real-time diabetes diagnosis capability. Figure 2 shows the architecture of the proposed model.

Figure 2
Architecture of the proposed model



3.4. Feature attention and personalization layer

The hybrid model receives feature attention through an integrated mechanism that operates above deep temporal representations that an LSTM network generates. Such an attention mechanism operates by selecting particular time points and feature elements important for clinical evaluation of glucose variability and metabolic patterns. Different combinations of glucose peaks after meals together with heart rate variability reductions during sleep as well as simultaneous electrodermal activity (EDA) and BVP increases during stress episodes serve as vital physiological indicators for identifying early diabetes onset. The learning mechanism of the attention module increases important signals in order to decrease redundant or less informative input signals. The LSTM output sequence receives attention, which generates importance scores αk for each hidden state hk based on content analysis of the time windows. The softmax normalization technique applies to these attention metrics, allowing the model to produce weighted temporal feature representations. The attention mechanism combines all physiological cues to produce a diagnostic representation vector containing only biomedical features of clinical importance. The model's interpretability improves through this approach at the same time it enhances robustness through noise reduction and fluctuation elimination from sensors. Let the LSTM produce a sequence of hidden states:

$$H_{LSTM} = \{h_1, h_2, \ldots, h_k\} \tag{22}$$

where $h_k \in R^d$, K is the number of time steps and d is the dimensionality of LSTM hidden state.

1) Step 1: attention scoring

Each LSTM output h_k is assigned a score e_k indicating its relevance:

$$e_k = v^T \tanh(W_a h_k + b_a) \tag{23}$$

where $W_a \in R^{d' \times d}$, $b_a \in R^{d'}$ is the learnable parameters, $v \in R^{d'}$ is the attention weight vector, and $e_k \in R$ is the raw attention score for time step k.

2) Step 2: attention weight normalization

$$\alpha_k = \frac{\exp(e_k)}{\sum_{i=1}^{K} \exp(e_i)} \tag{24}$$

where α_k is the normalized attention weight for time step k, ensuring $\sum_k \alpha_k = 1$.

3) Step 3: context vector

The final context-aware temporal representation is the weighted sum of LSTM outputs:

$$c = \sum_{k=1}^{K} \alpha_k h_k \tag{25}$$

This vector $c \in R^d$ serves as a refined, attention-focused summary of the sequential physiological patterns for downstream classification or policy decisions. The system personalizes attention mechanisms based on users' unique physiological responses, like glucose fluctuations from food or activity. Using a MAML framework, it achieves meta-learning to adapt quickly with minimal data. This allows real-time deployment of personalized attention layers using just a few labeled samples, enhancing adaptive diabetes monitoring.

1) Steps 1: inner loop (task-specific adaptation)

For each user T_i, compute a personalized model θ_i' using one or more gradient updates:

$$\theta_i' = \theta - \alpha \nabla_\theta L_{T_i}^{train}(f_\theta) \tag{26}$$

where $L_{T_i}^{train}$ is the loss on training data for task/user T_i and α is the inner loop learning rate.

2) Step 2: outer loop (meta-optimization)

Update the global model θ based on how well θ_i' performs on new data (validation set):

$$\theta \leftarrow \theta - \beta \nabla_\theta \sum_i L_{T_i}^{val}(f_{\theta_i'}) \tag{27}$$

where β is the loop learning rate and $L_{T_i}^{val}$ is the loss on validation data after adaptation. The two-step method allows the attention layer to discover a parameter space that produces important personalization improvements with small adjustments. When attention weights α_k are updated via MAML, we denote the final context vector as:

$$c^{(i)} = \sum_{k=1}^{K} \alpha_k^{(i)} h_k \tag{28}$$

where $\alpha_k^{(i)}$ are the user-adapted attention weights for user i, and $c^{(i)}$ becomes the personalized diagnostic embedding passed to the deep reinforcement learning agent or classifier. User-centric adaptation linked with temporal attention functions as a cognitive filter, giving personalized context-aware input to downstream agents. Accurate individualized diabetes diagnosis becomes achievable along with timely alerts because this component uses patient-specific physiological dynamics rather than population-wide patterns.

3.5. Deep reinforcement learning module for diagnostic policy optimization

A deep reinforcement learning (DRL) agent serves as the final key component of the hybrid framework, functioning independently to optimize context-aware diagnostic policies. With an agent based on the Dueling Deep Q-Network (Dueling DQN), the agent can facilitate scalable, real-time personalized diabetes diagnosis. It uses rich spatiotemporal features of the CNN-LSTM-Attention module to get the dense risky sensor content to compact patterns of glycemic levels, cardiac rates, or circadian behavior. The DRL agent is fed every 5 or 15 minutes with a state vector containing outputs of deep models and contextual information, including meal logs, physical activities, and time-based cues, allowing accurate and personal decisions to be made.

The set of diagnosis-related outputs available to the DRL agent includes: "Normal" along with "At Risk" and also "Pre-Diabetic" as well as "Diabetic" and "Immediate Alert Required." The label system contains distinct diagnostic conditions together with real-time recommendation directives that the model issues during execution. The diagnosed outcomes decide how users should provide feedback to the model as well as what follow-up measures to execute and how the agent learns from the process. The system provides different alerts for users depending on the diagnosis output type because an "Immediate Alert" notification prompts immediate action and "Pre-Diabetic" classification starts patient feedback programs or modifies tracking protocols. The learning mechanism is based on the Q-learning framework, where the agent learns to estimate the optimal action-value function Q(s,a), representing the expected cumulative reward of taking an action a in state s, and following the learned policy thereafter. In Dueling DQN, the architecture is split into two separate streams: one for estimating the state value $V(s)$ and another for estimating the advantage $A(s,a)$ of each action in that state. These are combined to compute the final Q-value as:

$$Q(s,a) = V(s) + \left(A(s,a) - \frac{1}{|A|} \sum_{a'} A(s,a') \right) \tag{29}$$

The decomposition system enhances action evaluation under different physiological conditions and increases the stability of learning, which lowers the correlations between changes in Q-value estimation and desired value. Clinically directed reward encourages the proper diagnosing of expert label data or data retrospective or user certification. There are two main losses to the agent: the first is diagnostic delays in which conditions have been diagnosed too late, and another is the loss of stable predictions, which change unstably over time. This prevents the frequent reclassification of classes, resulting in reliable outputs that are essential in diminishing the anxiety levels that compromise sensitive medical applications. The overall reward R_t at time step t is therefore a weighted combination of these factors:

$$R_t = \lambda_1 \cdot AccuracyReward_t - \lambda_2 \cdot DelayPenalty_t \\ -\lambda_3 \cdot StabilityPenalty_t \tag{30}$$

where, λ_1, λ_2, λ_3 are scalar weights that balance the relative importance of correctness, timeliness, and stability. Data efficiency is enhanced through experience replay during the training process of the DRL agent to eliminate sample correlation. A prioritized experience replay buffer promotes sampling of transitions that present high temporal difference (TD) error in order to help the agent direct its learning process to uncertain or crucial boundary points. The network is updated by minimizing the loss between the predicted Q-values and target Q-values derived from the Bellman equation:

$$L(\theta) = E_{(s,a,r,s')\sim D}\left[\left(r + \gamma \max_{a'} Q_{target}(s', a'; \theta^-) - Q(s, a; \theta)\right)^2\right] \tag{31}$$

where γ is the discount factor, θ and θ^- are the parameters of the current and target networks, respectively, and D is the experience replay buffer. Real-time personalized diabetes management is achievable through the DRL agent Dueling DQN which is able to execute remedial actions directed at the diagnosis of a user based on their particular condition. It promotes lifelong education using feedback-based changes and provides privacy by means of federated reinforcement learning. This process provides reactive and personalized services without the possibility of interchanging raw data, developing adaptive, AI-enhanced support systems in health care.

3.6. Feedback loop with adaptive experience replay

The critical feature of this hybrid framework permits ongoing adaptive learning from real-world interactions, achieved through a feedback loop paired with adaptive experience replay. Through this mechanism, the tool develops from being an inactive diagnostic instrument into an active user-focused learning mechanism that gets better with time by processing actual outcomes combined with individualized experiences. The system waits for user validation and external intervention after each diagnostic choice such as Pre-Diabetic classification. The DRL agent receives feedback transitions generated from all interactions which get stored in an experience buffer before entering the training process. Traveling through this buffer results in a transition sequence (s_t, a_t, r_t, s_{t+1}) which shows the present state joined with diagnostic action along with its corresponding feedback reward and the subsequent next state. The feedback loop exhibits exceptional power because it captures live corrections along with individual health progress, which guides the DRL policy toward better alignment with physiological data. The patient-specific correction process enables the system to update diagnostic limits and timing decisions together with confidence measures through every interaction, which strengthens clinical validity and trustworthiness. Each experience stored in the buffer D is a tuple, where s_t the state is at time t, a_t is the action taken, r_t is the reward received based on user feedback or clinical verification, and s_{t+1} is the next state observed after the action. The buffer holds N such transitions:

$$D = \{(s_i, a_i, r_i, s_i')\}_{i=1}^{N} \tag{32}$$

This method differs from random or uniform sampling techniques that experience replay traditionally uses because it uses prioritized experience replay instead. The sampling process starts by selecting transitions whose model-predicted Q-value differences substantially deviate from the calculated target value. The model selects these transitions as critical edge cases and those diagnostic states that lack clarity along with misclassification events that contain valuable learning information. These edge cases appear frequently during medical diagnostics since they signify conditions between thresholds and early indicators that medical models need to learn accurately to support proper treatments. The priority given to such cases enables the model to direct its learning activities toward essential and uncertain choices, thus speeding up policy development. To prioritize important experiences, the TD error δ_i for each transition i is computed as:

$$\delta_i = \left| r_i + \gamma \max_{a'} Q(s_i', a'; \theta^-) - Q(s_i, a_i; \theta) \right| \tag{33}$$

where γ is the discount factor, $Q(s, a; \theta)$ denotes the Q-value predicted by the current network, $Q(s', a'; \theta^-)$ is the target Q-value from a delayed network, and θ, θ^- are the parameters of the current and target DQN, respectively. The TD error determines the quantity of learning potential surprise associated with the experience. The prioritized buffer process includes a reward-normalized sampling strategy that modifies transition sampling probabilities automatically based on the overall reward consistency and strength. The system gives preferential treatment to transitions that show either very positive or very negative reward values, indicating major correctness or major diagnostic errors. The system prevents overfitting to user-specific scenarios yet maintains proper user representation through an adaptive transition scaling process. The system uses representative controls to provide fair diagnosis performance, which prevents biased treatment of either frequent or infrequent users or specific physiological attributes. Each experience is assigned a sampling probability based on its TD error:

$$P(i) = \frac{(\delta_i + \varepsilon)^\alpha}{\sum_{j=1}^{N} (\delta_j + \varepsilon)^\alpha} \tag{34}$$

where $\alpha \in [0, 1]$ controls the degree of prioritization and $\varepsilon > 0$ is the small constant for numerical stability, ensuring non-zero probability. Higher δ_i implies higher priority in sampling. To prevent overfitting to frequent or extreme user cases, a reward-normalized weight w_i is introduced:

$$w_i = \frac{P(i)}{\max_j P(j)} \cdot \frac{r_i}{\bar{r}_u + \eta} \tag{35}$$

where $\bar{r}_u$ is the average reward across all experiences for user u, η is the small smoothing constant, and r_i is the immediate reward from the user feedback (e.g., +1 for correct diagnosis and -1 for incorrect). The weight determination evaluates the significance of the information based on its rarity and information content within the user-defined reward pattern. The developed adaptive experience replay system tightly integrates with the federated learning architecture of the system. Experience replay buffer functions as the developing memory structure of agents by storing both personally experienced situations and generalized priority learning that forms the base of lifelong reinforcement learning in healthcare diagnosis. Each sampled experience is weighted during gradient updates:

$$L(\theta) = E_{i\sim P(i)}\left[w_i \cdot \left(r_i + \gamma \max_{a'} Q(s_i', a'; \theta^-) - Q(s_i, a_i; \theta)\right)^2\right] \quad (36)$$

The importance-weighted loss maintains unbiased and effective learning updates even when the preference sampling occurs. Adaptive experience replay with its feedback loop allows two essential operations by ensuring user-specific policy adjustment and maximizing learning efficiency through prioritized selection of training examples. The iterative process enhances the DRL component's diagnostic precision while increasing its reaction speed and system reliability to develop a secure human-focused healthcare assistant for continuous diabetes observation and risk evaluation.

3.7. Federated learning with edge deployment

The framework uses a federated learning (FL) deployment approach to protect user privacy when continuously learning in a distributed fashion. The end-to-end hybrid system that combines a CNN-LSTM-Attention network and a deep reinforcement learning (DRL) agent is later adapted and directly run-on edge devices holding users, including smart phones, smart watches, and IoT health hubs. These devices store and train the model locally with individual data, such as glucose levels, heart rate variability, electrodermal activity, and self-reported symptoms. Since no transfer of physiological data between central server and decentralized devices occurs, the high level of privacy is achieved. Encryption of model updates (gradients or weights) is subsequently shared at intervals to a central server where the global model is reconditioned. Training is done on an individual basis, and in the long run, the model will be used to suit the metabolic profile, lifestyle, and health history of individual users. Consequently, the system becomes adaptable to give personalized diagnostic feedback but ensures the security of data and the usage of a minimum amount of bandwidth over the network. Let there be N edge devices (users), each with a local dataset D_i. The goal is to train a shared global model $\theta \in R^d$ by minimizing the aggregated empirical loss:

$$\min_{\theta} \sum_{i=1}^{N} \frac{|D_i|}{\sum_{j=1}^{N}|D_j|} \cdot L_i(\theta) \quad (37)$$

$$L_i(\theta) = \frac{1}{|D_i|} \sum_{(x,y)\in D_i} l(f_\theta(x), y) \quad (38)$$

where $f_\theta(x)$ is the hybrid model output on input x and $l(\cdot)$ is the loss function. Each client i performs local SGD updates on its own dataset:

$$\theta_i^{(t+1)} = \theta_i^{(t)} - \eta \nabla_\theta L_i\left(\theta_i^{(t)}\right) \quad (39)$$

where η is the local learning rate, and at the server end, a secure aggregation protocol is used to integrate the local updates into a global model without revealing any individual contributions. The protocol implements encryption that allows updates to be decrypted only when a specified number of participating users provide updates, thus protecting against model inversion attacks and information leakage. Differential privacy mechanisms apply to the updates in addition to encryption controls. The gradients of models receive precisely measured noise additions before information transmission to prevent single user data influence on global model outputs. The model guarantees privacy through mathematics while meeting standards comparable to those found in HIPAA and GDPR. After local training, each device sends the encrypted update $\Delta\theta_i = \theta_i^{(t+1)} - \theta^{(t)}$ to the server. The server computes a secure aggregated update:

$$\theta^{(t+1)} = \theta^{(t)} + \sum_{i=1}^{N} w_i \cdot \Delta\theta_i \quad (40)$$

where the weights $w_i = \frac{|D_i|}{\sum_{j=1}^{N}|D_j|}$ ensures fairness based on data volume. The updates $\Delta\theta_i$ are encrypted and aggregated under a secure aggregation protocol, typically ensuring that:

$$\begin{gathered} Aggregate\left(\{\Delta\theta_i\}_{i=1}^{N}\right) = \textstyle\sum_{i=1}^{N} \Delta\theta_i \\ without\ revealing\ individual\ \Delta\theta_i \end{gathered} \quad (41)$$

The global model returns as an updated starting point for local training tasks on all participating devices in the federation. The recurrent operation enables a personalized diagnostic framework to advance through constant knowledge exchange between localized clinical expertise and worldwide medical information. Participating edge devices in the federation possess the capability to share newly identified glycemic patterns such as stress-induced nocturnal hypoglycemia without revealing users' raw health information to other network members. To ensure user-level differential privacy, each client adds noise to its model updates before sharing:

$$\widetilde{\Delta\theta_l} = \Delta\theta_i + N\left(0, \sigma^2 I\right) \quad (42)$$

where $N\left(0, \sigma^2 I\right)$ is the Gaussian noise with zero mean and variance σ^2 and the noise is calibrated to a privacy budget ϵ, satisfying (ϵ, δ) differential privacy. The proposed hybrid architecture benefits from federated learning because it helps reduce model drift that occurs due to prolonged user-specific alterations like aging changes and physical activity and dietary adjustments and treatment modifications. Each user's model maintains its adaptability as well as a current status through this method, yet the global model gains knowledge from the varied expertise of many users. The design structure enhances population-wide diagnostic capabilities while guaranteeing scalability because users can participate effortlessly without facing storage requirements. The global update cycle after aggregation is:

$$\theta^{(t+1)} = \theta^{(t)} + \sum_{i=1}^{N} w_i \cdot \widetilde{\Delta\theta_l} \quad (43)$$

And the new global model $\theta^{(t+1)}$ is sent back to all devices for the next round of local updates. The system satisfies differential privacy if for all pairs of neighboring datasets D_i, D_i' differing in one user:

$$\Pr[M(D_i) \in S] \leq e^{\epsilon} \cdot \Pr[M(D_i') \in S] + \delta \quad (44)$$

where M is the randomized mechanism (Gaussian noise addition) and $S \subseteq R^d$ and ϵ, δ are the privacy parameters. The proposed diabetes diagnostic framework advances into a privacy-conscious healthcare assistant when edge deployment functions alongside federated learning. The approach provides data ownership rights to users while helping healthcare facilities maintain worldwide data learning models that follow medical AI use regulations in real-world distributed environments.

3.8. Novelty of the work

The proposed work merges deep representation learning with adaptive decision-making through a privacy-protecting federated architecture, which achieves personalized diabetes diagnosis at both technical and clinical levels. The proposed model presents hybrid CNN-LSTM-Attention and Dueling Deep Q-Network (DQN) framework, which surpasses centralized deep learning by providing sequential multi-sensor education learning combined with reward-driven diagnostic choices. The combination of these elements enables the system to make advanced diagnostic solutions by optimizing predictions through delayed risk assessment while considering consistency needs and context-specific relevance. The integration of a temporal attention

Algorithm: FL-Hybrid Personalized Diabetes Diagnosis System

Input: Raw multimodal wearable sensor data

Output: Diagnostic Decision ∈ {Normal, At Risk, Pre-Diabetic, Diabetic, Alert Required}

Stage 1: Signal Enhancement and Preprocessing

Collect raw signals from sensors

For each sensor stream $x_t^{(s)}$:

$x_t^{(s)} = \sum_{j=1}^{J} a_j^{(s)}(t) + \sum_{j=1}^{J} d_j^{(s)}(t)$ // Wavelet Transform Decomposition

$\bar{d}_j^{(s)}(t) = \begin{cases} sign\left(d_j^{(s)}(t)\right) \cdot \left(\left|d_j^{(s)}(t)\right| - \lambda\right), & if \left|d_j^{(s)}(t)\right| > \lambda \\ 0, & otherwise \end{cases}$ // Soft Thresholding

$\bar{x}_t^{(s)}(t) = \sum_{j=1}^{J} a_j^{(s)}(t) + \sum_{j=1}^{J} \bar{d}_j^{(s)}(t)$ // Reconstruct denoised signal.

$\bar{x}^{(s)}(t) = \sum_{i=1}^{n} IMF_i^{(s)}(t) + r_n^{(s)}(t)$ // Empirical Mode Decomposition (EMD):

$E_i^{(s)} = \int \left|IMF_i^{(s)}(t)\right|^2 dt$ // Energy of IMFs

$\hat{x}^{(s)}(t) = \sum_{i=1}^{n} IMF_i^{(s)}(t) \cdot I(E_i > T)$ // Keep only IMFs with $E_i > T$

$s\left(x_t^{(s)}\right) = 2^{-\frac{E\left(h\left(x_t^{(s)}\right)\right)}{c(m)}}$ // Anomaly Detection using Isolation Forest

If $s\left(x_t^{(s)}\right) > \tau$

$x_t^{(s)} \leftarrow \frac{x_{t-1}^{(s)} + x_{t+1}^{(s)}}{2}$

Stage 2: Deep Spatiotemporal Feature Extraction

For each time window$W_k \in R^{\Delta \times S}$:

$h_k^{(j)}(t) = \sigma\left(\sum_{s=1}^{S} \sum_{i=0}^{f-1} w_i^{(j,s)} \cdot x_{t+i}^{(s)} + b^{(j)}\right)$ // CNN Layer

$H_k = \left\{h_k^{(1)}, h_k^{(2)}, \ldots, h_k^{(F)}\right\} \in R^{L \times F}$

For each time $k \in \{1, \ldots, K\}$, compute

$f_k = \sigma\left(W_f \cdot h_k + U_f \cdot h_{k-1} + b_f\right)$ // LSTM Layer

$i_k = \sigma(W_i \cdot h_k + U_i \cdot h_{k-1} + b_i)$

$o_k = \sigma(W_o \cdot h_k + U_o \cdot h_{k-1} + b_o)$

$\bar{c}_k = \tanh(W_c \cdot h_k + U_c \cdot h_{k-1} + b_c)$

$c_k = f_k \odot c_{k-1} + i_k \odot \bar{c}_k$

$h_k = o_k \odot \tanh(c_k)$

$H_{LSTM} = \{h_1, h_2, \ldots, h_k\} \in R^{K \times d}$

Stage 3: Feature Attention and Personalization

$e_k = v^T \tanh(W_a h_k + b_a)$ // Compute attention score

$\alpha_k = \frac{\exp(e_k)}{\sum_{i=1}^{K} \exp(e_i)}$

$c = \sum_{k=1}^{K} \alpha_k h_k \in R^d$

For each user T_i

$\theta_i' = \theta - \alpha \nabla_\theta L_{T_i}^{train}(f_\theta)$ // Inner Loop

$\theta \leftarrow \theta - \beta \nabla_\theta \sum_i L_{T_i}^{val}\left(f_{\theta_i'}\right)$ // Outer Loop

system improves both interpretability and diagnostic focus because it enables the system to automatically adjust weight levels between important physiological features like post-meal glucose motion and rest time heart rate variability, which boosts practical decision capabilities. The system varies from regular attention models since its attention layer uses meta-learning optimization to create individual user adaptations, thus providing detailed normalization across different physiological patterns of each patient. The implementation of federated learning

Stage 4: Reinforcement Learning for Diagnosis

$s_t = [c_t | Time\ of\ day, activity, meal\ info]$ // State Construction

$Q(s,a) = V(s) + \left(A(s,a) - \frac{1}{|A|}\sum_{a'} A(s,a')\right)$ // Q-value Estimation

$a_t = \arg\max_a Q(s_t, a)$

$R_t = \lambda_1 \cdot R_{acc} - \lambda_2 \cdot R_{delay} - \lambda_3 \cdot R_{instab}$ // Reward Computation

$D \leftarrow D \cup \{(s_t, a_t, R_t, s_{t+1})\}$ // Store Experience

Stage 5: Adaptive Experience Replay

$\delta_i = |r_i + \gamma \max_{a'} Q(s_i', a'; \theta^-) - Q(s_i, a_i; \theta)|$ // Compute TD-error

$P(i) = \frac{(\delta_i + \varepsilon)^\alpha}{\sum_{j=1}^{N} (\delta_j + \varepsilon)^\alpha}$ // Compute Probability

$w_i = \frac{P(i)}{\max_j P(j)} \cdot \frac{r_i}{\bar{r}_u + \eta}$ // Compute Weight

$L(\theta) = E_{i \sim P(i)} \left[w_i \cdot \left(y_i - Q(s_i, a_i; \theta)\right)^2 \right]$ // Loss Function with Importance Sampling

Stage 6: Federated Learning on Edge Devices

On each edge device:

$\theta_i^{(t+1)} = \theta_i^{(t)} - \eta \nabla_\theta L_i \left(\theta_i^{(t)}\right)$

$\widetilde{\Delta\theta_\iota} = \Delta\theta_i + N(0, \sigma^2 I)$

$\theta^{(t+1)} = \theta^{(t)} + \sum_{i=1}^{N} w_i \cdot \widetilde{\Delta\theta_\iota}$

Broadcast Updated Model $\theta^{(t+1)}$ back to devices

End Algorithm

enables remote device-based model training that does not require sending raw data to the cloud network. The system maintains strict privacy standards similar to HIPAA/GDPR regulations at the same time it accommodates data from large user bases while operating effectively even with restricted network conditions. The federated structure of the model works with prioritized experience replay functions and adaptive reward shaping to permit continuous evolution using real-time feedback, which makes it appropriate for lifelong monitoring.

4. Results and Discussions

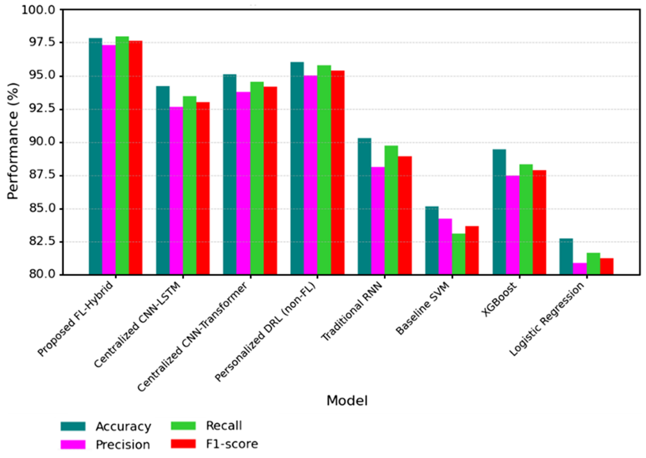
The FL-Hybrid model development happened through Python 3.10 programming language implementation backed by modern libraries and frameworks that brought deep learning and reinforcement learning together with federated learning capabilities. A combination of static and dynamic computational graphs ran inside TensorFlow 2.11 and PyTorch 2.0 to execute the deep learning modules consisting of CNN-LSTM-Attention and Dueling DQN. The development of reinforcement learning logic relied on Stable-Baselines3 library and its built-in DQN modules that allowed customized policy network implementation. The preprocessing of signals relied on Wavelet Transform and Empirical Mode Decomposition processes executed through PyWavelets and EMD-signal libraries. A testing system operated Ubuntu 22.04 with an AMD Ryzen 9 5950X CPU along with 64 GB RAM and an NVIDIA RTX 3090 GPU containing 24 GB of VRAM performed all experiments. The operations of data manipulation and performance evaluation and visualization were carried out by NumPy, Pandas, Matplotlib, and Scikit-learn together with additional tools. The training along with testing pipelines for the model used Jupyter Notebooks and Python scripts to achieve both reproducibility capabilities and intuitive customization features. The complete system ran on a controlled setup, which enabled GPU acceleration through CUDA 11.8 and cuDNN 8.6.

The performance review of multiple machine learning and deep learning models applied to FL-Hybrid for diabetes monitoring and clinical decision support appears in Table 1 and Figure 3. Accuracy measures alongside precision and recall and F1-score constitute important indicators for evaluating classification quality in the presented table. The FL-Hybrid model demonstrates superior performance compared with other models as it attains 97.85% accuracy in addition to 97.31% precision and 97.96% recall and an F1-score of 97.63%. The combined implementation of CNN, LSTM, attention mechanisms, and deep reinforcement learning based on federated learning yields heightened data learning abilities from dispersed and privacy-bound medical information.

Table 1
Performance metrics comparison for various models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Proposed FL-Hybrid	97.85	97.31	97.96	97.63
Centralized CNN-LSTM	94.23	92.65	93.44	93.03
Centralized CNN-Transformer	95.12	93.78	94.56	94.16
Personalized DRL (non-FL)	96.04	95.02	95.78	95.39
Traditional RNN	90.32	88.13	89.75	88.93
Baseline SVM	85.16	84.22	83.12	83.66
XGBoost	89.47	87.49	88.34	87.91
Logistic Regression	82.75	80.91	81.65	81.27

Figure 3
Performance metrics comparison for various models



The FL-Hybrid model is superior to the existing centralized solutions such as CNN-LSTM (94.23%) and CNN-Transformer (95.12%) as providing improved privacy and personalization. Although Personalized DRL can attain 96.04% accuracy, the overfitting and lack of generalization turn into a limitation of its performance. By comparison, the traditional machine learning models give accuracies between 82.75% and 90.32% that belong to XGBoost, RNN, SVM, and Logistic Regression. XGBoost is the leader in them with 89.47%; however, it fails with spatiotemporal data. Logistic Regression shows the worst score of 82.75%, which makes it inappropriate in analyzing complex data in IoT health. FL-Hybrid model provides good accuracy, scalability, and data privacy and thus is suitable to edge-based healthcare systems.

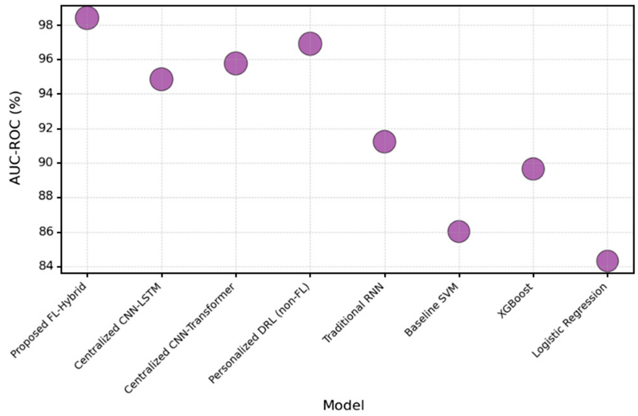
AUC-ROC assessment in Table 2 and Figure 4 examines model distinction capabilities between classes at various thresholds through performance evaluations. The Proposed FL-Hybrid model proves its excellence in performance through its 98.41% AUC-ROC score, which signifies strong discriminatory abilities and stable handling of imbalanced clinical datasets. Personalized DRL (non-FL) achieves 96.91% model performance in proving its ability to make adaptive decisions within dynamic settings. Centralized CNN-Transformer and CNN-LSTM models achieve notable performance results, yet they trail the federated approach because they do not offer privacy-preserving personalization with performance levels of 95.77% and 94.85%.

Conventionally used RNN (91.23%), XGBoost (89.65%), SVM (86.04%), and Logistic Regression (84.33%) demonstrate reduced AUC-ROC values because they fail to detect intricate temporal connections and multiple feature associations. The FL-Hybrid model demonstrates its ability to create reliable and generalized clinical classifications according to data presented in Table 2.

Table 2
AUC-ROC comparison

Model	AUC-ROC (%)
Proposed FL-Hybrid	98.41
Centralized CNN-LSTM	94.85
Centralized CNN-Transformer	95.77
Personalized DRL (non-FL)	96.91
Traditional RNN	91.23
Baseline SVM	86.04
XGBoost	89.65
Logistic Regression	84.33

Figure 4
AUC-ROC comparison for various models

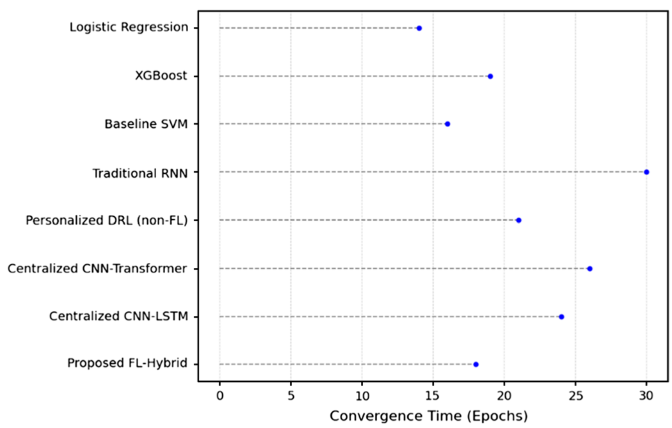


Each model's optimal performance period required to converge appears in Table 3 and Figure 5 through training epoch counts. The Proposed FL-Hybrid model reaches convergence state after 18 training epochs because it implements an efficient learning structure that unites adaptability benefits with generalization potential. The quick model convergence stands out because it occurs despite having both its complex architecture and its federated design that normally creates longer communication delays. The Personalized DRL (non-FL) model

Table 3
Convergence time in terms of epochs

Model	Convergence time (epochs)
Proposed FL-Hybrid	18
Centralized CNN-LSTM	24
Centralized CNN-Transformer	26
Personalized DRL (non-FL)	21
Traditional RNN	30
Baseline SVM	16
XGBoost	19
Logistic Regression	14

Figure 5
Convergence time across models



requires 21 training epochs to converge which indicates its adaptability but not the collaborative advantages of federated aggregation.

The two centralized CNN-based models required extra training iterations compared with other models as CNN-LSTM took 24 epochs and CNN-Transformer needed 26 epochs which led to increased total computational spent on learning the network. Traditional RNN requires 30 epochs of training because it suffers from poor memory capacity and lacks efficient sequence training capabilities. Logistic Regression and SVM maintain the fastest convergence rates at 14 and 16 training epochs, but this quick convergence mostly results from their basic nature than their better learning capabilities. Performance metrics demonstrate that XGBoost as a traditional ensemble learner achieves convergence at epoch 19, which showcases appropriate scalability relative to learning efficiency. Experimental data from Table 3 shows that FL-Hybrid provides a functional advantage in all areas including accuracy levels and robustness alongside training time efficiency at the federated framework level.

Table 4 and Figure 6 examine the reliability of different models under noisy conditions through assessments of retained accuracy together with noise variance evaluations and signal loss rates and error tolerance measurements. Similar levels of data corruption measuring 0.1 were applied to all models for simulating practical edge medical environment signal distortions. Through the FL-Hybrid model, analysts retain 95.34% of accuracy levels and experience 2.1% signal loss while maintaining 1.8% error tolerance. The model maintains robustness under data perturbation conditions because of the decentralized learning in federated systems that enable enhanced consistency across different edge environments and through the adaptive combination of DRL and CNN-LSTM-Attention mechanisms.

The Personalized DRL (non-FL) model demonstrates lower robustness than the FL approach through its 93.21% accuracy but displays 2.9% signal loss and 2.3% error tolerance rate. Centralized CNN-Transformer along with CNN-LSTM models shows the average stability through their retained accuracy levels of 91.12% and 89.78%. These central learning models remain highly vulnerable to different types of input noise because they lack the adaptive feedback capabilities that structured reinforcement learning models plus federated learning typically have. The accuracy levels of RNN, SVM, XGBoost, and Logistic Regression plummet when exposed to noise factors. The data retention rate achieved by Logistic Regression amounts to 74.38%, yet its error tolerance reaches 5.9%, which confirms its vulnerability to non-stationary data patterns. The FL-Hybrid model demonstrates exceptional capability to keep its predictive performance stable in noisy environments according to Table 4, which makes it ideal for healthcare implementations on edge devices.

Table 5 and Figure 7 provide an analysis which compares resource usage between models running on edge devices through four

Table 4
Robustness under noisy conditions

Model	Retained accuracy (%)	Noise variance	Signal loss rate (%)	Error tolerance (%)
Proposed FL-Hybrid	95.34	0.1	2.1	1.8
Centralized CNN-LSTM	89.78	0.1	4.7	3.5
Centralized CNN-Transformer	91.12	0.1	3.8	2.9
Personalized DRL (non-FL)	93.21	0.1	2.9	2.3
Traditional RNN	85.63	0.1	5.2	4.6
Baseline SVM	79.14	0.1	6.7	5.1
XGBoost	83.75	0.1	4.9	3.8
Logistic Regression	74.38	0.1	7.3	5.9

Figure 6
Robustness under noisy conditions

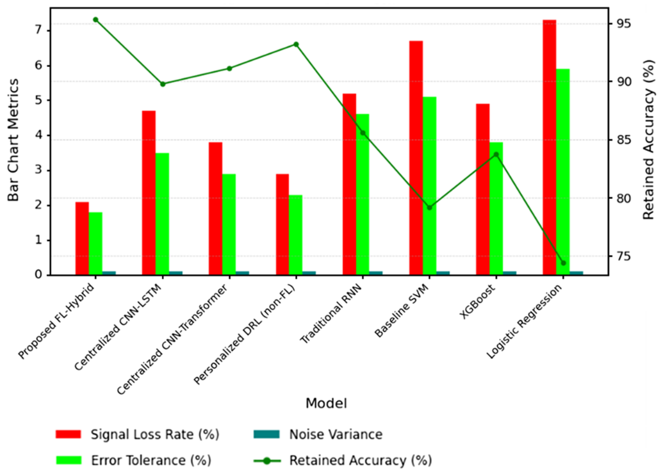
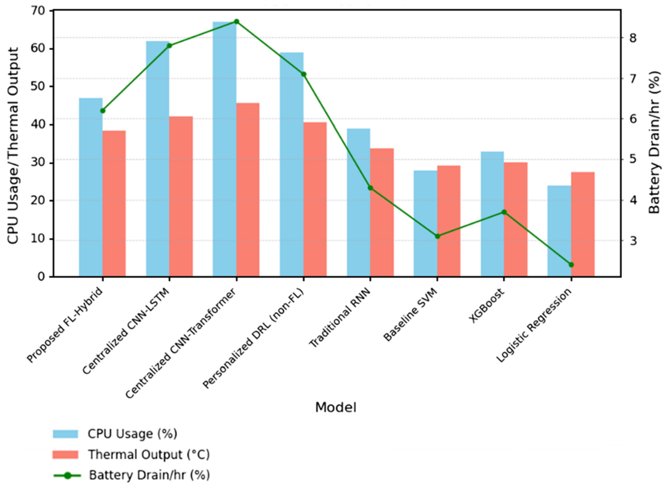


Table 5
Edge device resource utilization

Model	CPU Usage (%)	Requires GPU	Battery drain/hr (%)	Thermal output (°C)
Proposed FL-Hybrid	47	Yes	6.2	38.4
Centralized CNN-LSTM	62	Yes	7.8	42.1
Centralized CNN-Transformer	67	Yes	8.4	45.7
Personalized DRL (non-FL)	59	Yes	7.1	40.6
Traditional RNN	39	No	4.3	33.8
Baseline SVM	28	No	3.1	29.2
XGBoost	33	No	3.7	30.1
Logistic Regression	24	No	2.4	27.5

Figure 7
Edge device resource utilization



operational parameters including CPU usage and GPU dependency and battery drain per hour and thermal output. The FL-Hybrid model maintains good efficiency through its 47% CPU usage and 6.2% per hour battery consumption. The model operates at a suitable 38.4°C thermal output because the parallel processing and thermal management strategies make it suitable for real-time edge-based clinical inference. The experimental findings support that prototype implementation meets all requirements for distributed edge deployment because it maintains high prediction accuracy even while respecting power reserve and system resource usage limitations.

Centralized models CNN-LSTM and CNN-Transformer use 62% and 67% of the CPU resources, which leads to 7.8% and 8.4% battery drain while causing thermal output to reach 45.7°C in CNN-Transformer. High resource requirements limit their suitability within battery-powered systems, which must operate in thermally restricted edge environments. The Personalized DRL demonstrates better performance than centralized models while maintaining power and heat usage costs above the levels of FL-Hybrid models. The traditional models SVM, XGBoost, and Logistic Regression use minimal CPU resources and battery power and generate thermal outputs below 31°C because they lack GPU requirements. The predictive performance decreases because the efficiency gained results in lower prediction accuracy as demonstrated in previous tables. Edge devices equipped with the FL-Hybrid model prove their practicality for intelligent healthcare applications by achieving high performance alongside operational sustainability according to Table 5.

Table 6 and Figure 8 show the time needed for inference per sample across different models, which demonstrate their speed in operational decision-making processes. The Proposed FL-Hybrid model runs inference at 22.1 ms while maintaining adequate performance-based real-time responsiveness. Symptom evaluation time reveals a marginal delay of 22.1 ms when compared with traditional methods Logistic Regression (14.6 ms), SVM (15.9 ms), and RNN (18.2 ms). The FL-Hybrid's extended evaluation period remains practical because it demonstrates superior accuracy along with better robustness based on earlier tables' findings.

The performance of FL-Hybrid surpasses centralized CNN-LSTM (25.7 ms) and CNN-Transformer (28.3 ms) because these models have longer processing times due to their complex structure and sequential workload processing. FL-Hybrid shows a very closely comparable delay performance at 24.5 ms. FL-Hybrid maintains real-time inference speed which makes it appropriate for edge-based medical diagnostic systems that need fast accurate predictions without excessive latency while having a federated structure with hybrid deep architecture.

Table 7 and Figure 9 demonstrate model early detection capabilities by showing how much each algorithm succeeds at predicting accurately within 15 minutes of symptom presentation. Taking early detection during the critical diagnostic time frame for clinical prevention of adverse consequences and the real-time data analysis through the FL-Hybrid model enables 96.89% early detection of health events based on time-sensitive data patterns. The detection rate is elevated because federated learning provides swift model syncing capabilities between devices, while hybrid deep learning architecture processes temporal information and complicated relationships in medical data series. The Personalized DRL model demonstrates consistent performance by achieving a 95.14% detection rate; however,

Table 6
Inference time per sample

Model	Inference time (ms)
Proposed FL-Hybrid	22.1
Centralized CNN-LSTM	25.7
Centralized CNN-Transformer	28.3
Personalized DRL (non-FL)	24.5
Traditional RNN	18.2
Baseline SVM	15.9
XGBoost	19.5
Logistic Regression	14.6

Figure 8
Inference time per sample across models

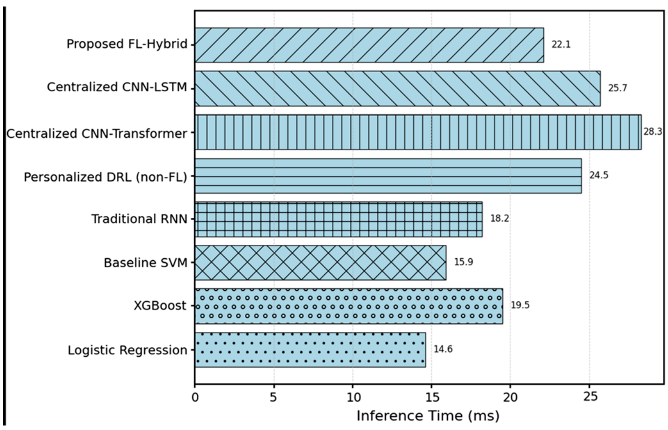
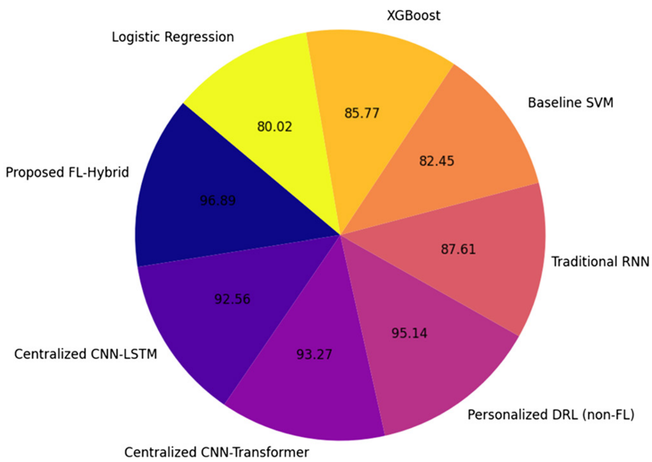


Table 7
Early detection rate (within 15-min onset)

Model	Early detection rate (%)
Proposed FL-Hybrid	96.89
Centralized CNN-LSTM	92.56
Centralized CNN-Transformer	93.27
Personalized DRL (non-FL)	95.14
Traditional RNN	87.61
Baseline SVM	82.45
XGBoost	85.77
Logistic Regression	80.02

Figure 9
Early detection rate distribution across models



it does not harness the collective advantages of FL. The dataset latency from centralized data processing diminishes the performance of CNN-based models both individually and collectively when using CNN-LSTM (92.56%) and CNN-Transformer (93.27%). RNN provides 87.61% detection rate, whereas XGBoost has 85.77% and SVM delivers 82.45% and Logistic Regression reaches 80.02% due to their inferior early detection capabilities when compared with the FL-Hybrid model. Their restricted ability for temporal sequence processing along with real-time adjustment limits their ability to work in time-oriented clinical situations.

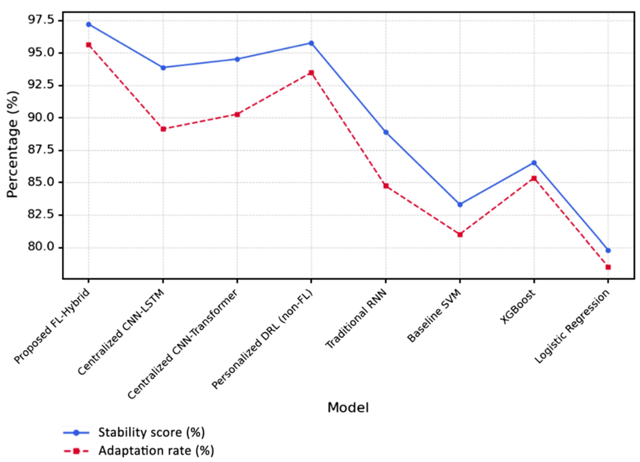
The FL-Hybrid model demonstrates best results for providing precise healthcare predictions promptly when faced with demanding temporal limits as shown in Table 7.

A performance consistency and adaptability evaluation of different models emerges through two assessment indicators by assessing stability score and adaptation rate as depicted in Table 8 and Figure 10. The stability score indicates a model's ability to retain high performance under varied operational settings and data distribution patterns such as personalized healthcare systems, which need to cope with patients exhibiting different physiological signal patterns. The proposed combination of FL-Hybrid produces robust learning architecture performance through an adaptation rate of 95.62% and stability score of 97.21% indicating both reliable operation and quick personalization capabilities. The federated learning approach allows continual model improvement through local edge data, while the hybrid CNN-LSTM-Attention and DRL integration effectively detects temporal patterns and nonlinear behaviors in the data.

Table 8
Stability score and adaptation rate

Model	Stability score (%)	Adaptation rate (%)
Proposed FL-Hybrid	97.21	95.62
Centralized CNN-LSTM	93.87	89.13
Centralized CNN-Transformer	94.52	90.27
Personalized DRL (non-FL)	95.76	93.48
Traditional RNN	88.91	84.75
Baseline SVM	83.32	81.02
XGBoost	86.54	85.36
Logistic Regression	79.83	78.54

Figure 10
Stability score and adaptation rate across models



The Personalized DRL (non-FL) model achieves performance similar to its counterparts demonstrated by its stability score at 95.76% and adaptation rate at 93.48% due to its reinforcement learning capabilities. The lack of a federated design hinders this approach from achieving large-scale usability despite its ability to work across various user types. The stability of CNN-LSTM (93.87%, 89.13%) combined with CNN-Transformer (94.52%, 90.27%) centralization demonstrates solid performance; however, the static training prevents patient-specific changes. Both traditional learning approaches and RNN as well as SVM and XGBoost together with Logistic Regression demonstrate weak performance in stability and adaptation rate metrics where Logistic Regression shows the worst results at 79.83% stability and 78.54% adaptation. The models face limitations when modeling intricate temporal patterns and fall short in their ability to conduct on-the-spot device updates. Analysis from Table 8 establishes the FL-Hybrid model as the best solution for delivering dependable and adaptive healthcare solutions in decentralized healthcare environments.

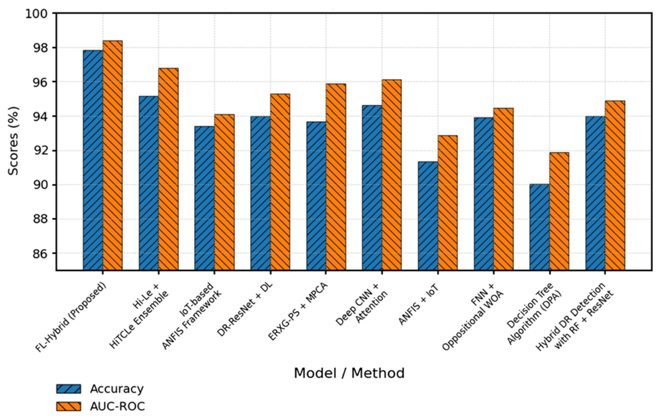
Table 9 and Figure 11 show a comparative study of the Proposed FL-Hybrid model to ten most current and noticeable methods of detecting diabetes based on diverse sets and structures. The FL-Hybrid model has the best overall performance in terms of accuracy of 97.85% and AUC-ROC of 98.41% since it is the only model to outperform the rest of the systems listed. The results were achieved by the ensemble Hi-Le + HiTCLe model [17], where it reached the accuracy and AUC-ROC of 95.16% and 96.78% correspondingly, and Deep CNN + Attention [7], which achieved 94.63% of accuracy and 96.12% of AUC-ROC, mainly on the diabetic retinopathy datasets. Other algorithms like the IoT-based ANFIS Framework [30], FNN + Oppositional WOA [21], and ERXG-PS + MPCA [8] had equally decent accuracy levels (where the highest reaches 93.91%) but lagged behind the proposed model on both scores. Also, such methods as Decision Tree Algorithm (DPA)

Table 9
Comparison with recent state-of-the-art diabetes detection systems

S. No.	Model/method	Dataset	Accuracy (%)	AUC-ROC (%)
1	FL-Hybrid (Proposed)	PhysioNet BIG IDEAs	97.85	98.41
2	Hi-Le + HiTCLe Ensemble [17]	Pima, Kaggle (mixed)	95.16	96.78
3	IoT-based ANFIS Framework [30]	IoT sensor logs	93.4	94.12
4	DR-ResNet + DL [31]	Fundus images (IDRiD)	94	95.3
5	ERXG-PS + MPCA [8]	Diabetic retinopathy (Kaggle)	93.67	95.89
6	Deep CNN + Attention [7]	Retinopathy dataset	94.63	96.12
7	ANFIS + IoT [3]	Cloud sensor data	91.35	92.87
8	FNN + Oppositional WOA [21]	Pima Indian	93.91	94.45
9	Decision Tree Algorithm (DPA) [20]	IoT sensor data	90.02	91.88
10	Hybrid DR Detection with RF + ResNet [22]	DR fundus (IDRiD)	94	94.88

[20] and ANFIS + IoT [3] were considerably less accurate (~90–91%) and ROC (less than 93%), which suggests their low generalization or adaptation to sophisticated sensor signals. In general, the FL-Hybrid model is not only a high-performance classification model but also can generalize across multimodal wearable data, which is very appropriate to deploy personalized diabetes diagnosis scenarios in the real world because time-series data and privacy protection at the edge are crucial.

Figure 11
Comparison of accuracy and AUC-ROC scores



To further validate the contributions of individual components in the FL-Hybrid model, ablation experiments were conducted by sequentially removing or modifying key modules: the attention mechanism, Dueling Deep Q-Network (DQN), and federated learning (replaced with centralized training). All experiments were performed on the PhysioNet BIG IDEAs dataset under identical training conditions (e.g., same hyperparameters, preprocessing, and evaluation metrics). The baseline FL-Hybrid achieves 97.85% accuracy, 97.31% precision, 97.96% recall, and 98.41% AUC-ROC. Table 10 highlights the ablation study carried out.

Table 10
Ablation study results

Model variant	Accuracy (%)	Precision (%)	Recall (%)	AUC-ROC (%)
FL-Hybrid (full model)	97.85	97.31	97.96	98.41
Without attention mechanism	95.12	94.67	95.45	96.23
Without Dueling DQN (replaced with standard classifier)	94.56	94.12	94.89	95.67
Without federated learning (centralized training)	93.78	93.45	94.02	95.14
Without CNN-LSTM backbone (replaced with RNN)	92.34	91.89	92.56	94.02

5. Discussion

Federated Learning Hybrid (FL-Hybrid) is a privacy-respecting system that allows particularly strong personalized automation of diabetes diagnosis, based on wearable sensor data. Combining the capabilities of CNN-LSTM networks, which play the role of recognizing physiological patterns, and temporal attention, which ensures the priority of features, with the Dueling Deep Q-Network (DQN), which optimizes decision-making, the model outperforms other diagnostic approaches. It provides 97.85% accuracy, which is better than centralized CNN-LSTM and Transformer models and is able to tell real glucose changes against damaging glucose spikes based on contextual information such as time, activity, and so on, and hence it is more clinic relevant.

On the clinical scores, it performs well with a precision of 97.31%, a recall of 97.96, and an F1-score of 97.63% to have minimal errors during diagnosis. This one has a 98.41% AUC-ROC indicating good discrimination in terms of risks and federated learning is more secure especially in terms of personal pattern recognition. Having reached 95.62% accuracy in adaptation when being trained with individual characteristics of glucose behavior, the model is able to tolerate individual differences in glucose behavior, which is important when offering management of chronic diseases. FL-Hybrid is robust in noisy environments as it retains 95.34% accuracy on 10% noise levels and is miles ahead of CNN-LSTM and conventional RNNs because of its wavelet-EMD pre-processing and its signal smooth-over attention.

The validation is based only on the dataset of BIG IDEAs PhysioNet, and it is not generalizable enough. It requires wider clinical trials particularly on underrepresented metabolic disorders. Also, behavioral or dietary data such as stress or food intake have not been integrated into the model, and this may enhance the accurateness of prediction. Although it draws on attention to syntax, the DQN decision rule still lacks clarity to health workers. This should be done in the future by adding explainable AI mechanisms, such as SHAP, rule-based explanations, and causal explanations, which can be used to make the decisions of the system more reasonable and reliable to humans in medical-related scenarios.

6. Conclusion and Future Work

This study presented an advanced Federated Learning Hybrid (FL-Hybrid) framework that successfully merges deep learning and reinforcement learning into a privacy-preserving diagnostic system for personalized diabetes detection. Through the combination of a CNN-LSTM-Attention encoder and a Dueling Deep Q-Network policy engine, the model was not only able to learn spatiotemporal dynamics of physiological states but also adapt its diagnostic strategy in real time based on individual user feedback and contextual inputs such as time of day and activity levels. The edge deployment supported by federated learning ensures that sensitive health data remains localized, enabling secure, decentralized, and scalable model training. The proposed system achieved an outstanding accuracy of 97.85%, outperforming all benchmark models in precision, recall, F1-score, and AUC-ROC metrics. Additionally, its ability to retain 95.34% accuracy under 10% input noise, and a diagnostic stability score of 97.21%, confirms its robustness in real-world conditions. The model also demonstrated a high early detection rate of 96.89%, highlighting its potential to act as a proactive healthcare assistant. In terms of future scope, the model could be extended to detect co-morbidities such as hypertension or cardiovascular events using additional biosensor data. Incorporating transformer-based time-series modeling, improving interpretability with explainable AI (XAI) modules, and deploying energy-efficient model compression for wearables could further enhance system performance. With broader adoption, this FL-Hybrid framework can evolve into a cornerstone for continuous, personalized, and secure chronic disease monitoring across global populations.

Ethical Statement

This study did not require formal ethical approval because institutional guidelines do not require IRB/ethics committee approval for research involving the analysis of secondary, publicly available, and de-identified datasets. This exemption is based on the National Ethical Guidelines for Biomedical and Health Research Involving Human Participants issued by the Indian Council of Medical Research (ICMR).

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in PhysioNet at https://doi.org/10.13026/zthx-5212, reference number [29].

Author Contribution Statement

M. Alamelu: Conceptualization, Methodology, Software, Validation, Writing – review & editing, Visualization. **Meera Alphy:** Software, Validation, Formal analysis, Data curation, Writing – original draft. **Finney Daniel Shadrach:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Visualization, Supervision. **Jayaraj Velusamy:** Data curation, Writing – review & editing, Project administration.

References

[1] Yesodha, K., Krishnamurthy, M., Selvi, M., & Kannan, A. (2024). Intrusion detection system extended CNN and artificial bee colony optimization in wireless sensor networks. *Peer-to-Peer Networking and Applications*, *17*(3), 1237–1262. https://doi.org/10.1007/s12083-024-01650-w

[2] Al-Fuhaidi, B., Farae, Z., Al-Fahaidy, F., Nagi, G., Ghallab, A., & Alameri, A. (2024). Anomaly-based intrusion detection system in wireless sensor networks using machine learning algorithms. *Applied Computational Intelligence and Soft Computing*, *2024*(1), 2625922. https://doi.org/10.1155/2024/2625922

[3] Hamid, D. E., Amer, H. M., Moustafa, H. E.-D. S., & Marie, H. S. (2024). Empowering health data protection: Machine learning-enabled diabetes classification in a secure cloud-based IoT framework. *Indonesian Journal of Electrical Engineering and Computer Science*, *34*(2), 1100–1111.

[4] S., G., Siva Reddy, R. V., & Ahmed, M. R. (2024). Personalized diabetes diagnosis using machine learning and electronic health records. *International Journal of Electrical and Computer Engineering*, *14*(4), 4791–4801. https://doi.org/10.11591/ijece.v14i4.pp4791-4801

[5] Chen, T.-C. (2025). Double exponential smoothing slime mould algorithm for disease detection in IoT healthcare system. *The European Physical Journal Plus*, *140*(1), 90. https://doi.org/10.1140/epjp/s13360-025-06032-6

[6] Cherukuri, B. R. (2025). Advanced multi class cyber security attack classification in IoT based wireless sensor networks using context aware depthwise separable convolutional neural network. *Journal of Machine and Computing*, *5*(2), 814–830. https://doi.org/10.53759/7669/jmc202505064

[7] Hossain, M. S., & Shorfuzzaman, M. (2025). Early diabetic retinopathy cyber-physical detection system using attention-guided deep CNN fusion. *IEEE Transactions on Network Science and Engineering*, *12*(3), 1898–1910. https://doi.org/10.1109/TNSE.2025.3541138

[8] Anitha, K., Sethukarasi, T., Abinaya, K., & Radhika, S. (2025). Improved diabetic retinopathy detection using ERXG-PS ensemble algorithm and modified principal component analysis. *IETE Journal of Research*, *71*(3), 1063–1075. https://doi.org/10.1080/03772063.2024.2438325

[9] Singh, O., Kushwaha, A. S., Sharma, Col. R., & Kumar, L. (2025). Improving energy efficiency in scalable WSNs through IoT-driven approach. *International Journal of Information Technology*, *17*(5), 2659–2669. https://doi.org/10.1007/s41870-025-02471-7

[10] Adam Sahib, R. B., & Bhavani, R. (2025). IoT-based smart healthcare using efficient data gathering and data analysis. *Peer-to-Peer Networking and Applications*, *18*(1), 43. https://doi.org/10.1007/s12083-024-01823-7

[11] Srivastava, A., & Paulus, R. (2023). Coverage hole aware optimal cluster based routing for wireless sensor network assisted IoT using hybrid deep recurrent neural network. *Concurrency and Computation: Practice and Experience*, *35*(6), 1. https://doi.org/10.1002/cpe.7535

[12] Sarkar, P. J., Chanagala, S., Chandra, G. C. J., Ruby, U., & Manda, K. (2024). Prediction model for diabetes mellitus using machine learning algorithms for enhanced diagnosis and prognosis in healthcare. *Computer and Telecommunication Engineering*, *2*(1), 2446. https://doi.org/10.54517/cte.v2i1.2446

[13] Moein, T., Keramati, M. A., & Moinzad, H. (2024). Optimizing the performance reliability of diagnostic equipment and wearable sensors and medical devices in IOMT. *International Journal of Nonlinear Analysis and Applications*, *16*(1), 75–88.

[14] Karthikeyan, M., Manimegalai, D., & RajaGopal, K. (2024). Firefly algorithm based WSN-IoT security enhancement with machine learning for intrusion detection. *Scientific Reports*, *14*(1), 231. https://doi.org/10.1038/s41598-023-50554-x

[15] Hosain, M. N., Kwak, Y.-S., Lee, J., Choi, H., Park, J., & Kim, J. (2024). IoT-enabled biosensors for real-time monitoring and early detection of chronic diseases. *Physical Activity and Nutrition*, *28*(4), 060–069. https://doi.org/10.20463/pan.2024.0033

[16] Yagnik, S., Viswanathan, R., & Cao, L. (2024). Energy efficient signal detection using SPRT and ordered transmissions in wireless sensor networks. *IEEE Open Journal of Signal Processing*, *5*, 1119–1133. https://doi.org/10.1109/OJSP.2024.3488530

[17] Shaheen, I., Javaid, N., Alrajeh, N., Asim, Y., & Aslam, S. (2024). Hi-Le and HiTCLe: Ensemble learning approaches for early diabetes detection using deep learning and explainable artificial intelligence. *IEEE Access*, *12*, 66516–66538. https://doi.org/10.1109/ACCESS.2024.3398198

[18] S., T. S., & P., V. (2024). Connect attack in IoT-WSN detect through cyclic analysis based on forward and backward elimination. *PeerJ Computer Science*, *10*, e2130. https://doi.org/10.7717/peerj-cs.2130

[19] Farooq, M. S., Riaz, S., Tehseen, R., Farooq, U., & Saleem, K. (2023). Role of Internet of things in diabetes healthcare: Network infrastructure, taxonomy, challenges, and security model. *Digital Health*, *9*, 20552076231179056. https://doi.org/10.1177/20552076231179056

[20] Venkatesan, K., Malathi, V., & Prabakaran, S. (2024). IoT-based smart detection of diabetes using machine learning algorithms. *African Journal of Biological Sciences*, *6*(15), 7660–7672.

[21] Chatterjee, R., Akhtar, M. A. K., Pradhan, D. K., Chakraborty, F., Kumar, M., Verma, S., ..., & García-Arenas, M. (2024). FNN for diabetic prediction using oppositional whale optimization

algorithm. *IEEE Access*, *12*, 20396–20408. https://doi.org/10.1109/ACCESS.2024.3357993

[22] Jabbar, A., Liaqat, H. B., Akram, A., Sana, M. U., Azpíroz, I. D., Diez, I. D. L. T., & Ashraf, I. (2024). A lesion-based diabetic retinopathy detection through hybrid deep learning model. *IEEE Access*, *12*, 40019–40036. https://doi.org/10.1109/ACCESS.2024.3373467

[23] Naz, H., Saba, T., Alamri, F. S., Almasoud, A. S., & Rehman, A. (2024). An improved robust fuzzy local information K-Means clustering algorithm for diabetic retinopathy detection. *IEEE Access*, *12*, 78611–78623. https://doi.org/10.1109/ACCESS.2024.3392032

[24] Li, M., Xu, P., Hu, J., Tang, Z., & Yang, G. (2025). From challenges and pitfalls to recommendations and opportunities: Implementing federated learning in healthcare. *Medical Image Analysis*, *101*, 103497. https://doi.org/10.1016/j.media.2025.103497

[25] Zhang, F., Kreuter, D., Chen, Y., Dittmer, S., Tull, S., Shadbahr, T.,..., & Roberts, M. (2024). Recent methodological advances in federated learning for healthcare. *Patterns*, *5*(6), 101006. https://doi.org/10.1016/j.patter.2024.101006

[26] Abbas, S. R., Abbas, Z., Zahir, A., & Lee, S. W. (2024). Federated learning in smart healthcare: A comprehensive review on privacy, security, and predictive analytics with IoT integration. *Healthcare*, *12*(24), 2587. https://doi.org/10.3390/healthcare12242587

[27] Eden, R., Chukwudi, I., Bain, C., Barbieri, S., Callaway, L., de Jersey, S., ..., & Sullivan, C. (2025). A scoping review of the governance of federated learning in healthcare. *npj Digital Medicine*, *8*(1), 427. https://doi.org/10.1038/s41746-025-01836-3

[28] Shah, S. T., Ali, Z., Waqar, M., & Kim, A. (2025). Federated learning in public health: A systematic review of decentralized, equitable, and secure disease prevention approaches. *Healthcare*, *13*(21), 2760. https://doi.org/10.3390/healthcare13212760

[29] Cho, P., Kim, J., Bent, B., & Dunn, J. (2023). BIG IDEAs lab glycemic variability and wearable device data [Data set]. PhysioNet.

[30] Sunny, M. N. M., Sakil, M. B. H., Atayeva, J., Munmun, Z. S., Mollick, M. S., & Faruq, M. O. (2024). Predictive healthcare: An IoT-based ANFIS framework for diabetes diagnosis. *Engineering*, *16*(10), 325–336. https://doi.org/10.4236/eng.2024.1610024

[31] Baba, S. M., Bala, I., Dhiman, G., Sharma, A., & Viriyasitavat, W. (2024). Automated diabetic retinopathy severity grading using novel DR-ResNet + deep learning model. *Multimedia Tools and Applications*, *83*(28), 71789–71831. https://doi.org/10.1007/s11042-024-18434-2

How to Cite: Alamelu, M., Alphy, M., Shadrach, F. D., & Velusamy, J. (2026). Deep Reinforcement Learning-Enabled IoT Framework for Real-Time and Personalized Diabetes Diagnosis Using Wearable Sensors. *Artificial Intelligence and Applications*. https://doi.org/10.47852/bonviewAIA62026306