Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Cyberattack detection and prevention framework for the healthcare sector using machine learning techniques

Ahmad Alshammari [1,*], Ali Alqarni [2]

[1]Department of Computer Sciences, Faculty of Computing and Information Technology, Northern Border University, Rafha, Saudi Arabia
[2]Department of Computer Science, College of Computing and Information Technology, University of Bisha, Bisha, Saudi Arabia

ARTICLE INFO

Article history:
Received 27 July 2025
Received in revised form
23 November 2025
Accepted 8 December 2025

Keywords:
Cybersecurity
Healthcare systems
Machine learning
Attack detection
Framework design

ABSTRACT

This paper presents a complete machine-learning framework for detecting and preventing cyberattacks in the healthcare sector. Because healthcare systems are highly vulnerable and data breaches can cause serious harm, the study seeks to address gaps in current solutions by developing an end-to-end model. Using a design science research approach, the framework includes five connected stages: data collection and preprocessing, data cleaning and feature selection, model training and evaluation, implementation and deployment, and continuous monitoring and improvement. The paper argues that this comprehensive approach, supported by comparisons with existing studies and an empirical analysis, offers a more effective and sustainable solution for healthcare cybersecurity than models that focus only on specific types of attacks.

1. Introduction

Cybercrime is a rapidly evolving field, frequently covered in the news, as it has been striving to stay ahead over the past decade. Although cyber-attacks have become more sophisticated with new techniques, their core intentions have remained unchanged (Ratta et al., 2021). Nowadays, all developing and developed countries are intensely focused on enhancing their healthcare systems, as healthcare directly impacts people's lives (Jalali et al., 2019). Healthcare is becoming an increasingly important area of concern as the sector continually evolves, integrating emerging, innovative, and computer-based technologies to enhance the efficiency of healthcare systems.

In today's increasingly digital world, society faces greater vulnerability to cybercrimes. Fig. 1 illustrates the various cyberattacks that pose a threat to the security and integrity of patient data. For instance, a Denial of Service (DoS) attack, rated with a high severity of 9 and considered a critical risk, can interrupt hospital operations by disabling networks and delaying essential care (Huseinović et al., 2020). Just like the former, Zero-Day attacks take advantage of weaknesses in a system, which allows them to access patient data without being authorized, which is a considerable risk factor. While malware and SQL injections pose moderate risks, both can enable the purchase of stolen patient data. Man-in-the-Middle (MITM) attacks can also intercept private communications between medical devices and a database, compromising data confidentiality. Additionally, phishing attacks, although less severe than ransomware, can deceive vulnerable healthcare professionals into revealing private information, resulting in data breaches. More than any other type of organization, healthcare providers find these cyber risks particularly challenging to confront, making it essential to address these issues to ensure smooth and seamless patient care.

Table 1 shows how cyber frameworks have been adopted in healthcare from 2000 to 2025. In 2000, basic data security measures were put in place to protect sensitive information and block unauthorized access on specific networks. Early on, security relied on tools like firewalls, antivirus programs, and simple access controls to prevent breaches. Regulations like HIPAA, introduced in 2005, along with stricter standards later, addressed these security issues. Today, healthcare data is generally encrypted, with access controls and audit trails regularly used to enhance data security.

As technology advanced, by 2010, healthcare systems faced increasingly sophisticated cyber

* Corresponding Author.
Email Address: ahmad.almkhaidsh@nbu.edu.sa (A. Alshammari)
https://doi.org/10.21833/ijaas.2026.01.001
Corresponding author's ORCID profile:
https://orcid.org/0009-0000-2051-2757

threats, including ransomware and advanced persistent threats (APTs). These new challenges required organizations to develop stronger security measures to detect and respond to such attacks quickly. One of the first instances of the WannaCry ransomware attack was its disruption of the healthcare system, which occurred in 2017. Since then, these attacks have exposed weaknesses in detecting potential threats within systems. In response to the 2019 attack, individuals created and implemented real-time, 24/7 networks and monitoring systems capable of detecting and responding to threats. Over the years, healthcare systems strengthened their defenses and began adopting the Zero Trust model. This means that any network, boundary, or access point could be a potential target. As insider threats emerged, a system began applying a closed breach approach for unapproved admissions and multi-factor verification. Over time, sensor-based systems have collected information, and breach systems have operated with integrated AI and machine learning under the Zero Trust model, leading to widespread success. By 2025, these systems are predicted to be fully integrated and functioning seamlessly within organized networks. AI is expected to be a core component of breach detection and prevention systems for healthcare, combined with integrated extraction systems. Your description of the evolution of threat detection and response accurately captures the ongoing efforts to improve security posture and the organization's security perimeter.

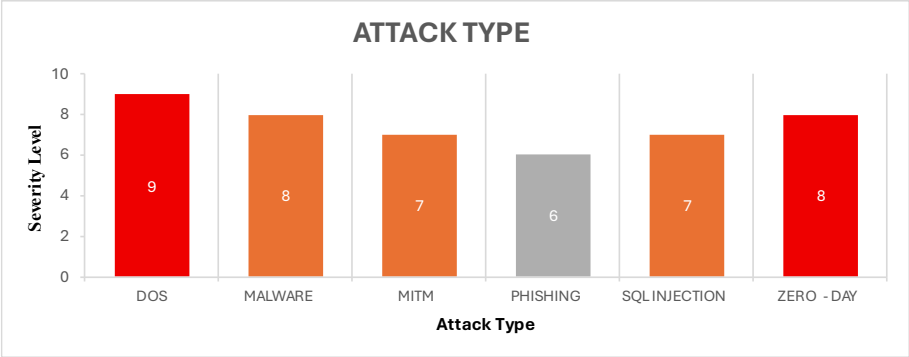

Fig. 1: Various cyberattacks can pose serious risks to patient information

Table 1: History of cybersecurity frameworks in healthcare from 2000 to 2025

Year	Event/development	Description
2000	Early frameworks	Initial cybersecurity frameworks focused on basic data protection and network security in healthcare.
2005	HIPAA updates	HIPAA introduced stricter rules for safeguarding patient data and electronic health records.
2010	Advanced threats emerge	Healthcare systems faced new challenges like ransomware and advanced persistent threats.
2015	Machine learning adoption	Machine learning began to be used for detecting and preventing cyberattacks in healthcare.
2017	WannaCry attack	Global ransomware attacks highlighted vulnerabilities in healthcare cybersecurity systems.
2019	Continuous monitoring	Frameworks emphasized real-time monitoring and proactive threat detection in healthcare.
2021	Zero-trust model	Zero Trust architecture gained traction, focusing on strict access controls and verification.
2023	Ai-driven frameworks	AI-powered systems became integral for analyzing threats and automating responses in healthcare.
2025	Integrated frameworks	Comprehensive frameworks that combine AI, machine learning, and Zero Trust principles have emerged.

Thus, a cybersecurity strategy helps protect computers, servers, mobile devices, electronic systems, networks, and interconnected infrastructures against malicious attacks. Recent attacks have disrupted the lives of hospital patients, including those caused by WannaCry and Ransomware. In addition to the IT vulnerability in infrastructures, social engineering exploits human weaknesses as well (Nifakos et al., 2021).

The three main pillars of authenticity for any network or device are confidentiality, integrity, and availability. In any industry, cybersecurity aims to protect these qualities. Patients' health information, especially their personally identifiable data, must be shielded from unauthorized access by individuals in the healthcare sector. Along with accuracy and trustworthiness, data integrity is maintained. We ensure that systems for processing and storing data are reliable. Security isn't just about protecting data; it's also about preserving patient safety, privacy, and trust. Healthcare systems, therefore, need to be secured by cybersecurity. This study has two main objectives: first, to examine different types of cyberattacks and the role of machine learning; and second, to develop a detailed framework for identifying and preventing cyberattacks in healthcare using machine learning methods. The research employs the design science methodology. Its primary contribution is the development of a comprehensive framework tailored to the healthcare sector, aiming to bridge the gap between existing cybersecurity measures and the specific needs of healthcare organizations. Existing cybersecurity frameworks for healthcare are frequently narrowly focused on specific attack types or systems, leading to a fragmented landscape of solutions that lack universal applicability.

In contrast, our framework presents an innovative, comprehensive approach through an end-to-end lifecycle management system. Its primary innovation does not rely on a single algorithm but instead on its holistic structure that seamlessly integrates five critical stages, from data preparation to deployment and continuous

adaptation, into a unified and practical blueprint for sustainable decisions.

2. Related works

2.1. Types of cyberattacks

This section provides a detailed explanation of the various types of cyberattacks illustrated in Fig. 2. There are seven (7) main common types of cyberattacks: denial of service attack (DOS), distributed denial of service attack (DDoS), malware, MITM, phishing, SQL injection, identity theft, zero-day exploits, social engineering, and insider threats.

A Denial-of-Service (DoS) attack attempts to prevent legitimate users from accessing a server or resource on a network. This type of attack is typically carried out in a distributed manner, involving an attacker taking control of multiple hosts, often without the victims' awareness, and instructing them to target the victim simultaneously (Oke et al., 2007). The DoS attacks come in several forms, as shown in Fig. 2: SYN Flooding, UDP Flooding, ICMP Flooding, DNS Amplification, HTTP Flooding, Volume-based Attacks, and Resource Exhaustion. The most popular and effective DDoS attack is SYN flooding (Mughaid et al., 2024). The UDP flooding attack occurs when a malicious system (bot) sends multiple UDP datagrams to the network simultaneously (Bijalwan et al., 2015). Many UDP datagrams flood the network at the same time, causing congestion when they reach a system. An ICMP flood attack is a common type of DoS attack where attackers overwhelm the target's resources with ICMP echo requests, potentially leading to resource exhaustion (Roshani and Nobakht, 2022). According to Kim et al. (2017), DNS amplification was one of the most malicious and disruptive attacks we must deal with. In March 2013, a massive DDoS attack of 300 Gbps was launched against the Spamhaus website, an organization dedicated to clearing spam from email inboxes. One of the most common DDoS attacks is HTTP flooding, which occurs when a group of attackers sends numerous HTTP requests to overwhelm a targeted server with an overwhelming number of requests from different directions (Mohammadi et al., 2023). Volume-based attacks are the most common type of DoS attacks. These attacks send many requests or data to the victim's server to overwhelm its bandwidth capability.

Unavailability is a significant consequence of this attack (Mohammadi and Babagoli, 2021). A resource exhaustion attack is a DoS attack that attempts to exhaust the resources of a target system or network (Hristozov et al., 2020). This attack compromises the system's performance, stability, and functionality by consuming all available resources, including CPU, memory, disk space, and network bandwidth.

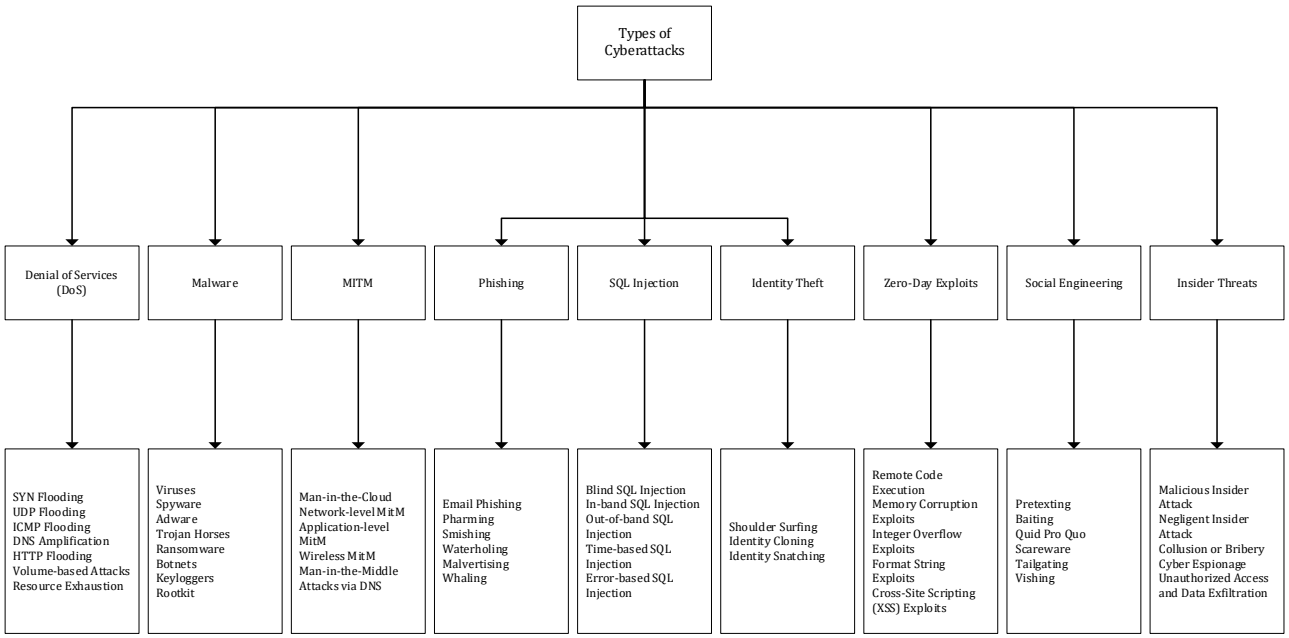

Fig. 2: Types of cyberattacks

Another common type of cyberattack is malware. The term "malware" refers to a type of cybersecurity risk in which malicious software (malware) infects a computer system, network, or device to cause damage, steal sensitive information, or disrupt normal operations. Malware attacks come in several forms, as shown in Fig. 2: viruses, spyware, adware, trojan horses, ransomware, botnets, keyloggers, and rootkits. Viruses are self-replicating programs designed to replicate themselves and infect other hosts, just like biological viruses do when they infect (Munjal and Puri, 2024). A spyware program is a type of malware that infects a victim's computer to gather and steal information about the victim and then transmit it back to the malware author without the victim's knowledge. Adware is software that makes advertisements, such as pop-up and banner advertisements, appear on your computer. As ransomware evolves from low levels of sophistication to more sophisticated tools used by

cybercriminals today, it targets and locks data, so the owner must pay to unlock it. Keyloggers and rootkits are serious cyber threats that can compromise your personal information and the security of your devices. It is essential to stay vigilant and take preventive measures to protect yourself from these malicious programs.

Another kind of cyberattack is the MITM. A man-in-the-middle attack occurs when someone listening in on the communication between two trusted parties steals sensitive information like passwords, credit card numbers, and personal identification numbers and misuses them (Sivasankari and Kamalakkannan, 2022). Several attack frames are illustrated in Fig. 2: Man-in-the-Cloud (MitC), Network-level MITM, Application-level MITM, Wireless MITM, and DNS-based Man-in-the-Middle attacks. The MitC attack involves an attacker obtaining unauthorized access to a system or network by exploiting cloud services (Medhioub and Hamdi, 2019). This attack enables the attacker to stay concealed and unnoticed, making it extremely difficult to defend against. The network-level MITM attack is carried out by delivering immediate illegal or incomplete data. In contrast, in service-level attacks, the target is filled with completed requests for services provided by the cloud service provider but with malicious intent. These two types of attack traffic must be filtered out at different levels.

Another kind of cyberattack is a phishing attack. A phishing attack refers to a type of network response attack where an attacker creates a fake copy of an existing website to fool an online user into providing their details (Gupta et al., 2016). A phishing attack comes in several forms: Email Phishing, Pharming, Smishing, Water-holing, Advertising, and Whaling. Email phishing is primarily known for attempting to steal personal and financial information. However, it is also used to compromise computers and IT networks at individual, business, and national levels. Pharming is a phishing attack wherein an attacker wants to steal sensitive information of internet users (Li et al., 2015). Smishing attacks typically involve scammers impersonating legitimate entities, such as financial institutions, government agencies, or well-known brands. This is a specific type of phishing attack that targets executive-level employees, such as directors, senior managers, etc. These 'Whales' will have higher privileges or access to vital data related to companies.

2.2. Machine learning

This section covers the history of machine learning. It is a core aspect of artificial intelligence, focused on developing algorithms that analyze historical data trends. Machine learning is utilized across many fields, such as bioinformatics, intrusion detection, information retrieval, gaming, marketing, malware detection, and image deconvolution.

In this study, the authors explore different types of machine learning and their associated algorithms, as shown in Fig. 3: Supervised learning, unsupervised learning, reinforcement learning, semi-supervised learning, transfer learning, active learning, and ensemble learning. For instance, supervised learning includes identifying patterns in data with algorithms and using these to predict future outcomes of dependent variables based on independent variables (Tiwari, 2022). The supervised learning algorithm is widely used in many fields, including Natural Language Processing, image analysis, and medical analysis. Unsupervised learning is an algorithm designed to create a learning paradigm for the sake of learning. In addition to having a rich, inherently structured dataset, unsupervised learning algorithms require sparse ground truth and metrics for training, as they are typically trained on sparse data. A significant branch of machine learning that is highly like how humans learn is reinforcement learning, which follows specific instructions. The Reinforcement Learning algorithm, first introduced by Samuel in 1959, has been effectively used in checkers, where it determines actions using a value function represented linearly. Semi-supervised learning is a machine learning technique that combines labelled and unlabeled data to learn and predict labels for unlabeled data. It uses labelled data to guide its learning process and incorporates unlabeled data to broaden its knowledge base and enhance prediction accuracy. Transfer learning involves machine learning that leverages supplementary information from related tasks, which is not present in typical training datasets.

2.3. Problem statement

Healthcare cyberattacks pose a significant threat to patient safety, privacy, and the quality of care. They can cause data breaches, unauthorized access to sensitive information, and harm medical devices. To address these risks, the healthcare sector requires a strong framework that employs machine learning to detect and prevent cybersecurity threats learning.

3. Methodology

This study's detection and prevention framework was developed using a design science research methodology, which encompasses a broad range of research methods linked to the design science paradigm (Al-Mugerrn et al., 2023). Considering a variety of research methodologies, which range from IT to anthropology and others, the research methodology encapsulates several guidelines for project evaluation and iteration that are specific to each discipline. In this study, the author adapted the development methodology from Alotaibi et al. (2022). Fig. 4 displays the adapted methodology.

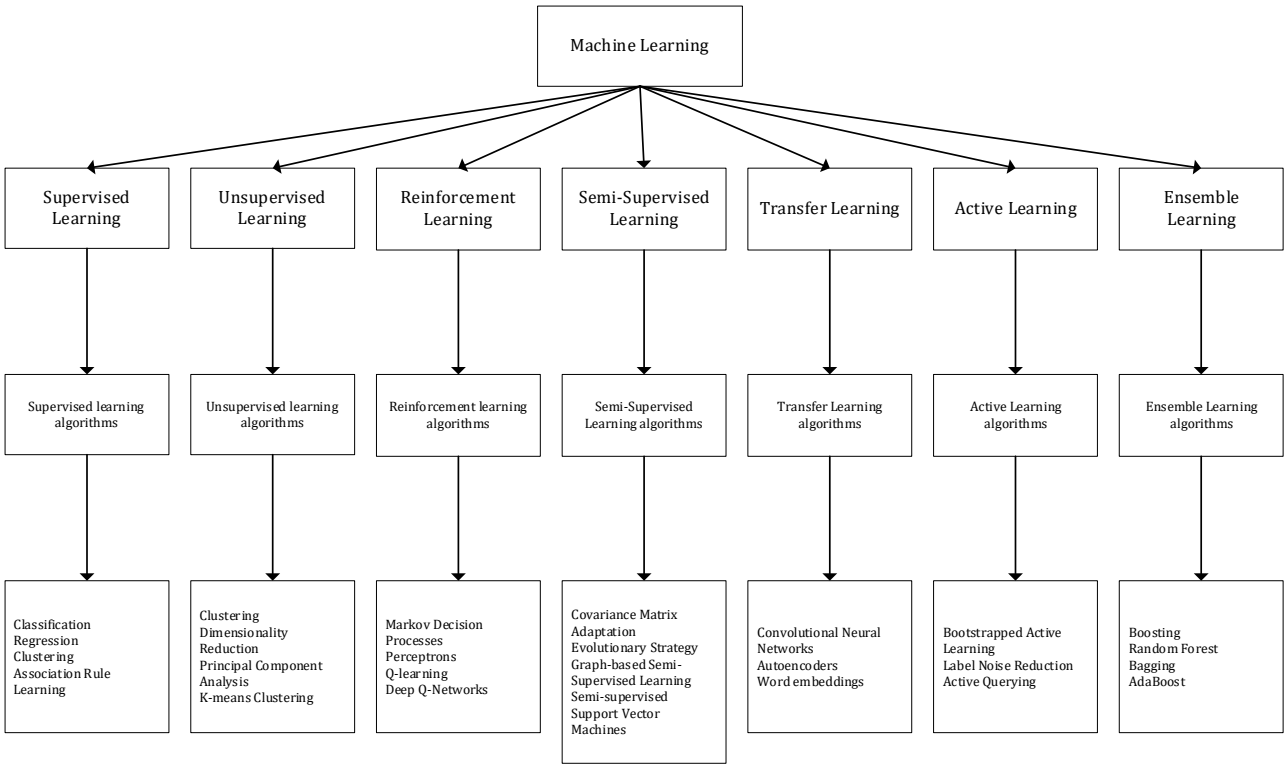

Fig. 3: Types of machine learning

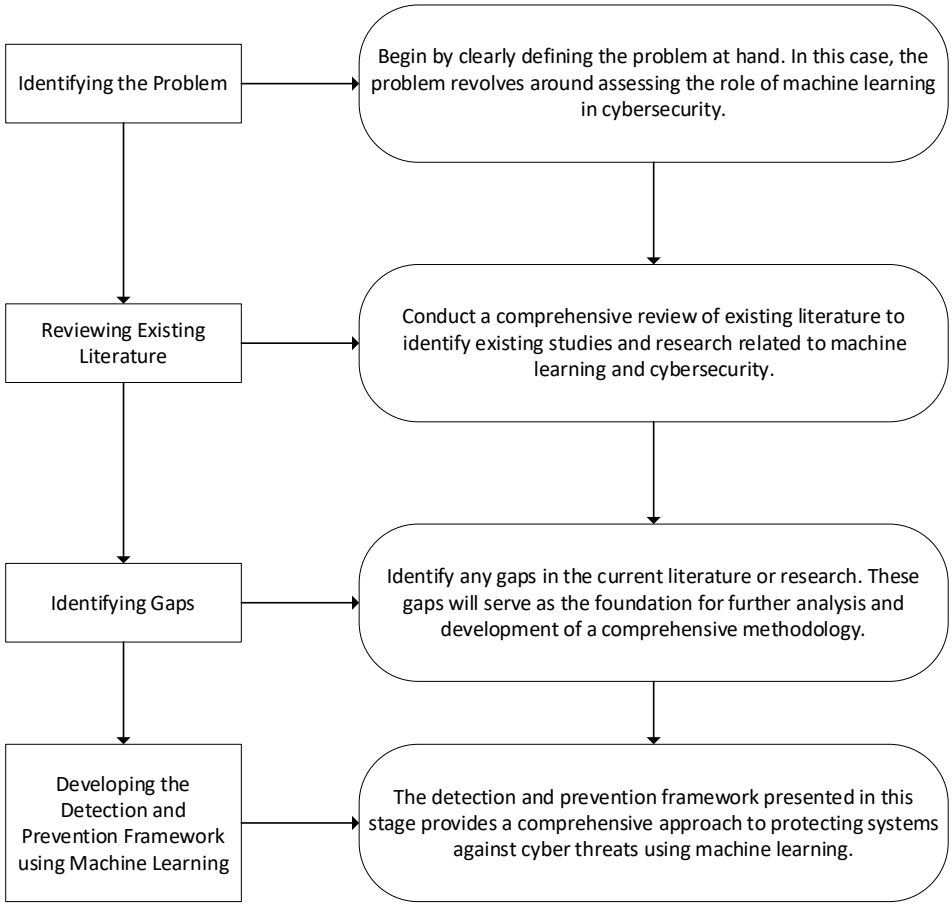

Fig. 4: Adapted methodology (Alotaibi et al., 2022)

1. Identifying the Problem: This stage begins the design science research methodology. It identifies the research problem, along with the justification and significance of the proposed solution. As outlined in Section 2, which aligns with the study's objectives, this phase provides an overview of the main research background topics discussed in the paper. The review begins with an examination of cyberattacks and the machine learning field to identify core issues. The root cause of the problem was identified during this process. The machine learning domain was then examined further,

focusing on models, mechanisms, approaches, methods, tools, processes, activities, operations, concepts, and terminology. Knowledge in machine learning is recognized as valuable, diverse, complex, and scattered across various sources such as journals, conferences, experts, books, chapters, magazines, dissertations, reports, and online resources

2. Reviewing Existing Literature: During this phase, the authors conduct a comprehensive review of existing literature on how machine learning is used to detect and prevent cyber-attacks. To gather resources, they utilized five online databases: Scopus, IEEE Xplore, Web of Science, Springer, and Google Scholar. They collected various types of literature, including review articles, research papers, conference papers, book chapters, and reports from these sources. The study excludes duplicate articles, articles without results, and low-quality publications from the analysis.
3. Identifying Gaps: Based on the analysis of the focused articles on detecting and preventing cyber-attacks with machine learning, the study found that traditional methods are inadequate for identifying and stopping various types of attacks. A summary of the selected articles is available in Table 2.
4. Developing the Detection and Prevention Framework using Machine Learning: This section presents the proposed framework for identifying and preventing healthcare cyberattacks using Machine Learning.

The framework consists of five stages, as illustrated in Fig. 5. Each stage includes multiple tasks and activities.

1. Data Collection and Preprocessing: The goal of this phase is to gather diverse healthcare cybersecurity data and convert it into a structured format for analysis.
2. Data Cleaning and Feature Selection: Data cleaning involves identifying and rectifying any inconsistencies, missing values, or irrelevant information. Feature selection consists of selecting the most relevant features from raw data to improve the performance of machine learning algorithms.
3. Model Training and Evaluation: Training consists of processing training data and applying different ML algorithms to detect patterns and connections that indicate healthcare cybersecurity attacks. Evaluation involves testing these models on new, unseen data to measure their performance effectiveness.
4. Implementation and Deployment: Successfully deploying ML models necessitates integrating them with current security infrastructure and systems. Continuous monitoring and updates are crucial to keep the models effective against emerging threats and attack methods.
5. Continuous Monitoring and Improvement: Healthcare cybersecurity is a continuous effort. Regularly monitoring the performance of ML models and adapting to emerging threats and attack patterns are essential to keeping the framework effective.

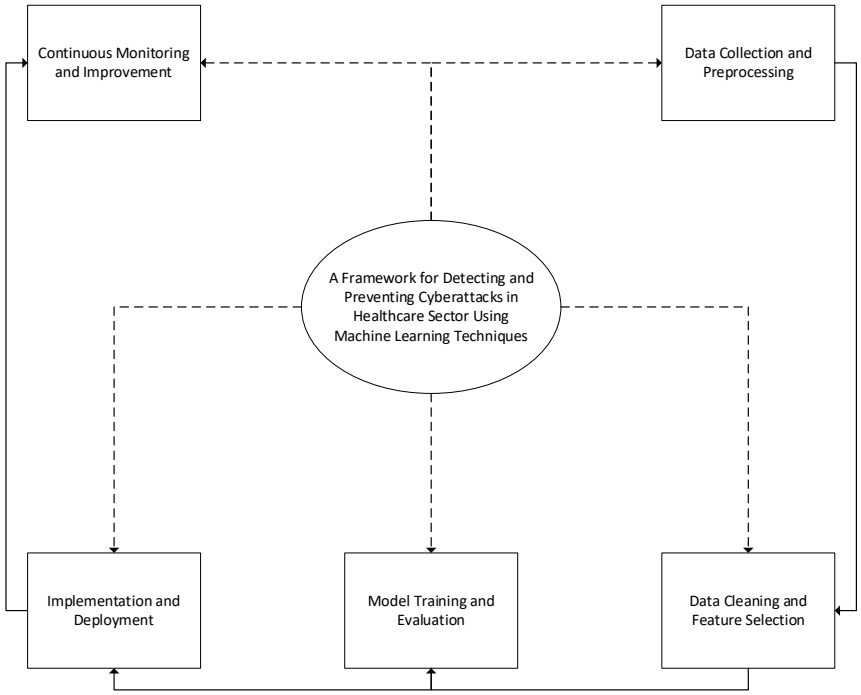

Fig. 5: A framework for detecting and preventing healthcare cyberattacks using machine learning

Table 2: A summary of the articles discussing healthcare cyberattack detection and prevention

Reference	Strengths	Limitations	Methodology	Contributions
Sharma et al. (2022)	Explains the link between digital marketing and computer science and proposes a machine learning framework to improve marketing strategies.	Few studies exist in this area, and privacy and data collection remain challenging.	Analysis of digital marketing techniques using deep learning concepts.	Identifies key technical and business decision-making factors using deep learning.
Banks et al. (2023)	Combines machine learning, biophysical modeling, and automation to address neural engineering problems.	Limited to VHA patients, with possible bias from retrospective data.	Machine learning applied to EHR data and biophysical neuron modeling.	Develops a prediction tool for opioid use disorder and a realistic amygdala model.
Haddad et al. (2022)	Reviews AI and blockchain integration in healthcare with a focus on security, privacy, and interoperability.	Mainly based on literature with limited real-world implementation.	Systematic review of 113 studies using PRISMA guidelines.	Presents benefits, security challenges, and a taxonomy for AI–blockchain EHR systems.
Pardakhe and Deshmukh (2019)	Combines blockchain security with hybrid deep learning for healthcare data analysis.	Scalability, privacy, interoperability, and regulatory challenges remain.	Permissioned blockchain integrated with hybrid deep learning.	Improves security, scalability, and decision-making in healthcare systems.
Pandey and Litoriya (2020)	Improves medical record security, privacy, and access using blockchain.	Mostly theoretical with limited real-world testing.	Hyperledger-based permissioned blockchain EMR system.	Proposes a secure blockchain-based EMR management framework.
Alsinglawi et al. (2022)	Develops an explainable ML model for ICU length-of-stay prediction.	Small dataset and lack of external validation.	Supervised ML with feature selection, imbalance handling, and SHAP explainability.	Introduces an explainable framework for ICU stay prediction.
Papadopoulos et al. (2021)	Improves privacy and trust in federated learning using decentralized identity.	Machine learning tuning and classification are not fully addressed.	Hyperledger-based decentralized identity and federated learning framework.	Establishes secure trust architecture for federated healthcare data sharing.
Chaithra and Vagdevi (2021)	Reviews blockchain-based EHR security and privacy solutions.	No evaluation of specific algorithms.	Systematic review from four databases (2015–2020).	Classifies EHR security challenges and protection mechanisms.
Kempa-Liehr et al. (2020)	Combines process mining and probabilistic learning to predict patient recovery.	Limited dataset and unverified modeling assumptions.	ProM process mining with probabilistic regression.	Predicts recovery times with interpretable care pathways.
Taylor et al. (2020)	Provides a scalable machine learning pipeline for medical predictions.	Data access restrictions limit cost analysis.	Random forest-based ML data pipeline.	Supports machine learning-based clinical decisions.
Gupta et al. (2023)	Uses high-throughput ML to predict prescriptions and healthcare costs.	Prediction accuracy decreases over longer time windows.	Random forest ML pipeline on de-identified EHRs.	Demonstrates ML-based healthcare cost estimation.
Arjun and Kumar (2020)	Highlights the role of machine learning in diagnostics, treatment, and drug discovery.	Data quality, bias, and regulation remain major challenges.	Comprehensive literature review on ML in healthcare.	Summarizes key ML applications and future research directions.
Wang and Sun (2022)	Generates realistic and privacy-preserving synthetic EHR data.	Limited evaluation against advanced adversarial privacy attacks.	Language model-based synthetic EHR generation.	Introduces the PromptEHR framework for secure EHR generation.
Yeng et al. (2020)	Compares ML methods for detecting anomalies in EHR access logs.	Uses simulated logs with low detection precision.	Eight ML classification methods for anomaly detection.	Evaluates ML performance for healthcare security monitoring.
Dutta and Bandyopadhyay (2021)	Reviews AI, ML, DL, and IoT integration in healthcare.	Data scarcity, security, and validation challenges remain.	Survey of ML, DL, IoT, and imaging-based healthcare systems.	Proposes integrated AI-driven healthcare frameworks.
Kumar et al. (2022)	Proposes an ensemble ML model for early breast cancer detection.	Validation limited to a single public dataset.	Optimized stacking ensemble using a genetic algorithm.	Improves diagnostic accuracy for breast cancer detection.
Rani et al. (2023)	Reviews AI and quantum ML techniques for heart disease diagnosis.	Data bias, privacy, and interpretability challenges.	Literature survey on AI and QML heart disease models.	Summarizes AI and QML advances in cardiovascular diagnosis.
Kumar and Padmapriya (2016)	Proposes a matrix-based method for disease information extraction.	Strong reliance on structured ICD-10 data.	Matrix-based disease extraction framework.	Introduces the DIECTA disease extraction technique.
Sun et al. (2022)	Uses EHR data to identify pneumonia risk factors.	Does not include free-text clinical data.	ML using CART and logistic regression.	Identifies key predictors of pneumonia.
Chen et al. (2022)	Predicts hypertension risk using ML in a Chinese population.	Cross-sectional design limits causal conclusions.	XGBoost, RF, and logistic regression models.	Identifies lifestyle and demographic hypertension risk factors.
Hussain et al. (2015)	Combines clinical knowledge and ML for heart disease diagnosis.	Focuses mainly on chest pain cases.	Ontology-based and ML-driven decision support system.	Accurately classifies cardiac and non-cardiac conditions.
Khan et al. (2022)	Uses blockchain and ML to enhance healthcare data security.	Relies mainly on simulations.	Blockchain architecture with ML-based SGD optimization.	Proposes secure blockchain–ML healthcare architecture.
Tenepalli and Thandava Meganathan (2023)	Comprehensive review of IoT, blockchain, and ML in healthcare.	Data storage, connectivity, and security challenges remain.	Review of 65 studies on IoT, blockchain, and ML integration.	Summarizes combined technology applications in healthcare.
Kumawat et al. (2022)	Reviews ML applications and ethical challenges in healthcare.	Theoretical focus without empirical validation.	Systematic review of 33 studies on ML in healthcare.	Proposes ethical and regulatory guidance for ML healthcare use.
Tumpa and Dey (2022)	Reviews ML applications in diagnosis, telemedicine, and drug discovery.	Lacks technical depth and empirical testing.	Design Science Research and systematic review.	Links ML methods to healthcare problem-solving.
Nasayreh et al. (2025)	Combines LSTM, PCA, and KNN for high-accuracy IoMT cyberattack detection.	Dataset scarcity and computational complexity issues remain.	Hybrid LSTM–PCA–KNN model.	Introduces an effective cyberattack detection model for IoMT.
Abbas et al. (2025)	Combines blockchain and ML to strengthen healthcare security.	Blockchain scalability, cost, and data imbalance issues.	Survey-based ML analysis using RF and SVM.	Proposes a hybrid blockchain–ML security framework.

4. Validation

This section aims to validate the developed framework for its completeness and comprehensiveness. Comparison with other models is used to evaluate the development. Table 3 provides a comparison between the developed framework and existing works. Analysis of Table 3 reveals a clear trend in current research on machine learning in healthcare cybersecurity. Nearly all cited studies focus on the early stages of the machine learning process, especially data collection and preprocessing. This focus is essential for any machine learning project, as it ensures access to diverse datasets that can be prepared for analysis. However, a significant gap is apparent in the discussion of later phases.

Although many studies address data cleaning and feature selection, only a few show significant engagement in this area. This stage involves identifying inconsistencies, managing missing data, and choosing the most relevant features, which are essential for improving model performance. All papers include model training and evaluation. However, there is a notable lack of research on the practical aspects of implementing and deploying models in real-world healthcare settings, and none examine primary processes like continuous monitoring and improvement.

Compared to the comprehensive framework that includes all previously described steps, the papers cited in Table 3 reveal numerous individual models and techniques. However, a complete approach covering the entire cybersecurity system lifecycle, from data collection to real-world deployment and ongoing maintenance, is missing. This finding supports a key goal for the article. The analysis shows that although many studies focus on specific ML methods for healthcare cybersecurity, the critical tasks of implementation, deployment, and continuous upkeep are largely ignored. It highlights the importance of the developed framework as a more integrated solution and underscores the need for future research to bridge the gap between theory and practical application.

Table 3: Compare the developed framework with the existing works

ID	Reference	Data collection and preprocessing	Data cleaning and feature selection	Model training and evaluation	Implementation and deployment	Continuous monitoring and improvement
1	Sharma et al. (2022)	Yes	No	Yes	No	No
2	Banks et al. (2023)	Yes	Yes	Yes	No	No
3	Haddad et al. (2022)	Yes	No	Yes	No	No
4	Pardakhe and Deshmukh (2019)	Yes	No	Yes	No	No
5	Pandey and Litoriya (2020)	Yes	No	Yes	No	No
6	Alsinglawi et al. (2022)	Yes	Yes	Yes	No	No
7	Papadopoulos et al. (2021)	Yes	No	Yes	No	No
8	Chaithra and Vagdevi (2021)	Yes	No	Yes	No	No
9	Kempa-Liehr et al. (2020)	Yes	No	Yes	No	No
10	Taylor et al. (2020)	Yes	No	Yes	No	No
11	Gupta et al. (2023)	Yes	No	Yes	No	No
12	Arjun and Kumar (2020)	Yes	No	Yes	No	No
13	Wang and Sun (2022)	Yes	No	Yes	No	No
14	Yeng et al. (2020)	Yes	Yes	Yes	No	No
15	Dutta and Bandyopadhyay (2021)	Yes	No	Yes	No	No
16	Kumar et al. (2022)	Yes	No	Yes	No	No
17	Rani et al. (2023)	Yes	No	Yes	No	No
18	Kumar and Padmapriya (2016)	Yes	No	Yes	No	No
19	Sun et al. (2022)	Yes	No	Yes	No	No
20	Chen et al. (2022)	Yes	No	Yes	No	No
21	Hussain et al. (2015)	Yes	No	Yes	No	No
22	Khan et al. (2022)	Yes	No	Yes	No	No
23	Tenepalli and Thandava Meganathan (2023)	Yes	No	Yes	No	No
24	Kumawat et al. (2022)	Yes	No	Yes	No	No
25	Tumpa and Dey (2022)	Yes	No	Yes	No	No
26	Nasayreh et al. (2025)	Yes	Yes	Yes	No	No
27	Abbas et al. (2025)	Yes	No	Yes	No	No

5. Empirical validation

This section demonstrates the practical effectiveness of the proposed framework through an empirical study using a real-world healthcare cybersecurity dataset. The study follows the main stages of the framework, beginning with data collection and ending with model evaluation. The dataset used in this study is the Waikato Internet Traffic Storage Dataset obtained from the University of Waikato. It contains real network traffic collected from a healthcare environment and includes both benign and malicious activities that were generated in a controlled and ethical setting. The dataset consists of packet capture files and corresponding flow-based records generated using CICFlowMeter. It covers several cyberattack types relevant to healthcare systems, including distributed denial-of-

service attacks, brute-force attacks on patient portals, and web-based intrusions. In total, the dataset contains approximately 2.5 million network flow records, each described by more than 80 features such as flow duration, protocol type, packet size statistics, flags, and timing information. Each flow is labeled as either benign or as a specific attack type such as DDoS, brute-force, or web attack.

The proposed framework is applied in three main stages. In the first stage, data collection and preprocessing were carried out by converting the raw packet capture files into structured flow-based records using CICFlowMeter. This process transformed raw network traffic into a machine learning–ready CSV format, in which each row represents a unique network flow and its corresponding characteristics. In the second stage, data cleaning and feature selection were performed. During cleaning, records containing missing or infinite values were removed, which represented about 1.5 percent of the total data. All numerical features were then scaled using the RobustScaler to reduce the influence of outliers. For feature selection, the large number of available features created a risk of overfitting, so a two-step process was applied. First, correlation analysis was conducted and features with a correlation coefficient greater than 0.95 were removed to eliminate redundancy. Second, a Random Forest model was trained to rank feature importance, and the top 30 most informative features were retained for classification. These features included backward packet length maximum, flow duration, total length of forward packets, and subflow forward bytes. In the third stage, model training and evaluation were conducted using four different machine learning algorithms to demonstrate the flexibility of the framework. These models included Random Forest, XGBoost, a multi-layer perceptron neural network, and a support vector machine. The dataset was divided into 70 percent for training and 30 percent for testing. A five-fold cross-validation strategy was applied to the training data for hyperparameter tuning using GridSearchCV. Because the dataset contains a class imbalance with more benign traffic than malicious traffic, the models were evaluated using accuracy, precision, recall, and F1 score to ensure a reliable performance assessment. The optimized models achieved strong results on the test set, as reported in Table 4. For the best-performing model, which was the Random Forest classifier, the confusion matrix shows that 547,200 benign flows were correctly classified as benign, while 2,300 benign flows were misclassified as attacks. At the same time, 196,400 attack flows were correctly detected, and 4,100 attack flows were incorrectly classified as benign. The Random Forest model achieved an F1-score of 98.1 percent, indicating an excellent balance between precision and recall. The high recall rate of 97.8 percent is especially critical in healthcare environments, where failing to detect an attack could result in serious risks to patient safety and data privacy. The strong performance of all tested models also confirms the effectiveness of the feature selection stage and shows that the selected features were highly discriminative. These results provide strong empirical support for the claim that machine learning techniques are highly effective for detecting cyberattacks in healthcare networks.

Although full real-time deployment is beyond the scope of this study, a proof-of-concept implementation was designed for the next stages of the framework. The trained Random Forest model was serialized and integrated into a simple Python-based REST API using the Flask framework. This API is capable of receiving real-time flow data from a network sensor, processing it through the trained model, and returning immediate predictions. For continuous monitoring, an MLOps-based pipeline is proposed to regularly log model predictions and confidence scores, detect performance degradation on new unseen data, and automatically trigger model retraining using newly collected data. This process enables continuous improvement and adaptation to evolving cyber threats.

Overall, this empirical validation confirms that the proposed framework is not only theoretical but also practical and effective for building robust machine learning–based cybersecurity systems in healthcare environments. The high level of performance achieved on real-world data highlights the strong potential of this approach to significantly enhance patient data security.

Table 4: Model Performance on Healthcare Network Intrusion Detection Task

Model	Accuracy	Precision	Recall	F1-Score
Random forest (RF)	99.2%	98.5%	97.8%	98.1%
XGBoost (XGB)	99.0%	98.1%	97.5%	97.8%
Multi-layer perceptron (MLP)	98.5%	97.2%	96.0%	96.6%
Support vector machine (SVM)	97.8%	96.0%	95.2%	95.6%

6. Discussion and analysis

This paper convincingly argues for a paradigm shift in how ML is used in healthcare cybersecurity. The main point is that current research is too fragmented, mainly focusing on the early, theoretical parts of the ML process, specifically data collection and model training. This narrow focus results in a significant "theory-practice gap," where advanced models are developed but rarely transition into practical, real-world security systems. To address this, the authors introduced a comprehensive five-stage framework meant to cover the entire ML lifecycle. This framework is the paper's key contribution, designed to guide the process from initial data collection and processing through feature selection, model training, and, most importantly, into the critical phases of practical implementation,

deployment, and ongoing monitoring and improvement. This holistic approach is presented as an integrated plan, intended to move beyond isolated algorithms toward a strong, sustainable defense system.

This framework's validation has two main parts, forming the empirical basis of the discussion. First, a systematic review of 27 recent studies highlights the research gap. This review shows that while all studied papers involve data collection and model training, only a few focus on feature selection, and none extensively cover implementation, deployment, or ongoing monitoring. This notable gap emphasizes the need for the authors' broader approach. Second, the practical effectiveness of the first three stages of the framework is validated through an empirical study using the Waikato Internet Traffic Storage (WITS) dataset, a real-world healthcare network data set. The findings are impressive: A Random Forest model achieves 99.2% accuracy and a 98.1% F1-Score, demonstrating the framework's capability for effective cyberattack detection.

A detailed analysis indicates that, although the paper's scholarly contribution is significant, its practical claims are somewhat premature. Its primary strength lies in effectively identifying a major flaw in the field and proposing a logical, well-structured architectural solution. The emphasis on the full ML lifecycle (MLOps) is both appropriate and essential. Nonetheless, the framework's validation highlights key limitations. The implementation stage is presented merely as a proof-of-concept API, which is far from the complex task of integrating an ML model into legacy clinical systems such as Epic or Cerner, while ensuring HIPAA compliance and maintaining real-time, low-latency performance without disrupting critical care. Similarly, the continuous monitoring stage is depicted as a theoretical concept rather than an empirically tested component, ignoring the significant complexities involved in automated retraining and drift detection. Moreover, the framework's evaluation is limited to network intrusion scenarios, leaving its applicability to other major healthcare threats, such as insider threats or medical device hijacking, unverified.

7. Conclusions

This study presented a comprehensive machine learning framework for detecting and preventing cyberattacks in healthcare. Its main contribution is an end-to-end lifecycle approach that encompasses five essential stages, from data collection to continuous monitoring, addressing a gap in existing solutions that often focus solely on specific parts of the process. Validation with a real healthcare network dataset demonstrated high effectiveness, with models like Random Forest achieving 99.2% accuracy and an F1 score of 98.1%. The study highlights that, while many existing works focus on model development, the practical aspects of implementation, deployment, and ongoing adaptation are crucial for real-world success and are frequently overlooked. The framework offers a valuable, scalable model for enhancing cybersecurity in healthcare, thereby protecting patient data, safeguarding privacy, and ensuring critical operations remain secure against emerging cyber threats. Future efforts will involve full deployment in real-world settings and further enhancements to continuous monitoring and retraining systems.

List of abbreviations

AI	Artificial intelligence
API	Application programming interface
APTs	Advanced persistent threats
CART	Classification and regression trees
CICFlowMeter	A tool used for network flow analysis from the University of Waikato
CPU	Central processing unit
CSV	Comma-separated values
DDoS	Distributed denial of service
DIECTA	Disease information extraction using CTA matrix
DL	Deep learning
DNS	Domain name system
DoS	Denial of service
DSR	Design science research
EHR	Electronic health record
EMR	Electronic medical record
F1-score	Harmonic mean of precision and recall
FL	Federated learning
GridSearchCV	Grid search cross-validation
HIPAA	Health insurance portability and accountability act
HTTP	Hypertext transfer protocol
ICMP	Internet control message protocol
ICU	Intensive care unit
IoT	Internet of Things
IoMT	Internet of medical things
IT	Information technology
KNN	K-nearest neighbors
LOS	Length of stay
LR	Logistic regression
LRTI	Lower respiratory tract infection
LSTM	Long short-term memory
MITM	Man-in-the-middle
ML	Machine learning
MLDPS	Machine learning-driven prognostic system
MLOps	Machine learning operations
MLP	Multi-layer perceptron
ODCRARS	Ontology-driven clinical risk assessment rule-based system
OSEL	Optimized stacking ensemble learning
PCA	Principal component analysis
pcap	Packet capture
QML	Quantum machine learning
REST	Representational state transfer
RF	Random forest
RTI	Respiratory tract infection
SDN	Software-defined networking
SGD	Stochastic gradient descent
SLR	Systematic literature review
SMOTE	Synthetic minority over-sampling technique
SQL	Structured query language
SVM	Support vector machine
SYN	Synchronize
UDP	User datagram protocol

VHA	Veterans health administration
WITS	Waikato internet traffic storage
XGB / XGBoost	Extreme gradient boosting

Acknowledgment

The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA, for funding this research work through the project number "NBU-FFR-2025-2990-08. "The authors are thankful to the Deanship of Graduate Studies and Scientific Research at the University of Bisha for supporting this work through the Fast-Track Research Support Program.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Abbas R, Ogunsanya VA, Nwanyim SJ, Afolabi R, Kagame R, Akinsola A, and Clement T (2025). Leveraging machine learning to strengthen network security and improve threat detection in blockchain for healthcare systems. International Journal of Scientific and Management Research, 8(2): 147-165. https://doi.org/10.37502/IJSMR.2025.8211

Al-Mugerrn R, Al-Dhaqm A, and Othman SH (2023). A metamodeling approach for structuring and organizing cloud forensics domain. In the International Conference on Smart Computing and Application (ICSCA), IEEE, Hail, Saudi Arabia: 1-5. https://doi.org/10.1109/ICSCA57840.2023.10087425

Alotaibi FM, Al-Dhaqm A, Al-Otaibi YD, and Alsewari AA (2022). A comprehensive collection and analysis model for the drone forensics field. Sensors, 22(17): 6486. https://doi.org/10.3390/s22176486 **PMid:36080945 PMCid:PMC9460793**

Alsinglawi B, Alshari O, Alorjani M, Mubin O, Alnajjar F, Novoa M, and Darwish O (2022). An explainable machine learning framework for lung cancer hospital length of stay prediction. Scientific Reports, 12: 607. https://doi.org/10.1038/s41598-021-04608-7 **PMid:35022512 PMCid:PMC8755804**

Arjun KP and Kumar KS (2020). Machine learning-A neoteric medicine to healthcare. International Journal on Emerging Technologies, 11(3): 195-201.

Banks TJ, Nguyen TD, Uhlmann JK, Nair SS, and Scherrer JF (2023). Predicting opioid use disorder before and after the opioid prescribing peak in the United States: A machine learning tool using electronic healthcare records. Health Informatics Journal, 29(2). https://doi.org/10.1177/14604582231168826 **PMid:37042333 PMCid:PMC10158959**

Bijalwan A, Wazid M, Pilli ES, and Joshi RC (2015). Forensics of random-UDP flooding attacks. Journal of Networks, 10(5): 287-293. https://doi.org/10.4304/jnw.10.5.287-293

Chaithra MH and Vagdevi S (2021). A detailed survey study on various issues and techniques for security and privacy of healthcare records. In: Raj JS, Palanisamy R, Perikos I, and Shi Y (Eds.), Intelligent Sustainable Systems. Lecture Notes in Networks and Systems, 213: 181–189. Springer, Singapore, Singapore. https://doi.org/10.1007/978-981-16-2422-3_15

Chen N, Fan F, Geng J, Yang Y, Gao Y, Jin H, Chu Q, Yu D, Wang Z, and Shi J (2022). Evaluating the risk of hypertension in residents in primary care in Shanghai, China with machine learning algorithms. Frontiers in Public Health, 10: 984621. https://doi.org/10.3389/fpubh.2022.984621 **PMid:36267989 PMCid:PMC9577109**

Dutta S and Bandyopadhyay SK (2021). Diabetes prediction using machine learning approaches. In: Roy S, Goyal LM, and Mittal M (Eds.), Advanced prognostic predictive modelling in healthcare data analytics: 179-202. Springer, Singapore, Singapore. https://doi.org/10.1007/978-981-16-0538-3_9

Gupta S, Nama GF, and Deivasigamani S (2023). Real-time monitoring of patient activity using IoT and machine learning in healthcare. International Journal of Intelligent Systems and Applications in Engineering, 11(7s): 51-57.

Gupta S, Singhal A, and Kapoor A (2016). A literature survey on social engineering attacks: Phishing attack. In the International Conference on Computing, Communication and Automation (ICCCA), IEEE, Greater Noida, India: 537-540. https://doi.org/10.1109/CCAA.2016.7813778

Haddad A, Habaebi MH, Islam MR, Hasbullah NF, and Zabidi SA (2022). Systematic review on AI-blockchain based e-healthcare records management systems. IEEE Access, 10: 94583-94615. https://doi.org/10.1109/ACCESS.2022.3201878

Hristozov S, Huber M, and Sigl G (2020). Protecting RESTful IoT devices from battery exhaustion DoS attacks. In the IEEE International Symposium on Hardware Oriented Security and Trust (HOST), IEEE, San Jose, USA: 316-327. https://doi.org/10.1109/HOST45689.2020.9300290

Huseinović A, Mrdović S, Bicakci K, and Uludag S (2020). A survey of denial-of-service attacks and solutions in the smart grid. IEEE Access, 8: 177447-177470. https://doi.org/10.1109/ACCESS.2020.3026923

Hussain A, Farooq K, Luo B, and Slack W (2015). A novel ontology and machine learning inspired hybrid cardiovascular decision support framework. In the IEEE Symposium Series on Computational Intelligence, IEEE, Cape Town, South Africa: 824-832. https://doi.org/10.1109/SSCI.2015.122

Jalali MS, Razak S, Gordon W, Perakslis E, and Madnick S (2019). Health care and cybersecurity: Bibliometric analysis of the literature. Journal of Medical Internet Research, 21(2): e12644. https://doi.org/10.2196/12644 **PMid:30767908 PMCid:PMC6396074**

Kempa-Liehr AW, Lin CYC, Britten R, Armstrong D, Wallace J, Mordaunt D, and O'Sullivan M (2020). Healthcare pathway discovery and probabilistic machine learning. International Journal of Medical Informatics, 137: 104087. https://doi.org/10.1016/j.ijmedinf.2020.104087 **PMid:32126509**

Khan AA, Laghari AA, Shafiq M, Cheikhrouhou O, Alhakami W, Hamam H, and Shaikh ZA (2022). Healthcare ledger management: A blockchain and machine learning-enabled novel and secure architecture for medical industry. Human-Centric Computing and Information Sciences, 12: 55. https://doi.org/10.22967/HCIS.2022.12.055

Kim S, Lee S, Cho G, Ahmed ME, Jeong J, and Kim H (2017). Preventing DNS amplification attacks using the history of DNS queries with SDN. In: Foley S, Gollmann D, and Snekkenes E (Eds.), Computer Security – ESORICS 2017. Lecture Notes in Computer Science, 10493: 135–152. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-66399-9_8

Kumar LS and Padmapriya A (2016). Disease information extraction from healthcare records using CTA matrix. Australian Journal of Basic and Applied Sciences, 10(2): 141-149.

Kumar M, Singhal S, Shekhar S, Sharma B, and Srivastava G (2022). Optimized stacking ensemble learning model for breast cancer detection and classification using machine learning.

Sustainability, 14(21): 13998. https://doi.org/10.3390/su142113998

Kumawat V, Umamaheswari B, Mitra P, and Lavania G (2022). Machine learning for health care: challenges, controversies, and its applications. In: Kumar R, Ahn CW, Sharma TK, Verma OP, and Agarwal A (Eds.), Soft computing: Theories and applications: Proceedings of SoCTA 2021: 253-261. Springer Nature, Singapore, Singapore. https://doi.org/10.1007/978-981-19-0707-4_24

Li Y, Chu S, and Xiao R (2015). A pharming attack hybrid detection model based on IP addresses and web content. Optik, 126(2): 234-239. https://doi.org/10.1016/j.ijleo.2014.10.001

Medhioub M and Hamdi M (2019). An identity-based cryptographic scheme for cloud storage applications. International Journal of Grid and Utility Computing, 10(2): 93-104. https://doi.org/10.1504/IJGUC.2019.10018608

Mohammadi R, Lal C, and Conti M (2023). HTTPScout: A machine learning based countermeasure for HTTP flood attacks in SDN. International Journal of Information Security, 22: 367-379. https://doi.org/10.1007/s10207-022-00641-3

Mohammadi S and Babagoli M (2021). A hybrid modified grasshopper optimization algorithm and genetic algorithm to detect and prevent DDoS attacks. International Journal of Engineering, 34(4): 811-824. https://doi.org/10.5829/ije.2021.34.04a.07

Mughaid A, Alnajjar A, El-Salhi SM, Almakadmeh K, and AlZu'bi S (2024). A cutting-edge intelligent cyber model for intrusion detection in IoT environments leveraging future generations networks. Cluster Computing, 27: 10359-10375. https://doi.org/10.1007/s10586-024-04495-3

Munjal G and Puri T (2024). Analysis of malicious executables and detection techniques. In: Mahajan S, Khurana M, and Estrela VV (Eds.), Applying artificial intelligence in cybersecurity analytics and cyber threat detection: 1-18. John Wiley & Sons, Hoboken, USA. https://doi.org/10.1002/9781394196470.ch1

Nasayreh A, Khalid HM, Alkhateeb HK, Al-Manaseer J, Ismail A, and Gharaibeh H (2025). Automated detection of cyber attacks in healthcare systems: A novel scheme with advanced feature extraction and classification. Computers and Security, 150: 104288. https://doi.org/10.1016/j.cose.2024.104288

Nifakos S, Chandramouli K, Nikolaou CK, Papachristou P, Koch S, Panaousis E, and Bonacina S (2021). Influence of human factors on cyber security within healthcare organisations: A systematic review. Sensors, 21(15): 5119. https://doi.org/10.3390/s21155119 **PMid:34372354 PMCid:PMC8348467**

Oke G, Loukas G, and Gelenbe E (2007). Detecting denial of service attacks with Bayesian classifiers and the random neural network. In the IEEE International Fuzzy Systems Conference, IEEE, London, UK: 1-6. https://doi.org/10.1109/FUZZY.2007.4295666

Pandey P and Litoriya R (2020). Securing and authenticating healthcare records through blockchain technology. Cryptologia, 44(4): 341-356. https://doi.org/10.1080/01611194.2019.1706060

Papadopoulos P, Abramson W, Hall AJ, Pitropakis N, and Buchanan WJ (2021). Privacy and trust redefined in federated machine learning. Machine Learning and Knowledge Extraction, 3(2): 333-356. https://doi.org/10.3390/make3020017

Pardakhe NV and Deshmukh VM (2019). Machine learning and blockchain techniques used in healthcare system. In the IEEE Pune Section International Conference, IEEE, Pune, India: 1-5. https://doi.org/10.1109/PuneCon46936.2019.9105710

Rani S, Pareek PK, Kaur J, Chauhan M, and Bhambri P (2023). Quantum machine learning in healthcare: Developments and challenges. In the IEEE International Conference on Integrated Circuits and Communication Systems (ICICACS), IEEE, Raichur, India: 1-7. https://doi.org/10.1109/ICICACS57338.2023.10100075

Ratta P, Kaur A, Sharma S, Shabaz M, and Dhiman G (2021). Application of blockchain and internet of things in healthcare and medical sector: Applications, challenges, and future perspectives. Journal of Food Quality, 2021: 7608296. https://doi.org/10.1155/2021/7608296

Roshani M and Nobakht M (2022). HybridDAD: Detecting DDoS flooding attack using machine learning with programmable switches. In the Proceedings of the 17th International Conference on Availability, Reliability and Security, Vienna, Austria: 1-11. https://doi.org/10.1145/3538969.3538991 **PMid:34053955**

Sharma DK, Chakravarthi DS, Boddu RSK, Madduri A, Ayyagari MR, and Khaja Mohiddin M (2022). Effectiveness of machine learning technology in detecting patterns of certain diseases within patient electronic healthcare records. In: Yadav S, Haleem A, Arora PK, and Kumar H (Eds.), Proceedings of Second International Conference in Mechanical and Energy Technology. Smart Innovation, Systems and Technologies, 290: 73–81. Springer, Singapore, Singapore. https://doi.org/10.1007/978-981-19-0108-9_8

Sivasankari N and Kamalakkannan S (2022). Detection and prevention of man-in-the-middle attack in IoT network using regression modeling. Advances in Engineering Software, 169: 103126. https://doi.org/10.1016/j.advengsoft.2022.103126

Sun X, Douiri A, and Gulliford M (2022). Applying machine learning algorithms to electronic health records to predict pneumonia after respiratory tract infection. Journal of Clinical Epidemiology, 145: 154-163. https://doi.org/10.1016/j.jclinepi.2022.01.009 **PMid:35045315**

Taylor A, Kleiman R, Hebbring S, Peissig P, and Page D (2020). High-throughput approach to modeling healthcare costs using electronic healthcare records. Arxiv Preprint Arxiv:2011.09497. https://doi.org/10.48550/arXiv.2011.09497

Tenepalli D and Thandava Meganathan N (2023). A review on machine learning and blockchain technology in e-healthcare. In: Abraham A, Pllana S, Casalino G, Ma K, and Bajaj A (Eds.), Intelligent systems design and applications. ISDA 2022: Lecture notes in networks and systems: 338-349. Springer Nature, Cham, Switzerland. https://doi.org/10.1007/978-3-031-35510-3_33

Tiwari A (2022). Supervised learning: From theory to applications. In: Pandey R, Khatri SK, Singh NK, and Verma P (Eds.), Artificial intelligence and machine learning for EDGE computing: 23-32. Academic Press, Cambridge, USA. https://doi.org/10.1016/B978-0-12-824054-0.00026-5

Tumpa ES and Dey K (2022). A review on applications of machine learning in healthcare. In the 6th International Conference on Trends in Electronics and Informatics (ICOEI), IEEE, Tirunelveli, India: 1388-1392. https://doi.org/10.1109/ICOEI53556.2022.9776844

Wang Z and Sun J (2022). PromptEHR: Conditional electronic healthcare records generation with prompt learning. In the Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, Abu Dhabi, UAE: 2873-2885. https://doi.org/10.18653/v1/2022.emnlp-main.185

Yeng PK, Fauzi MA, and Yang B (2020). Comparative analysis of machine learning methods for analyzing security practice in electronic health records' logs. In the IEEE International Conference on Big Data (Big Data), IEEE, Atlanta, USA: 3856-3866. https://doi.org/10.1109/BigData50022.2020.9378353