Journal of Pioneering Medical Sciences
Received: January 15, 2026 | Accepted: May 25, 2026 | Published: July 05, 2026
Volume 15, Issue 06, Pages 214-228

DOI https://doi.org/10.47310/jpms2026150628

A Differentiable Hybrid Classical-Quantum Circuit with PCA Feature Compression for Medical Disease Classification in Healthcare Datasets

S. Megala[1], P. Ragukumar[2*] and Neelu Khare[3]
[1,2]Department of Mathematics, Vellore Institute of Technology, Vellore 632014, India
[3]Department of Software and Systems Engineering, Vellore Institute of Technology, Vellore 632014, India

Author Designation: [1]Research Scholar, [2]Assistant Professor, [3]Professor

*Corresponding author: P. Ragukumar (e-mail: ragukumar.p@vit.ac.in).

Abstract Machine Learning (ML) approaches have achieved high performance by discovering underlying data patterns and structures. Diabetes is a life-threatening disease that affects vital organs such as the eyes (retinopathy), nerves (neuropathy), heart and kidneys (nephropathy). Although various ML models have been proposed for diabetes prediction, most struggle to effectively handle large, multi-feature datasets. Quantum Machine Learning (QML), which integrates quantum computing principles with machine learning, offers a promising solution for complex classification tasks. However, effective feature extraction remains challenging on current NISQ (Noisy Intermediate-Scale Quantum) devices due to hardware limitations. In this study, a Differentiable Quantum Circuit with Classical PCA Feature Compression is proposed for medical disease classification using a SMOTE-ENN resampled dataset and a penalized binary cross-entropy cost function to enhance quantum learning. The efficient ansatz designed in this work plays a crucial role in determining the model's learning capability and feature extraction performance. To evaluate the robustness and generalization capability of the proposed model, three categories of medical datasets are considered. The primary focus is on diabetes classification using four diabetes-related datasets, while additional experiments are conducted on liver disease and chronic kidney disease datasets to assess the adaptability of the proposed framework across multiple healthcare domains. Six datasets from UCI, Mendeley and IEEE repositories are used: PIMA Indian Diabetes, Type-2 Diabetes, a private RTML Insulin dataset, Indian Liver Patient Dataset (ILPD), Early-Stage Diabetes Risk Prediction Dataset (ESDRPD) and Chronic Kidney Disease (CKD). The proposed quantum model is simulated using the PennyLane quantum simulator and compared against preprocessed classical classifiers. Experimental results demonstrate remarkable efficiency and rapid convergence within 10--20 epochs, achieving high F1-scores of 96.15% (PIMA), 97.50% (Type-2), 98.04% (RTML), 89.47% (ILPD), 98.66% (ESDRPD) and 97.84% (CKD). These findings demonstrate that the proposed quantum model efficiently extracts relevant features and effectively handles medical disease classification tasks, highlighting its strong potential for practical healthcare applications.

Key Words Quantum Machine Learning, Diabetes Classification, PCA Feature Compression, SMOTE ENN Balancing Technique, Parameterized Quantum Circuit

INTRODUCTION

According to the World Health Organization (WHO), diabetes mellitus (including kidney disease deaths attributed to diabetes) ranked as the eighth leading cause of death globally in the most recent WHO report [1]. The WHO highlights that diabetes is responsible for millions of deaths annually and contributes significantly to the global burden of chronic diseases. Similarly, the International Diabetes Federation (IDF) Diabetes Atlas (11th Ed., 2025) reports that 11.1% of the global adult population (aged 20-79 years)-equivalent to approximately 1 in 9 adults-is currently living with diabetes [2].

Figure 1 illustrates the top countries or territories ranked by the number of adults aged 20-79 years living with diabetes in 2024, along with projections for 2050 [2]. These projections indicate that if current trends continue, the diabetes burden will nearly double in several regions, underscoring the urgent need for stronger public health

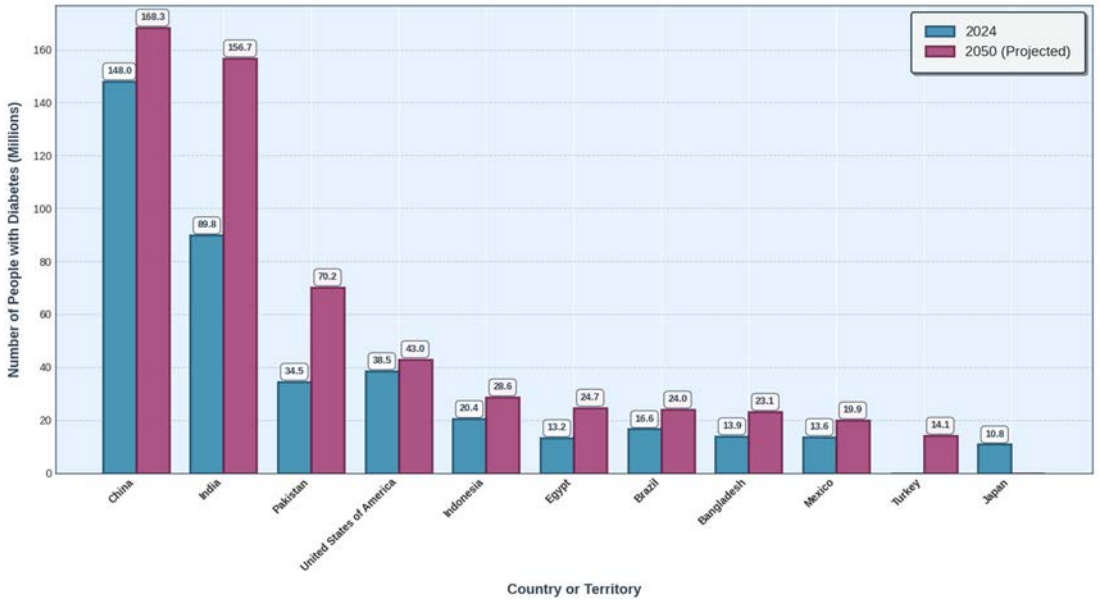

Figure 1: The Top Countries or Territories Based on the Number of Adults (Aged 20-79) Living with Diabetes in 2024 and 2050 Projections

policies, improved healthcare system responses and effective population-level prevention strategies. Moreover, diabetes leads to long-term damage of the human body parts causing serious, life-threatening and disabling complications such as heart attacks, strokes, kidney failure, vision loss and lower limb amputations [1]. Early prediction of diseases such as diabetes can play a crucial role in controlling the condition and preventing life-threatening complications, thereby saving human lives. Early risk prediction of diabetes can be effectively achieved using machine learning techniques. Recent studies have demonstrated promising results in predicting the risk of diabetes mellitus by leveraging various supervised learning algorithms.

This success is primarily attributed to the computational efficiency and strong data-handling capabilities of these algorithms [3]. In Sadhu and Jadli [3], the authors studied various machine learning algorithms using the Early Stage Diabetes Risk Prediction dataset. They employed K-Nearest Neighbors (KNN), Logistic Regression (LR), Support Vector Machines (SVM) with an RBF kernel, Naive Bayes (NB), Decision Tree (DT), Random Forest (RF) and Multi-Layer Perceptron (MLP) classifiers. Among these, the Random Forest classifier achieved the highest accuracy of 98.07%. In Ihnaini *et al.* [4], a smart healthcare recommendation system was developed for diabetes prediction using deep ensemble learning with data fusion. The study utilized the Hospital Frankfurt Germany and Pima Indians Diabetes datasets, achieving 91 and 72.73% accuracy before data fusion and an improved 99.64% accuracy after applying data fusion. Chronic Kidney Disease (CKD) is a critical health condition that can lead to severe complications if not identified and treated early. The early detection through routine blood or urine tests is critical, enabling timely medication and monitoring to prevent disease progression to advanced stages [5]. In this study, twelve supervised machine learning classifiers were evaluated for CKD prediction using the Chronic Kidney Disease dataset. Both the original dataset and a dimensionally reduced version obtained via Principal Component Analysis (PCA) were analyzed. The XGBoost classifier demonstrated superior performance, achieving an accuracy of 98.3% on the original dataset and 99.16% on the PCA-reduced dataset, highlighting the positive impact of feature dimensionality reduction on classification performance. The Rownak Textile Mills Limited (RTML) dataset comprises medical records collected from female employees of Rownak Textile Mills Ltd., Dhaka, Bangladesh [6]. In the reported study, the RTML dataset was combined with the Pima Indian Diabetes dataset to increase sample diversity.

To mitigate class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was applied. Eight classical machine learning classifiers including DT, LR, kNN, AdaBoost, RF and Naïve Bayes (NB)-were evaluated within a supervised learning framework. The XGBoost classifier achieved the best performance, with an accuracy of 83.22% and an area under the ROC curve (AUC) of 0.86 on the combined dataset [6].

The challenge of imbalanced datasets in liver disease prediction has been effectively addressed through hybrid ensemble learning approaches. A comprehensive study by Donaire *et al.* [7] implemented and evaluated multiple machine learning models. The research developed novel hybrid ensemble models that integrate Recursive Feature Elimination (RFE) for optimal feature selection, SMOTE-ENN (Synthetic Minority Over-sampling Technique with Edited Nearest Neighbors) for data balancing and ensemble learning techniques to handle class imbalance effectively.

These hybrid models demonstrated superior performance compared to standalone classifiers, achieving a remarkable accuracy of 93.2% on the Indian Liver Patient Dataset (ILPD) and 95.4% on the BUPA Liver Disorder Dataset. The study's findings underscore the critical importance of combining data balancing techniques with ensemble methods to enhance diagnostic accuracy in imbalanced medical datasets, providing a robust framework for reliable liver disease prediction systems.

Although classical machine learning algorithms have demonstrated superior performance and effective learning capabilities, the exponential growth of medical healthcare data presents new challenges and opportunities. Larger training datasets typically yield more accurate predictions, enabling early diagnosis that can be crucial for saving human lives. However, to fully leverage these massive datasets and achieve further advancements in predictive accuracy, additional computational paradigms such as quantum computing are required. Since the foundational proposal of quantum computing by Feynman [8], the field has undergone significant theoretical and experimental advancements, offering promising solutions for handling large-scale medical data analysis and complex pattern recognition tasks.

The study by Hancco-Quispe *et al.* [9] applied quantum machine learning techniques to diabetes classification using the PIMA dataset. The research implemented various classical classifiers including LR, KNN and NB, alongside preprocessing techniques such as PCA and LDA for dimensionality reduction. Two quantum algorithms were evaluated: Quantum Support Vector Classifier (QSVC) and Variational Quantum Classifier (VQC). The experimental results demonstrated that the combination of LDA for dimensionality reduction with quantum classifiers achieved superior performance, highlighting the potential of quantum machine learning approaches in medical diagnosis applications.

In Safriandono *et al.* [10], Quantum Feature Engineering (QFE) was applied in conjunction with SMOTE-Tomek class balancing on the ILPD dataset. The study evaluated classical machine learning models under two configurations: without QFE and with QFE preprocessing. Using traditional preprocessing with SMOTE-Tomek, with QFE-SVM achieved the highest accuracy of 72%. When QFE was incorporated with SMOTE-Tomek balancing, XGBoost attained an improved accuracy of 81%, demonstrating a substantial 9% enhancement and validating the effectiveness of quantum-enhanced feature engineering for imbalanced medical datasets.

Donaire *et al.* [7] proposed a hybrid quantum-classical model, QML-Liver, for liver disease classification on the ILPD dataset. The model combines three classical neural network layers with a single quantum layer using angle embedding and basic entanglement on one qubit. Trained for 100 epochs, QML-Liver without SMOTE achieved 83% accuracy, while the SMOTE-balanced version attained 75% accuracy but a higher precision of 89%, highlighting its effectiveness in reducing false positives for medical diagnostics. A comprehensive summary of key findings from the literature on diabetes prediction, encompassing classical, quantum and hybrid machine learning models evaluated across various datasets with different data balancing techniques, is presented in Table 8.

Motivation and Contributions

Advanced machine learning methods, including support vector machines, neural networks and decision tree classifiers, are frequently employed for disease diagnosis in healthcare systems [3-6]. However, classical models face significant challenges including requirements for large balanced datasets, computational limitations and difficulty capturing complex medical patterns [7].

Quantum machine learning offers a promising alternative by exploiting superposition and entanglement to learn complex patterns from smaller datasets through efficient ansatz representations. Despite this potential, current quantum systems face critical limitations: restricted qubit availability, noise and decoherence errors, challenges in encoding classical data into quantum states and difficulties in optimizing quantum neural network parameters [7]. These gaps between quantum computing's theoretical advantages and practical constraints motivate the development of resource-efficient quantum algorithms for medical diagnostics. In this work, we design and optimize a Differentiable Quantum Circuit with Classical PCA Feature Compression for diabetes classification. Our model integrates classical advantages in the preprocessing phase with quantum computing's capability to learn complex patterns from training data through efficient quantum state representations. The main contributions of this study are as follows:

- In the pre-processing stage, PCA is used to process all the medical healthcare datasets to effectively incorporate all of them. This dimensionality reduction aids in effective utilization of quantum resources by representing key features of the input into a quantum system with a reduced number of qubits
- An Efficient Ansatz Design is proposed for binary diabetes classification. Since the performance of quantum models strongly depends on the structure of the Ansatz, the design is carefully optimized to enhance learning capability and model expressiveness
- The proposed Ansatz significantly reduces the number of training epochs required for convergence. This improvement is particularly beneficial for (NISQ) devices, where shorter training times help mitigate noise and decoherence. With reduced qubit requirements due to PCA, the Ansatz efficiently learns feature patterns
- To prevent overfitting and ensure balanced learning of target classes, a combination of SMOTE-ENN resampling and a penalized cost function is employed, leading to improved model robustness
- The generalization capability of the proposed quantum architecture is validated through extensive experiments on six diabetes datasets of varying sizes and complexities, covering both small and large sample distributions

- The performance of the proposed Differentiable Hybrid Classical-Quantum Circuit is benchmarked against a comprehensive set of classical baseline classifiers, categorized into four groups: linear-based classifiers (Logistic Regression, Linear Discriminant Analysis), tree-based classifiers (Decision Tree, Extra Tree), ensemble-based classifiers (Random Forest, AdaBoost, Gradient Boosting, XGBoost) and probabilistic-based classifiers (Naïve Bayes), using standard evaluation metrics accuracy, precision, recall and F1-score, across all six datasets to demonstrate the superiority and efficiency of the proposed quantum model

The paper follows the following structure: Sec. 3 presents the Proposed Model Design, detailing the Classical Preprocessing Phase, Quantum Training Phase, Prediction and Inference Phase, Parameter Optimization Phase and Test Data Validation Phase. Sec. 4 provides Experimental Setup and Performance Analysis, including Dataset Information, Comparative Methods, Results & discussion and Sec. 5 Concludes the paper.

Model Design

Figure 2 illustrates the overall workflow of the proposed Differentiable Quantum Circuit with Classical PCA Feature Compression for Diabetes Classification. The workflow is divided into two major phases: (i) The Classical Preprocessing Phase and (ii) The Quantum Main Phase. In the Classical Preprocessing Phase, the raw medical dataset is transformed into a suitable format for quantum computation. The Quantum Main Phase is responsible for the core learning process. Each of these phases and modules is discussed in detail in the following sections.

Classical Preprocessing Phase

In this phase, the raw medical data is transformed into a suitable input format for quantum processing. Proper preprocessing is essential, as it directly influences the performance of the quantum circuit during training. This stage typically includes tasks such as Data Cleaning, Feature Engineering, Handling Class Imbalance and Quantum Feature Scaling.

Data Cleaning

The raw diabetes dataset is represented as:

$$\mathcal{D} = \left\{\left(\mathbf{x}^{(i)}, y^{(i)}\right) \middle| \mathbf{x}^{(i)} \in \mathbb{R}^d,\ y^{(i)} \in \mathcal{Y},\ i = 1,2,\dots,N\right\}$$

where, d denotes the number of features, $\mathcal{Y} = \{0,1,\dots,C-1\}$ is the set of class labels, C is the number of classes and N is the total number of samples.

The first step in preprocessing is to remove any duplicate entries $(x^{(i)}, y^{(i)})$ from the dataset. Following this, missing values (denoted as NaN) and biologically invalid zeros are handled using median imputation, ensuring that all features contain meaningful and valid values.

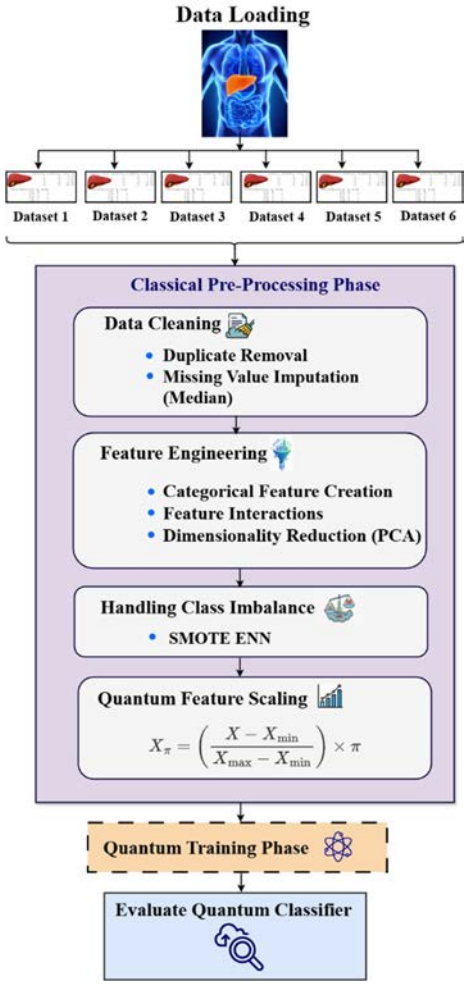

Figure 2: Overall Workflow of the Proposed Quantum Diabetes Classification Algorithm

Let the dataset after duplicate removal be denoted as:

$$\mathcal{D}' = \left\{\left(x^{(i)}, y^{(i)}\right)\right\}_{i=1}^{N'}$$

For each feature $j \in \{1,2,\dots,d\}$, missing or invalid entries in the j-th feature vector are replaced by the median of the observed valid values for that feature. Mathematically, the imputation for $x_j^{(i)}$ (the j-th feature of the i-th sample) can be expressed as:

$$x_j^{(i)} = \begin{cases} \text{median}(x_j), & \text{if } x_j^{(i)} = 0 \text{ or } x_j^{(i)} = \text{NaN}, \\ x_j^{(i)}, & \text{otherwise} \end{cases} \quad (1)$$

Equation 1 ensures that the dataset is free from duplicates and invalid entries, providing a reliable foundation for subsequent quantum feature encoding and learning.

Feature Engineering

Feature engineering is applied to enhance the predictive capability of the dataset by generating new meaningful

features from existing ones. This process includes the creation of categorical variables, interaction features and domain-specific transformations, which help the quantum model learn more discriminative patterns. Categorical feature creation converts continuous variables into clinically interpretable classes.

For example, Body Mass Index (BMI), Glucose level and Age are grouped into medically relevant categories such as normal, overweight, obese or diabetic. In addition to this, interaction features are generated by combining two or more significant risk factors associated with diabetes prediction. As discussed in Changpetch *et al.* [12], classification models that include interaction features perform better than those using only individual features, as interactions capture complex relationships between variables. Likewise, the study in Phongying and Hiriote [13] demonstrates that incorporating interactions among key diabetes risk factors-such as BMI, glucose levels and family history-significantly improves prediction accuracy. These enriched features, as presented in Table 1, provide deeper insights into the dataset, thereby enhancing both classical preprocessing and the learning performance of the proposed quantum model.

Two interaction features were defined to capture combined effects of key attributes:

$$\text{BMI\_Age\_Interaction} = \frac{\text{BMI} \times \text{Age}}{1000}$$

and:

$$\text{Glucose\_BMI\_Interaction} = \frac{\text{Glucose} \times \text{BMI}}{1000}$$

These newly created categorical and interaction features are added to the dataset before training, enabling the model to capture more comprehensive relationships among variables.

The next step in preprocessing is dimensionality reduction. In this study, PCA is employed, which is a widely used statistical technique for reducing the dimensions. PCA identifies a new set of orthogonal variables, called principal components, that capture the maximum variance present in the data [14]. The objective of PCA is to retain the most informative features while eliminating redundant or less significant ones, thereby preserving essential information without substantial loss. This reduction is particularly important in quantum machine learning, as it minimizes the number of qubits required during the quantum training phase. By applying PCA, the number of qubits required in the quantum circuit is significantly reduced. This enables efficient quantum training while still retaining essential information from all input features in the form of linear combinations.

After applying PCA, a set of d new principal component features is obtained. From these, the top p principal components that contribute the most to the total variance are selected for training. The optimal number of principal components is determined through hyperparameter tuning to ensure the best trade-off between accuracy and computational efficiency.

Handling Class Imbalance

To effectively address the issue of class imbalance in the dataset, the Synthetic Minority Oversampling Technique combined with Edited Nearest Neighbors (SMOTE-ENN) is applied. This hybrid approach, proposed by Batista and Prati in Batista *et al.* [15], integrates the strengths of both SMOTE and ENN. SMOTE generates synthetic minority class samples to balance the dataset; however, this may also introduce noisy or unreliable samples. To overcome this, ENN is used to clean the dataset by removing misclassified and ambiguous samples using the nearest neighbor rule. As a result, the training dataset becomes more robust, enabling the quantum model to learn more discriminative and reliable patterns for diabetes classification.

Quantum Feature Scaling

After completing all previous preprocessing steps, the dataset is prepared for quantum processing by scaling all feature values to the range $[0, \pi]$ using Min-Max normalization. This ensures that each input value $x_j^{(i)}$ (corresponding to the j-th feature of the i-th sample) satisfies: $0 \le x_j^{(i)} \le \pi$. The normalization is performed according to the following Equation 2:

$$x_j^{(i)} = \frac{(x_j^{(i)} - x_{\min})}{(x_{\max} - x_{\min})} \times \pi \tag{2}$$

Scaling features to this range is essential for quantum circuits, as it allows each value to be directly encoded as a rotation angle in quantum gates. This ensures that the quantum state can fully capture the rotational impact of all input features, facilitating effective quantum feature

Table 1: Categorical Feature Creation for BMI, Glucose and Age

Feature	Category	Condition
BMI	BMI_Normal	BMI<25
	BMI_Overweight	25≤BMI<30
	BMI_Obese	BMI≥30
Glucose	Glucose_Normal	Glucose<100
	Glucose_Prediabetic	100≤Glucose<126
	Glucose_Diabetic	Glucose≥126
Age	Age_Young	Age<30
	Age_Middle	30≤Age<50
	Age_Senior	Age≥50

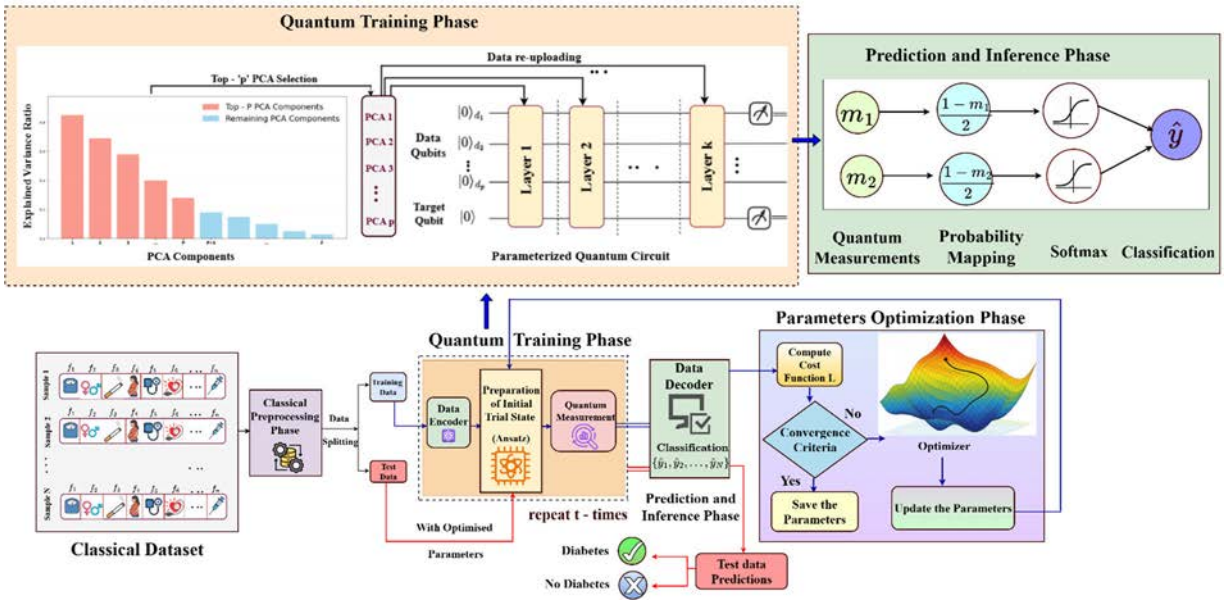

Figure 3: Workflow of The Proposed Quantum Training Framework for Diabetes Classification
The Diagram Depicts the Complete Quantum Learning Process, Where the Blue Path Represents the Training Pipeline and the Red Path Indicates the Testing (Inference) Phase

encoding and improving the model's learning performance. After normalization, the dataset is divided into training and testing sets, where 70% of the data is used for training and the remaining 30% for testing.

Quantum Main Phase

In this stage, the preprocessed classical data is encoded into a quantum circuit for further processing. The proposed model, consisting of a sequence of quantum gates with trainable parameters that are optimized during the learning process. The architecture is structured into four key phases: The Quantum Training Phase, Prediction and Inference Phase, Parameters Optimization Phase and Test Data Validation Phase.

The qubit register for this model is divided into two distinct parts: d feature qubits, which extract information from the input data and c target qubits, which are utilized for classification tasks. Figure 3 illustrates the detailed workflow of the proposed model for diabetes classification across various diabetes datasets, providing a clear representation of the overall process.

Quantum Training Phase

The Quantum Training Phase is responsible for learning patterns from the input features to extract meaningful knowledge from the data. The proposed framework for this phase consists of six core modules: Data Encoder, Data Qubit Gates, Data Qubit Dynamic Entanglements, Data & Target Qubit Dynamic Entanglement, Target Qubit Gates and Measurement on Data and Target Qubits.

Each module performs a distinct role in the quantum learning process, collectively enabling efficient feature encoding, meaningful interaction between data and target qubits and accurate measurement for classification. The detailed design and operation of each module are presented in the following subsections.

Data Encoder

Let d denote the number of feature qubits, where $d \in \{d_1, d_2, d_3, \dots, d_p\}$ and p represents the number of principal components selected by hyperparameter optimization.

The top p PCA-transformed features are encoded and fed into the quantum circuit layers. Let c denote the number of target qubits (here $c = 1$), k the number of circuit layers and t the number of training epochs. The classical input vector after PCA transformation is given by: $x = (x_0, x_1, \dots, x_{d_p-1}) \in \mathbb{R}^{d_p}$. The initial quantum state is prepared in the all-zero computational basis state: $|\psi_0\rangle = |0\rangle^{\otimes(d_p+c)}$. The data encoding is performed using the angle embedding unitary operator $U_{enc}(x)$, defined as:

$$U_{enc}(x) = \otimes_{i=0}^{d_p-1}(RY_i(x_i) \cdot RX_i(x_i)) \otimes I^{\otimes c} \quad (3)$$

where the gate product is evaluated from right to left, meaning $RX_i(x_i)$ is applied first, followed by $RY_i(x_i)$. Here $RX_i(x_i)$ and $RY_i(x_i)$ are single-qubit rotation gates around the X-axis and Y-axis, respectively, applied to the i-th qubit. These are defined as:

$$RX_i(x_i) = e^{-i\frac{x_i}{2}X}, \quad RY_i(x_i) = e^{-i\frac{x_i}{2}Y}$$

where, X and Y denote the Pauli-X and Pauli-Y operators, respectively. The general single qubit rotation gate is defined as:

$$R_{\alpha,i}(x_i) = \exp\left(-i\frac{x_i}{2}\sigma_\alpha\right), \quad \alpha \in \{X, Y, Z\}$$

where, $R_{\alpha,i}(x_i)$ is the single-qubit rotation around the α-axis applied to the i-th qubit and σ_α denotes the corresponding Pauli matrix. The identity operator $I^{\otimes c}$ acts on the c target qubits, which remain unaltered during this encoding step.

Data Qubit Gates

To extract meaningful features from the input data distribution, a parameterized quantum gate block is applied to the feature qubit register. Let the learnable parameters of the ℓ-th layer be denoted as $\theta^{(\ell)} \in \mathbb{R}^{3\times d_p}$, where d_p represents the number of feature qubits and $\ell \in \{0,1,\ldots,k-1\}$. The parameter matrix for the ℓ-th layer is expressed as:

$$\theta^{(\ell)} = \begin{bmatrix} \theta_{0,0}^{(\ell)} & \theta_{0,1}^{(\ell)} & \cdots & \theta_{0,d_p-1}^{(\ell)} \\ \theta_{1,0}^{(\ell)} & \theta_{1,1}^{(\ell)} & \cdots & \theta_{1,d_p-1}^{(\ell)} \\ \theta_{2,0}^{(\ell)} & \theta_{2,1}^{(\ell)} & \cdots & \theta_{2,d_p-1}^{(\ell)} \end{bmatrix}$$

The corresponding unitary transformation applied to the feature qubits is defined as:

$$U_{Dgate}^{(\ell)}(\theta^{(\ell)}) = \left[\bigotimes_{i=0}^{d_p-1}\left(RY_i\left(\theta_{2,i}^{(\ell)}\right)RZ_i\left(\theta_{1,i}^{(\ell)}\right)RY_i\left(\theta_{0,i}^{(\ell)}\right)\right)\right] \otimes I^{\otimes c} \quad (4)$$

where, $I^{\otimes c}$ denotes the identity operation acting on the c target qubits. This block introduces $3 \times d_p$ learnable parameters per layer and enables the circuit to capture the input features through tunable single-qubit rotations.

Data Qubit Dynamic Entanglements

This module introduces dynamic learnable entanglement among the data qubits. The purpose is to enable controlled information sharing between neighboring feature qubits, guided by tunable entanglement strengths. Only data qubits participate in these interactions. The Ising entangling gate between two qubits q_0 and q_1 is defined as:

$$R_{pq}(\phi) = Ising_{pq}(\phi)_{(q_0,q_1)} = \exp\left(-i\frac{\phi}{2}\sigma_p \otimes \sigma_q\right), \quad p, q \in \{X, Y, Z\}$$

where, σ_p and σ_q are the Pauli matrices corresponding to the axes p and q, respectively [16]. Let $\phi^{(\ell)} \in \mathbb{R}^{d_p}$ be the vector of learnable entanglement parameters for the ℓ-th layer, where:

$$\phi^{(\ell)} = \left[\phi_0^{(\ell)}, \phi_1^{(\ell)}, \ldots, \phi_{d_p-1}^{(\ell)}\right]$$

The unitary operator for the data-qubit entanglement module in the ℓ-th layer is defined in Equation 5:

$$U_{Dent}^{(\ell)}(\phi^{(\ell)}) = \prod_{j=0}^{d_p-1} Ising\left(\phi_j^{(\ell)}\right)_{(j,(j+1) \bmod d_p)} \quad (5)$$

This configuration ensures that each data qubit is entangled with its nearest neighbor, including a wrap-around entanglement between the last and the first qubit. Hence, this module introduces d_p trainable entanglement parameters per layer.

Data and Target Qubit Dynamic Entanglement

This module establishes dynamic interactions between the data qubits and the target qubits using parameterized two-qubit entangling gates. These interactions allow information to flow from selected wrapped data qubits to the target qubits during each layer of the circuit. Let $\gamma^{(\ell)} \in \mathbb{R}^c$ be the vector of learnable entanglement parameters for the ℓ-th layer, defined as:

$$\gamma^{(\ell)} = \left[\gamma_0^{(\ell)}\right]$$

The unitary operation for this module is defined in Equation 6:

$$U_{DTent}^{(\ell)}(\gamma^{(\ell)}) = Ising^{(\ell)}\left(\gamma_0^{(\ell)}\right)_{(d_0,c)} \quad (6)$$

where, the Ising interaction is applied between the first data qubit d_0 and the target qubit c. The parameter $\gamma_0^{(\ell)}$ controls the entanglement strength of this interaction. In this way, information is effectively transferred from the feature qubits to the target qubits, with the flow being modulated by the learnable parameter $\gamma^{(\ell)}$. This module introduces one learnable parameter per layer.

Target Qubit Gates

This module applies unitary transformations exclusively to the target qubit(s) c. Since in this architecture c = 1, these gates enable the network to further process and reshape the entangled information received from the feature qubits. Let $\beta^{(\ell)} \in \mathbb{R}^{1\times 3}$ denote the learnable parameters of this module for the ℓ-th layer, defined as:

$$\beta^{(\ell)} = \left[\beta_{0,0}^{(\ell)} \quad \beta_{0,1}^{(\ell)} \quad \beta_{0,2}^{(\ell)}\right]$$

Each row corresponds to the rotation angles applied to a single target qubit. The unitary operator for this module is defined in Equation 7:

$$U_{Tgates}^{(\ell)}(\beta^{(\ell)}) = I^{\otimes d} \otimes Rot\left(\beta_{0,0}^{(\ell)}, \beta_{0,1}^{(\ell)}, \beta_{0,2}^{(\ell)}\right)_c \quad (7)$$

where, $c = d_p + 1$ denotes the index of the target qubit in the full quantum register, $I^{\otimes d}$ is the identity operator acting on the d feature qubits and Rot(·) represents a general single-qubit rotation gate defined as: $Rot(\theta, \phi, \lambda) = RZ(\theta) \cdot RY(\phi) \cdot RZ(\lambda)$. This module allows the target qubit to capture non-linear combinations of the entangled information, enhancing the circuit's capability to model complex decision boundaries. This module introduces 1×3 learnable parameters per layer.

Final Cumulative Quantum States

The quantum circuit begins with an initial quantum state : $|\psi_0\rangle = |0\rangle^{\otimes(d_p+c)}$. By sequentially applying the modules, the cumulative unitary for the ℓ-th layer (Ansatz) is defined as using Equation 3-7: where, $U^{(\ell)} = U^{(\ell)}_{\text{Tgates}}(\beta^{(\ell)}) \cdot U^{(\ell)}_{\text{DTent}}(\gamma^{(\ell)}) \cdot U^{(\ell)}_{\text{Dent}}(\phi^{(\ell)}) \cdot U^{(\ell)}_{\text{Dgate}}(\theta^{(\ell)}) \cdot U_{\text{enc}}(x)$. Here, the product is evaluated from right to left, reflecting the order of unitary gate applications in quantum circuits. Here the data x is re-encoded at each layer, similar to the data re-uploading strategy used in variational quantum circuits [17]. This mechanism increases the expressive power of the model and allows better feature interaction across layers, analogous to how neurons in one layer of a classical neural network connect to all neurons in the next layer. The total Learnable Parameter per layer is: $4(d_p + 1)$. The final quantum state after k layer is presented in Equation 8:

$$|\psi_{\text{final}}\rangle = \left(\prod_{\ell=0}^{k-1} U^{(\ell)}\right)|\psi_0\rangle \tag{8}$$

Measurement on Data and Target Qubits

After applying all the preceding unitary modules, the final quantum state $|\psi_{\text{final}}\rangle$ encapsulates the transformed information distributed across both feature (data) and target qubits. To extract class-relevant information, measurements are performed on the first data qubit d_0 and the target qubit c, indexed as $d_p + 1$, using the Pauli-Z observable. The expectation value of a qubit q with respect to the Pauli-Z operator is defined as:

$$\langle Z_q\rangle = \langle\psi_{\text{final}}|Z_q|\psi_{\text{final}}\rangle \tag{9}$$

Here, Z_q denotes the Pauli-Z operator acting only on qubit q, while the identity operator I is implicitly applied to all other qubits. The quantum measurement outcomes are represented as: $m_1 = \langle Z_{d_0}\rangle$, $m_2 = \langle Z_c\rangle$ where, both m_1 and m_2 take values within the range $[-1,1]$. These expectation values serve as the quantum outputs, encoding the probabilistic information of the respective classes. The resulting values are subsequently passed to the classical post-processing layer for final decision-making or classification.

Prediction and Inference Phase

Once the final quantum state $|\psi_{\text{final}}\rangle$ is measured on the data and target qubits, the resulting values represent the expectation values of the Pauli-Z measurements on the respective qubits, as described in Equation 9. To convert these values into a usable positive scale, each component is rescaled to the interval $[0,1]$ using the following transformation:

$$\tilde{z}_q = \frac{1 - \langle Z_q\rangle}{2} \tag{10}$$

The rescaled output vector becomes $\tilde{m}_1 = \tilde{z}_{d_0}$ and $\tilde{m}_2 = \tilde{z}_c$. To interpret these rescaled outputs as class probabilities, the softmax activation function is applied, producing a normalized probability distribution across the classes:

$$p_{d_0} = \frac{\exp(\tilde{m}_1)}{\exp(\tilde{m}_1) + \exp(\tilde{m}_2)}, \quad p_c = \frac{\exp(\tilde{m}_2)}{\exp(\tilde{m}_1) + \exp(\tilde{m}_2)} \tag{11}$$

By using Equation 11, the resulting probability vector is expressed as $p = [p_{d_0}, p_c]$, which satisfies the normalization condition $p_{d_0} + p_c = 1$. Finally, the predicted class label is determined by selecting the index corresponding to the highest probability value:

$$\hat{y} = \arg\max_{j\in\{0,1\}} \; p_j \tag{12}$$

Parameter Optimization Phase

The Quantum Neural Network (QNN) is trained by optimizing all tunable parameters, including those associated with data-qubit rotations, target-qubit rotations, intra-data entanglement and data-target entanglement. The primary objective is to minimize the prediction error between the network's output probabilities and the true class labels.

The model outputs a probability vector $p(x;\Theta) = [p_{d_0}, p_c]$ (Equation 11), where $\Theta = \{\theta, \phi, \gamma, \beta\}$ represents the complete set of trainable parameters across all circuit layers. For a given training sample $(x^{(i)}, y^{(i)})$, where, $x^{(i)} \in \mathbb{R}^{d_p}$ and $y^{(i)} \in \{0,1\}$, the penalized binary cross entropy loss function is defined as:

$$\mathcal{L}(x^{(i)}, \Theta) = -\frac{1}{B}\sum_{i=1}^{B}\left[y_i \log\left(p_c^{(i)}\right) + (1 - y_i)\log\left(p_{d_0}^{(i)}\right)\right] + \lambda\sum_{j=1}^{M}\omega_j^2 \tag{13}$$

Here, $p_{d_0}^{(i)}$ and $p_c^{(i)}$ represent the predicted probabilities for Class 0 and Class 1, respectively, corresponding to the input sample $x^{(i)}$. The final term introduces an L_2 regularization over the trainable parameters $\omega = \{\omega_1, \omega_2, \ldots, \omega_M\}$ to prevent overfitting and improve generalization. To minimize this cost function, the pennylane.AdamOptimizer() is employed. This optimizer supports hybrid quantum-classical training through automatic differentiation. Gradients of the quantum circuit with respect to 'Θ' are computed using the parameter-shift rule, which ensures full compatibility with both simulation and real quantum hardware. During each training epoch t, the parameters are iteratively updated to minimize the loss and enhance classification accuracy. The key hyperparameters used during Adam optimization are summarized in Table 2.

Table 2: Hyperparameters used for Training the Quantum Model During Optimization

Hyperparameter	Value
Learning rate (α)	0.01
β_1	0.9
B_2	0.999
ϵ	1×10^{-8}
Regularization coefficient (λ)	0.001
Batch size (B)	16

Test Data Validation Phase

After completing the parameter optimization, the final optimized parameters $\hat{\Theta}$ are employed to evaluate the model's performance on the unseen test dataset. Here, the optimized parameter set is defined as:

$$\hat{\Theta} = \{\hat{\theta}^{(l)},\ \hat{\phi}^{(l)},\ \hat{\gamma}^{(l)},\ \hat{\beta}^{(l)}\}_{l=1}^{K}$$

This evaluation phase assesses the generalization capability of the variational quantum circuit, how effectively it has learned the underlying data distribution during training and how well it can predict the class labels of new, previously unseen inputs based on the learned parameters $\hat{\Theta}$.

Experimental Setup and Performance Analysis

Dataset Information: Six publicly available and private datasets are used in this study to evaluate the generalization capability of the proposed algorithm, encompassing three disease classification categories: diabetes disease classification, liver disease classification and chronic kidney disease classification. The descriptions of the independent features from six different diabetes datasets are provided in Table 3 and 4.

Diabetes Disease Classification

- The first dataset, the PIMA Indian Diabetes Dataset, contains 768 instances and is publicly available on Mendeley at https://data.mendeley.com/datasets/7zcc8v6hvp/1
- The second dataset, the Type-2 Diabetes Dataset, consists of 2000 patient records collected from Frankfurt Hospital in Germany and can be accessed through IEEE DataPort at https://ieee-dataport.org/documents/type-2-diabetes-dataset
  The third dataset, referred to as the RTML with Insulin Dataset, comprises 109 medical records of female patients from Rownak Textile Mills Ltd., Dhaka, Bangladesh and is available at https://github.com/tansin-nabil/Diabetes-Prediction-Using-Machine-Learning
- The fourth dataset, the Early-Stage Diabetes Risk Prediction Dataset (ESDRPD), contains 520 instances and is sourced from the UCI Machine Learning Repository at https://archive.ics.uci.edu/

Liver Disease Classification

- The fifth dataset, the Indian Liver Patient Dataset (ILPD), includes 583 instances and is obtained from the

Table 3: Descriptions of Independent Features from Six Datasets PIMA, RTML, Type-2 and ILPD Features (Table 1/2)

Attribute	Description	Type	Dataset
Pregnancies	Number of times pregnant	Integer	PIMA, RTML, Type-2
Glucose	Plasma glucose concentration at 2 hours in an oral glucose tolerance test (mg/dL)	Integer	PIMA, Type-2, RTML
BloodPressure	Diastolic blood pressure (mm Hg)	Integer	PIMA, CKD, RTML, Type-2
SkinThickness	Triceps skin fold thickness (mm)	Integer	PIMA, RTML, Type-2
Insulin	2-Hour serum insulin (μU/ml)	Integer	PIMA, Type-2, RTML
BMI	Body Mass Index (weight in kg/(height in m)²)	Decimal	PIMA, Type-2, RTML
Diabetes Pedigree Function	Diabetes pedigree function	Decimal	PIMA, Type-2
Age	Age (years)	Integer	PIMA, Type-2, RTML, ILPD, ESDRPD, CKD
Gender	Gender of the patient (Male/Female)	Categorical	ILPD, ESDRPD
Total Bilirubin	Total bilirubin level in blood (mg/dL)	Decimal	ILPD
Direct Bilirubin	Direct bilirubin level in blood (mg/dL)	Decimal	ILPD
Alkaline Phosphatase, Alanine Aminotransferase, Aspartate Aminotransferase	Alkaline phosphatase enzyme level in blood (IU/L); Alanine aminotransferase enzyme level in blood (IU/L); Aspartate aminotransferase enzyme level in blood (IU/L)	Integer	ILPD
Total Proteins	Total protein level in blood (g/dL)	Decimal	ILPD
Albumin	Albumin level in blood (g/dL)	Decimal	ILPD, CKD
Albumin and Globulin Ratio	Ratio of albumin to globulin in blood	Decimal	ILPD

Table 4: Descriptions of Independent Features from Six Datasets PIMA, RTML, Type-2 and ILPD Features (Table 2/2)

Attribute	Description	Type	Dataset
sg	Specific gravity (1.005-1.025)	Decimal	CKD
Polyuria, Polydipsia, Sudden Weight Loss, Weakness, Polyphagia, Genital Thrush, Visual Blurring, Itching, Irritability, Delayed Healing, Partial Paresis, Muscle Stiffness, Alopecia, Obesity	Indicates the presence or absence of each symptom (Yes = 1, No = 0)	Categorical	ESDRPD
su, bgr, bu, sod, pcv, wc	Sugar (0-5); Blood glucose random (mg/dL); Blood urea (mg/dL); Sodium (mEq/L); Packed cell volume; White blood cell count (cells/cumm)	Integer	CKD
sc, pot, hemo, rc	Serum creatinine (mg/dL); Potassium (mEq/L); Hemoglobin (gms); Red blood cell count (millions/cmm)	Decimal	CKD
Red Blood Cells, Pus Cell	Normal = 1, Abnormal = 0	Categorical	CKD
Pus Cell Clumps, Bacteria	Present = 1, Not Present = 0	Categorical	CKD
Hypertension, Diabetes Mellitus, Coronary Artery Disease, Pedal Edema, Anemia	Yes = 1, No = 0	Categorical	CKD
Appetite	Good = 0, Poor = 1	Categorical	CKD

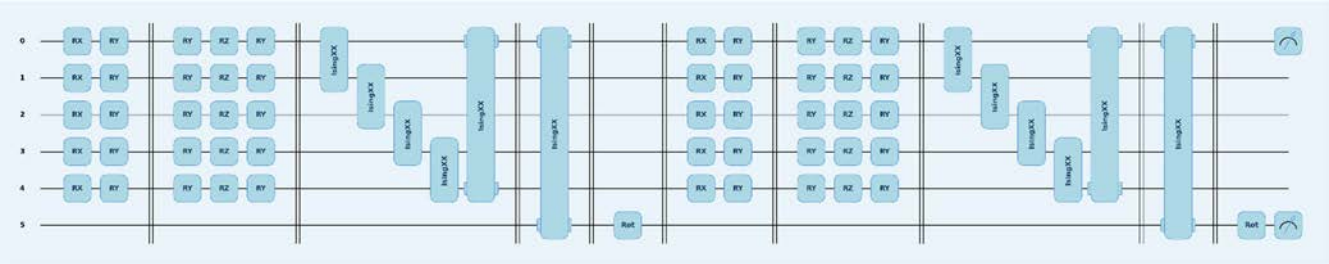

Figure 4: Architecture of the Proposed DQC Circuit with 5 PCA Components and 2 Variational Layers

Table 5: Experimental Results of the Proposed Trainable Quantum Circuit with Different PCA Components and Circuit Layers Across Multiple Diabetes Datasets. The best-performing configurations are highlighted in bold

Dataset	No. of Epochs (t)	PCA Components	No. of Layers (k)	No. of Qubits	Training Acc. (%)	Test Acc.(%)
PIMA	20	3	5	4	91.71	94.67
	20	4	3	5	92.29	96
Type-2	20	3	7	4	94.03	97.35
	20	4	5	5	90.59	88.76
RTML	10	3	5	4	95.55	95.83
	10	4	4	5	98.82	97.92
ILPD	20	3	9	4	89.95	90.24
	20	4	9	5	88.83	90.24
ESDRPD	10	3	3	4	98.24	98.64
	10	4	5	5	97.6	98.64
CKD	10	3	9	4	99.09	97.9
	10	4	6	5	96.42	93.06

UCI Machine Learning Repository at https://archive.ics.uci.edu/dataset/225/ilpd+indian+liver+patient+dataset

Chronic Kidney Disease Classification

- The sixth dataset, the Chronic Kidney Disease (CKD) Dataset, includes 400 instances and is obtained from the UCI Machine Learning Repository [18]

Experimental Results

The simulation experiments were conducted using the PennyLane framework with the default.qubit device as the quantum simulator backend. Table 5 presents the experimental outcomes of the proposed trainable quantum circuit, including hyperparameter tuning across multiple diabetes datasets. The tuned hyperparameters include the number of PCA components and the number of circuit layers, with the corresponding training and testing accuracies reported for each dataset. Various combinations of these parameters were systematically explored to identify optimal configurations and the best-performing settings are highlighted in bold. The proposed quantum model exhibits effective pattern learning across all six datasets, as reflected by the balanced training and testing accuracies, demonstrating strong generalization without overfitting.

The model achieves convergence within 10-20 epochs, indicating stable and efficient training behavior. The circuit architecture employs a two-register configuration: one register encodes the PCA-transformed input features, while the other consists of a single class qubit responsible for classification output. The total number of qubits required is computed as: Total Qubits = Number of PCA Components + 1 (class qubit).

From the hyperparameter tuning experiments, the optimal configurations and their corresponding results were determined for each dataset. The PIMA dataset, using 4 PCA components, 5 qubits and 3 layers, attained a test accuracy of 96%. The Type-II Diabetes dataset, with 3 PCA components, 4 qubits and 7 layers, achieved 97.35% accuracy. The RTML dataset, configured with 4 PCA components, 5 qubits and 4 layers, reached 97.92% accuracy. For the ILPD dataset, using 3 PCA components, 4 qubits and 9 layers, the model achieved 90.24% accuracy. Since the configuration with 4 PCA components required 5 qubits and yielded similar accuracy, the setup with fewer qubits was preferred for efficiency.

The ESDRPD dataset, with 3 PCA components, 4 qubits and 3 layers, achieved the highest test accuracy of 98.64%, while the CKD dataset, configured with 3 PCA components, 4 qubits and 9 layers, attained 97.90% accuracy. The model consistently converged within 10-20 epochs across all datasets. Notably, the proposed framework effectively learned feature patterns using only 4-5 qubits, demonstrating its resource efficiency and scalability for medical data classification.

Figure 4 illustrates the proposed Differentiable Hybrid Classical-Quantum Circuit with five PCA components and two variational layers, including the measurement process. This framework consists of all six proposed modules.

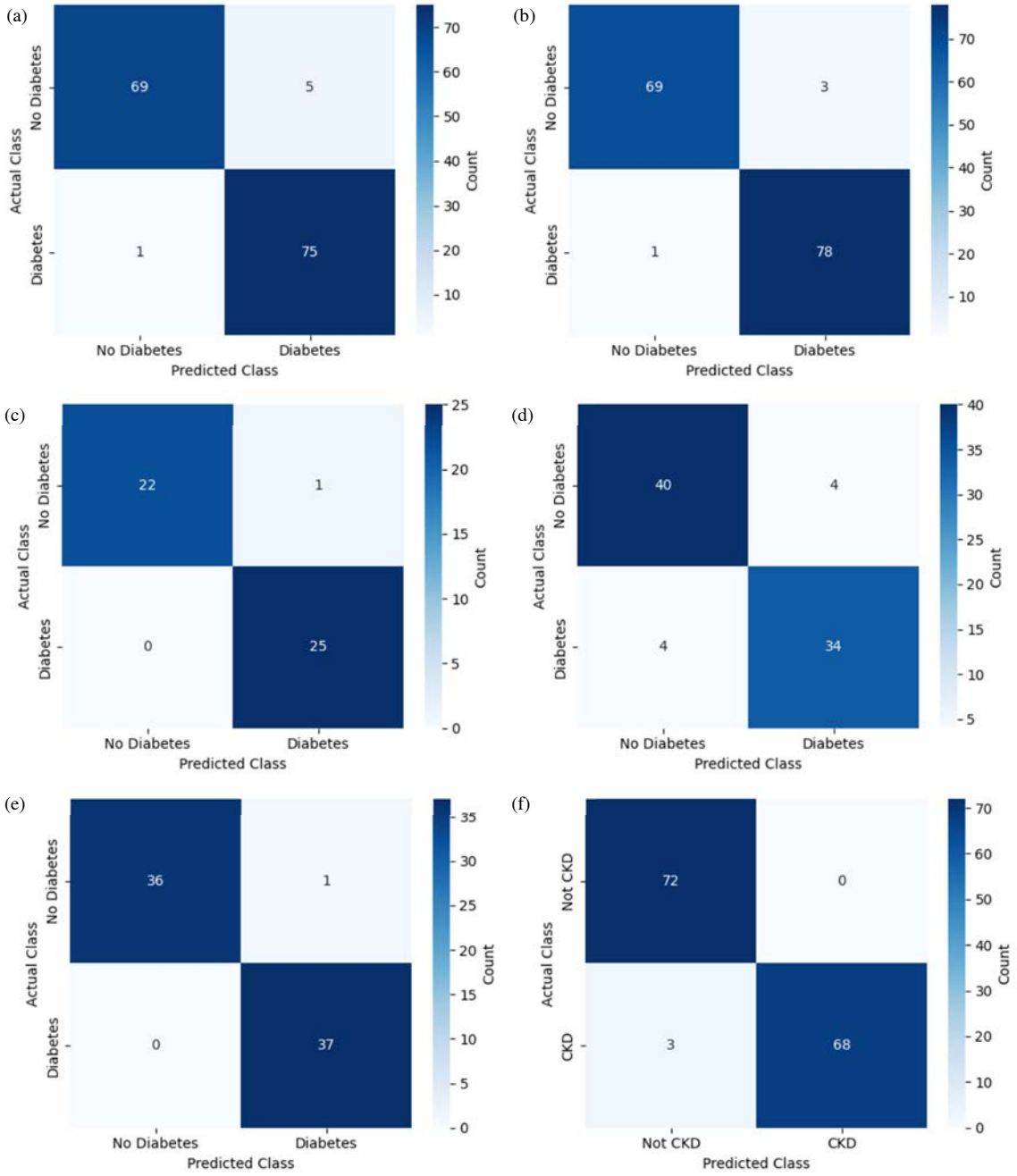

Figure 5(a-f): Confusion Matrices of the Proposed Quantum Classification Results across Six Diabetes Datasets

Comparative Methods and Discussions

The performance of the proposed quantum algorithm was evaluated using confusion matrices for all six diabetes datasets, as illustrated in Figure 5. These matrices represent the detailed classification outcomes of the quantum model across each dataset. Furthermore, comprehensive performance metrics were computed to quantitatively assess the classification process. These include Accuracy, Precision, Recall, F1-Score and the Area Under the Curve of the Receiver Operating Characteristic (AUC-ROC). Table 6 summarizes the performance metrics of the proposed Differentiable Quantum Circuit for diabetes classification. The F1-scores for all six datasets exceed 95%, except for the ILPD dataset, which achieved an accuracy of 90.24%. Notably, the AUC-ROC scores for all datasets are above 0.9, demonstrating the strong discriminative ability of the proposed model, with the ILPD dataset achieving an AUC-ROC of 0.9408.

For the classical baseline, Nine standard classifiers-SVM, LR, KNN, RF and Naive Bayes (NB), LDA, DT, Adaptive Boosting Technique (AdaB.) and Extreme Gradient Boosting (XGB.) were implemented using the classical preprocessing phase described in Figure 2. The corresponding Test Accuracy, Precision, Recall and F1 Score results for each classifier are summarized in Table 7.

Table 6: Classification Performance Metrics of Proposed Differentiable Quantum Circuit Across Various Diabetes Datasets

Dataset	Train Accuracy (%)	Test Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC-ROC
PIMA	97.85	96	96.13	96	96.15	0.9651
Type-2	98.5	97.35	97.38	97.35	97.5	0.9856
RTML	98.9	97.92	98	97.92	98.04	0.9826
ILPD	92.1	90.24	90.24	90.24	89.47	0.9408
ESDRPD	99.2	98.65	98.69	98.69	98.66	1
CKD	98.4	97.9	97.99	97.9	97.84	0.9924

Table 7: Experimental Results of Classical Classifiers with Preprocessing Phase Applied to Different Diabetes Datasets

Dataset	Metric	SVM	LR	KNN	RF	NB	LDA	DT	AdaB.	XGB.
PIMA	Accuracy	88.83	88.27	93.3	93.85	85.47	89.38	88.88	94.97	89.94
	Precision	86.54	87.13	93.75	92.93	88.04	90.52	91.3	93.93	94.94
	Recall	93.75	91.67	93.75	95.83	84.38	89.58	87.5	96.87	97.91
	F1 Score	90	89.34	93.75	94.36	86.17	90.05	89.36	95.38	96.41
Type-2	Accuracy	92.45	91.19	91.19	91.82	91.82	91.19	93.08	92.45	94.96
	Precision	92.39	92.22	88.78	91.4	94.25	95.23	90.72	90.62	93.61
	Recall	94.44	92.22	96.67	94.44	91.11	88.88	97.77	96.66	97.77
	F1 Score	93.41	92.22	92.55	92.9	92.66	91.95	94.11	93.54	95.65
RTML	Accuracy	96.15	96.15	88.46	96.15	96.15	96.15	94.23	94.23	96.15
	Precision	96.15	96.15	81.25	92.86	96.15	96.15	92.59	92.59	96.15
	Recall	96.15	96.15	100	100	96.15	96.15	96.15	96.15	96.15
	F1 Score	96.15	96.15	89.66	96.3	96.15	96.15	94.33	94.33	96.15
ILPD	Accuracy	87.22	86.47	81.2	84.96	85.71	81.95	85.71	87.21	88.72
	Precision	85.71	84.71	77.08	93.85	80.85	78.49	90.27	92.85	94.28
	Recall	93.51	93.51	96.1	79.22	98.7	94.8	84.41	84.41	85.71
	F1 Score	89.44	88.89	85.55	85.92	88.89	85.88	87.24	88.43	89.48
ESDRPD	Accuracy	85.53	76.32	73.68	86.84	85.53	73.68	75	78.95	80.26
	Precision	93.62	97.22	100	100	93.62	99.98	82.35	95	95.12
	Recall	84.62	67.31	61.54	80.77	84.62	61.54	80.77	73.08	75
	F1 Score	88.89	79.55	76.19	89.36	88.89	76.19	81.55	82.61	83.87
CKD	Accuracy	94.17	93.33	93.33	96.67	96.67	93.01	92.86	93.57	96.43
	Precision	100	100	100	100	100	99.98	92.65	99.98	97.01
	Recall	90.67	89.33	89.33	94.67	94.67	86.11	92.65	86.76	95.59
	F1 Score	95.1	94.37	94.37	97.26	97.26	92.54	92.65	92.91	96.3

Table 8: Summary of Key Literature Results on Various Diabetes Datasets

Reference	Dataset Used	Preprocess Techniques	Classical Classifier	Test Acc. (%)	Quantum Classifier	Test Acc. (%)
Islam *et al.* [5]	Chronic Kidney Disease	Applied	AdaBoost	98.3	Not Applied	None
			XGBoost	99.1		
			SVM	96.7		
			KNN	59		
			NB	88.3		
Donaire *et al.* [7]	Combined Pima and RTML	Applied	LR	75.6	Not Applied	None
			KNN	68.18		
			XGB	83.22		
			RF	80.68		
Safriandono *et al.* [10]	Indian Liver Patient (ILPD)	Applied	Without QFE		With QFE	
			SVM	72	QFE-SVM	75
			RF	69	QFE-RF	78
			XGB	69	QFE-XGB	81
Rubini *et al.* [18]	Indian Liver Ptient (ILPD)	Applied (SMOTE-ENN)	LR	79.72	Not Applied	None
			SVM	78.37		
			RF	87.83		
			KNN	91.89		
			XGB	90.54		
Feynman [8]	ILPD	Applied	-	-	QML-Liver+ SMOTE	75
		Not Applied	-	-	QML-Liver	83
Proposed Algorithm	PIMA	Applied	RF	93.85	Proposed Quantum Circuit	96
	Type 2		SVM	92.45		97.35
	RTML		SVM, LR, RF, NB	96.15		97.92
	ILPD		SVM	87.22		90.24
	ESDRPD		RF	86.84		98.65
	CKD		RF, NB	96.67		97.9

Some of the literature key performance metric comparisons between classical and existing quantum algorithms across the six diabetes datasets are presented in Table 8.

Using the proposed preprocessing phase, the classical standard classifiers achieved the following results: For the PIMA dataset, Adaptive Boosting Technique attained the

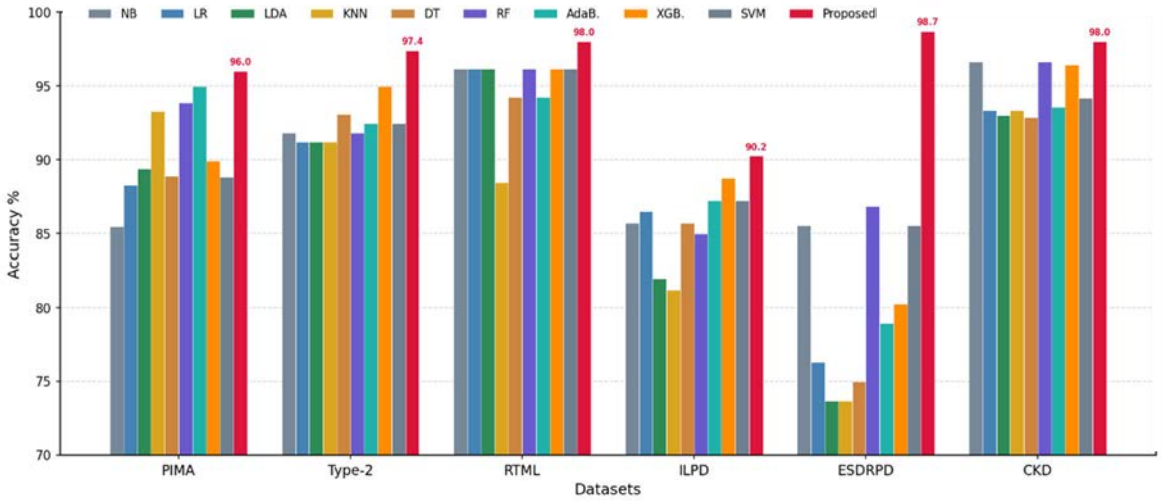

Figure 6: Classification Accuracy Between Classical Models and the Proposed Quantum Method

Table 9: Performance Comparison Between Best Classical and Proposed Quantum Classifier

Dataset	Category	Best Classical	Metric (%)	Score (%)		Epoch (Proposed)
				Classical	Proposed QML	
PIMA	Diabetes Classification	AdaBoost	Accuracy	94.97	96	20
			F1 Score	95.38	96.15	
Type-2	Diabetes Classification	XGBoost	Accuracy	94.96	97.35	20
			F1 Score	95.65	97.5	
RTML	Diabetes Classification	SVM	Accuracy	96.15	97.92	10
			F1 Score	96.15	98.04	
ESDRPD	Diabetes Classification	Random Forest	Accuracy	86.84	98.65	10
			F1 Score	89.36	98.66	
ILPD	Liver Disease Classification	XGBoost	Accuracy	88.72	90.24	20
			F1 Score	89.48	89.47	
CKD	Kidney Disease Classification	Random Forest	Accuracy	96.67	97.9	10
			F1 Score	97.26	97.84	

highest accuracy and F1 Score of 94.97 and 96.87%, respectively. In the case of the Type-2 Diabetes dataset, Extreme Gradient Boosting achieved an accuracy of 94.96% and an F1 Score of 95.65%.

For the RTML dataset, all classifiers performed competitively with an accuracy of around 96.15%, except KNN, which achieved a comparatively lower accuracy of 88.46%; however, in terms of F1 Score, RF achieved the highest value of 96.30%. For the ILPD dataset, Extreme Gradient Boosting recorded the best accuracy of 88.72% and an F1 Score of 89.48%. In the ESDRPD dataset, RF again outperformed other classifiers with 86.84% accuracy and 89.36% F1 Score. Finally, for the CKD dataset, both RF and Naive Bayes (NB) achieved the highest performance, with 96.67% accuracy and 97.26% F1 Score. The detailed description of results for all classifiers Table 9: Performance Comparison between best classical classifier baseline and the proposed quantum circuit across the datasets are summarized in Table 7.

The graphical representation comparing the proposed quantum circuit's accuracy with the preprocessed classical classifiers is shown in Figure 6. Similarly, the visual histogram representation comparing the proposed quantum circuit's F1 Score with the preprocessed classical classifiers is illustrated in Figure 7.

By referring to the statistical metrics and the clear visualizations in Figure 6 and 7, we conclude that the results highlight the effectiveness of the proposed Differentiable Quantum Circuit algorithm, which achieves the highest F1 Score with a minimal number of false positives and false negatives in diabetes classification. This architecture, incorporating data re-uploading, dynamic entanglement and separate target qubits, demonstrates strong capability in learning complex data distributions and generalizing to unseen data. Table 9 presents the performance comparison of Accuracy and F1 Score between the best-performing classical baseline classifier and the proposed Differentiable Hybrid Classical-Quantum Circuit across diabetes, liver disease and chronic kidney disease classification datasets, along with the number of epochs required for convergence by the proposed model. Bold values indicate the superior score in each metric.

Classical machine learning models' performance is heavily relied upon by the distribution and characteristics of features in underlying data of individual datasets, that have different performance in various medical classification tasks. In contrast, the proposed differentiable hybrid classical-quantum framework shows conceptually that Great performance stability and consistently best performance in various categories of datasets. This work was mainly geared towards diabetes classification, but some experiments are also done on a different application, which is experiments performed on liver disease and chronic kidney disease datasets further validates the robustness, adaptability and generalization ability of the proposed quantum framework.

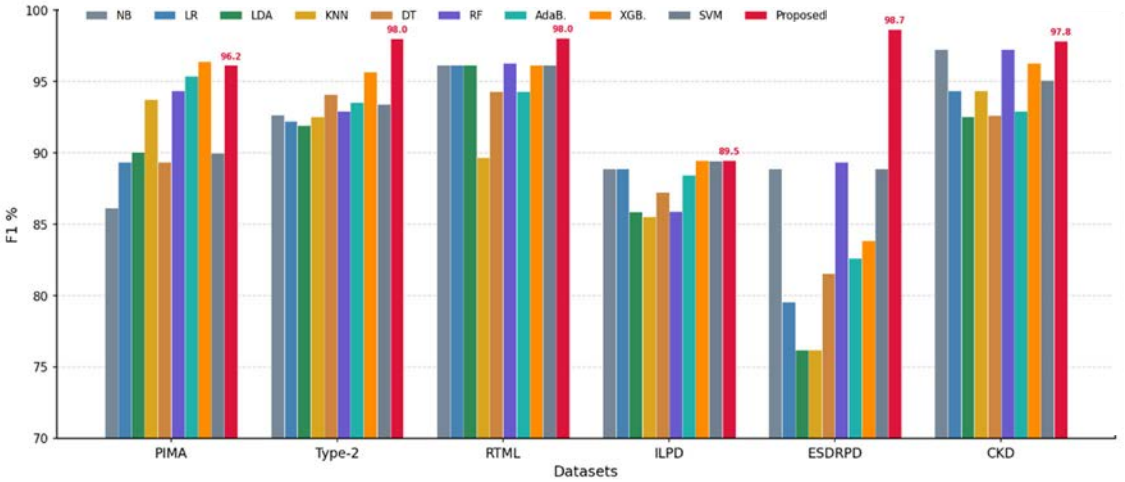

Figure 7: Classification F1 Score Between Classical Models and the Proposed Quantum Method

CONCLUSIONS

The proposed Differential Quantum Circuit (DQC) framework with the Classical PCA feature compression classifies diabetes disease with resampled dataset as shown in this study. For six diabetes datasets, the proposed model is used to find structural patterns in the features that are projected onto the PCA and a penalized binary cross-entropy cost function is used to prevent overfitting. Specifically, the model has demonstrated rapid convergence, converging to high accuracy within 10-20 training epochs, showcasing excellent generalization performance and learning of the underlying data distributions with few iterations. The high classification accuracy and F1-scores confirm the effectiveness of the proposed variational ansatz to improve the feature extraction and decrease the quantum circuit execution. This efficiency is the most beneficial when implementing on NISQ devices that have a limited number of qubits. In addition, the proposed model is always superior to the classical classifiers with preprocessed data on all assessment sets. Its good AUC-ROC scores also validate its strong discriminative power for distinguishing diabetic from non-diabetic cases. Putting all these results together suggests that the proposed quantum framework could be useful in the development of practical quantum machine learning real life applications.

Limitations

Although the proposed Differentiable Hybrid Classical-Quantum Circuit demonstrates outperformed multiple widely-used classical baseline classifiers, the current implementation is solely tested with the PennyLane simulator. A practical quantum device is a device that is practical for intermediate scale quantum technologies, which is known as the NISQ era. An open quantum system can be disturbed by environmental noise and decoherence because they are in a state of superposition, resulting in information loss [19]. If noise reduction and quantum error correction methods don't work, the claimed performance could suffer in quantum devices. So, the proposed framework is currently more reliable in the simulated closed quantum system containing error correction strategies.

Acknowledgement

The first author expresses gratitude to Vellore Institute of Technology, Vellore, for providing financial support that enabled this research work. The authors also thank the anonymous reviewers for their valuable comments and suggestions.

REFERENCES

[1] *World Health Organization*. "Diabetes." November 2024. *World Health Organization*, https://www.who.int/news-room/fact-sheets/detail/diabetes.

[2] *International Diabetes Federation. IDF Diabetes Atlas*. 11th ed., 2025, Brussels, Belgium.

[3] Sadhu, A. and A. Jadli. "Early-Stage Diabetes Risk Prediction: A Comparative Analysis of Classification Algorithms." *International Advanced Research Journal in Science, Engineering and Technology*, vol. 8, no. 2, 2021, pp. 193-201. https://doi.org/10.17148/IARJSET.2021.8228.

[4] Ihnaini, B. *et al.* "A Smart Healthcare Recommendation System for Multidisciplinary Diabetes Patients with Data Fusion Based on Deep Ensemble Learning." *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, 2021. https://doi.org/10.1155/2021/4243700.

[5] Islam, M.A. *et al.* "Chronic Kidney Disease Prediction Based on Machine Learning Algorithms." *Journal of Pathology Informatics*, vol. 14, 2023. https://doi.org/10.1016/j.jpi.2023.100189.

[6] Bardia, V. and E. Sophiya. "Diabetes Prediction Using Machine Learning Algorithm: A Comparative Analysis." *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 1, March 2024, pp. 1973-1979. IEEE, https://doi.org/10.1109/ICACCS60874.2024.10717264.

[7] Donaire, L.M. *et al.* "A Hybrid Quantum-Classical Approach for Liver Disease Detection Using Quantum Machine Learning." *Engineering Applications of Artificial Intelligence*, 2025. https://doi.org/10.1016/j.engappai.2025.113240.

[8] Feynman, R.P. "Simulating Physics with Computers." *Feynman and Computation*, CRC Press, 2018, pp. 133-153. https://doi.org/10.1201/9780429500459-11.

[9] Hancco-Quispe, J.K. *et al.* "Quantum Machine Learning Applied to the Classification of Diabetes." *arXiv*, 2022. arXiv:2301.00109.

[10] Safriandono, A.N. *et al.* "Analyzing Quantum Feature Engineering and Balancing Strategies Effect on Liver Disease Classification." *Journal of Future Artificial Intelligence and Technologies*, vol. 1, no. 1, 2024, pp. 51-63. https://doi.org/10.62411/faith.2024-12.

[11] Changpetch, P. *et al.* "Integrating Data Mining Techniques for Naïve Bayes Classification: Applications to Medical Datasets." *Computation*, vol. 9, no. 9, 2021. https://doi.org/10.3390/computation9090099.

[12] Phongying, M. and S. Hiriote. "Diabetes Classification Using Machine Learning Techniques." *Computation*, vol. 11, no. 5, 2023. https://doi.org/10.3390/computation11050096.

[13] Maćkiewicz, A. and W. Ratajczak. "Principal Components Analysis (PCA)." *Computers & Geosciences*, vol. 19, no. 3, 1993, pp. 303-342. https://doi.org/10.1016/0098-3004(93)90090-R.

[14] Batista, G.E. *et al.* "A Study of the Behavior of Several Methods for Balancing Machine Learning Training Data." *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, 2004, pp. 20-29. https://doi.org/10.1145/1007730.1007735.

[15] Pérez-Salinas, A. *et al.* "Data Re-Uploading for a Universal Quantum Classifier." *Quantum*, vol. 4, 2020. https://doi.org/10.22331/q-2020-02-06-226.

[16] Boyle, A.O.D. and R. Nikandish. "A Hybrid Quantum-Classical Generative Adversarial Network for Near-Term Quantum Processors." *IEEE Access*, 2024. https://doi.org/10.1109/ACCESS.2024.3433383.

[17] Rani, R. *et al.* "Enhancing Liver Disease Diagnosis with Hybrid SMOTE-ENN Balanced Machine Learning Models-An Empirical Analysis of Indian Patient Liver Disease Datasets." *Frontiers in Medicine*, vol. 12, 2025. https://doi.org/10.3389/fmed.2025.1502749.

[18] Rubini, L. *et al. Chronic Kidney Disease* [Dataset]. UCI Machine Learning Repository, 2015. https://doi.org/10.24432/C5G020.

[19] Preskill, J. "Quantum Computing in the NISQ Era and Beyond." *Quantum*, vol. 2, 2018. https://doi.org/10.22331/q-2018-08-06-79.