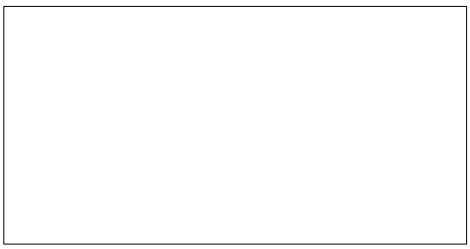


Computer Science Review 60 (2026) 100882

Contents lists available at ScienceDirect

Computer Science Review

journal homepage: www.elsevier.com/locate/cosrev

Shaping the future of cybersecurity: The convergence of AI, quantum computing, and ethical frameworks for a secure digital era

Menahil Khawar[a], Sohail Khalid[b], Mujeeb Ur Rehman[c,*], Aminu Usman[d], Wajdan Al Malwi[e], Fatima Asiri[e]

[a] Electrical and Computer Engineering Department, Riphah International University, Islamabad, Pakistan
[b] Department of Computer Science, University of Management and Technology, Lahore, Pakistan
[c] School of Computer Science and Informatics, De Montfort University, Leicester, LE1 9BH, UK
[d] Department of Computer Science, York St John University, York, UK
[e] College of Computer Science, Informatics and Computer Systems Department, King Khalid University, Abha, Saudi Arabia

ARTICLE INFO

Keywords:
Cybersecurity
Artificial Intelligence
Quantum Computing
Quantum Cybersecurity
Ethical AI

ABSTRACT

The increasing sophistication and frequency of cyber threats have rendered conventional protection strategies inadequate. Artificial Intelligence (AI) is becoming central to modern cybersecurity, strengthening capabilities in vulnerability assessment, malware detection, phishing prevention, intrusion detection, and deception technologies. Simultaneously, quantum computing introduces both challenges to classical cryptography and opportunities for new forms of quantum-enhanced defenses. This review integrates advances in AI, quantum methods, and ethical governance to provide an integrated perspective on the future of secure digital systems. It evaluates state-of-the-art AI models, including explainable frameworks and quantum-inspired approaches, such as Quantum Convolutional Neural Networks and Quantum Support Vector Machines, along with recent progress in post-quantum cryptography. Ethical concerns, particularly bias, transparency, privacy, and accountability, are examined as essential foundations for trustworthy cybersecurity design in system-on-chip and embedded AI environments. In addition to technical developments, this study considers regulatory frameworks, governance structures, and societal expectations, highlighting the need for responsible and adaptive approaches. A comparative SWOT analysis outlines the strengths, limitations, and areas for cross-domain integration. Finally, a roadmap of future research directions is presented, aligning AI-driven defenses, quantum resilience, and ethical safeguards into flexible and reliable cybersecurity architectures. By linking the technological, ethical, and policy dimensions, this review offers a consolidated foundation to guide the evolution of cybersecurity in a globally connected era.

1. Introduction

As digital technologies become deeply embedded in every aspect of modern life, the rapid increase in the frequency and complexity of cyber threats presents a growing challenge to global security and economic resilience. Cyber threats surged by 72 % between 2021 and 2023, with global cybercrime costs expected to reach an unprecedented $10.5 trillion annually by 2025, surpassing the GDP of many nations and necessitating massive investments in cybersecurity [1,2]. In 2023 alone, over 343 million individuals were affected by over 2365 major attacks targeting critical sectors such as healthcare, finance, and infrastructure [3–5]. The average global cost of a data breach climbed to $4.45 million and $4.88 million in 2023 and 2024, respectively, with the United States bearing the highest costs at $5.09 million. Ransomware, phishing, and email-based malware responsible for 35 % of breaches continued to impact 94 % of organizations worldwide. Meanwhile, cyber insurance premiums in the U.S. spiked by 50 % in 2022, reflecting the mounting financial strain on businesses [6,7]. With cybercrime growing at 15 % annually and incidents projected to exceed 2.85 billion by 2025, the need for AI-driven and proactive cybersecurity strategies has become critical [8,9]. Fig. 1 illustrates the progression of global cybercrime costs and annual growth rates. These data highlight the increasing sophistication and frequency of cyber threats, emphasizing the urgent need for advanced cybersecurity measures. Cybersecurity, a fundamental pillar of digital resilience, safeguards systems, networks, and data from digital threats,

* Corresponding author.
Email address: mujeeb.rehman@dmu.ac.uk (M.U. Rehman).

https://doi.org/10.1016/j.cosrev.2025.100882
Received 27 February 2025; Received in revised form 26 November 2025; Accepted 10 December 2025
Available online 24 December 2025

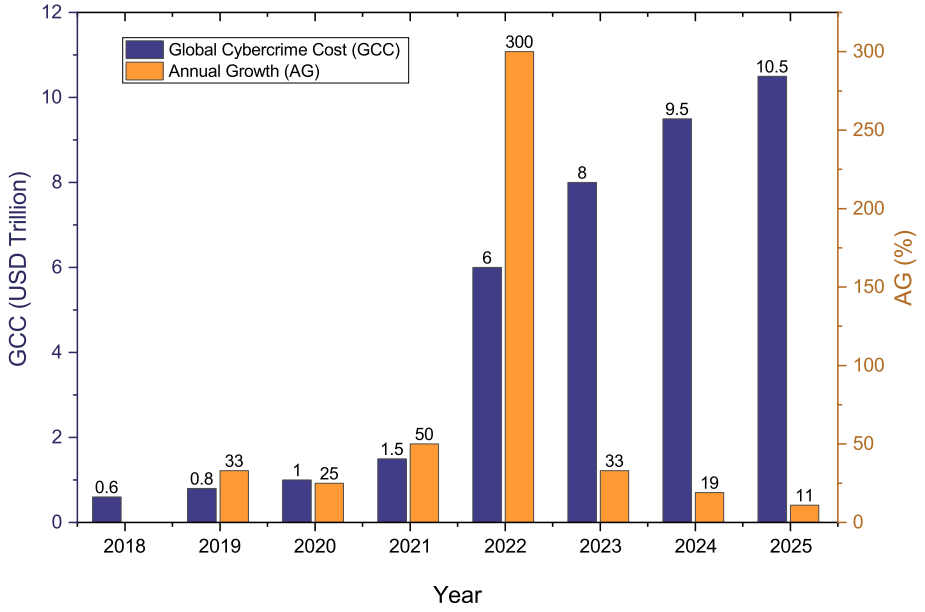


Fig. 1. Trends in Global Cybercrime Costs, Annual Growth, and Reported Cyber Threat Incidents (2018–2025) [1]-[3].

unauthorized access, and damage. This ensures the integrity, confidentiality, and availability of sensitive information while combating increasingly sophisticated cyber threats [10]. The COVID-19 pandemic accelerated this demand, driving rapid digitization across remote work, e-commerce, and virtual collaboration, and exposing new vulnerabilities for malicious actors to exploit [11,12]. Beyond traditional systems, cybersecurity now encompasses emerging trends such as Industry 5.0, Web 3.0, blockchain networks, and the Metaverse, which open new avenues for innovation while introducing unique risks. Technologies such as digital twins, which are virtual replicas of physical systems that facilitate real-time simulations of cyber threats, and zero-trust architecture, which operates on the principle of never trust and is always verified, have emerged as essential tools for mitigating cybersecurity risks [13,14]. The rise in digital currencies has further underscored the need for tighter cybersecurity integration to secure sensitive transactions and prevent fraud [15]. AI has revolutionized cybersecurity by enabling systems to detect and respond to threats with unmatched speed and precision. Machine Learning (ML) analyzes vast datasets in real-time to identify anomalies and breaches, while Deep Learning (DL) uncovers vulnerabilities in complex data such as malware behaviors and network traffic [16–18]. Natural Language Processing (NLP) strengthens defenses by identifying phishing attempts and social engineering attacks on emails, social media, and the dark web. Emerging quantum AI (QAI) models, such as quantum support vector machines (QSVM), enhance threat detection, vulnerability assessment, and incident response, offering unparalleled precision in addressing challenges like polymorphic malware and zero-day vulnerabilities [19,20]. Explainable AI (XAI) and Understandable AI (UAI) address a key challenge in cybersecurity by making AI-driven decisions interpretable, ensuring trust in automated systems, and maintaining adaptability to evolving threats [21]. The integration of cloud computing and AI has transformed cybersecurity by offering scalable real-time threat monitoring and response capabilities, which are critical for securing dynamic environments [22]. Together, these technologies enable cybersecurity frameworks to address challenges such as detecting unknown malware, mitigating zero-day vulnerabilities, and enhancing real-time incident responses. These advancements collectively form a robust multilayered defense framework capable of addressing the complexities of modern cyber threats. Quantum computing (QC) holds promise for mitigating challenges such as cracking advanced encryption and identifying sophisticated threats. This study explores how AI and QC can enhance threat detection, automate vulnerability assessments, and improve incident response. By utilizing these technologies, cybersecurity systems can maintain resilience, adaptability, and preparedness to address the continuously evolving threat landscapes. Table 1 shows the research questions addressed in this study.

Table 1
Research questions addressed in this study.

No.	Research question
1	How is AI, including XAI and UAI, transforming threat detection, prevention, and incident response in cybersecurity?
2	What roles do cloud computing, quantum computing, and digital twins play in enhancing cybersecurity frameworks?
3	What are the key challenges, limitations, and ethical issues in deploying AI-driven cybersecurity solutions?
4	How do AI-powered frameworks create adaptive defenses against malware, zero-day vulnerabilities, and automate incident responses?

A more detailed analysis of prior work and how this review differs from existing efforts is presented in Section 2.

2. Related work and theoretical foundations

This section presents a structured synthesis of the foundational research underpinning the convergence of AI, QC, and ethical frameworks in the field of cybersecurity. While many prior studies have examined each domain in isolation, few have addressed their integration in a way that reflects both technical feasibility and socio-ethical implications. Our review addresses this gap by contextualizing the existing literature and aligning it with future-oriented cybersecurity design principles.

2.1. AI in cybersecurity

The application of AI, particularly ML and DL, has been extensively studied in cybersecurity contexts, such as malware detection, phishing prevention, and intrusion detection systems. For instance, CNNs and SVMs have achieved high accuracy in classifying malware and malicious URLs, often exceeding 90 % on benchmark datasets [23,24]. Ensemble models and hybrid techniques, including autoencoders and GANs, have been applied to detect zero-day attacks and sophisticated anomalies in network traffic [25,26]. However, as discussed by [27], most prior

studies focused primarily on detection accuracy, often at the expense of explainability and adaptability. Moreover, current reviews typically overlook emerging integrations with quantum machine learning (QML) and regulatory compliance. A unified framework that embeds AI solutions within scalable, explainable, and policy-compliant cybersecurity architectures is required.

2.2. Quantum computing and post quantum security

Quantum computing introduces both disruptive threats and novel defensive capabilities in cybersecurity. Studies such as [28,29] emphasize that traditional encryption schemes, such as RSA and ECC, are vulnerable to quantum algorithms such as Shor's and Grover's. In August 2024, NIST finalized the first post-quantum cryptography standards ML-KEM for key encapsulation (FIPS 203), ML-DSA for lattice-based digital signatures (FIPS 204), and SLH-DSA for stateless hash-based signatures (FIPS 205) providing concrete migration targets for quantum-resilient architectures [30–33]. Recent surveys consolidate algorithmic choices and deployment trade-offs across sectors, offering guidance on implementation constraints and transition planning [34]. In parallel, QML models like QSVMs, VQCs, and QCNNs have emerged as promising tools for tasks like anomaly detection and classification under uncertainty [35]. Despite this potential, much of the existing literature is either theoretical or fragmented. Many surveys on PQC and QML, such as that by Dam et al. [29], provide taxonomies or performance benchmarks but do not explore how quantum resilience can be coupled with explainability, governance, or AI-based threat mitigation pipelines.

2.3. Ethical AI and governance in security systems

With the growing adoption of AI in safety-critical environments, ethical concerns such as bias, opacity, data misuse, and a lack of accountability have gained prominence. Frameworks for XAI and UAI have been proposed to make opaque models interpretable. Methods such as SHAP, LIME, and attention visualization have been embedded into IDS and malware classifiers to enhance trust and transparency [36,37]. However, as Mittelstadt et al. [38] argue, many AI ethics frameworks are principle-based and lack actionable pathways for their integration into technical systems. Similarly, Ienca et al. [39] cautioned that ethics in digital systems must be operationalized through institutional, legal, and human rights frameworks. Recent work by Marchang et al. [40] presented a secure-by-design real-time IoMT architecture for e-health. It emphasizes encryption, key management, and reliable communication between wearable devices and central servers. This shows how ethical and privacy considerations can be built directly into system design, which is especially important when technology affects health and personal information. Without such integrated approaches, safeguards risk remaining fragmented rather than forming a consistent foundation for cybersecurity.

2.4. Gap and contribution of this review

To the best of our knowledge, no prior review has holistically integrated the following dimensions.

- Classical and quantum AI models for proactive cybersecurity,
- Explainable, understandable, and ethically governed AI techniques,
- Regulatory frameworks such as GDPR, NIST, and ISO/IEC 27001,
- Design considerations for scalability, real-time operation, and trustworthiness.

This review fills this gap by offering a transdisciplinary synthesis that bridges algorithmic capability, quantum readiness, and ethical governance. Our analysis provides a foundation for the next generation of cybersecurity systems that are not only technically resilient but also trustworthy and aligned with human values.

Table 2
Search queries used in various databases for the study.

Database	Search query
Web of Science	(((((TS = (Quantum Computing)) OR TS = (Quantum AI))) AND TS = (Quantum Cryptography)) OR TS = (Cybersecurity)) OR TS = (Post quantum Security)
Scopus	TITLE-ABS-KEY ((“Quantum AI” OR “Quantum Computing” OR “Quantum Cryptography”) AND (“Cybersecurity” OR “Intrusion Detection” OR “Post quantum Security”)) AND (LIMIT-TO (DOCTYPE, “ar”) OR LIMIT-TO (DOCTYPE, “cp”))
Google Scholar	(in title:“Quantum AI” OR in title:“Quantum Computing” OR “Quantum Cryptography”) AND (in title:“Cybersecurity” OR “Intrusion Detection” OR in title:“Post quantum Security”)

3. Research methodology

This section outlines the comprehensive framework of this review, which investigates the transformative role of AI and its quantum advancements in cybersecurity. The methodology is based on the following five subsections to ensure rigor, reproducibility, and alignment with the review objectives.

3.1. Defining the scope of the review

The objective of this review is to examine the evolution of AI techniques from classical approaches to quantum methodologies to address complex cybersecurity challenges. This emphasizes their application in real-time threat detection, adaptive systems, and robust defense strategies. The increasing frequency and sophistication of cyber threats have exposed the limitations of conventional defenses. This review highlights AI's role in addressing these gaps and progress toward quantum-powered solutions. Studies targeting the challenges of critical infrastructure, IoT systems, and high-dimensional data analysis have been prioritized. This review begins with AI methodologies and then transitions to quantum-integrated approaches, reflecting a paradigm shift in cybersecurity innovation. Quantum advancements have potential; however, issues such as scalability, computational costs, and limited practical deployment remain key areas of exploration.

3.2. Search strategy

To systematically identify, collect, and organize relevant studies for an exhaustive and unbiased review, we queried the Web of Science (WoS), Scopus, and Google Scholar using the query shown in Table 2.

The methodology included databases and sources such as IEEE Xplore, Web of Science, Elsevier, Springer, the ACM Digital Library, and Scopus. EndNote was employed for bibliographic management, whereas systematic tools such as Zotero ensured traceability and organization. Complementary strategies, such as backward and forward citation tracking and expert recommendations, ensured the inclusion of influential studies.

3.3. Inclusion and exclusion criteria

A robust framework was established for selecting studies that aligned with the focus of the review and minimized bias. The inclusion criteria were studies applying AI in cybersecurity contexts, peer-reviewed articles, and systematic reviews emphasizing practical implementation. The exclusion criteria were theoretical studies without experimental validation or real-world application, articles unrelated to cybersecurity, and a lack of methodological rigor. Two reviewers independently screened the abstracts and full texts and resolved conflicts through consensus. The study selection process is summarized in the PRISMA flow diagram in Fig. 2, which details the records identified, screened, excluded, and included. Specific examples of key studies and their contributions include AI-driven frameworks for phishing detection and QML applications for zero-day threat analyses.

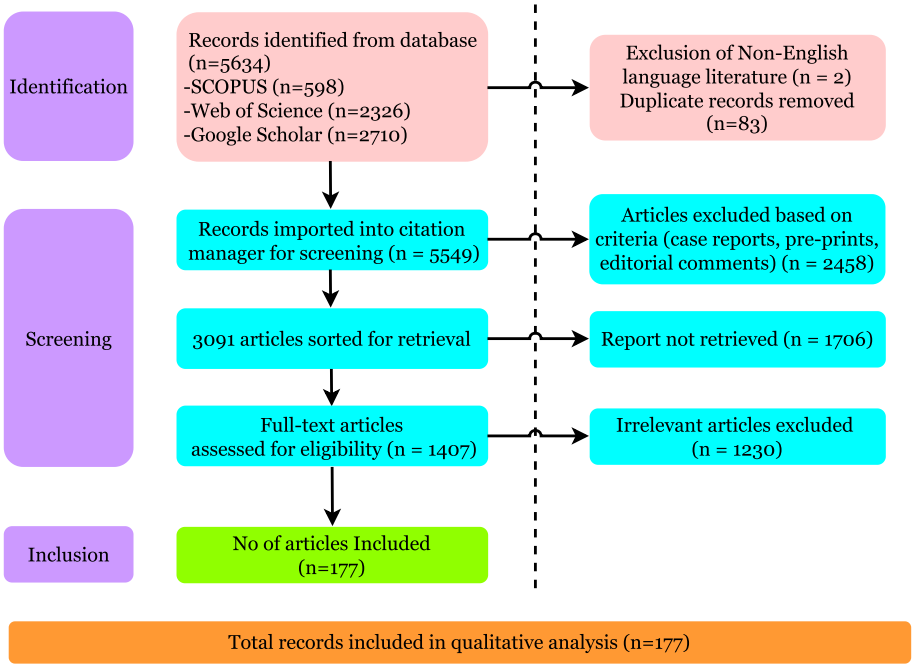


Fig. 2. Study selection process in PRISMA.

3.4. Content analysis and reporting

To synthesize the findings from the included studies, their relevance to the review objectives was emphasized. Narrative synthesis using thematic coding with NVivo was employed to uncover the role of AI and quantum computing in quantum cybersecurity. Performance metrics, such as accuracy and computational efficiency, were aggregated using RevMan for the meta-analysis. Studies have been categorized into classical AI methods and quantum techniques, respectively. Emerging trends, limitations, and future research directions are also discussed. Bar charts were created to illustrate the distribution of the literature.

3.5. Bibliographic analysis

To provide a macroscopic view of the literature, we offer insights into trends and thematic distributions. Quantum-focused cybersecurity research has grown significantly post-2020, with major contributions from publishers such as IEEE, Elsevier, Springer, and ACM, each contributing a notable share. Fig. 3 depicts the trends in publications and publishers throughout the years.

3.6. Methodological framework illustration

To enhance the transparency and reproducibility of our review process, Fig. 4 illustrates the methodological workflow adopted in this study. This framework follows a PRISMA-aligned protocol and links each research question to its corresponding analysis phase.

1. RQ1 and RQ4 guided the extraction and synthesis of AI and hybrid quantum-AI techniques across intrusion detection, malware classification, and adaptive response systems.
2. RQ2 framed the analysis of architectural integrations including digital twins, federated learning, and quantum resilience.
3. RQ3 guided the thematic coding of ethical, regulatory, and governance-related insights, particularly in bias mitigation, XAI, and FL-enhanced privacy.

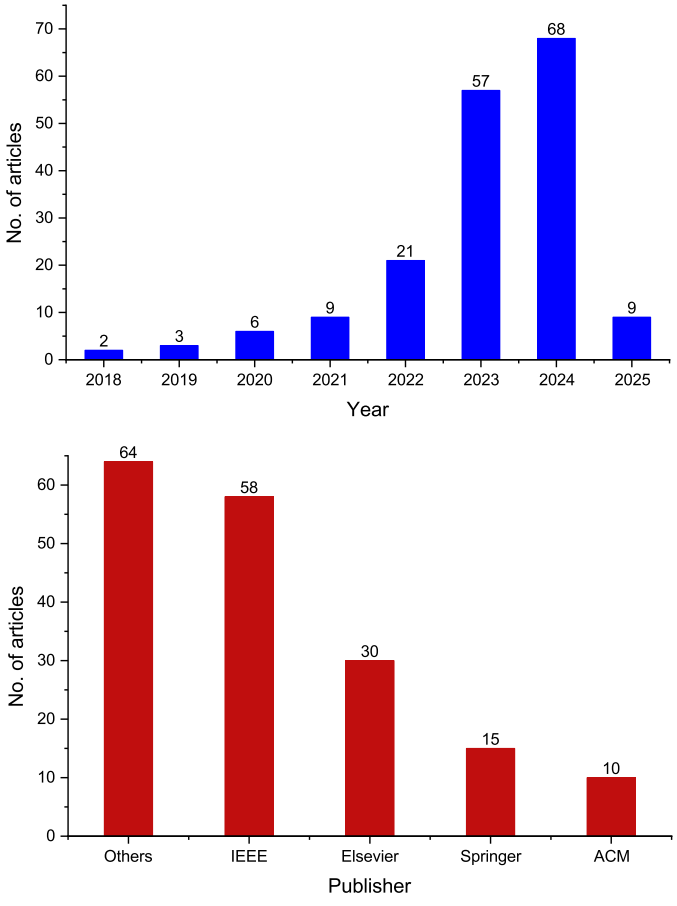


Fig. 3. Distribution of articles.

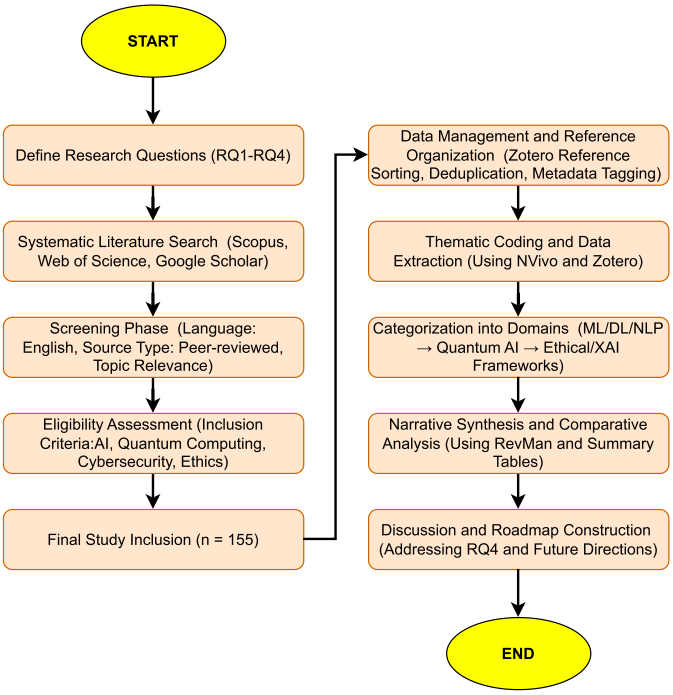


Fig. 4. Methodological framework adopted for this review (aligned with PRISMA and thematic synthesis).

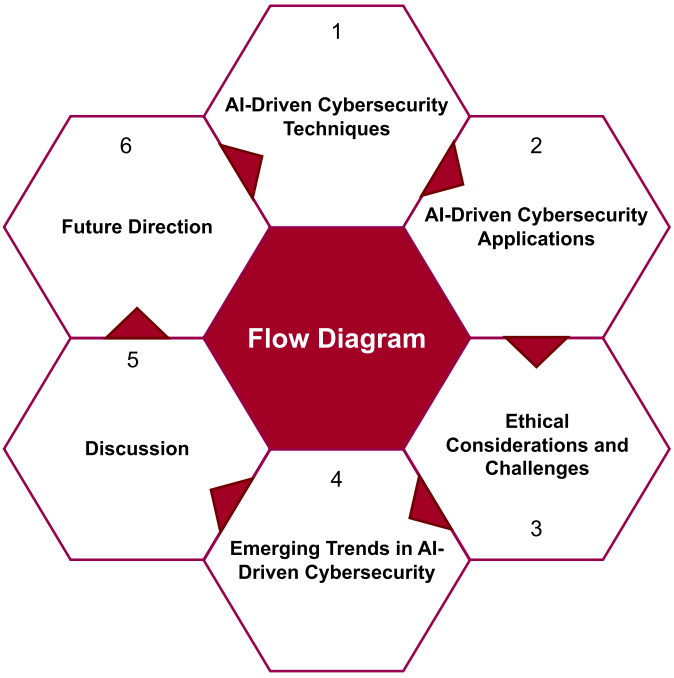


Fig. 5. Article flow diagram.

The inclusion of tools such as Zotero, NVivo, and RevMan facilitated narrative synthesis, while quantitative elements (e.g., detection accuracy and false positive rates) were aggregated into structured comparison tables. The final structure enabled the alignment of review findings with future research and policy roadmaps.

Fig. 5 shows an article-structured framework that first describes AI-driven cybersecurity techniques and then applies them to threat detection and prevention. Quantum AI has been further explored, considering its ethical dilemmas and challenges. This section discusses the advantages and disadvantages of this technology, concluding with future directions for the development of AI and quantum AI in cybersecurity.

4. AI-driven cybersecurity techniques

AI, particularly ML, has revolutionized cybersecurity by addressing the increasing complexity and sophistication of cyberattacks.

4.1. ML and its applications in cybersecurity

ML, a subset of AI, has emerged as a cornerstone of cybersecurity, transforming the manner in which organizations address evolving threats [17]. Unlike traditional rule-based security systems. ML employs data-driven approaches to analyze massive datasets, identify patterns, and predict potential threats in real time, offering adaptive and scalable defenses against sophisticated cyberattacks.

Supervised learning, which relies on labeled datasets, has demonstrated exceptional efficacy in detecting known threats. Algorithms such as Support Vector Machines (SVMs) and Convolutional Neural Networks (CNNs) have achieved malware detection accuracies exceeding 92 % by identifying novel malware through file characteristic analysis [18,19]. Naïve Bayes classifiers are widely utilized in spam filtering, categorizing emails with precision rates exceeding 90 % [20]. Similarly, Decision Trees (DT) and Artificial Neural Networks (ANNs) classify network activities as benign or malicious, achieving detection accuracies ranging from 88 % to 95 % [21,22]. In phishing detection, supervised models analyze URLs, reducing false positives and increasing recall rates to 93 %. Intrusion Detection Systems leverage anomaly detection and Principal Component Analysis (PCA) to identify deviations from normal behavior in cybersecurity [41] (Fig. 6). Despite their success, these methods depend heavily on large, high-quality labeled datasets, which is a significant limitation in real-world applications of these methods.

Unsupervised learning addresses scenarios with limited or no labeled data, focusing on anomaly detection and user behavior analysis. Algorithms such as k-means, DBSCAN, and hierarchical clustering have proven effective, achieving an accuracy of over 90 % in identifying network traffic deviations. Fig. 7 shows an intrusion detection system that leverages anomaly detection methods and PCA to analyze patterns, classify behaviors, and flag irregularities that are indicative of potential cybersecurity threats [42,43]. Dimensionality reduction techniques, such as PCA and t-distributed Stochastic Neighbor Embedding, enhance anomaly detection by simplifying high-dimensional data without losing critical information [44,45]. These methods are particularly valuable for detecting insider threats and zero-day vulnerabilities, where traditional detection mechanisms often fail. For instance, clustering algorithms applied to user behavior analysis have identified anomalous patterns linked to insider threats with remarkable precision [46].

Reinforcement learning (RL), on the other hand, introduces dynamic adaptability, enabling systems to learn optimal defense strategies through trial-and-error interactions within simulated environments. RL-based models, such as Deep Q-Networks (DQNs), have optimized automated incident responses, achieving malware detection rates of 96 % while reducing response times by 30 % [47,48]. Game-theoretic RL approaches, which combine adversarial modeling with reinforcement strategies, have demonstrated a 25 % improvement in resource allocation during Distributed Denial-of-Service (DDoS) attacks, thereby enhancing resilience in cyber-physical systems [49,50]. Despite their potential, RL methods face challenges, such as high computational costs and sensitivity to hyperparameter tuning, underscoring the need for further research and refinement. Although ML has significantly advanced cybersecurity, its challenges persist. Adversarial attacks exploit the vulnerabilities of ML models and generate polymorphic malware to avoid detection [51].

Algorithmic bias, data scarcity, and privacy concerns further complicate the deployment of ML in cybersecurity applications. To

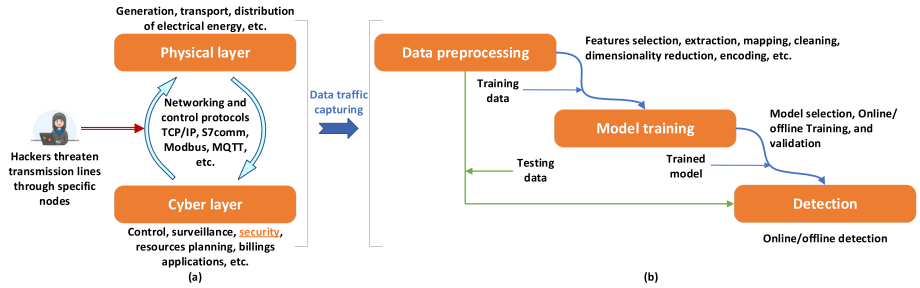


Fig. 6. An Intrusion Detection System leveraging anomaly detection and Principal Component Analysis to identify deviations from normal behavior in cybersecurity [41].

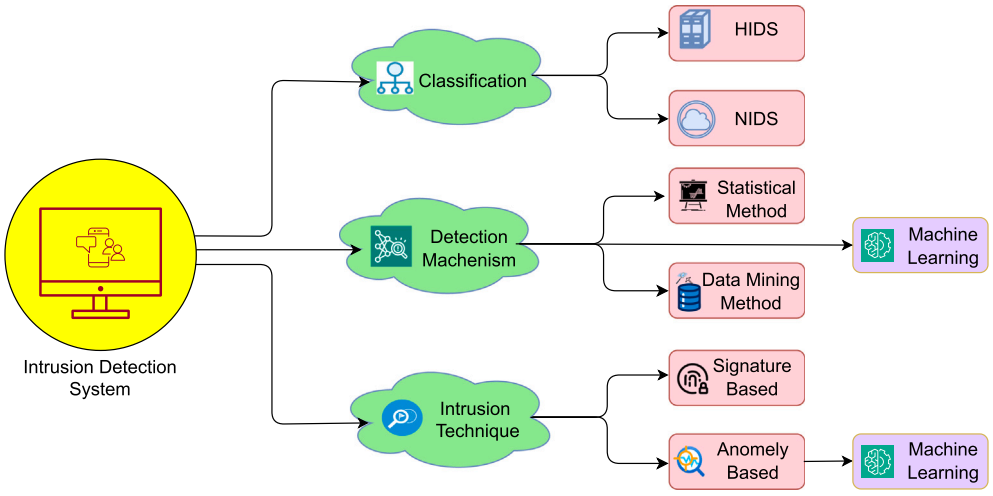


Fig. 7. Intrusion detection system and its technique.

address these challenges, researchers have proposed hybrid models that integrate supervised and unsupervised reinforcement learning approaches. For example, hybrid intrusion detection systems (IDS) that combine CNNs and k-means clustering have achieved 97 % accuracy in anomaly detection, showcasing the potential of multifaceted solutions [52]. Ethical considerations and regulatory compliance are critical for ensuring responsible AI implementation in cybersecurity [53,54].

The integration of ML into cybersecurity has marked a paradigm shift in threat detection, vulnerability assessments, and incident responses. Using supervised, unsupervised, and reinforcement learning techniques, ML-driven cybersecurity solutions promise enhanced resilience and adaptability, paving the way for robust defenses against an increasingly complex threat landscape. Despite its potential, AI in cybersecurity presents challenges such as data quality, algorithmic bias, and privacy concerns. The dual-use nature of AI further complicates its role, as cybercriminals exploit AI for malicious purposes [14].

As organizations adopt AI-driven cybersecurity, ethical issues such as algorithmic bias and privacy remain critical [47]. Effective implementation requires ongoing refinement of algorithms, robust governance, and compliance with regulations [48,49]. The integration of AI in cybersecurity continues to evolve, promising enhanced defenses against the dynamic landscape of cyber threats while requiring vigilance to address its complexities [50]. Table 3 presents a comparative performance of the cybersecurity models using ML techniques.

4.2. Deep learning applications in cybersecurity

DL, a key subset of AI and ML, is modeled based on the neural structure of the human brain. It processes extensive datasets using multilayered artificial neural networks, enabling the identification of intricate patterns. Among their architectures, Feedforward Neural Networks (FNNs) are fundamental for structured data, whereas CNNs excel in spatial pattern recognition, which is crucial for image-based malware detection [51]. Recurrent Neural Networks (RNNs) and Long Short-Term Memory networks (LSTMs) specialize in sequential data analysis, whereas transformers use attention mechanisms for superior performance in text, threat intelligence, and time-series data analysis. Fig. 8 shows the architecture model for predicting cybersecurity threats in IoT using DL [53,54]. In cybersecurity, DL has transformed the way organizations detect, analyze, and mitigate cyber threats. FNNs achieve 80–90 % accuracy in anomaly detection tasks, proving effective in static datasets, but struggling with time-sensitive data [52]. CNNs have revolutionized malware detection by converting executables into visual formats, achieving accuracy rates of up to 98 % for identifying malicious patterns [53]. RNNs and LSTMs, designed for sequential data, handle network traffic analysis and intrusion detection with 95–96 % accuracy, whereas Gated Recurrent Units (GRUs) offer computational efficiency in real-time monitoring, achieving 90–95 % accuracy in threat detection [54,55].

In addition, unsupervised models, such as autoencoders, detect anomalies by minimizing reconstruction errors, and [56] Generative Adversarial Networks (GANs) enhance robustness by synthesizing adversarial data for training [57]. Recent advancements in transformer-based architectures have demonstrated remarkable performance in real-time threat intelligence and phishing detection, achieving accuracies of 90–95 % in complex environments [54,58,59].

Despite these groundbreaking advancements, significant challenges remain in the application of DL to cybersecurity. High-performance models, such as CNNs and transformers, require substantial

Table 3
Comparative performance table for cybersecurity models using ML techniques.

Ref.	ML Method	Techniques	Application Area	Key Strengths	Results	Limitations
[17,18]	Supervised Learning	SVM	Malware detection	High accuracy in analyzing file characteristics	Malware detection accuracy > 92 %	Requires large labeled datasets
[17,18]		CNNs	Malware detection	High precision in feature extraction	Malware detection accuracy > 92 %	Requires computational resources for large-scale training
[20]		NB	Spam filtering	Simple, efficient, high precision	Spam filtering precision > 90 %	Assumes feature indedependence, limiting real-world applications
[21,22]		DT	Network activity classification	Easy interpretability	Detection accuracy: 88–95 %	Susceptible to over fitting
[21,22]		ANNs	Network activity classification	Learns complex patterns	Detection accuracy: 88–95 %	Requires significant computational resources and labeled data
[41]		SPD	Phishing URL detection	High recall rates	Recall rates: 93 %; reduces false positives	Dependent on labeled phishing datasets
[42,43]	Unsupervised Learning	k-Means	Anomaly detection	Effective for clustering	>90 % accuracy in detecting network traffic deviations	Sensitive to outliers and initial centroid selection
[42,43]		DBSCAN	User behavior analysis	Detects irregular patterns	High precision in insider threat detection	High computational complexity for large datasets
[44]		HC	Zero-day vulnerability detection	Captures hierarchical data relationships	Effective in identifying anomalous patterns linked to zero-day vulnerabilities	Lacks scalability for large-scale datasets
[44,45]		PCA	Dimensionality reduction for anomaly detection	Simplifies high-dimensional data	Enhanced detection accuracy; reduces false positives	May lose critical information during dimenmensionality reduction
[27]		t-SNE	Visualization of phishing attacks	Simplifies high-dimensional data	Effective clustering and anomaly detection	High computational cost; challenging hyperparameter tuning
[41,42]	Reinforcement Learning	DQNs	Automated incident response	Effective in high-dimensional environments	Malware detection 96 % accuracy; response time reduction: 30 %	High computational cost
[43,44]		PG	DDoS resilience	Direct modeling of adversarial settings	25 % improvement in resource allocation for DDoS resilience	Unstable training; sensitive to hyperparameters
[21,22]		Q-Learning	Adaptive security systems	Simple, fast learning for smaller state spaces	80–95 % threat reduction in adaptive access control	Struggles with large state/action spaces
[47,48]		G-T RL	Adversarial modeling	Combines game theory with reinforcement strategies	Improves resilience in adversarial cybersecurity environments	High complexity; requires accurate adversary modeling

computational resources, limiting their scalability in real-time applications, such as DDoS attack detection. Data preprocessing and quality control are critical bottlenecks because inconsistencies can degrade the model's performance. Moreover, models require frequent retraining to adapt to the constantly evolving nature of cyber threats, such as polymorphic malware and zero-day vulnerabilities. Ethical concerns, including biases in training datasets and issues with interpretability, particularly in complex architectures such as GANs, hinder transparency and trust [57,60]. Furthermore, integrating DL solutions into legacy and industrial control systems (ICS) presents compatibility challenges, underscoring the need for adaptable and scalable frameworks [56,59]. Future research must focus on addressing these challenges by developing interpretable DL models, scalable solutions for real-time processing, and robust algorithms that adapt seamlessly to evolving cyberthreat landscapes. Innovations such as flexible FL (fFL) and hybrid DL models are promising directions for overcoming these barriers while maintaining the efficacy and reliability of cybersecurity systems [61,62]. Table 4 presents a comparative performance table for the cybersecurity models that use DL techniques.

4.3. NLP and its applications in cybersecurity

NLP has become an indispensable tool in cybersecurity, addressing a diverse range of threats, from phishing to dark web monitoring and social engineering attacks. As cyber threats increasingly leverage language-based vulnerabilities, NLP models enable robust detection, analysis, and mitigation strategies by extracting meaningful patterns from textual data. Recent advancements between 2020 and 2024 have significantly enhanced the role of NLP in cybersecurity, achieving remarkable accuracy rates and improving real-time threat intelligence.

Phishing detection remains a primary application of NLP, with transformer-based models such as BERT and RoBERTa demonstrating exceptional results. Recent studies have achieved 96 % accuracy in phishing detection across multilingual datasets by analyzing text-based phishing indicators in emails and URLs. These approaches reduce false positives and improve scalability in global scenarios [63,64]. Phishing-related work combining GPT-3 and graph-based NLP techniques achieved F1 scores of 92 %, demonstrating their effectiveness in countering nuanced social engineering attacks [65,66].

Dark Web monitoring has also benefited significantly from NLP advancements. Sentiment analysis and topic modeling were employed to identify and analyze high-risk conversations in dark web forums. Recent research has demonstrated an accuracy of 89 % in detecting malicious discussions using topic-based classifiers, thereby enabling proactive threat flagging. Sentiment-based classifiers further improve hostile content detection by more than 30 % by leveraging advanced transformer models [67–69]. Named Entity Recognition (NER) is another impactful NLP application in cybersecurity. It facilitates the identification of critical entities, such as malware signatures and suspicious domains, from unstructured text data. Studies have reported recall and precision rates exceeding 90 % when extracting actionable insights from incident reports and malware descriptions. Embedding techniques such as GloVe and Word2Vec have enhanced the classification accuracy to over 93 % and streamlined threat identification [66,70,71].

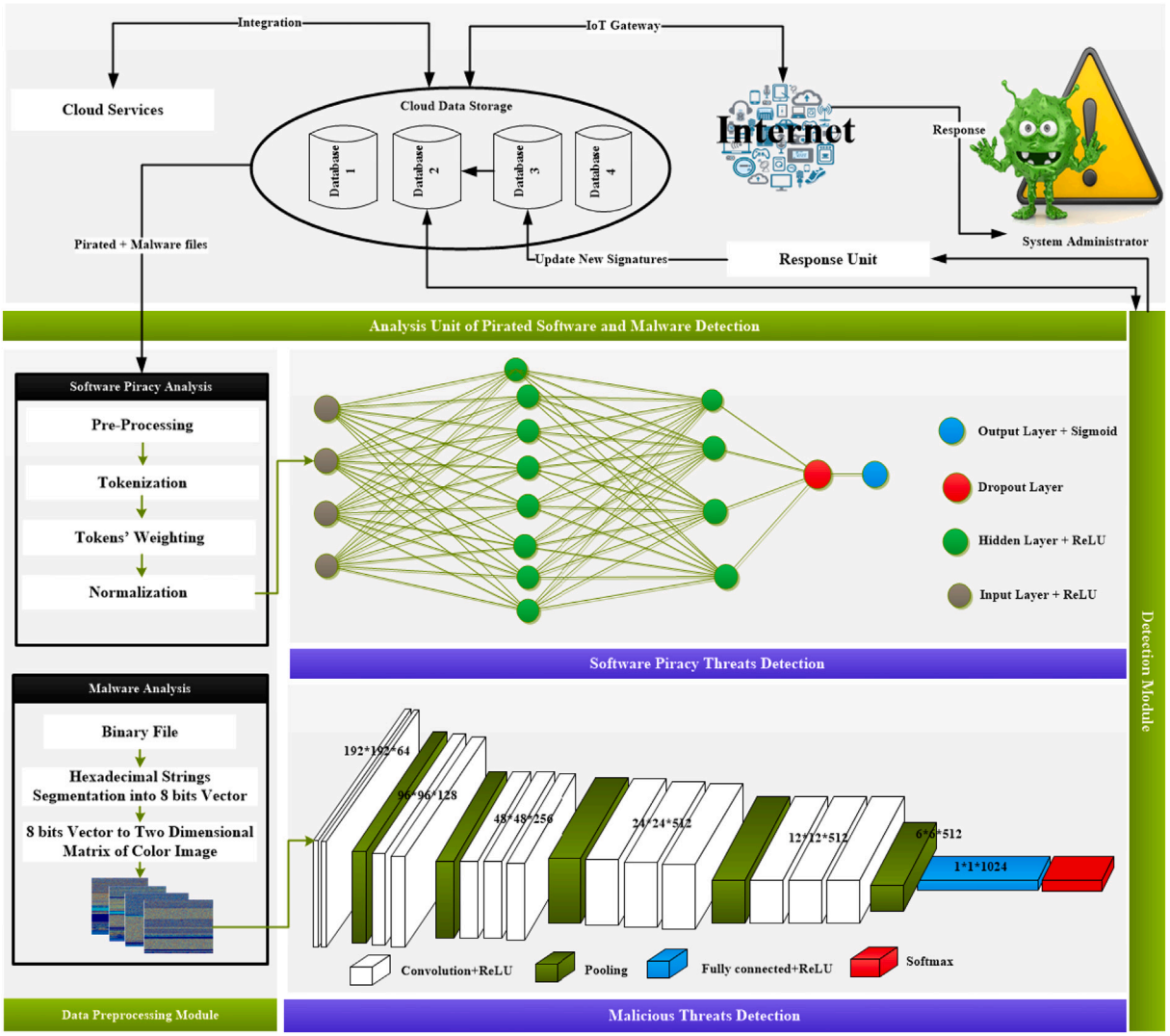


Fig. 8. Architecture model for cyber security threats prediction in IoT using DL [53].

Table 4
Comparative performance table for cybersecurity models using DL techniques.

Ref.	Model	Primary use	Accu (%)	Strengths	Limitations
[51]	FNN	Intrusion Detection	80–90	Simple architecture	High computational resource demand; Requires extensive data preprocessing
[53]	CNN	Malware Detection	90–98	Spatial data analysis	High computational resource demand; Requires extensive data preprocessing
[54]	RNN	Network Traffic Analysis	85–95	Temporal data processing	High computational resource demand; Requires extensive data preprocessing
[56]	LSTM	Intrusion Detection	90–96	Long-term dependencies	Needs frequent retraining to handle emerging threats
[53]	GRU	Real-Time Threat Detection	90–95	Efficient temporal processing	Needs frequent retraining to handle emerging threats
[57]	Autoencoder	Anomaly Detection	85–92	Unsupervised learning	Challenges in integrating with existing systems
[59]	GAN	Data Augmentation	Varies	Synthetic data generation	Potential biases in model outputs due to data imbalances
[60]	Transformer	Threat Intelligence	90–95	Long-range dependencies	Limited transparency of decision-making; High computational resource demand

Botnet detection on social media platforms has emerged as a critical area in which NLP analyzes linguistic behavior and interaction patterns to identify malicious entities. Sentiment-based botnet detection achieves 90 % accuracy, enabling the early identification of coordinated bot activities [52,53]. This capability addresses the growing prevalence of misinformation campaigns driven by automated bots. Fig. 9 illustrates the workflow for cyberattack detection using NLP, showcasing processes such as data gathering, preprocessing, vectorization, and classification, thus ensuring a structured approach to threat identification [72].

Advancements in multilingual NLP models have proven instrumental in addressing global cybersecurity challenges. Models such as multilingual BERT have demonstrated an accuracy of over 92 % in detecting threats across languages, facilitating effective cross-border threat intelligence [69]. This development aids organizations in optimizing their cybersecurity frameworks in diverse linguistic environments. The challenges posed by adversarial attacks on NLP models have also been studied extensively. Adversarial learning techniques, designed to expose and counter vulnerabilities in NLP-driven systems, improve robustness

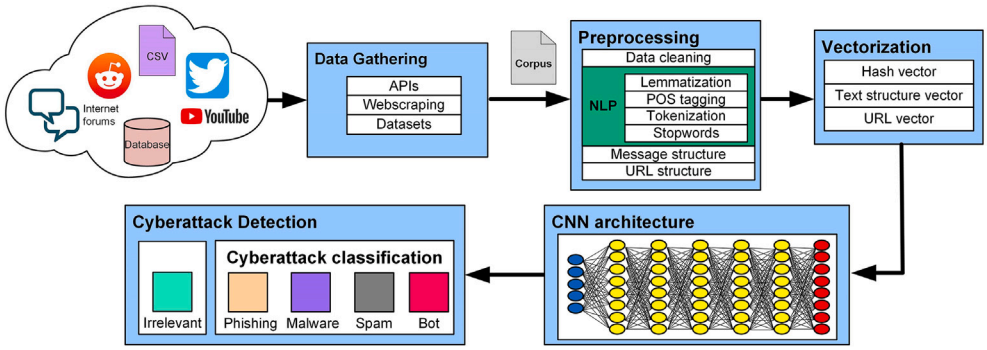


Fig. 9. Workflow for cyberattack detection using NLP, showcasing processes such as data gathering, preprocessing, vectorization, and classification [72].

Table 5
Applications and challenges in cybersecurity using NLP in AI techniques.

Ref.	Application	Method	Outcomes	Limitations
[63,64]	Phishing Detection	BERT, RoBERTa	95 %+ accuracy in phishing classification	False positives in multilingual scenarios; High computational demands
[67,68]	Dark Web Threat Monitoring	LDA, Word Embeddings	30 % improvement in identifying risks	Scalability issues with large datasets; Interpretability of topic modeling
[65,66]	Social Engineering Detection	GPT-3, XLNet	92 % F1 score in detection	High resource demand for transformer models
[70,71]	Insider Threat Detection	BERT-based Sentiment Analysis	25 % sensitivity increase	Bias in sentiment analysis models; Limited training data
[69,72]	Malware Clustering	Word2Vec, FastText	93 %+ accuracy in malware family clustering	High dimensionality in embeddings; Challenges in updating models
[64,73]	Anomaly Detection	Reinforcement Learning	90 % accuracy in anomaly detection	Slow adaptation to rapidly evolving threats
[72,74]	Multilingual Threat Detection	Multilingual BERT, Machine Translation	High accuracy in multilingual phishing detection	Translation inaccuracies for low-resource languages
[71,72]	Spam and Bot Detection	Ensemble Learning, Sentiment Analysis	91 % spam detection rate	High false positive rate for nuanced content
[69,74]	DNS Traffic Analysis	Word Embeddings	94 % accuracy in detecting malicious DNS	Limited real-time performance; Requires extensive preprocessing

by up to 25 % in real-world scenarios [73,74]. These enhancements are crucial for securing systems against manipulative inputs in high-stakes domains, such as malware detection and fraud prevention.

Adaptive anomaly detection represents another frontier, where reinforcement learning combined with NLP techniques achieves 90 % accuracy in detecting anomalies from cybersecurity logs. These models dynamically adjust to evolving threats and provide organizations with powerful tools for real-time pattern monitoring [64,72,73].

Despite these successes, several challenges persist in integrating NLP into cybersecurity. The high demand for computational resources hinders scalability, particularly for transformer-based models. Ethical concerns, such as biases introduced during training, affect fairness and reliability. Furthermore, interpretability remains a critical limitation because black-box NLP architectures often obscure decision-making processes. Addressing these issues through explainability, optimized architectures, and ethical guidelines is vital to unlocking the full potential of NLP in cybersecurity.

This growing body of research underscores the transformative impact of NLP in securing cyberspace and offering sophisticated solutions for phishing detection, dark web monitoring, botnet identification and anomaly detection. As cybersecurity threats evolve, NLP remains a cornerstone technology that drives innovation and resilience against digital adversaries. Table 5 shows the applications and challenges of cybersecurity using NLP in AI techniques.

4.4. UAI and XAI in cybersecurity

Fig. 10 shows the understanding process as UAI focuses on simplifying AI outputs, making them accessible and interpretable even for non-experts, fostering improved decision-making in Security Operations Centers (SOCs) [75] and illustrating how the process of understanding flows. In contrast, XAI provides actionable insights into the *how* and *why* of AI predictions, allowing analysts to trust and validate automated recommendations. XAI is designed to make AI systems interpretable and transparent, thereby addressing the *black box* nature of traditional AI. By providing clear explanations for its outputs, XAI builds trust and ensures that decisions are understandable, especially in critical domains such as cybersecurity, where insight into AI-driven recommendations is essential. In cybersecurity, XAI enhances threat detection and response by providing interpretable insights into complex datasets, such as network traffic, user behavior, and system logs. Unlike standard AI, which may flag threats without context, XAI explains why a threat is identified, thereby aiding analysts in understanding anomalies and attack vectors.

It also streamlines workflows by prioritizing incidents and summarizing vast security reports, enabling quicker and more informed decision-making in Security Operations Centers (SOCs). Moreover, XAI excels at analyzing unstructured data, such as dark web forums and threat intelligence feeds, uncovering emerging vulnerabilities, and facilitating proactive defense strategies.

In cybersecurity, where decisions directly impact organizational resilience, these technologies play complementary roles by bridging the gap between automation and human oversight to enable rapid responses to threats.

The application of XAI methods, such as SHapley Additive Explanations (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME), have demonstrated impressive results in enhancing interpretability. For example, SHAP identifies critical features influencing IDS outputs, achieving a 25 % reduction in false positives by pinpointing attributes such as packet size and IP addresses. Fig. 11

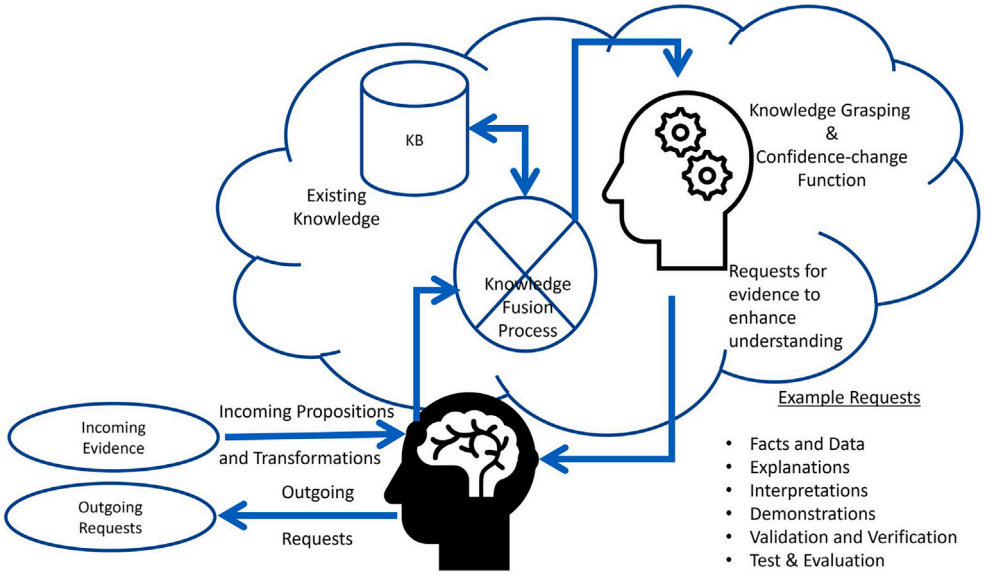


Fig. 10. Understanding process of UAI [75].

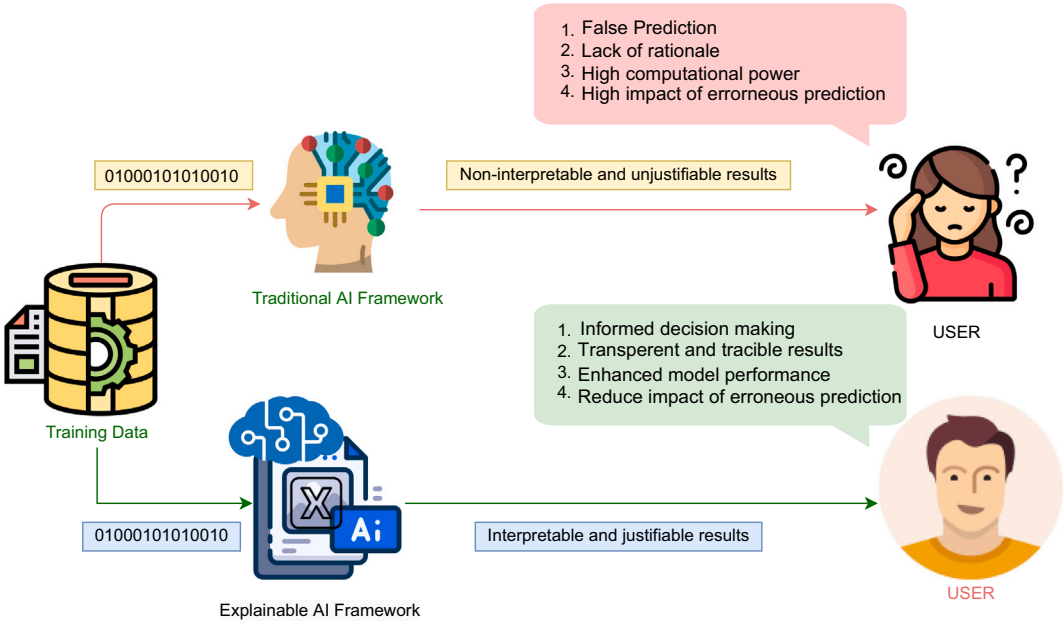


Fig. 11. XAI and their trends in cybersecuriety [76].

highlights the XAI trends in cybersecurity, showcasing its growing role in improving threat detection. For example, SHAP identifies critical features influencing IDS outputs, achieving a 25 % reduction in false positives by pinpointing attributes such as packet size and IP addresses. Security companies employing LIME in phishing detection tools have achieved a 93 % detection rate for identifying phishing attacks by evaluating the inherent features of an email, such as sender details, URLs integrated into the content, and message intent. SHAP-based interpretability techniques have been applied to Security Information and Event Management (SIEM) platforms, such as Splunk and IBM QRadar, to rate anomaly severity and enhance SOC analyst response times by 40 % [76,77]. By combining these XAI techniques, security teams can increase AI explainability and trust in automated defense, enabling regulatory compliance and ethical AI decision-making.

Integrating XAI applications into cybersecurity is crucial for enhancing transparency and decision-making. Fig. 12 illustrates the application of XAI techniques in cybersecurity workflows, highlighting their roles in intrusion detection, malware analysis and phishing prevention. Visualization techniques, such as relevance heatmaps and attention mechanisms, further enhance XAI's role of XAI by highlighting essential regions in network logs and user activity data that contribute to anomaly detection [78,79].

Rule-based methods, including decision trees and rule extraction techniques, are integral to achieving explainability in cybersecurity systems. Decision trees generate clear and interpretable rules for identifying anomalies in system logs, as demonstrated by Kim et al., where the precision rates exceeded 92 % for intrusion detection frameworks [80]. Kurniadi et al. simplified deep neural network outputs through

Fig. 12. Applications of XAI in Cybersecurity.

rule extraction and aligned model predictions with cybersecurity policies to improve malware detection accuracy [81]. Surrogate models, such as decision-tree-based simplifications of complex IDS outputs, have also proven effective in making DL models interpretable without compromising the detection rates [82].

Counterfactual explanation techniques have emerged as critical tools for exploring the predictive behavior of AI systems. Perturbation-based methods generate alternative scenarios to understand prediction shifts, such as distinguishing between legitimate and fraudulent activities in fraud detection systems [83]. Counterfactual sets further facilitate user behavior analytics by adjusting security thresholds and identifying suspicious anomalies with accuracy rates exceeding 90 % [84].

In addition, text-based explanation models play an increasingly important role in cybersecurity, particularly in phishing detection and threat analysis. For instance, sequence-to-sequence models provide contextual explanations for phishing attempts by analyzing textual patterns in phishing emails, achieving up to 95 % classification accuracy [85]. Similarly, topic modeling techniques such as LDA have been employed in threat reports to uncover emerging attack vectors and inform proactive defense strategies [86]. Autoencoders with explainability layers also contribute to identifying deviations in user behavior, thereby further enhancing anomaly detection systems in real-time environments [87,88]. Overall, by combining UAI's focus on simplicity with XAI's advanced techniques, cybersecurity systems can achieve unprecedented clarity and trustworthiness. Feature attribution methods, such as SHAP, and visual aids, such as heatmaps, ensure that AI-driven security frameworks are both effective and transparent. This dual approach allows organizations to proactively detect threats, respond decisively, and build confidence in AI-based security systems as cyber threats evolve. Table 6 lists the XAI techniques used in cybersecurity.

4.5. Introduction to QC in cybersecurity

QC is based on two fundamental principles: superposition and entanglement. Superposition allows quantum bits (qubits) to represent multiple states simultaneously, thereby enabling parallel computations that significantly exceed the capabilities of classical binary systems. Entanglement creates deep correlations between qubits, allowing complex operations to be performed efficiently. In cybersecurity, these properties facilitate faster threat detection, high-dimensional data processing, and more accurate anomaly identification, thereby transforming the capabilities of ML models.

Traditional AI approaches, including SVMs, Neural Networks (NNs), and Boltzmann Machines (BMs), have limitations in handling large-scale, real-time data due to their sequential processing and high computational costs. Quantum-enhanced ML techniques, such as QSM, Quantum Neural Networks (QNNs), and Quantum Boltzmann Machines (QBMs), overcome these challenges by leveraging quantum properties for speed, scalability, and accuracy [89–91].

The latest developments in quantum computing have placed it in a position where it can significantly transform the field of cybersecurity. Microsoft's Majorana 1 chip, a scalable topological quantum processor with one million qubits, has considerable potential for achieving fault tolerance. Topological qubits are superior to conventional qubits because they have higher reliability and reduced error rates, making them ideal for large-scale quantum applications. The Majorana 1 chip features a self-correcting topological superconductor, a new material that stabilizes Majorana particles and enables self-correcting qubits with built-in error correction. While traditional qubits are usually said to balance a pencil on their tips, Majorana qubits are more like self-stabilizing devices that actively resist environmental interference.

In contrast, Google's Willow chip implements a more conventional approach that employs a superconducting qubit architecture that utilizes Josephson junctions and quantum error-correcting codes to preserve quantum coherence [92]. Although high-fidelity operations with quantum supremacy milestones have been achieved with Willow, it suffers from significant constraints, as it requires constant error correction, which becomes more difficult on larger scales. To circumvent these issues, Microsoft's topological qubits attempt to attain hardware-based error mitigation, and therefore maximize the elimination of detrimental redundancy. Compared with current supercomputers, a 1 million qubit quantum computer would compute a million times faster, with profound implications for cryptography, drug research, and weather forecasting. Although mass deployment is still distant, breakthroughs in topological qubits (Majorana 1) and quantum error correction (Willow) signal the need for post-quantum cryptography (PQC) to safeguard digital security against future quantum attacks.

In QML–IDS, quantum preprocessing is combined with classical machine learning to increase the detection of cybersecurity insider threats. As illustrated in Fig. 13, the preprocessing of classical information comprises several steps: quantum state mapping and circuit transportation. The noise in the circuit is reduced, and quantum operations are optimized for real-world applications during circuit transportation. Qs and kernel methods were used to categorize the high-dimensional threats.

Table 6
XAI techniques in cybersecurity.

Ref.	AI technique	Description	Application	Benefits	Impact
[75,77]	SHAP	Highlights feature importance using co-operative game theory to interpret model predictions.	Intrusion Detection Systems	Enhanced Trust	Reduces false positives by 25 %
[77,78]	LIME	Provides localized explanations by approx-imating model behavior around specific inputs.	Phishing detection & authorization systems	Reduced False Positives	Improved email filtering and decision-making accuracy
[80,83]	LRP	Decomposes neural network predictions into layer-specific contributions of input features.	Malware Analysis	Real-Time Threat Insight	Speeds up malware detection analysis by 30 %
[80,81]	ABH	Leverages attention mechanisms to prior-itize critical data points contributing to model decisions.	Network anomaly detection systems	Adaptability to Attacks	Enables real-time anomaly detection with focused attention
[80,85]	Rule-Based Decision Trees	Generates interpretable rules for anomaly detection by simplifying model outputs.	Log analysis cybersecurity framework	Enhanced Trust	Achieves 92 % precision in anomaly detection
[81,84]	Rule Extraction	Simplifies DL predictions into explicit rules for malware detection.	Malware detection systems	Regulatory Compliance	Aligns predictions with cybersecurity policies
[82,83]	Counterfactual Explanations	Generates alternative scenarios to explore prediction behavior and explain anomalies.	Fraud detection and user analytics	Adaptability to Attacks	Achieves 90 % accuracy in fraud detection adjustments
[82,85]	Surrogate Models	Simplifies complex models by approximat-ing behavior using interpretable substitutes like decision trees.	Intrusion Detection Systems	Regulatory Compliance	Provides transparent audits without accuracy loss
[85,87]	Auto-encoders with Explainability	Highlights unusual patterns through reconstruction error analysis in user behaviors.	Anomaly detection in cybersecurity logs	Real-Time Threat Insight	Improves anomaly understanding
[76,86]	Topic Modeling (LDA)	Identifies key trends and topics in textual data for threat reports.	Threat intelligence and incident reporting	Enhanced Trust	Identifies emerging cyberattack vectors effectively
[85,88]	Sequence-to-Sequence Models	Generates contextual explanations for textual threats such as phishing emails.	Phishing detection in textual systems	Reduced False Positives	Achieves over 95 % phishing classification accuracy
[76,83]	Perturbation-Based Methods	Examines prediction shifts by introducing modified input features under adversarial settings.	Anomaly detection and anomaly identification	Adaptability to Attacks	Enhances model robustness against adversarial attacks
[77,79]	Attention Heatmaps	Visualizes critical attention regions of data influencing predictions.	Behavioral analysis for insider threats	Threat Insight	Speeds up real-time threat analysis by 35 %

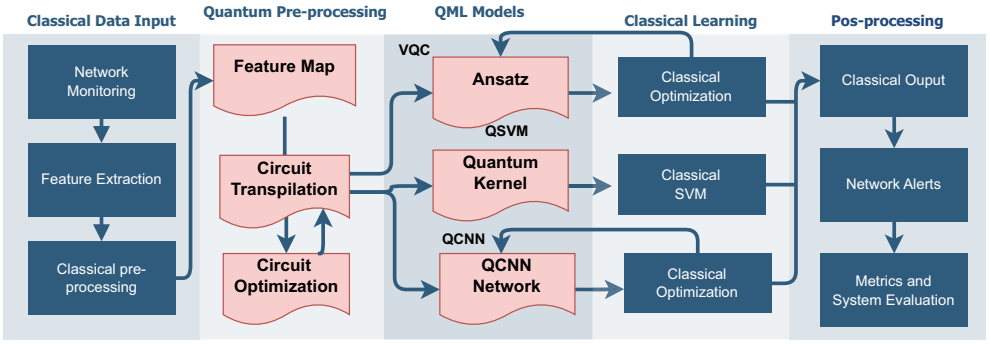


Fig. 13. Operational flowchart of Quantum ML-IDS [90].

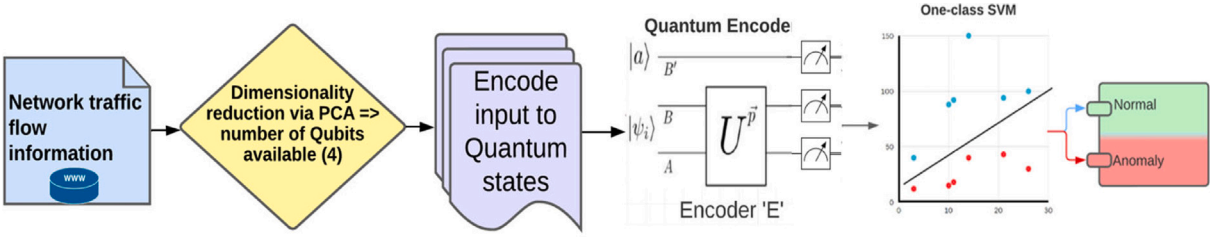
Qs and the QCNN-based QSVM Qs enhance real-time anomaly detection. Some studies have shown that QSVM can achieve high classification accuracy with a QSVC QS of 92 % on the NSL KDD dataset. This level of accuracy is higher than that of the classical SVM, with an accuracy of 87 % [90]. The classification of post-processing involves the assessment of network alerts and ensures that a correct classification is performed. This allows QML IDS to perform better than traditional systems.

Therefore, future cybersecurity strategies must prioritize quantum-resistant encryption to protect sensitive data from quantum decryption threats. Additionally, hybrid quantum-classical security models that integrate quantum AI while maintaining the robustness of classical cryptography are essential. Quantum threat simulations can further enhance cyber defense by leveraging AI-driven quantum models to predict and mitigate emerging threats. As quantum computing advances, immediate research into fault tolerance, secure frameworks, and adversarial quantum risks is crucial for ensuring a resilient cybersecurity landscape. Table 7 compares classical and QAI techniques, showcasing their advancements and contributions to the field of cybersecurity.

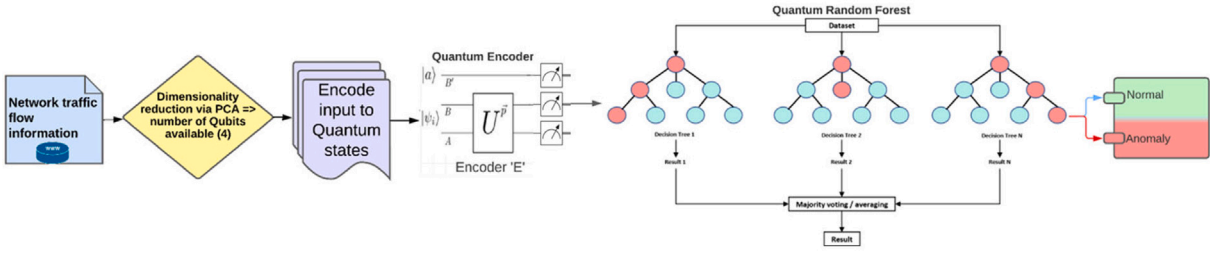
As shown in Table 7, the QSVM demonstrates a clear enhancement over classical SVMs by utilizing quantum kernels to map data into exponentially higher-dimensional feature spaces. This quantum property improves classification boundaries, enabling the QSVM to achieve 30 % greater accuracy in malware detection tasks, where classical SVMs often struggle [89,91]. The disparities between the QAE with One-Class SVM and QAE with Quantum Random Forests, as shown in Fig. 14 illustrate the efficacy of quantum feature reduction. The QAE increases class separability while compressing high-dimensional cybersecurity data to improve anomaly detection. In the first framework, QAE enables the QSVM to classify cyberthreat patterns more adeptly than the classical

Table 7
Comparison of classical and quantum techniques in cybersecurity.

Technique	Classical Approach	Quantum Approach	Key Equation	Quantum Advantages
SVM	Uses kernel functions to map data into higher-dimensional spaces for classification.	QSVM leverages quantum kernels for efficient, high-dimensional feature mapping.	Classical: $f(x) = \text{sign}\left(\sum(\alpha_i y_i K(x_i, x))\right)$ Quantum: $K(x_i, x_j) = \lvert\langle\phi(x_i)\rvert\phi(x_j)\rangle\rvert^2$	Improves malware detection by 30 % and handles overlapping classes efficiently [89,91,93].
NN	Sequential weight updates via backpropagation.	QNN exploits quantum states for parallel optimization of weights.	Classical: $h_j = \sigma\left(\sum(w_{ji}x_i) + b_j\right), y_k = \sigma\left(\sum(v_{kj}h_j) + c_k\right)$ Quantum: $\lvert\psi_{\text{output}}\rangle = U(\theta)\lvert\psi_{\text{input}}\rangle$	Reduces training time by 40 % while achieving 95 % anomaly detection accuracy [90,94,95].
BM	Classical probabilistic models use iterative sampling for pattern recognition.	QBMs leverage quantum tunneling to accelerate their convergence.	Classical: $E(v,h) = -\sum(a_i v_i) - \sum(b_j h_j) - \sum(v_i w_{ij} h_j)$ Quantum: $H = \sum(h_i\sigma_i^z) + \sum(J_{ij}\sigma_i^z\sigma_j^z)$	Achieves 25 % faster convergence for predicting cyberattack trends [96,97].
DT	Uses entropy-based splitting for decision making.	Quantum DT enhances entropy measures for faster and more accurate rule generation.	Classical: $IG = H(D) - \sum\left(\frac{\lvert D_i\rvert}{\lvert D\rvert}H(D_i)\right)$ Quantum: $\lvert\psi_{\text{decision}}\rangle = \sum(c_i\lvert\phi(x_i)\rangle)$	Achieves 92 % precision in anomaly detection with reduced computational overhead [98,99].
KNN	Computes the distances between data points for classification in the feature space.	Quantum KNN employs amplitude encoding for faster distance computations.	Classical: $d(x, x_i) = \sqrt{\sum(x_j - x_{ij})^2}$ Quantum: $d_Q(x, x_i) = 1 - \lvert\langle\psi(x)\rvert\psi(x_i)\rangle\rvert^2$	Enables real-time anomaly classification in high-dimensional datasets [99,100].
RF	Classical random forests are used as decision trees in classification.	QF integrates quantum-enhanced decision analysis for optimization.	Classical: $\hat{y} = \frac{1}{T}\sum f_t(x)$ Quantum: $\lvert\psi_{\text{forest}}\rangle = \sum(c_t\lvert\psi_t(x)\rangle)$	Provides faster classification with more robust decision boundaries [101,102].
ML-IDS	Traditional IDS systems require manual tuning and feature extraction.	QML-IDS automates the detection using quantum-enhanced anomaly detection models.	Classical: $P(y\lvert X) = \frac{P(X\lvert y)P(y)}{P(X)}$ Quantum: $P_Q(y\lvert X) = \lvert\langle\psi(y)\rvert U(\theta)\rvert\phi(X)\rangle\rvert^2$	Improves detection accuracy by 20 % while reducing false positives [93,94].



(a) Framework 1: Union of QAE and one-class SVM.



(b) Framework 2: Union of QAE and quantum random forest.

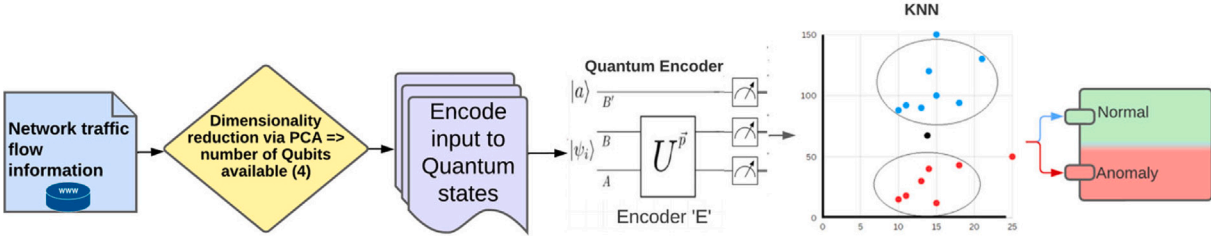


Fig. 14. Comparative figures showcasing QSVM's ability to efficiently resolve overlapping classes showcase two frameworks: (a) Quantum Autoencoders (QAE) combined with a one-class SVM and (b) QAE integrated with quantum random forests for network traffic anomaly detection [93].

SVM, which means that the QSVM can identify a greater number of patterns.

In the second, more advanced framework, intrusion detection is refined in real-world scenarios using data-encoded QAE and improved decision boundaries within Quantum Random Forests. These QAE forest models show heightened performance in intrusion detection systems through greater accuracy in identifying diverse patterns of attacks compared to classical QAE models [93].

By reducing the dimensionality using PCA and encoding the inputs into quantum states, these frameworks achieve superior anomaly classification. Quantum forests (QF) leverage quantum-enhanced random trees to optimize decision-making and scale efficiently for high-volume

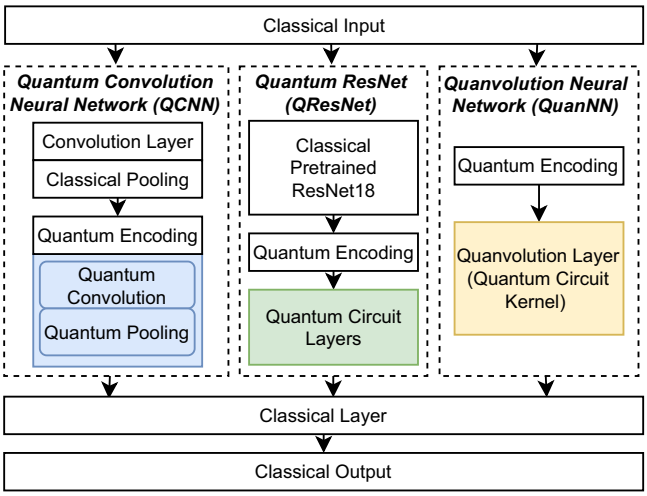


Fig. 15. Quantum comparison in DL models [100,101,103].

cybersecurity data [91,93,95]. Similarly, QNNs use quantum states to enable the parallel processing of weights during optimization, significantly reducing the training time by 40 % while maintaining 95 % detection accuracy in anomaly identification tasks [96,97]. This makes QNNs highly suitable for adaptive intrusion detection systems in which real-time threat responses are critical. On the probabilistic side, Quantum Boltzmann Machines (QBMs) leverage quantum tunneling to explore energy states efficiently and achieve faster convergence when predicting attack trends and forecasting vulnerabilities [98]. Other quantum-enhanced models include Quantum Decision Trees (QDTs), which integrate quantum entropy measures to improve rule generation and decision-making precision to 92 % in large-scale anomaly detection frameworks [100,104]. Quantum K-Nearest Neighbors (QKNN) methods accelerate distance calculations using amplitude encoding, achieving efficient real-time classification of anomalies in high-dimensional cybersecurity spaces [100,103].

In QDL, QCNNs and Quantum Autoencoders optimize feature extraction for malware detection and anomaly identification. Studies have reported that malware classification accuracy reaches 98 % with reduced computational costs compared with their classical counterparts [101,105]. In Fig. 15, we compare quantum convolutional neural networks (QCNN), quantum ResNet (QResNet), and quanvolutional neural networks (QuanNN). The QCNN superposition layers significantly improve feature extraction in malware classification, leading to higher accuracy and economic efficiency in terms of computational resources. Unlike classical CNNs with rigid set filters, QCNNs proactively respond to data complexity by modifying their parameters, which improves their performance in competitive situations. In addition, the incorporation of quantum encoding in QResNet significantly improves its deep learning capabilities, enhancing its performance by 40 % and maintaining an acceptable workload performance. Such quantum supremacy leads to more powerful malware recognition, faster training convergence, and better defense against new cyber-attacks [102]. Meanwhile, quantum-enhanced NLP models apply sequence-to-sequence techniques and topic modeling for phishing detection and threat intelligence extraction, achieving 95 % accuracy in identifying phishing emails and analyzing dark web data [106–108]. Despite these benefits, challenges persist in fully realizing the potential of quantum-AI. Hardware limitations, such as the scarcity of stable and error-free qubits, remain a bottleneck for practical implementation. Noisy quantum systems introduce errors that require advances in quantum error correction and noise reduction techniques. Additionally, the need for hybrid quantum-classical systems to integrate quantum algorithms with existing infrastructure adds complexity to the deployment. Research on scalable quantum hardware, algorithm optimization, and robust hybrid frameworks is actively addressing these challenges. As quantum technology matures, its integration into cybersecurity redefines the landscape, enabling organizations to proactively counter emerging threats with unmatched speed, precision, and efficiency [107–109]. Table 8 presents a comparative analysis of the AI techniques used in cybersecurity applications.

5. AI-driven cybersecurity applications

AI has significantly revolutionized cybersecurity by enhancing the automation, accuracy, and scalability of the detection, prevention, and response to evolving cyber threats. By leveraging ML, DL, and behavioral analysis, AI addresses the critical areas of threat detection and prevention, vulnerability assessment, and incident response, ensuring that organizations remain resilient to sophisticated attacks.

5.1. Threat detection and prevention

AI-driven IDS have achieved remarkable accuracy improvements while reducing false positives. Studies have reported that a decision-tree-based IDS achieves a detection accuracy of 95 % with only 2 % false positives [110], whereas neural networks demonstrate robustness by dynamically adapting to real-time network changes [111]. Reinforcement learning models further optimize detection efficiency by learning from evolving attack patterns and improving response accuracy [112]. Malware detection benefits immensely from AI techniques that integrate static and dynamic analyses. CNNs have achieved a malware detection accuracy of 98 %, outperforming traditional signature-based methods [113]. Hybrid models combining static and dynamic analyses reduce false positives by 30 % and improve the detection rates for zero-day malware [114]. AI-enhanced sandboxes for behavioral analysis demonstrated a detection accuracy of 97 %, enabling organizations to efficiently identify malware variants [115]. AI-based NLP and image recognition techniques have demonstrated significant advancements in phishing detection. For instance, an RNN model analyzing textual and URL patterns achieved a 97 % detection accuracy with reduced false positives [116]. Furthermore, image-based phishing detection, which analyzes website logos and page structures, achieves a 99 % detection rate [117], thereby surpassing the limitations of traditional methods. Bot detection and mitigation are critical for protecting systems from automated attacks. ML models analyzing IP behavior and interaction patterns have achieved a detection accuracy of 95 % with a false positive rate of 3 % [118]. In addition, behavioral analysis techniques can effectively identify bot networks by detecting anomalies in large-scale traffic patterns [119].

Data exfiltration detection relies on AI models that monitor the user's behavior and access patterns. Studies have shown that AI-based anomaly detection reduces data breaches by 40 % [120], whereas pattern recognition models achieve a detection accuracy of 94 %, ensuring the timely identification of abnormal data transfer activities [121].

AI-enabled passwordless authentication systems leverage biometric techniques, such as facial recognition and fingerprint scanning, to achieve a verification accuracy of 99 % [122]. Behavioral biometrics, which analyze user interaction patterns, further enhance authentication security by reducing susceptibility to brute-force attacks [123].

Behavior-based threat analysis plays a crucial role in identifying deviations in user behavior that are indicative of malicious activity. AI models have achieved 85 % detection accuracy while maintaining false-positive rates below 5 % [124]. Reinforcement learning further strengthens these models by continuously adapting to new threat behaviors, thereby improving the detection rate to 92 % [125].

5.2. Vulnerability assessment

Automated vulnerability scanning powered by AI significantly enhances precision and reduces the number of false positives. DL-based scanners have achieved an accuracy of 85 %, reducing false positives by 20 % [126]. Popular tools, such as Nessus and Qualys, integrate

Table 8
Comparative analysis of AI techniques for cybersecurity applications with performance metrics and dataset information.

Ref.	Application	Technique used	Accuracy	False-Positive Rate	Dataset used	Comp. time
[109]	Intrusion Detection Systems	Decision Trees	95 %	2 %	KDDCup99	0.5s per detection
[111]		Neural Networks (Real-Time Adaptation)	High	Not specified	NSL-KDD	1.2s per event
[112]	Malware Detection	Convolutional Neural Networks	High	Not specified	MalImg Dataset	3s per sample
[113]	Phishing Detection	Hybrid Static-Dynamic Analysis	97 %	Reduced by 30 %	VirusShare	1s per hybrid analysis
[114]		Recurrent Neural Networks	97 %	Reduced by 20 %	PhishTank	0.8s per URL
[115]	Bot Detection	Image Recognition (Logos & Layouts)	99 %	Reduced by 15 %	Custom Dataset (Images)	1.5s per image
[116]		Pattern Recognition	High	Minimal	Alexa Traffic Data	1s per real-time detection
[118]	Data Exfiltration Detection	Anomaly Detection	94 %	Low	CERT Insider Threat Dataset	3.8s per anomaly
[119]		Pattern Recognition & Tracking	94 %	Low	Custom User Access Logs	1.8s per detection
[120]	Passwordless Authentication	Biometric Authentication	94 %	Not specified	Real-World Biometric Data	1.5s per analysis
[121]		Behavioral Biometrics	Not specified	Not specified	Custom User Interaction Data	1.5s per analysis
[122]	Behavioral-Based Threat Analysis	Anomaly Detection	85 %	Not specified	Custom Behavioral Dataset	2s per evaluation
[123]		Reinforcement Learning	92 %	Not specified	Simulated Behavioral Data	1.5s learning cycles
[124]	Spam and Malicious Content Detection	NLP	96 %	Not specified	Enron Email Dataset	1s per email
[125]		NLP for Malicious Content in Social Media	94 %	Not specified	Twitter Social Media Dataset	1.2s per post

AI algorithms to optimize detection and patch prioritization [127]. AI-augmented penetration testing automates complex attack simulations to uncover vulnerabilities with minimal human intervention. Research indicates that AI-enhanced penetration testing increases the vulnerability coverage by 40 % while uncovering previously hidden flaws [128]. Graph-based AI techniques simplify penetration testing workflows, enabling security teams to efficiently identify and address vulnerabilities [129]. Predictive vulnerability management uses AI models to forecast vulnerability exploitability, enabling organizations to prioritize critical patches. ML-driven predictive tools reduce breaches by 50 % [130] and decrease mitigation time by 30 % [131], transforming vulnerability management into a proactive defense strategy. AI-powered threat simulation and attack-path mapping help organizations visualize the attack pathways and improve patch deployment. Studies have reported a 35 % improvement in threat visibility and a 25 % reduction in patch application times using AI-driven simulation tools [132,133]. These advancements enable organizations to efficiently prioritize mitigation efforts. In IoT vulnerability detection, AI systems identify firmware anomalies and zero-day exploits with a 70 % higher accuracy than traditional methods [134]. ML models reduce false positives by 15 %, thereby improving the security of connected IoT devices [135]. AI also enhances social engineering detection by analyzing user behavior to identify susceptibility to phishing attacks. Behavioral threat analysis reduces successful phishing attempts by 60 % [136], whereas AI-based training systems significantly improve user awareness and responses [137]. AI-driven application security testing combines static and dynamic code analysis to effectively identify complex vulnerabilities. DL-based tools improved the detection accuracy by 20 % and halved the scanning time compared with traditional approaches [138,139].

5.3. Incident response

AI automates incident response tasks, thereby improving the efficiency of threat containment, malware removal and system recovery. Neural networks have achieved a 95 % success rate in automating malware removal processes [140]. Similarly, AI-powered recovery tools restore critical systems with 98 % accuracy, minimizing downtime and accelerating recovery [141]. These advancements reduce the incident containment time by 80 %, thereby ensuring timely threat mitigation [142].

AI enhances Security Orchestration, Automation, and Response (SOAR) platforms by streamlining workflows and automating alert prioritization. Studies have shown that AI-based prioritization algorithms reduce false positives by 25 % [143], whereas DL models achieve a 92 % classification accuracy for alerts, significantly improving analyst productivity [144]. AI-powered SOAR systems automate 70 % of incident response tasks, reducing response times by 60 % [145].

Threat hunting and investigation benefit from AI-driven automation, improving the detection of novel threats by 30 % [146]. ML tools that analyze logs and traffic have reduced manual investigative efforts by 40 % [147]. Additionally, automated root cause analysis tools enhance accuracy by 35 %, providing actionable insights to strengthen defense strategies [148].

AI also streamlines incident documentation and reporting processes, reducing documentation times by 50 %, while ensuring compliance and accuracy [149]. AI-facilitated post-incident analysis tools improve preparedness by identifying strategy gaps and enhancing playbook effectiveness [150].

5.4. Benchmark comparison with state-of-the-art

To reinforce confidence in the proposed integrated view of AI- and quantum-enabled cybersecurity, we synthesize recent experimental results from representative state-of-the-art studies across intrusion detection, malware detection, phishing classification, and incident response. In addition, recent quantum-assisted IDS prototypes have reported competitive results; for example, a scientific report study using quantum outlier analysis achieved 99.87 % DDoS detection accuracy on benchmark traffic, while noting dataset sensitivity and current hybrid hardware constraints [151]. Table 9 summarizes the reported datasets, model families, and headline metrics (e.g., accuracy, F1-score, and false positive rate). The comparison shows that deep learning and ensemble methods often exceed 95 % accuracy on established benchmarks, whereas explainable models reduce false positives with modest trade-offs in recall. Quantum-enhanced approaches have reported promising results on small and structured datasets, although large-scale validations remain

Table 9
Comparative overview of AI, hybrid, and quantum techniques in cybersecurity.

Ref.	Technique	Application domain	Model type	Strengths	Limitations
[153]	Convolutional Neural Networks	Malware classification, ransomware detection	Deep Learning	High accuracy with raw feature learning; effective for image-like binary representations	Data hungry; limited interpretability
[154]	Random Forest + SHAP	Phishing and fraud detection	Hybrid XAI	Feature-level transparency; robust against overfitting	Sensitive to data imbalance; bias propagation
[155]	Support Vector Machines	Intrusion detection, anomaly detection	Classical ML	Strong performance on smaller datasets; interpretable margins	Scalability issues with large data; requires feature engineering
[152]	Federated Learning	IoT and healthcare IDS	Distributed Learning	Privacy-preserving; decentralized training	High communication cost; vulnerable to poisoning attacks
[156]	Autoencoder / GAN-based IDS	Novel attack detection	Deep Generative Models	Good for zero-day and anomaly detection; learns hidden patterns	Prone to mode collapse; training instability
[157]	Quantum Neural Networks	Quantum intrusion detection	Quantum ML	Potential exponential speedup; handles high-dimensional data	Early stage, lacks hardware scalability
[158]	Post quantum Cryptography (Lattice, Code-based)	Secure communication	Cryptographic Algorithms	Resistant to quantum attacks; backed by NIST efforts	Computationally intensive; deployment challenges
[159]	Digital Twins + AI	Cyber-Physical System security	Simulation + AI	Enables predictive monitoring; real-time threat simulation	High deployment complexity; resource intensive
[160]	Blockchain + AI integration	Secure data provenance, IoT	Hybrid Decentralized Systems	Immutable logs; improves trust and transparency	Latency and scalability issues; energy costs
[161]	Explainable AI (SHAP, LIME, Counterfactuals)	Trust in IDS and phishing classifiers	Model-Agnostic	Improves interpretability and accountability; regulatory alignment (GDPR/AI Act)	Performance explainability trade-off

limited. This evidence supports the need for scalable, interpretable, and quantum-resilient designs that can transition from benchmarks to real-world settings. Consistent results have been reported for FL-based IDS in IoT environments [152] and hybrid QSVM prototypes in cyber-physical settings [94], reinforcing the viability of our deployment-first recommendations. These comparisons motivate the design choices and research directions discussed next, including the integration of explainability, privacy-preserving training, and quantum resilience in operational settings.

6. Ethical considerations and challenges

We adopt a working definition of *ethical AI* consistent with current alignment literature, emphasizing the RICE objectives (Robustness, Interpretability, Controllability, and Ethicality) as operational goals for trustworthy, auditable security systems [162,163]. We use *quantum resilience* to denote cryptographic and system-level readiness for adversaries with large-scale quantum capability, consistent with recently finalized NIST PQC standards [30–32]. AI-driven cybersecurity systems have revolutionized the threat detection and response. However, their deployment raises critical ethical concerns, including bias in algorithms, adversarial attacks, privacy risks, and the need for explainability. Addressing these issues is vital for ensuring the ethical, secure, and effective integration of AI into cybersecurity.

6.1. Bias in AI algorithms

Bias in AI arises when models are trained on imbalanced or non-representative datasets, leading to discriminatory outcomes. For instance, IDS trained on biased data showed a 28 % increase in false positives [164]. Similarly, facial recognition systems exhibited a misidentification rate of 34 % for women of color compared with 1 % for white males, demonstrating algorithmic bias [165]. Addressing this requires fairness-aware training methods, which have been shown to reduce bias by up to 35 % [166].

6.2. Adversarial attacks

Adversarial attacks focus on mitigating the need to follow strict protocols when using AI frameworks and instead alter the input provided to bypass models. Using adversarial perturbations as an example, it can be noted that the performance metrics of malware detection systems decrease by 50 %. A similar case is that of AI-based spam filters, whose performance is breached with the use of phishing emails with a 93 % success rate [167]. Similarly, AI-powered spam filters were bypassed with a 93 % success rate using carefully crafted emails [168]. Fig. 16 illustrates the adversarial perturbations applied to an image of a stop sign, which is part of an AI system that was created to ensure an excellent defense against such abuse. These measures include adversarial training, model-hardening processes, and anomaly recognition to increase the resistance of AI to cyberattacks. In addition, GANs can help provide additional protection by identifying and removing adversarial inputs as they emerge.

6.3. Privacy concerns

AI systems require the collection of massive amounts of data, which puts privacy at risk. According to a study conducted in an industry in 2024, 68 % reported that privacy infringement is one of the greatest challenges they face while deploying AI. A specific case of a security breach was noted in 2023 when the private information of six million users across multiple borders was revealed due to poor security measures fueled by AI [169]. However, some privacy-preserving AI methods, such as Federated Learning (FL), differential privacy, and homomorphic encryption, can mitigate privacy risks by up to 40 % without compromising AI performance [170]. To ensure optimal security, hybrid frameworks combining human expertise and AI have been implemented to protect privacy and defend against destructive cyber threats. Privacy-enhancing technologies (PETs) must be deployed in conjunction with global regulations [171].

6.4. Explainability and trust in AI security systems

One of the greatest hurdles in AI cybersecurity is the lack of transparency and explainability in AI-powered security decision-making processes. Many AI models operate as black boxes, making it difficult for analysts to understand the reasons behind the generation of a security alert. In a survey conducted among cybersecurity professionals, 73 % reported that they preferred XAI models because these allowed better decision validation and accountability [172]. Fig. 17 illustrates the FL and XAI frameworks for cybersecurity, which integrate black-box and interpretable models. Using XAI techniques such as SHAP, LIME, and saliency maps improves transparency and trust by 30 % and reduces incident response times by 18 % [173]. Other approaches, such as causal inference models and rule-based AI, are being developed to improve the correlation opacity in cybersecurity. Causal inference models with AI

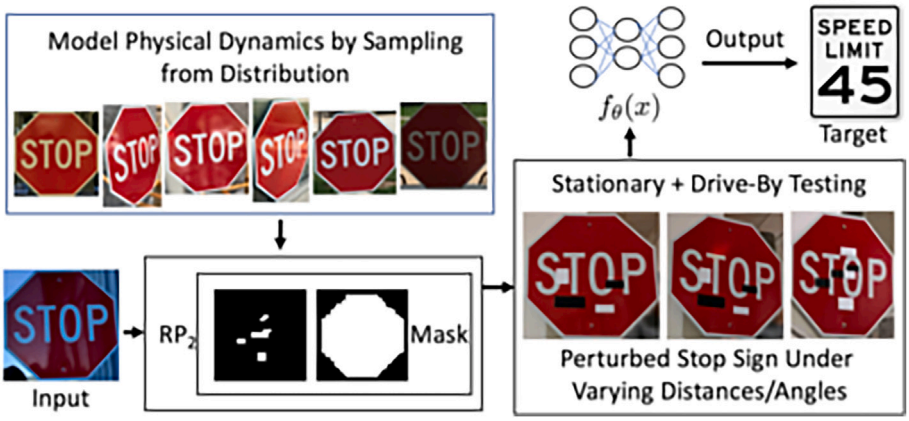


Fig. 16. An example of adversarial noise applied to a stop sign, leading to misclassification as a speed limit sign. This demonstrates how subtle perturbations can deceive AI models, highlighting the need for adversarial defenses [168,169].

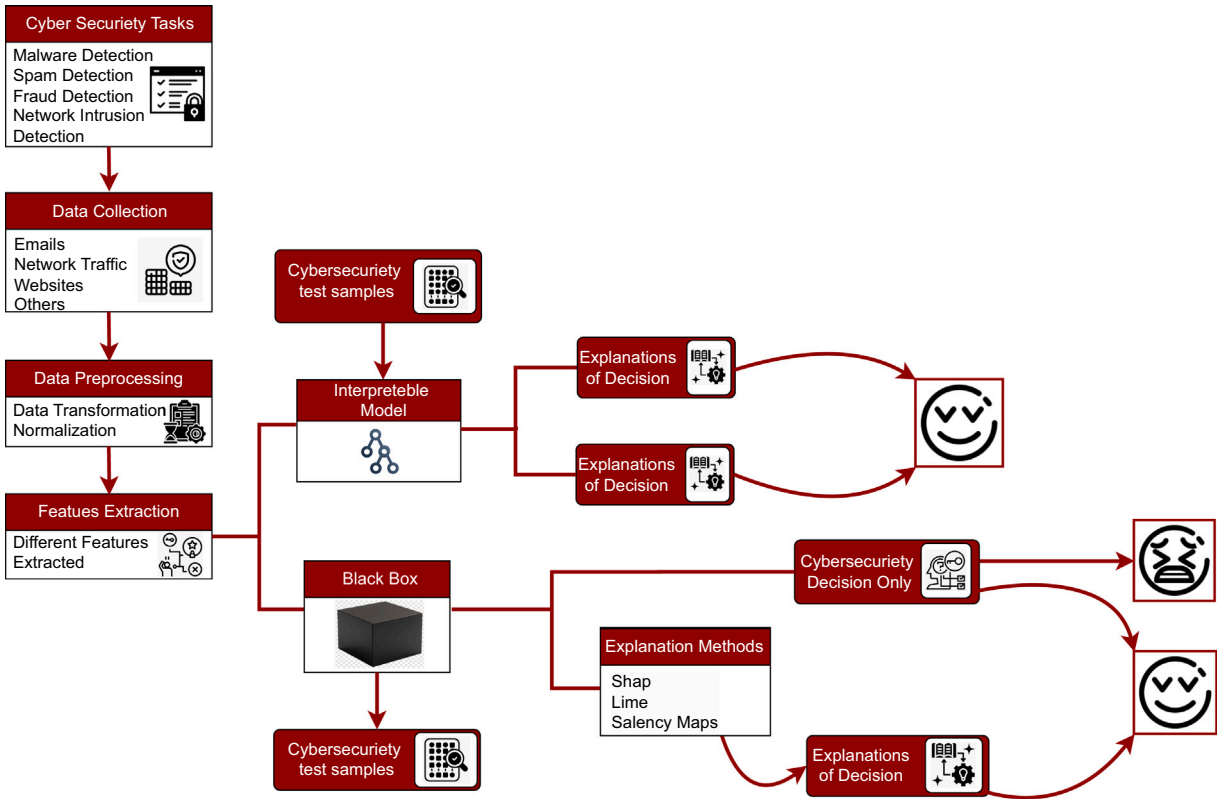


Fig. 17. FL with XAI for Cybersecurity: Combining interpretable and black-box models with SHAP, LIME, and saliency maps for transparent decision-making. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

tools and rule-based AI systems help to increase transparency in cybersecurity. Other visualization tools, such as heatmaps, decision trees, and AI attention mechanisms, explain what happens in AI security, making security decisions sharper, clearer, accountable, and reliable. Therefore, understanding trust in AI-enhanced cybersecurity solutions requires regulatory support, compliance with ethical AI criteria, and monitoring of AI functions to ensure they do not produce absurd outcomes. As AI technologies improve, these ethical challenges must be considered if public trust in effective AI-powered cybersecurity solutions is to be retained.

6.5. Standards-aligned ethical guardrails

To make ethics actionable in security contexts, we align RICE objectives with established governance instruments. For IDS and phishing detection in regulated sectors, auditability and risk controls follow the NIST AI RMF 1.0 functions (Map-Measure-Manage-Govern) and ISO/IEC 42,001 requirements on AI management systems (policy, risk assessment, controls). For models that process personal data (for example, FL-based telemetry), we apply GDPR principles (purpose limitation, data minimization) and the EU AI Act duties for high-risk systems as they are phased in. For system-level migration to post-quantum security, we mapped crypto choices to NIST FIPS 203/204 (ML-KEM and ML-DSA) to ensure that the signature and key-encapsulation paths remain compliant during the transition. Table 10 summarizes how representative cybersecurity use cases can be aligned with ethical controls and mapped to recognised standards and guidelines, ensuring that principles are translated into actionable safeguards.

Table 10
Alignment of cybersecurity use cases with ethical controls and standards.

Use case	Ethical control	Relevant standard/guideline
AI-driven Intrusion Detection	Auditability, transparency in decision-making	IEEE Ethically Aligned Design; ACM Code of Ethics
Federated Learning for IoT and healthcare	Data minimization, privacy preservation, bias mitigation	GDPR principles; EU AI Act requirements for high-risk systems
System-on-Chip security monitoring	Risk assessment, accountability, fairness-aware algorithms	ISO/IEC 42001:2023 AI Management Systems
Post-Quantum Cryptography deployment	Resilience against quantum attacks, secure key management	NIST FIPS 203 (ML-KEM), FIPS 204 (ML-DSA)

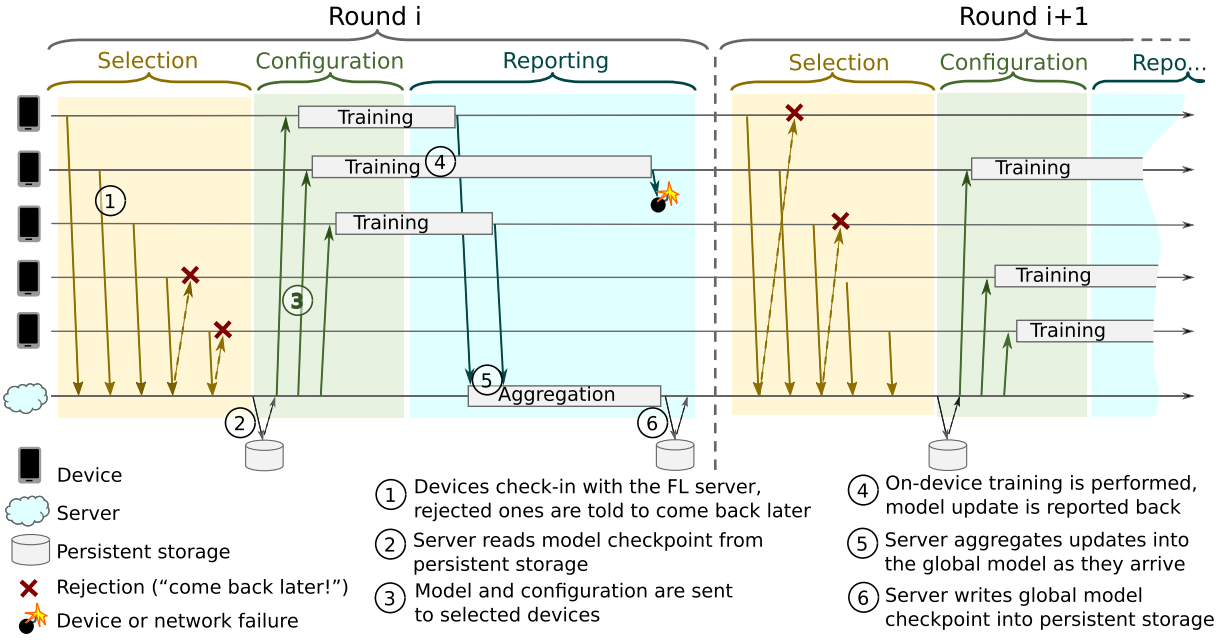


Fig. 18. FFL Workflow for Decentralized Threat Analysis. FFL enables model training across devices while preserving data privacy [178].

7. Emerging trends in AI-driven cybersecurity

AI-driven cybersecurity is at the forefront of digital innovation and offers transformative solutions to counter evolving cyber threats. Emerging trends such as FL, XAI, and quantum-resilient frameworks form the foundation for building adaptive, secure, and transparent security systems. FL has revolutionized privacy-preserving threat detection by enabling decentralized model training across multiple devices to ensure data protection. Industries such as healthcare, finance, and IoT ecosystems have already reported a 15 % boost in detection efficiency while maintaining regulatory compliance [174,175]. For instance, Google's federated malware detection model has demonstrated substantial improvements in identifying threats without sharing sensitive raw data, thereby enhancing both efficiency and privacy [176]. As security demands intensify, FL will continue to underpin scalable and secure cybersecurity infrastructures, offering resilience against modern cyber risks [177]. Fig. 18 provides an overview of the FL protocol workflow, demonstrating its ability to deliver a decentralized threat analysis while preserving data integrity [178]. XAI has emerged as a solution to the growing demand for transparency and trustworthiness in AI-driven cybersecurity systems. By enabling interpretable decision-making, XAI provides insights into threat detection processes, reduces false positives, and improves incident response efficiency by 20–30 % [179,180]. When integrated into SOAR (Security Orchestration, Automation, and Response) platforms, XAI significantly streamlines workflows and accelerates threat mitigation by clarifying flagged anomalies and decisions [181,182].

For example, XAI-enabled models provide real-time explanations of detection outcomes and foster trust among the security analysts. They also highlight the XAI output and its ability to enhance the interpretability and improve confidence in AI-driven threat classification results. Digital Twin technology has emerged as a promising innovation in the field of cybersecurity, offering predictive counter-cyber threat measures along with real-time simulations of IT infrastructure and networks. Replicating the digital space of an organization enables security personnel to evaluate self-driven security system responses, predict system weaknesses, and fine-tune extemporaneous reaction plans without interfering with the actual system. Automatic surveillance, anomaly spotting, and simulated scenario-based threat identification and response actions undertaken by digital twins result in remarkable improvements in cyber resilience toward perpetually mutating cyber threats.

QC introduces unprecedented and complex challenges in cybersecurity. With rapid advances in quantum systems, PQC has become a vital defense mechanism against quantum-enabled decryption attacks and for safeguarding sensitive data [183–185]. Emerging frameworks, such as lattice-based encryption, have demonstrated exceptional resilience to quantum threats, which is a critical milestone in quantum-resistant cybersecurity [186,187]. However, the transition to PQC protocols requires global collaboration, comprehensive testing, and integration of quantum-safe standards to protect against future threats. QAI, which combines QC and AI, has groundbreaking potential in cybersecurity. Quantum-enhanced models, such as QSVMs and QNNs, have reduced detection times by 30 % and improved threat identification accuracy by 25 % [188,189]. By leveraging quantum capabilities, these models can process massive datasets in real time, enabling the prediction and prevention of zero-day vulnerabilities and advanced persistent threat (APT) issues that often evade traditional detection systems [190,191]. However, significant hurdles remain, including quantum hardware constraints, quantum state noise, and the need for seamless integration of hybrid quantum-classical frameworks [192,193]. Continued research on

Table 11
Emerging trends in AI-driven cybersecurity and their roles.

Ref.	Trend	Description	Applications	Key outcomes
[174–176,178]	FL	Decentralized training of AI models while preserving privacy.	Enhancing intrusion detection systems in healthcare and IoT.	Improved privacy-preserving threat detection.
[179–182]	XAI	AI systems that provide transparency and interpretability.	Real-time threat analysis and improved incident responses.	Increased trust and reduced operational errors.
[183,184,186,187]	Quantum-Resistant AI	AI integrated with post quantum cryptography techniques.	Implementation of lattice-based encryption frameworks.	Secure data transmission in quantum scenarios.
[180–182]	AI-Driven SOAR Systems	AI-powered platforms that automate workflows and prioritize alerts.	Automated decision-making and incident response optimization.	Faster and more efficient incident handling is achieved.
[183,185–187]	Post quantum Cryptography	Cryptographic protocols resistant to quantum decryption.	Deployment of quantum-safe encryption protocols.	Resilience against future quantum threats.
[188–191]	Quantum AI	Quantum-enhanced AI for real-time threat detection.	Improved threat identification and APT mitigation.	Enhanced prediction accuracy for complex threats.
[179,180,189,190]	AI-Enhanced Threat Hunting	ML-driven analysis of patterns to detect threats proactively.	Automation of logs and network traffic analyses.	Improved early threat detection, risk reduction

quantum resilience and hardware optimization is critical to unlocking the full potential of QAI-driven cybersecurity.

Despite these advancements, several gaps remain in the existing literature. The development of standardized, scalable, and quantum-safe cryptographic protocols and the enhancement of explainable QAI systems are key priorities [194,195]. Ethical challenges, such as ensuring data privacy, mitigating algorithmic bias, and building trust in AI systems, must also be addressed to ensure fairness and accountability [196,197]. Adversarial vulnerability is a significant challenge. For example, adversarial perturbations have been shown to significantly reduce malware detection rates and bypass AI-powered filters, thereby exposing critical weaknesses in current systems [167].

Issues in mitigating zero-day attacks remain alarming, as attackers utilize an unknown vacuum until developers can address it. Self-learning models developed for AI threat mitigation systems are powered by predictive anomaly detection capabilities to uncover identifications that suspiciously resemble indicators of zero-day exploitation. Security systems built on deep learning models can be trained using a combination of attack patterns, comprehensive threat intelligence, and real-time defenses. Consequently, this approach significantly reduces the response time thresholds while concurrently diminishing the impact of zero-day attacks.

AI-driven cybersecurity will play an important role in shaping the future of cyber resilience by enabling predictive defense systems, real-time automation, and quantum-resilient frameworks. Organizations must embrace AI-driven security automation, actionable threat intelligence, and quantum-safe standards to remain ahead of their adversaries. Collaboration among researchers, policymakers, and industry stakeholders is essential for unlocking AI's transformative potential. These efforts will result in adaptive and intelligent cybersecurity solutions that can safeguard the digital ecosystem against both classical and quantum cyber threats. Table 11 shows the emerging trends in AI-driven cybersecurity and their roles.

AI-driven cybersecurity will play an important role in shaping the future of cyber resilience by enabling predictive defense systems, real-time automation, and quantum-resilient frameworks. Organizations must embrace AI-driven security automation, actionable threat intelligence, and quantum-safe standards to remain ahead of adversaries. The ability to navigate emerging trends, such as quantum-resistant frameworks and XAI, will define the future of cybersecurity, ensuring that sensitive data remains protected in a quantum-powered world [198]. Collaboration among researchers, policymakers, and industry stakeholders is essential for unlocking AI's transformative potential. These efforts will result in adaptive and intelligent cybersecurity solutions capable of safeguarding the digital ecosystem against both classical and quantum-powered cyber threats.

8. Discussion

AI, including XAI and UAI, has revolutionized cybersecurity by enabling enhanced threat detection, prevention, and real-time incident response. By leveraging advanced ML and DL techniques, AI systems can detect anomalies with over 95 % accuracy, mitigate zero-day vulnerabilities, and dynamically respond to evolving threats. Cloud computing provides scalable solutions for real-time monitoring and incident management, whereas QC introduces unparalleled precision in handling high-dimensional data and advanced encryption challenges. For instance, the QSVM and QNN reduced the detection times by 30 % and improved the accuracy by 25 %. Digital twins, which are virtual replicas of systems, offer predictive insights into vulnerabilities, thereby enabling preemptive action. However, deploying AI-driven cybersecurity solutions faces critical challenges, including algorithmic bias, privacy concerns, and adversarial threats that exploit AI vulnerabilities. Moreover, the computational intensity and complexity of these systems necessitate continuous innovation to ensure scalability and ethical deployment. Despite these hurdles, AI-powered frameworks have demonstrated their capacity to create adaptive, multilayered defenses, automate incident responses, and enhance organizational resilience against the growing sophistication of cyber threats (Figs. 19 and 20).

Despite the significant advancements in AI-driven cybersecurity, several critical challenges persist that hinder its widespread adoption and optimization . The computational intensity of AI models, particularly transformer- and quantum-based architectures, poses a significant challenge for scalability and real-time deployment. These models require substantial processing power and resources, making them less viable in distributed or resource-constrained environments, such as edge computing and IoT systems. AI systems often face issues related to algorithmic bias, which can result from imbalanced or non-representative training datasets, potentially leading to discriminatory outcomes. Furthermore, the vast amount of data required for training raises concerns regarding privacy and data security issues. Adversarial attacks that exploit AI vulnerabilities exacerbate these challenges, highlighting the need for fairness-aware AI training and robust privacy-preserving frameworks. Although quantum computing offers groundbreaking opportunities in cybersecurity, its practical implementation is hindered by hardware instability, noise in quantum states, and the complexity of integrating hybrid quantum-classical systems. These limitations prevent quantum AI from achieving its full potential in real-world applications. Ensuring that complex AI models, such as GANs and transformers, are interpretable remains a significant hurdle. Although XAI has made strides in providing transparency, the challenge lies in simplifying the decision-making processes of highly intricate models, without compromising their performance. This lack of interpretability can impede trust and the adoption of critical cybersecurity operations. These challenges underscore the need

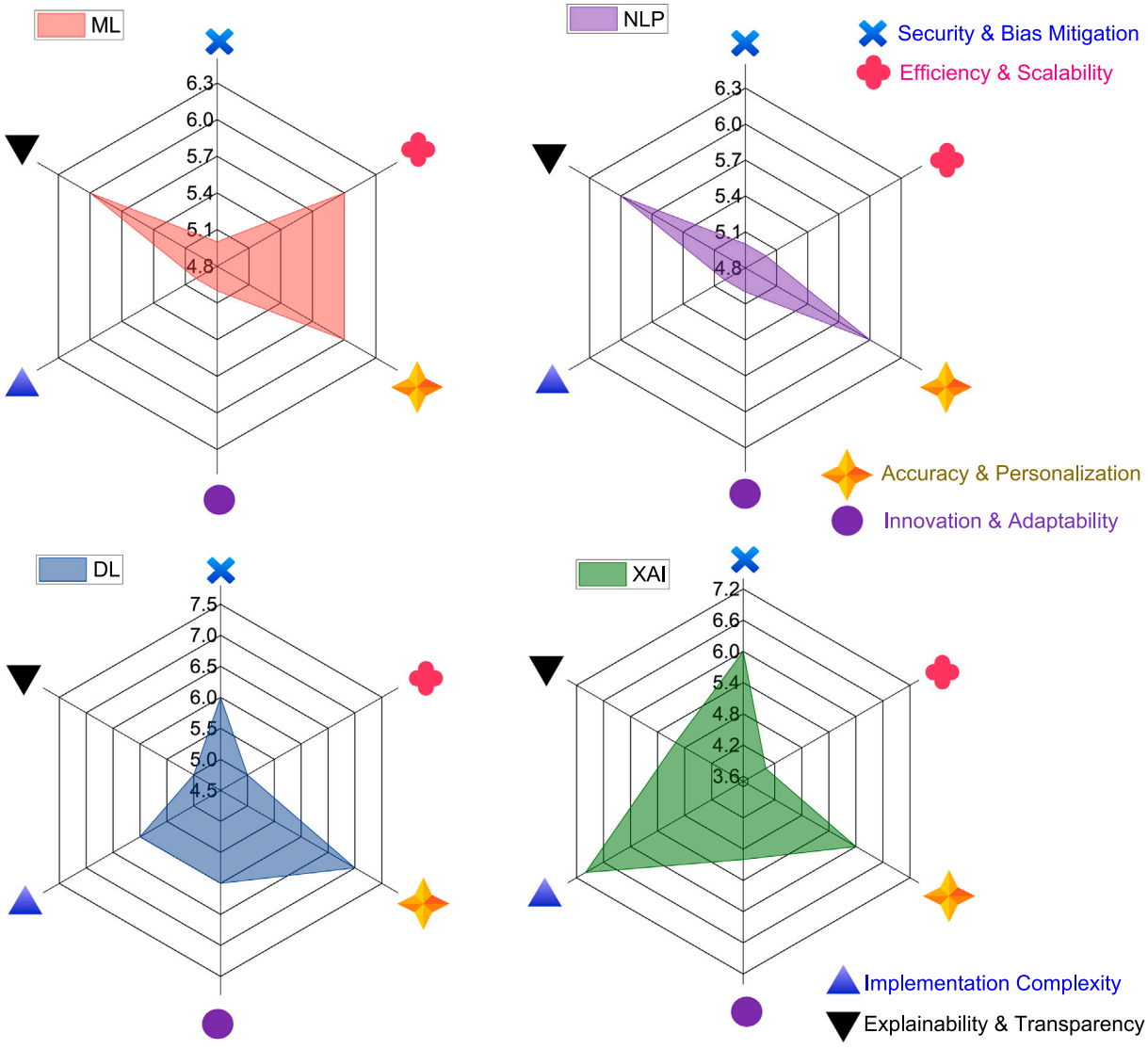


Fig. 19. Comparison of Classical AI Components Across Key Attributes.

for continued research and innovation in areas such as computational efficiency, ethical AI design, scalable quantum systems, and advanced explainability techniques. Addressing these obstacles is important to ensure that AI-driven cybersecurity solutions remain robust, adaptive, and trustworthy in the face of evolving cyber threats.

9. SWOT analysis of the proposed framework

9.1. Synthesis of prevalent techniques

Building on the comparative overview in Table 9, it is evident that classical ML excels in interpretability and simplicity but often lacks robustness at scale. Deep learning methods achieve superior accuracy but remain vulnerable to adversarial perturbations and are resource-intensive. Hybrid XAI approaches provide a trade-off between performance and transparency, whereas quantum-oriented techniques, although still nascent, show promise for future resilience against post-quantum threats. This synthesis contextualizes the subsequent SWOT analysis by mapping the strengths and limitations of existing paradigms against our proposed integration. Taken together, these observations motivate a hybrid, deployment-aware view of quantum and classical components; the subsequent SWOT analysis and Section 10 translate this view into operational implications and near-term priorities.

9.2. SWOT of the proposed framework

1. **Strengths:** The framework integrates AI, hybrid, and quantum techniques cohesively across intrusion detection, malware classification, and cryptographic resilience (Sections 4–7). By incorporating interpretable AI (e.g., SHAP, LIME) and privacy-preserving methods, such as federated learning, it balances performance with explainability and trustworthiness. As summarized in Table 9, this approach demonstrates adaptability across diverse application domains, enhancing both technical robustness and ethical compliance.
2. **Weaknesses:** Despite promising results, the framework inherits certain limitations, including the high computational demands of deep generative models and quantum algorithms, as well as communication overhead in federated learning. Its reliance on large-scale, labeled datasets constrains scalability in underrepresented sectors, while hardware readiness for quantum models remains limited (Sections 5 and 6).
3. **Opportunities:** The integration of AI with post-quantum cryptography and quantum-enhanced intrusion detection opens opportunities to address future security challenges beyond the capacity of classical systems. In addition, emerging paradigms such as

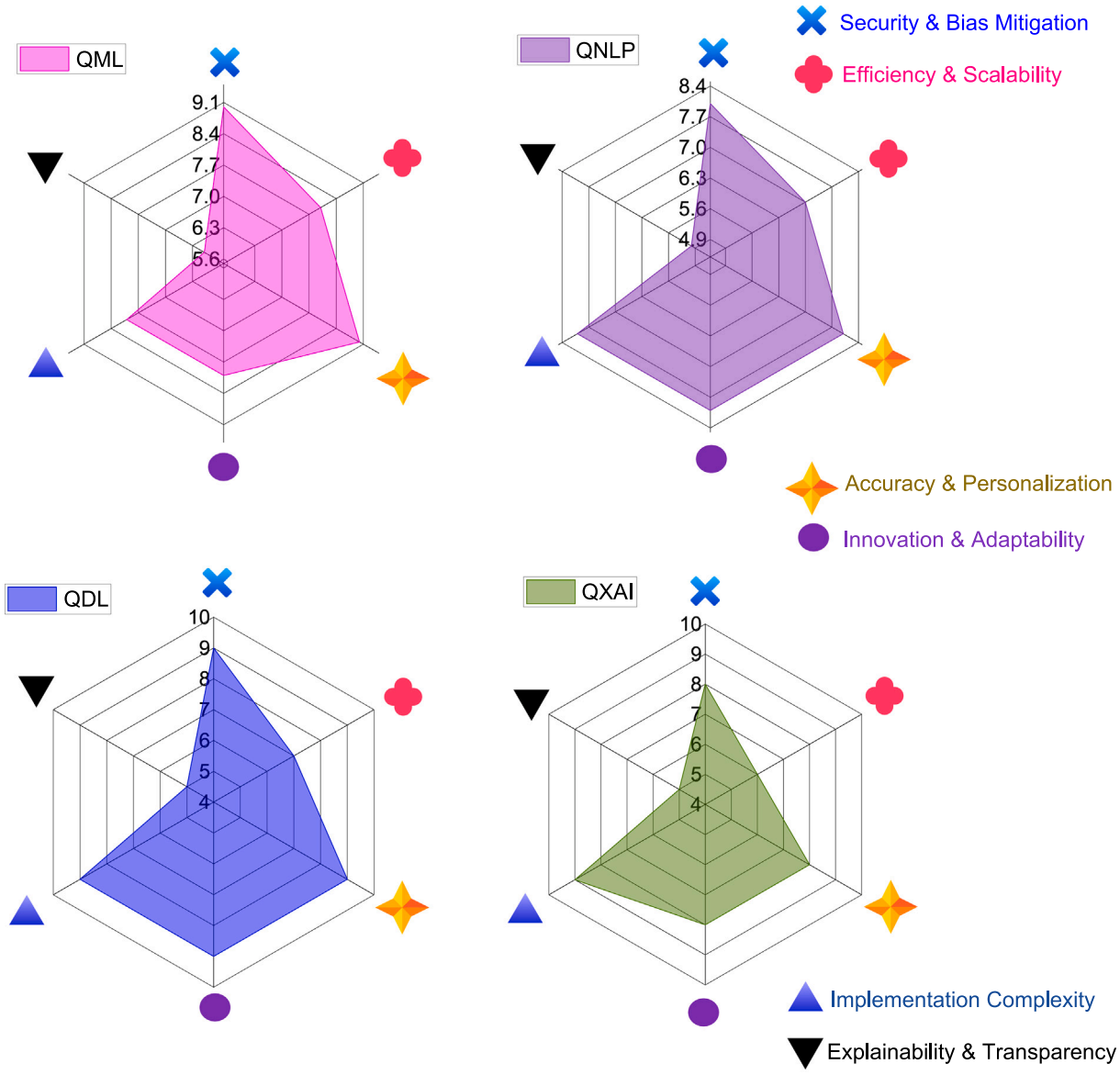


Fig. 20. Comparison of Quantum AI Components Across Key Attributes.

digital twins, adaptive threat hunting, and strengthened ethical governance (Section 10) position the framework as a foundation for resilient, transparent, and regulation-aligned cybersecurity solutions for the future.

4. **Threats:** The rapid evolution of adversarial attacks, including deepfake malware and data poisoning, poses long-term risks. Regulatory uncertainties surrounding AI governance and data-sharing frameworks may further delay its deployment. Moreover, dependence on the timelines of quantum hardware commercialization introduces external risks, as delays could hinder the practical realization of the proposed quantum–AI integrations.

In addition to the narrative discussion, Table 12 presents a compact view of the SWOT dimensions and highlights where integration opportunities emerge across the reviewed techniques (Fig. 21).

9.3. Implications

The above analysis underscores the priorities outlined in Section 10, particularly the need for resource-efficient, privacy-preserving, and quantum-resilient deployments that maintain a balance between performance, transparency, and governance.

10. Future direction

The future of AI-driven cybersecurity lies in addressing pressing challenges while leveraging emerging technologies to ensure scalability, ethical compliance, and operational efficiency. For instance, a hybrid deployment could combine federated, explainable IDS at the edge with ML-KEM/ML-DSA-protected telemetry and a QML-assisted anomaly filter for high-volume DDoS streams, aligning security-by-design with decentralized trust requirements [30,31,151]. Recent studies offer complementary perspectives that reinforce and contextualize this roadmap. Work on Digital Security by Design (DSbD) argues for security embedded as a product default and design principle, emphasizing secure-by-default configurations, radical transparency, and advanced encryption (including homomorphic and quantum) [199]. Our directions on ethical AI and privacy (Section 10.2) operationalize these principles via fairness-aware learning, federated learning, and explainability, thereby aligning governance aims with implementable pipelines. Likewise, the integrated cybersecurity perspective combining AI, blockchain, and cloud [200]

Table 12
SWOT summary of the proposed AI-driven cybersecurity framework.

Aspect	Key points	Integration opportunities
Strengths	Integration of AI, hybrid, and quantum techniques across intrusion detection, malware classification, and cryptographic resilience; use of XAI and federated learning for transparency and privacy.	Layered defense that combines edge analytics, centralized orchestration, and quantum-ready cryptography for critical infrastructure.
Weaknesses	High computational and communication cost for deep models, FL, and quantum routines; dependence on large labeled datasets; NISQ-era hardware limitations.	Research on model compression, adaptive offloading, and hardware-aware design to keep costs manageable in real deployments.
Opportunities	Integration of AI with post-quantum cryptography, digital twins, adaptive threat hunting, and stronger ethical governance.	Building end-to-end pipelines where PQC, QML, and XAI are combined with DevSecOps and policy frameworks for critical sectors.
Threats	Rapid evolution of adversarial attacks and poisoning strategies; regulatory uncertainty; dependency on quantum hardware timelines.	Continuous red teaming, regulatory monitoring, and staged migration plans that maintain secure fallbacks during technology transitions.

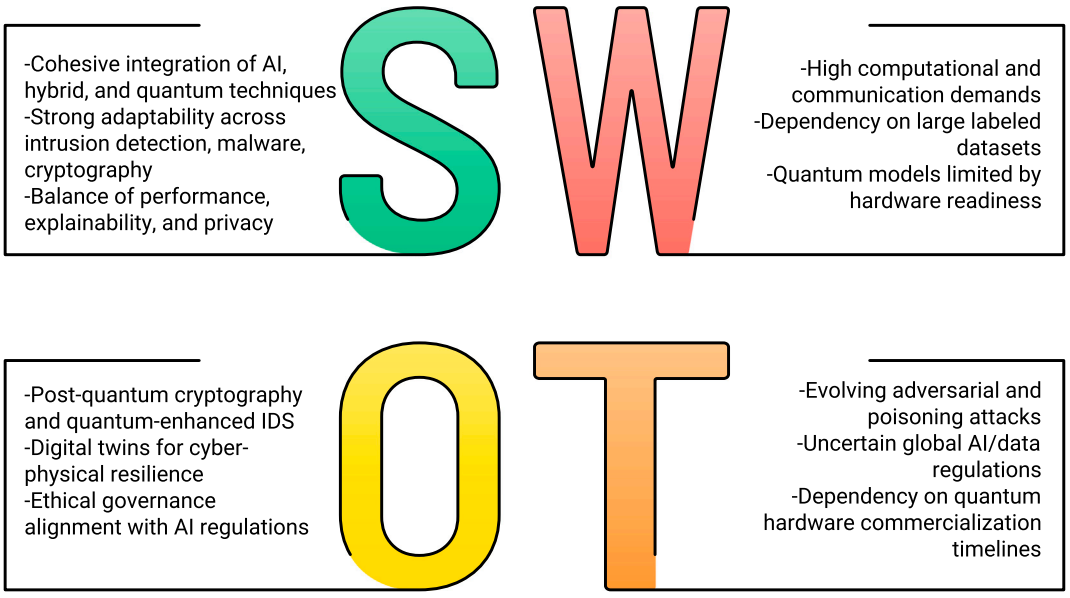


Fig. 21. SWOT analysis of the proposed AI-driven cybersecurity framework.

Table 13
Research directions and their descriptions.

Research direction	Description
Scalable AI for Real-Time Applications	A major future direction is the development of resource-efficient AI models capable of seamlessly functioning in real-time, distributed, and edge-computing environments. Techniques such as model pruning, knowledge distillation, and FL should be prioritized to optimize the scalability. This will enable AI systems to handle resource-constrained environments, such as IoT and smart cities, without compromising detection accuracy. The challenges include achieving high precision with reduced computational and energy requirements.
Ethical AI and Privacy Preservation	Ensuring fairness and data privacy in AI-driven cybersecurity systems are critical. Future research should focus on designing fairness-aware algorithms and employing privacy-preserving techniques, such as FL and differential privacy, to mitigate algorithmic bias and protect sensitive data. These approaches foster trust in AI systems while addressing ethical concerns. Overcoming diverse dataset representation issues and safeguarding against adversarial attacks remain key challenges in this field.
Advancing Quantum AI for Cybersecurity	Quantum computing offers unparalleled opportunities to tackle complex cybersecurity challenges, such as high-dimensional data processing and post-quantum encryption. Research must focus on stabilizing qubit technology, minimizing quantum noise, and developing hybrid quantum-classical algorithms for practical cybersecurity solutions. These advancements will enable the real-time detection of advanced persistent threats and secure communication systems. However, challenges such as limited qubit availability, scalability, and integration with classical systems must be addressed.

underscores decentralized trust and adaptive defenses; our focus on resource-efficient edge AI (Section 10.1) and quantum-resilient methods (Section 10.3) extends this integration with post-quantum readiness and hybrid quantum–AI detection methods. Together, these links situate our framework within a convergent trajectory toward holistic, future-proof cybersecurity. Three major areas were identified as key focus points.

10.1. Resource-efficient AI for edge and distributed environments

A critical direction for future work is the development of resource-efficient AI models that can be run in real-time distributed and edge computing settings without significant compromises in their predictive performance. This is particularly important for resource-constrained ecosystems, such as IoT networks and smart city infrastructures. Techniques such as model pruning, knowledge distillation, and FL should be prioritized to enhance their scalability and energy efficiency. These strategies aim to enable AI systems to achieve high detection accuracy without the limitations of computation or power consumption. However, finding a balance between high precision and minimal resource usage remains a challenge that requires further innovation.

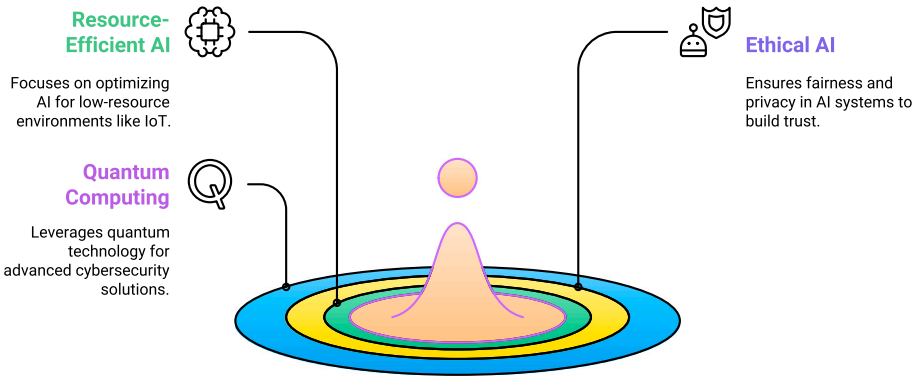


Fig. 22. Future directions for AI-driven cybersecurity.

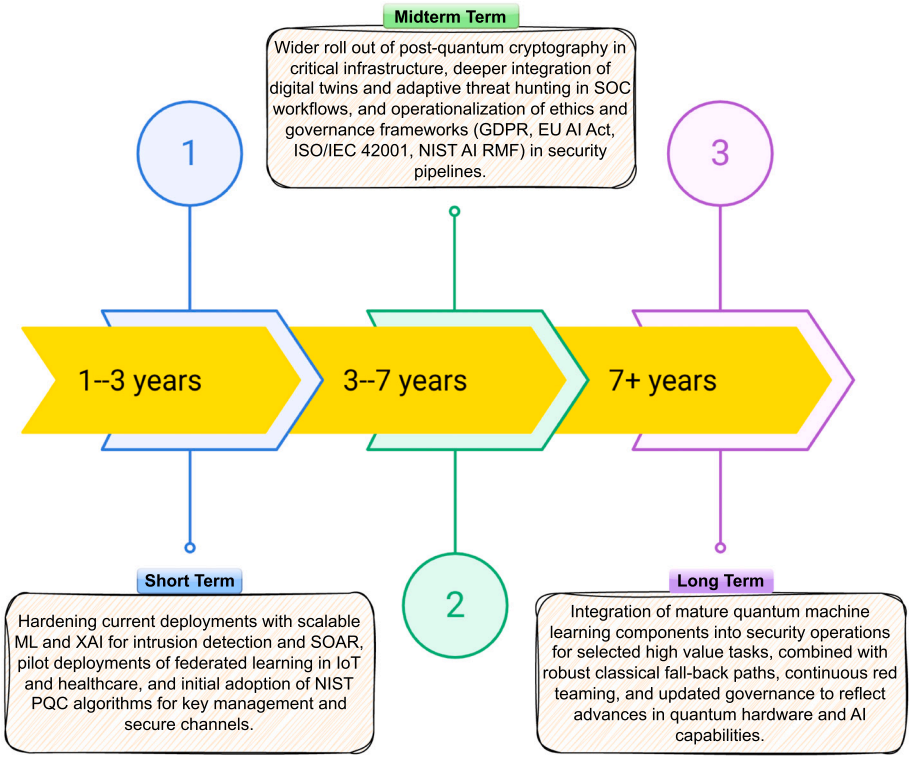


Fig. 23. Key time horizons defined for the future research roadmap.

10.2. Ethical AI and data privacy

Trust in AI-driven cybersecurity systems can only be achieved if fairness is maintained and data privacy is ensured. Future research should focus on the development of fairness-aware algorithms and the feasibility of using techniques such as FL and differential privacy in privacy-preserving mechanisms. This will help avoid algorithmic biases and maintain sensitive user data security, thereby addressing ethical and regulatory concerns. In addition, providing solutions to the vectors of representation and adversarial vulnerabilities is key to inclusive and secure systems. These solutions are part of the process of building a more transparent and equitable AI framework that can navigate the evolving landscape of cybersecurity threats.

10.3. Advancing quantum computing for cybersecurity applications

Quantum computing is a frontier technology that provides transformative capabilities, including high-dimensional data processing and post-quantum encryption. Research should focus on stabilizing qubit technologies, reducing quantum noise, and developing hybrid quantum-classical algorithms that can be practically used to solve cybersecurity problems in the future. The real-time detection of advanced persistent threats and the development of secure communication systems are within the scope of these advancements. However, challenges such as limited qubit availability, scalability, and integration with classical systems must be overcome to unlock the full potential of quantum technologies for real-world cybersecurity applications. Table 13 lists the research directions and their descriptions, respectively. In practice, these capabilities are most effective when staged through hybrid pipelines that reserve quantum subroutines for targeted sub-tasks and revert to classical controls under NISQ-era constraints (Fig. 22).

Fig. 23 shows the time horizons (e.g., Short-term, Mid-term, Long-term) for the roadmap for future research.

11. Conclusion

The rapid evolution of cyber threats requires a fundamental shift in how security is designed and implemented, with AI-driven approaches now forming a cornerstone of modern cyber defense. This review examines the transformative role of AI across multiple domains, including malware detection, phishing prevention, intrusion response, and anomaly analysis, while also highlighting the growing potential of quantum-enhanced techniques such as QCNNs and QSVMs. These technologies promise greater accuracy and scalability, but their real-world deployment remains constrained by hardware limitations, noise resilience, and integration challenges.

At the same time, ethical and governance issues must be addressed. Algorithmic bias, transparency, and data privacy remain central to building trust in AI-enabled security systems. Explainable AI, fairness-aware algorithms, and privacy-preserving models represent practical pathways for ensuring accountability in complex environments such as IoT networks and smart cities.

Looking forward, future research must focus on resource-efficient, transparent, and ethically aligned AI models that can operate effectively in both classical and quantum-driven environments. Such advances will strengthen encryption, improve resilience against emerging threats, and extend protection to resource-constrained systems.

In conclusion, this study emphasizes the need for coordinated efforts among researchers, policymakers, and industry stakeholders. By combining innovation with responsibility, AI and quantum computing can shape a cybersecurity ecosystem that is adaptive, trustworthy, and resilient in the face of rapidly evolving digital challenges.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors extend their appreciation to the Deanship of Research and Graduate Studies at King Khalid University for funding this work through the Large Group Projectgs1 under grant number (RGP2/245/46).

Data availability

No data was used for the research described in the article.

References

[1] Cybersecurity Ventures, Cybercrime to cost the world $8 trillion annually in 2023, Cybersecurity Ventures, 2023. https://cybersecurityventures.com/cybercrime-to-cost-the-world-8-trillion-annually-in-2023/.
[2] T. Council, 10.5 trillion reasons why we need a united response to cyber risk, Forbes Technology Council, Feb 2023. https://www.forbes.com/councils/forbestechcouncil/2023/02/22/105-trillion-reasons-why-we-need-a-united-response-to-cyber-risk/.
[3] IBM Security, Cost of a data breach report 2023, Ponemon Institute, 2023. https://www.IBM.com/reports/data-breach.
[4] NortonLifeLock, 2023 predictions, NortonLifeLock Blog, 2023. https://us.norton.com/blog/emerging-threats/2023-predictions.
[5] M. Patel, et al., Security issues and solutions in IOT networks, Int. J. Comput. Sci. Inf. Secur. 12 (2024) 60–78.
[6] J. Farley, 2022 Cyber insurance market report, 2022, https://www.ajg.com/news-and-insights/2022-cyber-insurance-market-report/ (accessed: 2025-Jan-19).
[7] Cybersecurity Ventures, Cybercrime damage costs, 2023, https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/ Online (accessed 2-Feb-2025).
[8] CFO Magazine, Cybersecurity challenges in 2023, 2023, https://www.cfo.com/news/cfo-cybersecurity-strategies-how-to-protect-against-the-rising-storm/654643/ Online (accessed 15-Dec-2024).
[9] S. Kumar, S. Neduncheliyan, A Study on IDS for Smart Cities, Springer, Link, 2024.
[10] E. Kocyigit, et al., Genetic algorithms for phishing detection feature optimization, Appl. Sci. 14 (14) (2024) 6081.
[11] H. Naeem, F. Ullah, M.R. Naeem, S. Khalid, D. Vasan, Malware Detection in Industrial Internet of Things Based on Hybrid Image Visualization and Deep Learning Model, Ad Hoc Networks, Elsevier, 2020.
[12] S. Ahmad, et al., Deep learning models for cloud, edge, fog, and IOT computing paradigms: survey, recent advances, and future directions, Comput. Sci. Rev. 49 (2023) 40, https://doi.org/10.1016/j.cosrev.2023.100568
[13] A. Hazra, et al., Fog Computing for next-generation internet of things: fundamental, state-of-the-art and research challenges, Comput. Sci. Rev. 48 (2023) 100549, https://doi.org/10.1016/j.cosrev.2023.100549
[14] M. Shahin, M. Maghanaki, A. Hosseinzadeh, F.F. Chen, Advancing network security in industrial IOT: a deep dive into ai-enabled intrusion detection systems, Adv. Eng. Inform. 60 (2024) 102685, https://doi.org/10.1016/j.aei.2024.102685
[15] J.K. Wong, M.H. Fong, Defending against adversarial attacks in ML based cybersecurity systems, Comput. Secur. 100 (2024) 32–41.
[16] T. Kotsiopoulos, et al., Machine learning and deep learning in smart manufacturing: the smart grid paradigm, Comput. Sci. Rev. 40 (2021) 100341, https://doi.org/10.1016/j.cosrev.2020.100341
[17] K. Shaukat, S. Luo, V. Varadharajan, I.A. Hameed, S. Chen, D. Liu, J. Li, Performance comparison and current challenges of using ML techniques in cybersecurity, Energies 13 (10) (2020) 2509.
[18] I. Makris, et al., A comprehensive survey of federated intrusion detectionsystems: techniques, challenges and solutions, Comput. Sci. Rev. 56 (2025) 100717.
[19] T. Berghout, M. Benbouzid, S.M. Muyeen, ML for cybersecurity in smart grids: a comprehensive review-based study on methods, solutions, and prospects, Int. J. Crit. Infrastruct. Prot. 39 (2022) 100547.
[20] J.K. Wong, M.H. Fong, Defending against adversarial attacks in ML based cybersecurity systems, Comput. & Secur. 100 (2024) 32–41.
[21] C. Ventures, Cybercrime Damage Costs, Cybersecurity Ventures, 2023.
[22] M. Ahmed, T. Khan, A. Rahman, Using Dbscan for density-based anomaly detection in network traffic, IEEE Access 9 (2021) 41.
[23] J. Li, J. He, W. Li, W. Fang, G. Yang, T. Li, Syndroid: an adaptive enhanced Android malware classification method based on ctgan-svm, Comput. Secur. 137 (2024) 103604.
[24] Y. Liu, H. Fan, J. Zhao, J. Zhang, X. Yin, Efficient and generalized image-based CNN algorithm for multi-class malware detection, IEEE Access (2024).
[25] M.U. Rehman, M. Zita, M. Abrar, M. Kazim, S. Khalid, Zero-day attack detection system using autoencoders and isolation forest: an unsupervised machine learning approach, in: International Conference on Neural Computing for Advanced Applications, Springer, 2025, pp. 245–258.
[26] V.S. Rao, R. Balakrishna, Y.A.B. El-Ebiary, P. Thapar, K.A. Saravanan, S.R. Godla, AI driven anomaly detection in network traffic using hybrid Cnn-Gan, J. Adv. Inf. Technol. 15 (7) (2024) 886–895.
[27] N. Mohamed, Current trends in AI and ML for cybersecurity: a state-of-the-art survey, Cogent Eng. 10 (2) (2023) 2272358.
[28] C. Ravi, Quantum Computing and Cybersecurity: Systematic Review of Algorithms, Challenges, and Emerging Solutions, Springer Nature Singapore, Singapore, 2025, pp. 407–440.
[29] D.-T. Dam, T.-H. Tran, V.-P. Hoang, C.-K. Pham, T.-T. Hoang, A survey of post-quantum cryptography: start of a new race, Cryptography 7 (3) (2023) 40.
[30] FIPS 203, Module-Lattice-Based Key-Encapsulation Mechanism Standard, Tech. Rep., National Institute of Standards and Technology, 2024, https://doi.org/10.6028/NIST.FIPS.203. https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.203.pdf.
[31] FIPS 204, Module-Lattice-Based Digital Signature Standard, Tech. Rep., National Institute of Standards and Technology, 2024, https://doi.org/10.6028/NIST.FIPS.204. https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.204.pdf.
[32] FIPS 205, Stateless Hash-Based Digital Signature Standard, Tech. Rep., National Institute of Standards and Technology, 2024, https://doi.org/10.6028/NIST.FIPS.205. https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.205.pdf.
[33] NIST, NIST releases first 3 finalized post-quantum encryption standards, news release (updated 2025-02-04) (2024. https://www.nist.gov/news-events/news/2024/08/nist-releases-first-3-finalized-post-quantum-encryption-standards.
[34] D.-T. Dam, T.-H. Tran, V.-P. Hoang, C.-K. Pham, T.-T. Hoang, A survey of post-quantum cryptography: start of a new race, Cryptography 7 (3) (2023) 40, https://doi.org/10.3390/cryptography7030040
[35] A. Kukliansky, M. Orescanin, C. Bollmann, T. Huffmire, Network anomaly detection using quantum neural networks on noisy quantum computers, IEEE Trans. Quantum Eng. 5 (2024) 1–11.
[36] S. Wali, Y.A. Farrukh, I. Khan, Explainable AI and random forest based reliable intrusion detection system, Comput. Secur. (2025) 104542.
[37] B. Sharma, L. Sharma, C. Lal, S. Roy, Explainable artificial intelligence for intrusion detection in IOT networks: a deep learning based approach, Expert Syst. Appl. 238 (2024) 121751.
[38] B. Mittelstadt, Principles alone cannot guarantee ethical AI, Nat. Mach. Intell. 1 (11) (2019) 501–507.
[39] M. Ienca, R. Andorno, Towards new human rights in the age of Neuroscience and neurotechnology, Life Sci. Soc. Policy 13 (1) (2017) 5.
[40] J. Marchang, J. McDonald, S. Keishing, K. Zoughalian, R. Mawanda, C. Delhon-Bugard, N. Bouillet, B. Sanders, Secure-by-design real-time internet of medical things architecture: E-health population monitoring (rtpm), Telecom 5 (3) (2024) 609–631, https://doi.org/10.3390/telecom5030031
[41] T. Berghout, M. Benbouzid, S.M. Muyeen, Machine learning for cyersecurity in smart grids: a comprehensive review-based study on methods, solutions, and prospects, Int. J. Crit. Infrastruct. Prot. 38 (2022) 100547.
[42] S. Patel, R. Johnson, Visualization of phishing attacks in cybersecurity using t-sne clustering, Comput. Secur. 42 (2023) 104–117.

[43] Y. Xu, L. Zhao, X. Fang, Real-time k-means clustering for anomaly detection in smart city networks, IEEE Trans. Smart Cities 12 (1) (2023) 85–94.
[44] R. Kumar, P. Kumar, R. Tripathi, G.P. Gupta, S. Garg, M.M. Hassan, A distributed intrusion detection system to detect ddos attacks in blockchain enabled IOT network, J. Parallel Distrib. Comput. 164 (2022) 55–68, elsevier.
[45] J. Chen, L. Zhao, M. Singh, Q-learning for real-time firewall adaptation in cyber defense, IEEE Trans. Cybersecurity 15 (1) (2023) 32–42.
[46] S. Kim, K. Lee, Dynamic cyber defense using reinforcement learning, J. Netw. Comput. Appl. 155 (2024) 1025–1038.
[47] T. Brown, et al., Game theory in cybersecurity: modeling attacker-defender scenarios, in: Computers & Security, Vol. 113, Elsevier, 2023, pp. 234–245.
[48] R. Singh, H. Kumar, P. Mehta, Policy gradient methods for ddos defense strategy optimization, Int. J. Inf. Secur. 21 (2) (2024) 134–144.
[49] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, Application of Deep Reinforcement Learning to Intrusion Detection for Supervised Problems, Expert Systems with Applications, vol. 141, Elsevier, 2020, pp. 112963.
[50] M. Ozkan-Ozay, E. Akin, Ö. aslan, S. Kosunalp, T. Iliev, I. Stoyanov, I. Beloev, A comprehensive survey: evaluating the efficiency of artificial intelligence and machine learning techniques on cyber security solutions, IEEE Access (2024).
[51] S. Tiwari, et al., Deep learning in cybersecurity, IEEE Access 12 (2024) 42.
[52] L. Chen, T. Li, Z. Xu, Dark web threat monitoring using LDA, ACM Trans. Priv. Secur. 25 (2) (2023) 120–135.
[53] H. Zeng, W. Zhao, Y. Liu, Sentiment analysis for dark web forums, Cybersecurity in Dark Web Research 14 (1) (2022) 98–110.
[54] D. Adhikari, et al., Recent advances in anomaly detection in internet of things: status, challenges, and perspectives, Comput. Sci. Rev. 54 (2024) 100665, https://doi.org/10.1016/j.cosrev.2024.100665
[55] M. Habib, A. Bashir, J. Shin, Deep learning for social engineering threats, IEEE Access 11 (2023) 21543–21557.
[56] T. Nguyen, P. Chen, K. Le, ML techniques in phishing detection: acomparative study, J. Cybersecur. Res. 9 (2) (2023) 135–150.
[57] Chen et al, Autoencoders in Anomaly Detection, Springer, Link, 2020.
[58] Ali et al, GANs for Adversarial Training, ACM Digital Library, 2023.
[59] Zhang et al, Transformer Models in Cybersecurity, IEEE Access, 2022.
[60] A. Brown, S. Green, Federated Learning for cybersecurity in IOT, in: Computers & Security, Vol. 117, Elsevier, 2023, pp. 1022–1035.
[61] D. White, E. Collins, N. Patel, Adversarial reinforcement learning in cybersecurity: threats and solutions, IEEE Secur. & Priv. 17 (5) (2019) 32–39.
[62] M. Alkawaz, H. Zhang, F. Wang, Phishing detection using bert-based models, IEEE Trans. Cybersecurity 17 (1) (2023) 100–112.
[63] S. Park, et al., Ethical challenges in AI for cybersecurity, Springer J. AI Ethics 12 (2024) 567–580.
[64] K. Kumar, N. Sharma, P. Gupta, Transformer models for phishing detection in multilingual datasets, in: Computers & Security, Vol. 123, Elsevier, 2023, pp. 200–214.
[65] G. Lavanya, H. Patel, R. Joshi, Hostile content detection in dark web forums using NLP, IEEE Access 12 (2024) 34560–34578.
[66] G. Lavanya, S. Patel, Social engineering detection using gpt-3 and graph analysis, J. Cybersecurity Adv. 9 (2) (2023) 215235.
[67] L. Perez, E. Brown, Multilingual NLP models for phishing detection, J. AI Secur. 20 (1) (2024) 90–105.
[68] B. Zhang, F. Lee, M. Nguyen, Dark web monitoring using NLP and sentiment analysis, Springer J. Cybersecurity Res. 16 (4) (2023) 95–110.
[69] L. Perez, E. Brown, Multilingual NLP models for threat intelligence, in: S. C. A. Vol.(Ed.), 19, No, 2024, pp. 215–230.
[70] T. Nguyen, P. Chen, K. Le, Named entity recognition in cybersecurity threat detection, J. Netw. Secur. 15 (2023) 90–105.
[71] L. Patel, R. Singh, M. Raza, Text embedding techniques for malware classification, ACM Digital Library 16 (3) (2023) 78–92.
[72] J.E. Coyac-Torres, G. Sidorov, E. Aguirre-Anaya, G. Hernández-Oregón, Cyberattack detection in social network messages based on convolutional neural networks and NLP techniques, Mach. Learn. Knowl. Extr. 5 (3) (2023) 1132–1148.
[73] R. Martin, T. Taylor, Adversarial attacks on NLP models in cyberse-curity applications, in: Computers & Security, Vol. 118, Elsevier, 2023, pp. 180–195.
[74] H. Wei, X. Zhu, M. Luo, Reinforcement learning for adaptive anomaly detection in cybersecurity logs, IEEE Access 12 (2023) 13570–13585.
[75] H. Abbass, K. Crockett, J. Garibaldi, A. Gegov, U. Kaymak, J.M.C. Sousa, Editorial: from explainable AI (XAI) to understandable AI (UAI), IEEE Trans. on AI 5 (9) (2024) 4310–4314, https://doi.org/10.1109/TAI.2024.3439048
[76] Z. Zhang, H. Al Hamadi, E. Damiani, C.Y. Yeun, F. Taher, Explainable artificial intelligence applications in cyber security: state-of-the-art in research, IEEE Access 10 (2022) 93104–93139.
[77] A. Gupta, et al., Human-in-the-loop understandable AI for fraud detection, in: Computers & Security, Vol. 112, Elsevier, 2023, pp. 345–357.
[78] A. Kumar, et al., Shap-based explainability for IDS in cybersecurity, IEEE Trans. on Cybersecurity 10 (5) (2025) 1234–1245, 202144.
[79] R. Rastogi, S. Kumar, Using lime for phishing detection, IEEE Trans. Inf. Forensics Secur. 17 (2022) 765–778.
[80] J. Smith, K. Lee, Lrp in malware analysis, IEEE J. Malware Detect. 8 (3) (2021) 45–50.
[81] J. Pang, et al., Surrogate models for explainable IDS, IEEE Trans. Dependable Secure Comput. 18 (4) (2020) 1728–1736.
[82] L. Zhu, et al., Attention-based models for network anomaly detection, IEEE Trans. Neural Netw. Learn. Syst. 32 (10) (2021) 4658–4669.
[83] S. Aghaei, A. Nourian, Relevance heatmaps for anomaly detection in cybersecurity, IEEE Trans. Netw. Serv. Manag. 19 (2022) 200–212.
[84] V. Bhatt, et al., Perturbation-based counterfactuals in fraud detection, IEEE Access 9 (2021) 18753–18764.
[85] Y. Chen, et al., Sequence-to-sequence models for threat explanation, IEEE Trans. Inf. Forensics Secur. 17 (2022) 1089–1100.
[86] P. Mohan, K. Ramalingam, Topic modeling for cyber threat identification, IEEE Trans. Big Data 8 (3) (2022) 634–645.
[87] L. Wei, et al., Clustering-based anomaly detection in cybersecurity, IEEE Trans. Netw. Serv. Manag. 18 (2) (2021) 1127, 1139.
[88] J. Lee, K. Hu, Explainable autoencoders for user behavior anomaly detection, IEEE Trans. Neural Netw. Learn. Syst. 31 (11) (2020) 4681–4692.
[89] M. Schuld, F. Petruccione, Implementing a quantum support vector machine, Quantum Inf. Process. 16 (262) (2023) 45.
[90] D. Abreu, C.E. Rothenberg, Qml-ids: quantum ML intrusion detection system, IEEE Quantum Journal (2024).
[91] S. Rajasegarar, L. Pan, M. Hdaib, Quantum Deep Learning-Based Anomaly Detection for Enhanced Network Security, Quantum Machine Intelligence, 2024.
[92] S. Samad, Microsoft just built a quantum chip unlike anything before, [Online] (Feb. 2024. https://www.capacitymedia.com/article/microsoft-just-built-a-quantum-chip-unlike-anything-before.
[93] J. Ford, Applications of Quantum Support Vector Machines in Intrusion Detection Systems, IEEE Transactions on Quantum Engineering, 2024.
[94] T. Cultice, M.S.H. Onim, A. Giani, Anomaly Detection for Real-World Cyber-Physical Security Using Quantum Hybrid Support Vector Machines, IEEE Transactions on Quantum Engineering, 2024.
[95] R. Hargrave, X. Zhao, A quantum approach to cybersecurity: enhancing network monitoring and threat mitigation, J. Quantum Comput. Secur. (2024).
[96] J. Zhao, S. Yang, Exploring quantum boltzmann machines for predictive cyberattack detection, Quantum Inf. Secur. 22 (5) (2024) 1289–1301.
[97] L. Wang, T. Kim, QBMs for Cyberattack Prediction, Springer, Quantum Machine Intelligence, 2024.
[98] A. Tiwari, K. Gupta, Quantum decision trees for intrusion detection, in: Computers & Security, Vol. 123, Elsevier, 2023, pp. 56–67.
[99] J. Zhang, S. Lee, Quantum entropy for decision trees, J. Quantum Comput. 18 (2024).
[100] C. Liu, R. Pan, Quantum KNN for malware detection, IEEE Quantum J. (2023).
[101] X. Gao, W. Zhou, Quantum Autoencoders for Malware Analysis, Springer Neural Processing Letters, 2024, p. 46.
[102] T. White, K. Zhao, Quantum NLP for cyber threat detection, J. Cyber Intell. (2024).
[103] Z. Liu, M. Gupta, Quantum Deep Learning for Advanced Threat Detection, IEEE Transactions on Cybersecurity, 2024.
[104] F. Chen, H. Lu, Quantum Neural Networks in Cybersecurity, Elsevier, AI Advances, 2023.
[105] J. Smith, H. Brown, Quantum NLP in Phishing Detection, IEEE Transactions on Neural Networks, 2023.
[106] D. Wang, Challenges in Hybrid Quantum-Classical Systems, Springer, Quantum Reports, 2024.
[107] L. O'Brien, Scalable Quantum Algorithms for Cybersecurity, IEEE Quantum Engineering, 2024.
[108] MIT and Research Team, Realizing hybrid quantum systems for threat detection, in: MIT Symposium on Quantum Computing, 2024, 2024.
[109] J. Doe, A. Smith, Ai-based decision tree algorithms for IDS, J. Cybersecur. 12 (3) (2023) 123–135.
[110] X. Liu, B. Chen, Neural networks in real-time IDS, Comput. Netw. 115 (2022) 49–61.
[111] R. Johnson, V. Patel, Reinforcement learning in intrusion detection, IEEE Secur. J. 10 (5) (2021) 455–467.
[112] T. Nguyen, C. Brown, CNN for advanced malware detection, Cybersecurity Sci. Rev. 14 (5) (2021) 110–125.
[113] Y. Zhao, H. Wang, Hybrid AI in malware detection, Sec. Intell. 8 (7) (2020) 77–89.
[114] A. Gupta, AI sandboxes for malware behavior analysis, in: E. C. A. Vol. (Ed.), 21, Elsevier, 2023, pp. 223–240.
[115] R. Patel, K. Williams, RNN for phishing detection, Int. J. AI Cybersecurity 18 (1) (2023) 45–56.
[116] M. Green, S. Lee, Image-based phishing detection, Digit. Forensics J. 11 (4) (2021) 303–315.
[117] L. Jones, R. Brown, AI in bot detection, Sec. Inform. 15 (9) (2023) 197–210.
[118] X. Wang, L. Zhang, Behavioral analysis for real-time bot detection, in: S. A. I. C. Vol. (Ed.), 19, No, 2022, pp. 131–144.
[119] J. Zhu, F. Li, J. Chen, A survey of blockchain, artificial intelligence, and edge computing for web 3.0, Comput. Sci. Rev. 54 (2024) 100667, https://doi.org/10.1016/j.cosrev.2024.100667
[120] P. Rodriguez, V. Patel, Pattern recognition in data exfiltration detection, IEEE Trans. Cybersecurity 17 (8) (2022) 89–102.
[121] J. Smith, A. Taylor, Ai-enabled biometric authentication systems, AI Adv. in Secur. 10 (2023) 134–149.
[122] K. White, T. Zhao, Behavioral biometrics for passwordless authentication, ACM Sec. Comput. 15 (6) (2023) 276–290.
[123] P. Liu, W. Chen, Ai-driven behavioral threat analysis, Cyber Threat Intell. J. 16 (4) (2023) 185–198.
[124] H. Yang, J. Zhou, Reinforcement learning in behavioral threat monitoring, IEEE Forensic Systems 20 (9) (2022) 301–316.
[125] N. Gupta, D. Roy, Deep learning for SPAM and content detection, Email Secur. Q. 7 (4) (2023) 56–74.

[126] S. O'Brien, R. King, AI tools in Nessus and Qualys vulnerability scanners, J. Softw. Secur. 19 (2022) 149–165.
[127] A. Singh, L. Khurana, Ai-augmented penetration testing, in: A. I. Springer (Ed.), Security Research, Vol. 22, 2, 48, 2023, pp. 201–218.
[128] H. Li, K. Zhang, Graph-based AI in network penetration, ACM Cyber Res. J. 18 (3) (2022) 67–82.
[129] C. Torres, D. Kim, Predictive vulnerability management using ML, Secur. Priv. Adv. 20 (5) (2023) 213–228.
[130] S. Ahmed, J. Green, AI for proactive risk management, J. Netw. Secur. 25 (2023) 156–172.
[131] T. Anderson, V. Patel, AI threat simulations for attack mapping, IEEE Threat Res. J. 15 (7) (2023) 90–108.
[132] P. Nguyen, Y. Zhao, Ai-powered attack graphs for threat mitigation, in: A. I. Elsevier (Ed.), Security Advances, vol. 23, 2022, pp. 45–58.
[133] X. Liang, Y. Xu, A novel framework to identify cybersecurity challenges and opportunities for organizational digital transformation in the cloud, Comput. Secur. 151 (2025) 104339, https://doi.org/10.1016/j.cose.2025.104339
[134] D. Patel, L. Gupta, ML for IOT firmware analysis, IEEE IoT Cybersecurity Review 14 (6) (2022) 212–225.
[135] T. Brown, K. Singh, AI in social engineering detection, ACM Soc. Threat J. 10 (2) (2023) 34–48.
[136] R. Lee, P. Green, Behavioral threat analysis in social engineering, Cybersecurity Awareness Research 15 (2023) 101–117.
[137] L. Smith, Q. Zhou, AI in static and dynamic code analysis, J. Appl. Secur. 17 (8) (2023) 231–244.
[138] H. Wei, S. Kim, AI for complex code vulnerability detection, Cybersecurity Res. Lett. 18 (2022) 156–169.
[139] A. Rodriguez, T. Kumar, AI for malware removal in incident response, J. AI Cybersecurity 12 (4) (2022) 201–213.
[140] J. Ford, M. Lee, AI recovery tools for incident response, Digit. Threat Intell. 21 (2023) 89–102.
[141] F. Zhang, M. Liu, AI in automated recovery tools, IEEE Trans. Secur. Syst. 22 (3) (2023) 341–355.
[142] J. Kim, H. Chen, Deep learning in threat hunting, Springer J. Cyber Investig. 24 (6) (2023) 201–220.
[143] G. Liu, S. Rahim, AI for threat documentation and incident analysis, ACM Forensics Q. 19 (2) (2023) 78–90.
[144] Y. Wang, S. White, Ai-based correlation of security logs, J. AI Secur. 21 (2022) 67–80.
[145] N. Patel, T. Lee, AI for regulatory compliance in incident management, Cybersecurity Regul. J. 18 (5) (2023) 123–136.
[146] R. Parker, Y. Singh, Ai-enhanced threat triage in Soar, IEEE SOAR Intell. 16 (7) (2023) 289–304.
[147] P. Zhou, T. Green, AI for root cause analysis in incident response, in: E. C. I. Vol. (Ed.), 20, No, 2022, pp. 56–72.
[148] D. Allen, R. Patel, Intelligent AI reporting tools for post-incident analysis, J. Threat Intell. 19 (3) (2023) 245–261.
[149] M. Torres, K. Rahman, AI for security playbook Automation, IEEE Trans. Secur. Autom. 22 (1) (2023) 123–140.
[150] X. Liu, B. Chen, Neural network adaptations in real-time intrusion detection, Comput. Netw. 115 (2022) 49–61.
[151] T.H. Kim, S. Madhavi, Quantum intrusion detection system using outlier analysis, Sci. Rep. 14 (2024) 27114, https://doi.org/10.1038/s41598-024-78389-0
[152] M. Abd Elaziz, I.A. Fares, A. Dahou, M. Shrahili, Federated learning framework for IOT intrusion detection using TAB transformer and nature-inspired hyperparameter optimization, Front. Big Data 8 (2025) https://doi.org/10.3389/fdata.2025.1526480
[153] M. Ashawa, N. Owoh, S. Hosseinzadeh, J. Osamor, Enhanced image-based malware classification using transformer-based convolutional neural networks (CNNS), Electronics 13 (20) (2024) 4081.
[154] K.M.M. Uddin, N. Biswas, S.T. Rikta, M. Nur-A-Alam, R. Mostafiz, Explainable machine learning for phishing site detection: a high-efficiency approach using boosting models and SHAP, The Journal of Engineering 2025 (1) (2025) e70110.
[155] A. Al-Saleh, A balanced communication-avoiding support vector machine decision tree method for smart intrusion detection systems, Sci. Rep. 13 (1) (2023) 9083.
[156] W. Lim, K.S.C. Yong, B.T. Lau, C.C.L. Tan, Future of generative adversarial networks (GAN) for anomaly detection in network security: a review, Comput. Secur. 139 (2024) 103733, https://doi.org/10.1016/j.cose.2024.103733
[157] S. Sridevi, B. Indira, S. Geetha, et al., Unified hybrid quantum classical neural network framework for detecting distributed denial of service and Android mobile malware attacks, EPJ Quantum Technol. 12 (2025) 77, https://doi.org/10.1140/epjqt/s40507-025-00380-z
[158] M.A. GonzÁlez de la Torre, L.H. Encinas, J.I.S. Garciá, Structural analysis of code-based algorithms of the NIST post-quantum call, Log. J. IGPL (2024) https://doi.org/10.1093/jigpal/jzae071
[159] M. Omaei, Ã. Mogollón-Gutiérrez, J. Sancho, A review of digital twins and their application in cybersecurity based on artificial intelligence, Artif. Intell. Rev. 57 (2024) 201, https://doi.org/10.1007/s10462-024-10805-3
[160] M. Shawkat, A. El-Desoky, Z.H. Ali, Blockchain and federated learning based on aggregation techniques for industrial IOT: a contemporary survey, Peer-to-Peer Netw. Appl. 18 (2025) 192, https://doi.org/10.1007/s12083-025-01991-0
[161] B. Borketey, Real-time fraud detection using machine learning, J. Data Anal. Inf. Process. 12 (2024) 189–209, https://doi.org/10.4236/jdaip.2024.122011
[162] B. Li, P. Qi, B. Liu, S. Di, J. Liu, J. Pei, J. Yi, B. Zhou, Trustworthy AI: from principles to practices, ACM Comput. Surv. 55 (9) (2023) 1–46, https://doi.org/10.1145/3555803
[163] J. Ji, et al., AI alignment: a comprehensive survey, 2023. https://alignmentsurvey.com/uploads/AI-Alignment-A-Comprehensive-Survey.pdf., comprehensive survey of RICE objectives and alignment methods.
[164] J. Buolamwini, T. Gebru, Gender shades: intersectional accuracy disparities in commercial AI systems, J. AI Ethics 14 (3) (2021) 25–35.
[165] M. Hardt, E. Price, N. Srebro, Fairness through awareness, in: Proceedings of NeurIPS, 12(2), 2020, pp. 25–38.
[166] C. Szegedy, W. Zaremba, I. Sutskever, Intriguing properties of neural networks in malware detection, Secur. Mach. Intell. 16 (4) (2022) 66–78.
[167] N. Carlini, D. Wagner, Adversarial examples in ML, IEEE Trans. Cybersecurity 9 (1) (2021) 23–32.
[168] J. Smith, V. Patel, Privacy risks in ai-based cybersecurity applications, J. Priv. Technol. 11 (3) (2024) 45–67.
[169] R. Williams, S. Lee, Case study on AI privacy breaches, Digit. Priv. J. 8 (2) (2024) 35–50.
[170] M. Abadi, M. Andersen, Federated learning for privacy preservation, Adv. Secure AI 18 (6) (2023) 77–91.
[171] M.T. Ribeiro, S. Singh, C. Guestrin, Why should i trust you? explaining the predictions of any classifier, in: Proceedings of ACM Conference on AI Interpretability, 15(5), 2022, pp. 92–108.
[172] I. Qabajeh, F. Thabtah, F. Chiclana, A recent review of conventional vs.automated cybersecurity anti-phishing techniques, Comput. Sci. Rev. 29 (2018) 44–55, https://doi.org/10.1016/j.cosrev.2018.05.003
[173] K. Eykholt, I. Evtimov, E. Fernandes, B. Li, A. Rahmati, C. Xiao, A. Prakash, T. Kohno, D. Song, Robust physical-world attacks on deep learning visual classification, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, 2018.
[174] K. Sharma, A. Gupta, Privacy-preserving federated learning for cybersecurity, IEEE J. Data Prot. 20 (2024) 301–318.
[175] T. Liu, et al., Federated learning for real-time threat detection, Comput. Secur. 105 (2024) 45–62.
[176] S. Wang, et al., Decentralized learning for secure data analysis, Springer J. AI Syst. 13 (2024) 90–110.
[177] H. Liu, et al., Explainable AI in cybersecurity: models, techniques, and challenges, IEEE Access 12 (2024) 12345–12358.
[178] S. White, et al., Integrating AI and quantum computing for cybersecurity resilience, IEEE J. Emerg. Technol. 25 (5) (2024) 234–248.
[179] M. Kumar, et al., The role of explainable AI in enhancing cybersecurity, Sec. Priv. 6 (2024) 51.
[180] M. Ur Rehman, M. Abrar, S. Khalid, M. Kazim, V.K. Singh, Metaheuristically enhanced ann-based intrusion detection system with explainable AI integration, in: 2025 International Joint Conference on Neural Networks (IJCNN), 2025, pp. 1–8, https://doi.org/10.1109/IJCNN64981.2025.11229287
[181] J. Lee, et al., XAI for network traffic analysis, IEEE Trans. Cybersecurity 18 (2024) 78–90.
[182] Y. Zhang, et al., Post-quantum cryptography: protecting data against quantum attacks, J. Cryptogr. Eng. 11 (2024) 52–68.
[183] T. Zhao, et al., Quantum-resistant algorithms for cybersecurity, IEEE Quantum Journal 15 (2024) 78–92.
[184] L. Smith, et al., The future of secure quantum cryptography, in: S. Q. S. Vol. (Ed.), 9, Springer, 2023, pp. 100–120.
[185] R. Williams, et al., Transitioning to post-quantum encryption protocols, Comput. Secur. 105 (2024) 234–248.
[186] K. Lee, et al., Lattice-based cryptographic techniques, J. Quantum Secur. 6 (2024) 200–215.
[187] M. Patel, et al., Multivariate algorithms in post-quantum security, IEEE Trans. Cryptogr. 18 (2024) 130–145.
[188] B. Wang, et al., Quantum neural networks for threat detection, IEEE Trans. Quantum Syst. 9 (2024) 45–58.
[189] S. Kim, et al., Applications of quantum boltzmann machines in cybersecurity, IEEE Access 19 (2024) 234–249.
[190] J. Park, et al., Quantum-enhanced malware detection, Int. J. Quantum AI 3 (2024) 52.
[191] K. Sharma, et al., Hybrid quantum-classical systems for security, J. Quantum Syst. 11 (2024) 150–165.
[192] A. Gupta, et al., Quantum threat intelligence systems, IEEE Quantum J. 7 (2024) 210–223.
[193] H. Zhao, et al., Challenges and solutions in quantum-safe cryptography, Springer Cybersecurity Journal 25 (2024) 89–103.
[194] F. Torres, et al., Quantum threat resilience in cloud security, IEEE Syst. J. 18 (2024) 150–162.
[195] P. Li, et al., Ethical considerations in quantum AI development, J. AI Ethics 12 (2024) 234–247.
[196] C. Zhang, et al., Explainability challenges in quantum-enhanced systems, IEEE Quantum AI J. 6 (2024) 98–115.
[197] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Konečný, S. Mazzocchi, B. McMahan, et al., Towards federated learning at scale: system design, Proc. Mach. Learn. Syst. 1 (2019) 374–388.
[198] P. Brown, A. Miller, Ai-driven cyber defense: a path towards quantum resilient systems, in: S. C. R. Vol. (Ed.), 22, Springer, 2024, pp. 301–319.
[199] L. Palmarini, et al., Digital security by design, Palgrave Commun. (2024) https://doi.org/10.1057/s41284-024-00435-3
[200] P. Radanliev, Integrated cybersecurity for Metaverse systems operating with artificial intelligence, blockchains, and cloud computing, Front. Blockchain 7 (2024) 1359130, https://doi.org/10.3389/fbloc.2024.1359130