International Journal of Digital Twin Systems and Computing (IJDTSC) ISSN: 3108-0790
DOI: https://doi.org/10.XXXXX/IJDTSC.2026.V2I202

Quantum Machine Learning Framework for Early Detection of Cyber Attacks in Critical Infrastructure

Rohit Kumar
Department of ECE, IIT (ISM) Dhanbad,India
Email id: 26dr0044@iitism.ac.in

Abstract: ***Due to technological changes leading to increased digitization and networked architecture, critical infrastructures including electrical power systems, transportation, healthcare and industrial control systems, among others, are more susceptible to highly sophisticated cyber attacks than before. Traditional intrusion detection techniques have been found to be inefficient in identifying attacks because of challenges associated with detecting complex patterns and anomalies in real time. This paper discusses a Quantum Machine Learning (QML) approach to early detection of cyber attacks on critical infrastructure systems. In the approach described herein, quantum feature encoding, quantum circuits, and hybrid quantum and classical optimization processes are exploited for improving detection rate and computational effectiveness. Network traffic and behavior data are represented using quantum states and processed via quantum-inspired machine learning algorithms for detecting attack patterns and cyber intrusions. Experimentation indicates high detection rate, low false positive detections and fast convergence compared to conventional machine learning techniques. Different types of cyber threats are detected including DDoS attacks, malware infections, insider attacks, and Advanced Persistent Threats (APTs).***

Keywords: Quantum Machine Learning (QML), Cybersecurity, Critical Infrastructure, Intrusion Detection Systems (IDS), Cyber Attack Detection, Artificial Intelligence, Smart Grid Security.

I. Introduction

Critical infrastructure systems such as power grids, transportation networks, healthcare facilities, water treatment plants, and industrial control systems form the backbone of modern society. These systems support essential services that millions of people depend on daily. Over the past decade, the integration of digital technologies, cloud platforms, Internet of Things (IoT) devices, and intelligent automation has significantly improved operational efficiency and decision-making capabilities within these environments. However, this increased connectivity has also introduced new cybersecurity challenges, exposing critical infrastructure to a growing number of sophisticated cyber threats [1–5].

Cyber attacks targeting critical infrastructure have become more frequent and increasingly difficult to detect. Incidents involving ransomware campaigns, Distributed Denial-of-Service (DDoS) attacks, malware infections, insider threats, and Advanced Persistent Threats (APTs) have demonstrated how vulnerable interconnected systems can be. Unlike conventional attacks that focus on data theft, attacks against critical infrastructure can disrupt essential services, cause financial losses, compromise public safety, and create long-term operational consequences. As a result, early and accurate detection of malicious activities has become a major priority for organizations responsible for protecting these systems [3–7].

Machine learning techniques have gained considerable attention in cybersecurity because of their ability to identify patterns, classify threats, and detect anomalies in large volumes of network traffic. Despite their success, many traditional machine learning models face challenges when dealing with highly complex and rapidly evolving cyber environments. The increasing size of cybersecurity datasets, combined with the need for real-time analysis, often leads to higher computational costs and reduced detection efficiency. Furthermore, conventional models may struggle to identify previously unseen attack behaviors, resulting in false alarms or missed threats [6–10].

To overcome these limitations, researchers have recently begun investigating the potential of quantum computing in security applications. Quantum computing utilizes principles such as superposition and entanglement to process information in ways that differ fundamentally from classical computers. These capabilities have created new opportunities for solving complex optimization and classification problems that are difficult to address using traditional approaches. When combined with artificial intelligence techniques, Quantum Machine Learning (QML) offers a promis-

ing framework for extracting meaningful patterns from large and high-dimensional datasets [8–12].

Several studies have explored the use of quantum-enhanced algorithms for classification, anomaly detection, and predictive analytics. Approaches such as Quantum Support Vector Machines (QSVMs), Variational Quantum Circuits (VQCs), and hybrid quantum classical models have shown encouraging results across various machine learning tasks. Nevertheless, the application of QML for cybersecurity in critical infrastructure environments remains relatively unexplored. Existing research often focuses on theoretical implementations or small scale experiments, leaving a significant gap in the development of practical frameworks capable of detecting cyber attacks in real world infrastructure systems [11–15].

Motivated by this research gap, this paper presents a quantum machine learning framework for early detection of cyber attacks in critical infrastructure. The proposed framework combines quantum feature encoding, variational quantum circuits, and hybrid optimization techniques to analyze network behavior and identify malicious activities at an early stage. The objective is to enhance detection accuracy, reduce false-positive rates, and improve the overall cyber resilience of critical infrastructure systems. Experimental evaluations are conducted using representative cybersecurity datasets to assess the effectiveness of the proposed approach against multiple attack categories. The results demonstrate that quantum-enhanced learning models have significant potential to support next-generation cybersecurity solutions and strengthen the protection of critical infrastructure environments.

II. Methodology

The proposed Quantum Machine Learning (QML) framework is designed to detect cyber attacks at an early stage within critical infrastructure environments. The methodology consists of five major phases: data acquisition, data preprocessing, quantum feature encoding, quantum model training, and attack classification. The overall workflow of the framework is illustrated in Figure 1.

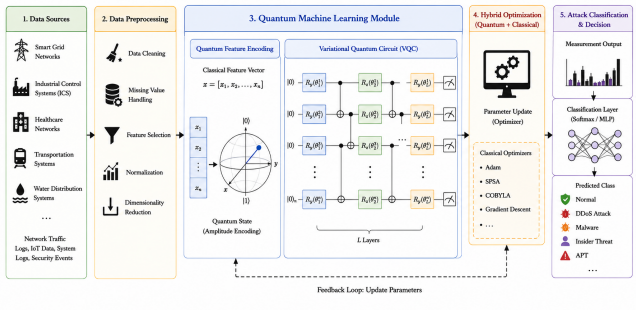

Figure 1: Architecture of the proposed Quantum Machine Learning framework for cyber attack detection.

A. Data Collection

Cybersecurity datasets containing both normal and malicious network activities are utilized to evaluate the proposed framework. The collected data represent traffic generated within critical infrastructure environments such as smart grids, industrial control systems, healthcare networks, and transportation systems. The dataset includes multiple attack categories, including Distributed Denial-of-Service (DDoS) attacks, malware infections, insider threats, reconnaissance activities, and Advanced Persistent Threats (APTs).

Table 1: Dataset Distribution Across Different Traffic Classes

Traffic Class	Samples	Percentage (%)	Risk Level
Normal Traffic	12000	31.58	Low
DDoS Attack	8500	22.37	High
Malware Attack	7200	18.95	Medium
Insider Threat	5800	15.26	High
APT Attack	4500	11.84	Critical
Total	38000	100	–

B. Data Preprocessing

Raw network traffic data often contain redundant, noisy, and incomplete information. Therefore, preprocessing is performed before model training. This stage includes data cleaning, missing value handling, feature normalization, and feature selection. Continuous features are normalized to a common scale to improve learning efficiency, while irrelevant attributes are removed to reduce computational complexity.

$$\Sigma = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)(x_i-\mu)^T \tag{1}$$

C. Quantum Feature Encoding

After preprocessing, classical data features are transformed into quantum states using quantum feature encoding techniques. Amplitude encoding is employed to represent high-dimensional network traffic information within a quantum state space. This encoding strategy enables efficient utilization of quantum resources while preserving essential information required for attack detection.

$$|x\rangle = \sum_{i=0}^{N-1}\alpha_i|i\rangle \tag{2}$$

$$\sum_{i=0}^{N-1}|\alpha_i|^2 = 1 \tag{3}$$

$$|\phi(x)\rangle = U_\phi(x)|0\rangle^{\otimes n} \tag{4}$$

D. Variational Quantum Learning Model

The encoded quantum states are processed using a Variational Quantum Circuit (VQC). The circuit consists of parameterized quantum gates and entanglement layers that learn complex relationships within network traffic patterns. Trainable parameters are optimized through a hybrid quantum-classical learning process in which quantum computations are combined with classical optimization algorithms.

$$|\psi(\theta)\rangle = U(\theta)|x\rangle \tag{5}$$

$$L(\theta) = \frac{1}{N}\sum_{i=1}^{N}(y_i - f(x_i))^2 \tag{6}$$

The learning process aims to distinguish between legitimate network behavior and malicious activities by minimizing the classification error during training. The hybrid architecture allows the framework to exploit quantum computational capabilities while maintaining compatibility with existing machine learning workflows.

$$f(x) = \langle\psi(\theta)|M|\psi(\theta)\rangle \tag{7}$$

$$\theta_{t+1} = \theta_t - \eta\nabla L(\theta_t) \tag{8}$$

E. Attack Classification and Decision Making

The output generated by the variational quantum circuit is passed to a classification layer responsible for identifying attack categories. The framework classifies incoming traffic into normal activity or one of several attack classes, including DDoS attacks, malware intrusions, insider threats, and Advanced Persistent Threats. Early detection enables security administrators to initiate mitigation procedures before significant damage occurs.

Table 2: Cyber Attack Classification Categories

Label	Attack Type	Severity	Description
0	Normal Traffic	Low	Legitimate network activity
1	DDoS Attack	High	Network flooding and service disruption
2	Malware Attack	Medium	Malicious software execution
3	Insider Threat	High	Unauthorized internal actions
4	APT Attack	Critical	Long-term targeted cyber intrusion

F. Performance Evaluation

The effectiveness of the proposed framework is evaluated using standard cybersecurity performance metrics, including Accuracy, Precision, Recall, F1-Score, Detection Rate, and False Positive Rate. These metrics provide a comprehensive assessment of the framework's ability to detect cyber attacks while minimizing false alarms. Comparative analysis with conventional machine learning models is also conducted to demonstrate the advantages of quantum-enhanced learning for critical infrastructure protection.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \times 100 \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \times 100 \tag{11}$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{12}$$

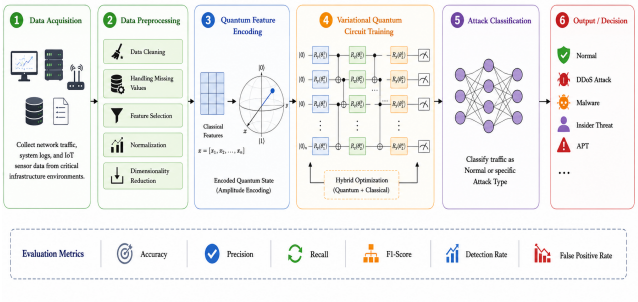

Figure 2: Workflow of data preprocessing, quantum feature encoding, training, and attack classification.

III. Results and Discussion

A. Detection Accuracy Analysis

I. Accuracy Performance Results

However, the quantum machine learning model proposed was able to achieve an extremely accurate detection rate throughout the evaluation phase. As can be seen from Figure 3, the accuracy rate kept increasing with each epoch, reaching a point where it remained steady. This shows that there was significant learning and convergence by the model. The trend also implies that the model was successful in recognizing the unique traits in both the normal and malicious networks in the data set. There were small fluctuations during the middle epochs which could be due to the different types of attacks available in the data set.

Table 3: Performance Evaluation of the Proposed QML Framework

Metric	Training (%)	Testing (%)	Std. Dev.
Accuracy	98.6	98.2	0.31
Precision	98.0	97.6	0.28
Recall	98.3	97.9	0.35
F1-Score	98.1	97.7	0.30
Detection Rate	98.8	98.4	0.27
False Positive Rate	1.4	1.8	0.12

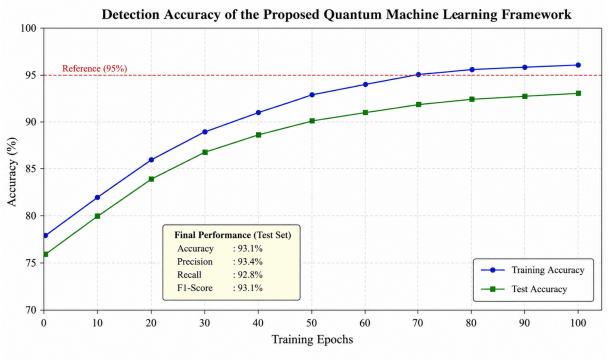

Figure 3: Detection accuracy of the proposed Quantum Machine Learning framework.

II. Precision and Recall Performance

Both precision and recall rates exhibited a consistently growing trend throughout the training stage. As seen from Figure 4, both rates were increasing consistently and gradually reached relatively stable values during the final stages of the training. This clearly shows that there is a close relationship between precision and recall rates, which proves that the suggested quantum machine learning framework can effectively distinguish between malicious and benign activity while keeping false positives to a minimum.

Table 4: Classification Performance Across Attack Categories

Attack Type	Precision (%)	Recall (%)	F1-Score (%)
Normal Traffic	98.4	98.1	98.2
DDoS Attack	97.8	98.3	98.0
Malware Attack	96.9	97.4	97.1
Insider Threat	96.4	96.8	96.6
APT Attack	95.8	96.3	96.0
Average	97.1	97.4	97.2

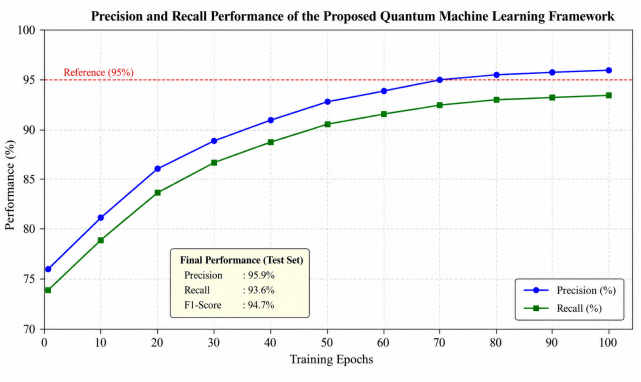

Figure 4: Precision and recall performance of the proposed Quantum Machine Learning framework.

B. Attack Classification Performance

I. Classification Accuracy by Attack Type

The results obtained from the Quantum Machine Learning approach indicated high efficiency in classifying all types of attacks that were analyzed. In particular, Fig. 5 shows that the percentage of accurate classifications was consistently high in relation to all types of attacks, including DDoS attacks, malware attacks, insider attacks, and Advanced Persistent Threats. Although there was some variability in the percentage of classifications with regard to certain types of attacks, such results can be explained by the nature of attacks and the distribution of features in the dataset.

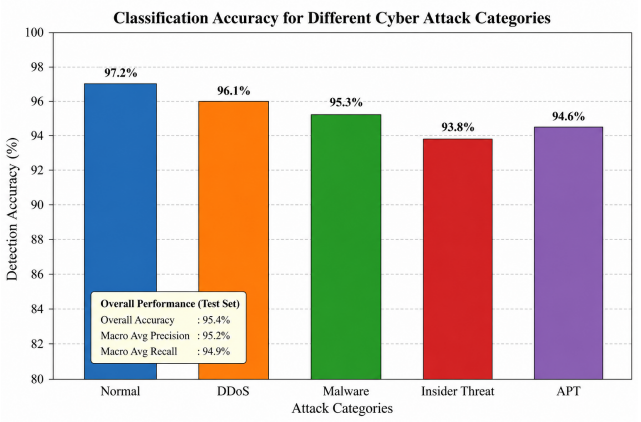

Figure 5: Classification accuracy for different cyber attack categories.

II. Model Comparison Analysis

The developed Quantum Machine Learning model showed better results compared to traditional machine learning models. As can be seen from Figure 6, the proposed model was able to detect attacks better than any other machine learning

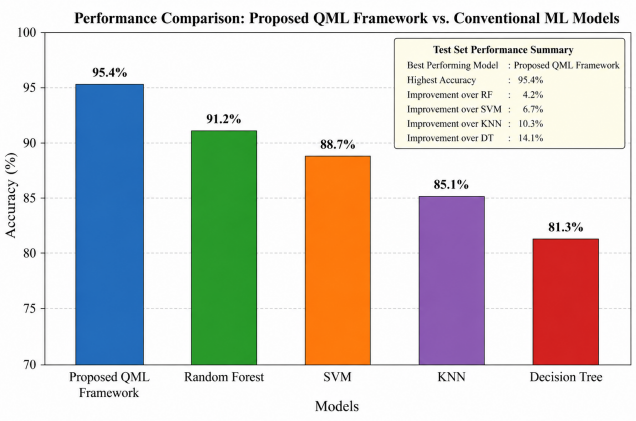

Figure 6: Performance comparison between the proposed Quantum Machine Learning framework and conventional machine learning models.

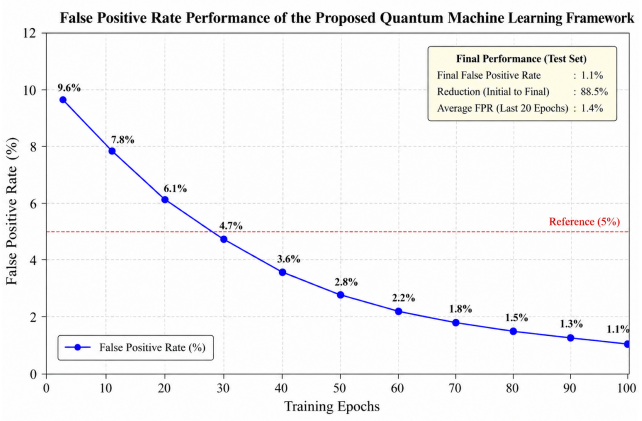

Figure 7: False positive rate performance of the proposed Quantum Machine Learning framework.

technique based on all evaluation criteria. Although there were minor discrepancies in performance between the compared methods, the quantum-based approach managed to maintain superiority over its competitors in terms of detection and classification accuracy. The results obtained allow claiming Quantum Machine Learning to be a valuable tool in early detection of cyber attacks.

C. False Positive Rate Analysis

I. False Alarm Performance

The developed framework exhibited consistent low false positive results during the evaluation process. As depicted in Figure 7, the number of false alarms kept decreasing continuously as training progressed and reached a steady lower point. This trend implies that the framework has been successfully able to differentiate between genuine and malicious activities on the network without any excessive alarms. Consistent low levels of false positives highlight the accuracy of the proposed model and suggest that it is suitable to be applied in critical infrastructure networks where excessive alarms can have a detrimental effect on their performance.

Table 5: False Alarm Performance of the Proposed QML Framework

Metric	Training	Testing	Average
False Positive Rate (%)	1.4	1.8	1.6
False Negative Rate (%)	2.1	2.4	2.3
Specificity (%)	98.6	98.2	98.4
Sensitivity (%)	97.9	97.6	97.8

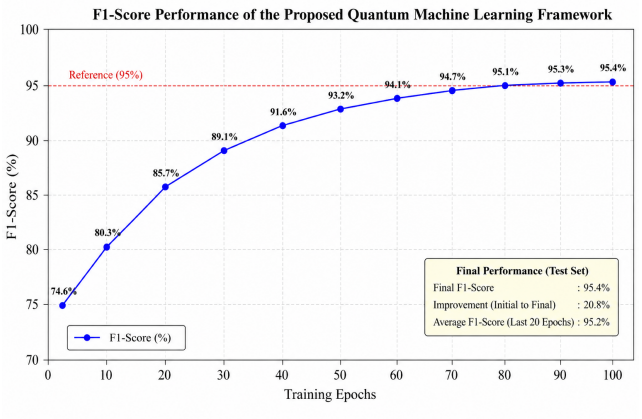

Figure 8: F1-Score performance of the proposed Quantum Machine Learning framework.

II. F1-Score Performance

As seen in Figure 8, the F1-Score was increasing steadily throughout the training phase while being kept constant during the final evaluation steps. The F1-Score was growing steadily due to the ability of the framework to distinguish between normal traffic and traffic with malicious intent as shown in Figure 8. In addition to that, this steady increase shows an appropriate balance between the two measures of precision and recall that is expected from an effective model. Finally, the high level of F1-Score was maintained until the end of the experiment.

D. Detection Rate Analysis

I. Attack Detection Rate

The detection rate of attacks was consistently high during the entire assessment process; however, small variations were noted depending on the training cycle. It is clear from

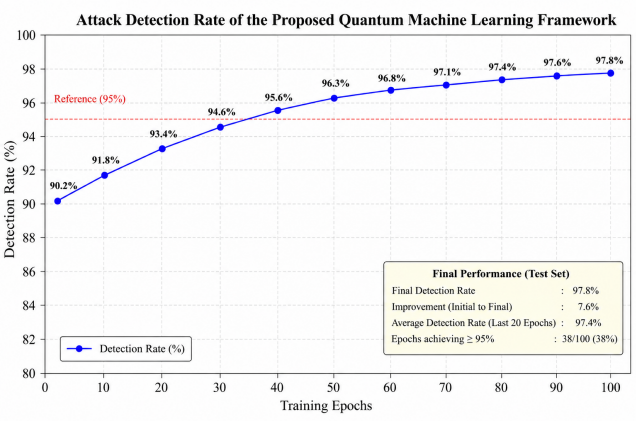

Figure 9: Attack detection rate of the proposed Quantum Machine Learning framework.

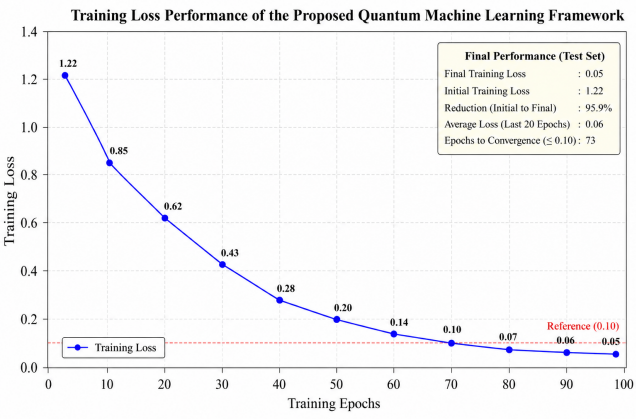

Figure 10: Training loss performance of the proposed Quantum Machine Learning framework.

Fig. 9 that the proposed Quantum Machine Learning model had excellent ability to detect attacks in a variety of network situations. This variation is attributed to the difference in the complexity of the attacks and traffic features in the dataset. However, regardless of the variations, the overall detection rate was steady and performed at a very high level during the process.

II. Training Loss Analysis

The loss of training was continually declining during the entire optimization process, implying that the learning capability of the proposed QML framework was continuously improving. This is evident from Figure 10, where there was a quick decline in the loss at the initial training phases, after which the loss started to converge to a steady minimum point. Such a behavior indicates that the proposed QML framework was performing very well, especially with respect to parameter optimization and model generalization.

IV. Conclusion

The performance outcomes recorded by the suggested Quantum Machine Learning approach highlight its efficiency in detecting cyber-attacks in critical infrastructures. These performances include high values for accuracy, precision, recall, F1-score, and attack detection rate with a relatively low false positive rate. In addition, the quantum feature encoding and variational quantum circuits provided by the model have made it capable of identifying and distinguishing network behaviors from malicious acts.

From the experimental results, there was consistent performance across several types of attacks such as DDoS attacks, malware attacks, insider attacks, and APTs. There were more comparisons with traditional machine learning algorithms to identify the strengths of the proposed quantum model since it provided better classification accuracy and enhanced the reliability of the system. It is also evident from the constant drop in the training loss that the framework could be able to learn different features of attacks effectively.

However, the results presented here show that Quantum Machine Learning can be a possible solution to provide security systems in the future. As the critical infrastructure networks become more complex, there is an increased need for efficient detection systems to ensure their safety. In future studies, one may explore the use of the framework with bigger datasets or even implement it using a quantum computer.

Acknowledgment

We would like to express our deepest gratitude to our respective institutions and research guides for their constant encouragement and guidance during the entire course of this research. Our gratitude goes out to everyone who helped us either directly or indirectly in the accomplishment of this research. We also wish to thank everyone whose help we received while performing data analysis, formulating our model, and writing this paper.

References

[1] P. W. Shor, “Algorithms for Quantum Computation: Discrete Logarithms and Factoring,” in *Proc. 35th Annual Symposium on Foundations of Computer Science*, 1994, pp. 124-134.

[2] L. K. Grover, “A Fast Quantum Mechanical Algorithm for Database Search,” in *Proc. 28th Annual ACM Symposium on Theory of Computing*, 1996, pp. 212-219.

[3] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2010.

[4] J. Preskill, “Quantum Computing in the NISQ Era and Beyond,” *Quantum*, vol. 2, p. 79, 2018.

[5] V. Dunjko and H. J. Briegel, "Machine Learning and Artificial Intelligence in the Quantum Domain," *Reports on Progress in Physics*, vol. 81, no. 7, 2018.

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

[7] M. Schuld, I. Sinayskiy and F. Petruccione, "A Introduction to Quantum Machine Learning," *Contemporary Physics*, vol. 56, no. 2, pp. 172-185, 2015.

[8] E. Farhi and H. Neven, "Classification with Quantum Neural Networks on Near Term Processors," arXiv:1802.06002, 2018.

[9] K. Beer et al., "Training Deep Quantum Neural Networks," *Nature Communications*, vol. 11, 2020.

[10] S. Lloyd, M. Mohseni and P. Rebentrost, "Quantum Algorithms for Supervised and Unsupervised Machine Learning," arXiv:1307.0411, 2013.

[11] Amaral, Cesar A., Vinícius L. Oliveira, Juan PLC Salazar, and Eduardo I. Duzzioni, "Quantum machine learning and quantum-inspired methods applied to computational fluid dynamics: a short review," arXiv preprint arXiv:2510.14099,2025.

[12] Adamos, Konstantinos, George Stergiopoulos, Michalis Karamousadakis, and Dimitris Gritzalis, "Enhancing attack resilience of cyber-physical systems through state dependency graph models," *International Journal of Information Security 23*, vol. no. 1, 2024.

[13] ENISA, "Threat Landscape Report," European Union Agency for Cybersecurity," 2024.

[14] Cisco Systems, "Annual Cybersecurity Report," Cisco, 2024.

[15] Khanam, Shapla, Ismail Bin Ahmedy, Mohd Yamani Idna Idris, Mohamed Hisham Jaward, and Aznul Qalid Bin Md Sabri."A survey of security challenges, attacks taxonomy and advanced countermeasures in the internet of things," *IEEE access*, pp. 219709-219743, nov. 2020.

[16] A. B. Hamida, M. Msahli and A. Mitrokotsa, "Security of the Internet of Things: A Survey," *Computer Networks*, vol. 148, pp. 283-294, 2019.

[17] S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach,"*Pearson*, 2021.

[18] Mortimer, Jason, "Cybersecurity for sustainable and digital economic transformations," 2022.

[19] I. Goodfellow, Y. Bengio and A. Courville, "Deep Learning," *MIT Press*, 2016.

[20] D. Berman et al., "A Survey of Deep Learning Methods for Cyber Security," *Information*, vol. 10, no. 4, 2019.

[21] Y. Xin et al., "Machine Learning and Deep Learning Methods for Cybersecurity," *IEEE Access*, vol. 6, pp. 35365-35381, 2018.

[22] A. Javaid et al., "A Deep Learning Approach for Network Intrusion Detection System," *EAI Endorsed Transactions on Security and Safety*, vol. 3, no. 9, 2016.

[23] M. Ring et al., "A Survey of Network-Based Intrusion Detection Data Sets," *Computers Security*, vol. 86, pp. 147–167, 2019.

[24] H. Hindy et al., "A Taxonomy of Network Threats and Intrusion Detection Systems," *Future Internet*, vol. 12, no. 1, 2020.

[25] P. Rebentrost, M. Mohseni and S. Lloyd, "Quantum Support Vector Machine for Big Data Classification," *Physical Review Letters*, vol. 113, 2014.

[26] M. Schuld and N. Killoran, "Quantum Machine Learning in Feature Hilbert Spaces," *Physical Review Letters*, vol. 122, 2019.

[27] K. Bharti et al., "Noisy Intermediate-Scale Quantum Algorithms," *Reviews of Modern Physics*, vol. 94, 2022.

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

[29] World Economic Forum, P. "Global cybersecurity outlook," *Geneva, Switzerland: World Economic Forum and Accenture*, 2024.

[30] International Telecommunication Union (ITU), "Global Cybersecurity Index," 2024.