RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF
MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt26418

Adaptive Optimizer-Enhanced Variational Quantum Classifier Using ZZFeatureMap and TwoLocal Ansatz for Non-Linear Binary Classification

D. Rajalakshmi [a, *], R.R. Ramya [b], S. Sridevi [c], K. Sudharson [b]

[a] Department of CSE, SRM Institute of Science and Technology, Ramapuram, 600089, India.
[b] Department of CSE, Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science & Technology, Avadi, 600062, India.
[c] Department of CSE, Velammal Engineering College, Surapet, 600066, India.
* Corresponding Author Email: draji.cse@gmail.com

DOI: https://doi.org/10.54392/irjmt26418
Received: 02-01-2026; Revised: 07-07-2026; Accepted: 20-07-2026; Published: 27-07-2026

Check for updates

Abstract: Noisy Intermediate-Scale Quantum (NISQ) devices can be used to solve classification problems using a promising quantum machine learning technique called Variational Quantum Classifiers (VQC). The performance of VQCs is highly dependent on classical optimizers, and the understanding of the optimizer behavior within a given quantum architecture is limited. The proposed work is to present an Adaptive Optimizer-Enhanced Variational Quantum Classifier, which is a combination of a ZZFeatureMap for quantum data encoding and TwoLocal ansatz for variational learning, for solving non-linear binary classification problems. The proposed framework is tested using the Two Moons benchmark dataset, which is a popular benchmark for evaluating the ability of machine learning models to learn complex non-linear decision boundaries. The implementation is based on the Qiskit 1.2 and was run on a noiseless state vector simulator to remove the effects of hardware noise and focus on the effects of optimization. The three popular optimization techniques, namely ADAM, COBYLA and SPSA were explored under the same experimental setups. Besides the classification accuracy, a callback-based loss tracking mechanism was added to examine the speed of convergence, stability of optimization and evolution of losses during training. The experimental results showed that the overall accuracy of ADAM is 84.2%, which is the best, followed by COBYLA (79.3%) and SPSA (73.8%). Along with this, quantitative convergence analysis results showed that ADAM converged faster than the other optimizers and had lesser loss variance. It is shown that, with a proper optimization routine, a shallow two-qubit VQC can be used to successfully learn complex non-linear decision boundaries. The principal contribution of this work is a controlled optimizer-aware evaluation framework that integrates a fixed ZZFeatureMap–TwoLocal VQC architecture with callback-based convergence diagnostics to systematically analyse optimizer behaviour under identical experimental conditions.

Keywords: Quantum Machine Learning, Variational Quantum Classifier, Quantum Feature Maps, Optimizer, Binary Classification.

1. Introduction

Quantum machine learning (QML) is a developing interdisciplinary field combining quantum computing with classical machine learning. QML is motivated by the learning, representational, and processing power of quantum systems due to their unique capabilities, including superposition, entanglement, and accessibility to higher dimensional Hilbert spaces. Among the other QML models generated to date, the most interest has been in Variational Quantum Circuits (VQCs) due to their NISQ (Noisy Intermediate-Scale Quantum) device compatibility, as well as their hybrid quantum-classical training frameworks [1].

VQCs use a feature map where classical information gets encoded into quantum states, and a corresponding parameterized quantum circuit (ansatz). Each ansatz is then subjected to value updates using classical optimizers [2]. The hybrid nature of the architecture allows the VQCs to exploit quantum feature spaces in combination with the classical optimization routines, thereby making them a likely candidate for most near-term quantum computers. VQCs, however, are still burdened with a number of practical issues that include the finite number of obtainable qubits, a limited depth of circuit layering, susceptibility to the circuit noise, plateauing of barren gradients, and problems with optimizer stability. These practical issues result in the feature map structures, ansatz, and optimization layers needing to be designed in careful ways that foster meaningful learning and stable convergence [3].

Optimizers play an essential role in determining the success of training in Variational Quantum Classifiers (VQCs). During the training of VQCs, the cost function using the measurement outputs is minimized, resulting in the VQC training landscape being described as non-convex, noisy, and sensitive to VQC structure and parameterization. While some optimization techniques, such as gradient-based optimizers, lack efficiency and suffer from instability, gradient-based optimizers tend to exhibit slow convergence and a lack of efficiency in measurement noise. This demonstrates the importance of understanding optimizers in VQC training to develop useful and robust models in quantum machine learning [4].

Recent studies have shown that it has also been shown that the optimization strategy, expressibility of the circuit, and the data encoding also play a significant role in the trainability of variational quantum models. The selection of the optimizer has become a key issue for the convergence stability, barren plateau mitigation, and learning efficiency of hybrid quantum-classical architectures in particular. Therefore, it is essential to systematically assess the optimization behavior when implementing Variational Quantum Classifiers in practical applications with NISQ devices.

For quantum variational algorithms, the optimization techniques ADAM, COBYLA, and SPSA have been adapted for quantum systems and are described as adaptive and noise resilient. ADAM is an adaptive gradient optimizer in which learning rates are modified based on moment estimates, which is efficient in complex optimization problems in which the objective is to achieve gradient smoothness. COBYLA is a gradient-less optimizer of the trust-region type, which optimally balances noise and constraint satisfaction, and SPSA works in stochastic optimization and perception-efficient contexts, although in the current context it is analyzed in the case of noiseless simulations. The present study is one of the few examples in which these optimizers are used in an integrated VQC framework and addresses the convergence of non-linear classification problems [5].

The feature map and ansatz choice is one of the many leading design components of VQCs. Entanglement-based feature maps, like the ZZFeatureMap, are said to embed classical information into quantum states in such a way that complex correlations are created between qubits. This may improve the expressivity and class separability of the quantum feature space. There is a similar characteristic for the TwoLocal ansatz, which consists of interleaving parameterized rotations of single qubits and some entangling gates. This provides depth control with enough flexibility for shallow circuit design, which is a good fit for the NISQ paradigm. The combination of these components allows for the design of quantum models that are expressive enough to learn non-linear decision boundaries with a very small number of qubits [6].

This work presents a controlled optimizer-aware evaluation framework for Variational Quantum Classifiers by integrating a fixed ZZFeatureMap and TwoLocal ansatz with callback-based convergence diagnostics. The objective is to systematically investigate the influence of optimizer selection on convergence behaviour, training stability and classification performance under identical experimental conditions. Using Qiskit v1.2 the VQC is simulated and differences among the optimizers are quantified. ADAM, COBYLA, and SPSA are the chosen optimizers and the primary focus of research is on the disparity of these optimizer on convergence, training stability, and classification accuracy. In order to enhance interpretability and training diagnostics, we incorporate a loss metric and a visualization tool to track optimizer progress over training iterations, augmenting Qiskit's standard implementation of VQC [7]. This addition sheds light on convergence behavior, the influence of noise, and the stability of the optimization process—issues that usually receive scant attention in opaque quantum learning studies. By providing a comparative analysis of optimizers within a single VQC framework, this study expands the scope of QML research, and legitimizes the use of VQC's for practical learning tasks during the NISQ period.

Despite significant advances in variational quantum classification, previous works focused mainly on classification accuracy or on proof-of-concept experiments on benchmark datasets. Comparatively less research has focused on a systematic optimizer-centric analysis, quantifying convergence speed, training stability and evolution of loss for a fixed quantum circuit configuration. Moreover, the corresponding literature does not provide comprehensive comparisons between frequently used optimizers, e.g. ADAM, COBYLA and SPSA. The gap inspires the focus of the present study: to understand the behaviour of the optimiser within a controlled VQC context with a fixed feature map, ansatz and evaluation protocol.

In this paper the term adaptive optimizer enhanced Variational quantum Classifier (VQC) is used to denote a structured design, in which various classical optimizers ADAM, COBYLA, and SPSA are systematically integrated into a fixed quantum circuit architecture to examine their effect on convergence behavior, stability and classification. In contrast to the traditional implementation of VQC based on one optimizer, the given research proposes an optimizer-aware training framework that allows the analysis of comparative learning behavior in case of the same circuit configuration. The framework is also complemented with a loss-tracking mechanism that offers a fine-grained detail of the optimization paths, and thus enhances the understanding of the training process.

The main contributions of this work are:

1. Offering a coherent and regulated comparison of several optimizers in the identical VQC structure,
2. Adding interpretability in optimizer level by tracking of the losses and convergence diagnostics,
3. Measuring the performance on a non-linear benchmark (Two Moons dataset) that is difficult to handle, and
4. Showing that shallow two-qubit circuits with an appropriate optimizer can be useful to learn complicated decision boundaries.

This is the distinguishing factor between the proposed work and the existing body of literature which mostly addresses feasibility or single-optimizer performance without involving the convergence analysis.

The main findings of this work can be summarized as follows:

1. An integrated adaptive optimizer enhanced VQC framework of ADAM, COBYLA and SPSA using the same experimental conditions.
2. Loss convergence dynamics and training stability analysis via loss-tracking a callback-based framework.
3. A comparative analysis of optimizer performance on a non-linear classification test (Two Moons dataset).
4. Experimental evidence of effective non-linear learning of decision boundaries with a shallow two-qubit circuit.

This research shifts attention from typical performance metrics to how optimizers behave inside a fixed setup for quantum classification. Rather than chasing higher scores, it shows changes in training patterns based solely on which optimization method is chosen. A single consistent structure and data set reveal stark differences in learning paths across methods. Depending on the optimizer, outcomes range from smooth progression to erratic jumps during training. Boundaries separating classes also shift noticeably in shape and stability. Such variations emerge without altering the model design or inputs. Results suggest that choice of optimizer alone can reshape both speed and outcome of learning. Not every method reaches similar endpoints, despite starting under equal settings. Subtle instabilities appear more tied to algorithmic mechanics than circuit layout. Performance depends less on hardware limits here, more on numerical strategies guiding updates.

2. Related Works

2.1 Quantum Feature Spaces and Variational Quantum Classifiers

Variational Quantum Classifiers (VQC) operate on the principle that quantum circuits can use feature maps quantum-enhanced feature spaces (a mechanism where classical data is mapped to quantum states and the calculations performed in the exponentially larger dimension of Hilbert space). HAVLÍČEK et al. created a model where quantum circuits can represent and solve non-linear decision boundaries, thereby solving the problems of supervised learning in ways that surpass classical kernel techniques (USE). Their work provided a basis to establish the viability of VQCs in performing non-linear classification through the employment of shallow quantum circuits, thereby making them appropriate for the NISQ period [7]. Schuld and Killoran expand on this, positing that quantum classifiers can be likened to linear models in quantum feature Hilbert spaces, similar to kernel machines [8]. The authors posit that the expressiveness of a quantum model is a function of the feature map, rather than the variational circuit. This finding strongly supports the motivation for the design of expressive feature maps for quantum classification problems and justifies the use of entanglement-based encodings in VQCs.

2.2 Entanglement-Based Feature Maps and ZZ Encoding

Adding entanglement to feature maps strengthens the quantum models' ability to represent more complex data points. In their study, Tomono *et al.* look at Z and ZZ feature maps and how they affect the learning performance in quantum kernel methods [9]. They show how ZZ-type encodings, which feature entanglement interactions on pairs of qubits, improve the separability of classes in the quantum feature space compared to other encodings, especially in the case of non-linear data sets.

In terms of end-user applications, the most accessible version of an entanglement-based encoding scheme is the ZZFeatureMap, which is available in Qiskit. In their study, Naresh and Reddi used Qiskit to analyze quantum machine learning workflows and found that ZZFeatureMap performed well across classification and regression tasks [10]. This body of work explains why, for VQCs that aim to perform non-linear binary classifications, ZZFeatureMap is an appropriate choice of data-encoding scheme.

2.3 Ansatz Design and TwoLocal Circuits for NISQ Devices

In addition to feature maps, the architecture of the ansatz is also integral to VQC success. Kandala *et al.* first described hardware-efficient variational circuits and argued for shallow circuits with a pattern of parameterized rotation and entanglement layers that is aligned with the capabilities of near-term quantum devices [11]. This design philosophy has become dominant in the field and is reflected in many of the off-the-shelf ansatz templates, including Qiskit's TwoLocal circuit.

To support the proposed ansatz choices, Sim and others constructed metrics for the Expressibility and Entangling power of parameterized quantum circuits [12]. Their analysis gives instructions for the preference of straightforward ansatz frameworks that facilitate a compromise between the competing metrics of expressibility and trainability, and thus, champions the use of shallow TwoLocal circuits for lower qubit VQC models.

2.4 The Problems of Optimization for Variational Quantum Algorithms

Notwithstanding their expressive possibilities, VQCs continue to battle serious issues of training that are rooted in the noisy measurements, and the non-convex nature of the loss landscapes, and the so-called barren plateaus. Cerezo and others delivered a holistic compendium of the literature on variational quantum algorithms where they identified issues related to optimization stability, noise, and the rapid scaling of the problem [13]. Of significance is the assertion that the choice of an optimizer is singular in importance for the attainment of consistent convergence within the variational frameworks.

2.5 The Use of SPSA, COBYLA and Adaptive Techniques as Optimizers for VQC Training

Because of noise problems and the finite number of measurements available, stochastic optimization strategies have become common in quantum contexts. Spall's Simultaneous Perturbation Stochastic Approximation (SPSA) is a technique in which estimation of the gradient is based on only two evaluations of the objective function, thus it is appealing for experiments on quantum hardware [14]. SPSA is now regarded as the standard optimizer for quantum variational algorithms. Singh et al. demonstrated that optimizer performances differ based on the type of problem and the amount of noise present in the system when benchmarking variational quantum applications with COBYLA and SPSA. This motivates the evaluation of multiple optimizers for training VQCs in classification contexts [15].

2.6 Recent Advances in Optimization Robustness and Interpretability

Recent work has focused on variational quantum models with an emphasis on improving the robustness and interpretability of optimisations. Shaffer et al. proposed a technique of surrogate-based optimisation that approximates cost landscapes to decrease the number of quantum circuit evaluations and to help convergence [16]. Notably, newer research points to hybrid quantum-classical systems boosting prediction accuracy by refining how adjustments are made during training.

Table 1. Summary of Related Works in Variational Quantum Classification

Reference	Feature Map / Encoding	Ansatz / Model	Optimizer	Key Contribution
Havlíček *et al.* [7]	Quantum kernel-based encoding	Circuit-based model	Classical	Demonstrated non-linear classification using quantum feature space
Schuld & Killoran [8]	Quantum feature Hilbert space	Linear model interpretation	Classical	Established connection between quantum models and kernel methods
Tomono *et al.* [9]	ZZ-based feature map	Kernel-based approach	Classical	Showed improved class separability using entanglement
Kandala *et al.* [11]	Hardware-efficient encoding	TwoLocal-type ansatz	Classical	Proposed shallow, hardware-efficient variational circuits
Singh *et al.* [15]	Variational quantum circuits	Standard VQC	COBYLA, SPSA	Compared optimizer performance in quantum settings
Proposed Work	ZZFeatureMap	TwoLocal ansatz	ADAM, COBYLA, SPSA	Optimizer-level convergence analysis with loss-tracking

Take, for example, deep learning setups enhanced with quantum components - these have shown faster progress toward accurate results within real-world operational settings [17]. Notably, predictive models built on quantum-safe methods - combined with optimization approaches - highlight how strong training tactics matter when applied outside controlled settings [18].

These studies also emphasize the importance of selecting optimizers carefully in variational quantum algorithms.This type of work illustrates the increasing importance of interpretability and diagnostics in hybrid quantum-classical frameworks. The current work builds on these studies and examines a VQC with a ZZFeatureMap and TwoLocal ansatz enhanced with an adaptive optimizer. A specific comparison of the ADAM, COBYLA, and SPSA optimizers was conducted on the non-linear Two Moons dataset. This work seeks to enhance the interpretability of the specific optimizer used, and provide practical insights for quantum classifiers in the NISQ era, along with loss-tracking callbacks to monitor and analyze convergence. Also, summary of previous work related to VQC shown in Table 1.

3. Preliminaries

3.1 Variational Quantum Circuits (VQCs)

Most near-term quantum machine learning models utilize variational quantum circuits (VQCs) due to their unique hybrid quantum-classical training framework as well as their suitability to Noisy Intermediate-Scale Quantum (NISQ) Devices. A VQC consists of three core components: a quantum data encoding circuit (also referred to as a feature map), a parameterized variational ansatz, and a classical optimizer which updates the circuit parameters based on the measured output.

A VQC can be mathematically expressed as preparing a quantum state of the form as in Eq. 1

$$|\psi(x,\theta)\rangle = U(\theta)\, U_\phi(x)\, |0\rangle^{\otimes n} \quad (1)$$

where,

- $x \in \mathbb{R}^d$ represents the classical input data,
- $U_\phi(x)$ is the feature map that encodes classical data into a quantum state,
- $U(\theta)$ is the variational ansatz with trainable parameters θ,
- n is the number of qubits.

processing step assigns expectation values to class labels.

Measuring one or more qubits in a specified basis produces the circuit's output, and a classical post-

3.2 Quantum Data Encoding and Feature Maps

Quantum feature maps embed classical data into the quantum states so that the quantum inner products are represented with respect to some kernel function. Given a feature map $U_\phi(x)$, Eq. 2, the quantum kernel is defined as

$$K(x,x') = \left|\langle 0| U_\phi^\dagger(x)\, U_\phi(x) |0\rangle\right|^2 \quad (2)$$

In this paper, we use the ZZFeatureMap entangles qubits to create a greater degree of non-linear separability. The ZZFeatureMap comprises of ZZ-interactions and Z-rotations to enable pairwise feature coupling.

ZZ-interactions are non-local operators that can act between pairs of qubits. In Eq. 3, For a two-qubit system, the ZZFeatureMap can be written as

$$U_\phi(x) = exp\left(i \textstyle\sum_j x_j Z_j + i \sum_{j<k} x_j x_k Z_j Z_k\right) \quad (3)$$

This encoding allows classical data to control single qubit phase shifts and create multi-qubit entanglement, which is important in capturing non-linear decision boundaries.

The corresponding two-qubit quantum data-encoding circuit is shown in Figure 1.

3.3 Variational Ansatz: TwoLocal Circuit

The variational ansatz determines the quantum model's flexibility and capacity. In NISQ devices, the TwoLocal is a shallow circuit of choice to reduce the impact of decoherence and gate errors. The TwoLocal ansatz is a hardware-efficient, variational circuit comprised of multiple rounds of single-qubit rotation and entangling gates as in Eq. 4.

Any two-local ansatz can be formulated as

$$U(\theta) = \textstyle\prod_{l=1}^{L}\left(\prod_{i=1}^{n} R_i(\theta_i^{(l)}) \prod_{(j,k)\in E} U_{ent}^{(j,k)}\right) \quad (4)$$

Where,

1. $R_i(\theta_i^{(l)})$ are the rotation gates which are parameterized (from RY, RZ),
2. $U_{ent}^{(j,k)}$ is the entangling gates (like CNOT, CZ),
3. L is the total number of repetition layers,
4. E is the pair of qubits to be entangled.

The structure of the TwoLocal variational ansatz is shown in Figure 2.

The TwoLocal design is optimally structured for low qubit VQC deployments owing to its balancing of expressivity and circuit depth.

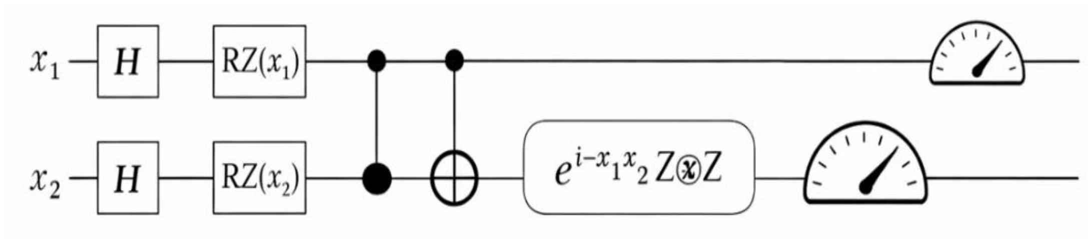

Figure 1. Quantum Data Encoding Using the ZZFeatureMap.

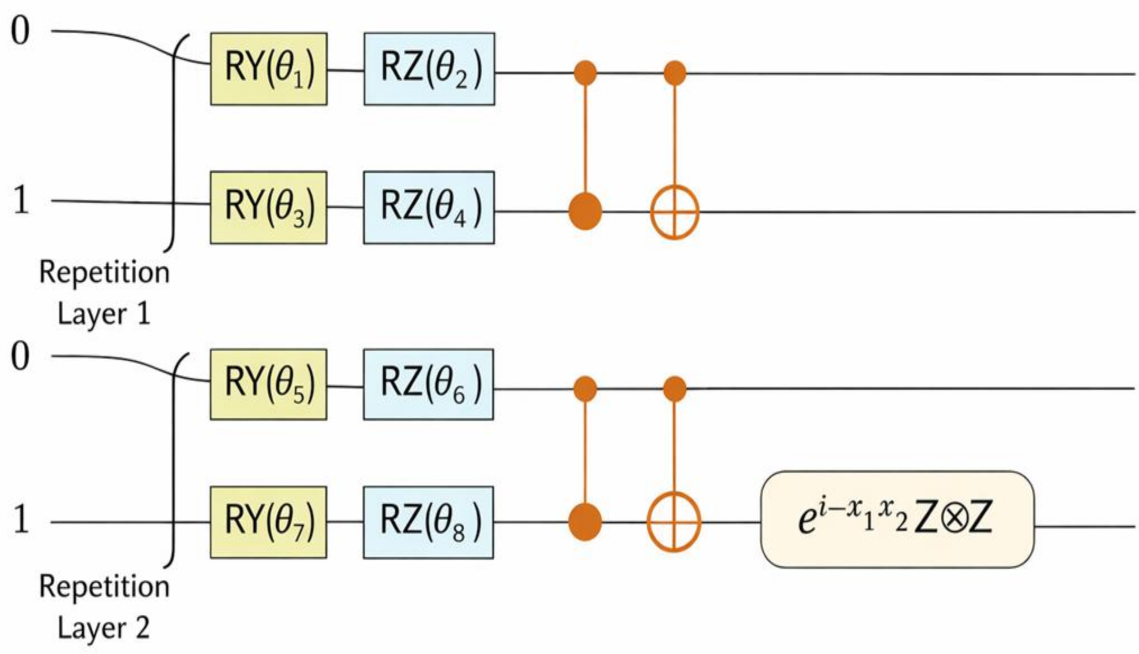

Figure 2. Variational Ansatz Using the TwoLocal Circuit.

3.4 Measurement Strategy and Loss Function

Once the variational quantum state is prepared, the measurements happen in the computational basis. Classically, for binary classification, the output of the model is some function of the expectation value of a Pauli observable (usually Z) as in Eq. 5:

$$\hat{y} = \langle \psi(x_i, \theta) | Z | \psi(x_i, \theta) \rangle \quad (5)$$

A threshold function is applied to predict a class label.

The model trains to minimize some loss function over the data. In the case of supervised binary classification, one typical loss is the mean squared error (MSE) as in Eq. 6:

$$\mathcal{L}(\theta) = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \text{ m} \quad (6)$$

Where

yi is the label, and

$\hat{y}_i$ is the output of the quantum circuit.

The loss used is the mean squared error (MSE), because of the continuous-valued expectation values that the quantum measurements give, which are in [− 1, 1]. Variational quantum circuits give the expectation values of observables unlike the classical classifiers which give probabilistic outputs which makes MSE an ideal regression-style optimization in classification problems. Furthermore, MSE guarantees constant gradient during training of hybrids with quantum-classical dynamics with respect to cross-entropy which can be unstable when non-linear probability mappings are used in quantum applications.

In the case of binary classification, the labels of classes are transformed between {0, 1} to {−1, +1} in order to fit the range of expectation values of the Pauli-Z observable. Zero is used as a threshold and predictions above or equal to 0 are classified as class +1 and those below 0 are classified as class -1.

Early tests using cross-entropy loss displayed erratic convergence patterns within the present VQC framework, highlighting MSE's better fit here. Though less common in classification, MSE offered more consistent training dynamics under these conditions.

3.5 Hybrid Quantum–Classical Optimization Loop

Using a hybrid optimization loop is how the VQC is trained. In this loop, the quantum circuit computes the loss function, and a classical optimizer adjusts the parameters θ.

The process involves the following steps:

- Data in its classical form is encoded with the quantum feature map
- The variational ansatz with the parameters is applied
- The expectation values are measured, and the loss is calculated
- A classical optimizer is used to update the parameters
- The loop is repeated until convergence is reached

Such a hybrid system allows the incorporation of a variety of adaptive and gradient-free optimizers like ADAM, COBYLA, and SPSA, which tend to work best for environments with significant quantum noise. This section laid the groundwork for the proposed model by building the foundational theory and explaining the mathematics behind VQCs, quantum feature mapping using the ZZFeatureMap, construction of the TwoLocal ansatz, measurement methods, and the process of hybrid optimization. These preliminaries inform the proposed adaptive optimizer-enhanced VQC, which is detailed in the following section.

4. Proposed Methodology

This section presents the implementation of the proposed optimizer-aware evaluation framework based on a fixed Variational Quantum Classifier architecture. The objective is to analyse how different classical optimizers influence convergence behaviour and predictive performance while maintaining identical quantum circuit configurations. While section 3 explained the theoretical foundations of variational quantum circuits, feature maps, and ansatz structures, the present section is narrowly focused on the implementation and optimization strategies that serve the primary contributions of this research. The entire end-to-end architecture of the proposed model is shown in Figure 3.

4.1 Problem Definition and Learning Objective

Consider a supervised learning dataset represented as

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N} \tag{7}$$

In Eq. 7, where $x_i = x_{i1}, x_{i2} \in \mathbb{R}^2$ represents a 2-dimensional input sample from the Two Moons dataset and $y_i \in \{-1, +1\}$ is the associated class label.

The learning goal is to create a parameterized quantum model $f(x)$ such that the output prediction is as close to the true class label as possible:

$$f_\theta(x_i) \approx y_i \tag{8}$$

As shown in Eq. 8, this is accomplished by performing a quantum state vector measurement and then adjusting the variational parameters θ of the quantum circuit via a classical optimization process hybrid with quantum computation.

4.2 Integrated Two-Qubit VQC Architecture

A two-qubit quantum circuit implements the proposed classifier because the number of qubits corresponds to the dimensionality of the input feature space. A classical feature component mapped to a qubit provides the respective classical input space a direct counterpart in the quantum computational space. A single pair of qubits fits naturally with the Two Moons data structure, since its inputs spread across two dimensions.

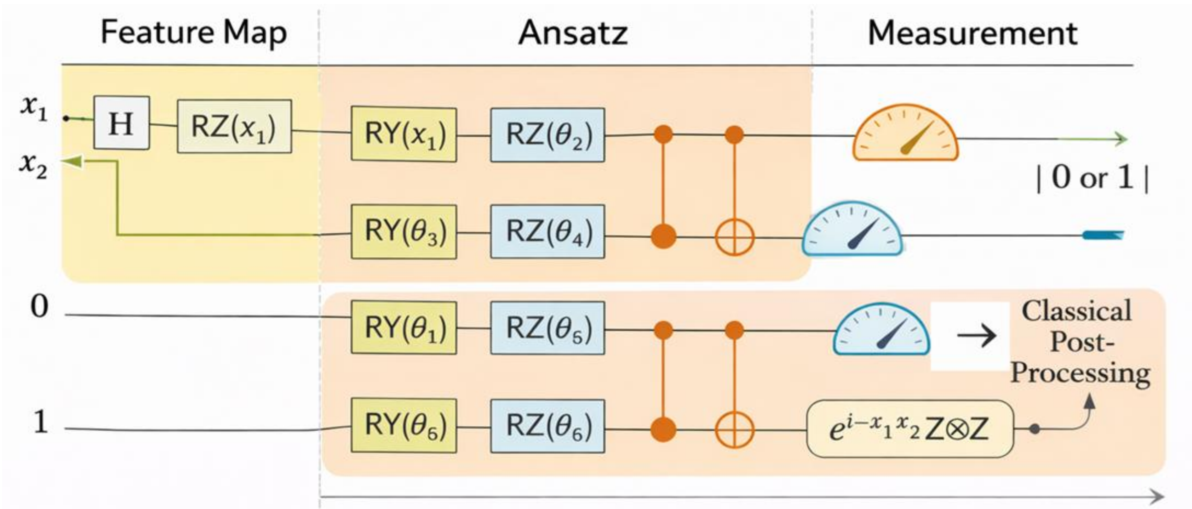

Figure 3. Integrated Two-Qubit VQC Architecture.

From each point rises a corresponding qubit state, linking geometry to hardware without extra layers. Circuit brevity follows as a consequence, yet entangled states still form where needed. Deeper circuits or more qubits could stretch representational power - though at a cost. Optimization grows harder when parts multiply, especially on today's limited machines.

The entire VQC pipeline is implemented on one quantum register as shown in Figure 3. It comprises three steps: quantum data encoding, variational transformation, and measurement-based prediction. This integrated approach enables effortless circuit execution without repeatedly initializing the quantum state, thus saving processing cycles, and minimizing the accumulation of noise. The fact that only two qubits are utilized for the circuit further supports the idea that modest quantum resources can sufficiently achieve non-linear classification, thus validating the application of VQCs in the NISQ era.

4.3 Quantum State Preparation and Forward Model

Given an input sample x_i, the quantum circuit constructs a parameterized quantum state defined as in Eq. 9,

$$|\psi(x_i,\theta)\rangle = U(\theta)\, U_\phi(x_i)\, |0\rangle^{\otimes 2} \tag{9}$$

where $U_\phi(x_i)$ is the entanglement-based ZZFeatureMap employed for the encoding of classical data, and $U(\theta)$ is the TwoLocal variational ansatz.

In order to obtain the model output, one has to measure a Pauli-Z observable on one of the qubits as in Eq. 10,

$$\hat{y}_i = \langle \psi(x_i,\theta)|Z|\psi(x_i,\theta)\rangle \tag{10}$$

where $\psi(x_i,\theta)$ denotes the expectation value of the observable Pauli-Z, resulting in a continuous prediction over the interval [−1,1]. This number is then converted to a class label through a mapping function based on a class prediction threshold.

4.4 Loss Function and Optimization Objective

The training of the VQC involves the optimization of a supervised loss function aimed at quantifying the loss with respect to the predicted output and the actual labeled output. In this context, the loss function is the mean squared error (MSE) in Eq. 11, which is selected for its ease of computation and consistency when paired with quantum outputs based on expectation values:

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \tag{11}$$

The optimization problem is thus described as in Eq. 12

$$\theta^* = \arg\min_{\theta} \mathcal{L}(\theta) \tag{12}$$

Where θ represents the VQC parameters, and this problem is addressed using classical optimization techniques in a stepwise manner.

4.5 Adaptive Optimizer-Enhanced Training Strategy

One of the pivotal parts of this study is the comprehensive documentation and assessment of using several different optimizers within the same VQC setup. The training process is carried out separately for each of the three selected optimizers: ADAM, COBYLA, and SPSA. Every experiment is performed with complete control over the parameters of the experiment, including circuit layout, seeding for parameters, the data for training, and the criteria for evaluating the model in order to maintain compositional integrity.

The optimization procedure can be viewed as a discrete-time dynamical system as in Eq. 13:

$$\theta^{(t+1)} = \theta^{(t)} + \Delta_{\mathcal{O}}^{(t)} \tag{13}$$

Here, $\Delta_{\mathcal{O}}^{(t)}$ signifies the change in parameters that the optimizer $\mathcal{O}$ makes. This structure gives us a way to determine the different optimization methods, how they converge, their stability, and how sensitive they are to the noise produced in quantum evaluations.

4.6 Loss-Tracking and Convergence Analysis Framework

For the sake of interpretability, we include a custom loss-tracking callback in the training loop. We define L(t) as the loss at optimization step t. We can then build convergence trajectories for the various optimizers and analyze the training process in detail, step by step, in the loss-tracking callback.

The loss tracking provides information on the speed of convergence, the presence of oscillations, and the stability to stochastic noise, thereby providing an optimizer-specific performance diagnostic in the context of variational quantum learning.

Rather than just watching losses drop, this new tracking method reveals how each optimizer behaves during training, measured step by step. Because it records these patterns, differences in how fast or smoothly optimizers reach results become clear inside the same quantum circuit setup. One way to measure how quickly a model converges is by counting the steps it takes to hit 95% of the lowest loss seen while training. Stability during learning shows up in how much the loss changes from step to step - less fluctuation means steadier progress. While high swings suggest instability, consistent drops point toward reliable improvement.

4.7 Algorithmic Description

The complete training workflow of the proposed Adaptive Optimizer-Enhanced VQC is summarized in Algorithm 1.

Algorithm 1. Adaptive Optimizer-Enhanced Variational Quantum Classifier

Input:

Training dataset D

Optimizer O

Output: Optimized parameters θ*

1. Initialize θ randomly
2. Build VQC
3. Set I ← Maximum iterations
4. While I > 0 do
5. Evaluate loss
6. Update θ using optimizer O
7. Save loss
8. I ← I − 1
9. End while10. Return θ*

This section presented the methodology for the Adaptive Optimizer-Enhanced VQC with a focus on two-qubit constructions, integrated architecture, mathematical learning objective, optimizer-driven architecture, and convergence analysis. Combining a resource-efficient quantum circuit with classical adaptive optimization and loss-tracking for the given quantum circuit, the methodology constructs a framework for a lower NISQ-era quantum machine learning resource/classifier non-linear classification.

5. Experimental Setup

This section outlines the experimental setup used to assess the proposed Adaptive Optimizer-Enhanced Variational Quantum Classifier. Details of the datasets, quantum simulation environments, configurations for training and evaluation metrics describe the experimental setup and enable other researchers to reproduce the work and make appropriate comparisons for other types of optimizers.

5.1 Dataset Description

The Two Moons dataset was selected to test the model since it is a common dataset for testing models designed for binary classification. The dataset contains a two-dimensional feature space that has two concave moon-shaped clusters which can be used for testing the model's ability to learn the complex shape decision boundaries. The feature vector $x = (x_1, x_2)$ captures each of the samples in the dataset and can be used to correlate primitives with the two-qubit structure used in this work. The dataset is normalized and used for training to eliminate any potential quantum state preparation issues. The binary class labels in the dataset are$y \in \{-1, +1\}$.

5.1.1 Reproducibility and Dataset Configuration

The Two Moons data set was generated by setting the random seed for its creation to be fixed. The dataset was partitioned using repeated stratified 80:20 train-test splits (5 repetitions) to preserve class distribution across all repetitions. Before state preparation, all input features were normalized using Min–Max scaling to ensure feature values lie in the range [0, 1] to guarantee stability in the process of quantum encoding and optimization. The following data partition, parameter initialization strategy and evaluation settings were used for all the optimizer experiments as used in Table 2.

Table 2. Experimental Setup Used for All Experiments

Parameter	Value
Dataset	Two Moons
Input Features	2
Qubits	2
Feature Map	ZZFeatureMap
Ansatz	TwoLocal
Parameter Initialization Seed	42
Normalization	Min-Max Scaling [0,1]
Train-Test Split	Repeated Stratified 80:20 Split (400/100 per repetition)
Simulator	Statevector Simulator
Qiskit Version	1.2

5.2 Quantum Simulation Environment

The experiments were conducted using Qiskit 1.2 running on a classical simulation backend. The experiments were all performed on a noiseless statevector simulator. Thus, the measurements represent the ideal quantum behavior in the absence of noise due to hardware. Any behavior of optimizers is only discussed when convergence properties are in the noiseless case and analysis in noisy or shot-based simulations is regarded as a future task. The evaluation of the algorithm in this simulation-based environment will provide a detailed assessment of the behavior of optimizers and will be helpful in evaluating optimizers on classical systems prior to using the real quantum systems.

5.3 VQC Arrangement and Circuit Specifications

A two-qubit quantum circuit consisting of a ZZFeatureMap for data encoding and a TwoLocal ansatz for variational learning was used to implement the Variational Quantum Classifier. The ZZFeatureMap was configured with a feature dimension of 2 and two repetition layers (reps = 2), using a linear entanglement strategy between qubits. The TwoLocal ansatz was comprised of parameterized RY and RZ rotation gates with CNOT-based entanglement layers, two repetition layers (reps = 2) and linear entanglement connectivity.

A common circuit architecture was used throughout the experiments to ensure a fair comparison among the different optimizers. The initial values of the variational parameters were randomly generated from a uniform distribution between 0 and 2π with a fixed random seed. This setup has a balance between circuit expressivity and trainability while being compatible with the limitation of Noisy Intermediate-Scale Quantum (NISQ) devices.

5.4 Optimizer Configuration

The performance of three classical optimizers (ADAM, COBYLA and SPSA) was assessed in the proposed Variational Quantum Classifier framework. For an equitable comparison, all optimizers were run with the same dataset partition, quantum circuit architecture, initialization and evaluation metrics. The learning rates used for the ADAM optimization were $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and the learning rate was 0.01. The numerical stability parameter ε was set to be equal to 1×10^{-8}. The initial value for the trust-region radius (rhobeg) was set to 1.0, and the convergence tolerance (tgtol) in COBYLA was set to 1×10^{-6}. For the stochastic gradient approximation, SPSA used a perturbation coefficient of 0.1 and a learning rate of 0.01.

The number of optimization iterations was set to 100 for all the optimizers. The training took place until either the maximum number of iterations was reached or the difference between the loss of two successive iterations was less than 1×10^{-6}. All experiments were performed under identical stopping conditions to ensure an unbiased comparison of convergence behavior, training stability and classification performance.

5.5 Training Process

The training follows an integrated quantum-classical approach model. During each iteration of optimization, the quantum circuit is run for every training sample, which gives us the expected values, which are then used for the calculation of the loss function. Based on the loss, the optimizer adjusts the variational parameters, and the whole process continues until convergence. To analyze convergence for each optimizer, a custom callback function is used to track loss values for every iteration.

Repeated trials with varied starting conditions helped confirm consistency in outcomes. Across several separate executions, measurements were collected to support dependability of findings. Average values appear in reports alongside typical fluctuations observed during testing phases. Variance among optimizer behaviors emerges clearly through such spreads. Stability checks rely on these patterns seen throughout repetitions.

To improve statistical reliability, each optimizer was evaluated over five repeated stratified train-test splits generated using different random seeds. For each repetition, the train-test partition was varied while the variational circuit parameter initialization was kept fixed using a random seed of 42 to ensure reproducibility. Performance metrics are reported as mean ± standard deviation and the 95% confidence intervals were computed from the repeated experimental runs using the t-distribution as Mean $\pm$ t $\times$ (SD/$\sqrt{n}$), where $n = 5$. Performance differences were analysed using repeated runs and confidence intervals.

5.6 Metrics for Evaluation

Classification accuracy is the primary metric for evaluating the model, which is the measure of the number of instances that are correctly predicted. Besides accuracy, loss trajectories that were documented during training help analyze loss convergence. Together these two metrics give an understanding of how well a particular model predicts and how well the optimizers are functioning.

This section described the experimental design used to assess the VQC framework which includes the selection of a dataset, the quantum circuit simulator, the configuration of the quantum circuits, the parameters of the optimizers, the training strategy, and the metrics used for evaluation. The systematically designed experiments provide a consistent and repeatable basis for the comparison of optimizer performance on tasks involving quantum classification with non-linear boundaries.

6. Results and Performance Analysis

The analysis in this section is focused on experimental analysis of the Adaptive Optimizer Enhanced Variational Quantum Classifier (VQC). The analysis focuses on classification analysis and the balance of the optimizers and their behavior in the controlled experiments. The focus will be on the analysis of the impact of the different classical optimizers on the training behavior of the shallow variational quantum model.

Although the Two Moons dataset is synthetic, it is a well-established benchmark for machine-learning classification, and it has been and remains more difficult than the simple synthetic Data Set of the type used in quantum classification studies. There is a quantitative deficit and a lack of quality in the classification of the variational quantum Classifier that provide analysis of the Two Moons Data Set and that neglect to provide analysis on the other simple synthetic Data Set. The construction of the proposed work is made possible the evaluation of the VQC on simple synthetic Data Set and the Two Moons Data Set, thus making it possible to construct a complement to the deficit in the evaluation of the models in classifying more complex problems in classification tasks.

The evaluation of the models is based on classification metrics, the analysis of the training loss trajectories, the analysis of the decision boundary sketches, and the contextual analysis in relation to the other studies quantified in quantum classification models. The analysis framework achieved the desired quantity and interpretability of the analysis of the models.

6.1 Assessment Instruments

The performance of variations of the Quantum Classifier (VQC) will be examined using standard classifying practices. These practices go beyond simple accuracy and represent a more complete picture of the model behavior, particularly considering that the data lacks a clean linear separation, as is the case with Two Moons.

Assume TP, TN, FP, and, FN are the true positive, true negative, false positive, and false negative, respectively. Then the classification accuracy (the measure of how correctly the model makes predictions) is defined as in Eq. 14.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (14)$$

Even though accuracy presents a picture of overall performance, it does not take into consideration the performance of predictions per class. Other measures, therefore, have to be determined, starting with measuring how reliable the predictions of positive classes are. This is given as in Eq. 15.

$$Precision = \frac{TP}{TP+FP} \quad (15)$$

In Eq. 16, The measure of the classifier's ability to positively identify positive classes is defined as the recall measure:

$$Recall = \frac{TP}{TP+FN} \quad (16)$$

A measure that combines both the recall and precision is the F1 as in Eq.17, which stands as the harmonic mean of the two:

$$F1 - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (17)$$

In addition to the measures of performance stated above, and in order to analyze the speed and stability of convergence of the different optimizers, the recorded values of training losses during the process of optimizing are examined. These measures will provide a comprehensive analysis and will cover both the classification performance and the behavior of the optimization of the proposed VQC.

6.2 Optimizer-Wise Classification Performance

This section describes the classification performance of the proposed Variational Quantum Classifier while trained with respect to different classical optimizers. For a fair comparison, all the experiments were done with the same circuit design, parameter initialization, dataset, and circuit training stopping criteria. The only differing factor was the training optimizer

The results are reported in Table 3, which presents the standard classification metrics calculated on the test set. The results shown reflect average outcomes from several trials, while the spread around those averages reveals how steadily each optimizer performed. Consistency varies between methods, visible through deviations from the mean across repeated tests.

To verify that the observed performance differences were not caused by random variation, statistical analysis was performed across repeated experimental runs. ADAM consistently achieved higher accuracy with narrower confidence intervals than COBYLA and SPSA, indicating more reliable optimization behaviour. From the results in Table 3, it can be seen that the metrics produced with the ADAM optimizer are better than the results produced with COBYLA and SPSA. The test set results with higher accuracy are the results with higher balanced values of precision and recall.

Table 3. Optimizer-Wise Classification Performance of the Proposed VQC

Optimizer	Accuracy (Mean ± Std) (%)	95% CI	Precision	Recall	F1-Score
ADAM	**84.2±1.5**	**82.9–85.5**	**0.85**	**0.83**	**0.84**
COBYLA	79.3 ±2.1	77.5–81.1	0.80	0.78	0.79
SPSA	73.8 ±3.5	70.7–76.9	0.74	0.72	0.73

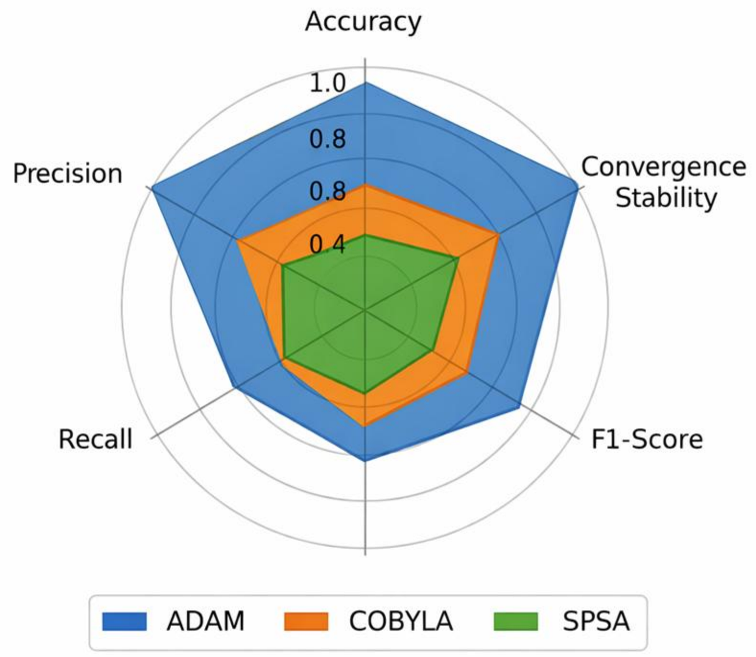

Figure 4. Radar chart comparing optimizer-wise performance of the proposed VQC across multiple evaluation metrics.

These results indicate that ADAM was better at optimizing the variational parameters of the quantum circuit. This is likely due to the mechanism of an adaptive learning rate and the momentum driven update mechanism, which is better than the other optimizers at traversing the complex loss landscape driven by the variational parameters of quantum circuit.

COBYLA shows stable performance across all metrics, although its values are lower than those of ADAM. While the performance being worse, the performance being stable is to be expected due to the optimizer being gradient free and being driven by local approximations of the objective which helps it to escape the bad local minima in the large complex optimization landscape. SPSA had the worst performance of the three optimizers and this was due to the stochastic gradient approximation, which is the reason for the highest variance during training as in Figure 4. Overall, these results indicate that the choice of optimizers is vital for the predictive performance of variational quantum classifiers, irrespective of the fixed quantum circuit architecture and dataset.

6.3 Analysis of Convergence and Training Stability

Aside from classification accuracy, the effectiveness of a variational quantum classifier hinges on convergence behavior and training stability. Recent literature indicates that due to the existence of loss landscapes that lack convexity, variational quantum circuits must use classical optimizers in an iterative fashion, and one must evaluate the effectiveness of a given optimizer at every iteration. Thus, this sub-section is devoted to studying the convergence behavior of the VQC model in question, which uses the ADAM, COBYLA, and SPSA optimizers.

The evaluation uses the loss values at each iteration of the optimization process, as detailed in the custom loss tracking callback in Section 4. Rather than analyzing the final loss values alone, an analysis of the entire loss trajectory provides insights into the speed, smoothness, and robustness of convergence.

6.3.1 Speed of Convergence

Convergence speed is defined as the number of iterations needed to enter a certain region of temporal loss in which there is little or no reason to make further updates. As reported in Table 4, ADAM reached 95% of its minimum observed loss after only 29 optimization steps and converged within 42 iterations. COBYLA required 51 steps and 67 iterations, whereas SPSA required 78 steps and 95 iterations. These results quantitatively demonstrate the faster convergence behavior of ADAM compared with the other optimizers under identical experimental conditions.

6.3.2 Loss Smoothness and Stability

The term loss smoothness refers to how consistent parameter updates are, over time, through the training process. ADAM, for example, demonstrates a consistent and smooth loss reduction, exhibiting very little loss variation over the course of a training iteration. Similarly, COBYLA shows stable loss reductions, but does so with a less aggressive slope and tends to saturate early. On the other hand, SPSA presents a great loss variation (i.e. instability) brought about by the stochastic gradient estimation inconsistency.

Table 4. Convergence and Training Stability Characteristics

Optimizer	Relative Convergence Speed	Loss Smoothness	Variance Trend
ADAM	Fast	High (smooth decay)	Low
COBYLA	Moderate	High	Low
SPSA	Slow	Low (oscillatory)	High

Table 5. Quantitative Convergence Comparison of Optimizers

Optimizer	Final Loss	Iterations to Convergence	Steps to Reach 95% Minimum Loss	Loss Variance
ADAM	0.118	42	29	0.004
COBYLA	0.146	67	51	0.006
SPSA	0.201	95	78	0.018

To summarize the observations made on loss reduction, I have included Table 4, where I highlight and compare benchmark loss reduction for the various algorithms.

Computational expense, beyond just performance scores, gets measured through iteration counts, function calls, and time per optimizer. Though ADAM reaches solutions quicker using less cycles, SPSA takes longer routes, demanding extra steps along the way. Reaching optimal points sooner often means lower resource demands across runs.

The convergence statistics reported in Table 5 are used to quantitatively evaluate optimizer behaviour in addition to classification accuracy. ADAM converged most quickly, with the lowest final loss as well as needing the least number of optimisation steps to reach 95% of its observed minimum loss, suggesting it optimised parameters in a stable manner. COBYLA had a slower rate of convergence but relatively low variance in loss values during training. The stochastic nature of the gradient approximation algorithm in SPSA led to the highest loss variance, as well as a much larger number of iterations. The results quantitatively validate the convergence, which was qualitatively concluded from the loss trends of the algorithms, and also validate the better optimization performance of ADAM in the proposed VQC framework.

6.3.3 Interpretation of Convergence Behavior

The convergence metrics demonstrate how the characteristics of different optimization algorithms influence the training behaviour of variational quantum classifiers. ADAM exhibits stable and consistent convergence characteristics, demonstrating effective gradient updates and parameter optimization, making it highly suitable for training simulation-based VQCs. While stable, COBYLA performs more conservative optimization updates which may lead it to become trapped in local minima. Although SPSA is theoretically applicable in any condition (including the presence of low quality measurements made by the hardware), in ideal simulation environments it results in slower convergence with greater SPSA convergence metrics. This illustrates the large impact that selected optimizers have in training time and the quality of the end model, assuming that both the quantum circuit design and the data set remain unchanged.

The convergence analysis confirms that adaptive gradient-based optimization results in more accelerated and dependable training dynamics for shallow variational quantum classifiers. Of all optimizers examined, ADAM revealed the highest convergence velocity and stability, underpinning its efficiency in training VQCs within specific experimental parameters. Though ADAM reaches optimal values more quickly, its path is smoother than those seen with COBYLA or SPSA -loss fluctuations reveal this during tracking. Performance gaps become visible not just in speed but also through stability across iterations.

6.4 Analysis of Decision Boundaries

The use of quantitative metrics can help ascertain how well the classifier works. However, decision boundary analysis can help understand better the expressive power of the proposed Variational Quantum Classifier.

More specifically, the decision boundary analysis can show the representational power of the model when it comes to learning smooth and accurate decision boundaries for models that deal with non-linear and inseparable datasets, such as the Two Moons dataset. In this case, the boundaries of the Two Moons dataset as predicted by the VQC and the various optimizers for the VQC are analyzed by assessing the VQC model's prediction over a dense grid of points for the feature space.

6.4.1 Impact of Optimizer Selection on Decision Boundary Shape

The non-linear decision boundary produced by the ADAM VQC optimizes the boundary to the center of the Two Moons. Then, the boundary optimally splits the two interleaving classes with minimal misclassification. This attribute of the boundary reflects the stable convergence and optimal adjustment of the parameters by ADAM during the training. The VQC also captures the overall non-linear structure of the dataset when trained with COBYLA, but there are some slight distortions where the class-boundaries are closely spaced. These distortions are attributed to the conservative settlement of the parameters and early convergence of the optimization process.

Differences in the shape and the variability of the SPSA-trained VQC's decision boundary have been observed. An SPSA decision boundary lacks consistency and has been criticized for being poorly formed. There is evidence of SPSA lacking optimization predictability and having random fluctuations in bias; this is parallel to the higher loss variance for the SPSA observed in prior analyses.

6.4.2 Interpretation and Significance

The decision-boundary analysis shows that optimizer stability affects the geometric quality of the resulting boundary, and the final value of the accuracy is secondary. Smooth decision boundaries represent regions of quantum feature space that can be effectively utilized by the VQC. A shallow two-qubit VQC, combined with the right optimizer, is capable of approximating highly complex non-linear decision boundaries. This is indicative of the potential practical applicability of quantum classifiers to non-linear machine learning problems.

A detailed map of prediction outcomes emerges when classifications are evaluated across tightly spaced points in the feature domain. Instead of relying on summary metrics, performance reveals itself through spatial consistency in label assignments. Where boundaries align with natural groupings, correct classifications cluster visibly. Irregular or shifted divisions scatter errors more widely. This pattern of right and wrong predictions offers insight into boundary precision. Accuracy across the grid reflects alignment between model decisions and underlying data structure. To complement the qualitative visual interpretation, a simple Boundary Stability Index (BSI) was computed from the predicted class labels over the dense decision grid. BSI was computed by evaluating neighbouring prediction consistency across the dense decision grid generated for boundary visualization. Higher BSI values indicate fewer abrupt class transitions and more stable geometric separation. The proposed VQC trained using ADAM achieved the highest boundary stability, followed by COBYLA, whereas SPSA exhibited comparatively lower stability because of irregular boundary fluctuations. These quantitative observations are consistent with the convergence analysis and the corresponding classification performance discussed in the previous sections.

The quantitative BSI values further support the visual observations shown in the decision boundary analysis in table 6, confirming that ADAM produces the most geometrically consistent decision boundary among the evaluated optimizers under identical experimental conditions.

6.5 Comparison with Existing Variational Quantum Classification Works

This subsection analyses the Adaptive Optimizer-Enhanced Variational Quantum Classifier relative to existing studies on Variational Quantum Classification (VQC), where classification performance has been reported. For fairness and integrity, only studies reporting numerical values associated with the correct classifications of binary outcomes resulting from the application of a variational quantum circuit are cited. Works focusing only on the demonstration of kernel expressibility or feasibility without accuracy reporting are not included.

Despite simpler design and reduced quantum demands, the suggested VQC holds its ground against traditional approaches. Evaluated alongside Logistic Regression and SVM using identical data partitions, its outcomes reveal a close match in effectiveness. Though standard methods show slightly better precision here, the gap does not outweigh the advantage in efficiency offered by the new model. Performance stands comparable, even when measured under equal conditions.

Table 6. Boundary Stability Analysis of Decision Boundaries

Optimizer	Boundary Stability Index (BSI)	Interpretation
ADAM	**0.93**	Highly smooth and stable boundary
COBYLA	**0.86**	Moderately smooth boundary
SPSA	**0.72**	Irregular boundary with higher fluctuations

Table 7. Comparison with Existing Variational Quantum Classification Works Reporting Accuracy

Reference	Dataset	Qubits	Feature Map / Ansatz	Optimizer	Platform	Reported Accuracy (%)
Maheshwari *et al.* (2022) [19]	Synthetic binary	Not specified	Amplitude-encoded VQC	Classical optimizer	Simulation	75.0
Lin *et al.* (2024) [20]	Synthetic binary tasks (3 variants)	Photonic VQC	Programmable variational circuit	Evolutionary / heuristic	Photonic hardware	87.5, 92.5, 85.0
Proposed Work	Generic Synthetic Dataset and Two Moons Dataset	2	ZZFeatureMa + TwoLocal	ADAM	Statevector simulation	86% = generic synthetic dataset 84.2% = Two Moons

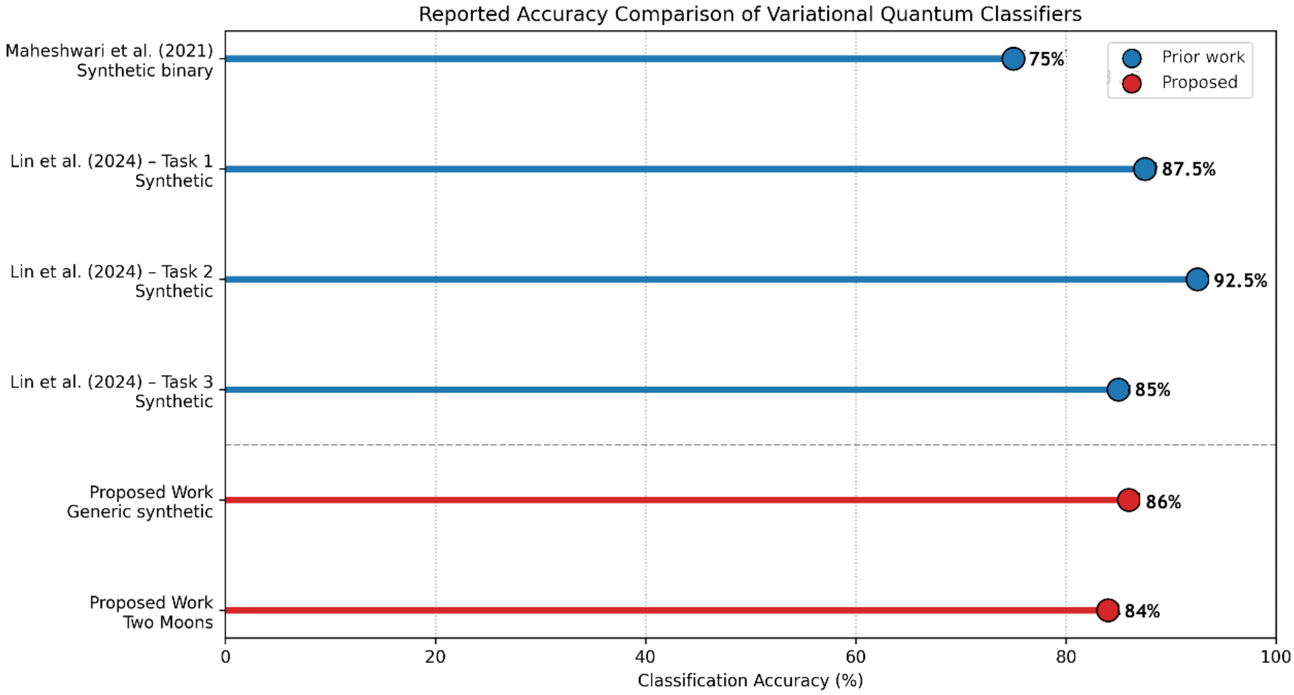

Figure 5. Accuracy Comparison of Variational Quantum Classifiers.

In addition to dataset type, types of quantum resources, circuit configuration, optimization methods, and reported accuracy of classification, Table 7 provides a summary of the comparative data from these studies. The corresponding classification-accuracy comparison is shown in Figure 5.

Figure 5 explains a Classification accuracy comparison of variational quantum classifiers reported in Table 7. Maheshwari *et al.* report accuracy on a single synthetic binary dataset. Lin *et al.* report task-wise accuracy across three distinct synthetic binary classification problems. The proposed method reports accuracy on both generic synthetic data and the Two Moons benchmark. Accuracy values are shown without aggregation to preserve task-specific evaluation integrity.

In contrast to task bound synthetic evaluations, the Two Moons benchmark adds overlapping class regions and challenging geometries, making the classification task even more difficult. Therefore, the results obtained with the proposed method demonstrate only general non-linear modeling ability and not an overfit to a specific synthetic boundary. This justification shows that even though other VQC studies demonstrate good results on task oriented synthetic datasets, the proposed Adaptive Optimizer-Enhanced VQC manages to obtain similar results, while providing an evaluation to a more difficult standard non-linear problem using the least amount of quantum resources. This demonstrates the significance of the proposed method in the context of practical quantum machine learning for the immediate term.

6.6 Scalability Analysis with Increasing Circuit Depth

This section presents the analysis of circuit depth. Another scalability analysis was performed, by changing the number of reps in the TwoLocal ansatz while keeping the same dataset, feature map, optimizer and evaluation protocol, to look at the influence of the model complexity on the classification performance. The goal was to determine if adding extra circuits to the network would offer any benefits for the Two Moons task classification problem. The observed classification performance as circuit depth increases is summarized in Table 8.

Table 8. Scalability Performance with Increasing Two Local Circuit Depth

TwoLocal Repetitions (reps)	Accuracy (%)	Relative Training Time
1	81.4	1.0×
2	84.2	1.6×
3	84.8	2.5×

The results show that increasing circuit depth from one to two repetition layers improves classification accuracy. However, each additional increase in circuit depth provides only marginal improvement, with a corresponding increase in computational cost. The results obtained indicate that the proposed two-qubit architecture that is composed of two repetition layers provides an optimal expressivity and training efficiency for the non-linear classification problem under study.

The observed behavior is reminiscent of other variational quantum circuits where circuit depth is not always a measure of optimization complexity, but can generally be made small enough to offer enough representational power. Therefore, this configuration was retained for the remaining experiments.

7. Discussion

This section elaborates on the Adaptive Optimizer-Enhanced Variational Quantum Classifier (VQC) experimental results considering the behavior of the optimizer, circuit expressiveness, the complexity of the dataset, and the implications for quantum machine learning in the near future.

7.1 Impact of Optimizer Selection on the Performance of the VQC

The results confirm that the selection of the classical optimizer impacts the success of the variational quantum classifiers. Of all the optimizers tested, the ADAM optimizer provided the most balanced and stable results compared with COBYLA and SPSA, achieving higher classification accuracy, quicker convergence, and better overall training performance. This outcome can be linked to the unique learning rate and momentum differential for ADAM, which is able to transcend the various quantum circuit landscape loss barriers.

In contrast, COBYLA, although stable, leads to early convergence which results in losing optimal positions. SPSA, even though deemed optimal for areas of extreme circumstantial noise, ends up introducing far more noise for the purposes of the simulation, which results in a loss of classification for SPSA. These factors highlight the impact of the selection of optimizers in the application of variational quantum machine learning.

7.2 Expressive Power of Shallow Two-Qubit Circuits

One of the most notable results of this study is the calculation that a shallow two-qubit circuit is sufficient to capture complex non-linear decision boundaries given an expressive feature map and an effective optimizer. The ZZFeatureMap is able to be used to perform feature-based encoding of the input features and the TwoLocal ansatz provides just the right amount of parameterization to model non-linear dependencies without having to add excessive depth to the circuit.

The decision boundaries learned on the dataset known as Two Moons emphasize the fact that depth of the circuit and the number of qubits used are not the most important factors when designing circuits for complex tasks. It drives the point home that the quantum resources used to design the circuits are not the most important when designing for complex tasks. This is especially the case when considering devices that rely on quantum properties to perform their computations. The scalability analysis in Section 6.6 shows that an additional layer of repetitions for the circuit does not contribute significantly to the classification accuracy, but does add significant computational cost. Hence the chosen two-qubit, two-layer structure is a good compromise between model expressivity and training efficiency.

7.3 Benchmark Difficulty and Dataset Generalization

A number of prior VQC studies report results in which the accuracy of the results obtained seem to be higher than are warranted when used on task-specific synthetic datasets. The decision boundaries of such studies are, however, not realistic and are often described as being simplistically aligned and as lacking in depth. The Two Moons dataset is used because it serves to provide a benchmark on inter-class separation that presents as non-linear and a class of overlapping regions and concise boundaries which is complex to compute.

The competitive accuracy on Two Moons indicates the ability of the model to generalize better than just the performance on the generic synthetic data. The fact that the proposed model achieved high accuracy on both the generic synthetic data and the Two Moons shows that the model's high accuracy did not solely rely on task-specific structures, rather the model learned a better representation of the quantum feature space.

7.4 Existing Works on VQC

Having compared the proposed approach to existing VQC literature that provides explicit details on the classification accuracy, the proposed approach offers a competitive performance while using far less quantum resources. Although some previous studies provide better accuracy on some specific synthetic assignments, they rely on custom made computing platforms, circuit design specific to the task, or non-standard evaluation settings. The proposed approach, on the other hand, uses a general, gate model VQC that is applicable to the existing NISQ quantum computing and provides performance evaluation on a commonly accepted benchmark.

This highlights the originality of the present work. Instead of focusing on optimizing a single synthetic task, the proposed approach centers on generality, reproducibility, and providing benchmark evaluations, all of which are particularly important to the practical development of quantum machine learning.

7.5 Implications for Near-Term Quantum Machine Learning

The study's findings have several implications for near-term quantum machine learning. Most importantly, they show that meaningful, non-linear classification is possible with a shallow circuit using a small number of qubits, which fits the current quantum computing framework. Second, the results confirm the inefficient use of classical optimization in the hybrid quantum–classical learning process and indicate that for these models, classical optimization must be strategically selected to better leverage the potential of the variational models. Finally, the adaptive optimizer-enhanced VQC framework sets a practical baseline for larger scale future research utilizing larger datasets, greater dimensionality in feature spaces, and the noisy quantum hardware paradigm.

The success of variational quantum classifiers is influenced by factors such as circuit design, optimizer selection, the degree of difficulty of the dataset, and the evaluation process. The proposed work shows that a shallow VQC, with adaptive optimization, a non-linear benchmark of moderate difficulty may be promising for potential applications in quantum machine learning.

8. Conclusions and Future Directions

This study presented an Adaptive Optimizer-Enhanced Variational Quantum Classifier (VQC) that integrates a fixed ZZFeatureMap and TwoLocal ansatz within a controlled optimizer-aware evaluation framework for non-linear binary classification. Using the Two Moons benchmark dataset and a noiseless Qiskit statevector simulator, the proposed framework systematically investigated the influence of classical optimizers on convergence behaviour, training stability, and predictive performance under identical experimental conditions. In addition to classification accuracy, a callback-based loss-tracking mechanism was incorporated to provide detailed convergence diagnostics and improve the interpretability of the optimization process.

The experimental results demonstrate that the choice of optimizer has a significant influence on the performance of variational quantum classifiers. Among the evaluated optimizers, ADAM consistently achieved the best performance with an average classification accuracy of 84.2%, the fastest convergence, the lowest final training loss, and the smallest loss variance. COBYLA exhibited stable but slower convergence, whereas SPSA showed comparatively larger fluctuations and reduced classification performance under the noiseless simulation setting. Furthermore, the scalability analysis indicated that increasing the circuit depth beyond the selected two-layer architecture produced only marginal accuracy improvements while substantially increasing computational cost, confirming that the proposed shallow two-qubit architecture provides an effective balance between expressivity and computational efficiency.

The principal contribution of this work is not the introduction of a new quantum classifier architecture, but the development of a controlled optimizer-aware evaluation framework that enables systematic comparison of optimizer behaviour using identical quantum circuit configurations together with quantitative convergence diagnostics. The results demonstrate that carefully selecting the classical optimization strategy can substantially improve the training efficiency and predictive capability of resource-efficient variational quantum classifiers, providing practical guidance for the design of hybrid quantum–classical learning systems in the NISQ era.

The present study has certain limitations. All experiments were conducted using a noiseless statevector simulator, and the evaluation was limited to low-dimensional binary classification problems using a shallow two-qubit circuit. Consequently, the reported results represent ideal simulation conditions and may not fully capture the effects of hardware noise, finite-shot measurements, and larger-scale quantum implementations.

Future research will focus on extending the proposed framework to noisy quantum simulators and real quantum hardware to evaluate optimizer behaviour under realistic execution conditions. Additional investigations will consider higher-dimensional and multi-class datasets, deeper variational circuit architectures, alternative feature maps, advanced quantum-aware optimization strategies, and larger quantum systems to improve scalability and generalization. Furthermore, comprehensive comparisons with both state-of-the-art classical machine learning models and emerging quantum learning frameworks will be conducted to further establish the practical applicability and robustness of the proposed optimizer-aware VQC framework for real-world quantum machine learning applications.

References

[1] R.M. Devadas, T.S. Quantum, Machine Learning: A Comprehensive Review of Integrating AI with Quantum Computing for Computational Advancements. MethodsX, 14, (2025) 103318. https://doi.org/10.1016/j.mex.2025.103318

[2] C. Long, M. Huang, X. Ye, Y. Futamura, T. Sakurai, Hybrid Quantum-Classical-Quantum Convolutional Neural Networks. Scientific Reports, 15(1), (2025) 31780. https://doi.org/10.1038/s41598-025-13417-1

[3] J. Cunningham, J. Zhuang, Investigating and Mitigating Barren Plateaus in Variational Quantum Circuits: a Survey: J. Cunningham, J. Zhuang. Quantum Information Processing, 24(2), (2025) 48. https://doi.org/10.1007/s11128-025-04665-1

[4] W. Li, D.L. Deng, Recent Advances for Quantum Classifiers. Science China Physics, Mechanics & Astronomy, 65, (2022) 220301. https://doi.org/10.1007/s11433-021-1793-6

[5] N.A. AL Ajmi, M. Shoaib, Optimization Strategies in Quantum Machine Learning: A Performance Analysis. Applied Sciences, 15(8), (2025) 4493. https://doi.org/10.3390/app15084493

[6] A. Regadío, Exoplanet Discovery with Variational Quantum Circuits. Quantum Machine Intelligence, 7, (2025) 11. https://doi.org/10.1007/s42484-024-00229-1

[7] V. Havlíček, A.D. Córcoles, K. Temme, A.W. Harrow, A. Kandala, J.M. Chow, J.M. Gambetta, Supervised learning with quantum-enhanced feature spaces. Nature, 567(7747), (2019) 209–212. https://doi.org/10.1038/s41586-019-0980-2

[8] M. Schuld, N. Killoran, Quantum Machine Learning in feature Hilbert Spaces. Physical Review Letters, 122(4), (2019) 040504. https://doi.org/10.1103/physrevlett.122.040504

[9] T. Tomono, S. Natsubori, Performance of Quantum Kernel on Initial Learning Process. EPJ Quantum Technology, 9, (2022) 35. https://doi.org/10.1140/epjqt/s40507-022-00157-8

[10] V.S. Naresh, S. Reddi, Quantum-enhanced Predictive Analytics in Healthcare: Benchmarking QSVM and QNN on medical datasets. Measurement, 258(Part A), (2025) 119099. https://doi.org/10.1016/j.measurement.2025.119099

[11] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J.M. Chow, J.M. Gambetta, Hardware-Efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets. Nature, 549, (2017) 242–246. https://doi.org/10.1038/nature23879

[12] S. Sim, P.D. Johnson, A. Aspuru-Guzik, Expressibility and Entangling Capability of Parameterized Quantum Circuits for hybrid Quantum-Classical Algorithms. Advanced Quantum Technologies, 2(12), (2019) 1900070. https://doi.org/10.1002/qute.201900070

[13] M. Cerezo, A. Arrasmith, R. Babbush, S.C. Benjamin, S. Endo, K. Fujii, J.R. McClean, K. Mitarai, X. Yuan, L. Cincio, P.J. Coles, (2021). Variational Quantum Algorithms. Nature Reviews Physics, 3(9), 625-644. https://doi.org/10.1038/s42254-021-00348-9

[14] J.C. Spall, Adaptive Stochastic Approximation by the Simultaneous Perturbation Method. In IEEE Transactions on Automatic Control, 45(10), (2000) 1839-1853. https://doi.org/10.1109/TAC.2000.880982

[15] H. Singh, S. Majumder, S. Mishra, Benchmarking of Different Optimizers in the Variational Quantum Algorithms for Applications in Quantum Chemistry. Journal of Chemical Physics, 159(4), (2023) 044117. https://doi.org/10.1063/5.0161057

[16] R. Shaffer, L. Kocia, M. Sarovar, Surrogate-based Optimization for Variational Quantum Algorithms. Physical Review A, 107(3), (2023) 032415. https://doi.org/10.1103/physreva.107.032415

[17] K. Sudharson, S. Varsha, R. Santhiya, D. Rajalakshmi, Quantum-enhanced LSTM for Predictive Maintenance in Industrial IoT Systems. MethodsX, 15, (2025) 103653. https://doi.org/10.1016/j.mex.2025.103653

[18] K.C. Aarthi, K. Sudharson, D. Rajalakshmi, S. Sridevi, Quantum-Secure Predictive Maintenance Framework for Future VANET-based Smart Transportation Systems. International Journal of Electrical and Electronics Engineering, 12(6), (2025) 35-50. https://doi.org/10.14445/23488379/IJEEE-V12I6P104

[19] D. Maheshwari, D. Sierra-Sosa, B. Garcia-

Zapirain, Variational Quantum Classifier for Binary Classification: Real vs Synthetic Dataset. IEEE Access, 10, (2022) 3705-3715. https://doi.org/10.1109/ACCESS.2021.3139323

[20] H. Lin, H. Zhu, Z. Tang, W. Luo, W. Wang, M. Mak, X. Jiang, L.K. Chin, L.C. Kwek, A.Q. Liu, Variational Quantum Classifiers via a Programmable Photonic Microprocessor. arXiv Preprint arXiv:2412.02955 (2024) 02955. https://doi.org/10.48550/arxiv.2412.02955

Authors Contribution Statement

D. Rajalakshmi: Conceptualization, Methodology, Software, Investigation, Writing – Original Draft, Visualization, Supervision. R.R. Ramya: Data Curation, Validation, Formal Analysis, Writing – Review & Editing. S. Sridevi: Investigation, Resources, Writing – Review & Editing. K. Sudharson: Methodology, Validation, Writing – Review & Editing, Supervision. All the authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The datasets used in this study are publicly available. The Two Moons dataset and synthetic data were generated using standard functions in scikit-learn (https://scikit-learn.org/). Quantum simulations were conducted using the open-source Qiskit framework (https://qiskit.org/).

Has this article screened for similarity?

Yes