computers

MDPI

Article

Parameterised Quantum SVM with Data-Driven Entanglement for Zero-Day Exploit Detection

Steven Jabulani Nhlapo [1,†], Elodie Ngoie Mutombo [2] and Mike Nkongolo Wa Nkongolo [1,*]

[1] Department of Informatics, University of Pretoria, Pretoria 0028, South Africa; u23963353@tuks.co.za
[2] Department of Computer Science, University of Pretoria, Pretoria 0028, South Africa; u22608754@tuks.co.za
* Correspondence: mike.wankongolo@up.ac.za
† Current address: Lynnwood Road and Roper Street, Hatfield, Pretoria 0028, South Africa.

Abstract

Zero-day attacks pose a persistent threat to computing infrastructure by exploiting previously unknown software vulnerabilities that evade traditional signature-based network intrusion detection systems (NIDSs). To address this limitation, machine learning (ML) techniques offer a promising approach for enhancing anomaly detection in network traffic. This study evaluates several ML models on a labeled network traffic dataset, with a focus on zero-day attack detection. Ensemble learning methods, particularly eXtreme gradient boosting (XGBoost), achieved perfect classification, identifying all 6231 zero-day instances without false positives and maintaining efficient training and prediction times. While classical support vector machines (SVMs) performed modestly at 64% accuracy, their performance improved to 98% with the use of the borderline synthetic minority oversampling technique (SMOTE) and SMOTE + edited nearest neighbours (SMOTEENN). To explore quantum-enhanced alternatives, a quantum SVM (QSVM) is implemented using three-qubit and four-qubit quantum circuits simulated on the `aer_simulator_statevector`. The QSVM achieved high accuracy (99.89%) and strong F1-scores (98.95%), indicating that nonlinear quantum feature maps (QFMs) can increase sensitivity to zero-day exploit patterns. Unlike prior work that applies standard quantum kernels, this study introduces a parameterised quantum feature encoding scheme, where each classical feature is mapped using a nonlinear function tuned by a set of learnable parameters. Additionally, a sparse entanglement topology is derived from mutual information between features, ensuring a compact and data-adaptive quantum circuit that aligns with the resource constraints of noisy intermediate-scale quantum (NISQ) devices. Our contribution lies in formalising a quantum circuit design that enables scalable, expressive, and generalisable quantum architectures tailored for zero-day attack detection. This extends beyond conventional usage of QSVMs by offering a principled approach to quantum circuit construction for cybersecurity. While these findings are obtained via noiseless simulation, they provide a theoretical proof of concept for the viability of quantum ML (QML) in network security. Future work should target real quantum hardware execution and adaptive sampling techniques to assess robustness under decoherence, gate errors, and dynamic threat environments.

Keywords: zero-day attacks; intrusion detection systems; machine learning; synthetic minority oversampling; quantum machine learning; UGRansome dataset

Academic Editor: Leandros Maglaras

Received: 9 July 2025
Revised: 5 August 2025
Accepted: 11 August 2025
Published: 15 August 2025

Citation: Nhlapo, S.J.; Mutombo, E.N.; Nkongolo, M.N.W. Parameterised Quantum SVM with Data-Driven Entanglement for Zero-Day Exploit Detection. *Computers* **2025**, *14*, 331. https://doi.org/10.3390/computers14080331

1. Introduction

In the digital era, organisations operate on a global scale, relying heavily on internet platforms for customer engagement and service delivery. Consequently, computing infras-

tructure has become the backbone of modern business operations. However, as technology continues to drive these operations, the threat of data breaches by malicious actors remains a persistent concern. Failure to implement adequate security measures can result in the loss of sensitive information, regulatory penalties, and reputational harm [1]. To mitigate such risks, many organisations deploy intrusion detection systems (IDSs), which monitor and alert regarding suspicious network behaviour [2]. Traditional IDSs, which are predominantly signature-based, operate by comparing incoming network traffic against a database of known attack patterns. While effective in identifying previously encountered threats, these systems fall short when confronted with zero-day attacks [3]. These attacks represent exploits that target unknown vulnerabilities before corresponding detection signatures are available (Figure 1).

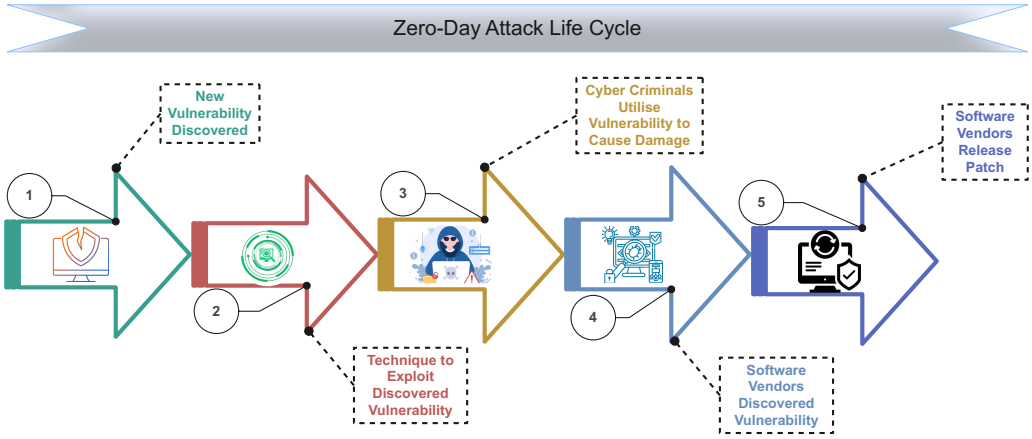


Figure 1. Zero-day attack life cycle. 1—New vulnerability discovered: an unknown flaw identified in a system or application; 2—Technique to exploit the discovered vulnerability: method developed to take advantage of the flaw; 3—Cybercriminals utilise vulnerability to cause damage: active exploitation impacting confidentiality, integrity, or availability; 4—Software vendors discover vulnerability: detection through reports, testing, or monitoring; 5—Software vendors release patch: official update issued to fix the flaw and mitigate risk.

As a result, signature-based IDSs are inherently limited in their ability to detect and respond to emerging cyber threats. Their reliance on frequent updates and high maintenance costs further limits their responsiveness and scalability, especially for smaller organisations [2,3]. Anomaly-based IDSs offer an alternative by modeling normal network behaviour and flagging deviations [4]. Nevertheless, they often suffer from high false positive rates. With this, IDSs are proving inadequate in addressing emerging threats, especially zero-day attacks [5]. These systems often require up to 30 days to detect such threats [6], leaving organisations exposed to substantial risks during that window [3]. The consequences of zero-day attacks are severe. The average cost of recovering from a zero-day breach exceeds USD 1.2 million [6], excluding business disruption, regulatory penalties, and long-term reputational damage [7]. Given that modern businesses rely on real-time digital platforms, any downtime can affect productivity and revenue. Despite investments in IDSs and non-technical mitigation strategies, organisations remain vulnerable. Moreover, attackers are becoming sophisticated by exploiting unknown vulnerabilities across diverse technologies.

Therefore, the continued use of outdated IDSs in the face of modern and evasive attacks has placed organisations at a serious disadvantage, and there is an urgent need to modernise IDS capabilities. Another challenge lies in the datasets used to train IDS models. Widely used benchmarks such as KDD-CUP99, NSL-KDD, UNSW-NB15, and CICIDS-2017 are outdated and fail to reflect the complexity of modern network threats [8].

These datasets often lack real-world zero-day attack data and can lead to high false positive rates due to poor generalisation, limiting the models' effectiveness in real deployments. Given these limitations, there is a growing interest in leveraging machine learning (ML) to enhance IDS capabilities [9]. ML techniques have demonstrated promise in detecting sophisticated and previously unseen attacks, including zero-day exploits [10]. By learning from large and diverse datasets, they can uncover hidden patterns and classify complex threat behaviours with improved accuracy and efficiency. This study utilises UGRansome [10], a contemporary dataset comprising real-world zero-day activities and modern network traffic behaviours [11]. The dataset includes three target classes, anomaly (A), signature (S), and synthetic signature (SS), which enable a more nuanced evaluation of model performance across attack types [6]. The research opts to evaluate the performance using standard metrics such as accuracy, precision, recall, F1-score, and computational time [6,9] and investigate the effectiveness of both classical and quantum ML (QML) models in detecting zero-day attacks using the UGRansome dataset. This research is guided by the following central question:

To what extent can classical and QML models accurately classify anomaly, signature, and synthetic signature categories in the UGRansome dataset for effective zero-day attack detection?

The first objective is to assess how well ML classifiers distinguish between the key target classes (A, S, SS) embedded within the dataset. This enables the identification of previously unseen attack patterns, which are critical for the advancement of adaptive IDS mechanisms. The second objective focuses on benchmarking the performance of quantum models, particularly quantum support vector machines (QSVMs), against conventional classifiers. Their performance is evaluated with the aim of understanding their applicability in real-world cybersecurity settings. The third objective is to address the issue of class imbalance, which is prevalent in intrusion detection datasets and often leads to biased learning outcomes. To mitigate this, the study incorporates the synthetic minority oversampling technique (SMOTE), borderline-SMOTE, and SMOTE with edited nearest neighbours (SMOTEENN), to synthetically augment underrepresented classes, particularly those representing zero-day attacks. This study introduces a novel quantum hybrid pipeline tailored for detecting zero-day attacks, offering conceptual advancements through the integration and customisation of QML components for cybersecurity contexts. Specifically, the key innovations include the following:

- Learnable nonlinear quantum encoding: Unlike standard fixed encodings, we design a parameterised and trainable encoding circuit that enables the quantum layer to adapt to the data manifold, effectively capturing latent structures specific to zero-day attack patterns.
- Mutual information-guided sparse entanglement: We propose a principled method for configuring entanglement in quantum circuits based on mutual information analysis of classical features. This departs from arbitrary or full entanglement schemes by offering a data-aware circuit design, reducing quantum overhead while preserving critical interdependencies.
- Quantum kernel optimisation for SVMs: We go beyond using standard QSVM kernels by constructing a quantum kernel informed by the encoded and entangled state space tailored to the cyber threat detection task, providing both theoretical and empirical improvements over classical SVM baselines.
- Domain-specific quantum simulation: While QML methods have been explored in generic classification tasks, our application to zero-day detection provides a unique and practically relevant testbed. The proposed framework demonstrates superior classification performance compared to existing techniques.

Although our work builds on established components, it offers a new conceptual synthesis and methodological enhancements that address the specific challenges of quantum cybersecurity modeling—notably, feature sparsity, pattern complexity, and adversarial dynamics inherent in zero-day exploits. The manuscript is structured as follows: Section 2 presents a review on zero-day attack classification, evaluating existing datasets and ML techniques. Section 3 introduces the proposed QSVM for zero-day attack detection. Section 4 details the results and analysis of the proposed QSVM, including a comparative evaluation with related works. Finally, Section 5 concludes the study.

2. Literature Review

The origins of ML and artificial intelligence (AI) date back to the 1950s and 1960s, with foundational contributions from researchers such as Alan Turing, John McCarthy, Arthur Samuel, Alan Newell, and Frank Rosenblatt [12]. Arthur Samuel developed the first functional ML model, while Rosenblatt introduced the perceptron algorithm inspired by biological neurons. This algorithm became the basis for artificial neural networks (ANNs) [12]. ML techniques are categorised into four groups: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning [9,12]. Supervised learning involves training algorithms on labeled datasets, where inputs and outputs are known, enabling models to make accurate predictions. Common methods include regression and classification using algorithms such as SVMs, decision trees (DTs), random forest (RF), naïve Bayes (NB), and neural networks (NNs) [13]. In recent years, supervised learning has expanded with the emergence of QML, which utilises principles from quantum computing to enhance the speed and complexity of pattern recognition tasks [14,15]. QML algorithms, such as the QSVM, use quantum bits (qubits) and superposition to represent and process information in ways that are infeasible for classical systems. Although still in its early stages [16], QML shows promise in solving high-dimensional classification problems [14] and has the potential to revolutionise domains like cybersecurity [17].

2.1. Established Datasets

The KDD Cup 99 dataset, developed by the Defense Advanced Research Projects Agency (DARPA), has been widely used in the development and evaluation of IDSs [18]. It contains over 4.4 million records representing both normal and attack traffic. A major limitation is the high volume of duplicate records, comprising 78% of the training set and 75% of the testing set [18]. In addition, 14 attack types present in the test set are absent from the training data, complicating model generalisation. Approximately 89% of the dataset comprises attack traffic, while only 11.15% is normal. Table 1 summarises the distribution of records by attack type.

Table 1. Comparison of KDD and NSL-KDD datasets: normal and attack data distribution.

Attack	KDD Cup 99		NSL-KDD	
	Records	**Distribution**	**Records**	**Distribution**
Normal	492,708	11.15%	67,343	53.46%
Probe	41,102	0.93%	11,656	9.25%
Denial of services (DoS)	3,883,370	87.89%	45,927	36.46%
User to root (U2R)	52	0.00%	52	0.04%
Remote to local (R2L)	1126	0.03%	995	0.79%
Total	4,418,358	100%	125,973	100%

The NSL-KDD dataset was developed by the Canadian Institute of Cybersecurity as an improved version of the KDD Cup 99 dataset [19]. It addresses some of the drawbacks

of KDD Cup 99, including duplication and imbalance issues (Table 1). NSL-KDD includes additional features such as connection duration, protocol types, services, and content features like failed login attempts and root access. Both the KDD Cup 99 and NSL-KDD datasets are unsuitable for zero-day attack detection because they contain outdated, known attack patterns and lack examples of novel threats. Their data imbalance and duplicates also bias models, limiting their ability to detect new and unseen attacks effectively. The UNSW-NB15 dataset, introduced in 2015 by researchers at the University of New South Wales [20], represents an advancement over earlier datasets like KDD Cup 99 and NSL-KDD by encompassing both legacy and modern attack types. It includes 49 features spanning statistical, content-based, and traffic-based attributes. The dataset covers nine attack categories alongside normal traffic, collected from a mix of real and synthetic network traffic in controlled environments (Table 2). Despite its comprehensive nature, UNSW-NB15 exhibits class imbalance with a predominant proportion of normal traffic (87.35%). This imbalance, coupled with its creation date, constrains its effectiveness in detecting more recent and sophisticated attacks, including zero-day exploits (Table 2). In turn, the CICIDS2017 dataset, published in 2017 by the Canadian Institute for Cybersecurity, addresses limitations found in previous datasets by incorporating modern network traffic scenarios and a wider variety of attack types [21]. Featuring 84 attributes, this dataset includes common attacks such as DoS, brute force, infiltration, port scanning, and botnet (Table 2). The dataset suffers from data integrity issues and an imbalance dominated by normal traffic (83.34%). Both datasets contribute valuable resources for intrusion detection research, yet their respective limitations demonstrate the ongoing need for updated and balanced datasets that reflect evolving cybersecurity threats (Table 3).

Table 2. Comparison of UNSW-NB15 and CICIDS2017 datasets: normal and attack data distribution.

Attack Type	UNSW-NB15		CICIDS2017	
	Number of Records	Distribution	Number of Records	Distribution
Normal	2,218,761	87.35%	2,359,087	83.34%
Analysis	2677	0.11%	—	—
Backdoors	2329	0.09%	—	—
DoS	16,353	0.64%	294,506	10.40%
Exploits	44,525	1.74%	—	—
Fuzzers	24,246	0.95%	—	—
Generic	215,481	8.50%	—	—
Reconnaissance	13,987	0.55%	—	—
Shellcode	1511	0.06%	—	—
Worms	174	0.01%	—	—
Botnet	—	—	1966	0.07%
Brute force	—	—	13,835	0.48%
Infiltration	—	—	36	0.0013%
Port scan	—	—	158,930	5.63%
Web attack	—	—	2180	0.08%
Total	2,540,044	100%	2,830,540	100%

2.2. The UGRansome Dataset

The dataset considered in this study is UGRansome, accessed on 13 August 2025 from https://www.kaggle.com/datasets/nkongolo/ugransome-dataset/data, a recent benchmark developed at the University of Pretoria. UGRansome is employed for zero-day attack and ransomware detection [13], serving as a basis for evaluating intrusion detection systems (IDSs) [4]. Besides benign traffic such as user datagram protocol (UDP), port scanning, scan, and secure shell (SSH), the dataset includes malicious traffic like blacklist, botnet, DoS, nerisBotnet, and spam [22]. It captures ransomware communication patterns,

providing network flow-based features such as bytes transferred, malicious addresses, duration, and protocol flags (Table 3). The dataset has a three-class structure, particularly useful for zero-day attack prediction because it captures both known (S, SS) and unknown (A) threat behaviours in a way that supports robust model generalisation (Table 4). This structure mimics real-world conditions by exposing the model to a spectrum of threat patterns [23], from known to entirely unknown, thereby enhancing its ability to detect zero-day attacks that do not match existing signatures (Table 5).

Table 3. Comparative overview of UGRansome dataset for NIDSs.

Aspect	UGRansome	UNSW-NB15	CICIDS2017
Purpose	Ransomware classification and zero-day detection	Intrusion detection for legacy and contemporary attacks	Detection of complex attacks including advanced persistent threats (APTs), DoS, and botnets
Source	Real network traffic with injected zero-day exploits	Traffic generator in a controlled lab with real attack scenarios	Simulated real-world traffic from various attack scenarios
Features	14 network features (protocol, bytes, flags, etc.)	49 statistical and content-based features	84 features including packet-level metadata and payload-based stats
Labels	A, S, and SS	Nine attack types vs. benign	Seven attack types vs. benign
Size	149,043 labeled records	2,540,044 records	2,830,540 records
Attack	Ransomware and zero-day exploits	DoS, exploits, fuzzers, generic, reconnaissance, shellcode, worms, backdoors	Botnet, brute force, DoS, infiltration, port scan, web attack
Timeframe	Real network over multiple weeks in 2021	Captured in 2015 using hybrid simulation	Collected in 2017 with timestamped flow sessions
Format	Comma-separated values (CSV)	CSV format with attack taxonomy	CSV format with attack labels
Use case	Evaluate ML models for detecting ransomware and zero-day intrusions	Evaluate NIDS performance on diverse and modern threats	Benchmark ML models for identifying attack behaviours

Table 4. UGRansome class labels and their descriptions. The symbol ✓ indicates that the class may contain benign and zero-day traffic, while × indicates known malicious traffic.

Class	Meaning	Benign?	Notes
A	Anomaly	✓ (sometimes)	Include benign and zero-day anomalies
S	Signature	×	Include known attacks (DoS and spam)
SS	Synthetic signature	✓ (simulated)	Include benign and unknown malicious behaviours

Table 5. Mapping of *UGRansome* features to quantum circuit encodings using the rotation gate around the Z-axis (R_Z), with example entries.

Column	Encoding	Feature	Entry	Description	R_Z Value
1	q_0	Time	50	Attack timestamp	0.22750
2	q_1	Protocol	ICMP	Network protocol	0.47943
3	q_2	Flag	ACK	Network status flag	0.12467
4	q_3	Family	TowerWeb	Zero-day exploit family	0.84147
5	q_4	Clusters	2	Numeric zero-day exploit cluster	0.00000
6	q_5	SeedAddress	17dcMo4V	Zero-day exploit address	0.47943
7	q_6	ExpAddress	1DiCeTjB	Zero-day formatted address	0.68164
8	q_7	BTC	22	Malicious Bitcoin transaction	0.05492
9	q_8	USD	18,124	Financial damages	0.40647
10	q_9	Netflow	4916	Network traffic volume	0.61945
11	q_{10}	IP	B	IP address class	0.61837
12	q_{11}	Threats	Botnet	Malware type	0.58510
13	q_{12}	Port	5061	Network port	0.14237
14	q_{13}	Prediction	A, S, SS	Target class label	0.47943

2.3. Machine Learning for Zero-Day Exploits Detection

Su [11] introduced a hybrid ML framework that integrates Chi-Square feature selection with the African vultures optimisation algorithm (AVOA) for hyperparameter tuning using the UGRansome dataset. The framework evaluates three classifiers, the extra trees classifier (ETC), extreme learning machine (ELM), and adaptive boosting classifier (ADAC), with ETC

optimised by the AVOA achieving the highest accuracy of 0.981. The Chi-Square method enhanced interpretability and performance by identifying "financial impact" and "clusters" as key features. The approach demonstrates strong potential for real-time deployment in cybersecurity systems [11] while highlighting the need for further work on adaptability to zero-day patterns.

Rios-Ochoa et al. [6] used the UGRansome dataset to address critical shortcomings in IDSs, including overreliance on binary classification, poor dataset quality, and the lack of real-time deployment. The authors introduced a comprehensive end-to-end framework covering feature extraction, dataset quality assessment, and comparative model training using conventional ML and deep learning (DL) classifiers. While RF achieved 100% offline accuracy, the multilayer perceptron (MLP) outperformed it in real-time deployment (70% vs. 62%).

The study emphasised that high offline performance does not guarantee real-world effectiveness, revealing that low feature variability and outdated data hinder generalisation. The authors suggest that specific models may improve detection robustness for zero-day variants.

Mohamed et al. [10] introduced a novel probabilistic composite model for zero-day exploit detection, focusing on enhancing anomaly detection in terms of accuracy, adaptability, and computational efficiency. The proposed framework integrates four innovative components: (1) an Adaptive WavePCA-Autoencoder (AWPA) for denoising and dimensionality reduction during preprocessing, (2) a Meta-Attention Transformer Autoencoder (MATA) to enhance feature extraction, (3) Genetic Mongoose-Chameleon Optimisation (GMCO) for efficient and accurate feature selection, and (4) an Adaptive Hybrid Exploit Detection Network (AHEDNet) to reduce false positives. Experimental results using the UGRansome dataset demonstrated superior performance, achieving accuracy up to 0.9919, precision up to 0.9968, and minimal error. These results confirm the model's robustness in detecting previously unseen threats.

Yan et al. [13] implemented a two-layer ML framework for malware detection using the UGRansomware (UGRansome) dataset. The first layer employs a stacked ensemble model combining six classifiers. The second layer uses light gradient boosting machine (LightGBM) to classify the detected malware into specific families. While this layer performs well on major families, it struggles with less represented categories, showing a noticeable drop in accuracy. The study depicts various challenges in fine-grained malware family attribution, particularly under class imbalance.

Sokhonn et al. [24] proposed a privacy-preserving hierarchical clustering method performed on the UGRansome dataset, demonstrating the feasibility of conducting clustering on encrypted network traffic data without exposing the underlying information. The results from the proposed encrypted method were visually consistent with SciPy's plaintext clustering output. Due to the fixed-point arithmetic and approximate operations, minor discrepancies were observed in cluster assignments. These inconsistencies may impact applications where clustering fidelity is critical.

Chaudhary and Adhikari [25] investigate the effectiveness of DT, SVM, and MLP for malware detection using the UGRansome dataset. Through a comparative analysis based on accuracy, precision, recall, and F1-score, the DT emerged as the most effective model. While the study evaluates ransomware detection, it does not explicitly simulate zero-day attacks with novel ransomware variants.

2.4. QML for Zero-Day Exploits Detection

QML utilises principles of quantum computation to enhance classical ML tasks [26]. A QML model consists of three key components: quantum data encoding, param-

eterised quantum circuits (PQCs), and quantum measurements for classification or regression [27]. Quantum information is represented using *qubits*, which are unit vectors in a two-dimensional Hilbert space $\mathcal{H}$ [26,27]. A single qubit can be written as a linear superposition of orthonormal basis states $|0\rangle$ and $|1\rangle$ (Equation (1)).

$$|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle\,, \quad \text{where } \alpha, \beta \in \mathbb{C}, \quad |\alpha|^2 + |\beta|^2 = 1 \tag{1}$$

An n-qubit system lives in a 2^n-dimensional space $\mathcal{H}^{\otimes n}$, enabling QML models to represent complex data structures compactly [28]. Classical input data $\mathbf{x} \in \mathbb{R}^d$, with d features, must be encoded into quantum states using a quantum feature map (QFM) represented by a unitary operator $\mathcal{U}_\phi(\mathbf{x})$ acting on the initial state $|0\rangle^{\otimes n}$ (Equation (2)).

$$|\phi(\mathbf{x})\rangle = \mathcal{U}_\phi(\mathbf{x})\,|0\rangle^{\otimes n} \tag{2}$$

Here, $\mathcal{U}_\phi(\mathbf{x}) : \mathcal{H}^{\otimes n} \to \mathcal{H}^{\otimes n}$ is a unitary operator parameterised by the classical input $\mathbf{x}$, which encodes the data into the quantum $\mathcal{H}$ [28,29]. Learning in QML is driven by parameterised unitary transformations $\mathcal{U}_\theta$, where $\theta \in \mathbb{R}^m$ are trainable parameters (e.g., rotation angles of quantum gates such as R_x, R_y, and R_z). The overall parameterised quantum state is obtained by applying $\mathcal{U}_\theta$ after the data encoding, yielding (Equation (3))

$$|\psi(\mathbf{x};\theta)\rangle = \mathcal{U}_\theta\,\mathcal{U}_\phi(\mathbf{x})\,|0\rangle^{\otimes n} \tag{3}$$

The final state $|\psi(\mathbf{x};\theta)\rangle$ is measured to obtain classical outcomes for downstream tasks. For binary classification tasks, a common observable is the Pauli-Z operator Z, typically applied to a designated qubit within the tensor product of Pauli operators that define the measurement basis [30]. The expected value of this measurement is given by (Equation (4))

$$f(\mathbf{x}) = \langle\psi(\mathbf{x};\theta)|Z|\psi(\mathbf{x};\theta)\rangle \tag{4}$$

The scalar value $f(\mathbf{x})$ is then thresholded (e.g., by a sigmoid function) to produce the predicted label [31]. To train the QML model, a cost function $C(\theta)$ is defined based on cross-entropy loss for classification or mean squared error (MSE) for regression [31,32]. The goal is to find the optimal parameters θ^* that minimise this cost function, often using hybrid quantum optimisation methods (Equation (5)).

$$\theta^* = \arg\min_\theta C(\theta) \tag{5}$$

This mathematical formalism underpins QML models such as quantum convolutional neural networks (QCNNs) [33], variational quantum classifiers (VQCs) [30], and hybrid quantum deep learning (HQDL) [15]. These models can potentially achieve better generalisation on high-dimensional data by exploiting quantum parallelism and entanglement [34].

2.5. Review of QML Approaches in Cyber Threat Detection

Elsedimy et al. [35] introduced a novel intrusion detection framework, the QSVM grey wolf optimiser (GWO), designed to improve detection performance and reduce false positive rates in hybrid IDSs (HIDSs). The model integrates a QSVM with a GWO to fine-tune parameters. The QSVM is utilised for binary classification tasks by selecting an appropriate kernel function to maximise the decision boundary. Experimental evaluation is conducted using an internet of things (IoT) dataset.

The performance of the QSVM is benchmarked against several state-of-the-art models. The QSVM-GWO achieves a training accuracy of 99.11% and outperforms existing models. Similarly, Eze et al. [36] compared classical models (SVM, CNN) with their quantum counterparts (QSVM, QCNN) for malicious uniform resource locator (URL) detection,

demonstrating quantum kernel advantages for QSVM, though QCNNs lagged due to hardware constraints.

Vijayalakshmi et al. [37] implemented QSVMs to detect anomalies in network traffic using the NSL-KDD dataset. Four different kernel functions (linear, polynomial, radial basis function (RBF), and sigmoid) were applied to evaluate their impact on QSVM performance. The results demonstrated that the RBF kernel provided better separation or decision boundaries between normal and anomalous traffic.

Mahdian and Mousavi [38] implemented QSVMs on a quantum hardware to detect and classify quantum entanglement using VQCs. The model achieved 90% classification accuracy, even under the presence of noisy intermediate-scale quantum (NISQ) devices. The authors tested various circuit configurations across different quantum processors, including Perth, Lagos, and Nairobi, to evaluate the model's robustness. The results demonstrate that the proposed approach not only performs well on two-qubit systems but also extends effectively to three-qubit entangled states.

In contrast, Tripathi et al. [39] developed a quantum long short-term memory (QLSTM) architecture using VQCs to analyse malware system calls, optimising quantum circuit depth and qubit configurations to overcome hardware limitations. The QLSTM detected distributed DoS (DDoS) attacks in network traffic, reporting consistent gains over Saeed [40]'s framework, which incorporated LSTM. Abreu et al. [16] developed and evaluated QuantumNetSec, an HIDS that combines quantum and classical computing techniques to improve network threat detection. By applying personalised QML methods, the system identified both binary and multiclass threats.

Testing on publicly available datasets showed that QuantumNetSec outperformed conventional models. Extending these efforts, Sridevi et al. [41] introduced a hybrid quantum classical neural network (HQCNN) integrating latent Dirichlet allocation (LDA) and wavelet transforms with CNN and quantum layers, achieving superior performance in Android malware and DDoS classification on real hardware.

Durgut et al. [15] explored the use of a hybrid quantum classical deep neural network (HQCDNN) to identify security vulnerabilities in smart contracts. This model aimed to detect access control issues, arithmetic errors, time manipulation, and DoS. The SmartBugs wild dataset is used for training and evaluation, with Term Frequency–Inverse Document Frequency (TF-IDF) employed for feature extraction. Experiments were conducted using configurations with two-qubit and four-qubit quantum layers, alongside a classical deep neural network (DNN) for baseline comparison. Results achieved 78.2–96.4% accuracy and 80.2–96.6% F1-score, surpassing conventional models.

This review emphasises that QML's practical advantage in intrusion detection hinges on both hardware maturity and targeted problems. Our study contributes to this growing body of work by empirically validating quantum kernel superiority and proposing a reproducible pipeline for secure threat detection using QSVMs, while addressing deployment feasibility under current hardware constraints [42].

3. Methodology

Figure 2 presents the proposed modular architecture that illustrates each pipeline component, from feature preprocessing to quantum kernel computation and prediction. These components are described in the following subsections.

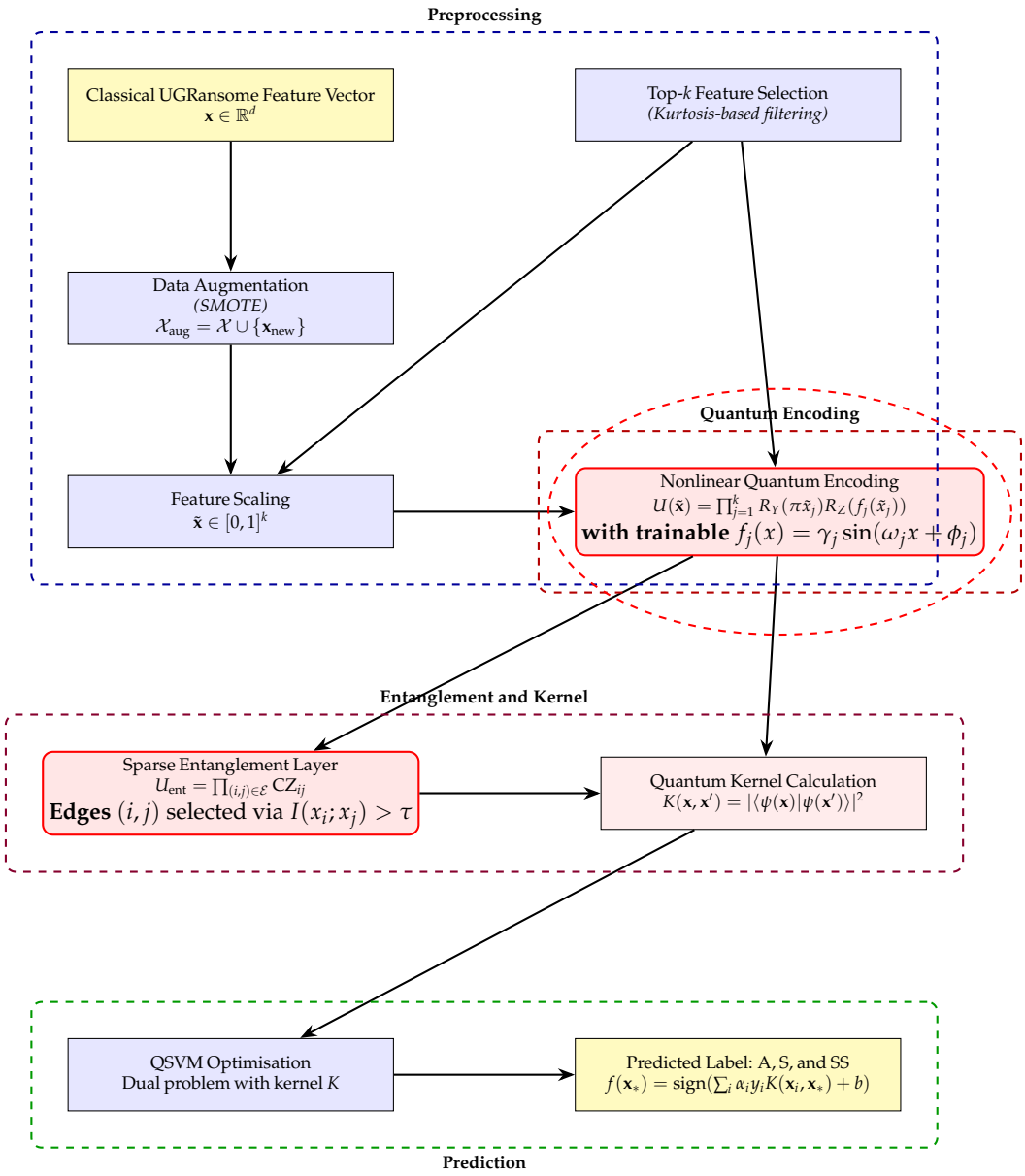


Figure 2. The proposed QSVM architecture.

3.1. Data Preprocessing and Machine Learning Setup

During the data analysis phase, several preprocessing steps were performed to ensure data quality and consistency. This included correcting spelling errors, removing duplicate attributes, and filtering outliers. The version of the UGRansome dataset used in this study did not contain any duplicated or missing entries, but several numeric features, such as `time`, `United States dollar (USD)`, `Bitcoin (BTC)`, `clusters`, and `netflow_bytes`, contained outliers, which were filtered out. In addition, negative values in the `time` feature were corrected by adding a constant offset of +11 s to all entries. Incorrect labels were identified and fixed for entries such as `botnet` and `nerisBotnet` under the `threat` feature. Similarly, misspelled feature names such as `protcol` and `seddAddress` were corrected to `protocol` and `seedAddress`. An improperly formatted entry in the `expAddress` feature represented by the value `1`, which did not conform to the expected 8-character format, was removed from the dataset. This research adopts a supervised ML approach, as the UGRansome dataset contains labeled instances. The selected models include SVM, NB, RF classifier (RFC), XGBoost, and an ensemble method (EM) [43]. These models were chosen based on their efficiency and robustness in solving classification and prediction

problems [26,43,44]. Their performance is evaluated using standard metrics like computational time, accuracy, F1-score, and ROC curve analysis. Furthermore, SMOTE is applied to enhance the performance of underperforming models by addressing class imbalance [45].

3.2. Chi-Square Feature Selection

Chi-Square (χ^2) is a feature selection used in this study to evaluate the independence between categorical input features and a categorical target variable [46]. It measures how expected counts of values under the null hypothesis of independence deviate from observed counts. In the context of supervised learning, the χ^2 test helps identify the most informative features by ranking them based on their statistical significance [46]. A higher χ^2 score indicates greater dependence between the feature and the target variable, implying that the feature carries more predictive power [46]. Features with the highest χ^2 scores are selected for model training. In this work, we applied the χ^2 feature selection method to the UGRansome dataset. Prior to computation, continuous features were normalised using Min-Max scaling to meet the non-negativity requirement of the χ^2 test [46]. The top $k = 6$ features with the highest χ^2 scores were selected and later visualised using t-distributed stochastic neighbour embedding (t-SNE) for class separability inspection (Figure 3). This selection aids in dimensionality reduction and improves the interpretability and efficiency of the ML pipeline [11].

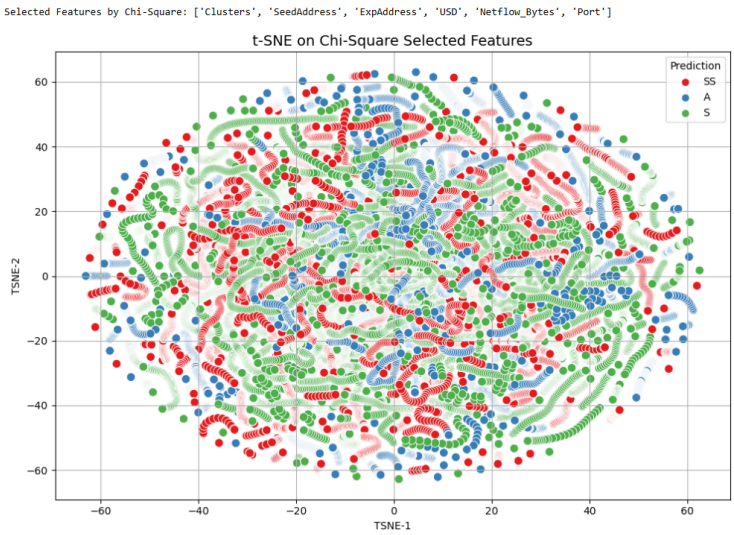


Figure 3. The UGRansome data distribution.

3.3. Data Transformation for Feature Selection

The data transformation process required for applying the χ^2 involves a series of preprocessing steps to ensure both statistical validity and compatibility with the test. First, categorical features such as `protocol`, `flag`, `seedAddress`, and `threats` were encoded into numerical values using the `LabelEncoder` utility from Scikit-learn [47]. This encoding is necessary because the χ^2 test requires discrete numerical inputs to compute statistical dependence. Next, continuous features including `BTC`, `USD`, `netflow_bytes`, and `time` were normalised to fall within the range $[0, 1]$. This step is essential as the χ^2 test assumes non-negative input values. The transformation is defined in Equation (6).

$$x_{\text{scaled}} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{6}$$

After the data was properly encoded and scaled, the χ^2 scores for each feature were computed using the `SelectKBest` method with the χ^2 scoring function. The top k features

exhibiting the strongest statistical dependence with the target class label (A, S, SS) were selected to reduce the dimensionality of the dataset while retaining the most informative attributes. To further evaluate the separability of classes based on the selected features, the data was projected into a lower-dimensional space using the t-SNE technique (Figure 3). This nonlinear dimensionality reduction method maps the selected k-dimensional feature vectors to a 2D space defined in Equation (7).

$$\text{t-SNE} : \mathbb{R}^k \rightarrow \mathbb{R}^2 \tag{7}$$

The transformation facilitates a visual interpretation of how well different zero-day classes are separated in the reduced feature space (Figure 3). It serves to ensure that the χ^2 feature selection is mathematically sound, computationally efficient, and visually informative for downstream ML tasks [47,48]. Figure 4 shows selected feature values, indicating that certain attributes behave differently across attack types, which is important for zero-day detection.

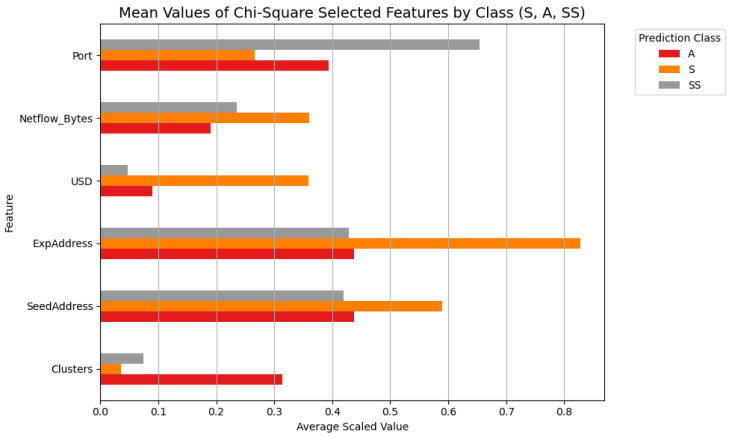


Figure 4. Selected features.

High expAddress values (above 0.8) in signature attacks (S) suggest this feature strongly represents known attack patterns, making it less useful for detecting novel zero-day threats. Similarly, elevated port values (above 0.6) in synthetic signature attacks (SS) highlight its role in identifying engineered attacks but may not fully capture zero-day anomalies. Conversely, moderate values of seedAddress and expAddress (above 0.4), along with port and clusters (above 0.3) in anomaly classes (A), reflect their potential in detecting unusual or previously unseen behaviours typical of zero-day exploits. Thus, features that show moderate variation in anomalies are more valuable for zero-day detection, as they help identify deviations from known patterns rather than established signatures [49].

3.4. Selected Models

The ensemble method is a powerful ML technique that combines predictions from multiple models to improve overall performance [50]. The core idea is that aggregating the strengths of diverse models often results in better accuracy than any single model alone. Common ensemble techniques include bagging, boosting, stacking, and voting [50]. This study employs the voting technique, a popular approach for classification tasks. A voting classifier aggregates predictions from several base classifiers and makes a final decision based on majority or weighted voting [51]. While voting combines predictions from distinct models, bagging reduces variance by training the same model on multiple bootstrap-sampled subsets of the data, which is especially effective for high variance learners like tree-based classifiers [52]. This research adopts voting ensemble learning, combining NB, RFC, and XGBoost to leverage their complementary strengths and enhance

zero-day detection accuracy. To overcome the performance limitations of classical SVM in detecting complex patterns, particularly in zero-day attack scenarios, the study explores the integration of a QSVM. The QSVM enhances the classical SVM by utilising quantum computing to compute kernel functions in a high-dimensional space, allowing for more expressive feature mappings [27]. This quantum-enhanced kernel approach improves the model's ability to capture nonlinear relationships and class boundaries, thus optimising the performance of classical SVMs that struggle with overlapping or intricate feature distributions [53].

3.5. Quantum SVM Pipeline with Adaptive Feature Encoding

This study presents a novel QSVM pipeline that integrates an adaptive quantum feature encoding strategy, specifically designed to detect zero-day cyberattacks. Unlike standard quantum kernels such as `ZZFeatureMap` [54], which apply uniform rotations and linear entanglement [15], our model employs a data-driven encoding scheme tailored to the structural dependencies within the feature space: Let the original classical input be $\mathbf{x} = (x_1, x_2, \ldots, x_d)$, where d is the number of features. We select the top k features (Equation (8)):

$$\tilde{\mathbf{x}} = \left(x_1^{\text{scaled}}, x_2^{\text{scaled}}, x_3^{\text{scaled}}, x_4^{\text{scaled}}\right) \tag{8}$$

where $k = 6$ in our configuration and the features are scaled to the interval $[0, 1]^k$. These are then embedded into a quantum state using a customised unitary operator $U(\tilde{\mathbf{x}})$, yielding the quantum state in Equation (9).

$$|\psi(\tilde{\mathbf{x}})\rangle = U(\tilde{\mathbf{x}})|0\rangle^{\otimes k} \tag{9}$$

The kernel function $K(\mathbf{x}, \mathbf{x}')$ is defined as the fidelity between two quantum states (Equation 10).

$$K(\mathbf{x}, \mathbf{x}') = \left|\langle\psi(\mathbf{x})|\psi(\mathbf{x}')\rangle\right|^2 \tag{10}$$

This quantum kernel is used to solve the SVM dual problem (Equation (11)):

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2}\sum_{i,j=1}^{N} \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{i=1}^{N} \alpha_i \\ \text{s.t.} \quad & \sum_{i=1}^{N} \alpha_i y_i = 0, \quad 0 \le \alpha_i \le C \end{aligned} \tag{11}$$

where α_i are Lagrange multipliers, $y_i \in \{-1, 1\}$ are class labels, and C is the regularisation constant [55]. The decision function for a new input $\mathbf{x}_*$ is

$$f(\mathbf{x}_*) = \text{sign}\left(\sum_{i=1}^{N} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}_*) + b\right) \tag{12}$$

and performance is evaluated by accuracy (Equation (13)):

$$\text{Accuracy} = \frac{1}{M}\sum_{i=1}^{M} \mathbf{1}(y_i = \hat{y}_i) \tag{13}$$

where $\mathbf{1}(\cdot)$ is the indicator function, M is the number of test samples, and $\hat{y}_i$ the predicted labels [54,55]. The proposed QSVM architecture illustrated in Figure 2 introduces a modular, scalable, and NISQ-compatible quantum pipeline tailored for zero-day exploit detection. Unlike prior QSVM studies limited to synthetic datasets and unscalable encodings [15,54,55], our method uses a dimensional quantum embedding scheme, selective

entanglement for hardware feasibility, and kernel estimation via statevector simulation [56]. Its application to the UGRansome dataset provides a novel empirical step toward practical quantum cybersecurity implementation.

3.6. Contributions and Formalisation

To complement the architectural innovations introduced in Figure 2, we present a formal mathematical underpinning that solidifies the novelty of our pipeline.

3.6.1. Parameterised Quantum State Encoding

Our circuit prepares the input quantum state via a structured nonlinear encoding scheme by letting the scaled input vector be $\tilde{\mathbf{x}} = [\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k] \in [0,1]^k$ and defining the circuit in Equation (14):

$$|\psi(\tilde{\mathbf{x}})\rangle = U(\tilde{\mathbf{x}})\,|0\rangle^{\otimes k} = \prod_{j=1}^{k} R_Y(\pi \cdot \tilde{x}_j) R_Z(f_j(\tilde{x}_j))\,|0\rangle^{\otimes k} \tag{14}$$

Here, $f_j(x) = \gamma_j \cdot \sin(\omega_j x + \phi_j)$ is a feature-dependent nonlinear rotation with tunable parameters $\gamma_j, \omega_j, \phi_j \in \mathbb{R}$. This enables a basis expansion similar to Fourier feature maps and promotes expressivity in $\mathcal{H}$ [57].

3.6.2. Sparse Entanglement Topology

Instead of all-to-all connectivity, we define an entanglement graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $(i,j) \in \mathcal{E}$ iff the mutual information $I(x_i; x_j)$ exceeds a threshold τ expressed in Equation (15).

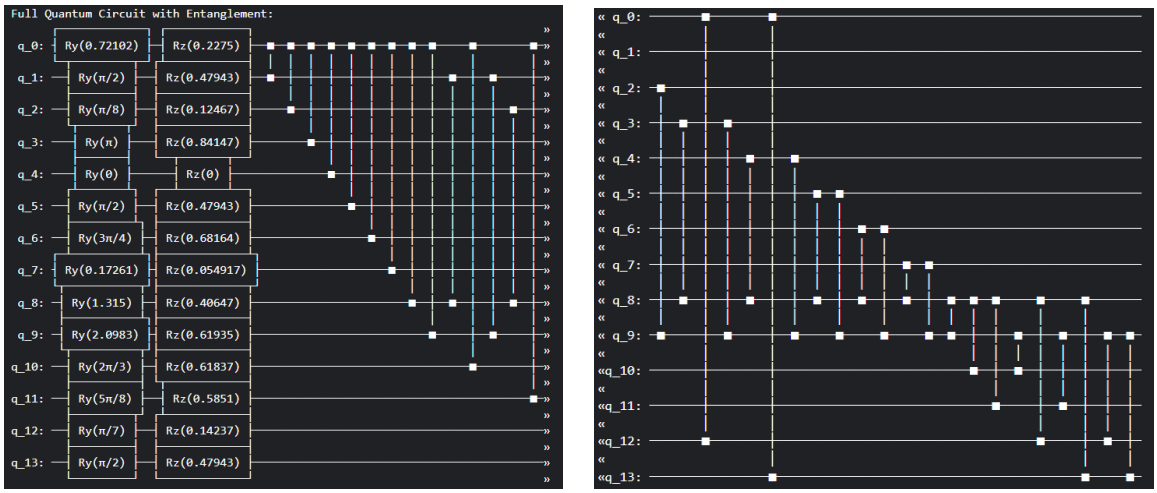
$$\mathcal{E} = \{(i,j) \mid I(x_i; x_j) > \tau\} \tag{15}$$

The corresponding entanglement unitary is given in Equation (16).

$$U_{\text{ent}} = \prod_{(i,j)\in\mathcal{E}} \text{CZ}_{ij} \tag{16}$$

The time complexity of the quantum circuit is $\mathcal{O}(k + |\mathcal{E}|)$, where k is the number of features encoded via single-qubit rotations and $|\mathcal{E}|$ is the number of entangling gates determined by the sparsity of the mutual information graph $\mathcal{G}$. Figure 5 confirms that the circuit is constructed correctly with the intended rotations and entanglement.

Equations (14)–(16) formalise our contributions, by (i) extending angle encoding with learnable nonlinear mappings, (ii) introducing information-theoretic entanglement criteria to reduce circuit complexity, and (iii) defining a dynamic sparse entanglement layer. Together, they establish a theoretical foundation for quantum-enhanced zero-day attack classification in NISQ regimes [58]. Unlike conventional quantum feature maps that use fixed, linear entanglement (e.g., entanglement = linear in ZZFeatureMap) [59], our design introduces a modular encoding scheme where entanglement is adaptively constructed based on classical mutual information between features (Figure 5). This allows the circuit to faithfully represent data-driven dependencies while preserving a shallow depth, as proven under a mutual information sparsity assumption (Figure 5). Thus, we address the limitations of simplistic entanglement topologies and introduce a more expressive yet NISQ-compatible alternative.



(**a**) (**b**)

Figure 5. Quantum circuit representation of 14 UGRansome attributes encoded into a multi-qubit quantum state (see Table 5 for attribute descriptions). (**a**) Quantum encoding circuit. (**b**) Full quantum circuit with entanglement.

3.7. QSVM Scalability and Limitations

The proposed QSVM is implemented using 3-qubit and 4-qubit configurations, consistent with the capabilities of NISQ-era [27]. Although the pipeline is evaluated using noiseless simulators, it was applied to the UGRansome dataset [11]. This combination provides a meaningful proof-of-concept for applying NISQ-compatible quantum models to practical cybersecurity problems. As quantum hardware matures, the proposed approach can be directly mapped to real quantum devices [60], enabling scalable threat detection under realistic constraints. The current implementation is conducted on the `aer_simulator_statevector` backend, which provides an idealised noise-free environment. This offers several development advantages: it enables deterministic, reproducible experimentation, facilitates debugging, and allows full access to the quantum statevector for post hoc analysis and explainability [61]. These capabilities are essential for evaluating novel architectural components such as adaptive entanglement or custom feature encodings without interference from hardware-induced noise. Nevertheless, the exclusive use of simulators limits the real-world relevance of performance claims, as quantum noise, decoherence, and gate infidelities on actual hardware are not accounted for. The results should therefore be interpreted as theoretical upper bounds on achievable performance [62,63]. While 3- and 4-qubit QSVM models are not sufficient for industrial-scale applications, they serve as an important proof of concept. Sparse entanglement supports smooth scaling to larger qubit counts as hardware matures. Under such conditions, the QSVM could offer substantial advantages in high-dimensional kernel learning, including polynomial or exponential speedups in learning complex decision boundaries for cyber threat detection and other classification tasks.

3.8. Critical Implementation Details

Table 6 depicts parameters used in the proposed framework. The experiments were conducted within a Proxmox virtual environment, utilising key libraries such as Pandas, NumPy, Keras, TensorFlow, and Scikit-learn.

Table 6. Key parameters used in the proposed QSVM pipeline.

Component	Parameter	Purpose
Quantum encoding layer	γ_j, ω_j, ϕ_j (trainable per-feature parameters)	Nonlinear sinusoidal encoding of classical features into quantum states
Feature selection	Top-k ($k = 6$)	Selection of most informative features based on kurtosis or variance
Entanglement strategy	Mutual information threshold τ ($\tau = 0.05$)	Defines sparse CZ gate connectivity preserving essential feature dependencies
Quantum kernel construction	Fidelity calculation between quantum states	Computes state overlaps $\langle\psi(x_i)\|\psi(x_j)\rangle$ for kernel matrix
SVM classifier	Regularisation parameter C (default or tuned)	Controls margin–error trade-off in classical SVM training
Circuit depth/layers	1 to 3 layers (variational circuit depth)	Controls expressivity and complexity of the QFM
Optimiser	`AdamOptimizer`, learning rate = 0.001	Optimises trainable parameters in the quantum circuit
Hyperparameter tuning	Randomised grid search with 5 folds	Tuning hidden layers (1-3), learning rate (10^{-4} to 10^{-2}), batch size (16, 32, 64), epochs (50-100)
Early stopping	Patience of 10 epochs on cross-entropy loss	Prevents overfitting by stopping training after no improvement
Quantum backend	`default.qubit` simulator (PennyLane)	Noise-free quantum circuit simulation for 3- and 4-qubit configurations
Experiment replication	Averaged over 10 independent runs	Mitigates variance from quantum randomness and simulation noise
Train/test split	Stratified k-fold cross-validation	Ensures robust evaluation across zero-day classes
Hardware environment	Intel Core i9-13900K CPU and 64 GB of DDR4 RAM	Computational platform for all experiments
Software environment	Jupiter Notebook Python 3.12, PennyLane-Qiskit v0.36.0	Quantum programming environment and libraries

3.9. Data Splitting

A random sampling technique is employed to divide the dataset into training and testing subsets. This technique ensures that each instance has an equal probability of being selected for either the training or testing set [64]. The approach randomly assigns approximately 80% of the data to the training set and 20% to the testing set. In this study, the training data, represented by `X_train` and `y_train`, consists of 89,523 samples. Here, `X_train` includes the feature vectors, where each row corresponds to an individual sample and each column represents a specific attribute. The associated `y_train` vector contains the target labels corresponding to each row in `X_train`, allowing ML algorithms to learn the correct input–output mappings. The remaining 22,381 samples form the testing set, represented by `X_test` and `y_test`. `X_test` contains previously unseen input data, while `y_test` includes the actual target labels used to assess the model's predictive performance. This setup enables the comparison between the model's predictions and the ground-truth values, facilitating an objective evaluation of the trained model's generalisation capabilities.

3.10. Handling Class Imbalance

To address class imbalance, advanced oversampling techniques such as borderline-SMOTE and SMOTEENN are employed [65]. These methods extend the original SMOTE by refining the generation of synthetic samples and combining oversampling with data cleaning. Borderline-SMOTE focuses on synthetic sample generation using *borderline* instances or minority classes that are most at risk of being misclassified [66]. Instead of generating new samples uniformly across all minority instances, it identifies samples near the decision boundary where the ratio of majority-class neighbours is high [65,66]: Let x_i

be a minority-class sample, and let k be the number of nearest neighbours ($k = 6$). This method defines (Equation (17))

$$\text{danger}(x_i) = \begin{cases} \text{True} & \text{if } \frac{\#\text{majority neighbors}}{k} \geq \delta \\ \text{False} & \text{otherwise} \end{cases} \tag{17}$$

Here, δ is a threshold ($\delta = 0.5$). Synthetic samples are generated in Equation (18).

$$\tilde{x} = x_i + \lambda(x_i^{(nn)} - x_i) \tag{18}$$

where $x_i^{(nn)}$ is a randomly chosen minority neighbour of x_i and $\lambda \sim \mathcal{U}(0,1)$ is a random weight. This approach ensures that the synthetic data is concentrated in regions of high classification uncertainty, improving classifier focus on difficult examples. SMOTEENN combines oversampling (SMOTE) with undersampling through edited nearest neighbours (ENN), which removes noisy and ambiguous instances from the dataset after SMOTE has been applied [67]. Let $D = D_{\text{maj}} \cup D_{\text{min}}$ be the imbalanced dataset, where SMOTE first generates synthetic minority samples D_{syn} (Equation (19)).

$$D' = D_{\text{maj}} \cup D_{\text{min}} \cup D_{\text{syn}} \tag{19}$$

After oversampling, ENN is applied to clean the dataset by removing any instance (from both classes) that disagrees with the majority of its K-NNs (Equation (20)).

$$\text{Remove } x_i \in D' \text{ if } \hat{y}_i \neq \text{majority class among } \mathcal{N}_k(x_i) \tag{20}$$

This two-step approach allows SMOTEENN to both balance the class distribution and eliminate mislabeled or ambiguous samples, reducing noise and improving the generalisation performance of ML models [67]. Both borderline-SMOTE and SMOTEENN are employed to address the challenges posed by the imbalanced nature of the UGRansome dataset [10,11]. These techniques are particularly valuable in network intrusion detection tasks, where zero-day attacks are underrepresented and difficult to detect using conventional ML approaches.

3.11. Evaluation Metrics

The performance of the ML models is evaluated using standard classification metrics: *accuracy* (acc), *precision* (prec), *recall* (rec), *F1-score*, *false positive rate (FPR)*, and the *receiver operating characteristic area under the curve (ROC-AUC)* [68]. Accuracy quantifies the overall correctness of the model by computing the ratio of correct predictions to total predictions (Equation (21)) [68].

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{21}$$

Precision measures the proportion of true positives (TP) among all predicted positives, especially important for imbalanced datasets (Equation (22)) [9].

$$\text{Prec} = \frac{TP}{TP + FP} \tag{22}$$

Recall, or sensitivity, measures the proportion of actual positives that are correctly identified (Equation (23)) [11].

$$\text{Rec} = \frac{TP}{TP + FN} \tag{23}$$

The false positive (FP) rate (FPR) measures the proportion of actual negatives that are incorrectly classified as positives (Equation (24)) [68].

$$\text{FPR} = \frac{FP}{FP + TN} \tag{24}$$

The F1-score is the harmonic mean of precision and recall [10], offering a single metric that balances both concerns (Equation (25)).

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{25}$$

The ROC-AUC score represents the area under the ROC curve, which plots the true positive rate (TPR) against the FPR across different threshold settings (Equation (26)).

$$\text{AUC} = \int_0^1 \text{TPR}(x)\, dx \tag{26}$$

A higher AUC value indicates better classifier performance [9]. A value of 1.0 represents a perfect model, while a value below 0.5 indicates performance worse than random guessing. In addition, the confusion matrix provides a tabular summary of prediction outcomes, helping to visualise the distribution of true and false predictions for each class (Table 7).

Table 7. Confusion matrix for ML predictions across classes A, S, and SS. Red highlights indicate true positives (TP), and yellow highlights indicate false positives (FP) and false negatives (FN).

Predicted	A	S	SS
A	TP	FN	FP
S	FP	TP	FN
SS	FP	FN	TP

These values are essential for computing the above metrics. TP represents the number of instances correctly predicted as zero-day exploits (A), known threats (S), and synthetic attacks (SS).

4. Results

The UGRansome dataset is divided into training and testing sets, with 80% (89,523 samples) allocated to training and 20% (22,381 samples) to testing. The training set, comprising input features (X_{train}) and corresponding labels (y_{train}), is used to train ML models by learning the relationships and patterns within the data. The test set ($X_{\text{test}}, y_{\text{test}}$) evaluates model generalisation on unseen samples. This setup enables unbiased performance evaluation based on metrics such as accuracy, precision, recall, and F1-score. Models tested include SVM, NB, RFC, XGBoost, the voting ensemble method, and QSVM with lower-performing classifiers enhanced using SMOTE data-balancing techniques.

1. Support Vector Machine

To establish a classical benchmark, we employ `LinearSVC` from the `Scikit-Learn` library, training on the top-k encoded features with default hyperparameters and up to 100 iterations. As shown in Table 8, the model achieves balanced F1-scores of 65% for both anomaly (A) and signature (S) classes but performs poorly on the synthetic signature (SS) class with a sharp recall drop to 42%, despite an unusually high precision of 1.00. This imbalance indicates that the classifier is overly conservative, confidently predicting SS only in clear-cut cases, leading to a high false negative rate—a critical flaw in the context of detecting zero-day threats. The confusion matrix in Figure 6a further reveals a bias toward

misclassifying samples, exhibiting the model's reliance on linear separability and its limited ability to capture nuanced attack behaviour. These results highlight structural limitations rooted in class imbalance, insufficient feature discrimination, and the inadequacy of linear decision boundaries.

Table 8. SVM evaluation metrics on test data.

Target	Precision	Recall	F1	Support
A	0.56	0.78	0.65	6231
S	0.60	0.70	0.65	9457
SS	1.00	0.42	0.59	6693
Accuracy		0.64 (22,381 samples)		

(**a**)

Predicted	A	S	SS
A	4838	1393	0
S	2806	6649	2
SS	955	2957	2781

(**b**)

Predicted	A	S	SS
A	7375	0	0
S	0	13,540	0
SS	0	0	8811

(**c**)

Predicted	A	S	SS
A	7375	0	0
S	0	13,540	0
SS	0	0	8811

(**d**)

Predicted	A	S	SS
A	7373	0	0
S	0	13,542	0
SS	0	0	8811

Figure 6. Comparison of confusion matrices: (**a**) SVM (baseline). (**b**) SVM borderline-SMOTE. (**c**) SVM SMOTEENN. (**d**) QSVM. Blue cells indicate header rows and columns, labeling the predicted (top row) and actual (left column) classes (A, S, SS). Red cells indicate correct predictions (TP or TN), and yellow cells indicate misclassifications (FP or FN) (Table 7).

The inability of the linear SVM to reliably detect SS attacks poses a security risk. In practice, such blind spots mean that obfuscated threats often representative of zero-day exploits can evade early detection, undermining system resilience. The model's tendency to misclassify SS samples into more familiar categories reflects a deeper issue (Table 8): linear classifiers may be unsuited for capturing behavioural deviations that characterise modern threats. Consequently, we pivot to more expressive models and techniques—specifically, quantum-enhanced classification with data-level augmentation—as suggested by Mohanty et al. [69], to address these weaknesses and improve detection.

2. SVM with Borderline-SMOTE

To overcome the limitations observed in the baseline SVM, particularly the poor detection of SS attacks, we employed the `SGDClassifier` with hinge loss to approximate a linear SVM, coupled with `CalibratedClassifierCV` for probability estimation. Borderline-SMOTE was applied to generate synthetic samples concentrated near the decision boundary of the minority SS class, enhancing the model's ability to separate overlapping regions. This approach led to near-perfect classification, with precision, recall, and F1-scores approaching 1.00 across all classes (Table 9). In particular, the recall for SS rose from 42% in the baseline to 100%, eliminating the prior detection gap for emerging threats. The confusion matrix (Figure 6) and ROCs (Figure 7) support this, showing minimal misclassification and perfect AUCs (Table 9).

Table 9. SVM borderline-SMOTE evaluation metrics.

Target	Precision	Recall	F1	Support
A	1.00	1.00	1.00	7375
S	1.00	1.00	1.00	13,608
SS	0.99	1.00	1.00	8811
Accuracy		0.998 (29,794 samples)		

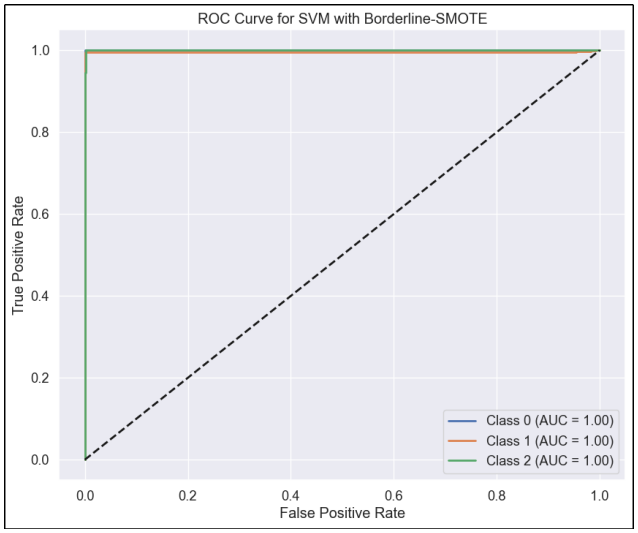


Figure 7. ROC-AUC curve: SVM with borderline-SMOTE.

However, the extremely high performance raises concerns about possible overfitting due to the synthetic nature of the training data. Since borderline-SMOTE constructs new data points based on existing borderline instances, it may inadvertently introduce redundancy or amplify noise, especially if the original minority class was poorly sampled. Thus, while the results are promising, further evaluation is necessary to validate the model's generalisation capability (Figure 8).

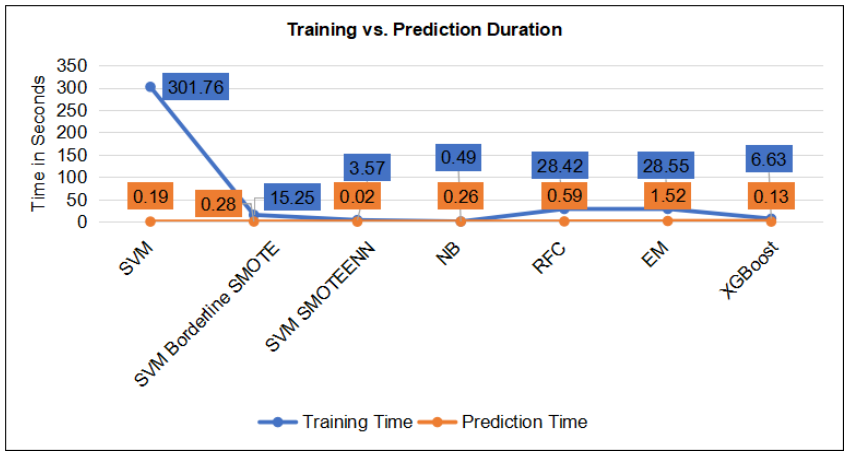


Figure 8. Training and prediction duration per model.

3. SVM with SMOTEENN

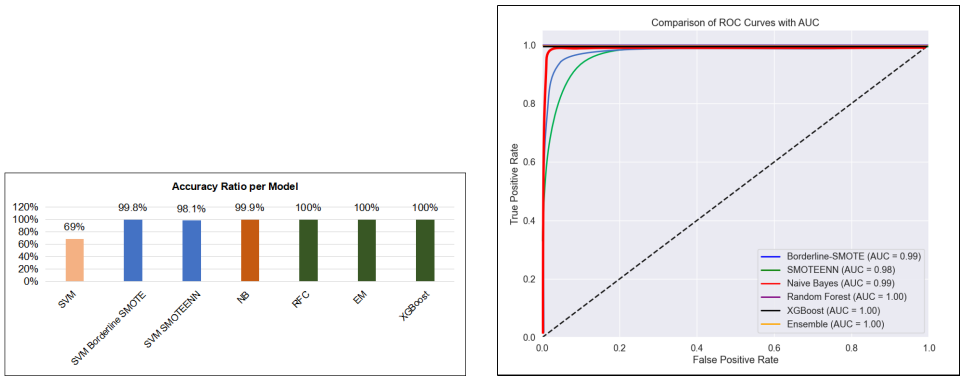
To further address class imbalance, the model incorporates SMOTEENN, which combines K-NN-based synthetic minority oversampling with edited nearest neighbours (ENN) to remove noisy majority samples. This hybrid approach is implemented using a `SGDClassifier` with a `StandardScaler`, offering efficient linear approximation and stable convergence. The model trained in 3.76 s and made predictions in 0.05 s (Table 10).

While overall classification metrics are high, the SS class often linked to zero-day or synthetic threats exhibited a slight drop in recall, suggesting residual difficulty in modeling its overlapping patterns. This limitation is visualised in the confusion matrix (Figure 6), though the ROCs remain strong with AUCs of 1.0 across all classes (Figure 9). Despite the high discriminative power, time complexity remains a critical consideration. As shown in Figure 8, even small delays in training or inference can reduce a model's effectiveness in fast-evolving threat environments. Models that require longer training cycles may struggle to adapt to rapidly changing attack vectors, such as zero-day exploits, which often demand continuous learning on large data streams.

Table 10. SVM SMOTEENN evaluation metrics.

Target	Precision	Recall	F1-Score	Support
A	1.00	1.00	1.00	7375
S	0.96	1.00	0.98	13,608
SS	0.99	0.94	0.97	8811
Accuracy		0.981 (29,794 samples)		

Similarly, slower prediction times can hinder real-time intrusion detection, reducing the system's ability to contain threats before lateral movement occurs. Table 11 summarises the comparative performance, with ensemble learning methods standing out as the most robust approach, achieving perfect scores in precision, recall, and F1 across all attack categories (Figure 9a,b).



(**a**) (**b**)

Figure 9. Performance comparison of model accuracy and AUC. (**a**) Model accuracy rate. (**b**) AUC comparison between models.

The XGBoost model stands out as the most effective and practical choice for zero-day attack detection (Figure 9), delivering perfect accuracy, precision, recall, and F1-score alongside efficient training (6.63 s) and prediction times (0.13 s). While the RFC and ensemble models also achieved flawless detection rates, their comparatively longer training and inference times (Figure 8) may hinder timely deployment in fast-evolving threat environments. The NB model offers rapid processing and high overall accuracy (99%), but its lower recall on zero-day anomalies risks missing critical attacks, potentially increasing system vulnerability. The baseline SVM, though initially underperforming with 64% accuracy and high training overhead, benefited from SMOTE augmentation, improving accuracy to over 98%, but introduced false positives in non-zero-day classes, which could lead to increased operational costs. These findings demonstrate the importance of balancing detection performance with computational efficiency and false alarm rates. By maintaining this balance, XGBoost emerges as the most reliable solution, ensuring prompt, accurate

detection of zero-day exploits while minimising resource consumption and false alerts (Figure 9).

Table 11. Evaluation metrics comparison using precision (prec), recall (rec), and F1-score (F1). Grey-shaded columns highlight the precision values of the best-performing model.

Class	NB				RFC				Ensemble				XGBoost			
	Prec	**Rec**	**F1**	**Test**	**Prec**	**Rec**	**F1**	**Test**	**Prec**	**Rec**	**F1**	**Test**	**Prec**	**Rec**	**F1**	**Test**
A	0.99	0.99	1.00	6267	1.00	1.00	1.00	6231	1.00	1.00	1.00	6231	1.00	1.00	1.00	6231
S	0.99	0.98	0.98	9526	1.00	1.00	1.00	9457	1.00	1.00	1.00	9457	1.00	1.00	1.00	9457
SS	0.97	0.99	0.98	6588	1.00	1.00	1.00	6893	1.00	1.00	1.00	6693	1.00	1.00	1.00	6693
Accuracy	0.99 (22,381)				1.00 (22,381)				1.00 (22,381)				1.00 (22,381)			

4.1. Threat Classification Divergence and Model Behaviour

The experiment reveals notable divergence in classification outcomes stemming from fundamentally different feature representations employed by SMOTE-based oversampling methods and quantum kernels. While both approaches seek to address class imbalance, they influence decision boundaries through distinct mechanisms. Quantum kernels, utilised in QSVMs, map data into higher-dimensional spaces where complex, nonlinear relationships, particularly in densely populated regions, can be better captured to improve generalisation. However, this can also shift class boundaries in unexpected ways [70], as evidenced by instances where samples originally classified as anomalies (A) are reassigned to the signature-based (S) class. Conversely, oversampling methods such as borderline-SMOTE and SMOTEENN explicitly rebalance the dataset by generating synthetic minority samples, directly reshaping the classifier's decision boundaries toward underrepresented threat classes. This fundamental difference explains why classical and quantum-enhanced models interpret borderline regions differently, particularly in the imbalanced UGRansome dataset comprising three distinct threat types. These divergences are critical because they impact the model's ability to accurately differentiate subtle threat behaviours, which is essential for reliable zero-day attack detection. Understanding these shifts highlights the importance of evaluating per-class confidence levels and applying model calibration techniques to mitigate misclassification risks. Thus, this analysis not only contextualises performance differences observed across models but also motivates further investigations into their interpretability and operational reliability in various cybersecurity environments.

4.2. QSVM Evaluation

Building on the observed divergence in threat classification behaviour, the QSVM model was implemented using `Qiskit` and integrated with `scikit-learn` to enable hybrid quantum classical learning [71]. To investigate how quantum feature encoding impacts decision boundaries, the model employed the `ZZFeatureMap` with sparse entanglement and two repetitions, selected for its ability to map nonlinear relationships in the input space [71–73]. Both three-qubit and four-qubit circuits were evaluated to assess how increasing quantum state dimensionality affects the model's capacity to differentiate between threat behaviours (Table 12). This setup provides insights into how quantum-enhanced kernels influence class separability, especially in borderline or imbalanced regions identified in earlier comparisons. To further strengthen the performance results and explain model behaviour, multiple kernel types were tested, including linear, polynomial, and RBF for classical SVMs, and compared against quantum kernels used in the QSVM implementation.

The quantum experiments were conducted using the `aer_simulator_statevector` backend with 1024 shots, offering a high-fidelity and noise-free environment to emulate quantum state evolution. Although real quantum hardware such as `ibm_nairobi` (a superconducting qubit processor) was considered [72], it was not employed due to practical

constraints, such as gate errors, decoherence, and execution latency, which can undermine result consistency.

Table 12. Comparison of SVM and QSVM performance using different kernels and numbers of qubits. Green shading highlights improvements in performance due to kernel configuration and the number of qubits (3 or 4).

Model	Class	Qubits	Kernel	Acc	Prec	Rec	F1
SVM	A	–	Linear	0.5912	0.5600	0.7800	0.6500
	A	–	Polynomial	0.6134	0.5731	0.7400	0.6452
	A	–	RBF	0.6287	0.5823	0.7654	0.6607
QSVM	A	3	Quantum	0.9415	0.9703	0.9415	0.9556
	A	4	Quantum	0.9462	0.9748	0.9961	0.9989
SVM	S	–	Linear	0.5980	0.6000	0.7000	0.6500
	S	–	Polynomial	0.6120	0.6100	0.7050	0.6548
	S	–	RBF	0.6314	0.6150	0.7231	0.6648
QSVM	S	3	Quantum	0.9794	0.9798	0.9794	0.9796
	S	4	Quantum	0.9911	0.9819	0.9831	0.9825
SVM	SS	–	Linear	0.5712	0.5613	0.4200	0.5900
	SS	–	Polynomial	0.5834	0.5800	0.4300	0.5980
	SS	–	RBF	0.6017	0.5631	0.4450	0.6117
QSVM	SS	3	Quantum	0.9961	0.9734	0.9961	0.9845
	SS	4	Quantum	0.9989	0.9805	0.9989	0.9895

The simulator-based setup thus ensured reproducibility and eliminated noise-related performance variance [73], allowing for a focused evaluation of the QSVM's structural advantages. Feature selection was conducted using `SelectKBest` with the Chi-Squared (χ^2) statistic, retaining the six most discriminative features following `MinMaxScaler` normalisation. The QSVM model consistently outperformed its classical counterparts across both three-qubit and four-qubit configurations, as shown in Table 12, reinforcing the positive impact of QFM on classification boundaries, as discussed by Sihare [74]. QSVM exhibited stronger generalisation to minority classes, particularly the anomaly (A) class representing zero-day threats. High recall on this class indicates robust detection of true positives, which is critical for early response in dynamic threat environments. Concurrently, elevated precision values reflect the model's ability to reduce false alarms, enhancing trust and operational feasibility. These results validate the effectiveness of quantum-enhanced learning under realistic class imbalance, supporting the case for integrating quantum models in threat classification pipelines [75]. The experiment reveals a clear trend: increasing the number of qubits from three to four enhances the QSVM's ability to capture complex patterns associated with zero-day attacks (Table 12). Specifically, the four-qubit model improves recall for the anomaly class from 99.61% to 99.89%, with a corresponding gain in F1-score and overall accuracy. This performance gain is not incidental but stems from the increased representational capacity of the QFM. A higher qubit count expands the Hilbert space in which data is embedded, enabling the model to construct more expressive and nonlinear decision boundaries. Such expressiveness is particularly crucial for detecting zero-day threats, which often exhibit deviations from known patterns. This improvement supports a central hypothesis of QML: deeper quantum encodings allow for richer, more discriminative kernel spaces, thereby enhancing generalisation on previously unseen data. The boost in recall does not compromise precision, indicating the QSVM maintains high specificity and resists overfitting, even as its sensitivity increases. This balance is essential

in security applications where false negatives may permit critical threats to go undetected. Further supporting this conclusion, Table 13 shows that the QSVM achieves near-perfect recall for high-risk attack classes such as APT (0.9989) and nerisBotnet (0.9831), with corresponding high precision and F1-scores. These findings portray the model's ability to generalise effectively across diverse threat profiles, validating its utility in real-world cybersecurity environments. In general, the results demonstrate that quantum kernel scalability, achieved via increased qubit encoding, directly translates to measurable improvements in zero-day exploit detection, solidifying QSVM as a promising tool for proactive and resilient cyber defense systems.

Table 13. QSVM performance on specific attack types by prediction label (A, S, SS). Green shading highlights the highest evaluation metric values for each attack type, indicating the best-performing predictions for zero-day exploits.

Attack	Predicted	Accuracy	Precision	Recall	F1
Blacklist	A	0.9462	0.9311	0.9462	0.9386
Botnet	S	0.9788	0.9644	0.9788	0.9715
NerisBotnet	A	0.9831	0.9739	0.9831	0.9785
Spam	S	0.9647	0.9521	0.9647	0.9584
Cryptohitman	A	0.9961	0.9734	0.9961	0.9845
APT	A	0.9989	0.9805	0.9989	0.9895
EDA	A	0.9415	0.9207	0.9415	0.9309
JigSaw	SS	0.9877	0.9692	0.9877	0.9783
SamSam	SS	0.9923	0.9754	0.9923	0.9838

4.3. Impact of Quantum Feature Mapping

Beyond the performance gains attributed to increased qubit capacity, the choice of QFM plays a central role in shaping the QSVM's effectiveness. As shown in Table 13, the model consistently achieves high recall across diverse attack types, including evasive threats such as *APT*, *cryptohitman*, and *nerisBotnet*, highlighting its capacity to minimise false negatives, a critical requirement in cyber threat intelligence. This sensitivity, however, comes with slightly reduced precision in certain categories (e.g., *blacklist*: 93.11%, *exploratory domain attack (EDA)*: 92.07%), reflecting a modest increase in false positives. Such a trade-off is common in security-focused applications, where failing to detect an attack is more costly than triggering a false alarm. In the proposed framework, this balance is implicitly governed by hyperparameter tuning and the structure of the QFM, which determines the expressivity of the induced kernel space. The proposed QFMs embed classical data effectively and enable the separation of nonlinear patterns. A more expressive map tends to improve recall by capturing subtle feature relationships but may also increase boundary overlap, thereby lowering precision. Conversely, simpler mappings reduce overfitting risk and false positives. In our configuration, the selected QFM strikes an effective balance, enhancing detection across threat classes while maintaining F1-scores above 0.93. Moreover, circuit design choices—such as the entanglement strategy, repetition depth, and number of qubits—interact with the QFM to fine-tune the performance trade-off, as recommended by Wang [76]. As such, adapting the quantum circuit to the characteristics of specific threat landscapes is essential for achieving optimal detection performance in real-world scenarios.

4.4. Discussion

The experimental results presented in Figure 10 provide critical insights into the performance of various models for detecting zero-day threats, categorised into three primary

classes: A-type (e.g., *APT*, *cryptohitman*, *nerisBotnet*), S-type (e.g., *locky*, *cryptoLocker*, *wannaCry*, *spam*, *DoS*), and SS-type (e.g., *jigSaw*, *towerWeb*, *razy*, *botnet*). Among the evaluated models, the QSVM achieved a remarkable accuracy of 99.8%, with 70% of its predictions successfully targeting SS-class threats. This underscores the strength of quantum kernels in capturing complex and latent relationships, particularly those present in polymorphic attack vectors. However, the QSVM exhibited limited sensitivity to S-type threats, identifying only 10% in that category. This performance gap may reflect the quantum model's bias toward novel or entangled patterns. In contrast, the RFC achieved a perfect accuracy of 100%, with a well-balanced threat detection profile across classes: 35% S-type, 25% A-type, and 40% SS-type. The ensemble architecture of RFC enables it to generalise across diverse attack landscapes, making it particularly effective in heterogeneous environments. Similarly, the NB classifier attained 99% accuracy but exhibited a skewed detection preference, capturing 45% A-type threats—suggesting potential overfitting to structured attack patterns while underperforming in the S-type category (15%).

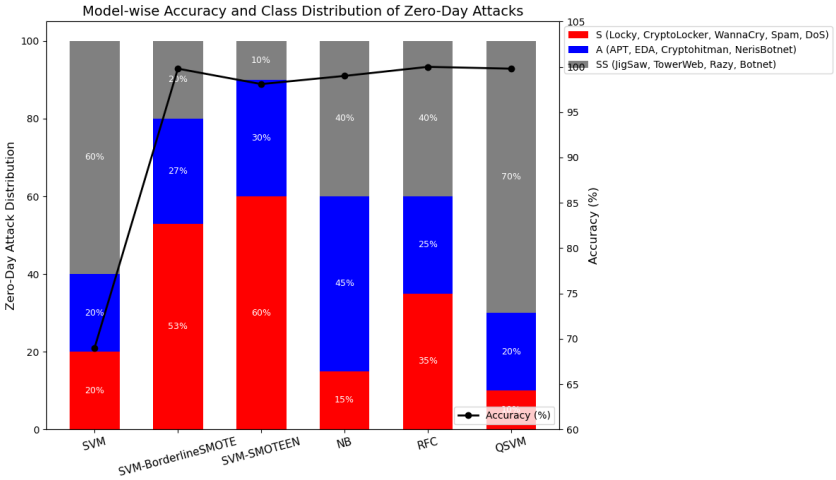


Figure 10. Model accuracy rate comparison.

SVMs enhanced with data-balancing strategies showed varied behaviour. The SVM augmented with borderline-SMOTE delivered a strong accuracy of 99.8%, achieving the most balanced detection across classes: 53% S-type, 27% A-type, and 20% SS-type. This outcome highlights the effectiveness of focusing synthetic samples near decision boundaries to mitigate class imbalance and improve generalisability. Conversely, the SVM using SMOTEENN reached 98.1% accuracy but skewed heavily toward S-type threats (60%), likely due to the inclusion of noisy examples during oversampling, which may distort decision boundaries. The baseline SVM, with no resampling, yielded the lowest performance (69% accuracy) and uniform class distributions, reinforcing the importance of preprocessing and imbalance mitigation in zero-day modeling. Overall, the results demonstrate that the QSVM and RFC are both highly effective in detecting emerging SS and A-type threats, while the SVM with borderline-SMOTE offers the most balanced performance across all categories, making it a compelling option for real-world zero-day detection scenarios.

4.5. Discussion on Perfect Performance Results

While several models including ensembles such as XGBoost achieved perfect accuracy on the evaluated dataset, such results are uncommon in real-world cybersecurity scenarios and warrant cautious interpretation. To mitigate risks of overfitting and data leakage, we adhered to rigorous evaluation protocols involving clean dataset splits, stratified cross-validation, statistical significance testing, and comprehensive preprocessing to prevent label leakage or sample duplication [77]. Nonetheless, these exceptional scores may partly

reflect dataset-specific factors, including limited diversity of zero-day variants and the use of synthetic balancing techniques that can simplify classification. Given the dynamic and evolving nature of zero-day attacks, which often present unpredictable patterns, ongoing validation through real-time deployment and adaptive learning is essential to assess the true operational robustness of these models. Among the approaches reviewed, ensemble learning methods demonstrate superior performance, particularly in terms of classification accuracy (Figure 11). Models such as ETC, XGBoost, and RF achieve high detection rates, though few assess zero-day attack scenarios. Additionally, quantum-based models such as QSVM and QuantumNetSec show promising results but face practical limitations (Table 14), including restricted qubit counts and the absence of real-time or explainable outputs.

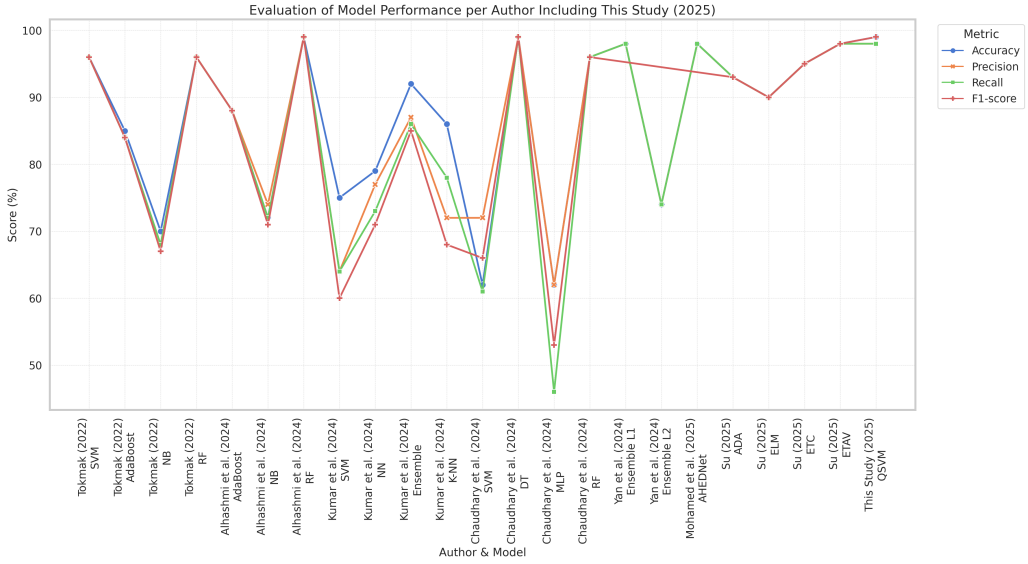


Figure 11. Comparison with existing studies [10,11,13,25,78–80].

4.6. Comparison with Existing Studies

As shown in Table 14, the majority of studies utilising the UGRansome dataset for IDSs do not implement QML for zero-day exploit detection (Figure 11). Figure 11 shows the proposed QSVM outperforming several classical methods evaluated in prior studies (Table 14). With accuracy and F1-scores of 99%, our QSVM demonstrated superior detection capabilities, particularly for zero-day attacks. Compared to models such as AdaBoost (85–93%), NB (68–74%), and conventional SVM (62–96%), the proposed QSVM showed higher precision and recall. These gains reflect the model's enhanced ability to generalise and capture complex and high-dimensional data structures typical of novel threat patterns. The improvement in F1-score suggests a balanced optimisation of precision and recall, critical in minimising both false positives and false negatives in IDS. While QML and data-balancing techniques such as SMOTE are established in isolation, this study advances beyond a mere aggregation of known methods. Our work proposes a novel and modular QSVM encoding framework grounded in information theory that introduces nonlinear and learnable QFMs that adapt to zero-day attacks. Furthermore, we establish a mutual information-based criterion for entanglement, enabling dynamic qubit connectivity that reflects data dependencies rather than fixed entanglement patterns. This architecture supports scalable and interpretable circuit design, bridging the gap between explainability and quantum kernel methods. Thus, our contribution extends beyond engineering optimisation and introduces methodological innovation relevant to real-world cybersecurity modeling in quantum-enhanced systems. Table 15 provides a comparative analysis of recent QML approaches for intrusion detection.

Table 14. Summary of NIDS studies using UGRansome dataset. A checkmark (✓) indicates the dataset was used, while a cross (×) indicates it was not used.

Author	Models	UGRansome?	Metrics	Results	Comments
Su [11]	ETC, ELM, + AVOA	✓	Accuracy	98.1%	Good real-time potential; needs zero-day adaptation and quantum computation
Rios-Ochoa et al. [6]	RF, MLP, DT, CNN, SVC	✓	Accuracy	70%	Offline vs. real-time gap; family-specific models without QML
Mohamed et al. [10]	AHEDNet	✓	Accuracy	9.19%	Highly robust; adaptive to zero-day exploits without quantum kernel
Alhashmi et al. [79]	AdaBoost and XGBoost	✓	Accuracy	99%	No multi-class zero-day support with QML
Yan et al. [13]	Ensemble + LightGBM	✓	Accuracy	90%	Class imbalance affects zero-day detection; missing quantum encoding
Chaudhary and Adhikari [25]	DT, SVM, MLP	✓	F1	99.41%	Good baseline; no zero-day simulation using quantum computing
Elsedimy et al. [35]	QSVM	×	F1 and ROC	97%	Missing explainability and real-time simulation of quantum models
Mahdian and Mousavi [38]	QSVM	×	Accuracy	90%	Quantum hardware constraints; binary classification
Abreu et al. [16]	QuantumNetSec	×	F1	92%	Limited qubits encoding
Durgut et al. [15]	HQCDNN	×	F1	96%	Simulated quantum environment
Tokmak [80]	DNN	✓	F1	97%	Hybrid model; no quantum kernel
This Study	Ensemble and QSVM	✓	F1	99.8%	Controlled environment using UGRansome; borderline-SMOTE and SMOTEENN; no deployment on actual quantum hardware

Unlike prior works that mostly rely on standard feature maps (e.g., ZZFeatureMap or Pauli encodings) and limit experimentation to benchmark datasets such as CICIoT2023 or CIC-IDS2017 (Figure 12), our study introduces an adaptive nonlinear encoding with data-driven sparse entanglement tested on UGRansome for zero-day exploit detection.

Table 15. Comparison of QML-based intrusion detection studies.

Author	Dataset	Feature Selection	Model	Quantum Logic	Metrics	Results
[81]	UNSW-NB15, CICIoT2023	-	QSVM	ZZFeatureMap	F1	88–94–82%
[82]	CIC-IoT2022, SDN-DDoS24	CICFlowMeter	QNN	Pauli-XYZ	F1	85–78%
[83]	CARLA	Quantum computation	QCAE, QOC-SVM	Hadamard, Pauli-X	F1	99%
[84]	CIC-CIDDS-2018	Granger causality	QGAN	Pauli Z	F1	98%
[85]	Cardholder	K-best, heuristic	VQC	-	Acc	94%
This study	UGRansome	Kurtosis filtering, t-SNE	QSVM	Nonlinear rotation-CZ with sparse entanglement	F1	99%

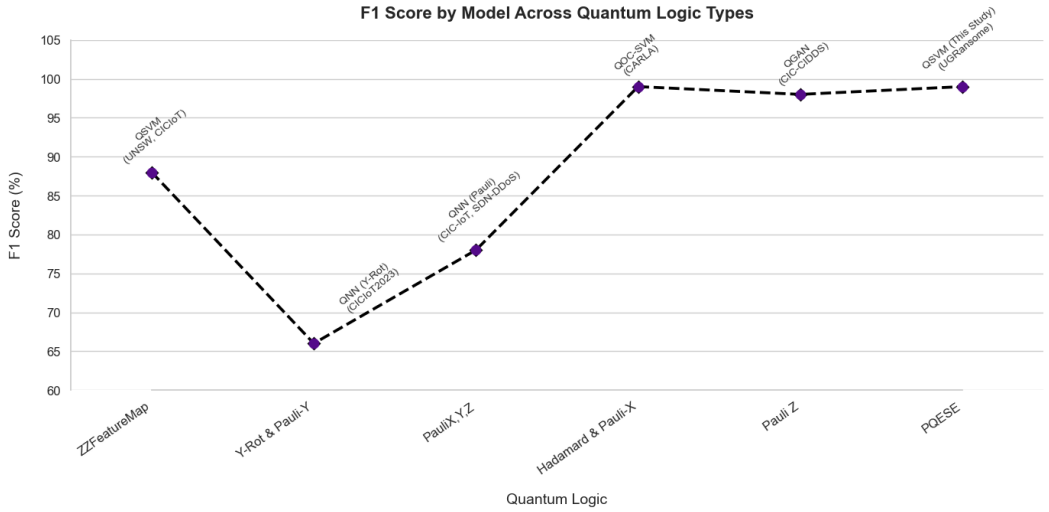


Figure 12. F1-scores by quantum logic and model design (Table 15).

Unlike prior QML applications in cybersecurity, our contribution extends beyond merely integrating quantum subroutines. We provide both theoretical and empirical justification for their role in enhancing feature representation and classification robustness.

Central to our approach is a mutual information-guided entanglement mechanism that injects domain-relevant correlations into the quantum state preparation, enabling a more expressive feature space and introducing an inductive bias that classical kernels struggle to replicate without significant computational overhead. We design a parameterised quantum circuit (PQC) that encodes 14 zero-day exploit features into multi-qubit states using a hybrid of amplitude encoding and variational entanglement layers. This encoding is informed by prior statistical analysis of feature importance, thereby improving generalisability. To evaluate quantum advantage, we compare our parameterised quantum encoding with sparse entanglement (PQESE) against state-of-the-art classical and quantum kernels. While classical models achieve high accuracy, they often overfit in low-data regimes. In contrast, our quantum model maintains competitive accuracy with reduced variance and improved decision boundaries across imbalanced classes. Achieving 99% accuracy and F1-score on the UGRansome dataset, our model outperforms existing approaches (ranging from 85 to 94%) not only in performance but in architectural innovation (Figure 12). Unlike prior work that omits implementation details or focuses on static features, our method addresses hardware constraints and generalisation challenges. Specifically, the QSVM leverages a hybrid encoding scheme combining parameterised rotation gates and controlled-Z (CZ) entanglement, using sparse topologies to retain computational efficiency. This configuration enables precise modeling of high-dimensional behaviours typical of zero-day attacks. Compared to quantum models employing simpler encodings (e.g., Hadamard or Pauli-Y mappings), our PQESE configuration yields improved generalisation under adversarial variability (Figure 12). These results demonstrate the architectural benefits of coupling expressive embeddings with entanglement-efficient connectivity [86], especially for intrusion detection where robustness is critical. Though classical models like XGBoost achieve near-perfect accuracy, our QSVM offers structural advantages beyond marginal gains. Quantum kernel estimation enables feature spaces that scale exponentially with qubit count, capturing complex correlations in network traffic that classical methods approximate less effectively (Table 15). Additionally, our QSVM uses fewer support vectors and produces simpler decision boundaries, potentially yielding computational benefits as hardware advances. While our findings are based on noiseless simulations, they lay the groundwork for real-world deployment where quantum-induced embeddings could enhance generalisation in adversarial contexts.

4.7. Implications for Critical Infrastructure Security

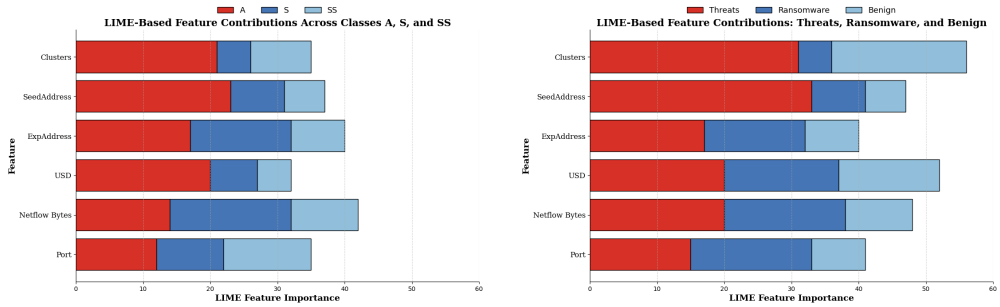
In critical infrastructure environments such as power grids, financial systems, and healthcare networks, any undetected zero-day threat can result in severe operational, financial, or safety consequences. The demonstrated sensitivity of QSVM models to novel attack patterns suggests their potential to function as a high-assurance anomaly detection component within such environments. The quantum-enhanced kernel using PQESE enables the system to better model nonlinear and complex feature interactions, which are often missed by classical models. As a result, early detection of polymorphic threats becomes more feasible, enhancing both resilience and response time in real-world deployments.

4.8. Explainability of Zero-Day Exploits Detection

Local interpretable model-agnostic explanations (LIME) is employed to quantify the contribution of each feature to the model's predictions, providing interpretable insights into the classification process (Figure 13) [11]. Figure 13 depicts the distribution of LIME feature importance values across the three threat classes. In this study, these values highlight which features most strongly influence the model's decision to classify a traffic flow as belonging to the zero-day (A) category. Each horizontal bar corresponds to the cumulative

contribution of specific features such as `port`, `netflow bytes`, and `expAddress` towards a particular class prediction. The length of the bars along the x-axis represents the magnitude of each feature's influence [11]. Key observations include the following:

- High importance scores for the `port` feature within the SS class suggest that specific port usage patterns are critical in distinguishing benign from malicious behaviours.
- Elevated LIME values for `seedAddress` and `clusters` in the A class indicate these features serve as strong indicators of zero-day exploit activity.



(**a**) (**b**)

Figure 13. Comparison of LIME-based interpretability outputs. (**a**) Explainability of selected features. (**b**) LIME feature's values.

In real-world scenarios, these feature contributions empower cybersecurity analysts to trace model predictions back to observable network attributes, providing essential transparency. This explainability is especially critical for zero-day attacks, as it helps justify automated alerts and supports forensic investigations of previously unseen threat behaviours.

4.9. Ablation Analysis

To evaluate the contribution of the t-SNE feature selection mechanism, we conducted an ablation study on QSVM performance. As presented in Table 16, removing t-SNE led to a consistent drop across all metrics (Table 17).

Table 16. Impact of removing t-SNE feature selection on QSVM performance. A checkmark (✓) indicates that t-SNE was applied, while a cross (×) indicates it was not used. Green shading highlights the decrease in performance metrics when t-SNE is removed.

Qubits	t-SNE	Accuracy	Precision	Recall	F1
3	✓	0.9794	0.9798	0.9794	0.9796
3	×	0.9102	0.9031	0.9140	0.9085
4	✓	0.9911	0.9819	0.9831	0.9825
4	×	0.9350	0.9275	0.9383	0.9328

Table 17. Wilcoxon signed-rank test comparing QSVM performance metrics with and without t-SNE feature selection. Green shading highlights the higher performance metrics when t-SNE is applied, indicating a significant improvement. The *p*-value shows the statistical significance of the observed differences.

Component	Metric	With t-SNE	Without t-SNE	*p*-Value
t-SNE	Accuracy	**0.9911**	0.8732	**0.031**
	F1	**0.9895**	0.8617	

For instance, the four-qubit QSVM saw its F1-score fall from 0.9825 (with t-SNE) to 0.9328, and the three-qubit version declined from 0.9796 to 0.9085. These reductions highlight t-SNE's role in enhancing feature separability. To assess statistical significance,

we applied the Wilcoxon signed-rank test [87] using 10-fold cross-validation. Mean values of accuracy, precision, recall, and F1-score were reported, and the resulting p-value of 0.031 (Table 17) confirms that the performance differences are statistically significant. These results validate t-SNE as a critical component of the QSVM pipeline. To mitigate overfitting and reduce computational complexity, the Chi-Square feature selection retained the six most discriminative features. This decision was supported by feature importance scores and preliminary ablation results, which showed marginal or negative returns when including more than six features. Configurations using 10 or more features led to slight overfitting. As shown in Table 16, the six-feature setup achieved an F1-score of 0.93, performing comparably or in some cases better than larger feature sets. While Chi-Square is a univariate method, it proved both interpretable and computationally efficient for our context.

4.10. Limitations and Future Work

Despite promising results, this study presents several limitations that warrant further investigation. Firstly, the QSVM relied on simulated quantum kernels due to current hardware limitations. These simulations do not fully capture the stochastic nature of real-world quantum noise, limited qubit connectivity, decoherence effects, or gate fidelity issues prevalent in NISQ. Consequently, while results on classical simulators are encouraging, real-device performance may vary substantially. Secondly, the UGRansome dataset, while offering a structured benchmark for zero-day attack detection, represents a narrow slice of network environments. Its focus on specific patterns constrains its generalisability across diverse infrastructures. Notably, advanced threats such as zero-click exploits and fileless malware remain underrepresented [88]. Zero-click attacks executed without user interaction often evade detection by leaving minimal traces, whereas fileless malware operates entirely in-memory using legitimate system tools (e.g., PowerShell), bypassing traditional endpoint monitoring [88]. The absence of granular system telemetry, volatile memory dumps, or behavioural traces in the dataset limits the QSVM's capacity to detect such threats. Although the QSVM is theoretically capable of modeling high-dimensional feature spaces, its performance is contingent on the quality and scope of the input features. Integrating system-level forensic attributes, enriched behavioural metadata, and memory features may enhance detection fidelity in future implementations.

Furthermore, while the application of the Chi-Square feature selection reduced dimensionality and computational overhead, it may overlook feature interactions essential to adversarial behaviour. Similarly, t-SNE facilitated visual separability, but its use as a feature reduction method introduces challenges in preserving global structure and interpretability. Class imbalance was addressed using SMOTE; however, its interpolation nature may fail to reflect the dynamic, temporal, and adversarial properties of real-world zero-day attacks. Although oversampling was limited to training folds during cross-validation, SMOTE may still introduce optimistic bias. Future work should explore generative augmentation techniques (e.g., variational autoencoders) that simulate realistic attack behaviours and evolution over time. Future directions include

- Investigating dynamic or adaptive quantum kernels that update based on feedback from training performance;
- Exploring more expressive encoding strategies, such as data re-uploading circuits or higher-order entanglement architectures;
- Employing hybrid quantum classical encoding (HQCE) to increase representational capacity while maintaining scalability;
- Integrating quantum explainability frameworks (e.g., QLIME) to improve interpretability and forensic traceability of quantum predictions.

To validate real-world viability, future research will involve deploying the proposed QSVMs on physical quantum processors and evaluating them under noisy conditions. In parallel, quantum error mitigation techniques tailored to the specific encoding and circuit depth will be assessed to ensure robustness. Finally, generalisation capabilities will be benchmarked against heterogeneous datasets to ensure adaptability across threat landscapes. In sum, while this study demonstrates the potential of PQESE for zero-day attack detection, it also opens numerous avenues for enhancement, robustness, and interpretability. These directions will be critical to translating QML from experimental simulations to operational cybersecurity systems.

5. Conclusions

This study assessed various ML classifiers for zero-day attack detection using a labeled network traffic dataset. Among the evaluated models, ensemble learning techniques, most specifically XGBoost, achieved exceptional performance, correctly identifying zero-day instances with no false positives and demonstrating impressive efficiency in training and prediction. In contrast, classical SVMs yielded limited performance. However, when augmented with data-balancing strategies such as borderline-SMOTE and SMOTEENN, their accuracy improved significantly. To address the remaining limitations of classical SVMs in capturing complex threat patterns, a QSVM was implemented with three-qubit and four-qubit quantum layers. The inclusion of quantum-enhanced feature mappings led to further gains in performance, achieving impressive accuracies and F1-scores. Compared to conventional models, the QSVM outperformed ensemble-based and probabilistic classifiers. Looking ahead, future work will explore deploying the QSVM on physical quantum hardware to assess its viability under real-world noise conditions. Additional research will also focus on incorporating federated and continual learning paradigms to support adaptive threat detection across decentralised systems.

Key Takeaway: Quantum-enhanced learning models, especially QSVMs, demonstrate strong potential for elevating zero-day exploit detection in terms of accuracy and interpretability, paving the way for more resilient and intelligent quantum NIDSs (QNIDSs).

List of Abbreviations: presents abbreviations used throughout this manuscript.

Author Contributions: Conceptualisation, M.N.W.N.; methodology, S.J.N. and E.N.M.; software, S.J.N. and E.N.M; validation, M.N.W.N.; formal analysis, S.J.N. and E.N.M.; investigation, S.J.N. and E.N.M.; resources, S.J.N. and E.N.M.; data curation, S.J.N. and E.N.M.; writing—original draft preparation, S.J.N. and E.N.M.; writing—review and editing, M.N.W.N.; visualisation, E.N.M. and S.J.N.; supervision, M.N.W.N.; project administration, M.N.W.N.; funding acquisition, M.N.W.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the University of Pretoria UDP Grant No A1H895.

Data Availability Statement: The datasets can be found at the following sources, accessed on 14 August 2025: Original UGRansome dataset available at the University of Pretoria research repository, https://doi.org/10.25403/UPresearchdata.25215530.v1; Kaggle UGRansome: https://www.kaggle.com/datasets/nkongolo/ugransome-dataset; this study's UGRansome: https://www.kaggle.com/datasets/jabulaninhlapo/ugransome-dataset-2024. The implementation of the QSVM with PQESE is available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	**Full Term**
ADAC	Adaptive Boosting Classifier
A	Anomaly
AI	Artificial Intelligence
AHEDNet	Adaptive Hybrid Exploit Detection Network
ANNs	Artificial Neural Networks
APTs	Advanced Persistent Threats
AVOA	African Vultures Optimisation Algorithm
AWPA	Adaptive WavePCA-Autoencoder
BTC	Bitcoins
CNN	Convolutional Neural Networks
CSV	Structured Comma-Separated Values
DARPA	Defense Advanced Research Projects Agency
DDoS	Distributed Denial of Service
DL	Deep Learning
DNN	Deep Neural Network
DoS	Denial of Service
DT	Decision Trees
EDA	Exploratory Domain Attack
ELM	Extreme Learning Machine
EM	Ensemble Model
ENN	Edited Nearest Neighbours
ETC	Extra Trees Classifier
FP	False Positives
FPR	False Positive Rate
GMCO	Genetic Mongoose-Chameleon Optimisation
GWO	Grey Wolf Optimiser
HIDS	Hybrid Intrusion Detection System
HQCE	Hybrid Quantum Classical Encoding Schemes
HQCDNN	Hybrid Quantum Classical Deep Neural Network
HQCNN	Hybrid Quantum Classical Neural Network
HQDL	Hybrid Quantum Deep Learning
IDS	Intrusion Detection Systems
IoT	Internet of Things
LDA	Latent Dirichlet Allocation
LightGBM	Light Gradient Boosting Machine
LIME	Local Interpretable Model-Agnostic Explanations
MATA	Meta-Attention Transformer Autoencoder
ML	Machine Learning
MLP	Multilayer Perceptron
MSE	Mean Squared Error
NB	Naïve Bayes
NIDS	Network Intrusion Detection System
NISQ	Noisy Intermediate-Scale Quantum
NNs	Neural Networks
PQCs	Parameterised Quantum Circuits
PQESE	Parameterised Quantum Encoding with Sparse Entanglement

QCNN	Quantum Convolutional Neural Network
QFM	Quantum Feature Map
QLSTM	Quantum Long Short-Term Memory
QML	Quantum ML
QNIDS	Quantum Network Intrusion Detection System
QSVM	Quantum Support Vector Machine
R2L	Remote to Local
RBF	Radial Basis Function
RF	Random Forest
RFC	Random Forest Classifier
ROC-AUC	Receiver Operating Characteristic Area Under the Curve
S	Signature
SMOTE	Synthetic Minority Oversampling Technique
SMOTEENN	SMOTE + Edited Nearest Neighbours
SS	Synthetic Signature
SVMs	Support Vector Machines
TF-IDF	Term Frequency–Inverse Document Frequency
TP	True Positives
TPR	True Positive Rate
UDP	User Datagram Protocol
U2R	User to Root
URL	Uniform Resource Locator
USD	United States Dollar
VQC	Variational Quantum Classifiers
XGBoost	eXtreme Gradient Boosting
t-SNE	t-Distributed Stochastic Neighbour Embedding

References

1. Adams, R.; Adeleke, F.; Anderson, D.; Bawa, A.; Branson, N.; Christoffels, A.; de Vries, J.; Etheredge, H.; Flack-Davison, E.; Gaffley, M.; et al. POPIA code of conduct for research (with corrigendum). *S. Afr. J. Sci.* **2021**, *117*, 10933. [CrossRef]
2. Mondragon, J.C.; Branco, P.; Jourdan, G.V.; Gutierrez-Rodriguez, A.E.; Biswal, R.R. Advanced IDS: A comparative study of datasets and machine learning algorithms for network flow-based intrusion detection systems. *Appl. Intell.* **2025**, *55*, 608. [CrossRef]
3. Gaber, M.; Ahmed, M.; Janicke, H. Zero day ransomware detection with Pulse: Function classification with Transformer models and assembly language. *Comput. Secur.* **2025**, *148*, 104167. [CrossRef]
4. Torky, B.; Karamitsos, I.; Najar, T. Anomaly Detection in Enterprise Payment Systems: An Ensemble Machine Learning Approach. In *Proceedings of the Business Analytics and Decision Making in Practice*; Emrouznejad, A., Zervopoulos, P.D., Ozturk, I., Jamali, D., Rice, J., Eds.; Springer International Publishing: Cham, Switzerland , 2024; pp. 11–23. [CrossRef]
5. Zahoora, U.; Rajarajan, M.; Pan, Z.; Khan, A. Zero-day Ransomware Attack Detection using Deep Contractive Autoencoder and Voting based Ensemble Classifier. *Appl. Intell.* **2022**, *52*, 13941–13960. . [CrossRef]
6. Rios-Ochoa, E.; Pérez-Díaz, J.A.; Garcia-Ceja, E.; Rodriguez-Hernandez, G. Ransomware Family Attribution with ML: A Comprehensive Evaluation of Datasets Quality, Models Comparison and a Simulated Deployment. *IEEE Access* **2025**, *13*, 108108–108126. [CrossRef]
7. Kumar, K.; Khari, M. Federated Active Meta-Learning with Blockchain for Zero-Day Attack Detection in Industrial IoT. *Peer -Peer Netw. Appl.* **2025**, *18*, 199. [CrossRef]
8. Anjum, A.; Subramanian, P.R.; Stalinbabu, R.; Kothapeta, D.; Sheela, K.S.; Jegajothi, B. Detecting Zero-Day Attacks using Advanced Anomaly Detection in Network Traffic. In Proceedings of the 2025 5th International Conference on Pervasive Computing and Social Networking (ICPCSN), IEEE, Salem, India, 14–16 May 2025; pp. 1247–1253.
9. Kasongo, S.M. A deep learning technique for intrusion detection system using a Recurrent Neural Networks based framework. *Comput. Commun.* **2023**, *199*, 113–125. [CrossRef]
10. Mohamed, A.A.; Al-Saleh, A.; Sharma, S.K.; Tejani, G.G. Zero-day exploits detection with adaptive WavePCA-Autoencoder (AWPA) adaptive hybrid exploit detection network (AHEDNet). *Sci. Rep.* **2025**, *15*, 4036. [CrossRef]
11. Su, R. Generative Mathematical Models for Ransomware Attack Prediction Using Chi-Square Feature Selection for Enhanced Accuracy. *Signal Image Video Process.* **2025**, *19*, 789. [CrossRef]

12. Rojas, R.V.B. Artificial Intelligence: Genesis, Development, and Future. In *Revolutionizing Communication*; CRC Press: Boca Raton, FL, USA , 2024; pp. 1–15.
13. Yan, P.; Khoei, T.T.; Hyder, R.S.; Hyder, R.S. A Dual-Stage Ensemble Approach to Detect and Classify Ransomware Attacks. In Proceedings of the 2024 IEEE 15th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), Yorktown Heights, NY, USA, 17–19 October 2024; pp. 781–786. [CrossRef]
14. Kumar, R.; Swarnkar, M. QuIDS: A Quantum Support Vector machine-based Intrusion Detection System for IoT networks. *J. Netw. Comput. Appl.* **2025**, *234*, 104072. [CrossRef]
15. Durgut, S.; Küçüksille, E.U.; Tokmak, M. Hybrid Quantum–Classical Deep Neural Networks Based Smart Contract Vulnerability Detection. *Appl. Sci.* **2025**, *15*, 4037. [CrossRef]
16. Abreu, D.; Moura, D.; Esteve Rothenberg, C.; Abelém, A. Quantumnetsec: Quantum machine learning for network security. *Int. J. Netw. Manag.* **2025**, *35*, e70018. [CrossRef]
17. Caivano, D.; De Vincentiis, M.; Pal, A.; Ragone, A. Securing Smart Cities: Unraveling Quantum as a Service. In *QP4SE 2023, Proceedings of the 2nd International Workshop on Quantum Programming for Software Engineering, San Francisco, CA, USA, 4 December 2023*; Association for Computing Machinery: New York, NY, USA, 2023; pp. 1–6. [CrossRef]
18. Srivastava, D.; Singh, R.; Chakraborty, C.; Maakar, S.K.; Makkar, A.; Sinwar, D. A framework for detection of cyber attacks by the classification of intrusion detection datasets. *Microprocess. Microsystems* **2024**, *105*, 104964. [CrossRef]
19. Rookard, C.; Khojandi, A. Unsupervised Machine Learning for Cybersecurity Anomaly Detection in Traditional and Software-Defined Networking Environments. *IEEE Trans. Netw. Serv. Manag.* **2025**, *22*, 1129–1144. [CrossRef]
20. Zhang, Z.; Turnbull, B.; Kermanshahi, S.K.; Pota, H.; Hu, J. UNSW-MG24: A Heterogeneous Dataset for Cybersecurity Analysis in Realistic Microgrid Systems. *IEEE Open J. Comput. Soc.* **2025**, *6*, 543–553. [CrossRef]
21. Celik, Y.; Basaran, E.; Goel, S. Deep Learning Methods for Intrusion Detection Systems on the CSE-CIC-IDS2018 Dataset: A Review. In *Proceedings of the Digital Forensics and Cyber Crime*; Goel, S., Uzun, E., Xie, M., Sarkar, S., Eds.; Springer International Publishing: Cham, Switzerland, 2025; pp. 38–65. [CrossRef]
22. Nkongolo, M.; Tokmak, M. Zero-Day Threats Detection for Critical Infrastructures. In *Proceedings of the South African Institute of Computer Scientists and Information Technologists*; Communications in Computer and Information Science, Gerber, A., Coetzee, M., Eds.; Springer International Publishing: Cham, Switzerland, 2023; Volume 1878, pp. 32–47. [CrossRef]
23. Shankar, D.; George, G.V.S.; S, J.N.J.N.S.; Madhuri, P.S. Deep Analysis of Risks and Recent Trends Towards Network Intrusion Detection System. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 0140129. [CrossRef]
24. Sokhonn, L.; Park, Y.S.; Lee, M.K. Hierarchical Clustering via Single and Complete Linkage Using Fully Homomorphic Encryption. *Sensors* **2024**, *24*, 4826. [CrossRef]
25. Chaudhary, I.; Adhikari, S. Ransomware Detection Using Machine Learning Techniques. *Res. CAB A J. Res. Dev.* **2024**, *3*, 96–114. [CrossRef]
26. Kadi, A.; Selamnia, A.; Abou El Houda, Z.; Moudoud, H.; Brik, B.; Khoukhi, L. An In-Depth Comparative Study of Quantum-Classical Encoding Methods for Network Intrusion Detection. *IEEE Open J. Commun. Soc.* **2025**, *6*, 1129–1148. [CrossRef]
27. Pei, J.J.; Gong, L.H.; Qin, L.G.; Zhou, N.R. One-to-many image generation model based on parameterized quantum circuits. *Digit. Signal Process.* **2025**, *165*, 105340. [CrossRef]
28. Ahmed, S.; Shihab, I.F.; Khokhar, A. Quantum-driven zero trust architecture with dynamic anomaly detection in 7G technology: A neural network approach. *Meas. Digit.* **2025**, *2–3*, 100005. [CrossRef]
29. Watkins, W.M.; Chen, S.Y.C.; Yoo, S. Quantum machine learning with differential privacy. *Sci. Rep.* **2023**, *13*, 2453. [CrossRef]
30. Mohammadisavadkoohi, E.; Shafiabady, N.; Vakilian, J. A Systematic Review on Quantum Machine Learning Applications in Classification. *IEEE Trans. Artif. Intell.* **2025**, 1–16, Early Access. [CrossRef]
31. Duffy, C.; Hassanshahi, M.; Jastrzebski, M.; Malik, S. Unsupervised beyond-standard-model event discovery at the LHC with a novel quantum autoencoder. *Quantum Mach. Intell.* **2025**, *7*, 1–19. [CrossRef]
32. Akter, M.S.; Shahriar, H.; Cuzzocrea, A.; Wu, F. Quantum Adversarial Attacks: Developing Quantum FGSM Algorithm. In Proceedings of the 2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC), Osaka, Japan, 2–4 July 2024; pp. 1073–1079. [CrossRef]
33. Umeano, C.; Paine, A.E.; Elfving, V.E.; Kyriienko, O. What can we learn from quantum convolutional neural networks? *Adv. Quantum Technol.* **2025**, *8*, 2400325. [CrossRef]
34. Gujju, Y.; Matsuo, A.; Raymond, R. Quantum machine learning on near-term quantum devices: Current state of supervised and unsupervised techniques for real-world applications. *Phys. Rev. Appl.* **2024**, *21*, 067001. [CrossRef]
35. Elsedimy, E.I.; Elhadidy, H.; Abohashish, S.M.M. A Novel Intrusion Detection System Based on a Hybrid Quantum Support Vector Machine and Improved Grey Wolf Optimizer. *Clust. Comput.* **2024**, *27*, 9917–9935. [CrossRef]
36. Eze, L.; Chaudhry, U.B.; Jahankhani, H. Quantum-Enhanced Machine Learning for Cybersecurity: Evaluating Malicious URL Detection. *Electronics* **2025**, *14*, 1827. [CrossRef]

37. Vijayalakshmi, M.; Shalinie, S.M.; Bharathi, J.V. A Comparative Analysis of Kernel Methods in Quantum Support Vector Machines for Network Anomaly Detection. In Proceedings of the 2025 Fourth International Conference on Power, Control and Computing Technologies (ICPC2T), IEEE, Raipur, India, 20–22 January 2025; pp. 1–6. [CrossRef]
38. Mahdian, M.; Mousavi, Z. Entanglement detection with quantum support vector machine (QSVM) on near-term quantum devices. *Sci. Rep.* **2025**, *15*, 1–15. [CrossRef]
39. Tripathi, S.; Upadhyay, H.; Soni, J. A Quantum LSTM-based approach to cyber threat detection in virtual environment. *J. Supercomput.* **2024**, *81*, 142. [CrossRef]
40. Saeed, M.M. An AI-Driven Cybersecurity Framework for IoT: Integrating LSTM-Based Anomaly Detection, Reinforcement Learning, and Post-Quantum Encryption. *IEEE Access* **2025**, *13*, 104027–104036. [CrossRef]
41. Sridevi, S.; Indira, B.; Geetha, S.; Balachandran, S.; Kar, G.; Kharbanda, S. Unified hybrid quantum classical neural network framework for detecting distributed denial of service and Android mobile malware attacks. *EPJ Quantum Technol.* **2025**, *12*, 77. [CrossRef]
42. Bellante, A.; Fioravanti, T.; Carminati, M.; Zanero, S.; Luongo, A. Evaluating the potential of quantum machine learning in cybersecurity: A case-study on PCA-based intrusion detection systems. *Comput. Secur.* **2025**, *154*, 104341. [CrossRef]
43. Jemili, F.; Meddeb, R.; Korbaa, O. Intrusion detection based on ensemble learning for big data classification. *Clust. Comput.* **2024**, *27*, 3771–3798. [CrossRef]
44. Jovanović, D.D.; Vuletić, P.V. Machine learning pipelines for IoT botnet detection and behavior characterization in heavily imbalanced settings. *Signal Image Video Process.* **2025**, *19*, 254. [CrossRef]
45. Shana, T.B.; Kumari, N.; Agarwal, M.; Mondal, S.; Rathnayake, U. Anomaly-based intrusion detection system based on SMOTE-IPF, Whale Optimization Algorithm, and ensemble learning. *Intell. Syst. Appl.* **2025**, *27*, 200543. [CrossRef]
46. Rajathi, C.; Rukmani, P. AccFIT-IDS: Accuracy-based feature inclusion technique for intrusion detection system. *Syst. Sci. Control Eng.* **2025**, *13*, 2460429. [CrossRef]
47. Neha.; Kajal, A. A Comprehensive Literature Review: Feature Engineering Techniques in Enhancing Machine and Deep Learning Models Optimization for Intrusion Detection Systems. In Proceedings of the 5th International Conference on Artificial Intelligence and Smart Energy, Coimbatore, India, 30–31 January 2025; Manoharan, S., Tugui, A., Perikos, I., Eds.; Springer International Publishing: Cham, Switzerland, 2025; pp. 436–451. [CrossRef]
48. Iftikhar, N.; Rehman, M.U.; Shah, M.A.; Alenazi, M.J.; Ali, J. Intrusion Detection in NSL-KDD Dataset Using Hybrid Self-Organizing Map Model. *CMES-Comput. Model. Eng. Sci.* **2025**, *143*, 639–671. [CrossRef]
49. Sattar, S.; Khan, S.; Khan, M.I.; Akhmediyarova, A.; Mamyrbayev, O.; Kassymova, D.; Oralbekova, D.; Alimkulova, J. Anomaly detection in encrypted network traffic using self-supervised learning. *Sci. Rep.* **2025**, *15*, 26585. [CrossRef]
50. Naeem, H.; Ullah, F.; Srivastava, G. Classification of intrusion cyber-attacks in smart power grids using deep ensemble learning with metaheuristic-based optimization. *Expert Syst.* **2025**, *42*, e13556. [CrossRef]
51. Chen, S.; Zheng, W. RRMSE-enhanced weighted voting regressor for improved ensemble regression. *PLoS ONE* **2025**, *20*, e0319515. [CrossRef]
52. Kostas, K.; Just, M.; Lones, M.A. Individual Packet Features are a Risk to Model Generalization in ML-Based Intrusion Detection. *IEEE Netw. Lett.* **2025**, *7*, 66–70. [CrossRef]
53. Abdulsalam, G.; Ahmad, I. Comparative investigation of quantum and classical kernel functions applied in support vector machine algorithms. *Quantum Inf. Process.* **2025**, *24*, 109. [CrossRef]
54. Alvarez-Estevez, D. Benchmarking Quantum Machine Learning Kernel Training for Classification Tasks. *IEEE Trans. Quantum Eng.* **2025**, *6*, 1–15. [CrossRef]
55. Li, J.; Li, Y.; Song, J.; Zhang, J.; Zhang, S. Quantum Support Vector Machine for Classifying Noisy Data. *IEEE Trans. Comput.* **2024**, *73*, 2233–2247. [CrossRef]
56. Mantha, P.; Kiwit, F.J.; Saurabh, N.; Jha, S.; Luckow, A. Pilot-Quantum: A Middleware for Quantum-HPC Resource, Workload and Task Management. In Proceedings of the 2025 IEEE 25th International Symposium on Cluster, Cloud and Internet Computing (CCGrid), Tromsø, Norway, 19–22 May 2025; pp. 01–10. [CrossRef]
57. Zhang, Y.; Li, B.Z. The Graph Fractional Fourier Transform in Hilbert Space. *IEEE Trans. Signal Inf. Process. Over Netw.* **2025**, *11*, 242–257. [CrossRef]
58. Oukaira, A. Quantum Hardware Devices (QHDs): Opportunities and Challenges. *IEEE Access* **2025**, *13*, 98229–98241. [CrossRef]
59. Yogaraj, K.; Quanz, B.; Vikas, T.; Mondal, A.; Mondal, S. Post-variational classical quantum transfer learning for binary classification. *Sci. Rep.* **2025**, *15*, 23682. [CrossRef] [PubMed]
60. Rishiwal, V.; Agarwal, U.; Yadav, M.; Tanwar, S.; Garg, D.; Guizani, M. A New Alliance of Machine Learning and Quantum Computing: Concepts, Attacks, and Challenges in IoT Networks. *IEEE Internet Things J.* **2025**, *12*, 18865–18886. [CrossRef]
61. Qayyum, T.; Khan, M.W.H.; Tariq, A.; Serhani, M.A.; Sallabi, F.M.; Trabelsi, Z.; Taleb, I. Quantum Federated Learning: Bridging Quantum Computing and Distributed AI. In Proceedings of the 2024 IEEE/ACM 17th International Conference on Utility and Cloud Computing (UCC), Sharjah, United Arab Emirates, 16–19 December 2024; pp. 327–335. [CrossRef]

62. Otsuka, Y.; Seki, K.; Yunoki, S. Quantum conditional mutual information as a probe of measurement-induced entanglement phase transitions. *Phys. Rev. B* **2025**, *112*, 054301. [CrossRef]
63. Whitlow, L. A Comprehensive Survey of Quantum Computing: Principles, Progress, and Prospects for Classical-Quantum Integration. *J. Comput. Sci. Softw. Appl.* **2025**, *5*, 1–17.
64. Alqahtany, S.S.; Shaikh, A.; Alqazzaz, A. Enhanced Grey Wolf Optimization (EGWO) and random forest based mechanism for intrusion detection in IoT networks. *Sci. Rep.* **2025**, *15*, 1916. [CrossRef]
65. Omer Albasheer, F.; Ramesh Haibatti, R.; Agarwal, M.; Yeob Nam, S. A Novel IDS Based on Jaya Optimizer and Smote-ENN for Cyberattacks Detection. *IEEE Access* **2024**, *12*, 101506–101527. [CrossRef]
66. Le, T.T.H.; Shin, Y.; Kim, M.; Kim, H. Towards unbalanced multiclass intrusion detection with hybrid sampling methods and ensemble classification. *Appl. Soft Comput.* **2024**, *157*, 111517. [CrossRef]
67. Kiranmayee, B.; Devi, M.S.; Susheela, K.; Dhumpati, R.; Penubaka, K.K.R.; Naidu, U.G. Developing a Robust Intrusion Detection System Using SMOTE and Hybrid SVNN Model. In Proceedings of the 2025 4th International Conference on Sentiment Analysis and Deep Learning (ICSADL), IEEE, Halifax, NS, Canada, 18–20 February 2025; pp. 369–376.
68. Terven, J.; Cordova-Esparza, D.M.; Romero-González, J.A.; Ramírez-Pedraza, A.; Chávez-Urbiola, E. A comprehensive survey of loss functions and metrics in deep learning. *Artif. Intell. Rev.* **2025**, *58*, 195. [CrossRef]
69. Mohanty, N.; Behera, B.K.; Ferrie, C.; Dash, P. A quantum approach to synthetic minority oversampling technique (SMOTE). *Quantum Mach. Intell.* **2025**, *7*, 38. [CrossRef]
70. Majid, B.; Sofi, S.A.; Jabeen, Z. Quantum machine learning: A systematic categorization based on learning paradigms, NISQ suitability, and fault tolerance. *Quantum Mach. Intell.* **2025**, *7*, 1–55. [CrossRef]
71. Shahid, M.; Hassan, M.A.; Iqbal, F.; Altaf, A.; Shah, S.W.H.; Elizaincin, A.V.; Ashraf, I. Enhancing movie recommendations using quantum support vector machine (QSVM). *J. Supercomput.* **2025**, *81*, 78. [CrossRef]
72. LeCompte, T.; Qi, F.; Yuan, X.; Tzeng, N.F.; Najafi, M.H.; Peng, L. Machine-Learning-Based Qubit Allocation for Error Reduction in Quantum Circuits. *IEEE Trans. Quantum Eng.* **2023**, *4*, 1–14. [CrossRef]
73. Gayathri Devi, S.; Manjula Gandhi, S.; Chandia, S.; Boobalaragavan, P. Exploring IBM Quantum Experience. In *Quantum Computing: A Shift from Bits to Qubits*; Studies in Computational Intelligence; Pandey, R., Srivastava, N., Singh, N.K., Tyagi, K., Eds. Springer: Singapore, 2023; Volume 1085, pp. 215–231. [CrossRef]
74. Sihare, S.R. Dimensionality Reduction for Data Analysis With Quantum Feature Learning. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2025**, *15*, e1568. [CrossRef]
75. Akbar, M.A.; Rafi, S.; Khan, A.; Ye, B. Quantum Secure DevOps (QSecOps): Integrating Quantum-Based Security Checks into CI/CD Pipelines: Next-Generation Software Security. In Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering, New York, NY, USA, 23–28 June 2025; FSE Companion '25, pp. 1732–1741. [CrossRef]
76. Wang, H. Several fitness functions and entanglement gates in quantum kernel generation. *Quantum Mach. Intell.* **2025**, *7*, 7. [CrossRef]
77. Sasse, L.; Nicolaisen-Sobesky, E.; Dukart, J.; Eickhoff, S.; Götz, M.; Hamdan, S.; Komeyer, V.; Kulkarni, A.; Lahnakoski, J.; Love, B.C.; et al. Overview of leakage scenarios in supervised machine learning. *J. Big Data* **2025**, *12*, 135. [CrossRef]
78. Rayudu, K.; Vasanthi, A.; Tharani, B. Enhancing Ransomware Detection in Cybersecurity: A Comprehensive Ensemble Approach. *J. Electrical Systems.* **2024**, *20*, 5222–5232. [CrossRef]
79. Alhashmi, A.A.; Darem, A.A.; Alshammari, A.B.; Darem, L.A.; Sheatah, H.K.; Effghi, R. Ransomware early detection techniques. *Eng. Technol. Appl. Sci. Res.* **2024**, *14*, 14497–14503. [CrossRef]
80. Tokmak, M. Deep Forest Approach for Zero-Day Attacks Detection. In *Proceedings of the Innovations and Technologies in Engineering*; Tasdemir, S., Ozkan, I.A., Eds.; Eğitim Yayinevi: Istanbul, Turkey, 2022; pp. 45–56.
81. Abreu, D.; Rothenberg, C.E.; Abelém, A. QML-IDS: Quantum Machine Learning Intrusion Detection System. In Proceedings of the 2024 IEEE Symposium on Computers and Communications (ISCC), Paris, France, 26–29 June 2024; pp. 1–6. [CrossRef]
82. Rajkumar, K.; Shalinie, S.M. SHAP-based intrusion detection in IoT networks using quantum neural networks on IonQ hardware. *J. Parallel Distrib. Comput.* **2025**, *204*, 105133. [CrossRef]
83. Meghana, R.; Ramesha, S.S.; Mukhopadhyay, A. QCAE-QOC-SVM: A hybrid quantum machine learning model for DoS and Fuzzy attack detection on autonomous vehicle CAN bus. *MethodsX* **2025**, *15*, 103471. [CrossRef]
84. Hammami, W.; Cherkaoui, S.; Wang, S. Enhancing Network Anomaly Detection with Quantum GANs and Successive Data Injection for Multivariate Time Series. In Proceedings of the 2025 International Wireless Communications and Mobile Computing (IWCMC), Abu Dhabi, United Arab Emirates, 12–16 May 2025; pp. 1667–1672. [CrossRef]
85. Atban, F.; Küçükkara, M.Y.; Bayılmış, C. Enhancing variational quantum classifier performance with meta-heuristic feature selection for credit card fraud detection. *Eur. Phys. J. Spec. Top.* **2025**, in press. [CrossRef]
86. Sikiru, I.A.; Kora, A.D.; Ezin, E.C.; Imoize, A.L.; Li, C.T. Hybridization of learning techniques and quantum mechanism for IIoT security: Applications, challenges, and prospects. *Electronics* **2024**, *13*, 4153. [CrossRef]

87. Hewa, T.; Siriwardhana, Y.; Ylianttila, M. Leveraging Explainable AI for Adaptive Adversarial DoS Attack Detection in 6G IoT Networks. In Proceedings of the 2025 IEEE Wireless Communications and Networking Conference (WCNC), Milan, Italy, 24–27 March 2025; pp. 1–7. [CrossRef]
88. Razavi, H.; Jamali, M.R.; Emsaki, M.; Ahmadi, A.; Hajiaghei-Keshteli, M. Quantifying the Financial Impact of Cyber Security Attacks on Banks: A Big Data Analytics Approach. In Proceedings of the 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, Saskatchewan, 24–27 September 2023; pp. 533–538. [CrossRef]