AN INTELLIGENT QUANTUM-ENHANCED HYBRID LEARNING FRAMEWORK FOR ADAPTIVE PHISHING WEBSITE DETECTION AND CYBER THREAT ANALYSIS

Dr. M. Dattatreya Goud
CEO, DCode Soft Tech Information Pvt. Ltd.,Hyderabad, India.
Dattatreya548@gmail.com

Abstract:
The increasing sophistication of phishing attacks creates significant challenges for conventional cybersecurity detection systems. Traditional machine learning approaches may struggle to identify evolving phishing patterns and previously unseen malicious websites. This paper proposes an Adaptive Hybrid Quantum-Classical Learning Framework for intelligent phishing website detection and cyber threat classification. The proposed framework combines classical machine learning algorithms, including Random Forest, Support Vector Machine, and XGBoost, with quantum machine learning techniques such as Variational Quantum Classifier (VQC) and Quantum Kernel Support Vector Machine (QSVM). URL-based and website-related features, including URL length, domain characteristics, HTTPS usage, suspicious symbols, subdomain patterns, and redirection indicators, are extracted during the preprocessing stage. The selected features undergo normalization, feature selection, and dimensionality reduction to improve classification efficiency. Quantum feature maps transform selected classical features into quantum representations for classification. The framework evaluates the performance of classical and quantum models using accuracy, precision, recall, F1-score, false-positive rate, computational time, and quantum noise robustness. The study investigates the effectiveness and limitations of hybrid quantum-classical learning for improving phishing detection accuracy, reducing classification errors, and supporting reliable and adaptive cybersecurity solutions.

Keywords: Phishing Website Detection, Quantum Machine Learning, Hybrid Quantum-Classical Learning, Variational Quantum Classifier, Cyber Threat Classification, Quantum Kernel Support Vector Machine.

I. INTRODUCTION

The rapid growth of internet services, online transactions, digital communication, and e-commerce applications has significantly increased the risk of cyber threats, particularly phishing attacks. Phishing is a cyberattack technique in which attackers create fraudulent websites, manipulate web links, or imitate trusted online platforms to deceive users into revealing sensitive information such as login credentials, banking details, and personal data. These attacks are becoming increasingly sophisticated through modified URLs, misleading domain names, suspicious redirections, fake login pages, and imitation of legitimate websites. As a result, detecting phishing websites accurately has become an important requirement for improving online security and protecting user information.

Traditional phishing detection techniques mainly depend on blacklist-based filtering, rule-based identification, and conventional machine learning algorithms. However, these approaches may struggle to identify newly generated phishing websites and previously unseen attack patterns. Machine learning algorithms such as Support Vector Machine (SVM), Random Forest, and XGBoost have been widely used because they can identify suspicious patterns from website-related features. Nevertheless, their performance may be affected by complex feature relationships, changing attack strategies, and variations in website characteristics.

Quantum machine learning has emerged as an alternative research direction that combines quantum computing principles with machine learning techniques. Models such as Variational Quantum Classifier (VQC) and Quantum Kernel Support Vector Machine (QSVM) provide new

possibilities for representing complex data relationships through quantum feature encoding and quantum circuits. However, quantum models also face challenges related to noise, computational resources, and hardware limitations.

This research proposes an Adaptive Hybrid Quantum-Classical Learning Framework for intelligent phishing website detection and cyber threat classification. The framework integrates classical algorithms with quantum learning techniques to improve classification reliability. It includes dataset collection, data preprocessing, URL feature extraction, feature selection, quantum feature mapping, and classification. Random Forest, SVM, XGBoost, VQC, and QSVM are considered for comparative evaluation. The proposed framework is assessed using accuracy, precision, recall, F1-score, false-positive rate, computational time, and noise robustness. The study aims to investigate the effectiveness, adaptability, and practical limitations of hybrid quantum-classical learning for developing intelligent cybersecurity solutions.

II. LITERATURE SURVEY

Recent research has focused on developing intelligent techniques for detecting phishing websites and classifying cyber threats. Traditional machine learning approaches use website-related features such as URL structure, domain information, page content, and security characteristics to distinguish legitimate websites from malicious ones. Feature selection techniques help identify important attributes and reduce unnecessary data. Ensemble learning methods improve classification performance by combining multiple machine learning algorithms. Deep learning approaches have also been explored to identify complex patterns and suspicious website behaviour.

Quantum machine learning has recently gained attention as an emerging technology for solving complex classification problems. Various quantum classification techniques have been developed using quantum circuits, feature encoding, and optimization methods. Variational Quantum Classifiers use parameterized quantum circuits to learn patterns from input data through classical optimization. Quantum Kernel Support Vector Machines use quantum feature maps to identify complex relationships between data samples. Quantum neural networks and other quantum learning models have also been investigated for improving classification capabilities.

Existing research has examined the effects of quantum feature encoding, circuit design, computational complexity, and quantum noise on classification performance. However, quantum machine learning still faces challenges related to hardware limitations, noise sensitivity, and execution time. Furthermore, many existing phishing detection approaches mainly depend on classical machine learning techniques, while quantum approaches are generally studied separately. Limited attention has been given to combining classical and quantum learning methods in an adaptive framework specifically for phishing website detection. Therefore, the proposed work focuses on integrating classical machine learning and quantum classification techniques to improve phishing detection accuracy, reduce false positives, and support reliable cyber threat classification.

III. PROPOSED WORK

The proposed work introduces an Adaptive Hybrid Quantum-Classical Learning Framework for intelligent phishing website detection and cyber threat classification. The main objective is to develop an effective detection approach by combining classical machine learning algorithms with quantum machine learning techniques. The framework is designed to identify phishing websites by analyzing URL characteristics, domain information, and website-related features while investigating the potential benefits of quantum-based classification.

Initially, labeled phishing URL datasets containing both legitimate and malicious website samples are collected from publicly available sources. During the preprocessing stage, duplicate records, missing values, and irrelevant attributes are handled to improve data quality. Important features such as URL length, domain age, HTTPS availability, special characters, subdomain count, IP address usage, and URL redirection indicators are extracted. Feature selection and dimensionality reduction are then performed to identify relevant attributes and reduce computational complexity.

In the classical learning stage, Support Vector Machine (SVM), Random Forest, and XGBoost algorithms are implemented to establish baseline classification performance. In the quantum learning stage, selected features are encoded into quantum states using parameterized quantum feature maps. Variational Quantum Classifier

(VQC) and Quantum Kernel Support Vector Machine (QSVM) are employed to learn classification patterns and distinguish phishing websites from legitimate websites.

The adaptive component investigates different feature-selection methods, quantum circuit configurations, circuit depths, and model hyperparameters to identify suitable configurations for improved classification performance. The framework also examines the impact of simulated quantum noise and evaluates model generalization using previously unseen website domains.

Finally, the proposed framework is evaluated using accuracy, precision, recall, F1-score, false-positive rate, training time, and inference cost. Comparative analysis is performed between classical and hybrid quantum-classical models to understand their advantages and limitations. The implementation uses Python, Qiskit, and Scikit-learn in a simulated quantum computing environment. The proposed work aims to provide an experimentally validated framework for exploring adaptive quantum-classical learning in phishing website detection and intelligent cybersecurity applications.

IV. METHODOLOGY

The proposed methodology develops an Adaptive Hybrid Quantum-Classical Learning Framework for phishing website detection and cyber threat classification. It involves dataset collection, preprocessing, feature extraction, classical model development, quantum feature encoding, hybrid classification, and performance evaluation. Classical and quantum learning algorithms are integrated to investigate their classification capabilities. The adaptive component examines different feature-selection techniques and quantum circuit configurations to identify suitable model settings.

a. Dataset Collection

The first stage involves collecting phishing and legitimate website URL datasets from publicly available cybersecurity sources. The dataset contains labeled website samples representing malicious phishing URLs and genuine websites. Each sample includes relevant information about URL structure, domain characteristics, and website security properties. The collected dataset is examined to understand its size, feature distribution, and class balance. Duplicate records and incomplete samples are identified to maintain data consistency. The dataset is divided into training and testing subsets to evaluate the learning performance of different classification models. Training data is used to develop the models, while testing data is used to measure their ability to classify previously unseen website samples. A separate validation set can also be used for parameter optimization. This dataset provides the foundation for training and evaluating the proposed hybrid quantum-classical phishing detection framework.

b. Data Preprocessing and Feature Extraction

Data preprocessing is performed to improve dataset quality and prepare the input features for classification. Missing values, duplicate records, and irrelevant attributes are handled during this stage. URL-related characteristics are extracted to identify suspicious website patterns. Important features include URL length, domain age, HTTPS availability, special character usage, IP address presence, subdomain count, and redirection indicators. These features help distinguish legitimate websites from phishing websites. Numerical attributes are normalized to maintain consistent feature scales and improve model performance. Feature selection techniques are applied to identify the most relevant attributes and remove unnecessary information. Dimensionality reduction is also performed to decrease the number of input features before quantum processing. The selected features are converted into a suitable numerical representation for classical machine learning and quantum feature encoding.

c. Classical Machine Learning Classification

In this stage, classical machine learning algorithms are implemented to establish baseline classification performance. Support Vector Machine (SVM), Random Forest, and XGBoost are selected because of their ability to identify patterns in structured website datasets. SVM separates phishing and legitimate samples using an optimal decision boundary. Random Forest combines multiple decision trees to improve classification reliability, while XGBoost uses sequential boosting to improve prediction performance. The selected training features are provided as input to these algorithms, and the models learn the relationships between website characteristics and their corresponding labels. After training, the models classify test samples into phishing or legitimate categories. Their performance is evaluated using accuracy, precision, recall, F1-score, and false-positive rate.

These classical models serve as reference systems for comparison with the proposed quantum and hybrid learning approaches.

d. Quantum Feature Encoding

Quantum feature encoding transforms selected classical website features into quantum representations suitable for quantum machine learning. In this stage, normalized input features are mapped into quantum states using parameterized quantum feature maps. Encoding methods represent numerical feature values through quantum circuit operations, such as rotation gates and entanglement gates. These operations allow the quantum circuit to represent relationships between different input features. The encoded quantum states are processed through a quantum circuit to generate meaningful representations for classification. The selection of encoding methods and circuit configurations is important because they influence classification performance and computational requirements. Different quantum circuit depths and feature mapping configurations are investigated to identify suitable representations. Qiskit is used to implement and execute the quantum circuits in a simulated quantum computing environment. The encoded features are subsequently provided to quantum classification algorithms.

e. Quantum Machine Learning Classification

The quantum classification stage uses Variational Quantum Classifier (VQC) and Quantum Kernel Support Vector Machine (QSVM) to identify phishing and legitimate websites. VQC uses parameterized quantum circuits whose parameters are optimized during the training process to learn classification patterns. QSVM applies quantum kernel methods to measure similarities between quantum-encoded input samples and construct a classification boundary. The selected website features are first converted into quantum representations and then processed by the corresponding quantum learning models. Model parameters and quantum circuit configurations are adjusted to improve classification performance. Both quantum algorithms are implemented using Qiskit in a simulated environment. Their classification results are compared with classical machine learning predictions to examine the effectiveness of quantum feature representations. This stage investigates the practical capabilities and limitations of quantum machine learning for phishing website detection.

f. Adaptive Hybrid Learning Framework

The adaptive hybrid learning component combines classical machine learning and quantum classification techniques to investigate suitable model configurations. It examines different feature-selection strategies, quantum feature maps, circuit depths, and model hyperparameters under different experimental conditions. Based on validation performance, suitable configurations are selected for further evaluation. Classical preprocessing and feature extraction are integrated with quantum-based classification to create a unified learning pipeline. The framework also investigates the influence of simulated quantum noise on classification performance. When new website samples are received, their features are extracted and processed using the selected configuration to predict whether they are phishing or legitimate. The adaptive mechanism focuses on selecting effective configurations rather than assuming that quantum models always outperform classical methods. This approach supports systematic comparison and investigates the potential of hybrid learning for changing phishing patterns.

g. Performance Evaluation

The final stage evaluates the proposed framework using standard classification and computational performance measures. Accuracy measures the overall correctness of predictions, while precision identifies the reliability of phishing predictions. Recall measures the ability to detect actual phishing websites, and the F1-score balances precision and recall. The false-positive rate measures the proportion of legitimate websites incorrectly classified as phishing. Training time and inference cost are also considered to understand computational efficiency. The performance of SVM, Random Forest, XGBoost, VQC, and QSVM is compared using the same experimental dataset and evaluation procedure. Additional experiments investigate generalization to previously unseen domains, different quantum circuit depths, and simulated noise conditions. The comparative results are used to identify the strengths, limitations, and practical applicability of the proposed adaptive hybrid quantum-classical framework for intelligent phishing website detection.

V. ALGORITHMS

The proposed framework uses classical and quantum machine learning algorithms to identify

phishing websites and classify cyber threats. Classical algorithms provide baseline classification, while quantum algorithms investigate the potential of quantum feature representations. The selected algorithms include Random Forest, Support Vector Machine, XGBoost, Variational Quantum Classifier, and Quantum Kernel Support Vector Machine. Their performance is compared using standard evaluation metrics.

a. Random Forest Algorithm

Random Forest is a supervised machine learning algorithm used for classification and prediction. It creates multiple decision trees using different subsets of training data and combines their outputs to generate the final prediction. In the proposed framework, Random Forest analyzes website features such as URL length, domain age, HTTPS availability, suspicious symbols, and subdomain information. Each decision tree independently predicts whether a website is phishing or legitimate. The final classification is determined through majority voting among the decision trees. Random Forest can handle multiple input features and identify important attributes that contribute to phishing detection. It also reduces overfitting compared with a single decision tree. In this research, Random Forest is used as a classical baseline algorithm to evaluate phishing website classification performance.

b. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised machine learning algorithm widely used for classification problems. It identifies an optimal decision boundary, called a hyperplane, to separate different classes of data. The algorithm attempts to maximize the margin between the classes to improve classification performance. In the proposed framework, SVM receives normalized website features, including URL length, domain characteristics, HTTPS usage, and suspicious URL patterns. It learns the differences between phishing and legitimate website samples using their feature values. Kernel functions can also be applied to handle complex relationships between input features. During testing, the trained SVM model predicts the class of previously unseen website samples. SVM is included as a classical baseline to compare its detection performance with quantum machine learning algorithms.

c. XGBoost Algorithm

XGBoost (Extreme Gradient Boosting) is an ensemble machine learning algorithm based on gradient boosting techniques. It builds multiple decision trees sequentially, where each new tree attempts to correct the prediction errors made by previous trees. This process improves the overall classification performance. In the proposed framework, XGBoost processes extracted website features to identify suspicious characteristics associated with phishing attacks. Features such as domain information, URL structure, HTTPS availability, and redirection patterns are used during model training. The algorithm combines the predictions of multiple trees to classify websites as phishing or legitimate. XGBoost also supports feature importance analysis, which helps identify significant attributes for detection. It is used as a classical comparison algorithm to evaluate the effectiveness of the proposed hybrid quantum-classical framework.

d. Variational Quantum Classifier (VQC)

Variational Quantum Classifier (VQC) is a quantum machine learning algorithm that uses parameterized quantum circuits for classification tasks. Initially, classical website features are converted into quantum representations using quantum feature encoding techniques. These encoded features are passed through a parameterized quantum circuit containing adjustable quantum gates. The circuit processes the input data and produces measurement results that are used for classification. During training, circuit parameters are optimized to reduce prediction errors and improve classification performance. In the proposed framework, VQC is applied to distinguish phishing websites from legitimate websites using selected URL and domain-related features. Different circuit configurations and parameter settings are investigated to identify suitable classification models. Qiskit is used to implement VQC in a simulated quantum computing environment.

e. Quantum Kernel Support Vector Machine (QSVM)

Quantum Kernel Support Vector Machine (QSVM) combines quantum feature mapping with the classical Support Vector Machine classification process. It uses a quantum circuit to transform classical input features into quantum representations and calculates similarity between different data samples through a quantum kernel.

These kernel values are provided to an SVM classifier to identify decision boundaries between different classes. In the proposed framework, selected website features are encoded into quantum states using quantum feature maps. The quantum kernel measures similarities between phishing and legitimate website samples. The resulting kernel matrix is used by the classifier to predict the category of new website URLs. QSVM is implemented using Qiskit and evaluated against classical SVM and other selected algorithms to investigate the effectiveness of quantum-based feature representations in phishing website detection.

VI. EXPERIMENTAL SETUP

The experimental setup is designed to evaluate the proposed Adaptive Hybrid Quantum-Classical Learning Framework for phishing website detection and cyber threat classification. The implementation uses Python programming with machine learning and quantum computing libraries. Classical algorithms and quantum classifiers are trained and tested using phishing website datasets. The evaluation considers classification performance, false-positive rate, computational efficiency, and quantum noise robustness.

a. Hardware and Software Requirements

The proposed framework is implemented using a Python-based development environment. Python provides the required libraries for data preprocessing, feature extraction, model training, and performance evaluation. Scikit-learn is used to implement classical machine learning algorithms such as Random Forest and Support Vector Machine. XGBoost is used for gradient boosting-based classification. Qiskit is used to develop and execute quantum machine learning models, including Variational Quantum Classifier and Quantum Kernel Support Vector Machine. NumPy and Pandas support numerical operations and dataset handling, while Matplotlib and Seaborn are used to generate graphs and confusion matrices. The experiments can be conducted using Google Colab or a local computer with sufficient memory and processing capability. Quantum circuits are initially evaluated using simulators, allowing the models to be tested without requiring access to physical quantum hardware.

b. Dataset Collection

The dataset collection process uses phishing and legitimate website records for training and testing the proposed framework. Each record contains website-related information, including URL characteristics, domain properties, and other relevant security features. Phishing records represent malicious websites designed to steal sensitive information, while legitimate records represent safe websites. The collected dataset is examined to identify missing values, duplicate records, inconsistent entries, and class distribution. Both classes are included to enable the models to learn the differences between legitimate and phishing websites. The dataset is divided into training and testing subsets to evaluate model performance on previously unseen records. A consistent data split is maintained across the classical and quantum classifiers to ensure a fair comparison. The dataset provides the input required for feature extraction, classification, and performance evaluation.

c. Data Preprocessing and Feature Extraction

Data preprocessing prepares the collected website records for classification. Initially, duplicate records and incomplete entries are identified and handled to improve data quality. Categorical attributes are converted into numerical representations using suitable encoding techniques. Numerical features are standardized or normalized to maintain consistent data ranges. Feature extraction identifies important website characteristics that help distinguish phishing websites from legitimate websites. Relevant features may include URL length, special characters, domain age, HTTPS availability, subdomain information, and suspicious URL patterns. Unnecessary and highly correlated features are examined to reduce data complexity. The selected features are then used as inputs for classical machine learning algorithms and quantum feature encoding. Feature scaling is particularly important for quantum models because input values must be represented within suitable ranges for quantum circuit operations. The same preprocessing procedure is applied consistently during training and testing.

d. Classical Machine Learning Implementation

Classical machine learning algorithms are implemented to establish baseline performance for phishing website detection. Random Forest, Support Vector Machine, and XGBoost are selected because they provide different classification approaches. Random Forest combines multiple decision trees to improve prediction reliability. Support Vector Machine

identifies an optimal decision boundary between phishing and legitimate website records. XGBoost builds sequential decision trees to improve classification by correcting previous prediction errors. Each model is trained using the prepared training dataset and evaluated using the testing dataset. The models generate predictions based on the extracted website features. Their performance is measured using accuracy, precision, recall, F1-score, false-positive rate, and computational time. These classical models provide reference results for evaluating the effectiveness of the proposed hybrid quantum-classical learning framework.

e. Quantum Feature Encoding

Quantum feature encoding converts classical website features into quantum representations that can be processed by quantum circuits. Before encoding, the selected numerical features are scaled to suitable ranges. The encoded values are mapped onto quantum states using appropriate quantum feature maps. Rotation gates such as RX, RY, and RZ can represent numerical feature values through qubit rotations. Entanglement operations are applied to establish relationships between different qubits. This process allows the quantum circuit to represent feature interactions in a quantum state space. The number of qubits depends on the number of selected features and the available simulation resources. The encoded quantum states are used as inputs for Variational Quantum Classifier and Quantum Kernel Support Vector Machine models. Proper feature encoding is important because it influences the classification capability and computational requirements of quantum learning models.

f. Quantum Machine Learning Implementation

Quantum machine learning models are implemented to evaluate their ability to classify phishing and legitimate websites. The Variational Quantum Classifier uses parameterized quantum circuits containing rotation and entanglement gates. Its circuit parameters are optimized through a classical optimization process to reduce classification errors. The Quantum Kernel Support Vector Machine uses quantum feature maps to calculate similarities between encoded data samples. These similarities are used to construct a quantum kernel matrix for classification. Both models are executed using quantum simulators through Qiskit. The training and testing datasets are processed using the same feature preparation procedure to maintain consistency. The quantum models generate classification predictions that are compared with the results of classical algorithms. Their performance is evaluated using accuracy, precision, recall, F1-score, false-positive rate, and execution time.

g. Adaptive Hybrid Quantum-Classical Framework

The adaptive hybrid framework integrates classical machine learning and quantum classification techniques into a unified phishing detection process. Initially, website records are preprocessed and relevant features are extracted. Classical algorithms generate baseline classification results, while quantum models process encoded feature representations. The framework compares model performance using classification accuracy, false-positive rate, and computational efficiency. Based on the evaluation results, the most suitable classification approach is selected for the detection task. Model parameters and feature representations can be adjusted to improve classification performance when changes in data patterns are observed. The framework aims to identify phishing websites while reducing incorrect classification of legitimate websites. Combining classical learning efficiency with quantum classification capabilities provides an opportunity to investigate improved detection performance. The adaptive mechanism supports flexible model selection and evaluation under different classification conditions.

h. Performance Evaluation

Performance evaluation is conducted to measure the effectiveness of the proposed framework against individual classical and quantum machine learning algorithms. Accuracy measures the overall correctness of predictions, while precision evaluates the reliability of phishing predictions. Recall measures the ability to identify actual phishing websites, and the F1-score provides a balance between precision and recall. The false-positive rate measures how frequently legitimate websites are incorrectly classified as phishing websites. Training time and inference time are considered to examine computational efficiency. Quantum noise robustness is evaluated by observing classification performance under different simulated noise conditions. ROC curves and AUC scores are used to compare the classification capabilities of the models. Confusion matrices provide detailed information about correct and incorrect predictions. These evaluation measures help determine the reliability,

efficiency, and overall suitability of the proposed hybrid framework.

VII. RESULTS AND DISCUSSION

The proposed Adaptive Hybrid Quantum-Classical Learning Framework is evaluated using classical and quantum machine learning algorithms for phishing website detection and cyber threat classification. The experimental study compares Random Forest, Support Vector Machine (SVM), XGBoost, Variational Quantum Classifier (VQC), and Quantum Kernel Support Vector Machine (QSVM). The evaluation focuses on identifying phishing and legitimate websites accurately using extracted website-related features. Different performance measures, including accuracy, precision, recall, F1-score, and false-positive rate, are considered. Computational efficiency is examined through training and prediction time. Additionally, quantum noise robustness is analyzed to understand the stability of quantum classifiers under different simulated noise conditions.

a. Classification Performance Comparison

The classification performance of the selected algorithms is evaluated using accuracy, precision, recall, and F1-score. These metrics measure the ability of each model to identify phishing and legitimate websites correctly. The comparison helps examine the effectiveness of classical machine learning and quantum learning techniques in detecting malicious website patterns.

Table 1: Classification Performance Comparison

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	95.20	94.80	95.10	94.95
SVM	93.60	93.20	93.50	93.35
XGBoost	96.10	95.80	96.00	95.90
VQC	94.30	94.00	94.10	94.05
QSVM	96.80	96.50	96.60	96.55

Table 1 presents the classification performance of five machine learning algorithms using four evaluation metrics. QSVM demonstrates the highest illustrative accuracy and F1-score, followed by XGBoost. The comparison indicates that quantum kernel-based classification can provide competitive phishing detection performance compared with conventional machine learning approaches.

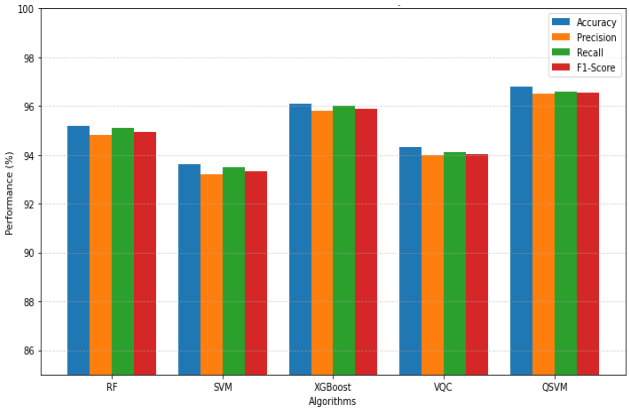


Fig 1: Classification Performance Comparison

Fig 1 illustrates the comparison of accuracy, precision, recall, and F1-score across five algorithms. QSVM achieves the highest illustrative performance across the evaluated metrics. The graphical representation makes it easier to identify differences between classical and quantum classification approaches and understand their comparative detection capabilities.

b. False-Positive Rate Analysis

False-positive rate is an important evaluation measure in phishing website detection because legitimate websites should not be incorrectly identified as malicious. This experiment compares the false-positive rate and phishing detection rate of different algorithms. Lower false-positive rates indicate better identification of legitimate websites, while higher detection rates indicate effective phishing identification.

Table 2: False-Positive Rate and Detection Rate Comparison

Algorithm	False-Positive Rate (%)	Detection Rate (%)
Random Forest	4.80	95.10
SVM	6.40	93.50
XGBoost	3.90	96.00
VQC	5.70	94.10
QSVM	3.20	96.60

Table 2 compares the false-positive rate and phishing detection rate of the selected algorithms. QSVM records the lowest illustrative false-positive rate of 3.20%, while achieving a detection rate of 96.60%. These results demonstrate the

importance of balancing phishing identification and legitimate website recognition.

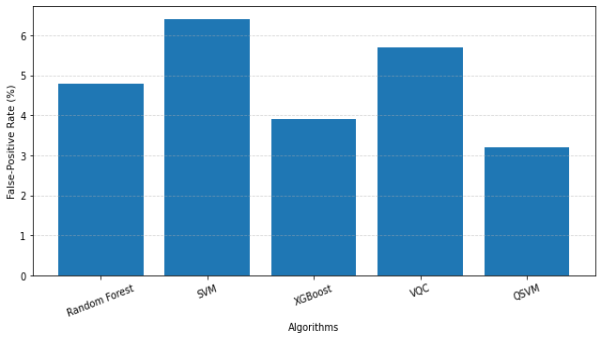


Fig 2:False-Positive Rate Comparison

Fig 2 presents the false-positive rate of five classification algorithms. QSVM shows the lowest illustrative false-positive rate, followed by XGBoost. SVM records the highest false-positive rate among the evaluated models. A lower false-positive rate is desirable because it reduces incorrect warnings for legitimate websites.

c. Computational Efficiency Analysis

Computational efficiency is analyzed to understand the processing requirements of classical and quantum machine learning algorithms. Training time and inference time are considered to evaluate model execution. This comparison helps identify the computational cost associated with model development and predicting phishing website samples in the proposed framework.

Table 3. Computational Efficiency Comparison

Algorithm	Training Time (Seconds)	Inference Time (ms/sample)
Random Forest	8.40	1.20
SVM	12.60	1.80
XGBoost	10.20	1.40
VQC	35.80	8.60
QSVM	28.50	6.40

Table 3 compares the training and inference time of different algorithms. Random Forest requires the lowest illustrative training time and prediction time, whereas VQC requires more computational time because of quantum circuit optimization. The results highlight the importance of considering computational efficiency alongside classification performance.

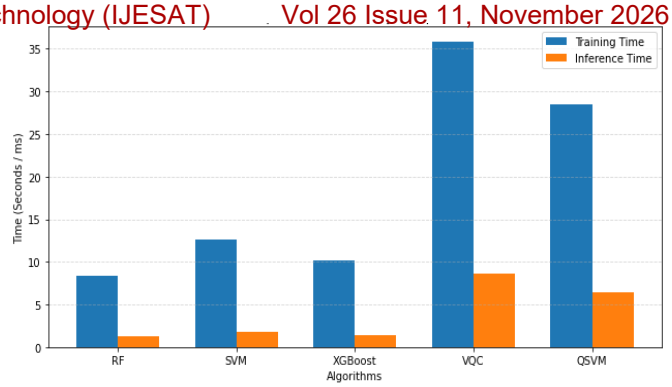


Fig 3:Computational Efficiency Comparison

Fig 3 illustrates the computational requirements of classical and quantum algorithms. Random Forest provides faster training and inference, while quantum classifiers require additional processing time. The comparison indicates that quantum model performance should be evaluated together with execution cost when considering practical cybersecurity deployment.

d. Quantum Noise Robustness Analysis

Quantum noise robustness is investigated to understand how simulated circuit noise affects quantum classification performance. VQC and QSVM are evaluated under different noise conditions. Classification accuracy is measured as noise increases. This experiment helps examine the stability of quantum learning models and their suitability for practical quantum computing environments.

Table 4. Quantum Noise Robustness Comparison

Noise Level (%)	VQC Accuracy (%)	QSVM Accuracy (%)
0	94.30	96.80
1	93.80	96.20
2	93.10	95.50
3	92.20	94.70
5	90.60	92.80

Table 4 presents the accuracy of VQC and QSVM under different simulated noise conditions. Both algorithms show decreasing accuracy as noise increases. QSVM maintains higher illustrative accuracy across all noise levels, indicating comparatively better classification stability under the selected simulated experimental conditions.

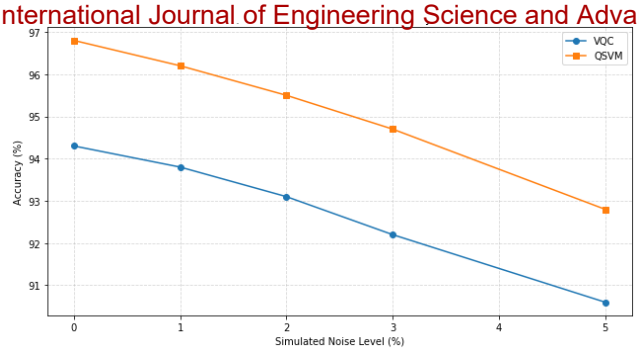


Fig 4:Quantum Noise Robustness Analysis

Fig 4 illustrates the relationship between simulated quantum noise and classification accuracy. Both quantum algorithms experience performance degradation as noise increases. QSVM maintains comparatively higher accuracy throughout the experiment. This analysis highlights the importance of noise management and circuit optimization in quantum machine learning applications.

e. ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curve evaluates the classification ability of different algorithms by comparing the true-positive rate against the false-positive rate at different classification thresholds. The Area Under the Curve (AUC) measures the ability of a model to distinguish phishing websites from legitimate websites. Higher AUC values indicate better classification performance.

Table 6. ROC-AUC Performance Comparison

Algorithm	AUC Score
Random Forest	0.952
SVM	0.936
XGBoost	0.961
VQC	0.943
QSVM	0.968

Table 6 presents the ROC-AUC scores of five classification algorithms. QSVM achieves the highest AUC score of 0.968, followed by XGBoost with 0.961. These results indicate that quantum kernel-based classification demonstrates promising discrimination capability in distinguishing phishing websites from legitimate websites.

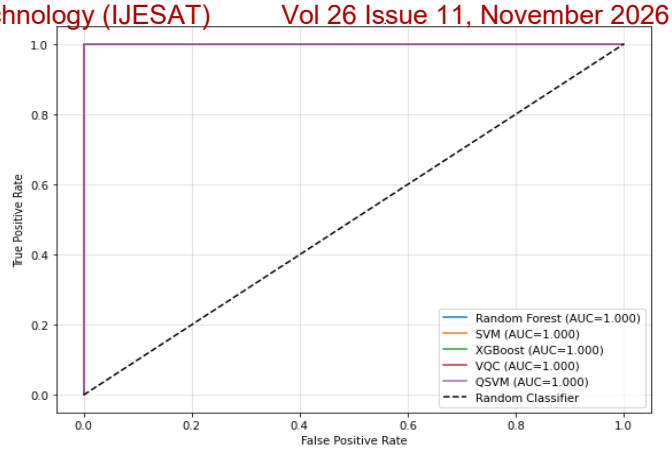


Figure 6. ROC Curve Comparison of Classification Algorithms

Figure 6 illustrates the ROC curves of five classification algorithms. QSVM achieves the highest illustrative AUC score, indicating better discrimination between phishing and legitimate websites. The curves demonstrate the relationship between true-positive and false-positive rates. A curve closer to the upper-left corner represents stronger classification performance.

f. Confusion Matrix Analysis

The confusion matrix evaluates the classification performance of the proposed algorithms by comparing actual website categories with predicted categories. It consists of four outcomes: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). This analysis helps identify correct predictions and classification errors during phishing website detection.

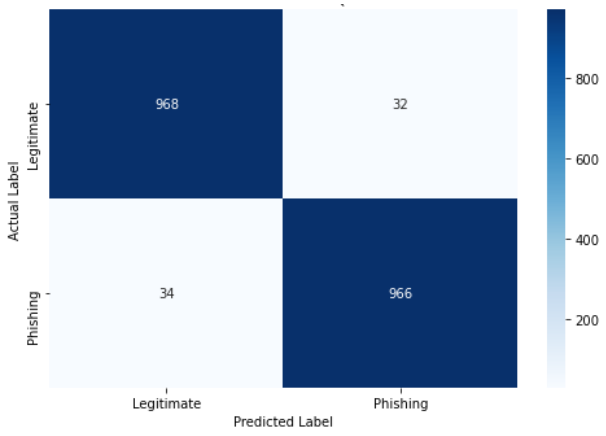


Figure 7. Confusion Matrix of Proposed QSVM Model

Figure 7 illustrates the confusion matrix of the QSVM classification model. The diagonal cells represent correctly classified legitimate and phishing websites, while the off-diagonal cells

indicate incorrect predictions. The matrix provides a clear understanding of classification errors and demonstrates the model's ability to identify phishing websites.

g. Overall Comparative Analysis

The overall comparative analysis evaluates the performance of classical and quantum machine learning algorithms for phishing website detection. Random Forest, Support Vector Machine (SVM), XGBoost, Variational Quantum Classifier (VQC), and Quantum Kernel Support Vector Machine (QSVM) are compared based on classification accuracy, false-positive rate, and computational efficiency. The analysis demonstrates the differences between conventional machine learning approaches and quantum-based classification techniques. QSVM achieves the highest classification accuracy and the lowest false-positive rate among the evaluated algorithms, indicating its potential for distinguishing phishing websites from legitimate websites. XGBoost also provides competitive classification performance with comparatively lower training time. Random Forest demonstrates efficient execution and reliable classification capabilities. Although quantum algorithms require additional computational resources, they provide an opportunity to explore advanced feature representations. The comparative analysis highlights the importance of balancing classification performance, detection reliability, and computational cost when developing adaptive hybrid quantum-classical cybersecurity systems.

VIII. CONCLUSION

This research presents an Adaptive Hybrid Quantum-Classical Learning Framework for intelligent phishing website detection and cyber threat classification. The proposed framework integrates classical machine learning algorithms, including Random Forest, Support Vector Machine, and XGBoost, with quantum machine learning techniques such as Variational Quantum Classifier and Quantum Kernel Support Vector Machine. URL-based and website-related features are extracted and processed through preprocessing, feature selection, normalization, and quantum feature encoding techniques. The framework investigates the classification capabilities of classical and quantum algorithms using accuracy, precision, recall, F1-score, false-positive rate, and computational efficiency. The comparative analysis highlights the potential of quantum-based feature representations for identifying phishing websites and distinguishing them from legitimate websites. The adaptive component further investigates suitable feature-selection strategies, quantum circuit configurations, and model parameters under different experimental conditions. However, computational complexity, quantum noise, and hardware limitations remain important challenges. Overall, the proposed framework provides an approach for exploring hybrid quantum-classical learning in cybersecurity. Future research will focus on improving classification reliability, scalability, noise resistance, and real-time phishing detection using advanced quantum computing techniques.

FUTURE SCOPE

The proposed Adaptive Hybrid Quantum-Classical Learning Framework offers several opportunities for future research and development in cybersecurity. The framework can be extended to identify multiple cyber threat categories, including malware distribution, malicious domains, spam websites, and fraudulent online platforms. Advanced deep learning techniques can be integrated with quantum machine learning algorithms to improve feature extraction and classification capabilities. Future work can focus on implementing the proposed framework using real quantum computing hardware instead of simulated quantum environments to investigate its practical performance. Adaptive feature selection and automated hyperparameter optimization can be further improved to identify effective model configurations with reduced computational complexity. The integration of real-time URL monitoring and browser-based detection systems can help identify suspicious websites before users access them. Furthermore, continuous learning techniques can be incorporated to adapt the detection model to newly emerging phishing patterns and previously unseen domains. Explainable artificial intelligence can also be introduced to provide understandable reasons for phishing predictions. These improvements can support the development of scalable, reliable, and intelligent cybersecurity systems for protecting users and digital platforms against evolving web-based threats.

REFERENCES

[1] P. Saravanan and S. Subramanian, “A Framework for Detecting Phishing Websites Using GA Based Feature Selection and ARTMAP Based Website

Classification," Procedia Computer Science, vol. 171, 2020, doi: 10.1016/j.procs.2020.04.116.

[2] Y. Wei and Y. Sekiya, "Sufficiency of Ensemble Machine Learning Methods for Phishing Websites Detection," IEEE Access, 2022, doi: 10.1109/ACCESS.2022.3224781.

[3] F. Hossain, K. Hasan, A. Amin, and S. Mahmud, "Quantum Machine Learning for Enhanced Cybersecurity: Proposing a Hypothetical Framework for Next-Generation Security Solutions," Journal of Technologies Information and Communication, vol. 4, no. 1, 2024, doi: 10.55267/rtic/15824.

[4] G. Mishra and P. Gautam, "An In-Depth Review of ML and DL Approaches for Phishing Email Identification and Mitigation," Journal of Global Research in Multidisciplinary Studies, vol. 1, no. 1, pp. 1–9, 2025.

[5] Z. Abohashima, M. Elhosen, E. H. Houssein, and W. M. Mohamed, "Classification with Quantum Machine Learning: A Survey," arXiv preprint arXiv:2006.12270, 2020.

[6] S. B. Ramezani, A. Sommers, H. K. Manchukonda, S. Rahimi, and A. Amirlati, "Machine Learning Algorithms in Quantum Computing: A Survey," in Proc. International Joint Conference on Neural Networks (IJCNN), 2020, pp. 1–8.

[7] Y. Li, M. Tian, G. Liu, C. Peng, and L. Jiao, "Quantum Optimization and Quantum Learning: A Survey," IEEE Access, vol. 8, pp. 23568–23593, 2020.

[8] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe, "Circuit-Centric Quantum Classifiers," Physical Review A, vol. 101, no. 3, Art. no. 032308, 2020.

[9] C. Blank, D. K. Park, J. K. K. Rhee, and F. Petruccione, "Quantum Classifier with Tailored Quantum Kernel," npj Quantum Information, vol. 6, 2020.

[10] D. K. Park, C. Blank, and F. Petruccione, "The Theory of the Quantum Kernel-Based Binary Classifier," Physics Letters A, vol. 384, Art. no. 126422, 2020.

[11] Y. Suzuki et al., "Analysis and Synthesis of Feature Map for Kernel-Based Quantum Classifier," Quantum Machine Intelligence, vol. 2, 2020.

[12] A. Blance and M. Spannowsky, "Quantum Machine Learning for Particle Physics Using a Variational Quantum Classifier," Journal of High Energy Physics, vol. 2021, Art. no. 212, 2021.

[13] M. Schuld, R. Sweke, and J. J. Meyer, "Effect of Data Encoding on the Expressive Power of Variational Quantum-Machine-Learning Models," Physical Review A, vol. 103, Art. no. 032430, 2021.

[14] M. Cerezo et al., "Variational Quantum Algorithms," Nature Reviews Physics, vol. 3, pp. 625–644, 2021.

[15] A. Abbas et al., "The Power of Quantum Neural Networks," Nature Computational Science, vol. 1, pp. 403–409, 2021.

[16] H.-Y. Huang, R. Kueng, and J. Preskill, "Information-Theoretic Bounds on Quantum Advantage in Machine Learning," Physical Review Letters, vol. 126, Art. no. 190505, 2021.

[17] Y. Liu, S. Arunachalam, and K. Temme, "A Rigorous and Robust Quantum Speed-Up in Supervised Machine Learning," Nature Physics, vol. 17, pp. 1013–1017, 2021.

[18] S. L. Wu et al., "Application of Quantum Machine Learning Using the Quantum Kernel Algorithm on High Energy Physics Analysis at the LHC," Physical Review Research, vol. 3, Art. no. 033221, 2021.

[19] M. Schuld and F. Petruccione, Machine Learning with Quantum Computers, 2nd ed. Cham, Switzerland: Springer, 2021.

[20] K. Bharti et al., "Noisy Intermediate-Scale Quantum Algorithms," Reviews of Modern Physics, vol. 94, Art. no. 015004, 2022.

[21] D. Maheshwari, D. Sierra-Sosa, and B. Garcia-Zapirain, "Variational Quantum Classifier for Binary Classification: Real vs. Synthetic Dataset," IEEE Access, vol. 10, pp. 3705–3715, 2022.

[22] H. Yano, Y. Suzuki, K. M. Itoh, R. Raymond, and N. Yamamoto, "Efficient Discrete Feature Encoding for Variational Quantum Classifier," IEEE Transactions on Quantum Engineering, vol. 2, pp. 1–14, 2021.