Chapter 1
Introduction to Artificial Intelligence and Cybersecurity for Industry

Shipra Rohatgi, B. Abinash, Sunidhi Joshi, Shami Sushant, and Sachil Kumar

Introduction

According to Atiku et al. [1], cybersecurity is a phenomenon that deals with ways to safeguard computer systems, networks, and electronic data from disclosure, theft, service disruption, and unauthorized access. *Cybersecurity* is defined broadly as "the study of the security of anything in the cyber environment," including information security, network security, operational security, application security, Internet of Things (IoT) security, cloud security, and infrastructure security [2]. Cybersecurity's primary goal is to shield consumers from threats as much as feasible. Additionally, it serves the purpose of promptly and effectively completing the requirements for detection before, handling during, and recovery following the accident. According to predictions, the cybersecurity sector may increase by more than $1 trillion USD between 2016 and 2021. The modern generation now relies heavily on the Internet in their daily lives. The amount of data we trade every day is massive and vast. However, the rate at which cyberattacks are occurring is also dramatically rising. Every few months, fraudsters reduce the cost of their customized attacks while doubling their effectiveness. Furthermore, cybersecurity becomes ineffectual as cyberattacks get more complex and automated [3, 4]. Traditional cybersecurity techniques, such as network protection and computer security systems, are ineffectual against cyberattacks' always-changing, transformative, and inventive attempts [5].

S. Rohatgi · S. Joshi · S. Sushant
Amity Institute of Forensic Sciences, Amity University Noida Campus, Sector-125, Noida, Uttar Pradesh, India

B. Abinash · S. Kumar (✉)
Department of Life Sciences, CHRIST (Deemed to Be University), Bengaluru, Karnataka, India
e-mail: sachil.kumar@christuniversity.in

V. Sarveshwaran et al. (eds.), *Artificial Intelligence and Cyber Security in Industry 4.0*, Advanced Technologies and Societal Change,
https://doi.org/10.1007/978-981-99-2115-7_1

Classification of Cyberattacks (See Fig. 1.1)

1. **Depending on the goal**: It can target a person or an organization. Government and commercial sector cyberattacks are more harmful since they cause enormous losses and frequently start with altering important data for financial gain.
2. **Web-based attack:** A hacker attempts to obtain unauthorized access to a website by exploiting vulnerabilities in it. It may be done by:
   - **SQL Injection Technique**—allows hackers to read, change, and delete tables from databases by manipulating a regular SQL query on a database-driven website.
   - **Phishing—**Hacker's spoof emails, in particular, to get recipients to accept them and follow instructions that typically request personal information. Some phishing scams involve the download of malware.
   - **Man in the Middle**—The hacker gains access to the information stream between the user's device and the website server. The hacker's computer assumes control of an IP address, so the communication channel between

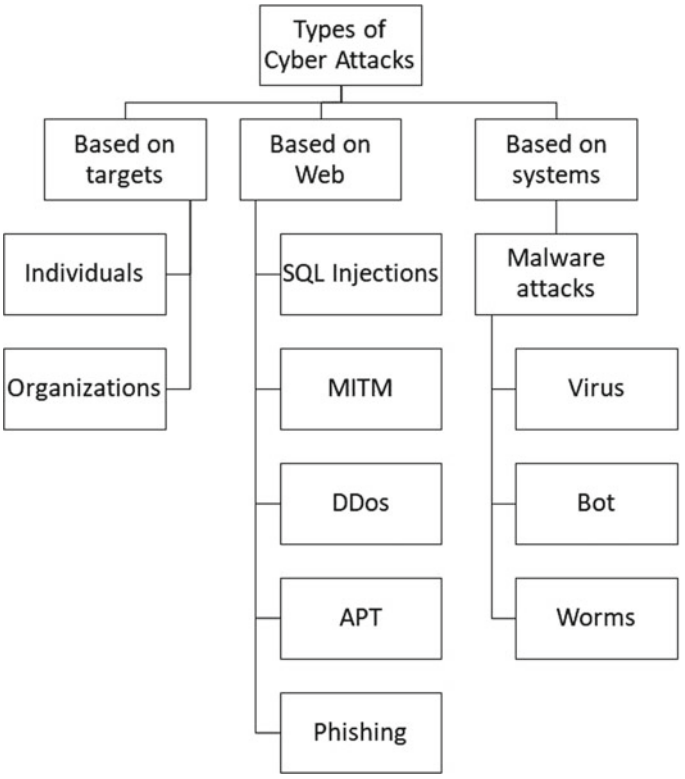

Fig. 1.1 Types of cyberattack

the user and the website is covertly interrupted. This frequently occurs on unprotected Wi-Fi networks.
- **Password Attack**—Typically, the system passwords are broken by utilizing either widely used passwords or by attempting all conceivable alphabetic combinations. It is among the most straightforward techniques to break into a system.
- **APT**—Hackers gain access to networks over an extended period to obtain sensitive information.

3. **A System-Based Attack**: It inserts itself into the target computer system, where it will replicate and compromise the system to harm the computer system or its network.

- **Malware Attack**—Typically, the hacker sends phony emails that seem to originate from a reliable source. This is done to implant malware or steal sensitive data, such as credit card numbers. Any form of malicious software, regardless of its operation, purpose, or delivery, is called malware. Some of them are as follows:
- **Virus**—In addition to their malicious behavior, viruses can spread to other systems and infect other programs. When a file is opened, the virus is launched along with it. The virus then moves, deletes, corrupts, or encrypts data and files. An enterprise-level antivirus solution gives you centralized management and visibility while safeguarding all your devices from one location regarding virus protection. Regularly do comprehensive scans and update your antivirus definitions.
- **Bots**—are computer programs that carry out automated activities without human interaction. Attacks can be executed by bots a lot quicker than by humans. A computer that has a bot infection can propagate the bot to other computers, forming a network of computers known as a botnet. This network of infected workstations may be managed and utilized to carry out extensive attacks like DDoS attacks and brute-force attacks. Many times, device owners are ignorant of their contribution to the attack. On specific hardware, bots are also employed for cryptocurrency mining. Using tools to distinguish traffic from actual users and bots is one technique to manage bots. For instance, you could add a CAPTCHA for her to stop bots from flooding your form.
- **Worms**—Like viruses, worms can replicate on different hardware and software. In contrast to viruses, worms do not require human intervention to spread over a network or system. Worms frequently target computer hard drives and memory. One should ensure that all their devices have the most recent fixes installed to protect themselves from worms. Technology can also assist in identifying files and links that can contain worms, such as firewalls and email filters.
- **Trojan Horse**—Trojans impersonate desired software or programs. The Trojan can hijack the victim's machine if it is downloaded by an unwary user and used for malicious purposes. Trojans can be inserted into phishing

email attachments, games, apps, and even software patches. Trojans are malicious programs that pose as trustworthy ones. Trojan horses must be spread by the victim, frequently via social engineering techniques like phishing, as they cannot spread themselves as viruses and worms can. Trojans rely on social engineering to proliferate, leaving the user responsible for defense. Unfortunately, 82% of breaches in 2022 had a human component. Security awareness training is crucial since employees are the first line of defense against these types of assaults and the target of Trojans.

Types of Cybersecurity

Hardware, software, and infrastructure are all part of cyberspace security. Consequently, it can also be classified as.

Hardware and Software Tools

For network security, shield infrastructure and networks from disruption, unlawful access, and abuse. Effective network security shields corporate assets from several risks inside and outside the business.

Information Security

Safeguards private data against review, modification, recording, and illegal activities like interruption or destruction. Its goal is to guarantee the security and privacy of sensitive data, including financial information, intellectual property, and customer account information.

Cloud Security

Utilizing cloud service providers like Amazon Web Services, Google, Azure, and Rack Space to build secure cloud architectures and apps for your business is referred to as cloud security.

Security for Applications

Application security entails putting in place a variety of defenses in a company's software and services against various attacks. This includes security testing by either white-box, black-box, or gray-box testing. This helps to improve the application's security aspects by making its code more secure, reducing vulnerability by reducing

the likelihood of unwanted access or alteration, or implementing strong data input validation.

Tools Used in Cybersecurity

Firewall

A *firewall* is a safety measure that guards a network against unauthorized access to sensitive information. A firewall establishes a barrier between your secure internal and unreliable external networks and safeguards your computer against harmful malware. Depending on the amount of security required by the customer, firewalls provide various levels of protection. Firewalls often welcome incoming connections that are authorized access to the network. Depending on the security rules in place, the security system either permits or denies data packets. Web traffic is filtered at checkpoints set up by firewalls. These solutions allow us to examine and respond to unauthorized network activity before it negatively impacts the network being attacked.

Honeypots

Honeypots are security tools or decoy systems used by security experts to lure or attract attackers by making the target vulnerable or attractive. They are used to detect, deter, and investigate unauthorized access to the target system. They can act as high-value targets for attackers, like a server online. The primary purpose of honeypots is to gather attacker data by allowing for early detection of infiltration attempts while giving a few footprints of attackers. To minimize the threats, honeypots often employ a secure operating system with added security features.

Penetration Testing

Simulation of a cyberattack on a computer system to identify its vulnerabilities is known as penetration testing or Pentesting. This technique's most commonly implemented use is to improve the security of web applications and their firewalls (WAFs). Pentesting targets various systems, such as front-end and back-end systems/servers. APIs and other components to detect any security weaknesses. This testing process also checks for potential injection vulnerabilities. The results of penetration testing can be used to enhance the security settings of WAFs and help address any vulnerabilities.

Encryption Tools

Data are transformed from readable to encrypted forms using encryption in cybersecurity. Only after it has been decrypted, an encrypted data can be read or processed.

A crucial element of data security is encryption. This is the most straightforward and crucial approach to stop someone from stealing or reading your computer system's data with harmful intent.

Data Protection Individual and large enterprises frequently employ encryption to safeguard user data between browsers and servers. Everything from payment information to personal information is included in this information. Data encryption software is used to create encryption schemes that, in theory, can only be broken by powerful computers. This software is sometimes referred to as encryption algorithms or ciphers. Symmetric and asymmetric encryptions are the two most popular types.

- Private key encryption also refers to using a symmetric encryption key. It is perfect for single-user and closed systems because the encryption key is also needed for decryption. If not, the recipient should receive the key. If a third party gains access, the risk of compromise rises. B. Cybercriminals are stopped. Compared to the asymmetric method, this one is quicker.
- Asymmetric cryptographic keys: These employ public and private keys that are mathematically tied to one another. Keys are essentially big numbers that are paired but not identical, hence the word "asymmetric." Public keys are distributed to approved recipients or made available to the public, while private keys are kept hidden by their owners.

Packet Sniffers

During a network's Transmission Control Protocol/Internet Protocol (TCP/IP) layer, a packet sniffer is a program or utility that reads data packets. These technologies are used by network administrators to "sniff" Internet traffic and keep an eye on data in real time. The data can be interpreted to evaluate and identify server, network, hub, and application performance issues. Network administrators can utilize one of several techniques to find sniffers on their networks when hackers use packet sniffing to monitor Internet activities illegitimately. Utilize this early warning and take action to safeguard your information from unauthorized listeners.

One strategy that can be employed successfully in cybersecurity is artificial intelligence (AI) [6]. AI is regarded as the intelligence being added in digital devices and computer-operated devices to perform n number of tasks like a human. It mainly focuses on studying the brain's cognitive process and activities to employ in developing devices and software with human intelligence. Fighting cyberattacks has been made much easier thanks to two recently developed fields of AI: machine learning (ML) and deep learning (DL). John McCarthy first proposed the concept of AI in 1956; it is the science and engineering of creating intelligent autonomous security systems. The primary goal of AI is to teach computers to think, learn, act, and behave intelligently and cognitively like humans.

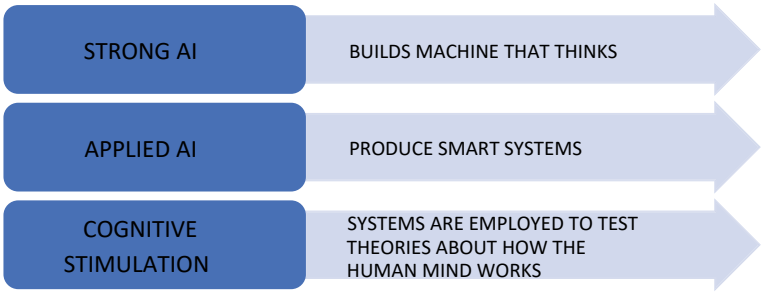

Fig. 1.2 Significance of AI

In the current technological period, AI offers services in many fields, including computer vision, pattern recognition, expert systems, language processing and translation, speech recognition, biometric systems, robotics, the Internet of Things (IoT), and other related domains [7]. A significant area of computer science called AI works with autonomous, intelligent systems resembling the human brain [8]. The importance of AI in cybersecurity is more significant than ever in the digital age (see Fig. 1.2). Cybersecurity is a popular topic. Several industries, including natural language processing, gaming, health care, manufacturing, and education, swiftly adopted artificial intelligence once it was initially presented. AI can analyze massive amounts of data quickly, effectively, and accurately thanks to its robust analytics capabilities. In contrast to previous systems, AI systems can anticipate upcoming cyberattacks based on current threats, even if those threats change. As a result, it is only possible to use AI to counter security risks.

Two Approaches in Artificial Intelligence: Symbolic Versus Connectionist

There are two basic approaches taken up in AI: the symbolic or top-down approach and the connectionist or bottom-up approach. The top-down approach is about analyzing the cognitive process and replicating it to analyze the symbols or symbolic labels. It is not dependent on the biological constitution of the brain. In contrast, the bottom-up or connectionist approach focuses more on brain structure. It involves the creation of an artificial neural network that imitates the brain.

It can be easily understood straightforwardly, considering the task of building a system that can recognize the letters of the alphabet. To demonstrate the distinction between these methods, a bottom-up approach is considered. A typical bottom-up strategy involves gradually "tuning" an artificial neural network to improve performance by presenting letters to it one at a time. The responsiveness of various neural pathways to various stimuli can be altered through tuning. A top-down approach, on the other hand, typically entails creating a computer program to compare each letter to geometric descriptions. The bottom-up approach relies on neural activities, whereas the top-down approach relies on symbolic descriptions.

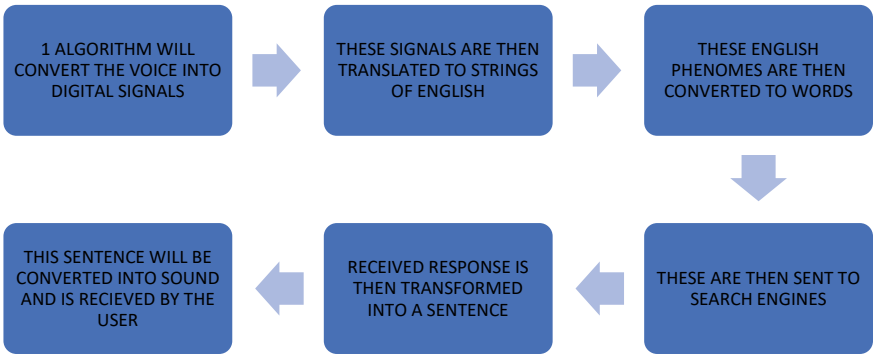

Fig. 1.3 Working mechanism

How Does Artificial Intelligence Work? (See Fig. 1.3)

AI works by implementing human intelligence and a set of algorithms in order to answer several questions. This combination allows AI to analyze the data and generate the results. These data are then further used to generate the results and build expertise.

For example, if someone asks an AI system—how to make chapatis?

AI Implementation Methods

Machine Learning

The ability to learn in AI is obtained through machine learning. It exploits the algorithms to obtain insight into patterns and generate a profile.

Deep Learning

Deep learning tries to mimic the event in layers of neurons to learn in a real sense to identify a pattern in the given text. These methods are only sometimes better than traditional supervised approaches as deep learning performance is subjected to the correct selection of algorithms and numbers of hidden layers and a feature representation technique. Deep learning models show a promising future in text mining as it depends entirely on the neural network with extra depth.

To emphasize the significance of AI systems and technologies as a defense against cybersecurity, this chapter is being offered. This chapter will provide a significant

response to the essential subject of how AI-based solutions might help to safeguard cybersecurity. Additionally, we have demonstrated the limitations of AI in cybersecurity as well as some potential future research topics in this chapter.

Background Information on AI Methods and Cybersecurity Applications

The opposite is also true: Cybercrimes and attacks are rising rapidly due to the rapid advancement of Internet technology and systems. We require AI-based strategies in our cybersecurity systems to improve the security of cyberspace more effectively in order to combat and defeat these crooks and their cunning methods.

Table 1.1 provides short descriptions of AI methods and applications (see Table 1.1, Fig. 1.4).

Table 1.1 Techniques and applications of artificial intelligence (AI) in cybersecurity

Techniques of AI	Applications in cybersecurity
Neural nets	1. For intrusion detection and prevention system 2. Very high-speed of operation 3. For Denial-of-service (D0S) detection 4. For forensic investigation 5. Warm detection 6. Fuzzy logic
Intelligent agents	1. Proactive and reactive 2. Agent communication language 3. Defense against Distributed Denial-of service (DDoS)
Expert systems	1. For network intrusion detection 2. For decision support 3. Knowledge base 4. Inference engine
Application of learning	1. Machine learning and deep learning 2. Data mining 3. Supervised and unsupervised learning 4. Intrusion detection and malware detection 5. Self-organizing maps

Fig. 1.4 Applications of AI in cybersecurity

The Significance of Cybersecurity

- Cybersecurity has grown significantly in importance in society due to the rise in cybercrime; antivirus software is no longer adequate to safeguard our system network and its data.
- The cybersecurity sector nowadays is primarily concerned with defending systems and devices against intruders. It is challenging to picture the bits and bytes that go into these efforts, but it is far simpler to consider their meaning. Due to frequent denial-of-service attempts, many websites would be almost entirely inoperable if not for the diligent efforts of cybersecurity experts.
- Only authorized individuals can access confidential data and operations thanks to cybersecurity, military secrets, for instance.

The integrity principle states that only authorized individuals and agents can add, modify, or delete sensitive information and functionality. Example: The database contains inaccurate data that a user entered.

Availability: According to the availability, concept, systems, functions, and data must always be accessible per predetermined guidelines based on service levels.

How AI Can Be Applied on Cybersecurity Issues

To address cybersecurity issues, AI offers many advantages, some of which are listed below:

1. As opposed to previous technology, which was primarily concerned with the past and solely relied on known cyberattacks, conventional systems fail to recognize

changes when a new cyberattack occurs, creating a blind space for unorthodox attacks. AI can recognize novel and intricate variations in attack flexibility. Future AI systems will better detect similar changes. AI machines are more capable of learning and adapting, and they can identify abnormal operations more quickly and accurately. This ability of AI systems is more critical when cyberattacks are becoming more refined, and cybercriminals are coming up with new and inventive methods [3, 4].

2. AI can handle many security data [3, 4], because AI has built-in security mechanisms that can recognize and respond to threats. Security employees must deal with an intolerable number of data breaches daily, but automatically identifying and responding to threats have lessened their workload. Furthermore, AI is the only approach that can effectively handle these intrusions. Network security analysts will find it increasingly challenging to detect and monitor attack elements accurately and promptly as more and more security data are generated and transferred over the network daily. AI can aid in this situation by boosting the frequency with which suspicious behavior is highlighted and recognized. This can help network security officials to respond to circumstances.
3. Over time, AI security systems examine application behavior, and routine network AI creates a baseline of typical patterns by detecting risks over time. The AI security system will identify the attacks if there is any modification or divergence from the usual routine.

AI Techniques Used for Cybersecurity

Some AI security models that can effectively counter threats and cyberattacks include neural networks, expert systems, machine learning, deep learning, and data mining. In the cybersphere, intelligent decisions can be made using AI-based techniques. All of these AI strategies have been emphasized in this essay and are briefly discussed in the parts that follow.

Artificial Neural Network

Neural networks represent deep learning and analysis of data by AI. Some scenarios and situations are too much and out of scope for machine learning algorithms to cope with. ANN can be connected to biological neurons within the body, which performs a specific function and carries stimulus resulting in action. It consists of layers of interconnected artificial neurons powered by different activation functions, which help in on/off mechanisms. Each neuron receives a unique version of input and random words, which are added with fundamental barriers and unique layers; this then passes to an activation function that depicts the final value of the neuron. The output is generated from the final neural layer, the loss function. It will input

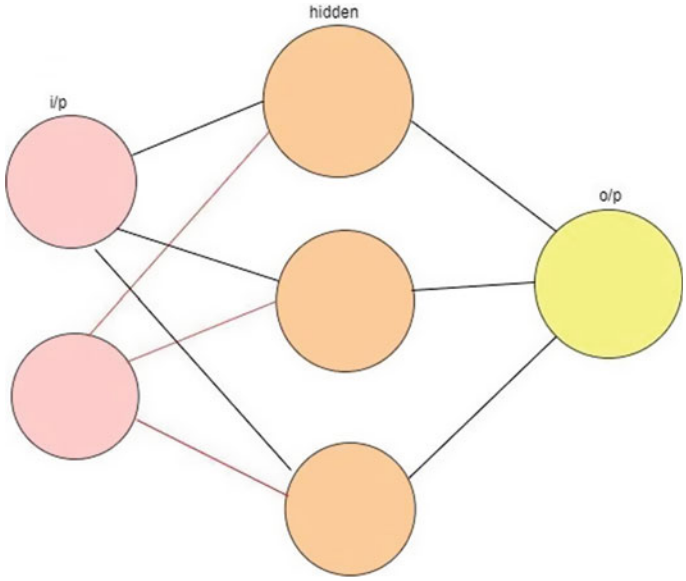

Fig. 1.5 Different layers of neurons in a neural network

versus output is calculated, and backward propagation of error is performed. The weights are designed to make the loss minimum.

1. **Weights**—values that are multiplied by the input. In the backpropagation of errors, they are modified to reduce the losses. Weeds are machine-learned values from the neural network (NN) that evaluate self-adjust depending on the difference between foreseen output and input. *Activation functions* are a mathematical formula that helps in the on–off of neurons.

   The following figure shows three main layers of a neural network (see Fig. 1.5):

- **Input layers** represent the dimensions of input vectors.
- **Hidden layer** represents the media notes that divide the input into the area with boundaries 8, takes weighted inputs, and produces output through an activation function.
- **Output layer** provides the output of the neural network.

Detection Using Neural Networks in Tweets

1. **CNN—**A convolutional neural network consists of three layers naming

   - Convolution layer.
   - Pool layer.
   - FC layer.

2. **The convolution layer**—This layer is an essential element in CNN. It extracts features from data (input).
3. **Pool layer—**This layer performs a function to reduce the dimensions of input it gives you say the number of parameters.

There are several types of functions.

- Max pooling.
- Average pooling.

Max pooling is used widely as it only takes some maximum value in the window.

For example, the view layer takes a 4 * 4 feature map as input from the C layer and performs max pooling with 2 * 2 window size I am Stroud of 2 for each window max pooling takes a maximum value of 4 consequently exchange rates 2 * 2 feature outputs which are sent to fully connected layer.

Security Expert Systems

A computer program known as an expert system, or AI, aids a human expert in making decisions. The system consists of an inference engine and a knowledge base, which together produces security rules [9]. Security standards provide the basis for judgments made by cybersecurity expert systems. Applications for expert systems' modeling can be found in cyberspace, finance, and medical diagnosis. Expert systems can come in various sizes, from simple hybrid systems to massive, complex ones that handle complex concerns and issues. The knowledge base phase of the cybersecurity expert framework describes domain knowledge and operational understanding of the rules governing security decisions. The inference engine phase retrieves information from the knowledge base and concludes new facts through expert systems. In one strategy known as the "case-based reasoning (CBR) approach," a particular problem is solved by recalling prior, analogous cases. Then, a solution is determined by adapting the previous solution to a new problem case. This method examines novel solutions to raise the system's accuracy and capacity for learning.

Another method for resolving issues is known as rule-based systems (RBSs), which rules established by experts characterize. The condition portion and the action are the two subsystems of the rule-based system. The difficulties are analyzed using condition part evaluation, and the appropriate course of action is decided. Cybersecurity expert system uses basic standards and rules to combat cyberattacks. For instance, it compares the process to the knowledge base; if the process is excellent and known, the security system considers it secure; otherwise, the system declares the process a threat and ends it. The system looks for the sets of rules in the inference engine to determine the machine's state if the knowledge base does not have such a procedure. Based on the machine's condition, followed by the inference identified by the knowledge base, the system informs the manager or user of the machine's status.

Thus, a rule-based cybersecurity expert system model can make decisions like a security expert in an intelligent cybersecurity framework designed to address challenging cybersecurity problems. Because of this, cybersecurity expert system modeling, based on its computational powers and capacity for reasoning, might be helpful in AI-based cybersecurity.

Intelligent Agents (IAs)

Intelligent agents (IAs) are autonomous systems with a decision-making process internal to them and a personal goal. Through sensors, it assesses risks, and actuators monitor the domain. It directs the activity until a specific goal is attained [10]. These systems exhibit proactivity and responsiveness, and when interacting with other autonomous agents, they can comprehend and adapt to changes in their environment. These intelligent agents can learn about and interact with their surroundings, making them adaptive. IAs are successful in thwarting distributed denial-of-service (DDoS) assaults. How can these agents be used to defend against decentralized cyberattacks? The solution is to create artificial "Digital police," which must comprise mobile intelligent agents. It was necessary to install infrastructure to give solid support.

Search

Every day, we employ the search technique as a heuristic for solving issues. Before using a search algorithm, a prior understanding of the search strategy is essential. Nearly, all intelligent programs now use or include these search algorithms, which favors the entire intelligent system. Several search security systems are used in AI, including the search estimation used in many projects. For computer chess, the search estimation was developed. For computer chess, the search estimation was developed. It uses the "isolate and vanquish" critical thinking technique, which is helpful in ad hoc leadership situations where two opponents decide on their most advantageous course of action.

Bio-inspired Computing Method

Using sophisticated algorithms and techniques, bio-inspired computing in artificial intelligence (AI) uses bio-inspired behaviors and attributes to address various challenging academic and environmental problems. These methods are frequently used in cyberspace and include Evolution Strategies (ESs), Ant Colony Optimization (ACO), Artificial Immune System (AIS), Particle Swamp Optimization (PSO), and Genetic Algorithms (GAs). This method is also employed to categorize computer malware. These methods are generally employed to improve the characteristics and parameters used by the classifiers in classifying computer malware. As an illustration, PSO and GA approaches were used to increase the effectiveness of the malware detection

system [11]. Another study employed fuzzy logic and GA to detect intrusions. Using glow analysis to forecast network traffic behavior for a given period, the GA was used to construct a digital signature of a network section. Additionally, the fuzzy logic method determined the network instance's anomalousness. A university's network traffic was used for the study, and the outcomes showed 96.53% accuracy and 0.56% false notice.

Machine Learning (ML) and Deep Learning (DL) Methods

Artificial intelligence, known as "machine learning," focuses on teaching computers how to learn new things and use algorithms to make data-based decisions. Machine learning is strongly tied to mathematical methods enabling data extraction, pattern detection, and conclusion drawing. Regression and classification are two of ML technology's most crucial techniques. Supervised, unsupervised, semi-supervised, and reinforcement learnings are the first four types of learning.

Machine learning, also called "deep learning," is a skill that uses data to teach computers how to perform tasks that humans previously could only perform. This is performed by modeling the mechanism of data interpretation in the human brain. Deep learning is predicated on the idea that more extensive neural networks perform better as we train them with more data and scale them up.

It has been demonstrated that ML and DL are crucial for solving cybersecurity problems. The security system can use ML approaches in a variety of ways. Spam filtering, network abnormalities analysis, botnet tracking, and user behavior anomalies' tracking are a few examples.

Deep Learning Detection for Misinformation

What is misinformation? (See Fig. 1.6)

Misinformation is false or inaccurate information that deceives people by obscuring the truth. It can also be referred to as untruth, ambiguity, or deception. The spread of misinformation can harm relationships and undermine trust by presenting false sensations, leading to a negative breach of expectations or trust in society and the people in it, which is harmful.

Misinformation has many terms:

- **Rumor** is a story of the circulation of information from one user to another whose authenticity status is doubtful.
- **Fake news** is an article that misleads its readers and is false.
- **Spam** can be an unsolicited text sent over the Internet to spread advertising malware and other unresourceful complete data.

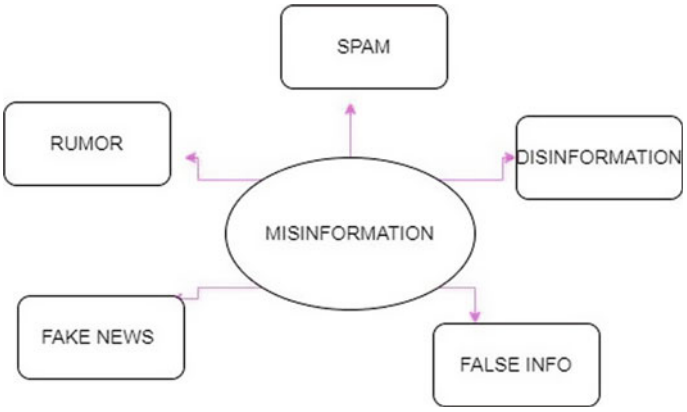

Fig. 1.6 Types of misinformation

- **Disinformation** is a piece of inaccurate information that people spread intentionally to mislead other readers
- **False information** is misinformation spread done intentionally.

Methodologies (see Fig. 1.7)

We have deep learning techniques in three main categories:

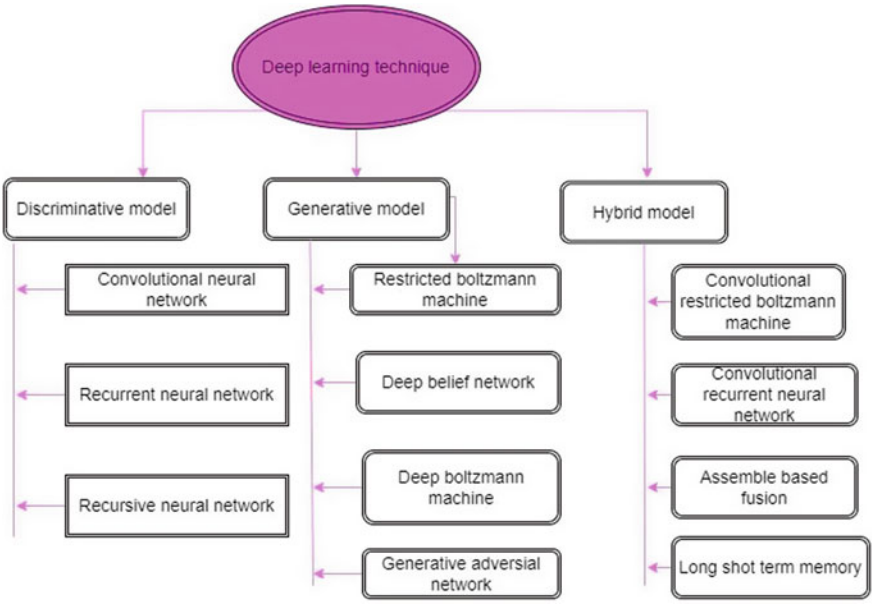

Fig. 1.7 Deep learning techniques for misinformation detection

1. **Descriptive model.**
2. **Generative model.**
3. **Hybrid model.**

Benefits of AI in Cybersecurity

Significant benefits have been received by those institutions that use AI techniques in their cybersecurity operations. Implementing AI in cybersecurity challenges, for instance, the ROI of some institutes has increased. Siemens AG developed the AI-based Siemens Cyber Defense Center (CDC), which is swift, autonomous, and adaptable. Amazon Web Services (AWSs) also utilize this system. The system was predicted to be under 60,000 attacks per unit of time due to AI application. With less than 12 individuals, the overall capability was quickly controlled, and system performance was well-maintained. Artificial intelligence in cybersecurity can detect emerging attacks by examining historical threat trends [12].

The time and effort spent on researching and identifying risks and assaults can be saved by using this AI approach. It has been discovered that AI is beneficial in identifying risks and responding to them at minimal cost (an average 12% cost decrease). AI may offer significant answers to cybersecurity concerns as the cybersecurity system shifts from conventional and manual procedures to automated algorithm mitigation. In contrast to traditional technology, AI can detect new, complex changes in the attack extensibility, which primarily relies on previously identified intruders and intrusions and thus leaves a blind spot during unusual intrusion activities. At present, AI technology has overcome the shortcomings of traditional security systems. For instance, it is now possible to monitor privileged Internet activity, and any change in how privileged access procedures are carried out is a potential threat. Security teams have an advantage thanks to AI predictive tactics, which are crucial to thwarting assaults before they cause any damage. For the detection of trends and dangers across a variety of sectors, including retail, manufacturing, energy, and transportation companies, Dark Trace (a UK company) uses ML technology. Through AI-based solutions, large amounts of data may be managed, as well as the development of network security systems. Security experts are overwhelmed by the vast number of open security issues. The workload of security organizations has lessened due to AI's autonomous detection and response to assaults. Security specialists need help for managing security data when it is produced and exchanged in large quantities daily. AI can therefore aid in accelerating the study of dubious procedures and activities.

Additionally, by eliminating the manual methods that take much time when responding to novel situations, security people can profit and react to new scenarios better. AI-based systems are equipped to learn over time and can react to threats and attacks more effectively. Using application characteristics and overall network activity, AI can help to identify attacks. As time passed, AI memorized the common

and typical traffic situation and established a cap for the usual activities. As a result, the attack is noted when there is any abnormal deviation.

Detecting False Information Using Neural Networks

This section covers various methods and procedures for identifying false rumors and deceptive material. It also thoroughly examines the five most common phrases for misinformation and how they deceive social media users.

Multiple models are put out for detection based on various methods and annotations. This research on socially essential data is used in fake news identification to tell real news from phony by comparing the two. Convolutional neural networks with gated recurrent units extended short-term memory networks, and other deep learning networks were investigated in this. We also looked into the benefits of feature extraction and feature embedding in deep neural networks.

Using Neural Networks to Find Objectionable YouTube Content

The transfer learning method is used to extract frame features from videos. The video features are then processed through directional LSTM, where the model learns efficient video representation and forms video classifications. The deep learning-based framework is proposed for inappropriate video content detection and classification. All analyses are carried out by utilizing manually compiled YouTube video clip datasets. The following are some advantages of a system for detecting child-appropriate content based on deep learning:

- It operates by taking into account real-time circumstances and processing video at a speed of twenty-two frames per second while utilizing an effective network and bidirectional extended short-term memory-based framework.
- Any video-sharing site can benefit from using it to flag or remove videos with questionable content.

Detecting Hate Speech Using a Neural Network

Although significant legislative and law enforcement efforts and millions of dollars in investments from social media firms, the spread of hate speech on online platforms has been significantly expanding, it is commonly acknowledged that automated data mining techniques are essential for developing efficient defenses against various threats. The techniques for classifying hate speech utilizing deep neural models incorporating CNN and LSTM, as well as GRU, to increase classification accuracy were first introduced. Second, we compare and contrast the most extensive public

dataset with the rest of the data to create new references for comparative research in the future. With all the information, distinct datasets categorized into different subclasses aid in classification.

Challenges Faced with Integration of AI in Cybersecurity

Classification Error

AI can possess severe classification errors; these errors can be due to the alteration of pixels. According to the published literature, classification errors of neural networks can occur with the change in pixels or modification of bytes. Hence, it can be concluded that if the data source is infected or altered, it can easily cheat the AI system.

Intensive Requirement for Resources

While AI can bring good automation for cybersecurity, one thing that can be a significant issue is the need and requirement for high-end equipment and servers for usage. Most small-scale businesses need help to afford this, which becomes a significant issue. Another thing is the need for personnel trained in using and maintaining AI.

Maliciously Modified Model

Implementing an AI model is a program that may have some vulnerabilities. These vulnerabilities may be due to the designer's unreasonable and careless design of the logical structure of the model. They may come from a specific high-level language, hardware-specific problems, or the back door embedded in the model. Gu et al. [13] implemented the backdoor in the neural network, which made the neural network's performance in the specific attacker sample very poor. These shortcomings also reflect that the given answers by the program are only sometimes accurate.

Cost

Integration of AI requires a large sum of investment to run and maintain. In the current market scenario, only a few companies in the IT field have been able to integrate AI into their systems.

Lack of Transparency

In the decision-making process of AI, all the participants, including programmers, do not know why the AI model gives the final decision results, i.e., the decision-making process of AI lacks transparency. The AI model is like a black box. In the process

of its creation and self-improvement, it can realize the automatic configuration and adjustment of parameters without too much staff intervention, thus saving human resources. Nevertheless, at the same time, the problem is that its decision-making process needs to be explained clearly. Although the AI model can achieve high accuracy, the tests are implemented in the test set. Therefore, whether the AI model can achieve such a high accuracy remains to be verified when facing unknown events. When there are objections to the decision-making results given by the AI model, it is difficult to explain the decision-making process, so some people will be skeptical of the decision-making results, that will not be conducive to the rapid judgment of the network situation or even cause irreversible consequences. Some research teams have begun to conduct in-depth research on this issue.

Public Perception

Popular media and sci-fi movies have painted AI in a sense that puts fear into the minds of people unfamiliar with AI and machine learning and how it works. For them, AI can bring about the end of humanity if it has access to security systems.

AI Use Cases

A major issue with AI implementation is the use case. The primary reason for failure is the improper implementation of the AI use case, and often companies try to implement AI without taking baby steps.

Malware Signature

The method by which AI works on the malware is through its known signature. These signatures are like fingerprints, and changes in these signatures cause the fingerprint to change. With Scripts changing every moment, it becomes complicated for the AI to identify and work on this malware. The AI models need a lot of data to complete the training. Before using data, they may do operations that mainly include a series of steps such as data noise reduction, normalization, missing value filling. If a supervised method is used, it is necessary to label the data manually. However, due to the substantial heterogeneity of cyberspace, different cyber structures may produce different high-risk events, which have sudden characteristics. Therefore, each possible high-risk event must be estimated in advance before the design of the models, and these high-risk events should be analyzed and labeled in advance. Meanwhile, AI models have a high demand for data, which may need more time to make a timely judgment.

Discussion

Cybersecurity and artificial intelligence overlap in various transdisciplinary fields (AI). Deep learning and other AI technologies can be used in cybersecurity to create intelligent models for malware categorization, intrusion detection, and threat intelligence sensing. On the other hand, AI models will be exposed to various cyber threats, which will interfere with their decision-making, learning, and sampling. Therefore, cybersecurity defense and protection solutions are required for AI models to counter adversarial machine learning, safeguard machine learning privacy, secure federated learning, etc. We examine the interaction of AI and cybersecurity considering the two factors mentioned earlier:

1. We review the research on using AI to defend against cyberattacks, including adopting conventional machine learning techniques and current deep learning solutions.
2. We examine the counterattacks that AI itself might encounter, examine their traits, and categorize the related defense strategies.
3. We elaborate on the existing research on creating a secure AI system from the perspectives of developing encrypted neural networks and implementing secure federated deep learning.

Conclusion

New problems for cybersecurity have developed and emerged along with the ICT's quick improvements. Modern cyberattacks and threats are so complicated and advanced that traditional techniques and strategies can no longer help them. New procedures and strategies that are optimal, scalable, adaptable, and flexible are required to combat these sophisticated cyberattacks. We have provided an overview of AI applications in cybersecurity in this paper. Data learning, security expert systems, and bio-inspired methodologies are a few of the well-researched AI-based cybersecurity solutions that have been covered.

Additionally, areas, where AI is used in cybersecurity, are examined, including the prediction, detection, and prevention of intrusion and malware, defenses against distributed denial-of-service (DDoS), a method where digital police is used, and many other areas. AI applications in cybersecurity were also highlighted, along with some of the benefits and difficulties. These advantages include managing massive amounts of data quickly and accurately, lowering the cost of using AI approaches to address cybersecurity concerns, and boosting the return on investment for AI-powered cybersecurity technologies, among others. Adversarial machine learning and human self-approval are two significant difficulties with using AI-based applications for cybersecurity. Although there are more advantages and disadvantages, AI-based security solutions are still used in cybersecurity. Many industry professionals concur that AI and cybersecurity must be combined because humans depend on cybersecurity.

References

1. Atiku, S.B., Aaron, A.U., Job, G.K., Fatim, S., Yakubu, I.Z.: Survey on the applications of artificial intelligence in cyber security. Int. J. Sci. Technol. Res. **9**(10), 165–170 (2020)
2. Sarker, I.H., Furhad, M.H., Nowrozy, R.: Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Comput. Sci. **2**(3), 1–18 (2021)
3. Truong, T.C., Diep, Q.B., Zelinka, I.: Artificial intelligence in the cyber domain: Offense and defense. Symmetry **12**(3), 410 (2020)
4. Truong, T.C., Zelinka, I., Plucar, J., Čandík, M., Šulc, V.: Artificial intelligence and cybersecurity: Past, presence, and future. In: Artificial intelligence and evolutionary computations in engineering systems, pp. 351–363. Springer, Singapore (2020)
5. Kabbas, A., Alharthi, A., Munshi, A.: Artificial intelligence applications in cybersecurity. IJCSNS Int. J. Comput. Sci. Netw. Secur. **20**(2), 120–124 (2020)
6. Soni, V.D.: Challenges and Solution for Artificial Intelligence in Cybersecurity of the USA. Available at SSRN 3624487 (2020)
7. Shamiulla, A.M.: Role of artificial intelligence in cyber security. Int. J. Innov. Technol. Exploring Eng. **9**(1), 4628–4630 (2019)
8. Helm, J.M., Swiergosz, A.M., Haeberle, H.S., Karnuta, J.M., Schaffer, J.L., Krebs, V.E., Spitzer, A.I., Ramkumar, P.N.: Machine learning and artificial intelligence: Definitions, applications, and future directions. Curr. Rev. Musculoskelet. Med. **13**(1), 69–76 (2020)
9. Tyugu, E.: Artificial intelligence in cyber defense. In: 2011 3rd International Conference on Cyber Conflict, pp. 1–11. IEEE (2011)
10. Wirkuttis, N., Klein, H.: Artificial intelligence in cybersecurity. Cyber Intell. Secur. **1**(1), 103–119 (2017)
11. Fatima, A., Maurya, R., Dutta, M.K., Burget, R., Masek, J.: Android malware detection using genetic algorithm based optimized feature selection and machine learning. In: 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), pp. 220–223 (2019)
12. LAZIĆ, L.: October. Benefit from Ai in cybersecurity. In: The 11th International Conference on Business Information Security (BISEC-2019), 18th October 2019, Belgrade, Serbia (2019)
13 Gu, F., Ma, B., Guo, J., Summers, P.A., Hall, P.: Internet of things and Big Data as potential solutions to the problems in waste electrical and electronic equipment management: An exploratory study. Waste. Manage. **68**, 434–48 (2017)