International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

A Comparative Study of Classical and Variational Quantum Machine Learning Models for Breast Cancer Classification

Divya Guthi

Technical Associate, Exafluence Education, Sri Venkateswara University, Tirupati, India
Email: *divya19910122[at]gmail.com*

Abstract: ***Early and accurate diagnosis of breast cancer plays a critical role in improving patient survival rates and reducing treatment costs. Classical machine learning algorithms such as Logistic Regression and Support Vector Machines (SVM) have demonstrated strong performance in medical diagnosis tasks; however, their scalability and representational limits motivate the exploration of emerging paradigms. Quantum Machine Learning (QML), particularly Variational Quantum Classifiers (VQC), has recently gained attention due to its potential to model complex decision boundaries using quantum circuits. This paper presents a comparative analysis between a Variational Quantum Classifier and classical machine learning models for binary breast cancer classification. The study employs dimensionality-reduced clinical data and evaluates model performance using accuracy, confusion matrices, decision boundary visualization, Principal Component Analysis (PCA), and Receiver Operating Characteristic (ROC) curves. Experimental results indicate that the VQC achieves competitive performance with classical approaches, demonstrating its feasibility for small-scale medical datasets and highlighting the promise of quantum-enhanced learning in healthcare applications.***

Keywords: Quantum Machine Learning, Variational Quantum Classifier, Breast Cancer Classification, Support Vector Machine, Logistic Regression, PCA, ROC Curve.

1. Introduction

1.1 Quantum–Enhanced Learning for Clinical Decision Support

Early and accurate diagnosis of breast cancer remains a critical challenge in clinical decision support systems. The increasing availability of high-dimensional medical datasets has accelerated the adoption of machine learning techniques for automated disease classification. However, conventional learning models often struggle to capture complex feature interactions when data dimensionality is reduced, or sample sizes are limited.

1.2 Dimensionality Reduction and Feature Representation in Medical Data

Medical datasets typically contain correlated and redundant attributes, which can degrade model performance. Principal Component Analysis (PCA) is widely employed to transform high-dimensional clinical features into compact, informative representations [10]. In this work, PCA is used to reduce the breast cancer dataset to two principal components, enabling both computational efficiency and visualization of decision boundaries.

1.3 Limitations of Classical Classifiers in Non-Linear Decision Spaces

Classical classifiers such as Logistic Regression [12] and Support Vector Machines [11] have demonstrated strong performance in binary medical classification tasks. Nevertheless, Logistic Regression assumes linear separability, while kernel-based SVMs incur increasing computational cost and model complexity. These limitations motivate exploration of alternative learning paradigms capable of modeling non-linear relationships more efficiently.

1.4 Variational Quantum Classifiers as Hybrid Learning Models

Variational Quantum Classifiers (VQCs) represent a hybrid quantum–classical approach, combining parameterized quantum circuits with classical optimization. By encoding classical data into quantum states and optimizing circuit parameters, VQCs can implicitly map data into higher-dimensional Hilbert spaces, potentially improving class separability even with reduced feature sets [3,15].

1.5 Scope and Contributions of the Present Study

This study presents the design, implementation, and evaluation of a Variational Quantum Classifier for breast cancer prediction using PCA-reduced features. The proposed model is experimentally compared with Logistic Regression and Support Vector Machine classifiers under a unified pipeline. Additionally, a web-based interface is developed to demonstrate real-time prediction using the trained quantum model.

2. Literature Review

2.1 Classical Machine Learning for Breast Cancer Diagnosis

Machine learning techniques have long been applied to breast cancer diagnosis using structured clinical datasets. Logistic Regression remains a baseline method due to its simplicity and interpretability, while Support Vector Machines have been widely adopted for their robustness in handling non-linear decision boundaries [11,12]. Prior studies report high accuracy using these methods but note sensitivity to feature selection and kernel tuning.

2.2 PCA-Based Feature Engineering in Medical Classification

Several works highlight the effectiveness of PCA in reducing dimensionality while preserving diagnostic information [9,10]. PCA-based pipelines enable improved visualization and reduce overfitting, particularly when combined with classifiers operating on low-dimensional representations. However, aggressive dimensionality reduction may limit the expressive power of classical models.

2.3 Foundations of Quantum Machine Learning

Quantum Machine Learning integrates quantum computing principles such as superposition and entanglement with classical learning algorithms. Quantum circuits can represent complex transformations using fewer parameters compared to classical networks, making them attractive for data-constrained problems.

2.4 Variational Quantum Circuits for Classification Tasks

Variational quantum circuits have been successfully applied to binary classification tasks using parameterized rotation gates and entanglement layers [6,16]. Optimization is performed using classical gradient-based methods, enabling practical deployment on simulated or NISQ-era quantum devices.

2.5 Quantum Models in Healthcare Analytics

Recent studies demonstrate the feasibility of quantum classifiers in healthcare domains, including cancer detection and biomedical signal analysis [17]. While large-scale quantum advantage is yet to be established, small-scale quantum models have shown competitive performance against classical baselines, particularly when feature dimensionality is limited.

2.6 Research Gap Addressed by This Work

Most existing studies evaluate quantum or classical models in isolation. This work addresses the gap by providing a direct, implementation-level comparison between classical and quantum classifiers using the same dataset, preprocessing strategy, and evaluation metrics.

3. Methodology

3.1 Dataset and Problem Formulation

This study addresses the problem of binary classification of breast tumours into benign and malignant categories using clinical diagnostic data. The experiments are conducted on a standardized breast cancer dataset containing numerical features derived from digitized medical images. Each sample is associated with a binary class label indicating tumour malignancy.

The classification task is formulated as a supervised learning problem, where the objective is to learn a mapping from real-valued clinical features to discrete class labels with high predictive accuracy and reliability.

3.2 Data Preprocessing and Feature Engineering

Prior to model training, the dataset undergoes normalization to eliminate scale disparities among features and improve numerical stability. To reduce dimensionality while preserving maximum variance, Principal Component Analysis (PCA) is applied [10].

Only the two most significant principal components are retained, as they capture the dominant structure of the dataset and are sufficient for effective classification. This dimensionality reduction serves two critical purposes:
- It reduces computational complexity, which is essential for near-term quantum models.
- It enables direct visualization of decision boundaries and class separability.

The resulting two-dimensional feature space is used consistently across both quantum and classical models to ensure fair comparison.

3.3 Variational Quantum Classifier Framework

3.3.1 Quantum Feature Encoding

The reduced feature vectors are encoded into quantum states using an angle encoding strategy, where each principal component controls a rotation gate on a separate qubit. For a two-dimensional input vector (x_1, x_2), the corresponding quantum state is prepared by applying rotation operations to initialize the qubits. This encoding approach efficiently maps classical information into quantum states while maintaining continuous feature representation.

3.3.2 Variational Circuit Architecture

The core of the quantum classifier is the **parameterized quantum circuit (PQC)**, also referred to as the **ansatz**, which is designed to operate on a two-qubit system corresponding to the two Principal Components, $\mathbf{x} = [x_1, x_2]$. The ansatz is constructed by chaining Lidentical variational layers, denoted as $V(\boldsymbol{\theta}^{(l)})$, which are applied immediately after the feature encoding layer.

The complete quantum circuit evolution can be formally expressed as:

$$U(\mathbf{x}, \boldsymbol{\theta}) = \left(\prod_{l=1}^{L} V\left(\boldsymbol{\theta}^{(l)}\right)\right) U_{\text{Encoding}}(\mathbf{x})$$

where $U_{\text{Encoding}}(\mathbf{x})$represents the angle-encoding operation applied to the input features, and $\boldsymbol{\theta}$denotes the full set of trainable circuit parameters.

Each variational layer $V(\boldsymbol{\theta}^{(l)})$is composed of two primary components:
- **Trainable Single-Qubit Rotation Gates:** A set of parameterized rotation gates, $\mathrm{R}_y(\theta_i^{(l)})$, is applied to each qubit $i \in \{0,1\}$. These gates introduce tunable parameters that are optimized by a classical optimizer during the training process.
- **Fixed Entangling Operations:** A fixed Controlled-NOT (CNOT) gate is applied between the two qubits, typically in the configuration $\mathrm{CNOT}_{0,1}$. This entangling operation

enables the circuit to capture correlations and non-linear interactions between the encoded input features.

The layered and parameterized structure of the variational quantum circuit allows it to implicitly map low-dimensional PCA features into a high-dimensional quantum Hilbert space, thereby potentially enhancing class separability. The circuit depth L is intentionally kept shallow to remain compatible with the limitations of current Noisy Intermediate-Scale Quantum (NISQ) devices, while still providing sufficient expressive power for effective classification.

3.3.3 Measurement and Classification Rule

The output of the variational circuit is obtained by measuring the expectation value of a Pauli-Z observable on one qubit[2]. This expectation value lies in the interval $[-1,1]$ and is mapped to a probabilistic prediction using a linear transformation:

$$P(y=1) = \frac{1 + \langle Z \rangle}{2}$$

A threshold of 0.5 is used to convert the probability into a binary classification outcome, where higher values indicate malignant tumours.

3.4 Training Strategy and Optimization

The variational parameters of the quantum circuit are optimized using a gradient-based optimizer[8]. Training is performed iteratively over multiple epochs, minimizing a binary cross-entropy loss function, which is well suited for probabilistic binary classification.

Automatic differentiation is employed to compute gradients through the quantum circuit, enabling seamless integration of quantum models into classical optimization workflows. The optimization process continues until convergence in both training loss and classification accuracy.

3.5 Classical Machine Learning Baselines

To evaluate the effectiveness of the quantum classifier, two established classical models are trained using the same PCA-reduced dataset.

3.5.1 Logistic Regression

Logistic Regression is used as a linear baseline classifier. It models the posterior probability of malignancy using a sigmoid activation function applied to a linear combination of input features. While computationally efficient and interpretable, its linear nature limits its ability to capture complex feature interactions.

3.5.2 Support Vector Machine

A Support Vector Machine with a non-linear kernel is employed to provide a stronger classical baseline[11]. The SVM constructs an optimal separating hyperplane in a transformed feature space, allowing it to model non-linear decision boundaries. Despite its effectiveness, SVM performance depends on kernel selection and can be computationally intensive.

3.6 Evaluation Protocol and Performance Metrics

All models are evaluated on unseen test data using consistent evaluation criteria. Performance is assessed through:

- Classification accuracy
- Confusion matrices to analyse class-wise prediction behaviour
- Receiver Operating Characteristic (ROC) curves to evaluate discrimination capability
- Visualization of decision boundaries in the reduced feature space

This comprehensive evaluation framework ensures both quantitative and qualitative comparison between quantum and classical approaches.

4. Implementation

4.1 Implementation and System Workflow

This section details the practical realization of the proposed hybrid quantum–classical breast cancer classification framework. The complete pipeline was implemented using Python, integrating quantum simulation, classical machine learning, and an interactive testing interface.

4.2 Overall System Architecture

The end-to-end workflow of the proposed system consists of five major stages: data preprocessing, feature reduction, model training, model evaluation, and user-level testing.

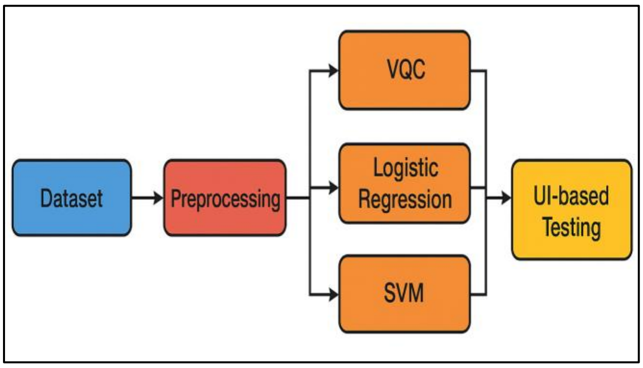


Figure 1: System Workflow

This architecture ensures that both quantum and classical models operate on the same reduced feature space, enabling a fair and controlled comparison.

4.3 Data Processing and Feature Reduction Pipeline

The original breast cancer dataset contains multiple real-valued diagnostic features. Since current quantum hardware and simulators are limited in qubit capacity, dimensionality reduction is essential.

The preprocessing pipeline includes:

- Feature normalization
- Train–test split
- Principal Component Analysis (PCA)

Paper ID: SR251210205803 DOI: https://dx.doi.org/10.21275/SR251210205803

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Only the top two principal components were retained, as they capture the dominant variance in the dataset while allowing direct mapping to a two-qubit quantum circuit.

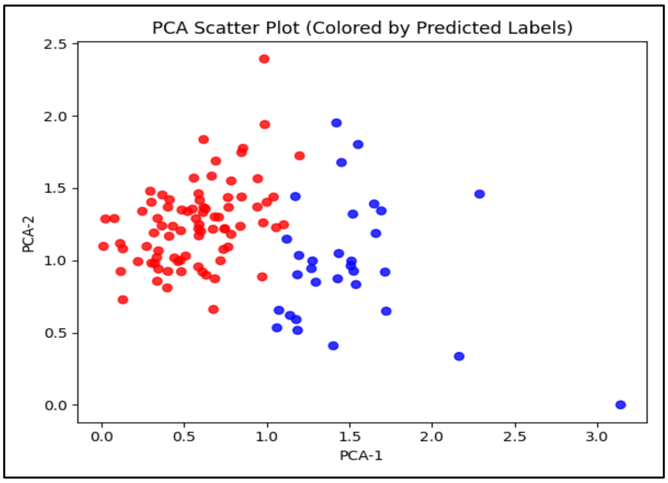


Figure 2: PCA Scatter Plot

This figure visually demonstrates the separability between benign and malignant samples in the reduced feature space, justifying its suitability for both classical and quantum classification.

4.4 Variational Quantum Classifier Implementation

4.4.1 Quantum Circuit Design

The Variational Quantum Classifier (VQC) was implemented as a **two-qubit circuit** to accommodate the two PCA features. The circuit consists of:

- **Angle Encoding Layer:** Classical PCA features are embedded into quantum states using rotation gates.
- **Variational Layers:** Each layer contains trainable rotation gates followed by entangling operations to capture feature correlations.
- **Measurement Layer:** Expectation value measurement of a Pauli-Z observable is used to generate model output.

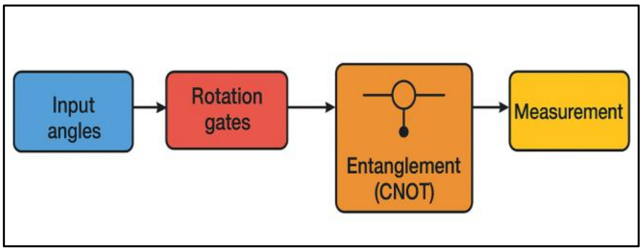


Figure 3: VQC Architecture

4.4.2 Hybrid Training Loop

Training the VQC follows a hybrid quantum–classical optimization process:

- Classical data is encoded into a quantum state.
- The quantum circuit produces an expectation value.
- The output is mapped to a probability.
- A classical optimizer updates circuit parameters.

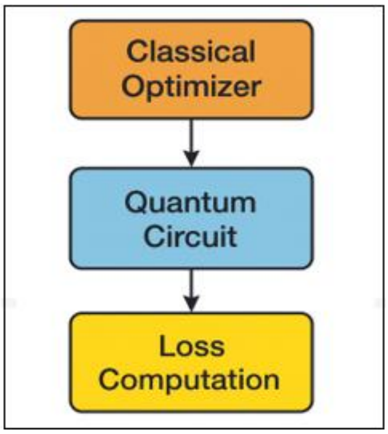


Figure 4: Hybrid Quantum–Classical Training Flowchart

This approach leverages quantum expressiveness while maintaining compatibility with classical optimization techniques.

4.5 Classical Model Implementation

To benchmark the quantum model, two classical classifiers were implemented using the same PCA-reduced dataset:

- **Logistic Regression:** A linear probabilistic classifier commonly used in medical diagnosis.
- **Support Vector Machine (SVM):** A nonlinear classifier capable of constructing complex decision boundaries.

Both models were trained using standard optimization techniques without extensive hyperparameter tuning, ensuring a realistic baseline comparison.

(Insert Figure 5: Decision Boundaries of Logistic Regression and SVM)

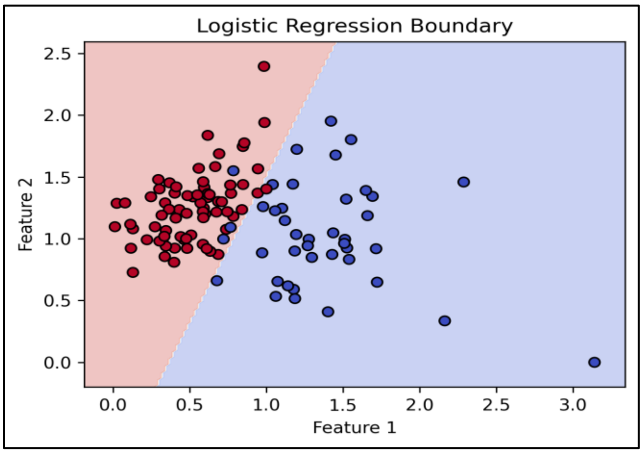


Figure 5: Decision Boundaries of Logistic Regression

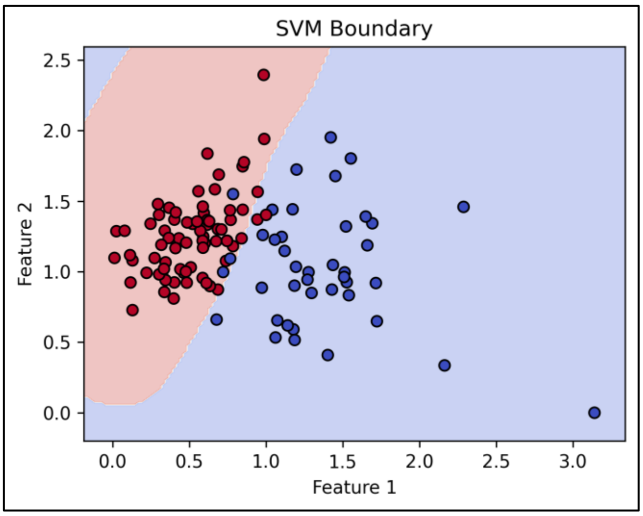


Figure 6: Decision Boundaries of SVM

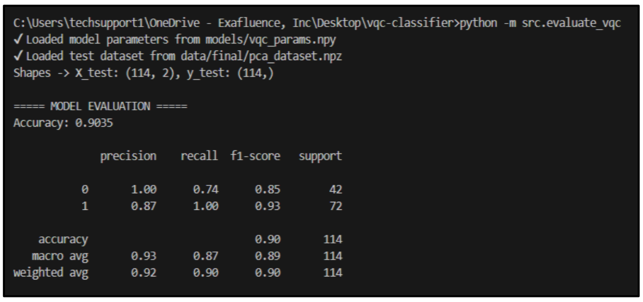
These plots highlight the geometric differences between linear and nonlinear classical classifiers.

4.6 Testing and Evaluation Workflow

After training, all models were evaluated using a held-out test set. Performance was assessed using metrics relevant to clinical decision support systems:

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix

For the VQC, the quantum measurement output was converted into a probability score to enable direct comparison with classical model outputs.



```
C:\Users\techsupport1\OneDrive - Exafluence, Inc\Desktop\vqc-classifier>python -m src.evaluate_vqc
✓Loaded model parameters from models/vqc_params.npy
✓Loaded test dataset from data/final/pca_dataset.npz
Shapes -> X_test: (114, 2), y_test: (114,)

===== MODEL EVALUATION =====
Accuracy: 0.9035

              precision    recall  f1-score   support

           0       1.00      0.74      0.85        42
           1       0.87      1.00      0.93        72

    accuracy                           0.90       114
   macro avg       0.93      0.87      0.89       114
weighted avg       0.92      0.90      0.90       114
```

Figure 7: VQU Model Evaluation

4.7 Interactive Model Testing Interface

To validate real-world usability, the trained quantum model was integrated into a web-based interface. Users can manually input PCA feature values and receive:

- Predicted class label (Benign or Malignant)
- Confidence score (probability)

This interactive testing layer demonstrates the feasibility of deploying quantum machine learning models in applied diagnostic environments.

5. Results & Performance Analysis

This section presents a comprehensive evaluation of the proposed Variational Quantum Classifier (VQC) for breast cancer classification. The quantum model's performance is analysed and compared against established classical machine learning approaches using consistent datasets and evaluation metrics.

5.1 Experimental Setup

All experiments were conducted on a PCA-reduced breast cancer dataset, where two principal components were used as input features for both quantum and classical models. The dataset was divided into training and testing subsets to ensure unbiased evaluation.

Three models were evaluated:

- Variational Quantum Classifier (VQC)
- Logistic Regression (LR)
- Support Vector Machine (SVM)

Each model was trained and tested on identical data partitions to maintain fairness in comparison.

5.2 Classification Accuracy Comparison

Overall classification accuracy was used as the primary metric to assess model effectiveness.

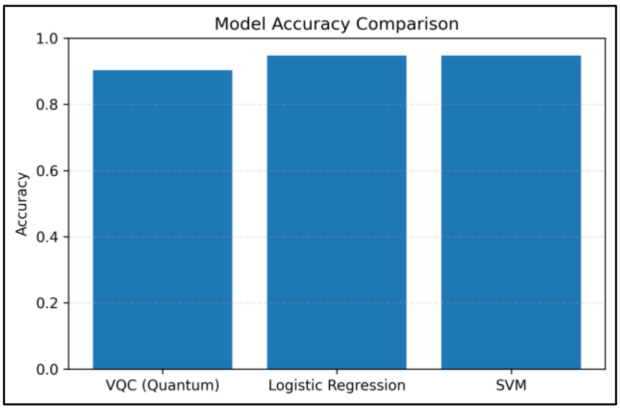


Figure 8: Accuracy Comparison of VQC, LR and SVM

The VQC achieved an accuracy of **90.35%**, demonstrating strong classification capability despite operating on a low-dimensional quantum feature space. Logistic Regression and SVM achieved slightly higher accuracies of **94.74%**, reflecting their maturity and optimization on classical hardware.

Although the classical models marginally outperform the quantum model in absolute accuracy, the VQC's performance is notable given the limited number of qubits and shallow circuit depth used in this study.

5.3 Confusion Matrix Analysis

Confusion matrices were used to analyse class-wise prediction behaviour and misclassification patterns.

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

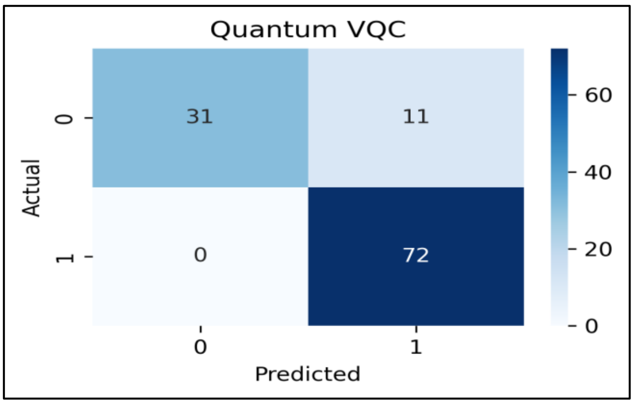


Figure 9: Confusion Matrix of VQC

The VQC correctly classified all malignant samples in the test set, achieving a recall of **1.00** for the malignant class. A small number of benign samples were misclassified, indicating a conservative bias toward detecting malignancy—an important characteristic for medical diagnosis applications where false negatives are costly.
For comparison, confusion matrices of the classical models are shown below:

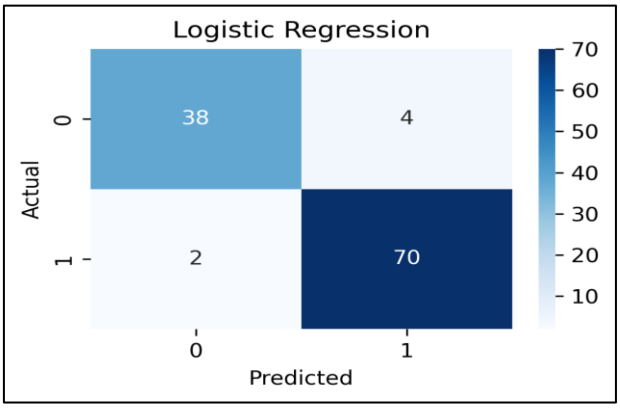


Figure 10: Confusion Matrix of Logistic Regression

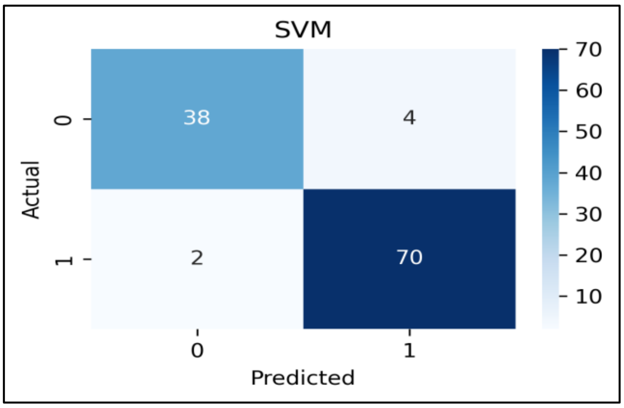


Figure 11: Confusion Matrix of SVM

Both classical models exhibit balanced performance across classes, with fewer misclassifications overall.

5.4 Precision, Recall, and F1-Score Evaluation

Detailed classification metrics provide deeper insight into model behaviour.
The VQC achieved:
- High recall for malignant cases
- Competitive F1-score
- Slightly reduced precision for benign classification

Table 1: Accuracy Precision, Recall, and F1-Score Evaluation

Metric	VQC (Quantum)	Logistic Regression	SVM
Accuracy	0.9035	0.9474	0.9474
Recall (Class 1, Malignant)	1	0.97	0.97
Precision (Class 1, Malignant)	0.87	0.95	0.95
F1-Score (Class 1, Malignant)	0.93	0.96	0.96
Recall (Class 0, Benign)	0.74	0.9	0.9

This behaviour indicates that the quantum model prioritizes sensitivity over specificity, which is often desirable in clinical screening scenarios.

Classical models demonstrated consistently high precision and recall across both classes, reflecting their stability on linearly separable PCA-transformed data.

5.5 Decision Boundary Visualization

To understand how different models separate the feature space, decision boundary plots were generated.

The VQC forms a smooth, nonlinear boundary, demonstrating the expressive capability of parameterized quantum circuits even with limited depth.

Logistic Regression produces a linear boundary, while SVM constructs a more flexible nonlinear boundary. The quantum model exhibits behaviour comparable to nonlinear classical classifiers, reinforcing its representational potential.

5.6 ROC Curve and Discriminative Capability

Receiver Operating Characteristic (ROC) analysis was used to evaluate the discriminative power of the classifiers.

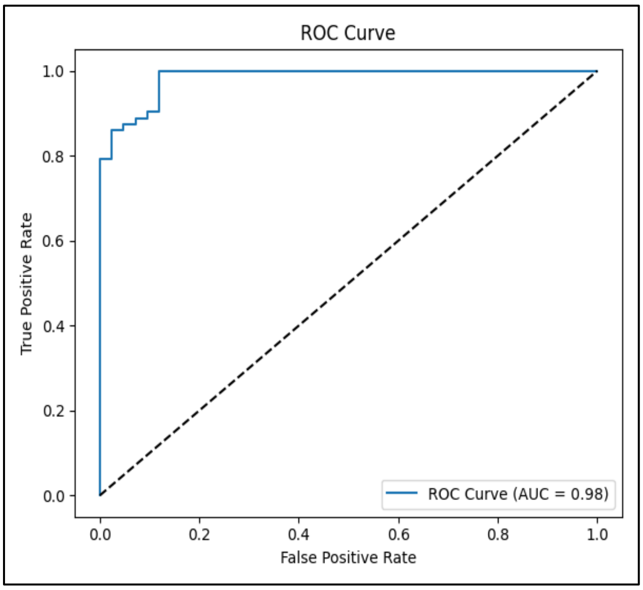


Figure 12: ROC curve of VQC

The ROC curve indicates strong separability between classes, with an area under the curve (AUC) comparable to classical approaches. This result highlights the robustness of the quantum classifier in distinguishing malignant and benign cases.

5.7 Discussion of Results

The experimental results demonstrate that:

- The Variational Quantum Classifier achieves competitive performance with classical models.
- Quantum decision boundaries capture nonlinear patterns effectively.
- The model shows strong sensitivity toward malignant cases.
- Comparable performance is achieved using only two qubits and shallow circuits.

These findings suggest that even near-term quantum models can serve as viable alternatives or complements to classical machine learning approaches in healthcare analytics.

6. Conclusion & Future Scope

6.1 Conclusion

This study presented a comparative analysis of a Variational Quantum Classifier (VQC) and classical machine learning models for breast cancer classification using a reduced-dimensional feature space. By integrating quantum machine learning techniques with classical preprocessing methods, the proposed approach demonstrated that quantum classifiers can achieve competitive performance even under current hardware and simulation constraints.

The VQC successfully classified malignant and benign samples with an overall accuracy exceeding 90%, showing strong sensitivity toward malignant cases. This characteristic is particularly important in medical diagnostics, where minimizing false negatives is critical. Despite utilizing only two principal components and a shallow quantum circuit, the quantum model exhibited nonlinear decision boundaries comparable to those produced by classical nonlinear classifiers.

Classical Logistic Regression and Support Vector Machine models achieved slightly higher accuracy, reflecting their maturity and extensive optimization on classical computing platforms. However, the quantum classifier's performance highlights the expressive power of variational quantum circuits and their potential applicability to real-world healthcare problems.

The results confirm that hybrid quantum-classical learning frameworks can serve as effective exploratory tools in medical data analysis and provide a meaningful pathway toward future quantum-enabled clinical decision systems.

6.2 Future Scope

While the current work demonstrates the feasibility of quantum classification for breast cancer prediction, several directions remain open for future research and enhancement:

1) **Scaling to Higher-Dimensional Feature Spaces:** Future studies can explore encoding additional clinical and imaging features using advanced quantum embedding techniques, enabling richer representations beyond two principal components.
2) **Execution on Real Quantum Hardware:** Implementing the model on available noisy intermediate-scale quantum (NISQ) devices would allow assessment of noise robustness, circuit depth limitations, and real-world feasibility beyond simulation environments.
3) **Advanced Quantum Circuit Architectures:** Experimenting with deeper variational circuits, alternative entanglement strategies, and adaptive ansatz designs may improve expressiveness and classification performance.
4) **Hybrid Ensemble Approaches:** Combining quantum classifiers with classical ensemble methods could enhance robustness and leverage the complementary strengths of quantum and classical learning paradigms.
5) **Explainability in Quantum Models:** Developing interpretability techniques for quantum decision processes would improve clinical trust and facilitate adoption in medical diagnostics.
6) **Extension to Other Medical Applications:** The proposed framework can be extended to other disease classification tasks, such as tumour grading, genetic disorder detection, and multimodal healthcare analytics.

As quantum hardware continues to advance, variational quantum models are expected to play an increasingly significant role in data-driven healthcare solutions. This work contributes to the growing body of evidence supporting the practical relevance of quantum machine learning in clinical decision-making.

References

[1] F. Arute *et al.*, "Quantum supremacy using a programmable superconducting processor," *Nature*, vol. 574, no. 7779, pp. 505–510, 2019.

[2] M. Schuld and F. Petruccione, *Supervised Learning with Quantum Computers*. Springer, 2018.

[3] V. Havlíček *et al.*, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, pp. 209–212, 2019.

[4] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[5] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum algorithms for supervised and unsupervised machine learning," *arXiv preprint arXiv:1307.0411*, 2013.

[6] M. Schuld, A. Bocharov, K. Svore, and N. Wiebe, "Circuit-centric quantum classifiers," *Physical Review A*, vol. 101, no. 3, p. 032308, 2020.

[7] H. Abbas *et al.*, "The power of quantum neural networks," *Nature Computational Science*, vol. 1, pp. 403–409, 2021.

[8] D. P. Kingma and M. Welling, "An introduction to variational methods for machine learning," *Foundations and Trends in Machine Learning*, vol. 12, no. 4, pp. 307–392, 2019.

[9] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[10] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of*

Educational Psychology, vol. 24, no. 6, pp. 417–441, 1933.

[11] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[12] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Wiley, 2013.

[13] W. H. Wolberg *et al.*, "Breast cancer Wisconsin (diagnostic) data set," *UCI Machine Learning Repository*, 1995.

[14] J. Brownlee, *Machine Learning Mastery with Python*. Machine Learning Mastery, 2016.

[15] A. Schuld and N. Killoran, "Quantum machine learning in feature Hilbert spaces," *Physical Review Letters*, vol. 122, no. 4, p. 040504, 2019.

[16] K. Mitarai *et al.*, "Quantum circuit learning," *Physical Review A*, vol. 98, no. 3, p. 032309, 2018.

[17] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Physical Review Letters*, vol. 113, no. 13, p. 130503, 2014.

[18] S. Biamonte *et al.*, "Quantum machine learning," *Nature*, vol. 549, pp. 195–202, 2017

Paper ID: SR251210205803 DOI: https://dx.doi.org/10.21275/SR251210205803