Hybridized Quantum Computing and Federated Learning for Multi-Disease Prediction

Anbarasi M. S.[1]; Aaron Pushparaj M. M.[2]; Boomika S.[3]; Gowtham[4]

[1;2;3;4]Department of Information Technology, Puducherry Technological University, Puducherry, India

Publication Date: 2026/05/09

Abstract: The growing occurrence of patients with multiple chronic diseases has increased the need for efficient multi-disease risk prediction systems in healthcare. However, clinical data is often distributed in various cloud, but it lacks in efficient privacy preservation. To address these challenges, this project proposes a hybridized framework for privacy-aware multi-disease risk prediction in distributed healthcare environments. nA key challenge in multi-disease data collection from heterogeneous clinical data may lag in normalization. So, cleaning and preprocessing is done on the first phase. Second challenge lies in capturing complex and non-linear relationships among clinical features within reduced feature spaces. To handle this, quantum-inspired feature encoding and variational quantum circuits are employed to enhance feature representations. Since data is distributed across cloud data privacy and confidentiality issues exist. To overcome this issue Federated Learning for multi-disease prediction is applied in the third phase. To address these challenges, this paper proposes a novel framework titled Hybridized Quantum Computing and Federated Learning for Multi-Disease Prediction (HQCF-MDRP). The proposed system integrates structured data preprocessing, quantum-assisted feature representation, and federated collaborative learning into a unified architecture. Initially, healthcare datasets are cleaned, normalized, and subjected to feature selection to improve data quality and reduce dimensionality. Selected features are then transformed using quantum-inspired angle encoding and Variational Quantum Circuits (VQC), enabling enhanced representation of hidden correlations among clinical attributes. Subsequently, Federated Learning with FedProx optimization is employed to train predictive models collaboratively across distributed healthcare institutions without sharing raw patient data. This ensures privacy preservation while improving model generalization over heterogeneous datasets. Experimental evaluation on multiple disease datasets demonstrates high predictive performance, achieving global accuracy of 98.38%, F1-score of 97.60%, and baseline model accuracy up to 99.9%. The proposed framework offers a scalable, secure, and efficient solution for next-generation intelligent healthcare analytics and distributed multi-disease risk prediction.

Keywords: *Multi-Disease Prediction, Federated Learning, Quantum Computing, Variational Quantum Circuit, Privacy-Preserving Healthcare, Machine Learning, Feature Selection, Distributed Learning, Clinical Data Analytics, Fedprox Optimization.*

How to Cite: Anbarasi M. S.; Aaron Pushparaj M. M.; Boomika S.; Gowtham (2026) Hybridized Quantum Computing and Federated Learning for Multi-Disease Prediction. *International Journal of Innovative Science and Research Technology*, 11(4), 3914-3920. https://doi.org/10.38124/ijisrt/26apr1596

I. INTRODUCTION

The proposed system is designed as a modular and intelligent framework for privacy-preserving multi-disease prediction using distributed healthcare data. It begins with data collection, where multiple clinical datasets from independent healthcare sources are provided as structured inputs. These datasets are then subjected to preprocessing operations such as cleaning, normalization, missing value handling, and validation to ensure reliable and consistent input for predictive analysis.

The system then performs feature selection using machine learning-based evaluation methods to identify the most relevant clinical attributes. This reduces dimensionality, removes redundant information, and improves overall learning efficiency. After feature selection, a quantum-assisted feature representation layer is applied, where classical healthcare features are encoded into higher-dimensional spaces using angle encoding and Variational Quantum Circuits. This enables the extraction of complex and non-linear relationships that may not be captured effectively through conventional techniques.

Once transformed, the system carries out privacy-aware collaborative model training using Federated Learning. Local models are trained independently at multiple healthcare nodes, while only model parameters are shared with a central server for aggregation. FedProx optimization is incorporated to handle data heterogeneity across institutions and improve convergence stability during global model training.

Additionally, the system includes evaluation, monitoring, and result analysis mechanisms, which provide predictive accuracy, F1-score, confusion matrices, and comparative performance insights. An integrated logging and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr1596

model management layer ensures systematic tracking of training rounds and prediction outcomes.

Overall, the proposed system transforms traditional healthcare prediction methods into a secure, scalable, and intelligent multi-disease analytics framework, addressing key limitations such as privacy risks, isolated learning, poor feature representation, and lack of collaborative intelligence.

II. RELATED WORKS

The literature survey focuses on major components of intelligent healthcare prediction systems, including clinical data preprocessing, feature selection, quantum-assisted learning, federated model training, and privacy-preserving multi-disease prediction.

Machine learning-based disease prediction has gained significant attention in recent years. Hidayaturrohman et al. [1] proposed an XGBoost-based framework for healthcare prediction using preprocessing techniques such as imputation, encoding, and standardization. Their work improved prediction accuracy, but it mainly focused on single-disease analysis rather than unified multi-disease prediction. Similarly, Zhou et al. [2] developed an XGBoost-based predictive model for hospital outpatient volume forecasting. Although the model achieved strong performance, it focused on demand estimation rather than disease risk prediction. Zheng et al. [3] proposed an XGBoost-based clinical prediction model using multicenter datasets. While effective, the approach relied on centralized learning and lacked privacy-preserving collaboration.

To improve feature-level understanding, quantum-assisted learning techniques are increasingly explored. Angelino et al. [4] proposed a Variational Quantum Classifier for predictive analytics, demonstrating that quantum models can achieve competitive classification performance. However, the study was limited to industrial datasets and did not address healthcare-specific challenges. Shirin et al. [5] reviewed quantum computing applications in healthcare and highlighted the potential of quantum machine learning for diagnostics and medical imaging. However, their work remained conceptual without practical deployment. In another study, Ovalle-Magallanes et al. [6] introduced angle encoding with learnable rotations for quantum-classical neural networks. Although the method improved feature representation, it was primarily evaluated on image classification tasks rather than structured clinical datasets.

In the area of privacy-preserving distributed learning, several studies focus on Federated Learning. Shahnazeer et al. [7] proposed a federated transfer learning framework for multi-disease prediction. The system improved collaborative healthcare analytics but did not include advanced feature enhancement methods. Wang et al. [8] developed a secure federated logistic regression framework using encryption techniques for privacy protection. While effective in preserving confidentiality, the model increased computational overhead. Li et al. [9] proposed the FedProx optimization algorithm to address data heterogeneity in federated environments. Their method improved convergence stability, but it was not specifically designed for multi-disease healthcare systems.

Hospital and healthcare analytics systems also employ traditional centralized machine learning models such as Random Forest, Logistic Regression, and Neural Networks. These approaches often achieve good predictive accuracy, but they require centralized data storage, which raises concerns related to privacy, ownership, and regulatory compliance.

Overall, existing works address preprocessing, quantum learning, or federated collaboration individually. However, they lack a unified framework that integrates structured preprocessing, quantum-assisted feature transformation, and privacy-preserving distributed model training for multi-disease prediction. The proposed system addresses these limitations by combining feature selection, quantum-enhanced learning, and federated optimization into a single intelligent healthcare pipeline.

III. PROPOSED WORK

Fig. 1 shows the architecture of the proposed system. The proposed system focuses on the design and implementation of an intelligent privacy-preserving multi-disease prediction system using distributed clinical datasets. It is developed as a complete pipeline that connects data preprocessing, quantum-assisted feature transformation, federated model collaboration, and predictive analysis into a single framework. The system converts raw healthcare records into meaningful diagnostic insights for accurate and secure disease prediction.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr1596

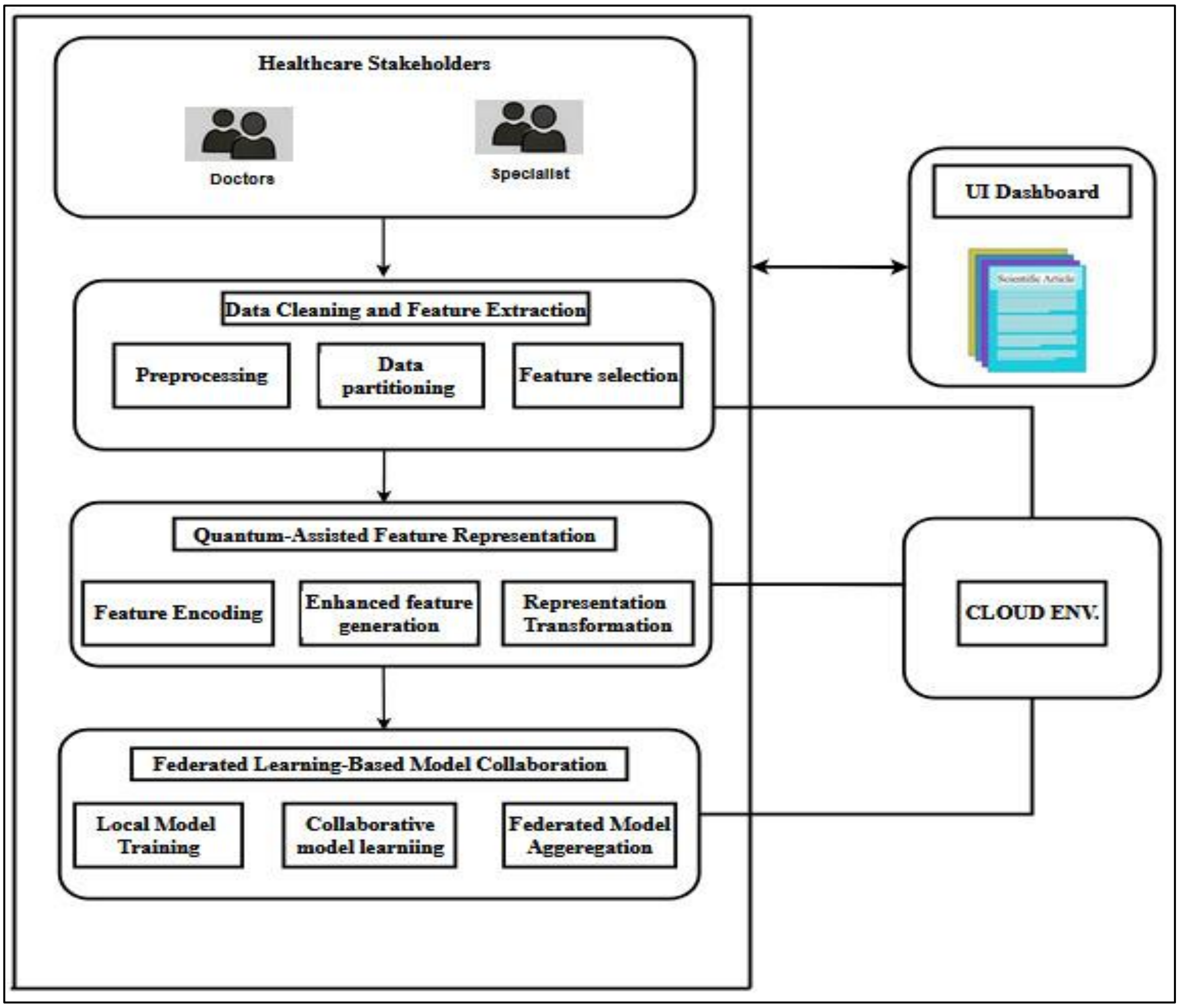

Fig 1 Architecture Diagram of the Proposed Model

The system is divided into multiple modules, where each module performs a specific stage in the overall workflow. This modular structure improves scalability, maintainability, and processing efficiency.

A. Module – I: Data Preprocessing and Feature Selection

This module, represented in Fig. 2, is responsible for collecting clinical datasets, preprocessing raw healthcare records, and selecting relevant features for predictive analysis. It converts raw medical input into structured and optimized datasets for further processing.

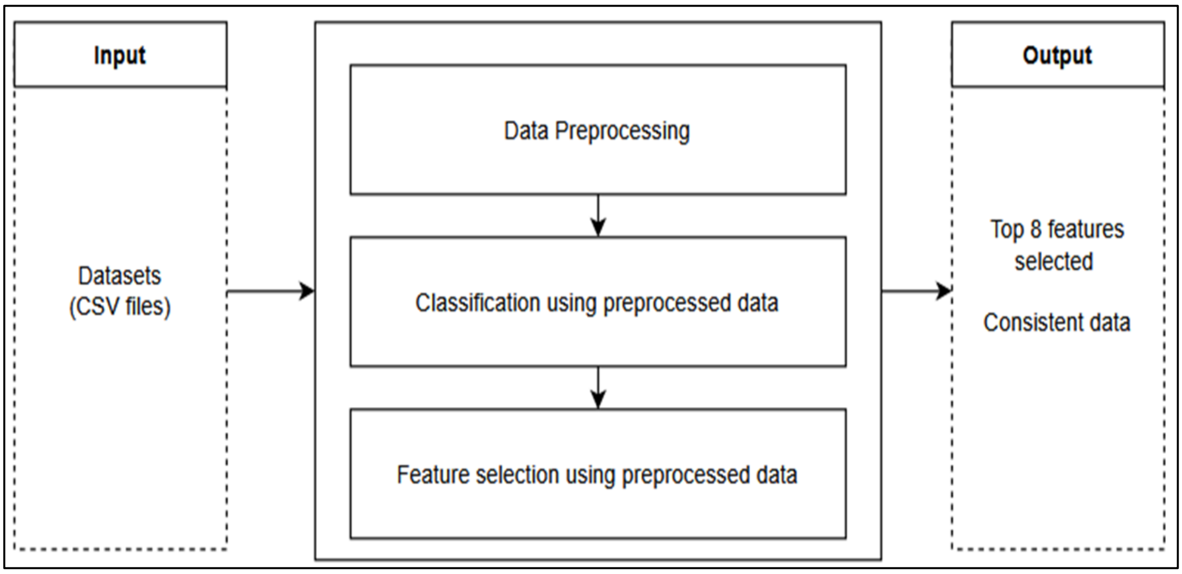

Fig 2 Data Preprocessing and Feature Selection

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr1596

➢ *Dataset Collection and Input Loading*

The system accepts multiple clinical datasets related to different diseases from distributed healthcare sources. These datasets are provided in structured formats such as CSV files containing patient records, medical attributes, and diagnosis labels.

➢ *Data Cleaning and Validation*

The loaded datasets undergo cleaning operations to remove duplicate records, inconsistent values, and incomplete entries. Healthcare-specific validation rules are applied to eliminate physiologically invalid data and improve dataset reliability.

➢ *Missing Value Handling*

Missing values present in numerical and categorical attributes are handled using suitable imputation techniques such as median or mode replacement. This ensures completeness of data before model training.

➢ *Data Transformation and Normalization*

Clinical attributes are transformed into machine-readable form using categorical encoding methods. Numerical features are normalized into a standard range to improve training stability and model efficiency.

➢ *Feature Evaluation Using XGBoost*

The system applies XGBoost-based evaluation to measure the predictive importance of each clinical attribute. Important features contributing strongly to disease prediction are ranked accordingly.

➢ *WEKA ClassifierAttributeEval*

To further refine feature relevance, WEKA ClassifierAttributeEval is employed to independently assess each attribute using classification performance. This improves the quality of selected inputs.

➢ *Final Feature Selection*

Based on ranking and evaluation scores, the most informative attributes are selected while redundant and less useful features are removed. This reduces dimensionality, lowers computational complexity, and enhances prediction accuracy for the next module.

B. Module– II: Quantum-Assisted Feature Representation

This module, represented in Fig. 3, is responsible for transforming selected clinical features into enhanced representations using quantum-inspired techniques. It converts classical healthcare attributes into richer feature spaces for improved predictive learning.

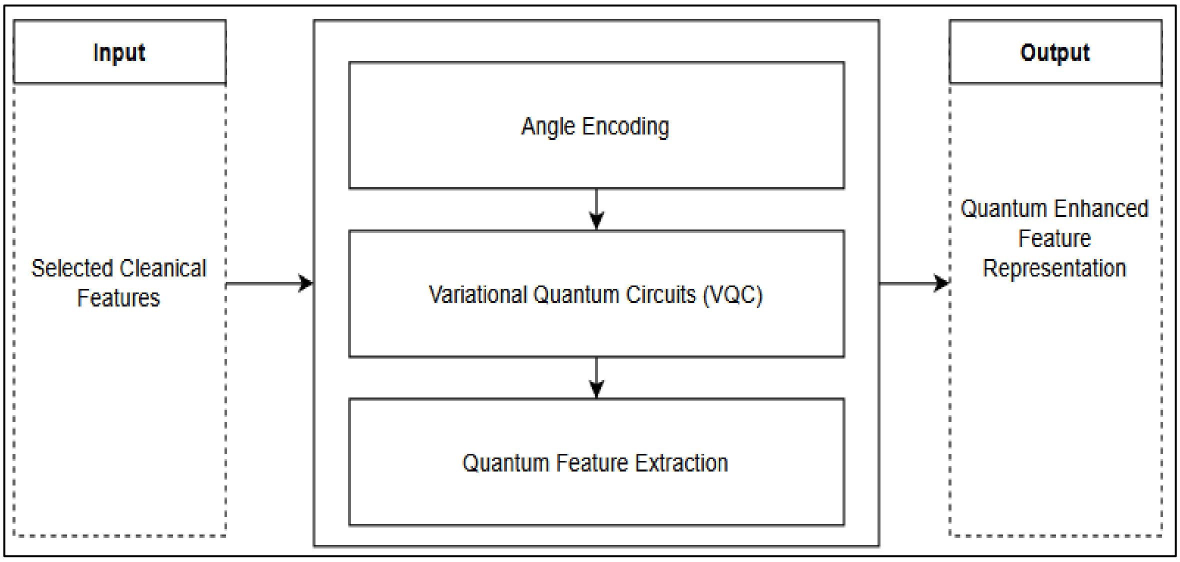

Fig 3 Quantum-Assisted Feature Representation

➢ *Input Feature Acquisition*

The selected features obtained from Module I are provided as inputs to this stage. Only the most relevant clinical attributes are considered for quantum transformation.

➢ *Feature Normalization*

All input features are normalized into a fixed numerical range to ensure compatibility with quantum encoding operations. This improves consistency during circuit execution.

➢ *Angle Encoding*

The normalized classical features are mapped into quantum states using angle encoding. Each feature value is converted into a rotation angle and assigned to a corresponding qubit.

➢ *Variational Quantum Circuit (VQC)*

The encoded inputs are processed through a Variational Quantum Circuit consisting of parameterized rotation gates and entanglement layers. This enables learning of complex and non-linear relationships among features.

➢ *Entanglement and Feature Interaction*

Quantum entanglement operations are applied between qubits to capture hidden dependencies and interactions among clinical attributes that are difficult to model classically.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr1596

➢ *Measurement Operation*

After circuit processing, measurement operations are performed to convert quantum states back into classical numerical outputs. These values represent transformed feature information.

➢ *Quantum Feature Generation*

The measured outputs are stored as quantum-enhanced features and forwarded to the next module. These transformed representations improve learning capability and predictive performance during collaborative model training.

C. Module– III: Federated Learning Based Model Collaboration

This module, represented in Fig. 4, is responsible for collaborative model training across multiple healthcare institutions while preserving patient privacy. It converts distributed local learning into a unified global prediction model without sharing raw clinical data.

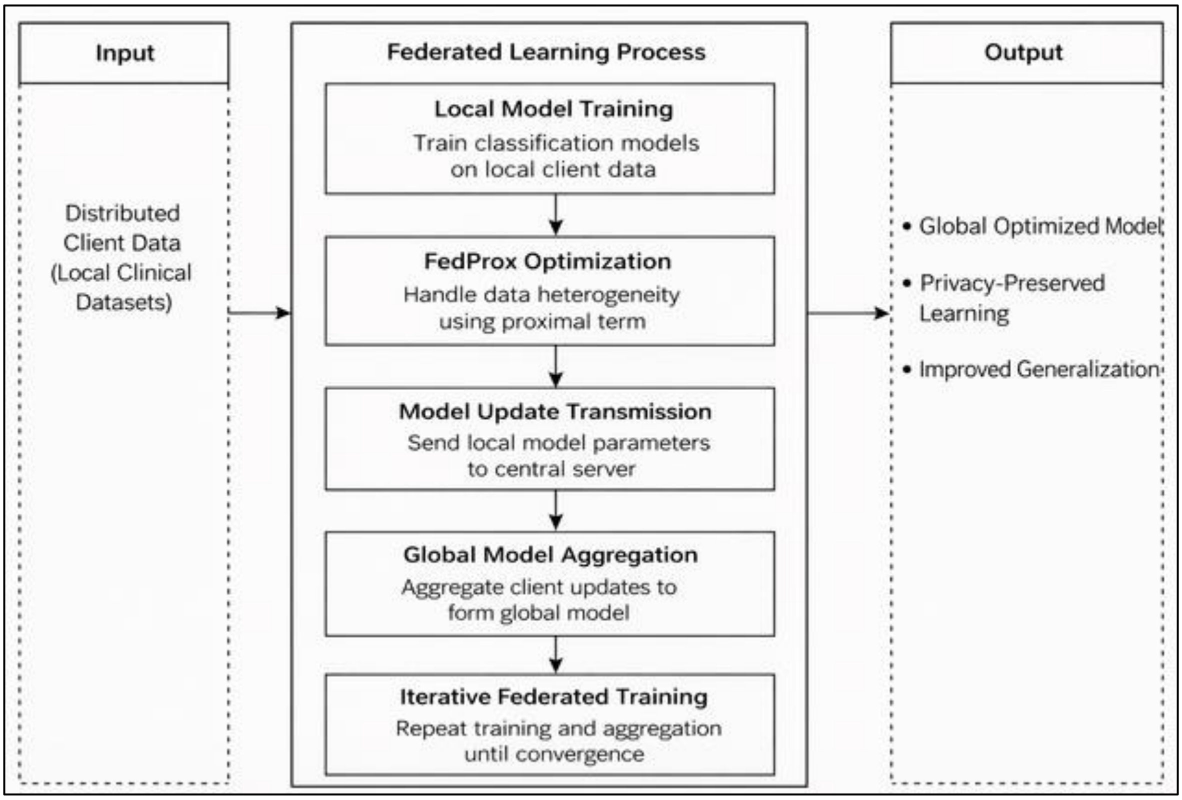

Fig 4 Federated Learning Model Collaboration

➢ *Client Initialization*

The system consists of multiple healthcare clients, where each client holds its own local clinical dataset. These datasets remain stored locally and are not transferred to external servers.

➢ *Local Model Training*

Each client independently trains a Logistic Regression model using its quantum-enhanced features received from the previous module. Local training enables learning from private healthcare data securely.

➢ *Model Parameter Update*

After local training, only model parameters such as weights and gradients are generated for transmission. Raw patient records are never shared, ensuring confidentiality.

➢ *FedProx Optimization*

FedProx optimization is applied during local learning to reduce divergence between local and global models. This improves convergence stability when data distributions differ across healthcare institutions.

➢ *Global Model Aggregation*

The central server collects model updates from all participating clients and performs weighted aggregation to create a global prediction model. This combines collective learning knowledge from distributed sources.

➢ *Iterative Federated Rounds*

The updated global model is redistributed to all clients for the next training round. This process is repeated iteratively until the system reaches stable and optimal performance.

➢ *Final Multi-Disease Prediction*

After convergence, the final global model is used for multi-disease risk prediction on unseen patient data. The system provides accurate, scalable, and privacy-preserving healthcare predictions.

IV. RESULTS ANALYSIS

The proposed system is evaluated based on prediction accuracy, privacy preservation, scalability, and distributed

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr1596

learning efficiency. The evaluation is carried out using multiple healthcare datasets under classical machine learning, quantum-assisted learning, and federated learning environments.

The comparative analysis shows that conventional centralized machine learning approaches provide good predictive performance but suffer from privacy risks and limited scalability. Traditional methods such as Logistic Regression and Random Forest require all healthcare data to be stored in a single repository, which is not suitable for distributed clinical environments.

The proposed HQCF-MDRP framework overcomes these limitations by integrating structured preprocessing, quantum-assisted feature transformation, and federated collaborative learning. Experimental results indicate that the global federated model achieved an accuracy of 98.38% with strong convergence across training rounds. In addition, the baseline Multi-Layer Perceptron model achieved 99.9% accuracy and F1-score, demonstrating highly reliable prediction capability.

The feature selection module successfully reduced irrelevant attributes and improved learning efficiency. Quantum-assisted feature representation enhanced the model's ability to capture non-linear relationships among clinical features. Federated Learning with FedProx optimization effectively handled heterogeneous client datasets and maintained stable convergence without sharing raw patient data.

The system also demonstrated high scalability by supporting prediction across four disease datasets in a unified framework. Since PennyLane simulation was used for quantum processing, no physical quantum hardware was required.

Standard evaluation metrics such as Accuracy, Precision, Recall, and F1-score confirm the strong performance of the proposed framework. These results validate that the system is capable of delivering accurate, privacy-preserving, and scalable multi-disease prediction in real-world healthcare applications.

Overall, the proposed system demonstrates reliable, scalable, and accurate multi-disease prediction performance. The integration of preprocessing, quantum-assisted learning, and federated collaboration significantly improves prediction capability while ensuring privacy and distributed intelligence.

V. CONCLUSION AND FUTURE ENHANCEMENTS

The proposed HQCF-MDRP framework presents an effective solution for privacy-preserving multi-disease prediction using distributed healthcare datasets. By integrating data preprocessing, feature selection, quantum-assisted feature representation, and federated learning, the system enables accurate disease prediction while maintaining patient data confidentiality. Unlike conventional centralized approaches, the proposed framework allows collaborative model training without transferring raw clinical records.

A key contribution of the system is the use of quantum-inspired feature transformation, which enhances the representation of complex and non-linear relationships among clinical attributes. In addition, the federated learning module with FedProx optimization improves convergence stability and supports learning across heterogeneous healthcare institutions. The modular design of the framework also improves scalability, flexibility, and suitability for real-world healthcare environments.

Experimental evaluation demonstrates that the proposed system achieves high predictive performance, strong privacy preservation, and efficient distributed collaboration. The framework therefore provides a reliable foundation for next-generation intelligent healthcare analytics and secure clinical decision support systems.

Future enhancements can focus on extending the framework to handle dynamically changing healthcare data through adaptive feature selection strategies. More advanced quantum-inspired or real quantum models may be integrated to further improve learning capability as computational resources evolve. Alternative federated aggregation techniques can also be explored to enhance robustness under highly non-IID data conditions.

In addition, explainable artificial intelligence methods can be incorporated to improve transparency and trust in clinical predictions. Validation using larger real-world hospital datasets and cloud-based deployment can further strengthen the practical applicability and scalability of the proposed framework.

REFERENCES

[1]. M. Kabir, M. Kaosar, and F. Sohel, "QTopic: A novel quantum perspective on learning topics from text," Neurocomputing, vol. 669, Art. no. 132483, 2026, doi: 10.1016/j.neucom.2025.132483.

[2]. P. Kottapalle, T. K. Tak, P. R. Kshirsagar, G. Ginnela, and V. K. Akula, "QHF-CS: Quantum-enhanced heart failure prediction using quantum CNN with optimized feature qubit selection with cuckoo search in skewed clinical data," Computers, Materials & Continua, vol. 84, no. 2, pp. 3857–3878, 2025, doi: 10.32604/cmc.2025.065287.

[3]. N. Moneesha, M. Sa, D. G. Naira, and J. J. Naira, "FedHybrid: Unifying aggregation strategies to optimize federated learning on non-IID dataset," in Proc. Int. Conf. Machine Learning and Data Engineering, Procedia Computer Science, vol. 258, pp. 3126–3134, 2025, doi: 10.1016/j.procs.2025.04.570.

[4]. S. C. K. Shahnazeer and G. Sureshkumar, "Federated transfer learning framework for multi-disease prediction," Procedia Computer Science, vol. 258, pp. 830–838, 2025, doi: 10.1016/j.procs.2025.04.315.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr1596

[5]. "Quantum Computing and Machine Learning in Medical Decision-Making: A Comprehensive Review," Algorithms, vol. 18, no. 3, Art. no. 156, 2025, doi: 10.3390/a18030156.
[6]. E. A. Radhi, M. Y. Kamil, and M. A. Alshujeary, "Quantum Machine and Deep Learning for Medical Image Classification: A Systematic Review," Iraqi Journal for Computer Science and Mathematics, vol. 6, pp. 107-138, 2025.
[7]. A. Wijesekara et al., "A systematic review of quantum machine learning for digital health," npj Digital Medicine, vol. 8, Art. no. 237, 2025, doi: 10.1038/s41746-025-01597-z.
[8]. "Quantum Convolutional Neural Network for Skin Cancer Classification with Federated Learning and Explainable AI," International Journal of Applied Mathematics, vol. 38, no. 6s, 2025.
[9]. Hidayaturrohman, Q. A., & Hanada, E. (2024). Impact of data pre-processing techniques on XGBoost model performance for predicting all-cause readmission and mortality among patients with heart failure. BioMedInformatics, 4, 2201–2212. https://doi.org/10.3390/biomedinformatics4040118
[10]. Zhou, L., Zhu, Q., Chen, Q., & Wang, P. (2025). Predicting hospital outpatient volume using XGBoost: a machine learning approach. Scientific Reports, 15, 17028.
[11]. Zheng, J., Li, J., Zhang, Z., Yu, Y., Tan, J., Liu, Y., Gong, J., Wang, T., Wu, X., & Guo, Z. (2023). Clinical data-based XGBoost algorithm for infection risk prediction of patients with decompensated cirrhosis: a multicenter retrospective study. BMC Gastroenterology, 23, 310

AUTHOR'S BIOGRAPHY

AARON PUSHPARAJ M M; A B.Tech Information Technology student at Puducherry Technological University with a strong interest in full-stack application development. Skilled in Python, C, and C++, along with frameworks and technologies such as Flask, React, and SQL, with a focus on designing scalable software systems and structured architectures. Focused on developing efficient applications that combine robust backend processing with smooth and interactive user interfaces.

BOOMIKA S; A B.Tech Information Technology student with a strong interest in software development and Artificial Intelligence/Machine Learning. Proficient in technologies such as React, Flask, Node.js, and MySQL, with a focus on building scalable and efficient web applications. Interested in developing intelligent systems and applying modern computing techniques to address real-world challenges while continuously strengthening technical knowledge.

GOWTHAM M; A B.Tech Information Technology student with a keen interest in web development and content-oriented digital solutions. Skilled in technologies including Python, React.js, Node.js, and MySQL, with a focus on designing scalable and user-friendly applications. Passionate about creating engaging digital platforms and applying modern tools to build practical and effective solutions.

Anbarasi M S; Professor in the Department of Information Technology, Puducherry Technological University. B.E., M.E., Ph.D., M.I.E.E.E., with specialization in Data Mining for Healthcare Predictions, Big Data Analytics, Cloud Computing, and Software Engineering. Actively involved in teaching and research in the areas of data analytics, cloud computing, and secure computing technologies.