Paradigm International Journal of Mathematics and Computer in Engineering 4(1) (2026) 11–32

International Journal of Mathematics and Computer in Engineering

https://reference-global.com/journal/IJMCE

**Review Study**

# The future of web application security: Opportunities and challenges for machine learning-based techniques

Bolanle Eunice Oduleye[1]†

[1]Department of Computer Engineering, Faculty of Engineering, University of Uyo, Uyo, 1017, Nigeria

Communicated by Haci Mehmet Baskonus; Received: 16.04.2024; Accepted: 28.11.2024; Online: 02.02.2026

## Abstract

In this paper, the future prospects and potential for machine learning-based solutions in web application security are examined. These include enhancing accuracy and efficiency, real-time detection, scaling, explainable Artificial Intelligence (AI), adversarial machine learning, and automated response. The article also gives a general overview of Machine Learning (ML) methods that are frequently applied to web application security, including supervised learning methods like decision trees, Support Vector Machines (SVMs), and neural networks, unsupervised learning methods like k-means clustering and Principal Component Analysis (PCA), and deep learning methods such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). The incorporation of machine learning-based approaches into security measures will be more necessary as online applications continue to develop in order to meet new threats.

**Keywords:** Supervised and unsupervised learning, real-time detection, scalability, explainable AI, adversarial ML, automated response.
**AMS 2020 codes:** 68M25.

## 1 Introduction

Threats to cybersecurity have increased as a result of the expansion of web applications. Web application security has been improved by ML-based approaches, however there are still issues that need to be resolved. Web apps have permeated every aspect of our lives in the digital era. These cutting-edge programs offer a wide range of services that we depend on every day, from social networking to e-commerce and online banking [1–3]. But as online apps expand, so do cybersecurity risks that might jeopardize user data, protected information, and even national security [4, 5]. According to Genge et al. [6], Tao et al. [7] and Alawida et al. [8], traditional security solutions struggle to keep up with the sophistication and frequency of cyberattacks. According to Cui et al. [9], Al-Garadi et al. [10], Nguyen and Reddi [11], Venketeswaran [12], Antoniadi et al. [13] and Macas and Fuertes [14] by enabling increased detection and response capabilities, ML-based solutions have shown considerable potential for increasing online application security. These methods are able to evaluate enormous volumes of data in real-time, spot hazards, and respond appropriately to lessen them [15]. The need to improve accuracy and efficiency, real-time detection, scalability, explainable AI, and adversarial ML are

†Corresponding author.
Email address: bolanle_ayomi@yahoo.com

Paradigm
ISSN 2956-7068 doi:10.2478/ijmce-2026-0002

just a few of the issues that still need to be resolved despite the potential of these technologies. With each passing year, the use of web-based apps has been rising quickly. There were roughly 367 million domain names representing static or dynamic web apps as of Q1 2020, while the precise number of web applications in existence worldwide is unclear [16]. These programs might transmit and handle private user information, which makes them a tempting target for nefarious attackers [17]. Web apps are becoming more and more important, both publicly and privately. Unfortunately, more than 90% of these systems lack security, and there are on average 13 flaws per software [18, 19]. Therefore, security is essential for guaranteeing that online applications may be used safely [5]. Web-based software has established itself as a helpful platform for a variety of enterprises, giving them the chance to enhance operations, save money, and expedite procedures [15]. Web applications have become a crucial component of the productive sector as a result of their deployment in the business and communication sectors [20–22]. Millions of organizations have started using the internet as a cost-effective platform for customer transactions, knowledge exchange, and communication during the past year. Advertisers may engage with potential consumers and offer tailored browsing experiences, product information request forms, and email subscriptions by using web-based services. With more than 3.88 billion users globally, the internet and other online service providers have emerged as key components of contemporary automation of traditional jobs [23]. However, breaches in Information Technology (IT) security have caused serious issues for people, organizations, corporations, governments, and nations [4, 5]. Information is lost every day due to cyber dangers, and millions of dollars are taken every year [24]. Although research on online vulnerabilities and cyber threats has been done, there is always a need for fresh approaches to risk mitigation and defense against malware, hackers, and other damaging assaults [25–27]. Evaluating ML models requires a comprehensive set of metrics to fully understand their performance. Relying solely on accuracy can be misleading, especially when dealing with imbalanced datasets where security threats are unevenly distributed. Precision measures how accurately the model identifies positive security alerts, focusing on minimizing false positives. This is crucial for ensuring that security interventions are targeted and effective, without overwhelming users with irrelevant alerts [28]. Recall assesses the model's ability to detect all actual threats within the dataset, emphasizing the importance of identifying every potential security issue, even if it results in some false positives [29]. The F1-score integrates precision and recall into a single metric, providing a balanced view that is particularly valuable in datasets with skewed class distributions. It ensures that both false positives and missed threats are accounted for in the assessment [30]. Area Under the Receiver Operating Characteristic Curve (AUC-ROC) evaluates the model's capability to differentiate between legitimate and malicious activities across various thresholds, offering insight into its overall effectiveness [31]. Together, these metrics offer a comprehensive view of ML-based web application security systems, highlighting their strengths and areas for improvement beyond simple accuracy. While many studies have been undertaken on the future of web application security and the application of ML methods, there are still some significant gaps in prior work and surveys. These research gaps identify areas in which further study is required to expand our understanding and implementation of ML-based approaches in online application security. Among the significant research gaps are:

1. Limited focus on adversarial attacks
2. Insufficient consideration of real-time detection
3. Limited evaluation on real-world datasets
4. Lack of interpretability and explainability
5. Inadequate consideration of ethical implications.

This article examines the potential applications of ML-based approaches in online application security as well as their future possibilities. It gives a general introduction to supervised, unsupervised, and deep learning algorithms—all of which are frequently used ML approaches in online application security. The essay also

looks at the difficulties of incorporating ML-based security solutions, such as the requirement for AI that can be explained and the possibility of adversarial assaults. The incorporation of ML-based approaches into security measures will become more and more necessary as online applications continue to develop in order to meet new threats. The study explores the future potential of ML in web application security, examining advancements in detection and scalability, the role of Explainable AI, challenges of Adversarial ML, and the importance of automated responses. It also emphasizes collaboration between data scientists and cybersecurity experts to address evolving cyber threats. ML is crucial for enhancing web application security as cyber threats grow more complex. Traditional methods are no longer sufficient, prompting the adoption of ML to predict and counteract attacks, optimise network performance, and ensure secure online services. Web applications, accessible via browsers, provide interactive functionalities. Tools like Sublime Text and Burp Suite help developers create and secure these applications, protecting them from vulnerabilities. Network protocols like TCP/IP and DNS facilitate data exchange but can be exploited by attackers. Understanding and securing these protocols is essential to preventing breaches. In 2022, cybercrime escalated, particularly with ransomware and social engineering attacks targeting businesses and blockchain projects. The rise in these attacks highlights the need for advanced security measures. ML-based web application security uses algorithms like decision trees and neural networks to detect and prevent attacks, offering robust defenses. However, these methods face challenges, such as vulnerability to adversarial attacks and the need for continuous adaptation. Future advancements may focus on improving accuracy, real-time detection, and integrating ML with existing security measures, ensuring more effective protection against evolving threats.

## 2 ML and web application security

The expansion of the internet has been considerably aided by the development of companies, technology, and intelligent applications [32]. As a result, in the field of cyber-human intelligence, computer networks and networking technologies are now crucial for both consumers and businesses in the delivery of information and online needs. However, as networks have grown in popularity, their security and Quality of Experience (QoE) have become more important [5, 33]. Industries increasingly want networks that are efficient, dependable, and capable of providing secure services [34, 35]. Despite these objectives, networks frequently experience limitations because of a combination of greater traffic demand and higher computing needs. Furthermore, due to the complexity of the interactions on the network and the sophistication of cyber criminals, identifying and stopping network assaults has grown more challenging but is still crucial [36, 37]. Every network has particular characteristics, organizational needs, and performance standards that are dynamic [15]. Therefore, it has become difficult to design strong methods and architectures for managing complicated scenarios for diverse networking-related use cases, especially those dealing with security challenges [38, 39]. Although ML techniques have been used in the networking industry, it is still unclear exactly what function (if any) ML may play in the networking industry. The constructed statistical approaches used in earlier networking research, which depended entirely on known port numbers to discover desired patterns in various datasets, turned out to be comparatively unproductive [40–42].

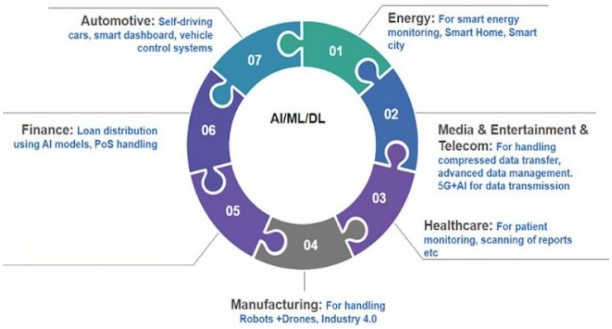


**Fig. 1** ML roles in various industry [41].

Because enabling networks are essential to the smooth operation of both enterprises and consumers, the use of ML in networking is still a pertinent and difficult problem. Many firms complain about not having enough staff to manage network assaults, and some small businesses fail to budget for network security, leaving them with antiquated systems that are more vulnerable to attack [4]. Furthermore, given that they may account for up to 25% of a nation's total internet traffic, attacks like Distributed Denial of Service (DDoS) continue to pose a danger to the majority of service providers. Monitoring network traffic is more important than ever to guarantee that customer data is safe from nefarious attackers. It is advantageous and cost-effective for all sorts of businesses to use solutions that can profit from the special capacity of ML to achieve greater optimization to address security challenges in their intrusion detection and network monitoring [5, 15]. Combining ML-based techniques with the expanding complexity and size of computer and mobile networks may spur academics and practitioners to make more investments in clever approaches to managing network security [43]. As a result, further study is required on the use of ML in networking, particularly in the context of network security. This will assist businesses in creating effective strategies and architectural frameworks for handling challenging network conditions. The advantages and uses of integrating ML in the networking area have been covered in several papers and articles. While some articles cover the process and potential for ML in networking, some articles concentrate on how specific ML algorithms may address various networking difficulties. The benefits of applying ML to networking, including improved decision-making, more generalized models, and approximated models of correctness [44]. Wang et al. [44], also provides a typical ML in networking workflow, albeit this workflow may need to take into account roadblocks to completely implementing ML in networking as shown in Figure 1. Research already conducted in this field mostly concentrate on lowering measurement costs, which is insufficient for solving networking concerns. According to Bhutani [45], networks can adapt to their environments and make judgments while still learning about them with the use of ML. The article also discusses some of the difficulties faced by wireless networks, including delivering the best QoE while working with constrained resources and inconsistent signal quality. For wireless networks, Bhutani [45] recommends a number of well-liked ML techniques, including Artificial Neural Networks (ANNs), Naive Bayes, and Logistic Regression. Finding high-quality datasets for networking is still difficult since the effectiveness of each ML approach in terms of resource friendliness is not usually taken into account in most studies. Usama et al. [46], concentrate on the use of unsupervised ML techniques in networking, claiming that unsupervised ML may successfully handle a variety of networking difficulties. Unsupervised ML approaches, which do not require labeled data, may handle unstructured data and give greater flexibility. Examples include anomaly detection, categorizing internet traffic, networking optimization, and analytic. A specific use of ML in networking is discussed in Fang et al. [47], specifically for intrusion detection. The authors implemented ML for intrusion detection using SVM and Elman Neural Network approaches. According to the results, their installation was very accurate, reduced the number of false alarms, and enhanced the capacity to recognize unusual network activity. By concentrating on the layers of the network protocol stack, Kulin et al. [48] study investigates the application of ML in networking. Insights into how ML may be used to forecast patterns and trends at various

network levels, such as Signal-to-Noise Ratio (SNR), latency, and energy usage, are provided by this technological method. Those who are not ML specialists can benefit much from the article's essential foundational information.

### 2.1 Web application

A web application is a software program that is accessible via a network using a web browser. It gives consumers functionality and interactive elements that allow them to complete activities or obtain information [16, 17] (see Figure 2). There are several technologies available for developing and securing web apps. The following are some regularly used tools: Sublime Text, Atom, Eclipse, Visual Studio Code, NetBeans, Notepad++, Postman, Burp Suite, ZAP OWASP. These tools let developers write code, test functionality, debug, manage versions, and ensure security measures are in place to safeguard web applications from possible threats and vulnerabilities. Developers may construct sturdy and secure online applications for customers to enjoy by skillfully leveraging these capabilities.

**Table 1** Comparing ML and web application security based on the tools, algorithms, applications, and types.

| Comparison | ML | Web Application Security |
|---|---|---|
| Tools | Scikit-learn, TensorFlow PyTorch, Keras, XGBoost, LightGBM, Cae, Theano, RapidMiner, KNIME, Weka | Web Application Firewalls (WAFs), ModSecurity, OWASP Zap, Acunetix, Burp Suite, Qualys, Nessus, OWASP Dependency Check Nikto, AppSpider, Nexpose, Vega, Wireshark, Zed Attack Proxy |
| Algorithms | Support Vector Machines (SVM), Random Forest (RF), Neural Networks (NN), Naive Bayes (NB), K-Nearest Neighbors (KNN), Decision Tress, Gradient Boosting Ma chine (GBM) | Cross-Site Scripting (XSS) Pre vention, SQL Injection Detection, Anomaly Detection, Genetic Algo rithms for security optimization |
| Application | Natural Language Pro cessing (NLP), Image Recognition, Fraud Detection, Predictivee Maintenance, Recom mender Systems, Speech Recognition, Time series Forecasting. | Secure Coding, Authentication, Ses sion Management, Input Validation, API Security, DDoS Detection and Mitigation, Intelligent Bot Protec tion, Code Review (Static and Dy namic Analysis), Web application security scanning (SAST, DAST), Cryptographic Implementation for Data Protection. |
| Type | Supervised Learning, Unsupervised Learning, Reinforcement Learn ing, Semi-Supervised Learning, Meta-Learning, Self-Supervised Learning | Preventive Measures, Detect and Respond, Secure Software Devel opment Lifecycle (SDLC), Incident Response, Continuous Monitoring, Threat Modeling and Simulation, Content Security Policy (CSP) Im-plementation |

Table 1 compares ML and web application security in terms of tools, methods, applications, and kinds.

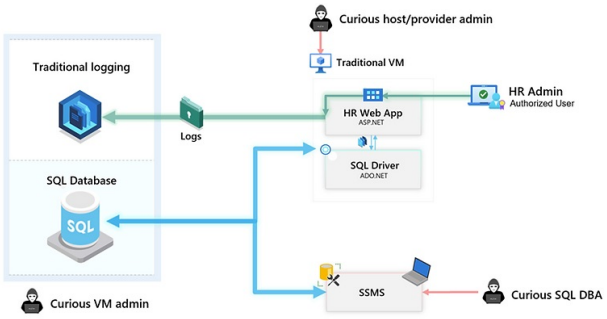


**Fig. 2** Web security typical architecture [16].

Popular tools in ML include Scikit-learn, TensorFlow, PyTorch, and Keras. ML algorithms include SVM, Random Forest, Neural Networks, and Naive Bayes. Web Application Firewalls (WAFs), Penetration Testing technologies, and Vulnerability Scanners are popular web application security technologies. Web application security methods address particular risks such as Cross-Site Scripting (XSS) protection, SQL Injection detection, and anomaly detection.

ML is used in a variety of fields, including Natural Language Processing (NLP), Image Recognition, and Fraud Detection. Web application security, on the other hand, includes features such as secure code, authentication, session management, and input validation. ML is classified into three types: supervised learning, unsupervised learning, and reinforcement learning. Web application security, on the other hand, involves preventative measures, detection and response methods, and the incorporation of security into the Software Development Lifecycle (SDLC).

## 3 Network protocols and their vulnerabilities

The exchange of information and communication between computers on a network depends on network protocols [26]. Because network assaults frequently target certain elements of these protocols, such as the network layers, hardware, and others, it is essential to take into account the many kinds of network protocols [27]. A protocol is a collection of guidelines that governs several network features, such as access techniques, topologies, and more [25]. Table 2 shows the overview of Network protocols and their various functions.

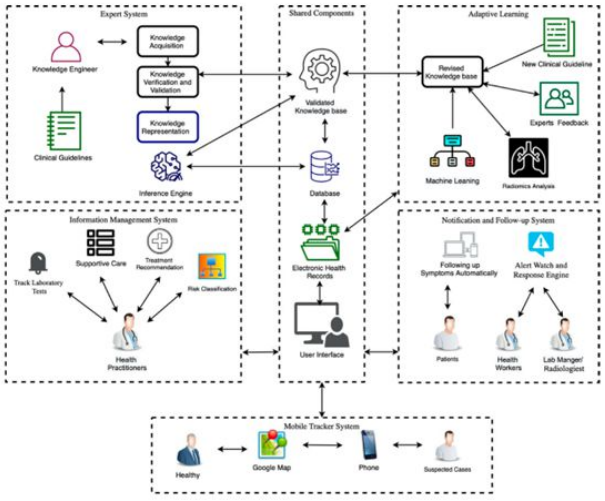


**Fig. 3** A brief overview of different types of networking attacks.

**Table 2** Overview of network protocols and their functions.

| Protocol | Level | Description |
|---|---|---|
| TCP/IP | Application, Transport, Internet, Network, Access | A protocol with four levels that is still in use. The Network Access level serves as the physical layer. The Internet Level, which contains several protocols including the Internet Protocol (IP), is in charge of transferring data for the whole network. The most used protocol for connecting devices to the internet, IPv4, has an addressing system that can only hold lit tle more than four million addresses. IPv6 can store a greater number of addresses [49]. |
| ICMP | Internet | Gives information about network issues and is used for network diagnostics [27]. |
| ARP | Internet | Employed by IPv4, it converts addresses with a specific number of bits into addresses with a larger num ber of bits. Functions at the OSI Models layers two and three [27]. |
| DNS | | Converts domain names into IP addresses so that web browsers may access resources. Frequently neglected when it comes to the infrastructure of network secu rity, and out-of-date defenses leave it open to assaults [25]. |
| BGP | | Transmits data over networks and enables routers to learn about IP address blocks and choose eective routes. Has not changed much over time and is open to abuse. Version 4 of BGP, the most recent release, was created in 2006 [4, 50]. |

Figure 3 provides a brief overview of different types of networking attacks. The image displays various attacks from left to right, which includes Syn Flood, Smurfing, UPD Flood, Internet Control Message Protocol (ICMP) Flood, Slowloris, NTP, HTTP Flood, and Hammer-Anvil. Networking attacks have considerably changed over time. Malicious software and malware for network assaults have really existed since the 1960s [51]. Large systems at universities were frequently the target of attackers in the 1970s who employed trojans to carry out their tricks and assaults. The Data Encryption Standard initially received approval in 1976. The frequency of incidents that became more widely publicized throughout the 1980s increased. The first DDoS assault took place in November 1988, garnering international attention [52, 53]. Many different networking attacks still take place today, and they focus on particular network protocol layers. Depending on the size and intensity of the assault, these expenditures might go into the millions of dollars for corporations. An overview of the various networking threats is presented in Figure 3 to aid in understanding them.

Businesses must be aware of these assaults and take precautions to safeguard their networks against them. Although the Syn Flood attack has been present since the 1980s, it first received widespread media notice in 1996 [53]. The TCP/IP protocol is used in this attack to interfere with network activities. In order to establish a connection with a server, a client typically sends a message, which the server must first acknowledge. In contrast, a syn flood attack bombards the server with messages, using up all of the network resources and blocking genuine traffic from reaching it. The device may respond slowly or not at all as a result of this assault, which will affect the performance of the entire network. While network ports are typically the main focus of a Syn flood attack, it may also be used to support other attacks. Smurfing is a different kind of hack that takes use of

broadcast networks and the ICMP [54]. Network managers can communicate about the condition of the network and ping other nodes to learn about their network status using the ICMP protocol. Every host on a broadcast network receives ping queries, which result in an automated response. Attackers can exploit this weakness by fabricating a bogus request with a faked source IP address, which is the address of the target server [55]. The request is subsequently relayed to all other hosts through the broadcast network, each of which responds to the faked source address with an ICMP message. It is possible to bring down the target server with enough ICMP answers. Network administrators and IT experts should be aware of these kinds of assaults and take the necessary precautions to avoid them. To identify and lessen such assaults, they might include firewalls, intrusion detection systems, and other security mechanisms [56]. Regular security audits and testing can also assist in locating weaknesses and ensuring that security precautions are successful in thwarting online attacks [4, 5]. An online service is rendered inaccessible to users during a DDoS assault, which stands for DDoS. Such assaults may be facilitated by Syn Flood assaults. DDoS assaults may be divided into several kinds, such as protocol-based attacks, volume-based attacks, and application layer attacks [57]. In application layer assaults, the attacker produces protocol packets that cause the server to have too many open sessions and run out of resources. While protocol-based attacks take use of flaws in certain protocols, like TCP, volume-based assaults entail flooding the target device with a tremendous amount of network data. One thing to keep in mind with DDoS assaults is that slower attacks are more difficult to notice with popular detection systems since they mimic ordinary network traffic [58]. UDP flooding, ICMP flooding, Ping of Death, Slowloris, NTP amplifying assault, HTTP flooding, and zero-day DDoS attacks are a few examples of typical DDoS attacks. UDP flooding is saturating haphazard network gates on a particular host with data-carrying IP packets. As a result, the system becomes overloaded and is unable to communicate. On the other side, ICMP flooding involves bombarding a victim computer with pings, which can have a detrimental effect on a network's incoming and outgoing channels. Robert Snake Hanson created the Slowloris attack, which is a straightforward DDoS assault that uses very little bandwidth [58]. Through the use of incomplete requests, it establishes several connections to the target host and maintains them open for a long time. Infected servers may produce even more connections, reaching capacity and blocking legitimate connection requests. This assault is challenging for conventional intrusion detection systems to identify because it uses incomplete packets. Another type of DDoS assault, known as NTP Amplification, takes use of the NTP protocol by delivering UDP traffic to a host. Because it requires obtaining a response from a server to a fake IP address, this kind of attack is referred to as a reflection attack. This leads in an increase in network traffic transfer and subpar performance for legitimate interactions [58]. Attacks using HTTP Flood pose a problem since they are hard to detect and virtually indistinguishable to regular network traffic [59]. By bombarding a victim with HTTP requests, these attacks generally target the application layer (layer 7) of the OSI framework. Typically, botnets are utilized to increase the attack's effect. Zero-day DDoS assaults take use of recently identified security flaws, and there are presently no effective defenses in place to stop them [58]. The use of cutting-edge deep learning techniques to defend against such security assaults seems promising. Cross-layer attacks target different network layers. A jamming device and a hacked endpoint are used in the assault known as Hammer and Anvil to obstruct communication. Modern wireless networks are the subject of MAC poisoning, which slows the node's performance. TCP-timeout causes a node's TCP flows to be interrupted by sending high-priority, brief breaches, which puts the target in a time-out state. In general, network attacks, particularly DDoS attacks, are still common, but ML, especially deep learning, shows promise in reducing these attacks [58]. The development of efficient methods for identifying security assaults across many domains is still being researched.

## 4 2022 cyber crime attacks: Businesses, governments and blockchain

We ascribe cyber crime attacks to rising cyber-tensions as seen in Figure 4. The growth of the cybercrime sector is also having a significant influence, as hackers continue to expand their unlawful operations. Attackers can utilize compromised user information to carry out harmful operations as a result of large-scale data dumps.

For the same reasons, the number of assaults will increase much more in future.

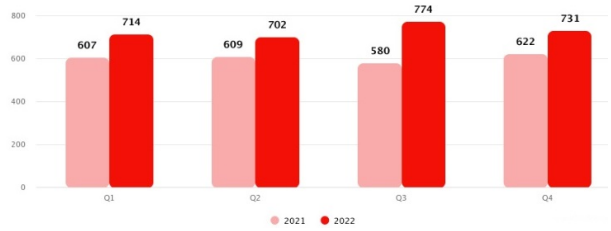


**Fig. 4** Number of incidents in 2021 and 2022 (by quarter) [60].

The percentage of events affecting corporate online resources climbed from 17% to 22% in 2021. Government institutions were affected the hardest, with the number of successful assaults on their websites more than tripling. Large-scale data breaches dominated the last year, with countless allegations of compromised data belonging to various corporations and their consumers. In 47% of cases, attackers were effective in stealing sensitive information from companies, while 64% were successful in stealing personal information from individuals. The use of spyware is increasing, especially in assaults against people. By the end of 2022, these sorts of malware had been utilized in half of all successful user assaults. Ransomware accounted for 51% of malware used in organizational assaults and is continually changing. More organizations began rewriting malware in cross-platform languages or developing versions that target both Windows and Linux computers in 2022. In 2022, we saw an increase in the use of data-wiping malware, some of which was disguised as ransomware.

Social engineering is still quite effective, accounting for 43% of successful assaults against organizations and 93% of successful attacks against individuals. The broad adoption of the phishing-as-a-service concept has led to the method's rising ubiquity. Attackers are increasingly using social media and messaging platforms to target people, while successful assaults on the second factor of authentication have occurred in situations involving corporations. Attacks on IT firms are increasingly having cross-industry ramifications, not just by compromising customer infrastructure but also by interrupting customer business operations due to service failures. The popularity of crypto currencies is growing, and a growing number of blockchain-based enterprises are developing. Attackers are close following, with the number of assaults on blockchain projects more than doubling since 2021.

Successful cyber attacks on businesses' web resources increased by 56%. While 17% of corporate online resources were targeted in 2021, this ratio grew to 22% in 2022. Cyber attacks strike organizations in a variety of industries, with government entities bearing the brunt of the damage. The number of events targeting these agencies more than doubled, with their proportion increasing from 23% to 41%.

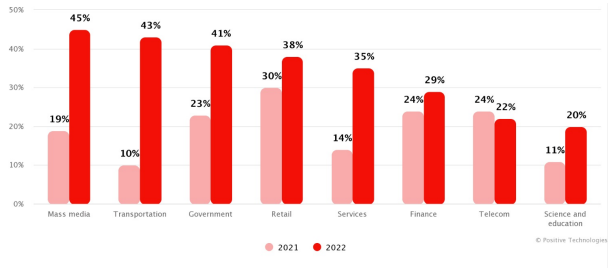


**Fig. 5** Percentage of incidents involving attacks on web resources in 2021 and 2022 [60].

Cyberspace tensions led to an increase in hacktivist assaults on organizations' websites in 2022, causing

interruptions in 53% of cases as shown in Figure 5. These assaults interrupt critical corporate operations while collecting important user information. Because of sensitive client information, retail remains an appealing target for cyber assaults. The increase in assaults is attributable to vulnerabilities in popular plugins such as WordPress and Magento, with the most often exploited vulnerabilities being Log4Shell, Spring4Shell, Adobe Commerce, October CMS, and WPGateway. The rise of the shadow market has had a significant influence on the amount of cyberattacks, which are expected to rise in near future for organizations that provide online services and gather user data. Large-scale data breaches plagued companies and individuals all throughout the world in 2022, including those in Russia as shown in Figure 6. With 82% of incidents involving sensitive information theft, medical institutions, scientific research and educational service providers, and merchants were the most affected. The average cost of such catastrophes hit a new high of $4.35 million, a 2.6% rise over the previous year. 47% of sensitive information was compromised in successful assaults on corporations, with personal data being the most prevalent. Attackers targeted trade secrets and account credentials in 64% of cases, while persons were targeted in 64% of cases. The proportion of stolen personal data grew in 2021, with companies increasing by 4 percentage points and individuals increasing by 8 percentage points.

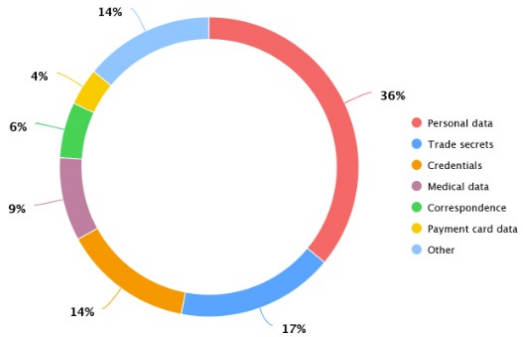


**Fig. 6** Types of data stolen (in successful attacks on organizations) [60].

Over the course of the year, there was a considerable increase in the use of spyware. In 2021, the use of spyware in attacks on companies was 12%, while it was 32% in assaults on individuals as shown in Figure 7. Throughout 2022, the number of events employing spyware increased steadily, resulting in an increase in the percentage of this sort of malware, with 13% used in attacks on businesses and 43% used in attacks on people.

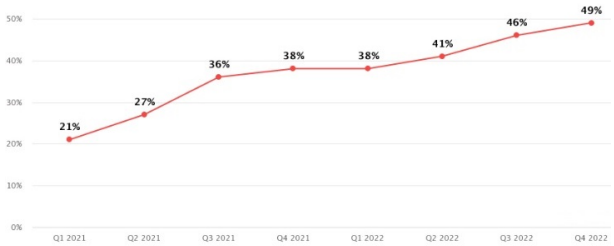


**Fig. 7** Use of spyware in attacks on individuals [60].

Phishing websites are the predominant distribution mechanism for spyware in individual-targeted assaults, accounting for 42% of all occurrences. Attackers used email as a vector of attack less frequently, accounting for only 20% of instances. Meanwhile, users of social media and messaging applications should exercise caution, since thieves used these channels to transmit malware in 14% of cases, while message and text services were used in 9% of assaults (see Figure 8). Throughout 2022, there were multiple reports of malware being discovered in legitimate app stores (10%). The vast majority of these assaults targeted Android mobile

device users. Accenture specialists discovered that the most widely utilized spyware infections were RedLine, Vidar, and Raccoon Stealer. New companies, such as BlueFox, Aurora, and Erbium, as well as regular upgrades and malware-as-a-service schemes, have made spyware a popular choice among attackers, decreasing the entrance threshold for criminality. The emergence of the shadow market has had a substantial influence on the proliferation of spyware assaults. Remote control malware and spyware are the most often mentioned forms of malware, accounting for 48% of malware-related letters. These programs frequently include stealth capabilities such as text message interception, monitoring user location, and screen capture.

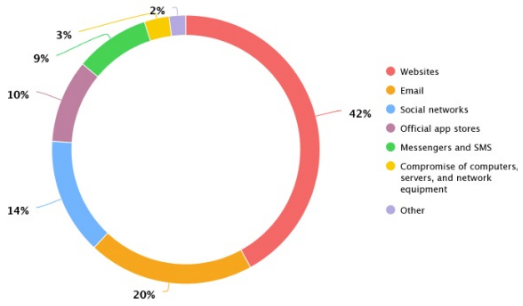


**Fig. 8** Methods of spyware distribution in successful attacks on individuals [60].

Prices for these tools start at $10, and others are given out for free by thieves. Criminals want to capture user data and sell it on shady websites, mostly for checking in to services, social networks, and messengers. Spyware can also compromise company credentials by infecting personal devices used by workers to connect to work resources. In 2022, ransomware continued to develop, with attackers integrating such tools in every other successful malware-based assault (51%). State institutions (15%), industrial businesses (15%), medical organizations (14%), and scientific and educational institutions (13%), were the most commonly targeted by ransomware operators (see Figure 9).

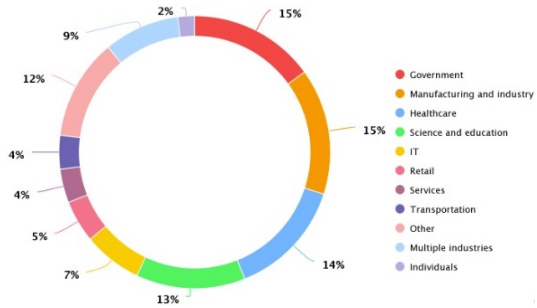


**Fig. 9** Distribution of ransomware incidents by industry [60].

Nine out of ten ransomware events caused disruptions to an organization's key operations and resulted in the destruction of infrastructure, data, and services. The most prevalent categories of private information stolen, accounting for 55% of assaults, were trade secrets and personal information. In 12% of cases, victims experienced financial losses. Ransomware organizations claimed a 40% drop in revenue in 2022, while a Coveware investigation revealed that the proportion of victims who handed over ransom to attackers dropped from 76% in 2019 to 41% as shown in Figure 10.

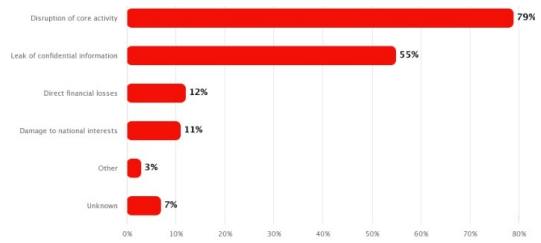


**Fig. 10** Consequences of ransomware attacks (percentage of successful attacks) [60].

Several ransomware organizations, including LockBit, Hive, Voice Society, BlackCat (ALPHV), and Conti, started to appear in 2022. The active ransomware organization known as LockBit targets vital infrastructure facilities and socially significant institutions with its aggressive actions. A decryptor was made available as a result of the Federal Bureau of Investigation (FBI) hacking Hive. Voice Society targets academic and research organizations, whereas BlackCat (ALPHV) is a more recent organization with vast ransomware deployment expertise. Due to intelligence agency targeting, Conti, a long-standing menace and industry leader in ransomware-as-a-service, was forced to quit the scene in May 2022. Ransomware organizations showed a greater interest in Linux-based systems in 2022, and there was a move toward cross-platform malware created with the Rust programming language. As a result of the lack of analysis tools, it is harder for defenses to find malware. Major players have released cross-platform solutions, including RansomEXX, Black Basta, and Hive. Attackers compromise as many devices as they can in a short amount of time, execute the virus, and swiftly encrypt data before it is discovered. This is how ransomware assaults work. In 2022, intermittent file encryption has gained popularity since it is quicker and less obvious to programs that look for suspicious activities. The unorthodox strategy employed by BlackCat group omits the encryption stage in favor of an exfiltration tool. The publication of the source codes for popular ransomware families like Conti and Yanluowang in 2022 may have slowed the spread of ransomware assaults. New ransomware organizations might, however, emerge and use the exposed source code to create their own encryption models and conduct attacks. The proportion of social engineering incidents in individual assaults climbed from 88% in 2021 to 93% in 2022. Despite the fact that the number of assaults on businesses has remained constant, the percentage of occurrences utilizing this strategy has decreased from 50% to 43%. In the face of many data dumps, attackers can plan assaults using compromised data, including credentials. By compromising credentials, attackers were able to get access to target systems and resources in 16% of successful assaults on businesses in 2022. This was accomplished by either mining passwords or utilizing compromised credentials obtained from data breaches.

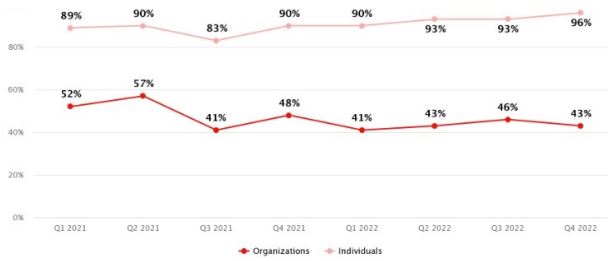


**Fig. 11** Share of incidents with the use of social engineering [60].

Malicious emails were used in roughly nine out of ten successful social engineering assaults against corporations. Fraudsters predominantly used phishing sites (56% of all assaults against people). We also saw in Figure 11 an increase in successful assaults using messengers and SMS texts (18%) and social networks (21%). Throughout 2022, the number of successful assaults on IT organizations climbed significantly, with the number of attacks in the fourth quarter roughly double that of the first. The most common occurrences at IT businesses

comprised sensitive information breaches, which accounted for 63% of total incidents, followed by interruption of core business (35% of cases), and the use of corporate resources to perform assaults (13% of instances) (see Figure 12).

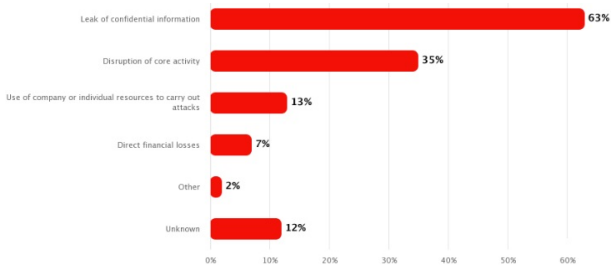


**Fig. 12** Consequences of attacks on IT companies (percentage of successful attacks) [60].

A series of assaults by the Lapsus$ organization on IT goods in 2022 resulted in 31% of private information being exposed. These attacks used stolen certificates to make malware look legal. IT solutions employed by diverse companies and people can interrupt service delivery and render enterprises unusable. In certain situations, attackers used IT goods and service providers to target people. Okta, a supplier of multifactor authentication systems, was the victim of a series of successful assaults that allowed attackers to access the data of over 300 corporate customers. In 2022, attackers continued to target cloud services and virtualization settings, propagating malware using libraries for popular frameworks. Attacks on software supply chains and IT firms' clients are predicted to continue in future. Developers must be on the lookout for these heinous events and implement adequate safeguards. Analyzing code for security vulnerabilities on a regular basis, extensively reviewing third-party libraries, and participating in bug bounty programs may all help find and address security concerns in their products. Attackers' interest in bitcoin exchanges and DeFi protocols surged considerably in 2022, with assaults on blockchain projects more than tripling from 2021. Attackers successfully took cash in 78% of situations, causing harm in certain circumstances. With $3.8 billion lost, Chainalysis revealed the largest financial losses caused by hacks on bitcoin organizations. Ronin sidechain ($617 million stolen), BSC Token Hub ($566 million stolen), and Wormhole ($326 million stolen) are examples of high-profile attacks. Smart contract vulnerabilities were the most typically exploited by attackers, with flash loan assaults being the most prevalent. Attackers distribute posts on social media and messengers about free token and NFT giveaways, enticing users to deposit cash and promising higher profits. The majority of thieves have already mastered the capacity to steal the credentials of popular bitcoin wallets. In future, the number of assaults against blockchain projects is likely to increase, as will the amount of fraud targeting bitcoin asset owners. Users should exercise care, undertake rigorous research before investing, and use two-factor authentication to safeguard their accounts. To uncover vulnerabilities, developers should audit smart contracts, use safe development methods, and participate in bug bounty schemes.

## 5 ML-based web application security

Online application security based on ML has become a potential method for boosting online applications' security [15]. Traditional security measures are no longer sufficient to ensure the security of online applications due to the rising number of cyber threats and assaults [50]. The capacity to recognize and counteract security risks, adjust to changing attack patterns, and enhance the overall security posture of online applications are all provided by ML algorithms [4,5]. Users access online applications using web browsers, which are sophisticated software programs that run on web servers. They let users to do transactions, engage with web content, and talk to other users. However, they are also susceptible to a variety of security risks, including as denial-of-

service (DoS) attacks, XSS attacks, and injection attacks. These assaults may result in data breaches, the loss of confidential information, and harm to a company's reputation [61]. In order to detect and stop security risks in online applications, ML-based web application security requires training ML algorithms [24]. Using patterns in online traffic and user behavior, these algorithms can spot suspicious or malicious activities. They may also examine the code of online applications to find holes that attackers could exploit. In order to create a layered defense against cyber-attacks, ML-based security solutions can be employed in conjunction with more conventional security measures like firewalls, intrusion detection and prevention systems (IDPS), and encryption [4, 56]. The capacity of ML-based web application security to change with evolving attack patterns is one of its main advantages [50, 61]. Attackers are always developing new strategies to get around established security measures due to the ongoing evolution of cyber threats. ML algorithms can adapt their behavior in response to new attack patterns, making it more challenging for attackers to use web application weaknesses. Real-time threat identification and response is another benefit of ML-based online application security. The manual analysis and intervention that are a common part of traditional security procedures might delay the identification and handling of security issues [24]. Data breaches and other security problems are less likely because to ML-based security systems' ability to scan web traffic in real-time and immediately react to security concerns. There are difficulties with ML-based online application security, though. The necessity for a lot of high-quality data to properly train ML algorithms is one of the main obstacles. To enable the algorithms to learn from a variety of security risks and attack patterns, this data must be broad, representative, and current. The possibility for false positives and false negatives in ML-based security solutions is another difficulty. False positives happen when harmless online traffic is mistakenly labeled as harmful, while false negatives happen when the algorithm fails to identify malicious traffic. These mistakes increase the likelihood of security events and might cause users to lose faith in the security solution [50]. A promising strategy for improving web application security is ML-based security. It provides the capacity to recognize and thwart security risks, adapt to changing attack patterns, and deliver real-time threat detection and response. But it also has drawbacks, such the requirement for a lot of high-quality data and the possibility of false positives and false negatives. ML-based web application security is going to play a bigger role in a complete security strategy for online applications as cyber threats continue to develop [4, 61].

### 5.1 Proposed algorithms for web application security

Web application security has undergone a revolution thanks to ML algorithms, which have made it much easier to identify and stop different kinds of intrusions [16, 62]. Table 3 discusses some of the most popular ML techniques for web application security in this part.

Web applications may now be protected against a variety of threats thanks to ML techniques [16]. It's crucial to remember that these algorithms are not infallible and can be subject to adversarial assaults. As a result, it's critical to keep creating and enhancing these algorithms to keep up with the development of cyber threats.

### 5.2 Performance evaluation and comparison of ML-based web application security techniques

In recent years, there has been a lot of interest in the use of ML methods to online application security [4]. Several ML-based strategies have been suggested by researchers to increase the security of online applications [50]. These methods try to spot and stop attacks like CSRF, SQL Injection, and cross-site scripting. To determine which strategies are the most successful, it is necessary to assess and compare their performance. However, the efficacy of various techniques differs [24]. A key component of ML-based web application security solutions is performance evaluation. Recent studies have assessed these strategies' effectiveness using a variety of measures, including accuracy, precision, recall, F1-score, and AUC-ROC. These metrics measure the effectiveness of the methods in terms of their capacity to distinguish between legitimate and malicious traffic. One study evaluated the performance of four different ML algorithms (Support Vector Machines, Random Forest, Multilayer Perceptron, and Naive Bayes) for detecting SQL injection attacks [64, 65]. The study used a dataset containing both benign and malicious SQL queries. The results

**Table 3** Machine learning algorithms in web application security.

| Learning Type | Algorithm | Description |
|---|---|---|
| Supervised | Decision Trees | Simple yet effective techniques that may be applied to classification and regression problems. They operate by repeatedly dividing the data into subgroups depending on the value of a feature until a stopping re-requirement is satisfied [16, 62, 63] |
| Supervised | SVMs | Binary classifiers that locate the hyperplane that most effectively distinguishes between the two classes. They may be taught to handle non-linear decision limits and are especially helpful for high-dimensional data [63]. |
| Supervised | Neural Network | A method for ML that is fashioned after the human brain. They are made up of layers of linked neurons, or nodes, that are capable of learning sophisticated representations of the incoming data [16, 62, 63]. |
| Unsupervised | K-means Clustering | A quick and effective technique that divides the data into k groups according to how similar the data points are to one another. It may be applied to finding malicious activity clusters and anomaly detection [16, 63]. |
| Unsupervised | PCA | A method for lowering the dimensionality of high-dimensional data while maintaining as much of the variance as feasible by projecting it onto a lower-dimensional subspace. |
| Deep Learning | CNNs | Particularly effective for image and video recognition tasks and have been used for detecting malicious behavior in network traffic [63]. |
| Deep Learning | RNNs | Well-suited for sequential data and have been used for detecting attacks that occur over time, such as SQL injection and cross-site scripting [63]. |

showed that SVM achieved the highest accuracy of 97.89%, followed by Random Forest with 97.37%, Multilayer Perceptron with 96.47%, and Naive Bayes with 94.25%. The study concluded that SVM and Random Forest are effective techniques for detecting SQL injection attacks [64–66]. Different studies evaluated the performance of six different ML algorithms (Support Vector Machines, Decision Tree, Random Forest, AdaBoost, Gradient Boosting, and XGBoost) for detecting cross-site scripting attacks [67, 68]. The study used a dataset containing both benign and malicious JavaScript code snippets. The results showed that XGBoost achieved the highest accuracy of 98.32%, followed by Gradient Boosting with 98.23%, Random Forest with 97.97%, AdaBoost with 97.90%, Decision Tree with 97.62%, and SVM with 96.86%. The studies concluded that XGBoost and Gradient Boosting are effective techniques for detecting cross-site scripting attacks [67–69]. In addition to evaluating the performance of individual ML algorithms, studies have also compared the performance of different feature selection techniques. Feature selection is an important aspect of ML-based web application security techniques, as it aims to identify the most relevant features for detecting attacks [4]. One study compared the performance of three different feature selection techniques (Chi-Square, Information Gain, and ReliefF) for detecting SQL injection attacks [70]. The study used a dataset containing both benign and malicious SQL queries. The results showed that ReliefF achieved the highest accuracy of 98.01%, followed by Chi-Square with 96.85%, and Information Gain with 95.29%. The study concluded that ReliefF is an effective feature selection technique for detecting SQL injection attacks [70]. Identification of the most efficient strategies requires performance evaluation and comparison of various approaches [62]. These

studies also emphasize how crucial it is to choose the right feature selection methods and ML algorithms for particular kinds of assaults. Identification of the most efficient strategies requires performance evaluation and comparison of various approaches. To further enhance the security of online applications, researchers should investigate novel strategies and assess their effectiveness.

### 5.3 Challenges and limitations in ML-based web application security

There is a lot of interest in using ML techniques to secure web applications. There are a lot of benefits, such as the ability to see complex attacks and anomalies, but there are also a number of challenges and limitations that must be addressed. In this section, we will discuss some of the most prevalent issues and limitations related to ML-based web application security [4]. One of the primary obstacles is the scarcity of high-quality training data. Effective ML algorithms require substantial volumes of data to learn and perform optimally. However, obtaining good training data for online application security is often challenging [24, 46]. This difficulty arises because high-quality, labeled data, especially for specific types of cyberattacks, is not always readily available. The dynamic nature of online applications and the continuously evolving threat landscape further complicate the task of keeping training data up-to-date, which can limit the effectiveness of ML models. Another critical issue is the interpretability of ML models. ML algorithms are frequently criticized for being "black boxes," where the decision-making process is opaque. This lack of transparency makes it difficult to understand how the algorithms reach their conclusions, which can complicate the identification of false positives or false negatives and hinder the ability to justify security decisions [54]. To improve interpretability, utilizing transparent ML models like decision trees or employing post-hoc interpretability techniques can be beneficial. Moreover, extensive data collection necessary for training ML models can raise privacy concerns. Handling sensitive information without proper consent can infringe on individual privacy rights, particularly in surveillance applications. The inherent "black box" nature of some ML models can also obstruct accountability, making it challenging to address ethical issues and ensure that privacy and data protection standards are met. To mitigate these concerns, it is crucial to establish robust ethical guidelines, ensure data protection, and promote transparency in ML applications. The performance of ML-based web application security techniques can also be affected by the quality of the features used for training the models [50]. The selection of appropriate features is critical in determining the effectiveness of the ML algorithms.

The effectiveness of ML algorithms can be impacted by the presence of noisy or irrelevant features that can lead to inaccurate results. Another challenge faced in ML-based web application security is the ability of attackers to evade detection. Attackers can use techniques such as obfuscation and evasion to bypass machine-learning-based security systems [4, 50]. An attacker can modify the attack vector or payload to evade detection, making it difficult for ML algorithms to detect such attacks. To address this challenge, security analysts must continuously monitor and update ML models to ensure that they can detect new and emerging attack techniques. In cybersecurity environments, models must not only detect and mitigate known threats but also withstand adversarial conditions and handle noisy or manipulated data. Adversarial attacks exploit weaknesses in ML algorithms by creating inputs designed to deceive models into making incorrect predictions or classifications [71]. Several approaches are needed to solve this problem, such as adding adversarial examples to the training dataset to make the model more resistant to attacks like these that happen in real life [72]. Robust evaluation metrics, like adversarial accuracy and perturbation sensitivity, are essential to assess model performance under adversarial conditions [71]. Noise-handling mechanisms—such as data augmentation, denoising autoencoders, and robust feature extraction—enhance a model's ability to maintain accuracy despite corrupted or altered input data [71]. Leveraging ensemble methods by combining multiple models can also improve robustness against adversarial attacks, reducing vulnerabilities and increasing overall reliability. Continuous model updating, which includes regular retraining with new data, ensures that models remain effective against evolving threats and manipulation tactics. Implementing comprehensive adversarial testing frameworks allows for the simulation of various attack scenarios, aiding in the assessment of model performance and the identification of potential weaknesses. Another significant challenge in the application of

ML-based security models is the need for continuous updating and maintenance as attack vectors evolve. Cyber threats are dynamic, and models trained on historical data can quickly become outdated if not regularly updated with new data, including the latest adversarial examples [73]. Continuous model updating, through strategies like incremental learning and automated retraining pipelines, is essential to ensure that these models remain robust against emerging threats. However, implementing such strategies presents its own set of challenges, including the risk of overfitting, the need for substantial computational resources, and the potential introduction of new vulnerabilities during updates [73–75]. Addressing these challenges is critical for the long-term efficacy of ML-based security systems and represents a key area for future research and development. In deploying ML models for web application security, real-time detection capabilities are critical. The challenge lies in processing data instantly to identify and mitigate threats as they arise. Latency in detection can result in delayed responses, giving attackers the opportunity to breach systems before defences are activated [76]. To address this, optimising algorithms for faster inference times is essential. This can be achieved through techniques such as model pruning, quantisation, and using lightweight architectures. Implementing hardware accelerators like GPUs and TPUs can significantly boost processing speed, ensuring that the system can keep pace with real-time demands [76]. Scalability is another crucial aspect, particularly for large-scale web applications that generate vast amounts of data, especially during peak usage. Security systems must efficiently handle this data influx without compromising performance [75]. Distributed computing approaches, including cloud-based services and parallel processing techniques, are vital for managing and analysing large datasets. These approaches allow for horizontal scaling, ensuring that the system can accommodate increasing data volumes without a drop in detection speed or accuracy. Integrating ML models with existing security infrastructure poses additional challenges, especially when compatibility with legacy systems is required. A gradual integration strategy is recommended, where hybrid models combining traditional rule-based systems with ML techniques are employed. This ensures a seamless operation alongside other security measures and maintains system integrity. Adopting interoperability standards can further ease the integration process. Operational feasibility also needs to be addressed, involving resource allocation, maintenance, and continuous monitoring. Automation tools for monitoring and updating ML models are essential for keeping systems current with evolving threats. Continuous training and model updates are crucial to maintaining the effectiveness of these security solutions over time.

### 5.4 Future directions and opportunities in ML-based web application security

As ML-based techniques continue to advance, there are many opportunities for their application in the field of web application security. Here are some potential future directions and opportunities in this area as shown in Table 4. In the area of online application security, there are a lot of fascinating potential for ML-based solutions. We may anticipate more accuracy and efficiency in detecting threats as well as greater integration with current security measures as algorithms evolve further and data sets get larger. Real-time detection, enhanced scalability, and increased accountability are also required for these approaches' operation. ML-based solutions can become more crucial in defending online applications against security risks by addressing these issues and possibilities.

## 6 Conclusion

ML-based web application security holds significant promise, with substantial opportunities for growth and enhancement. As ML algorithms evolve, we can expect improvements in scalability, real-time detection, accuracy, and efficiency. The integration of Explainable AI will be crucial for enhancing transparency and accountability in decision-making, while advancements in Adversarial ML will be essential to developing robust defences against attacks designed to manipulate ML systems. Future research should prioritize automated response mechanisms to enable rapid and effective responses to potential breaches. Collaborations between cybersecurity specialists and data scientists will be vital for ensuring that ML algorithms are precise,

**Table 4** Future directions in ML-based web application security.

| Research Area | Description |
|---|---|
| Improved accuracy and efficiency | One of the biggest challenges facing ML-based web application security techniques is achieving high accuracy while also maintaining efficiency. As algorithms continue to develop and data sets become larger, there may be opportunities to improve the accuracy and efficiency of these techniques. |
| Better integration with existing security measures | ML-based techniques can complement traditional security measures such as firewalls and intrusion detection systems. Future research may focus on findin ways to integrate ML-based techniques with these existing measures to improve overall security. |
| Greater focus on real-time detection | Many current ML-based web application security techniques are designed to detect threats in batch processing mode. However, as attacks become more sophisticated, there is a need for real-time detection and response. Future research may focus on developing real-time detection methods that can quickly identify and respond to potential threats. |
| Improved scalability | As web applications continue to grow in complexity and scale, ML-based techniques must be able to keep up. Future research may focus on developing techniques that can scale to handle larger and more complex web applications. |
| Integration of explainable AI | As ML-based techniques are increasingly adopted in the field of web application security, there is a need for greater transparency and accountability in how these techniques work. Explainable AI, which refers to the ability of algorithms to provide understandable explanations for their decisions, may be a useful tool for achieving greater transparency and accountability. |
| Adversarial ML | Adversarial ML involves designing attacks specifically to fool ML-based algorithms. This is a growing concern in the field of web application security, as attackers may use adversarial ML to circumvent security measures. Future research may focus on developing techniques that are more resistant to adversarial attacks. |
| Automated response | ML-based techniques can also be used to automate responses to security threats. ML algorithm could automatically block traffic from a particular IP address if it is determined to be malicious. Future research may focus on developing automated response systems that can quickly and effectively respond to potential threats. |

reliable, and capable of adapting to emerging threats. Despite the considerable promise of deep learning and other ML techniques, it is crucial to recognize their limitations and ensure they are continuously updated to address new vulnerabilities. Effective ML-based cybersecurity models must not only detect known threats but also be resilient to adversarial attacks and handle noisy data. Implementing strategies like adversarial training, robust evaluation metrics, and noise-handling mechanisms will enhance model resilience. Real-time detection and scalability are essential for large-scale applications, requiring ongoing optimization and distributed computing. Integrating ML with existing security systems and maintaining continuous updates are critical for long-term effectiveness. This paper underscores the potential of ML-based web application security, highlighting advancements in scalability, real-time detection, and accuracy. Key contributions include the integration of Explainable AI, focus on Adversarial ML, emphasis on automated responses, and the need for collaborative development of adaptive security solutions. ML-based solutions have the potential to significantly enhance web application security, offering a vital defense against evolving cyber threats. As the complexity and volume of these threats grow, the demand for innovative and effective security solutions will increase. Continued investment and research in ML algorithms are essential to address these challenges and advance the future of web application security.

## 7 Declarations

### 7.1 Conflict of interests:

The authors hereby declare that there is no conflict of interests regarding the publication of this paper.

### 7.2 Funding:

There is no funding regarding the publication of this paper.

### 7.3 Author's contributions:

O.B.E.-Conceptualisation, Methodology, Validation, Formal Analysis, Investigation, Resources, Data Curation, Writing. All authors read and approved the final submitted version of this manuscript.

### 7.4 Acknowledgement:

The author extends sincere thanks to the reviewers and editors for their constructive feedback and contributions.

### 7.5 Data availability statement:

This paper is a theoretical work and has been developed with data which is still being used for development and research.

### 7.6 Using of AI tools:

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

## References

[1] Badawy A.M., Technology management simply defined: A tweet plus two characters, Journal of Engineering and Technology Management, 26(4), 219–224, 2009.
[2] Ponnapalli H.K.B., Saxena A., A digital signature architecture for web apps, IT Professional, 15(2), 42–49, 2013.
[3] Abomhara M., Køien G.M., Cyber security and the internet of things: vulnerabilities, threats, intruders, and attacks, Journal of Cyber Security and Mobility, 4, 65–88, 2015.

[4] Kumar R., Khan A.I., Abushark Y.B., Alam M.M., Agrawal A., Khan R.A., An integrated approach of fuzzy logic, AHP, and TOPSIS for estimating usable-security of web applications, IEEE Access, 8, 50944–50957, 2020.
[5] Aydos M., Aldan Ç., Coşkun E., Soydan A., Security testing of web applications: A systematic mapping of the literature, Journal of King Saud University - Computer and Information Sciences, 34(9), 6775–6792, 2022.
[6] Genge B., Kiss I., Haller P., A system dynamics approach for assessing the impact of cyber-attacks on critical infrastructures, International Journal of Critical Infrastructure Protection, 10, 3–17, 2015.
[7] Tao F., Akhtar M.S., Jiayuan Z., The future of artificial intelligence in cybersecurity: A comprehensive survey, EAI Endorsed Transactions on Creative Technologies, 8(28), e3, 2021.
[8] Alawida M., Omolara A.E., Abiodun O.I., Al-Rajab M., A deeper look into cybersecurity issues in the wake of Covid-19: A survey, Journal of King Saud University–Computer and Information Sciences, 34(10), 8176–8206, 2022.
[9] Cui L., Yang S., Chen F., Ming Z., Lu N., Qin J., A survey on application of machine learning for internet of things, International Journal of Machine Learning and Cybernetics, 9, 1399–1417, 2018.
[10] Al-Garadi M.A., Mohamed A., Al-Ali A.K., Du X., Ali I., Guizani M., A survey of machine and deep learning methods for internet of things (IoT) security, IEEE Communications Surveys & Tutorials, 22(3), 1646–1685, 2020.
[11] Nguyen T.T., Reddi V.J., Deep reinforcement learning for cyber security, IEEE Transactions on Neural Networks and Learning Systems, 34(8), 3779–3795, 2021.
[12] Venketeswaran A., Lalam N., Wuenschell J., Ohodnicki P.R., Badar M., Chen K.P., Lu P., Duan Y., Chorpening B., Buric M., Recent advances in machine learning for fiber optic sensor applications, Advanced Intelligent Systems, 4(1), 2100067, 2021.
[13] Antoniadi A.M., Du Y., Guendouz Y., Wei L., Mazo C., Becker B.A., Mooney C., Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: A systematic review, Applied Sciences, 11(11), 5088, 2021.
[14] Macas M., Wu C., Fuertes W., A survey on deep learning for cybersecurity: Progress, challenges, and opportunities, Computer Networks, 212, 109032, 2022.
[15] Seng L.K., Ithnin N., Said S.Z.M., The approaches to quantify web application security scanners quality: A review, International Journal of Advanced Computer Research, 8(38), 285–312, 2018.
[16] https://www.alifconsulting.com/post/azure-vms-security-overview, Accessed: December 6, 2023.
[17] Bowen B.M., Hershkop S., Keromytis A.D., Stolfo S.J., Security and Privacy in Communication Networks, (Chapter 3: Baiting inside attackers using decoy documents), Security and Privacy in Communication Networks, 14–18 September 2009, Athens, Greece.
[18] AbouTrab M.S., Brockway M., Counsell S., Hierons R.M., Testing real-time embedded systems using timed automata based approaches, Journal of Systems and Software, 86(5), 1209–1223, 2013.
[19] Abolfazli S., Sanaei Z., Alizadeh M., Gani A., Xia F., An experimental analysis on cloud-based mobile augmentation in mobile cloud computing, IEEE Transactions on Consumer Electronics, 60(1), 146–154, 2014.
[20] Aazam M., Zeadally S., Harras K.A., Deploying fog computing in industrial internet of things and industry 4.0, IEEE Transactions on Industrial Informatics, 14(10), 4674–4682, 2018.
[21] Xu L.D., Xu E.L., Li L., Industry 4.0: State of the art and future trends, International Journal of Production Research, 56(8), 2941–2962, 2018.
[22] Oztemel E., Gursev S., Literature review of Industry 4.0 and related technologies, Journal of Intelligent Manufacturing, 31, 127–182, 2020.
[23] Kaka H., Zhang E., Khan N., Artificial intelligence and deep learning in neuroradiology: Exploring the new frontier, Canadian Association of Radiologists Journal, 72(1), 35–44, 2020.
[24] Sönmez F.Ö., Security qualitative metrics for open web application security project compliance, Procedia Computer Science, 151, 998–1003, 2019.
[25] Safianu O., Twum F., Hayfron-Acquah J.B., Information system security threats and vulnerabilities: Evaluating the human factor in data protection, International Journal of Computer Applications, 143(5), 8–14, 2016.
[26] Appiah V., Asante M., Nti I.K., Nyarko-Boateng O., Survey of websites and web application security threats using vulnerability assessment, Journal of Computer Science, 15(10), 1341–1354, 2019.
[27] Althunayyan M., Saxena N., Li S., Gope P., Evaluation of black-box web application security scanners in detecting injection vulnerabilities, Electronics, 11(13), 2049, 2022.
[28] Aminanto M.E., Ban T., Isawa R., Takahashi T., Inoue D., Threat alert prioritization using isolation forest and stacked auto encoder with day-forward-chaining analysis, IEEE Access, 8, 217977–217986, 2020.
[29] Lin Y.D, Liu Z.Q., Hwang R.H., Nguyen V.L., Lin P.C., Lai Y.C., Machine learning with variational AutoEncoder for imbalanced datasets in intrusion detection, IEEE Access, 10, 15247–15260, 2022.
[30] Mirza B., Haroon D., Khan B., Padhani A., Syed T.Q., Deep generative models to counter class imbalance: A model-metric mapping with proportion calibration methodology, IEEE Access, 9, 55879–55897, 2021.
[31] Hashemi S.K., Mirtaheri S.L., Greco S., Fraud detection in banking data by machine learning techniques, IEEE Access, 11, 3034–3043, 2022.

[32] D'Alconzo A., Drago I., Morichetta A., Mellia M., Casas P., A survey on big data for network traffic monitoring and analysis, IEEE Transactions on Network and Service Management, 16(3), 800–813, 2019.
[33] Natalino C., Schiano M., Giglio A.D., Wosinska L., Furdek M., Experimental study of machine-learning-based detection and identification of physical-layer attacks in optical networks, Journal of Lightwave Technology, 37(16), 4173–4182, 2019.
[34] Ray P.P., A survey on internet of things architectures, Journal of King Saud University–Computer and Information Sciences, 30(3), 291–319, 2018.
[35] Attaran M., The impact of 5G on the evolution of intelligent automation and industry digitization, Journal of Ambient Intelligence and Humanized Computing, 14(5), 5977–5993, 2023.
[36] Alshamrani A., Myneni S., Chowdhary A., Huang D., A survey on advanced persistent threats: Techniques, solutions, challenges, and research opportunities, IEEE Communications Surveys & Tutorials, 21(2), 1851–1877, 2019.
[37] Chaabouni N., Mosbah M., Zemmari A., Sauvignac C., Faruki P., Network intrusion detection for IoT security based on learning techniques, IEEE Communications Surveys & Tutorials, 21(3), 2671–2701, 2019.
[38] Bolla R., Bruschi R., Davoli F., Cucchietti F., Energy efficiency in the future internet: A survey of existing approaches and trends in energy-aware fixed network infrastructures, IEEE Communications Surveys & Tutorials, 13(2), 223–244, 2010.
[39] Negandhi P., Trivedi Y., Mangrulkar R., Emerging Research in Computing, Information, Communication and Applications (Chapter 43: Intrusion detection system using random forest on the NSL-KDD dataset), 27–28 July 2018, Karnataka, India.
[40] Sultana N., Chilamkurti N., Peng W., Alhadad R., Survey on SDN based network intrusion detection system using machine learning approaches, Peer-to-Peer Networking and Applications, 12, 493–501, 2019.
[41] https://www.opensourceforu.com/2020/05/machine-learning-and-deep-learning-a-perspective-on-the-future/, Accessed: May 13, 2020.
[42] Habtamu G.T., Kassahun A.Y., A systematic review of botnet detection system using deep learning and machine learning approaches, SSRN Electronic Journal, 4256438, 1–10, 2022.
[43] Boutaba R., Salahuddin M.A., Limam N., Ayoubi S., Shahriar N., Estrada-Solano F., Caicedo O.M., A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities, Journal of Internet Services and Applications, 9(16), 1–99, 2018.
[44] Wang M., Cui Y., Wang X., Xiao S., Jiang J., Machine learning for networking: Workflow, advances and opportunities, IEEE Network, 32(2), 92–99, 2017.
[45] Bhutani G., Application of machine-learning based prediction techniques in wireless networks, International Journal of Communications Network and System Sciences, 7(5), 131–140, 2014.
[46] Usama M., Qadir J., Raza A., Arif H., Yau K.L.A., Elkhatib Y., Unsupervised machine learning for networking: Techniques, applications and research challenges, IEEE Access, 7, 65579–65615, 2019.
[47] Fang W., Tan X., Wilbur D., Application of intrusion detection technology in network safety based on machine learning, Safety Science, 124, 104604, 2020.
[48] Kulin M., Kazaz T., Poorter E.D., Moerman I., A survey on machine learning-based performance improvement of wireless networks: PHY MAC and network layer, Electronics, 10(3), 318, 2021.
[49] Rodrigues J.J.P.C., Neves P.A.C.S., A survey on IP-based wireless sensor network solutions, International Journal of Communication Systems, 23(8), 963–981, 2010.
[50] Alzahrani F.A., Fuzzy based decision-making approach for estimating usable-security of healthcare web applications, Computers Materials & Continua, 66(3), 2599–2625, 2021.
[51] Olalere M., Abdullah M.T., Mahmod R., Abdullah A., A review of bring your own device on security issues, SAGE Open, 5(2), 1–11, 2015.
[52] Ahmad I., Namal S., Ylianttila M., Gurtov A., Security in software defined networks: A survey, IEEE Communications Surveys & Tutorials, 17(4), 2317–2346, 2015.
[53] Cerf V.G., The fragmentation of the internet, IEEE Internet Computing, 20(1), 88-c3, 2016.
[54] Brar H.S., Kumar G., Cybercrimes: A proposed taxonomy and challenges, Journal of Computer Networks and Communications, 2018(D 1798659), 1–11, 2018.
[55] Fonseca O., Cunha Í., Fazzion E., Meira W., Silva B.A.D., Ferreira R.A., Katz-Bassett E., Identifying networks vulnerable to IP spoofing, IEEE Transactions on Network and Service Management, 18(3), 3170–3183, 2021.
[56] Toprak S., Yavuz A.G., Web application firewall based on anomaly detection using deep learning, Acta Infologica, 6(2), 219–244, 2022.
[57] Patcha A., Park J.M., An overview of anomaly detection techniques: Existing solutions and latest technological trends, Computer Networks, 51(12), 3448–3470, 2007.
[58] Praseed A., Thilagam P.S., DDoS attacks at the application layer: Challenges and research perspectives for safeguarding web applications, IEEE Communications Surveys & Tutorials, 21(1), 661–685, 2018.
[59] Chang R.K.C., Defending against flooding-based distributed denial-of-service attacks: A tutorial, IEEE Communications Magazine, 40(10), 42–51, 2002.

[60] https://global.ptsecurity.com/analytics/cybersecurity-threatscape-2022-rundown, Accessed: April 16, 2024.
[61] Alenezi M., Agrawal A., Kumar R., Khan R.A., Evaluating performance of web application security through a fuzzy based hybrid multi-criteria decision-making approach: Design tactics perspective, IEEE Access, 8, 25543–25556, 2020.
[62] Iskandar A., Tuasamu M.R.F., Syamsu S., Mansyur M., Listyorini T., Sallu S., Supriyono S., Saddhono K., Napitupulu D., Rahim R., Web based testing application security system using semantic comparison method, IOP Conference Series: Materials Science and Engineering, 420, 012122, 19–20 July 2018, Medan, Indonesia.
[63] Choraś M., Kozik R., Machine learning techniques applied to detect cyber attacks on web applications, Logic Journal of IGPL, 23(1), 45–56, 2014.
[64] Santos R.J., Bernardino J., Vieira M., Approaches and challenges in database intrusion detection, SIGMOD Record, 43(3), 36–47, 2014.
[65] Alghawazi M., Alghazzawi D., Alarifi S., Detection of SQL injection attack using machine learning techniques: A systematic literature review, Journal of Cybersecurity and Privacy, 2(4), 764–777, 2022.
[66] Li X., Xue Y., A survey on server-side approaches to securing web applications, ACM Computing Surveys, 46(4), 1–29, 2014.
[67] Sabir B., Ullah F., Babar M.A., Gaire R., Machine learning for detecting data exfiltration: A review, ACM Computing Surveys, 54(3), 1–47, 2021.
[68] Chen Z., Liu J., Shen Y., Simsek M., Kantarci B., Mouftah H.T., Djukic P., Machine learning-enabled IoT security: Open issues and challenges under advanced persistent threats, ACM Computing Surveys, 55(5), 1–37, 2022.
[69] Ahmetoglu H., Das R., A comprehensive review on detection of cyber-attacks: Data sets, methods, challenges, and future research directions, Internet of Things, 20, 100615, 2022.
[70] Zhao J., Masood R., Seneviratne S., A review of computer vision methods in network security, IEEE Communications Surveys & Tutorials, 23(3), 1838–1878, 2021.
[71] Melacci S., Ciravegna G., Sotgiu A., Demontis A., Biggio B., Gori M., Roli F., Domain knowledge alleviates adversarial attacks in multi-label classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(12), 9944–9959, 2021.
[72] Xu M., Zhang T., Li Z., Zhang D., InfoAT: Improving adversarial training using the information bottleneck principle, IEEE Transactions on Neural Networks and Learning Systems, 35(1), 1255–1264, 2022.
[73] Yener B., Gal T., Cybersecurity in the era of data science: Examining new adversarial models, IEEE Security & Privacy, 17, 46–53, 2019.
[74] Vitorino J., Oliveira N., Praça I., Adaptative perturbation patterns: Realistic adversarial learning for robust intrusion detection, Future Internet, 14(4), 108, 2022.
[75] Rajan M., Choksey M., Jose J., Runtime Detection of Time-Delay Security Attack in System-an-chip, 15th IEEE/ACM International Workshop on Network on Chip Architectures, 2 October 2022, Chicago, USA.
[76] Rudrabhatla C.K., A quantitative approach for estimating the scaling thresholds and step policies in a distributed microservice architecture, IEEE Access, 8, 180246–180254, 2020.