IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 January 2024, accepted 27 February 2024, date of publication 1 March 2024, date of current version 8 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3372187

RESEARCH ARTICLE

A Situation Based Predictive Approach for Cybersecurity Intrusion Detection and Prevention Using Machine Learning and Deep Learning Algorithms in Wireless Sensor Networks of Industry 4.0

FATIMA AL-QUAYED[1], **ZULFIQAR AHMAD**[2], **AND MAMOONA HUMAYUN**[3]

[1]Department of Computer Science, College of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia
[2]Department of Computer Science and Information Technology, Hazara University, Mansehra 21300, Pakistan
[3]Department of Information Systems, College of Computer and Information Sciences, Jouf University, Sakaka 72388, Saudi Arabia

Corresponding authors: Fatima Al-Quayed (ffalquayed@ju.edu.sa) and Mamoona Humayun (mahumayun@ju.edu.sa)

This work was supported by the Deanship of Scientific Research at Jouf University through the Fast-track Research Funding Program.

ABSTRACT Industry 4.0 is fundamentally based on networked systems. Real-time communication between machines, sensors, devices, and people makes it easier to transmit the data needed to make decisions. Informed decision-making is empowered by the comprehensive insights and analytics made possible by this connectedness in conjunction with information transparency. Industry 4.0-based wireless sensor networks (WSNs) are an integral part of modern industrial operations however, these networks face escalating cyber-security threats. These networks are always vulnerable to cyber-attacks as they continuously collect data and optimize processes. Increased connections make people more susceptible to cyberattacks, necessitating the use of strong cybersecurity measures to protect sensitive data. This study proposes a predictive framework intended to intelligently prioritize and prevent cybersecurity intrusions on WSNs in Industry 4.0. The proposed framework enhances the cybersecurity of WSNs in Industry 4.0 using a multi-criteria approach. It implements machine-learning and deep-learning algorithms for cybersecurity intrusion detection in WSNs of Industry 4.0 and provides prevention by assigning priorities to the threats based on the situation and nature of the attacks. We implemented three models, i.e., Decision Tree, MLP, and Autoencoder, as proposed algorithms in the framework. For multidimensional classification and detection of cybersecurity intrusions, we implemented Decision Tree and MLP models. For binary classification and detection of cybersecurity intrusions in WSNs of Industry 4.0, we implemented Autoencoder model. Simulation results show that the Decision Tree model provides an accuracy of 99.48%, precision of 99.49%, recall of 99.48%, and F1 score of 99.49% in the detection and classification of cybersecurity intrusions. The MLP model provides an accuracy of 99.52%, precision of 99.5%, recall of 99.5%, and F1 score of 99.5% in the detection and classification of cybersecurity intrusions. The implementation of Autoencoder with binary classification yields an accuracy of 91%, a precision of 92%, a recall of 91%, and an F1 score of 91%. The benchmark models, i.e., Random Forest (RF) for multidimensional classification and Logistic Regression (LR) for binary classification, have also been implemented. We compared the performance of the benchmark models with the models implemented in the proposed framework, revealing that the models in the proposed framework

The associate editor coordinating the review of this manuscript and approving it for publication was Emanuele Lattanzi.

significantly outperformed the benchmark models. The framework presents an intelligent prioritizing methodology that is significant for effectively identifying and addressing high-risk intrusions. The proposed framework implements a proactive preventive system that functions as a strong defensive wall by quickly putting countermeasures in place to eliminate threats and increase network resilience.

INDEX TERMS Cybersecurity, WSN, detection, prediction, intrusions, machine learning and deep learning.

I. INTRODUCTION

Industry 4.0 denotes a paradigm shift in the manufacturing and industrial processes, defined by the combination of automation, data sharing, and digital technologies [1], [2]. It expands on previous industrial revolutions, which included the utilization of steam power and water, electricity, and computers, and now unites the digital, biological, and physical domains [2], [3]. Industry 4.0 is fundamentally based on networked systems. Real-time communication between machines, sensors, devices, and people makes it easier to transmit the data needed to make decisions. Informed decision-making is empowered by the comprehensive insights and analytics made possible by this connectedness in conjunction with information transparency [3], [4]. In Industry 4.0, technological innovations like artificial intelligence (AI), machine learning, and augmented reality are essential components [5]. These developments offer intelligent assistance, streamlining procedures and raising output for a range of industrial jobs. Cyber-physical systems autonomously decide what to do based on the information they collect, which makes industrial environments respond faster and more effectively [2]. Strengthened cybersecurity measures are required due to the growing connection. With the growing digital infrastructure, safeguarding systems and data from cyber threats becomes sensitive [5], [6], [7], [8], [9]. According to Statista, the highest number of cyber-attacks in the manufacturing industry between January 2022 and March 2023 was detected in May 2022 with 32 incidents as shown in Figure 1. In December 2022, the sector saw four attacks, the lowest number of incidents in the measured period. In January 2023, this number had an uptick, reaching 20 attacks [10].

Industry 4.0 has a broad impact on many different industries. Smart factories are starting to take shape in the manufacturing sector, using robotics, IoT, and AI to run autonomous and effective production lines [2]. Improved visibility and traceability help supply chain management by streamlining logistics and cutting down on waste. Industry 4.0 is utilized by sectors such as healthcare, automotive, and agriculture [1], [2], [9], [11], [12]. IoT gadgets and data analytics enhance patient care and equipment upkeep in the medical field [13]. In order to produce smarter and self-driving cars, the automotive industry uses automation and networking [14]. Precision farming is beneficial to agriculture because it maximizes crop yields and resource utilization through the use of smart sensors and data analysis. The core component of Industry 4.0 is Wireless Sensor Network (WSN), which is a network of interconnected sensors and equipment that communicate wirelessly in industrial environments [15], [16]. These networks play a key role in the collection, transmission, and analysis of real-time data that is essential for process optimization and decision-making. They serve as the industrial setup's nerve center, gathering data from multiple locations, including motion, temperature, pressure, and other characteristics. The capacity of WSNs to create smooth communication between devices, systems, and people is one of its main advantages. The Industrial Internet of Things (IIoT), which provides a thorough and integrated view of the entire manufacturing or industrial environment, is made possible by this interconnection [17], [18]. It serves as the cornerstone of an infrastructure that is linked and data-driven, facilitating effective data sharing and analysis. WSNs make it possible for employees to monitor and control industrial operations remotely, giving them greater authority. This feature is important for improving overall operational efficiency, decreasing downtime, and performing predictive maintenance. The gathered information provides information about production procedures, resource usage, and possible areas for optimization, which forms the foundation for well-informed decision-making [6], [13], [19], [20].

In contrast to conventional wired systems, WSNs provide more scalability and flexibility. They are easily expandable, reconfigurable, and suitable for a variety of industrial environments [21]. However, there are several security challenges with these networks. Increased connection makes people more susceptible to cyberattacks, necessitating the use of strong cybersecurity measures to protect sensitive data [22]. The performance of these networks is impacted by problems with signal interference and dependability in intricate industrial environments. The use of battery-powered sensors raises additional concerns because prolonged battery life and energy efficiency must be balanced for consistent and dependable operation [23]. Despite security challenges, WSNs are essential to the operation of Industry 4.0's data-driven, networked infrastructure. To fully utilize these networks in the context of the fourth industrial revolution, security problems are required to be addressed more efficiently [1], [2], [6], [21], [24].

Cybersecurity breaches within WSNs are a major concern in modern industrial environments. These networks are always vulnerable to cyber-attacks as these networks continuously collect data and process optimization [25]. The incorporation of diverse sensors and wirelessly connecting equipment has rendered these networks susceptible to cyber-attacks, hence posing a risk to safety in industrial environments, disrupting operations, and compromising data. These wireless sensor networks are vulnerable due to their interconnectedness, which is also a benefit for smooth data transfer [19], [26], [27]. Cybercriminals take advantage of these flaws to obtain private information or take over vital systems. Attacks of this kind have the potential to be extremely damaging, impacting manufacturing lines, jeopardizing quality, and even endangering worker safety [28], [29]. These networks are working continuously in real-time, thus it is imperative that cyber-attacks be found and stopped as soon as possible. Prolonged downtimes or irreversible damage to industrial systems could result from operations being negatively impacted by delayed detection or response to intrusions [27], [30]. Thus, to safeguard these wireless sensor networks in Industry 4.0, a strong cybersecurity framework with an efficient prioritization process is required. It is important to create proactive cybersecurity tactics using intelligent algorithms and predictive models. These models monitor network activity continually, searching for unusual or suspicious patterns that could point to possible security vulnerabilities. Prioritizing these risks according to their possible impact and severity enables quick and efficient actions, reducing the effect of any prospective incursions.

A. RESEARCH MOTIVATION

In industry 4.0, WSNs serve as the backbone of networked industrial operations [2], [16]. The development of a comprehensive predictive framework for cybersecurity intrusion detection and prevention with a feature of intelligent prioritized is the significant requirement in this context [6]. WSNs are essential communication components for real-time data transfer because to the convergence of digital technologies, however, they are also susceptible to cyberattacks [23], [31]. The understanding of the vital role these networks play in the operation of contemporary enterprises serves as the driving force. They are a prime target for cyberattacks because they enable the smooth flow of data that powers decision-making and process optimization. Beyond only compromising data, a cybersecurity breach within these networks might have far-reaching consequences. It could cause serious financial losses, interfere with operations, and jeopardize safety [5], [16], [19]. The dynamic and ongoing nature of cyber threats necessitates the creation of a framework with well-considered priorities. There is always a chance that certain incursions may be more dangerous than others and will have a major impact on operations [32]. Therefore, the driving force is to create a system that can differentiate between various threat levels so that a targeted and effective response may be made. By placing the most serious hazards at the front of the list, their impact on industrial operations is either prevented or minimized. A predictive framework is also intended to establish a proactive protection mechanism. The goal is to foresee and stop possible risks rather than responding to breaches after they happen. Through the analysis of past data and current network behavior, predictive models are able to identify patterns that indicate possible hazards before they manifest. This proactive strategy fits perfectly with Industry 4.0's fast-paced environment, where preventive actions are significant to protect against cyber threats.

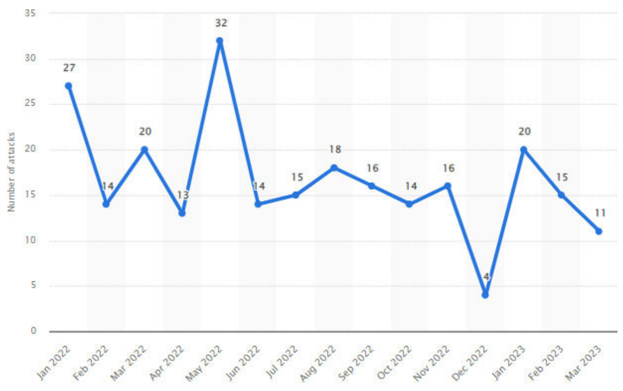


FIGURE 1. Monthly number of cyber-attacks in automotive production companies worldwide [10].

B. RESEARCH CONTRIBUTIONS

The proposed research presents an intelligently prioritized and robust predictive framework for cybersecurity intrusion detection and prevention in industry 4.0 based wireless sensor networks, with following research contributions.

- A framework for cybersecurity intrusion detection and prevention in industry 4.0 based wireless sensor networks will be developed.
- An AI-based detection mechanism will be implemented that recognize and classify cybersecurity intrusions. Three distinct machine learning models (multilayer perceptron, autoencoder, and decision tree) will be implemented for cybersecurity intrusions detection and classification within WSNs.
- An intelligent prioritization model will be implemented that can be used to give priorities to cyber threats based on their nature and impact.
- A prevention system will be implemented that can be used to efficiently and effectively mitigate the impact of cybersecurity intrusions.

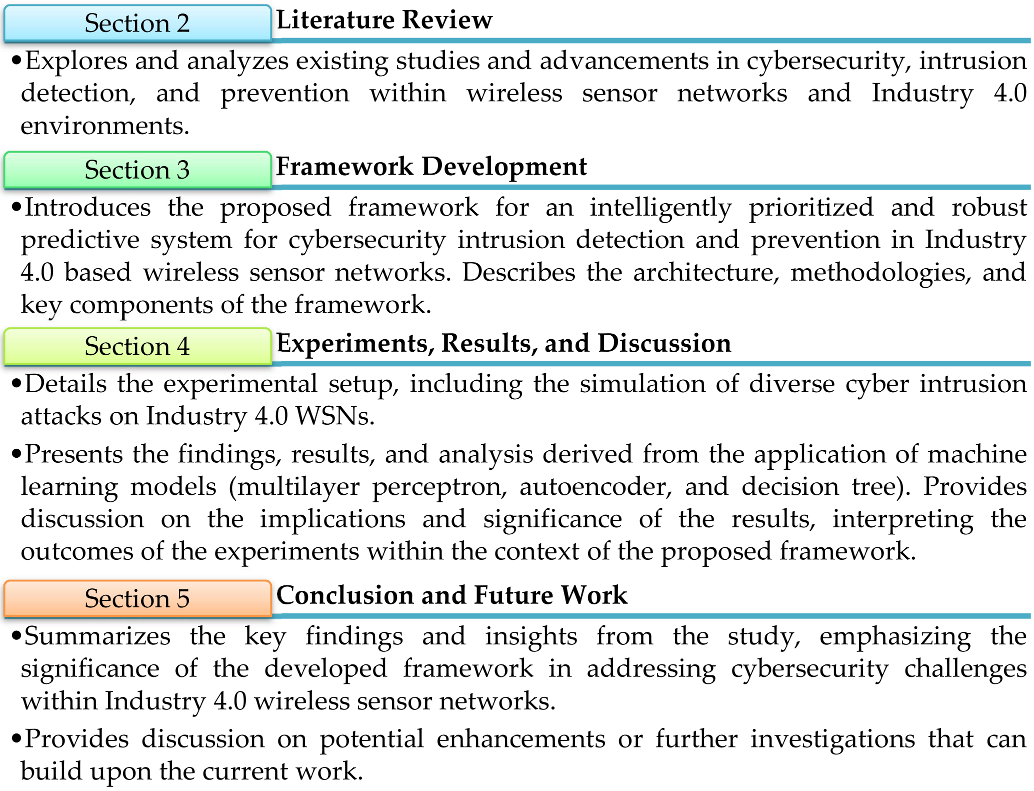
C. PAPER ORGANIZATION

The remaining part of the paper is structured as follows:

II. LITERATURE REVIEW

We explore and analyze the existing work in the fields of cybersecurity, WSNs, industry 4.0, threat detection, classification and prevention in WSNs.

Industrial machines have been a part of the market and manufacturing enterprises since the First Industrial Revolution and continue to be a part of the Fourth Industrial Revolution, also known as Industry 4.0. An increasing number of utilities are moving to Internet Protocol (IP)-based systems for wide-area communication as standardized protocols



Section 2 Literature Review

- Explores and analyzes existing studies and advancements in cybersecurity, intrusion detection, and prevention within wireless sensor networks and Industry 4.0 environments.

Section 3 Framework Development

- Introduces the proposed framework for an intelligently prioritized and robust predictive system for cybersecurity intrusion detection and prevention in Industry 4.0 based wireless sensor networks. Describes the architecture, methodologies, and key components of the framework.

Section 4 Experiments, Results, and Discussion

- Details the experimental setup, including the simulation of diverse cyber intrusion attacks on Industry 4.0 WSNs.
- Presents the findings, results, and analysis derived from the application of machine learning models (multilayer perceptron, autoencoder, and decision tree). Provides discussion on the implications and significance of the results, interpreting the outcomes of the experiments within the context of the proposed framework.

Section 5 Conclusion and Future Work

- Summarizes the key findings and insights from the study, emphasizing the significance of the developed framework in addressing cybersecurity challenges within Industry 4.0 wireless sensor networks.
- Provides discussion on potential enhancements or further investigations that can build upon the current work.

have grown in popularity. One of the standards that enables industries to obtain data directly from the machines through TCP/IP or RS323 communication is SECS/GEM [33]. The SECS/GEM protocol is mostly utilized within factories rather than in public spaces, some businesses might overlook its security characteristics. SECS/GEM communication is extremely vulnerable to several types of cyberattacks. The potential replay-attack hacks that could affect an SECS/GEM system are examined in [33]. This paper assumes an enemy who wishes to cause ongoing damage to an operation-based control system using replay attacks. In order to inject an external control input covertly, the adversary can intercept messages, watch and record their contents for a predefined period of time, record them, and then replay them when attacking. The purpose of the paper is to demonstrate the cyberattack vulnerability of SECS/GEM communication and to develop a detection system to guard against replay attacks. The findings show that replay attacks against SECS/GEM communications were identified and effectively stopped by the design mechanism.

Technological developments in the fields of digital electronics, wireless communications, and electro-mechanical systems have brought about a global revolution in society and economy. These developments have made it possible to develop sensor nodes that are inexpensive, power-efficient, and multifunctional [27]. By utilizing the sensing, data processing, and communication capabilities inherent in these nodes, sensor networks are realized. Despite the restricted energy capacity of wireless sensor network (WSN) nodes, the current intrusion detection systems inside WSN have even lower detection accuracy. The authors in [27] provide a hierarchical intrusion detection model that groups WSN nodes based on their functional roles in order to lower the energy consumption of nodes during detection processing. By evaluating and utilizing the multi-kernel function, the authors get the best linear combination and construct a multi-kernel extreme learning machine for WSN intrusion detection systems. According to simulation results, the system is ideal for WSNs with limited resources because it not only significantly shortens detection times but also ensures excellent detection accuracy.

Information security lapses and privacy violations are serious problems for both individuals and businesses, according to earlier research [34]. It is recognized that reducing risk in this area necessitates taking into account both the technological and human components of information security. Most of the risks to an organization's information assets are caused, whether on purpose or accidentally, by its employees. The study in [34] offers a novel conceptual framework that combines preventative and deterrent strategies to reduce the danger of insider attacks. Situational crime prevention factors motivate employees to stop information security misconduct, whereas deterrence factors dissuade them from acting improperly in terms of information security within organizations. The results demonstrate that people's attitudes are strongly influenced by their perceptions of the certainty and severity of consequences, which serves as a deterrent to information security malfeasance.

In order to create a new conceptual model of hybrid threats that incorporates deception techniques, the study [32] investigates the cyber-deception-based approach. Preventive techniques are the main emphasis of security programs since they keep hackers out of the network. In an effort to identify and thwart attackers before they can enter, these programs detect and block malicious activity in an effort to use hardened perimeters and endpoint defenses. The majority of businesses use layered preventive measures to strengthen their networks with defense-in-depth. Detection controls are not as frequently used for in-network threat detection as they are to support perimeter prevention. This architecture has detection gaps that are hard to cover with current security measures that are not tailored to that role. Defenders are implementing a more balanced approach that incorporates detection and response in place of relying just on prevention, a tactic that attackers have regularly been successful against [32]. The majority of businesses use next-generation firewalls or intrusion detection systems (IDS) to identify known threats by identifying patterns in the data. Other detection methods make use of behavioral analysis, traffic, or monitoring. Reactive defenses are meant to identify an attack once it occurs, however they frequently fall short. Their inability to detect attacks based on what appears to be authorized access or credential harvesting is another reason for their shortcomings. They contribute to analyst alert fatigue by being perceived as complicated and prone to false positives. Recent innovation in the security sector has concentrated on developing more precise methods of identifying hostile activities using technologies like big data, artificial intelligence (AI), deception, user and entity behavioral analytics (UEBA), and deception [32].

The IoT environment is made up of dispersed nodes, servers, and software for efficient communication and it is essential to many industries, including the automotive and medical tracking sectors [29]. Existing intrusion detection approaches are unable to withstand attacks that pose a threat to security and privacy, despite the fact that this IoT paradigm has been plagued by such threats and attacks. In order to counter these dangers and attacks, the sparse convolute network has been used to analyze the IoT infiltration threat. The internet is trained with sets of intrusion data, traits, and questionable activities to help detect and follow attacks, particularly Denial of Service (DDoS) attacks. In addition, the network is optimized by the application of evolutionary approaches that recognize and track error, regular, and intrusion efforts under various scenarios. Neurons in the sparse network evaluate complex hypotheses, and the resulting event stream outputs are routed to additional hidden layer processes. This procedure reduces the amount of intrusion involved in the transfer of IoT data. Standard and threat patterns are successfully classified in the network by the efficient use of training patterns [29]. The system's efficacy is assessed through the analysis of experimental findings and conversations. When it comes to network security, network intrusion detection systems outperform other forms of conventional network defense. Using an autoencoder network model and an enhanced evolutionary algorithm to detect intrusions, the research used an IGA-BP network to tackle the rising problem of Internet security in the big data era. It was constructed with MATLAB, which guarantees a performance ratio of 90.26%, a detection rate of 98.98%, and accuracy of 99.29% with little processing complexity. In the future, a meta-heuristic optimizer was employed to improve the system's capacity to predict attacks.

The Smart Grid uses digital information and control technology to improve the efficiency, safety, and dependability of the electric grid. Techniques for state estimation and real-time analysis are essential to guaranteeing correct control implementation [28]. However, because Smart Grid systems depend on communication networks, there is a serious risk to grid stability as a result of their susceptibility to cyberattacks. Effective intrusion detection and prevention systems are crucial for reducing such risks. In order to identify distributed denial-of-service attacks on the communication infrastructure of the Smart Grid, the authors in [28] suggests a hybrid deep-learning approach. Our approach combines recurrent gated unit algorithms with convolutional neural networks. Two datasets were used: a bespoke dataset created with the Omnet++ simulator and the Intrusion Detection System dataset from the Canadian Institute for Cybersecurity. For attack surveillance and resilience, the authors also created a Kafka-based dashboard for real-time monitoring. Results from simulations and experiments show that our suggested method obtains a high accuracy rate of 99.86%.

Malware, advanced persistent threats, and distributed denial of service (DDoS) attacks all actively jeopardize the security and availability of Internet services [35]. In order to detect DDoS attacks, study in [35] suggests an intelligent agent system that uses automatic feature extraction and selection. In our experiment, we employed a custom-generated dataset called CICDDoS2019, and we found that the system outperformed the state-of-the-art machine learning-based DDoS attack detection approaches by 99.7%. The authors created an agent-based mechanism for this system that blends sequential feature selection with machine learning methods. When the system dynamically identified DDoS attack traffic, the system learning phase picked the best attributes and rebuilt the DDoS detector agent. With the use of the most recent CICDDoS2019 custom-generated dataset and automatic feature extraction and selection, our suggested approach outperforms the current standard in processing speed while meeting the most advanced detection accuracy.

The technique of cyber-resilience in small and medium-sized businesses (SMEs) is examined in [26], and a complete solution is suggested for identifying newly emerging threats that makes use of open-source tools for prescriptive malware analysis, detection, and response. A system that is specifically made for SMEs with up to 250 employees is developed by utilizing open-source software and solutions, with an emphasis on the identification of new dangers. The approach's usefulness in increasing SMEs' cyber-defense skills and bolstering their overall cyber-resilience is proved through thorough testing and validation, along with effective algorithms and methodologies for safety, security, and anomaly detection [26]. The results demonstrate the viability and scalability of using open-source resources to address the particular cybersecurity issues that small and medium-sized businesses confront. The suggested solution finds and analyses harmful activity within SME networks by fusing real-time threat intelligence feeds with sophisticated malware analysis techniques. Through the use of behavior-based analysis and machine-learning algorithms, the system is able to identify and categorize even the most complex strains of malware. Using real-world facts and scenarios, comprehensive testing and validation were carried out to assess the system's efficacy [26]. The approach effectively recognizes new threats that conventional security methods frequently overlook, as evidenced by the results, which show notable gains in malware detection rates. The suggested system is a workable and expandable approach that makes use of containerized apps and is easily implementable by small and medium-sized businesses looking to strengthen their cyber-defense capabilities.

Theft of intellectual property or security information, fraud, sabotage, and other destructive acts by authorized users are examples of insider risks [36]. Insider threats can do a great deal of harm even though they are far less common than external network attacks. Insiders have intimate knowledge of an organization's systems, making it challenging to identify their harmful activity. Conventional insider-threat detection techniques emphasize rule-based strategies developed by subject matter experts; nevertheless, they lack both adaptability and resilience. In [36], the authors offer approaches for insider threat identification based on anomaly detection algorithms and user behavior modelling. The authors created three different kinds of datasets using user log data: the user's weekly email communication history, the user's daily activity summary, and the user's email contents subject distribution. Then, in order to find malicious activity, the authors used four anomaly detection methods and their combinations. The outcomes of the experiments suggest that the suggested structure can function effectively for unbalanced datasets with little insider threats and no knowledge provided by domain experts.

Agriculture 4.0, the impending revolution in agriculture, incorporates state-of-the-art information and communication technologies into current processes. Security researchers are becoming more and more interested in various cyber threats associated with the previously described integration [37]. Fighting such attacks can greatly benefit from the application of Machine Learning (ML) techniques for network traffic analysis and classification. In this direction, the research work presents and assesses several machine learning classifiers for the classification of network traffic, including Random Forest (RF), Stochastic Gradient Descent (SGD), Decision Tree (DT), K-Nearest Neighbours (KNN), Support Vector Classification (SVC), and Random Forest (KNN), along with an ensemble model that uses both soft and hard voting. Three NSL-KDD dataset variations—the original dataset, the under sampled dataset, and the oversampled dataset—were used in the context the suggested study [37]. In all three dataset modifications, the effectiveness of each individual machine learning algorithm was assessed and contrasted with the voting ensemble methods' effectiveness. When compared to the individual models, it was discovered that both the hard and the soft voting models performed better in terms of accuracy in the majority of cases.

A. GAP ANALYSIS

From SECS/GEM communication flaws to IoT infiltration threats and intrusion detection systems in Smart Grids, the literature review highlights a wide range of cybersecurity issues that affect industries, businesses and technologies. However, it is reflected from the existing that there is a lack a robust and intelligently prioritized predictive framework in order to find and prevent cybersecurity intrusions in Industry 4.0-based WSNs. The existing studies provide discussion on the problems that WSNs face and how important it is to have good attack detection systems, but it does not go into great detail about a predictive framework that uses prioritization intelligence. Industry 4.0 depends on devices and systems that are linked to each other, which makes WSNs an important part. It is important to deal with the unique problems they create. To make sure that industrial WSNs are safe and reliable, the system is required to include machine learning algorithms with the ability to find outliers, and real-time threat analysis.

III. A PREDICTIVE FRAMEWORK FOR CYBERSECURITY INTRUSION DETECTION AND PREVENTION IN INDUSTRY 4.0 BASED WIRELESS SENSOR NETWORKS

We propose an intelligently prioritized and robust predictive framework for cybersecurity intrusion detection and prevention in Industry 4.0 based wireless sensor networks. The framework includes several essential component as given in Figure 2. These components include industry 4.0, WSN, intrusion based cyber-attacks, AI-based detection and classification of cybersecurity intrusions, and intelligent prioritization and prevention system. The framework provides a specialized system designed to identify and prevent cybersecurity intrusions in wireless sensor networks. Three different machine learning models i.e., multilayer perceptron, autoencoder, and decision tree have been used in an AI-driven detection method. These models will make it possible to identify and categories various cybersecurity attacks, improving the network's capacity to quickly detect and address threats. The framework also presents an intelligent prioritization model, a significant component that rank various cyber threats according to their characteristics and their consequences. By focusing on and responding to high-risk intrusions first, this prioritization model helps the network to allocate resources more effectively to counter the most serious attacks. In addition, a proactive preventive system will be included to lessen the effects of cybersecurity breaches.

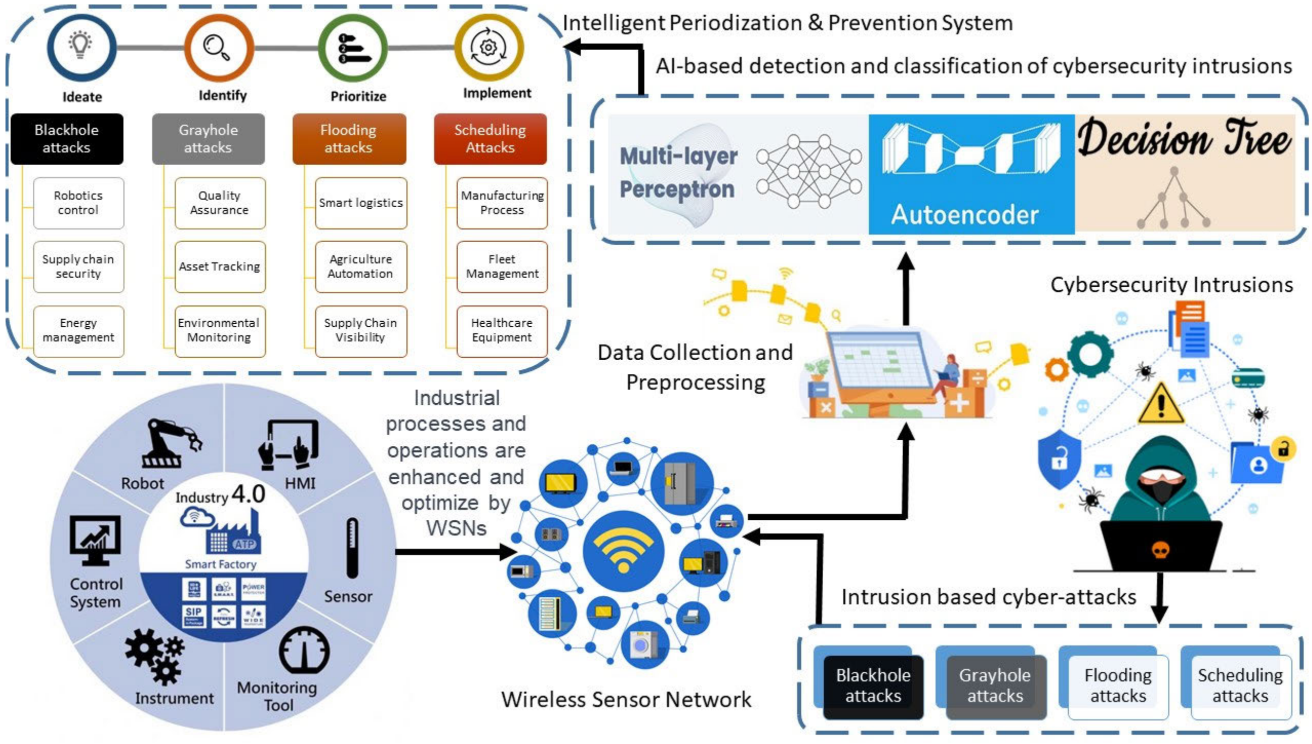


FIGURE 2. A predictive framework for cybersecurity intrusion detection and prevention in industry 4.0 based wireless sensor networks.

In addition to being able to identify threats, this system is built to serve as a strong defensive barrier, putting in place countermeasures that quickly and effectively eliminate any threats and increase the network's overall resilience. This all-encompassing framework aims to strengthen the security of industry 4.0-based wireless sensor networks through the integration of AI-based detection techniques, implements threat prioritization, and proactive preventive tactics.

A. COMPONENTS OF THE PROPOSED PREDICTIVE FRAMEWORK FOR CYBERSECURITY INTRUSION DETECTION AND PREVENTION IN WSN

1) INDUSTRY 4.0

The integration of modern digital technology into manufacturing and industrial processes is embodied in Industry 4.0, the fourth industrial revolution. Through the integration of cloud computing, AI, cyber-physical systems, and the IoT, it transforms traditional industries [5]. The IoT, a network of interconnected devices with sensors and actuators that gather and share data in real time, is a key component of Industry 4.0. AI and advanced analytics are then used to interpret and analyses data, allowing for autonomous decision-making, optimization, and predictive analysis. This revolution has an impact on many different industries. Smart factories use robotics, IoT, and data analytics to improve overall efficiency, forecast maintenance needs, and maximize productivity. IoT-enabled monitoring solutions in logistics and supply chain management offer visibility and efficiency all the way through the supply chain, cutting down on delays and enhancing inventory control [38]. IoT devices and data analytics are important in the healthcare industry as well, as they enhance patient care through individualized therapies, remote monitoring, and effective resource management in hospitals [13]. Industry 4.0 is comprised of three major sectors including smart manufacturing, supply chain management and healthcare. In smart manufacturing, factories are empowered by interconnected sensors and IoT devices that gather real-time data, enabling predictive maintenance and optimize production lines. In supply chain management, IoT devices are used for inventory tracking, monitoring transport conditions, and ensuring efficient delivery of goods. In healthcare sector, hospitals utilize IoT devices for remote patient monitoring, inventory management, and provision of healthcare services.

2) WSNs IN INDUSTRY 4.0

A key element of Industry 4.0 is Wireless Sensor Networks (WSNs) that enable real-time monitoring and data collection across a range of industrial applications. WSNs facilitate smart factories, streamline workflows, and enhance decision-making [39], [40]. These networks are made up of spatially dispersed, autonomous sensors that work together to monitor and collect data in a variety of environments via wireless communication [31], [41], [42]. There are following key feature of WSNs in Industry 4.0:

- WSNs make it easier to gather data in real time from sensors positioned across industrial environments. Temperature, pressure, humidity, and other factors that are essential for monitoring and controlling systems are frequently included.
- WSNs create a network that allows sensors to communicate to central systems and to each other. The smooth conveyance of data made possible by this connection promotes intelligent decision-making and process optimization.
- WSNs are made to be both scalable and flexible. They are adaptable for many Industry 4.0 use cases and may be expanded or changed to meet shifting industry requirements.
- WSNs are built with energy efficiency in mind. In order to provide continuous data collection and transmission, sensor nodes are frequently battery-powered and designed to last for long periods of time without requiring frequent maintenance.

In Industry 4.0, WSNs are applied across following sectors:

- **Smart Manufacturing:** WSNs make it possible to monitor manufacturing lines, inventory, and equipment, which guarantees preventative maintenance and streamlines workflows.
- **Predictive Maintenance:** Utilizing WSNs save downtime and increase operational efficiency by anticipating equipment breakdowns and maintenance requirements.
- **Environmental Monitoring:** In order to ensure regulatory compliance and create safer working environments, WSNs monitor environmental conditions in industrial environments.
- **Supply Chain Optimization:** WSNs provide real-time data to optimize supply chain management by tracking inventory and transit conditions.

3) INTRUSION-BASED CYBER-ATTACKS

Intrusion-based cyber-attacks comprise a diverse range of tactics employed to undermine the security and integrity of computer networks, systems, and information. These attacks are executed by taking advantage of holes or flaws in the systems that are being targeted. This category includes a number of different intrusion types, each with a unique approach and objective. Below is the detail of common intrusion based cyber-attacks:

- **Blackhole Attacks:** These attacks are also referred as packet drop attacks. Blackhole attacks happen when malicious nodes in a network discard or drop packets, preventing data from flowing normally. These nodes draw in network traffic, but instead of forwarding the packets, they drop them, which causes congestion on the network or information loss. This attack is especially harmful to WSNs because compromised nodes may drop packets in an attempt to save energy, losing important data in the process.
- **Grayhole Attacks:** Grayhole attacks are a variation of blackhole attacks. Instead of dropping all packets randomly, nodes in grayhole attacks drop packets selectively. The hacker uses network manipulation to intercept or disrupt particular data packets. Because this manipulation involves selective interference rather than a full packet drop, it can be more difficult to identify the malicious nodes.
- **Flooding Attacks:**In a flooding attack, an excessive volume of traffic is directed towards a system, preventing it from responding to valid requests. These attacks involve sending a lot of requests or data packets to the target, which overloads it and makes it unresponsive. Numerous systems are used to flood the target, increasing the impact of flooding attacks, such as Distributed Denial of Service (DDoS) attacks.
- **Scheduling attacks:** Scheduling attacks target the time synchronization mechanisms of WSNs. The goal of these attacks is to interfere with the network's scheduling or timing functions. Attackers affect overall functionality and reliability network by interfering with critical operations or creating anomalies in the network through timing manipulation. For example, in scheduling attacks the attacker might attempt to compromise the TDMA (Time Division Multiple Access)-based scheduling by manipulating the allocation of time slots/frames causing timing inconsistencies or collisions

When considering WSNs utilized in Industry 4.0 and IoT environments, each of the above intrusions presents a serious risk to the security and optimal operation of systems. Strong intrusion detection and prevention systems are required to find and prevent these intrusions.

4) DATA COLLECTION & PREPROCESSING

Preprocessing and data collecting are important steps in data-driven system. It entails obtaining raw data from several sources, organizing, and cleaning it to guarantee its quality and suitability. This stage is significant, particularly for cybersecurity frameworks and machine learning-based systems that seek to anticipate or identify attacks. Within a network, data can come from a number of locations. Within Industry 4.0 wireless sensor networks, the data comprise of sensor data, network traffic logs, system event logs, and pertinent information related to cybersecurity. The collected data include a variety of formats, including text from several sources, numerical sensor readings, and category system records. It is imperative to have access to both real-time streaming data and historical data. Historical data is useful for comprehending patterns and trends, whereas, real-time data is helpful in identifying persistent risks.

Data preprocessing involves cleaning and quality control. It filling in any gaps in the data, getting rid of duplicates, and fixing any discrepancies. During this stage, anomalies or outliers also recognized and dealt with. In order to bring all features to the same scale, data need to be normalized or transformed. This is important step for machine learning algorithms that are sensitive to different data scales. Producing pertinent characteristics from unprocessed data is necessary in cybersecurity. Extracting certain information from logs or sensor data is helpful to efficiently identify possible risks or anomalies. Reducing dimensions or getting rid of features that are unnecessary or less useful increase processing efficiency in big datasets without sacrificing important information. In order to detect patterns in the data for the learning algorithms, it must be labelled with the relevant classes or categories if the data is being used for supervised learning tasks like classification. The effectiveness of the analysis is greatly impacted by data preprocessing when it comes to cybersecurity intrusion detection. The accuracy and efficacy of machine learning models or detection systems are significantly influenced by the quality, relevance, and organization of the data.

5) AI-BASED DETECTION AND CLASSIFICATION OF CYBERSECURITY INTRUSIONS

The deployment of an AI-driven intrusion detection and classification system is used for safeguarding Industry 4.0 WSNs. Three different AI models i.e., Decision Tree, Multilayer Perceptron (MLP), and Autoencoder are implemented using publicly available WSN dataset. Cybersecurity intrusions are detected classified with a particular emphasis on flooding, scheduling, blackhole, and grayhole attacks. Each of the selected models has a distinct function in recognizing and categorizing cyber-attacks as given below:

- **Decision Tree:**This model organizes data into a tree-like structure and it makes decisions based on conditions. It is intuitive and can handle both numerical and categorical data which makes it suitable for classifying different types of attacks based on specific characteristics [43], [44].
- **Multilayer Perceptron (MLP):** The MLP is a type of neural network that is very good at finding complicated patterns and connections in data. The fact that it can learn from both structured and unstructured data makes it useful for finding both simple and complicated attack patterns [45], [46].
- **Autoencoder:**It is an unsupervised learning approach that is applied to data compression and feature learning. For anomaly detection, it is especially helpful. It recognizes anomalies or deviations from the learned patterns by recreating the input data [29], [47].

The system is specifically trained to recognize four primary types of common attacks in WSNs:

- **Blackhole Attacks:** Models aim to identify instances where packets are dropped or lost. It leads to data loss or network congestion.
- Grayhole Attacks: Models are used for identification of selective packet manipulation to disrupt specific data flows within the WSNs.
- **Flooding Attacks:** Models are used for recognition of patterns involving excessive traffic designed to overwhelm the network and disrupt normal operations.
- **Scheduling Attacks:** Models are used for detection of inconsistencies or manipulations in the WSNs' timing or scheduling mechanisms.

The publically available WSN dataset is structured, preprocessed, and used. This include feature engineering, cleaning, and dividing the dataset into subsets for testing and training. The dataset is used to train each AI model i.e., Decision Tree, MLP, and Autoencoder. By exposing the models to labelled data, they are able to pick up on and recognize patterns linked to various kind of attack. Metrics including accuracy, precision, recall, and F1 score are used to evaluate how well they detect and classify attacks. To increase the models' accuracy and resilience, their parameters are changed and fine-tuned. The objective is to develop a system that can precisely identify and categories various kinds of attacks within WSNs by utilizing these AI models. This strengthens the WSNs' resistance to different cybersecurity intrusions and is one of the core components of the proposed cybersecurity framework for Industry 4.0.

6) INTELLIGENT PRIORITIZATION AND PREVENTION SYSTEM

Intelligent Prioritization and Prevention System is the core component of the proposed framework. It sorts various kinds of attacks (blackhole, grayhole, flooding, and scheduling) by considering their importance in Industry 4.0 environments. Blackhole attacks are considered dangerous in sectors like robotics control, supply chain security and energy management. When these attacks happen, vital systems can be seriously affected, as dropped or lost packets can make communication difficult and even cause harm. Grayhole Attacks involve selectively changing packets and are most common in quality assurance, asset tracking, and environmental monitoring. They can have a direct effect on the accuracy of data, which can make quality control or tracking tasks difficult to complete. Flooding Attacks are sensitive in situations like smart logistics, farm automation, and supply chain visibility. They flood the network with too much traffic, which can seriously impede the smooth flow of data that is needed in these situations. Scheduling Attacks focus on manufacturing processes, fleet management, and healthcare equipment because they change the timing and scheduling systems. Timing problems can make important operations in these areas run late or not at all.

Let A be the set of attacks, a_i represents an individual attack, $I\left(a_i\right)$ be the importance of attack a_i in industry 4.0 environments, Category $\left(a_i\right)$ be the cateogry of attack a_i, Impact $\left(a_i\right)$ be the impact of attack $a_i$on the system. Then intelligent prioritization and prevention system can be represented by equation 1 and 2.

$$Attack_{Sorted} = Sort\left(A, I\left(a_i\right)\right) \tag{1}$$

$$Impact\left(a_i\right) = f\left(Category\left(a_i\right), Industry\ 4.0\left(Applications\right)\right) \tag{2}$$

In order to perform proactive prevention actions, we prioritize attacks based on their impact on Industry 4.0 scenarios. The following prevention tactics have been used to prevent cybersecurity threats:

- When the proposed framework detects blackhole attacks, it will prevent the system by activating and setting up an environment for verifying packets and making sure there are multiple paths for important communications in fields like robotics, supply chain management, and energy management.
- When the system detects grayhole attacks, it will utilize verification tools and data validation methods to prevent selective packet manipulation in quality assurance and asset tracking scenarios.
- In order to prevent flooding attacks, traffic analysis tools and rate-limiting methods will have been deployed to keep the network from getting too busy in smart logistics and supply chain visibility situations.
- Scheduling attack prevention includes the use of time synchronization procedures and backup plans to keep important time frames in manufacturing, fleet management, and healthcare equipment environments.

The intelligent prioritization and prevention system not only finds potential threats and ranks them, but it also makes sure that prevention plans are tailored to the unique weaknesses of each Industry 4.0 scenario. The goal of this proactive method is to make WSNs safer and more reliable.

B. PSEUDOCODE OF THE PROPOSED PREDICTIVE FRAMEWORK FOR CYBERSECURITY INTRUSION DETECTION AND PREVENTION IN WSN

Algorithm 1 shows the procedure of the proposed framework. It works with several components for cybersecurity intrusion detection and prevention in WSNs. The fundamental environment is Industry 4.0, which is the amalgamation of cloud computing, AI, cyber-physical systems, and the IoT. WSN is the core component of Industry 4.0 and is used for real-time data collection, transmission, scalability, and energy efficiency across a range of industrial applications. Blackhole, grayhole, flooding, and scheduling attacks are considered intrusion-based cyberattacks that jeopardize the security of WSNs. Data collection and preprocessing is performed by the process of collecting, cleaning, normalizing, and feature engineering data sources. Using labeled datasets and assessment criteria, AI-based detection and classification uses decision tree, multilayer perceptron, and autoencoder models to identify and classify cyberattacks. The Intelligent Prioritization and Prevention system develops customized preventative strategies to successfully minimize danger by classifying them according to how they affect various Industry 4.0 scenarios. With the use of proactive preventive measures, prioritization strategies, and AI-driven detection techniques, the proposed framework aims to strengthen the security of WSNs in Industry 4.0.

IV. EXPERIMENTS, RESULTS & DISCUSSIONS

We perform the simulations and evaluate the performance of proposed framework with respect to the cybersecurity intrusion detection and classification.

A. EVALUATION METRICS

We evaluated the performance of the models implemented in the proposed framework using accuracy, precision, sensitivity (recall), F1 score, specificity, and precision-recall curve [14]. For multidimensional classification and detection of cybersecurity intrusions, we used Decision Tree and MLP models. For binary classification and detection of cybersecurity intrusions in WSNs of Industry 4.0, we used Autoencoder model. We implemented the benchmark models, i.e., RF for multidimensional classification and LR for binary classification and compared the performance with the models implemented in the proposed framework. Specificity and precision-recall curves are the metrics applicable to binary classification models. Therefore, for multidimensional classification through Decision Tree, MLP, and RF models, we used accuracy, precision, sensitivity, and F1 score. Whereas, for binary classification through Autoencoder and LR models, we used specificity and precision-recall curve metrics in addition to accuracy, precision, sensitivity, and the F1 score. We calculate these performance metrics based on the following terms:

- True Positives (TP): The number of tuples that are really found to be intrusive at the end of the process.
- True Negatives (TN): The number of valid tuples that are found at the end of the detection process.
- False Positives (FP): The number of safe tuples that, at the conclusion of the detection process, are identified as intrusions.
- False Negatives (FN): The quantity of dangerous tuples that, at the conclusion of the detection process, are found normally.

When assessing the effectiveness of classification models, accuracy is a commonly used parameter. It assesses the overall accuracy of the model predictions by figuring out the proportion of correctly predicted cases among all the instances in the dataset [46]. Mathematically, it is represented by A and can be calculated with the help of equation 3.

$$A = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

Precision is a way to measure how well a classification model works. It checks how good the model is at making positive predictions by counting the number of true positives out of all positive predictions, or true positives plus fake positives [46]. Mathematically, it is represented by P and can be calculated with the help of equation 4.

$$P = \frac{TP}{TP + FP} \tag{4}$$

Sensitivity is a way to measure how well a classification model works. This number is also known as the recall or true positive rate. The sensitivity of the model measures how

Algorithm 1 A Predictive Framework for Cybersecurity Intrusion Detection and Prevention in WSNs

```
1.  Begin
2.  Input:    D_0:   Industry 4.0 based WSNs Data
3.  Output:   IDPA:  Intrusion Detection and Preventive Action
4.  Procedure:Cybersecurity intrusion detection and prevention (D_0)
5.  Industry 4.0 (I-4.0) Environment:
                    I-4.0 = {cloud computing, AI, IoT, WSNs}
6.  Wireless Sensor Networks (WSNs):
        WSNs = {Real time data collection, communication, scalability, energy efficiency}
7.  Intrusion-based Cyber Attacks (IBCA):
                $IBCA = \{Blackhole, Grayhole, Flooding, Scheduling\}$
                    $Attack_{Sorted} = Sort\ (A, I\ (a_i))$
8.  Data Collection & Preprocessing (S):
    $\leftarrow D_{Collected\&Preprocessing} = \{data\ source, cleaning, normalization, feature\ engineering$
9.  AI-based Detection & Classification (AIDC):
    while (Intrusion (I), WSNs (W)) do
10.     AIDC = {DT, MLP, AE, Evaluation Metrics, S}
    end while
11. Intelligent Prioritization and Prevention:
12. if (Threats Detected (TD)) do
13.     Prioritized Threat based on Industry 4.0 Environments (IE)
            $Impact\ (a_i) = f\ (Category\ (a_i), Industry4.0\ (Applications))$
        while (TD, IE) do
                Preventive Actions
        end while
14. end if
15. ReturnIDPA
16. end
```

well it can find every single positive case in the dataset [46]. Mathematically, it is represented by R and can be calculated with the help of equation 5.

$$R = \frac{TP}{TP + FN} \tag{5}$$

The F1 score demonstrates how well classification models perform when selecting between two choices. The F1 score is useful when there is a difference between accuracy and recall [46]. Mathematically, it is represented by F1-S and can be calculated with the help of equation 6.

$$F1 - S = 2 \times \frac{P \times R}{P + R} \tag{6}$$

Specificity is the performance metric specifically used in the evaluation of binary classification. It is used to measure the ability of a model to correctly identify negative instances out of all actual negatives. Mathematically, it is represented by S and can be calculated with the help of equation 7.

$$S = \frac{TN}{TN + FP} \tag{7}$$

A precision-recall curve is a graphical representation of the trade-off between precision and recall for different classification thresholds. The precision-recall curve is created by varying the classification threshold of the model and determining the precision and recall at each threshold. A higher area under the precision-recall curve (AUC-PR) indicates better performance for the model.

B. DATASET

In order to evaluate the working of proposed framework, WSN-DS: A dataset for intrusion detection systems in wireless sensor networks [48], a publically available dataset on a Kaggle website has been used. The dataset replicates many Denial-of-Service (DoS) attacks on WSN using the LEACH (Low Energy Adaptive Clustering Hierarchy) protocol. It includes Blackhole, Grayhole, Flooding, and Scheduling attacks, which are four different categories of attacks. The goal of these attacks is to determine how they affect network performance and what effects they have on the LEACH protocol. In the Blackhole attack, at the beginning of a round, an attacker assumes the identity of a Cluster Head (CH). When nodes connect to this fake CH, they unintentionally submit their data packets to it, which are then transmitted to the Base Station (BS). Data loss results from the Blackhole attacker's dropping or discarding of these packets rather than transmitting them. As with the Blackhole attack, attackers assume the identity of CHs in the Grayhole assault. These attackers may do this on the basis of the sensitivity of the data included in the packets they drop or delete. The goal of the flooding attack is to flood the network with too many high-transmission-power advertising CH messages. The sensor nodes' energy is depleted as they process the barrage of messages and choose which CH to join. The Scheduling attack takes place in the setup stage of the LEACH protocol. Assuming the role of CHs, attackers provide every node the same time slot for data transmission, which causes packet collisions and eventual data loss.

C. EXPERIMENTAL DESIGN

The models implemented in the proposed framework have been evaluated with "WSN-DS," a dataset for intrusion detection systems in wireless sensor networks. The dataset has been divided into two parts: the training set and the test set. The training set comprised 80% of the total records in the dataset. It was used to train the proposed models. On the other hand, the test set comprised 20% of the total number of records. It was used to test and validate the proposed model. Cross-validation was performed for Decision Tree and MLP through the "cross_val_score" function from scikit-learn. However, for Autoencoder, we validated the model using a train-test split with an 80:20 ratio since cross-validation is not applicable due to its unsupervised learning nature. All experiments are implemented in Python on a GPU based environment with 1.8 GHz CPU and 12 GB of RAM. Predefined machine learning packages and libraries, namely Pandas, Numpy, Seaborn, Sklearn, LabelEncoder, OneHoTencoding and Matplotlib have been implemented.

D. RESULTS AND EVALUATION

The experiments were performed by implementing three AI models, i.e., Decision Tree and MLP for multidimensional classification and the Autoencoder for binary classification. We also implemented the benchmark models, i.e., Random Forest (RF) [49] for multidimensional classification and Logistic Regression (LR) [50] for binary classification and compared the performance with the models implemented in the proposed framework. The evaluation results are highlighted below:

1) DECISION TREE AND RF

The Decision Tree model is used to detect cybersecurity intrusions with essential Python libraries and modules including Pandas and Sklearn. The Decision Tree approach performs well in classifying different kinds of WSN intrusions and achieve the accuracy of 99.48% as compared with Random Forest with an accuracy of 98%. Table 1 shows the values of accuracy, precision, recall, and F1 score of different kinds of attacks. The Decision Tree performs better at correctly identifying typical behavior, which is essential for cybersecurity. Its accuracy in differentiating between various intrusion kinds is strong. The macro average and weighted average rows show the overall model performance, which is consistently high across the dataset for Decision Tree model. This indicates a robust performance of Decision Tree in classifying instances within this multiclass classification problem.

In identifying typical network behavior, the Decision Tree model performs significantly well by attaining high precision, recall, and F1-scores as shown by Figures 3, 4 and 5. Figure 3 shows precision score per class, in which the decision tree model performs robustly in classifying all the instances. The model achieves a significant 99.77% precision for the "Normal" class, which is noteworthy. This suggests a high degree of precision in detecting typical occurrences, reducing false positives. With precisions of 98.82% and 97.89%, respectively, the model also performs well for the "Blackhole" and "Grayhole" classes, demonstrating its efficacy in correctly classifying them. The "Flooding" class observes 95.17% accuracy. For the "TDMA" class, the precision is 91.51%, indicating a little higher chance of false positives. The decision tree model performs well in all the classes, demonstrating its effectiveness. Figure 4 shows recall score per class, which indicates that the decision tree model performs well to accurately classify classes. With notable values for "Normal" (99.74%) and "Blackhole" (98.78%), the model demonstrates good recall across most classes, suggesting that it can catch most real occurrences for these classes. With recall score of 93.66% and 93.89%, respectively, "Flooding" and "TDMA" perform well, while the "Grayhole" class also exhibits great recall of 97.89%. These findings imply that the model minimizes false negatives by efficiently identifying instances of network penetration. Figure 5 shows the F1 score per class, which indicates a thorough assessment of the performance of the Decision Tree model. The results show a commendable balance between precision and recall across several classes. The model notably attains high F1-scores for "Blackhole" (98.80%) and "Normal" (99.75%), indicating a successful trade-off between reducing false positives and false negatives for these classes. With a balanced F1-score of 97.89%, the "Grayhole" class exhibits good recall and accuracy performance. The F1-scores of 94.41% and 92.68% for "Flooding" and "TDMA," respectively, show how well the model balances recall and precision for these classes. This well-balanced performance highlights that the model detects cybersecurity intrusions and handles failures in a comprehensive way. The efficacy of model in enhancing cybersecurity within the WSN is demonstrated by its capacity to effectively distinguish between typical and intrusive network activity.

The confusion matrix of Decision Tree is given by Figure 6 illustrates the performance of the classification model in differentiating among five discrete classes, namely TDMA, Blackhole, Flooding, Grayhole, and Normal. In this matrix, each row shows the real instances of a certain class, and each column shows the predicted value of that class. The model is good at finding instances of the "Normal" class; it correctly predicted 67787 cases. However, it shows difficulty in accurately distinguishing between classes like Blackhole and Grayhole, as well as between Normal and TDMA due to the similarities in their features or actions.

Figure 7 shows a graph indicating the true positive rate and false positive rate of the Decision Tree model for various attack types. It is represented by the ROC (Receiver Operating Characteristic) curve. The true positive rate is plotted on the y-axis, while the false positive rate is plotted on the x-axis. The area under the ROC curve is a measure of the overall performance of the model for the given attack types: blackhole, grayhole, flooding, TDMA, and normal. Higher values for all the classes indicate that the decision tree model performs significantly well in the identification and classification of cybersecurity intrusions.

TABLE 1. Accuracy, precision, recall and F1 score for decision tree and RF models.

Class	Precision		Recall		F1-score		Accuracy	
	DT	RF	DT	RF	DT	RF	DT	RF
Blackhole	98.82%	96%	98.78%	0.93%	98.80%	91%	99.48%	98%
Flooding	95.17%	90%	93.66%	0.94%	94.41%	100%		
Grayhole	97.89%	86%	97.89%	0.81%	97.89%	76%		
Normal	99.77%	99%	99.74%	0.99%	99.75%	100%		
TDMA	91.51%	100%	93.89%	0.93%	92.68%	88%		
Macro average	96.63%	94%	96.79%	0.92%	96.71%	91%		
Weighted average	99.49%	98%	99.48%	0.98%	99.49%	98%		

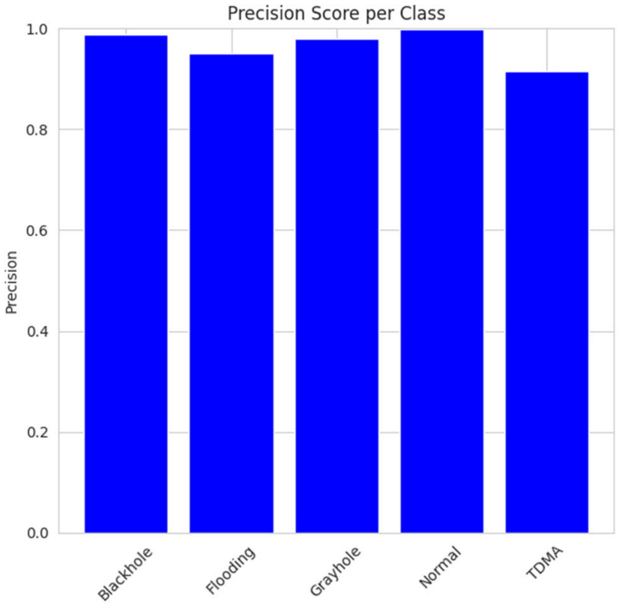


FIGURE 3. Precision score per class for decision tree model.

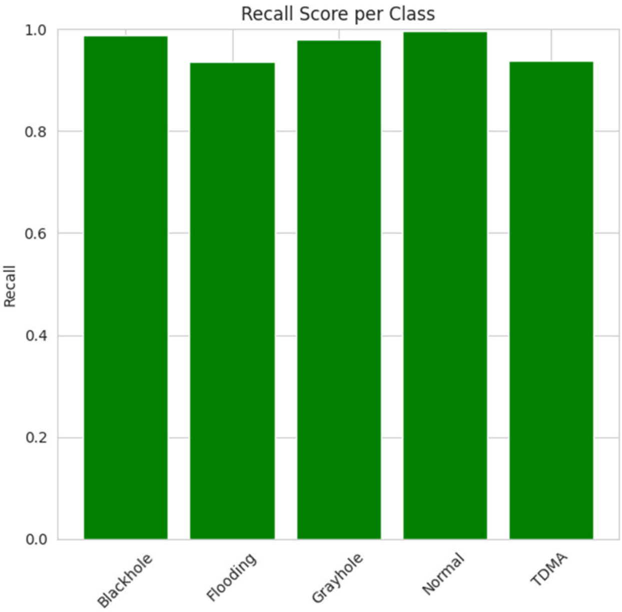


FIGURE 4. Recall score per class for decision tree model.

2) MLP AND RF

The MLP model is used to detect cybersecurity intrusions with essential Python libraries and modules including Pandas and Sklearn. The MLP approach performs well in classifying different kinds of WSN intrusions and achieve the accuracy of 99.52% as compared with RF model with an accuracy of 98%. Table 2 shows the values of accuracy, precision, recall, and F1 score of different kinds of attacks. The MLP model exhibits noteworthy precision rates, especially for the Normal and TDMA classes, suggesting a high degree of accuracy in the positive predictions. The TDMA class, on the other hand, shows a slight lower recall rate, indicating that some real TDMA instances are absent from the MLP model. The Blackhole, Flooding, and Grayhole classes show a good balance between precision and recall, as evidenced by their high F1-Score values of MLP model. The 'Macro Avg' and 'Weighted Avg' rows show the overall MLP model performance, which is consistently high across the dataset as compared with RF model. This indicates a robust performance of MLP model in classifying instances within this multiclass classification problem.

The MLP model ability to correctly identify instances within the Blackhole, Flooding, Grayhole, Normal, and TDMA categories is indicated by the Figures 8, 9 and 10. Figure 8 shows the precision score per class for the MLP model, which indicates that the ''Normal'' class has exceptional precision of 99.7%, suggesting that 99.7% of the time the model correctly predicts an instance to be ''Normal.'' For

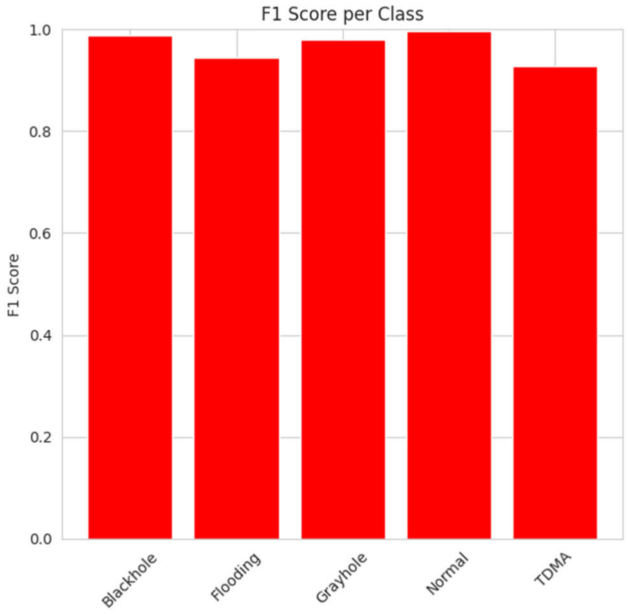


FIGURE 5. F1 score per class for decision tree model.

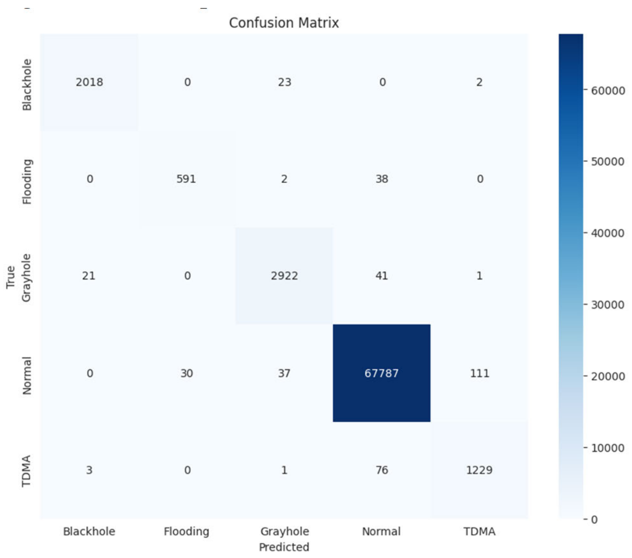


FIGURE 6. Confusion matrix for decision tree model.

each of these categories, the ''TDMA'' class similarly reaches a precision of 99.7%, demonstrating the resilience of the model in reducing false positives. With a precision of 98.1%, the ''Grayhole'' class comes in close behind, indicating the ability of the model to correctly categorize this category. The precision for the ''Blackhole'' (96.1%) and ''Flooding'' (93.6%) classes is still strong, although slightly lower. This suggests that the model effectively reduces false positives for these types of network intrusions. Figure 9 shows the recall score per class for the MLP model, which indicates that the model exhibits strong recall values of 96.8% for the ''Grayhole'' class. It shows the proficiency of the model in capturing the majority of actual instances in this category. The ''Normal'' class achieves an exceptional recall of 99.9%, indicating its ability to identify the vast majority of actual ''Normal'' instances with minimal false negatives. The ''Blackhole'' class follows closely with a recall of 99.2%, demonstrating the high sensitivity of the model. The ''Grayhole'' class also displays a strong recall of 96.8%, indicating the proficiency of the model in capturing the majority of actual instances in this category. The ''Flooding'' class achieves a good recall of 95.1%, and the ''TDMA'' class demonstrates a slightly lower recall of 90.4%. Figure 10 shows the F1 score per class for the MLP model, which indicates that with an F1-Score of 99.8%, the ''Normal'' class stands out as having a precisely balanced recall and precision for correctly identifying instances. Strong F1-Scores of 97.6% and 97.5%, respectively, are also displayed by the ''Blackhole'' and ''Grayhole'' classes, demonstrating the capacity of the model to successfully minimize both kinds of mistakes. Even though the F1 scores for ''Flooding'' and ''TDMA'' are a little bit lower at 94.3% and 94.8%, respectively, these numbers still show a well-rounded performance in keeping the balance between precision and recall.

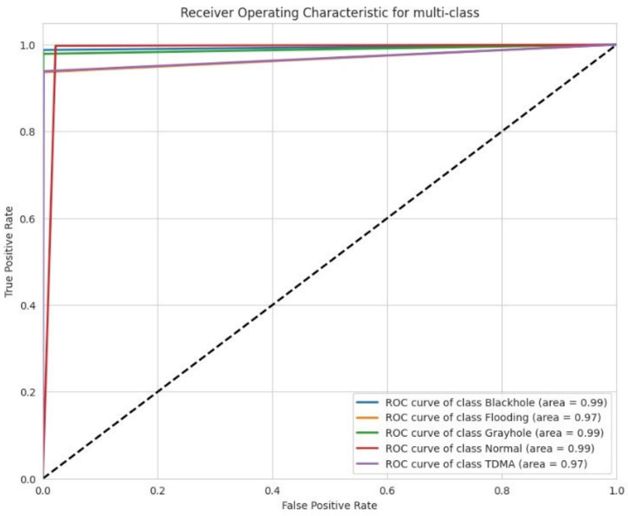


FIGURE 7. Decision tree receiver operating characteristic curve.

Figure 11 illustrates the confusion matrix for the MLP model, which indicates the performance in differentiating among five separate classes, i.e., TDMA, Blackhole, Flooding, Grayhole, and Normal. With no misclassifications, the model shows excellent prediction power for Blackhole, correctly recognizing 2022 instances. It also does a good job at identifying grayhole instances, correctly predicting 2765 instances, though with a little misperception. The model does a fair job of identifying instances of the Flooding and TDMA classes, but it has trouble telling normal instances apart from the other classes. In particular, it incorrectly labels cases of flooding as normal and normal instances as grayhole, TDMA, and flooding. The model shows a good capacity to find instances within each class, although it struggles to differentiate between flooding, normal, and grayhole classes.

TABLE 2. Accuracy, precision, recall, and F1 score for MLP and RF models.

Class	Precision		Recall		F1-score		Accuracy	
	MLP	RF	MLP	RF	MLP	RF	MLP	RF
Blackhole	96.1%	96%	99.2%	0.93%	97.6%	91%	99.52%	98%
Flooding	93.6%	90%	95.1%	0.94%	94.3%	100%		
Grayhole	98.1%	86%	96.8%	0.81%	97.5%	76%		
Normal	99.7%	99%	99.9%	0.99%	99.8%	100%		
TDMA	99.7%	100%	90.4%	0.93%	94.8%	88%		
Macro average	97.5%	94%	96.3%	0.92%	96.8%	91%		
Weighted average	99.5%	98%	99.5%	0.98%	99.5%	98%		

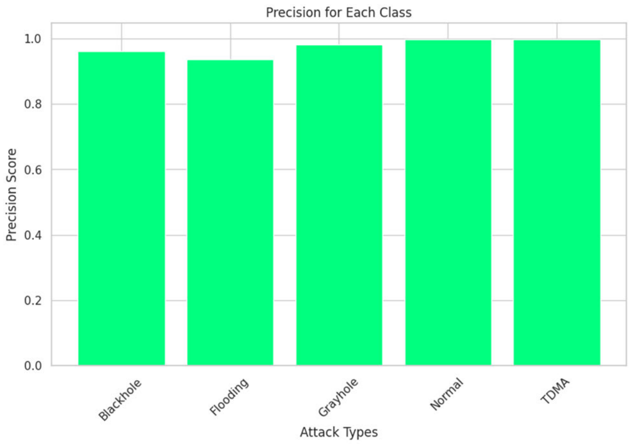


FIGURE 8. Precision score per class for MLP model.

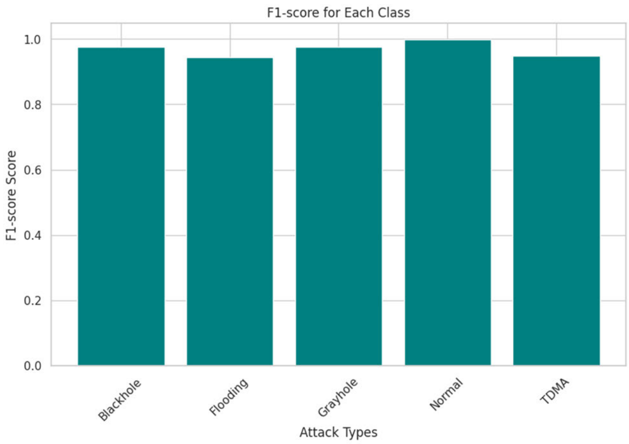


FIGURE 10. F1 score per class for MLP model.

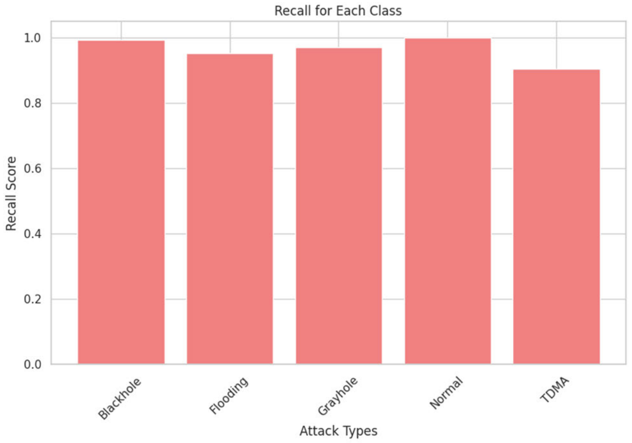
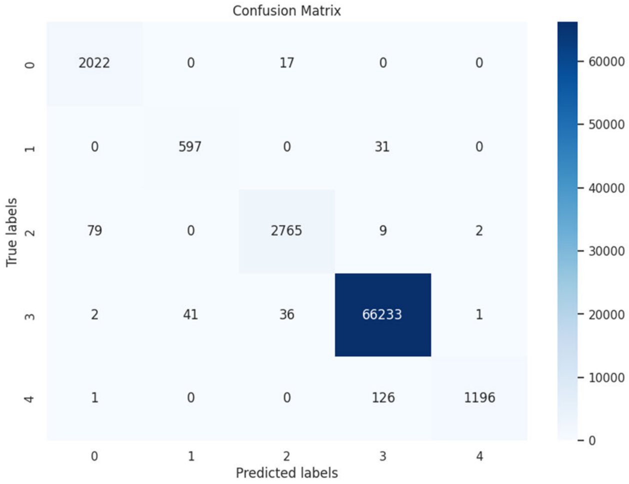


FIGURE 9. Recall score per class for MLP model.



	0	1	2	3	4
0	2022	0	17	0	0
1	0	597	0	31	0
2	79	0	2765	9	2
3	2	41	36	66233	1
4	1	0	0	126	1196

FIGURE 11. Confusion matrix for MLP model.

3) AUTOENCODER AND LR

The Autoencoder model is used to detect cybersecurity intrusions with essential Python libraries and modules, including Pandas and Sklearn. The results, as given in Table 3, indicate the performance metrics of a binary classification model distinguishing between normal and anomalous instances. The precision for both normal and anomalous instances is high at 0.95, and 0.88, respectively, for Autoencoder model. It signifies that when the model predicts an instance as normal and anomalous, it is correct 95% and 88% of the time, respectively. The LR model, on the other hand, performs exceptionally well in identifying the normal instances with a precision of 1.00; however, it struggles hard to identify the anomalous instances with a precision score of just 0.39. The recall for normal and anomalous instances is 0.88 and 0.95 for the Autoencoder model, indicating that the model captures 88% and 95% of the actual normal and anomalous instances, respectively. The recall values for the LR model are 84% and 99% for normal and anomalous instances, respectively, which are comparatively low in capturing the normal and slightly high in capturing the anomalous instances. The F1-score for the Autoencoder model stands at 0.91 and 0.92 for normal and anomalous instances, respectively. It signifies a good balance between the ability of the model to correctly identify normal and anomalous instances as compared with the LR model, with F1 scores of 0.91 and 0.55 for normal and anomalous instances, respectively. The overall accuracy of the Autoencoder model is 0.91, representing the proportion of correctly identified instances out of the total instances. It indicates that the overall performance of the model in classifying it as either anomalous or non-anomalous is 91% accurate as compared with the LR model, which has an accuracy of 85%. The sensitivity of the Autoencoder model is 88%, as compared with the LR model, which has a sensitivity value of 83%. It indicates that the Autoencoder model performs well in identifying positive instances out of all actual positive instances. On the other hand, the specificity of the Autoencoder model is 95% as compared with the LR model, which has a specificity value of 99%. It indicates that the LR model performs slightly well in identifying negative instances out of all actual negative instances. However, if we consider the sensitivity and specificity values of the Autoencoder model, it shows that the Autoencoder model has the great ability to correctly identify negative instances out of all actual negative instances and positive instances out of all actual positive instances.

The results given in Figure 12 suggest that the model shows better performance in identifying both normal and anomalous instances. While it maintains a high precision for normal instances, the recall for anomalous instances is comparatively higher, indicating that the model can identify a larger proportion of actual anomalous instances. The confusion matrix of Autoencoder model is given in Figure 13, which illustrates its performance in binary classification. The model accurately classified 6,072 instances as normal, with a small fraction of 866 normal instances as anomalies. This suggests that there are a very small number of false negatives, or cases where "normal" data is mistakenly classified as abnormal. The model identified 325 cases of anomalies as "normal" while properly identifying 6,575 incidents. This suggests that there are very few false negatives in the Anomaly class for the model. The higher percentage of true positives in both classes suggests that the model has the ability to correctly identify both "normal" and "anomaly" classes.

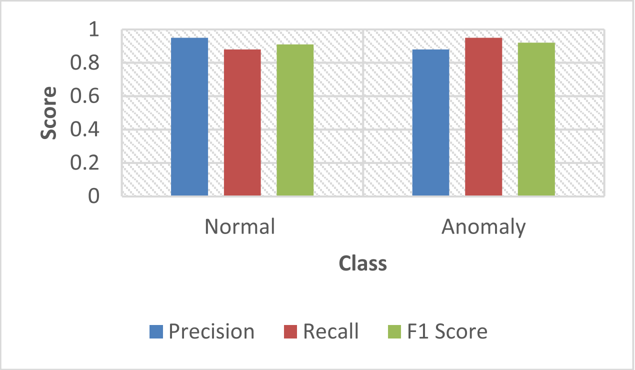


FIGURE 12. Classification metrics histogram of autoencoder Model.

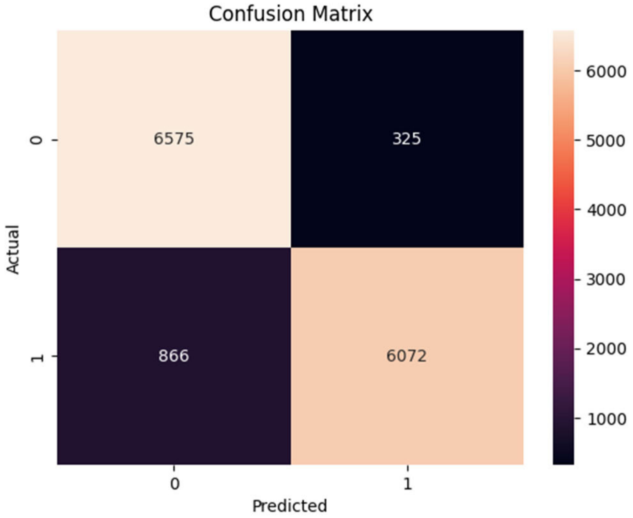


FIGURE 13. Confusion matrix of autoencoder model.

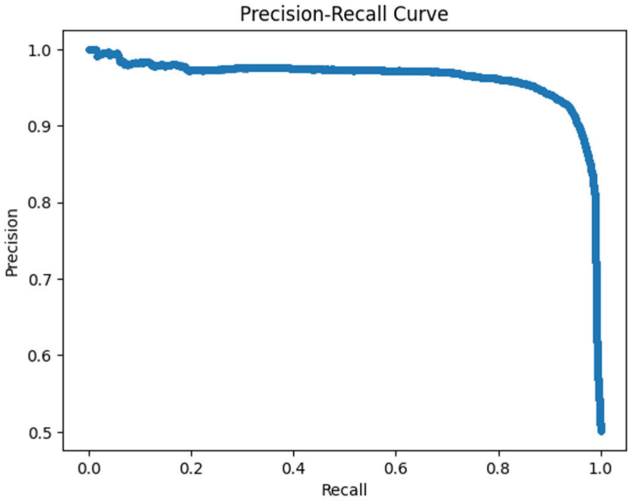


FIGURE 14. Autoencoder precision-recall curve.

TABLE 3. Accuracy, precision, recall and F1 score for Autoencoder and LR models

Class	Precision		Recall		F1-score		Accuracy		Sensitivity		Specificity	
	Autoencoder	LR	Autoencoder	LR	Autoencoder	LR	Autoencoder	LR	Autoencoder	LR	Autoencoder	LR
Normal	95%	100%	88%	84%	91%	91%	91%	85%	88%	83%	95%	99%
Anomaly	88%	39%	95%	99%	92%	55%						
Macro average	92%	69%	91%	91%	91%	73%						
Weighted average	92%	94%	91%	85%	91%	88%						

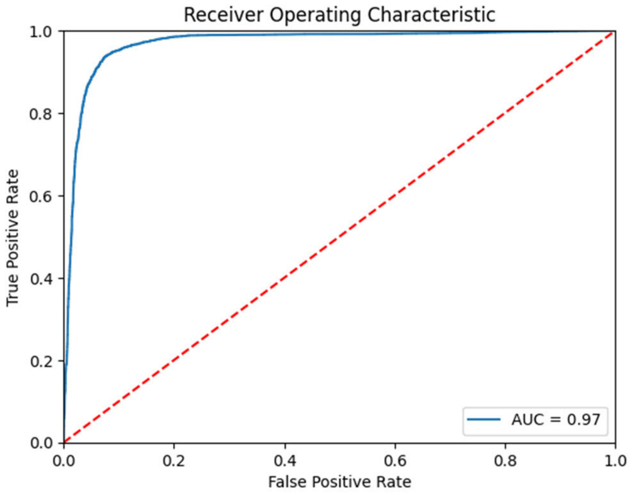


FIGURE 15. Autoencoder receiver operating characteristic curve.

The effectiveness of autoencoder model is assessed by precision and recall curve as given in Figure 14, a graph showing the precision and recall values for various threshold settings. Precision is plotted on the y-axis while recall is plotted on the x-axis. The average precision of model is represented by the area under the curve. The precision of the graph falls as the recall rises from its initial high precision and lower recall. With a high recall and low precision, the line terminates. The range of potential precision-recall curves for the autoencoder model is shown by the area under the curve. The model has a lower precision for high recall values and a high precision for low recall values.

Figure 15 shows a graph indicating the true positive rate and false positive rate for various threshold values for the Autoencoder model. The true positive rate is plotted on the y-axis, while the false positive rate is plotted on the x-axis. The area under the ROC curve is a measure of the overall performance of the model. A curve value of 0.97 indicates that the Autoencoder model performs well in finding the normal and anomaly classes.

E. DISCUSSIONS ON RESULTS

The proposed framework enhances the cybersecurity of WSNs in Industry 4.0 using a multi-criteria approach. It implements machine-learning and deep-learning algorithms for cybersecurity intrusion detection in WSNs of Industry 4.0 and provides prevention by assigning priorities to the threats based on the situation and nature of the attacks.

- In order to show the effectiveness of the proposed framework, we implemented three models, i.e., Decision Tree, MLP and Autoencoder, as proposed algorithms in the framework. For multidimensional classification and detection of cybersecurity intrusions, we implemented Decision Tree and MLP models. For binary classification and detection of cybersecurity intrusions in WSNs of Industry 4.0, we implemented Autoencoder model. Simulation results show that the Decision Tree model provides an accuracy of 99.48%, precision of 99.49%, recall of 99.48%, and F1 score of 99.49% in the detection and classification of cybersecurity intrusions. The MLP model provides an accuracy of 99.52%, precision of 99.5%, recall of 99.5%, and F1 score of 99.5% in the detection and classification of cybersecurity intrusions. The implementation of Autoencoder with binary classification yields an accuracy of 91%, precision of 92%, recall of 91%, and F1 score of 91%.
- To the best of our knowledge and as reflected in the literature review, no existing studies have implemented a multi-criteria approach for cybersecurity intrusion detection and classification in WSNs of Industry 4.0. Therefore, we implemented the benchmark models, i.e., Random Forest (RF) for multidimensional classification and Logistic Regression (LR) for binary classification. We compared the performance of the benchmark models with the models implemented in the proposed framework, revealing that the models in the proposed framework significantly outperformed the benchmark models.

The Decision Tree model exhibits a notable capacity to accurately classify instances as 'Normal,' which is a significant component in context with cybersecurity, at a rate of 99.48%. It does have trouble in classifying the difference between some types of attacks, such as ''Blackhole'' and ''Grayhole.'' Even though the accuracy is better, there are small deviation in the recall and precision rates between classes. This suggests that it is difficult to find certain types of intrusions accurately. The model is not very good at telling the difference between different types of intrusions, as shown by the small fluctuations in accuracy, recall, and F1-scores for classes like ''TDMA'' and ''Blackhole.'' The MLP model has a more equal performance across different classes, boasting an impressive accuracy of 99.52%. With respect to accurately identifying 'Normal' and 'Blackhole' situations, it excels in precision and recall rates. It does, however, have significant difficulty correctly distinguishing 'Normal' instances from other classes, as evidenced by some misclassifications in the confusion matrix. A reduced recall rate for ''TDMA'' cases indicates that the model missed some instances in its predictions. The Autoencoder model achieves an overall accuracy of 91% by using a binary classification approach. It obtains a good balance between precision and a recall trade-off for anomalous occurrences. Its strength is also its high precision for non-anomalous instances. Confusion matrix of the Autoencoder shows a minor percentage of misclassifications, particularly in differentiating between 'Normal' and 'Anomalies' cases.

Integration of the strengths of the Decision Tree, MLP, and Autoencoder models could lead to a more complete solution for building a smartly prioritized and strong predictive. The goal of the study is to find intelligent ways to prioritize things, and the Decision Tree is very good at finding ''Normal'' situations. The even success of MLP model across different classes gives us a full picture of intrusions. But it is important to talk about their own problems with telling the difference between different types of intrusion. A hybrid model that combines these strengths and makes up for their weaknesses by using feature engineering, ensemble methods, or even the Autoencoder binary classification strategy can be used to make the system much better at making predictions. The goal of this combination should be to create a strong multiclass classification system that can quickly tell the difference between normal and abnormal cases. It is very important for Industry 4.0-based WSNs to have an adaptive and proactive cybersecurity strategy that uses real-time threat intelligence and models that are constantly retrained and evaluated using new data. The proposed framework has the ability to be implemented in the real world through edge computing. Edge computing is a scalable distributed computing paradigm that provides computing and data storage services closer to the source of data generation. Since the proposed framework involves Industry 4.0 and IIoT, edge nodes implemented with WSNs in Industry 4.0 are the best choice for data processing, and cloud datacenters will be used for data storage and high-performance processing. We also made the assumption of constant model retraining in the proposed framework, which is based on the dynamic nature of cybersecurity threats within Industry 4.0 environments. There is also the challenge of obtaining relevant and diverse datasets. This can be solved by defining the strategies required to ensure the continuous supply of high-quality data for retraining. Retraining costs are another issue that needs to be resolved; however, the incorporation of edge computing aims to distribute the computational load efficiently, making constant retraining more practical and cost-effective. Although the use of AI-powered cybersecurity solutions raises a number of ethical considerations, privacy issues, and potential biases. However, we thoroughly scrutinized the models for any biases, implemented privacy-preserving techniques to safeguard sensitive information, and ensured transparency in the decision-making process. We established a robust and responsible framework with AI-enabled cybersecurity intrusion detection and prevention mechanisms in line with the ethical standards required in Industry 4.0 environments.

V. CONCLUSION

The proposed predictive framework is an intelligent and smart way to find and prevent cybersecurity attacks in WSNs based on Industry 4.0. The proposed framework combines important components including Industry 4.0, WSN, AI-driven detection, smart prioritization, and proactive safety measures. Using three different machine learning models i.e., Decision Tree, MLP and the Autoencoder, we make it possible to find and group different cybersecurity intrusions, which makes it easier for the network to quickly find and deal with these possible risks. Simulation results show that the Decision Tree model provides an accuracy of 99.48%, precision of 99.49%, recall of 99.48%, and F1 score of 99.49% in the detection and classification of cybersecurity intrusions. The MLP model provides an accuracy of 99.52%, precision of 99.5%, recall of 99.5%, and F1 score of 99.5% in the detection and classification of cybersecurity intrusions. The implementation of Autoencoder with binary classification yields an accuracy of 91%, precision of 92%, recall of 91%, and F1 score of 91%. The framework also includes an intelligent prioritization model that is key to quickly identifying and responding to high-risk intrusions by allocating resources in the best way to stop the worst attacks. Having a proactive preventive system in place makes the network more secure by quickly taking action to stop threats and making the whole thing more resistant to damage. The proposed framework is meant to make Industry 4.0-based WSNs safer by adding AI-based detection methods, ranking threats, and putting in place proactive defense strategies.

The proposed study is limited to the specific domain of WSNs in Industry 4.0 for detection and prevention cybersecurity intrusions. In order to safeguard the communication networks of other industries require further analysis. This study can be further enhanced through implementation of hybrid and customized AI models by considering the consequences of various types of attack. We aim to integrate Industry 4.0 standards including ISO/IEC 27001, NIST Cybersecurity Framework, and IEC 62443 with the proposed framework in future to make it more comprehensive and effective. We intend to strengthen the proposed framework in the future by incorporating dynamic threat intelligence tools. Dynamic threat intelligence tools are cybersecurity solutions designed to provide real-time, up-to-date information about potential and existing cybersecurity threats. These tools have the ability to continuously analyze and interpret data from various sources. However, these tools require more expert power, high-performance computational resources, and continuous training on updated datasets. By incorporating these tools, the system will be able to handle complex and new threats. By integrating advance behavioral analysis and anomaly detection methods, we intend to improve the framework performance and enable it to successfully manage new and complex cyberattacks.

ACKNOWLEDGMENT

This work was funded by the Deanship of Scientific Research at Jouf University through the Fast-track Research Funding Program.

REFERENCES

[1] R. S. Peres, X. Jia, J. Lee, K. Sun, A. W. Colombo, and J. Barata, ''Industrial artificial intelligence in Industry 4.0—Systematic review, challenges and outlook,'' *IEEE Access*, vol. 8, pp. 220121–220139, 2020, doi: 10.1109/ACCESS.2020.3042874.

[2] M. Ghobakhloo, ''Industry 4.0, digitization, and opportunities for sustainability,'' *J. Cleaner Prod.*, vol. 252, Apr. 2020, Art. no. 119869, doi: 10.1016/j.jclepro.2019.119869.

[3] M. van Geest, B. Tekinerdogan, and C. Catal, ''Design of a reference architecture for developing smart warehouses in Industry 4.0,'' *Comput. Ind.*, vol. 124, Jan. 2021, Art. no. 103343, doi: 10.1016/j.compind.2020.103343.

[4] P. Pop, B. Zarrin, M. Barzegaran, S. Schulte, S. Punnekkat, J. Ruh, and W. Steiner, ''The FORA fog computing platform for industrial IoT,'' *Inf. Syst.*, vol. 98, May 2021, Art. no. 101727, doi: 10.1016/j.is.2021.101727.

[5] H. Singh, ''Big data, Industry 4.0 and cyber-physical systems integration: A smart industry context,'' *Mater. Today: Proc.*, vol. 46, pp. 157–162, 2021, doi: 10.1016/j.matpr.2020.07.170.

[6] A. Corallo, M. Lazoi, and M. Lezzi, ''Cybersecurity in the context of Industry 4.0: A structured classification of critical assets and business impacts,'' *Comput. Ind.*, vol. 114, Jan. 2020, Art. no. 103165, doi: 10.1016/j.compind.2019.103165.

[7] J. Hajda, R. Jakuszewski, and S. Ogonowski, ''Security challenges in Industry 4.0 PLC systems,'' *Appl. Sci.*, vol. 11, no. 21, p. 9785, Oct. 2021, doi: 10.3390/app11219785.

[8] M. Humayun, N. Jhanjhi, B. Hamid, and G. Ahmed, ''Emerging smart logistics and transportation using IoT and blockchain,'' *IEEE Internet Things Mag.*, vol. 3, no. 2, pp. 58–62, Jun. 2020, doi: 10.1109/IOTM.0001.1900097.

[9] M. Humayun, M. S. Alsaqer, and N. Jhanjhi, ''Energy optimization for smart cities using IoT,'' *Appl. Artif. Intell.*, vol. 36, no. 1, Dec. 2022, Art. no. e2037255, doi: 10.1080/08839514.2022.2037255.

[10] A. Petrosyan. *Global Monthly Number of Cyber Attacks in Automotive Sector 2022–2023*. Accessed: Nov. 14, 2023. [Online]. Available: https://www.statista.com/statistics/1374790/biggest-automotive-cyber-attacks-worldwide/

[11] N. Verba, K.-M. Chao, J. Lewandowski, N. Shah, A. James, and F. Tian, ''Modeling Industry 4.0 based fog computing environments for application analysis and deployment,'' *Future Gener. Comput. Syst.*, vol. 91, pp. 48–60, Feb. 2019, doi: 10.1016/j.future.2018.08.043.

[12] I. Hussain, S. Tahir, M. Humayun, M. F. Almufareh, N. Z. Jhanjhi, and F. Qamar, ''Health monitoring system using Internet of Things (IoT) sensing for elderly people,'' in *Proc. 14th Int. Conf. Math., Actuarial Sci., Comput. Sci. Statist. (MACS)*, Nov. 2022, pp. 1–5, doi: 10.1109/MACS56771.2022.10023026.

[13] S. Kumar and R. R. Mallipeddi, ''Impact of cybersecurity on operations and supply chain management: Emerging trends and future research directions,'' *Prod. Oper. Manage.*, vol. 31, no. 12, pp. 4488–4500, Dec. 2022, doi: 10.1111/poms.13859.

[14] W. M. S. Yafooz, Z. B. A. Bakar, S. K. A. Fahad, and A. M. Mithun, ''Business intelligence through big data analytics, data mining and machine learning,'' in *Data Management, Analytics and Innovation*, vol. 1016. VIT Vellore, India: Springer, Jan. 2024, pp. 217–230, doi: 10.1007/978-981-13-9364-8_17.

[15] A. M. Riad, A. S. Salama, A. Abdelaziz, and M. Elhoseny, ''Intelligent systems based on loud computing for healthcare services: A survey,'' *Int. J. Comput. Intell. Stud.*, vol. 6, nos. 2–3, p. 157, 2017, doi: 10.1504/ijcistudies.2017.10010029.

[16] S. Zahoor and R. N. Mir, ''Resource management in pervasive Internet of Things: A survey,'' *J. King Saud Univ. Comput. Inf. Sci.*, vol. 33, no. 8, pp. 921–935, Oct. 2021, doi: 10.1016/j.jksuci.2018.08.014.

[17] B. Diène, J. J. P. C. Rodrigues, O. Diallo, E. H. M. Ndoye, and V. V. Korotaev, ''Data management techniques for Internet of Things,'' *Mech. Syst. Signal Process.*, vol. 138, Apr. 2020, Art. no. 106564, doi: 10.1016/j.ymssp.2019.106564.

[18] G. Fortino, A. Guerrieri, P. Pace, C. Savaglio, and G. Spezzano, ''IoT platforms and security: An analysis of the leading industrial/commercial solutions,'' *Sensors*, vol. 22, no. 6, p. 2196, Mar. 2022, doi: 10.3390/s22062196.

[19] I. H. Sarker, M. H. Furhad, and R. Nowrozy, ''AI-driven cybersecurity: An overview, security intelligence modeling and research directions,'' *Social Netw. Comput. Sci.*, vol. 2, no. 3, p. 173, May 2021, doi: 10.1007/s42979-021-00557-0.

[20] A. Corallo, M. Lazoi, M. Lezzi, and P. Pontrandolfo, ''Cybersecurity challenges for manufacturing Systems 4.0: Assessment of the business impact level,'' *IEEE Trans. Eng. Manag.*, vol. 70, no. 11, pp. 3745–3765, Nov. 2021, doi: 10.1109/TEM.2021.3084687.

[21] S. H. Zhu and P. Tang, ''A design and implementation of water surveillance system based on wireless sensor networks,'' *Appl. Mech. Mater.*, vols. 602–605, pp. 2305–2307, Aug. 2014, doi: 10.4028/www.scientific.net/amm.602-605.2305.

[22] T. Ali, M. Irfan, A. Shaf, A. S. Alwadie, A. Sajid, M. Awais, and M. Aamir, ''A secure communication in IoT enabled underwater and wireless sensor network for smart cities,'' *Sensors*, vol. 20, no. 15, p. 4309, Aug. 2020, doi: 10.3390/s20154309.

[23] M. Hanif, H. Ashraf, Z. Jalil, N. Z. Jhanjhi, M. Humayun, S. Saeed, and A. M. Almuhaideb, ''AI-based wormhole attack detection techniques in wireless sensor networks,'' *Electronics*, vol. 11, no. 15, p. 2324, Jul. 2022, doi: 10.3390/electronics11152324.

[24] D. Popescu, F. Stoican, L. Ichim, G. Stamatescu, and C. Dragana, ''Collaborative UAV-WSN system for data acquisition and processing in agriculture,'' in *Proc. 10th IEEE Int. Conf. Intell. Data Acquisition Adv. Comput. Systems: Technol. Appl. (IDAACS)*, vol. 1, Sep. 2019, pp. 519–524, doi: 10.1109/IDAACS.2019.8924424.

[25] K. Shaukat, S. Luo, V. Varadharajan, I. Hameed, S. Chen, D. Liu, and J. Li, ''Performance comparison and current challenges of using machine learning techniques in cybersecurity,'' *Energies*, vol. 13, no. 10, p. 2509, May 2020, doi: 10.3390/en13102509.

[26] L. F. Ilca, O. P. Lucian, and T. C. Balan, ''Enhancing cyber-resilience for small and medium-sized organizations with prescriptive malware analysis, detection and response,'' *Sensors*, vol. 23, no. 15, p. 6757, Jul. 2023, doi: 10.3390/s23156757.

[27] W. Zhang, D. Han, K.-C. Li, and F. I. Massetto, ''Wireless sensor network intrusion detection system based on MK-ELM,'' *Soft Comput.*, vol. 24, no. 16, pp. 12361–12374, Aug. 2020, doi: 10.1007/s00500-020-04678-1.

[28] U. AlHaddad, A. Basuhail, M. Khemakhem, F. E. Eassa, and K. Jambi, ''Ensemble model based on hybrid deep learning for intrusion detection in smart grid networks,'' *Sensors*, vol. 23, no. 17, p. 7464, Aug. 2023, doi: 10.3390/s23177464.

[29] M. H. Ali, M. M. Jaber, S. K. Abd, A. Rehman, M. J. Awan, R. Damaševičius, and S. A. Bahaj, ''Threat analysis and distributed denial of service (DDoS) attack recognition in the Internet of Things (IoT),'' *Electronics*, vol. 11, no. 3, p. 494, Feb. 2022, doi: 10.3390/electronics11030494.

[30] S.-F. Lokman, A. T. Othman, and M.-H. Abu-Bakar, ''Intrusion detection system for automotive controller area network (CAN) bus system: A review,'' *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, p. 184, Dec. 2019, doi: 10.1186/s13638-019-1484-3.

[31] M. Rabbat and R. Nowak, "Distributed optimization in sensor networks," in *Proc. 3rd Int. Symp. Inf. Process. sensor Netw.*, Apr. 2004, pp. 20–27, doi: 10.1145/984622.984626.

[32] W. Steingartner, D. Galinec, and A. Kozina, "Threat defense: Cyber deception approach and education for resilience in hybrid threats model," *Symmetry*, vol. 13, no. 4, p. 597, Apr. 2021, doi: 10.3390/sym13040597.

[33] M. A. Al-Shareeda, S. Manickam, S. A. Laghari, and A. Jaisan, "Replay-attack detection and prevention mechanism in Industry 4.0 landscape for secure SECS/GEM communications," *Sustainability*, vol. 14, no. 23, p. 15900, Nov. 2022, doi: 10.3390/su142315900.

[34] N. S. Safa, C. Maple, S. Furnell, M. A. Azad, C. Perera, M. Dabbagh, and M. Sookhak, "Deterrence and prevention-based model to mitigate information security insider threats in organisations," *Future Gener. Comput. Syst.*, vol. 97, pp. 587–597, Aug. 2019, doi: 10.1016/j.future.2019.03.024.

[35] R. Abu Bakar, X. Huang, M. S. Javed, S. Hussain, and M. F. Majeed, "An intelligent agent-based detection system for DDoS attacks using automatic feature extraction and selection," *Sensors*, vol. 23, no. 6, p. 3333, Mar. 2023, doi: 10.3390/s23063333.

[36] J. Kim, M. Park, H. Kim, S. Cho, and P. Kang, "Insider threat detection based on user behavior modeling and anomaly detection algorithms," *Appl. Sci.*, vol. 9, no. 19, p. 4018, Sep. 2019, doi: 10.3390/app9194018.

[37] N. Peppes, E. Daskalakis, T. Alexakis, E. Adamopoulou, and K. Demestichas, "Performance of machine learning-based multi-model voting ensemble methods for network threat detection in Agriculture 4.0," *Sensors*, vol. 21, no. 22, p. 7475, Nov. 2021, doi: 10.3390/s21227475.

[38] L. S. Vailshery. *Industry 4.0 Technologies to Have Greatest Impact on Organizations Worldwide 2020*. Accessed: Oct. 30, 2023. [Online]. Available: https://www.statista.com/statistics/1200006/industry-40-technology-greatest-impact-organizations-worldwide/

[39] F. A. Saputra, M. U. H. A. Rasyid, and B. A. Abiantoro, "Prototype of early fire detection system for home monitoring based on wireless sensor network," in *Proc. Int. Electron. Symp. Eng. Technol. Appl. (IES-ETA)*, Sep. 2017, pp. 39–44, doi: 10.1109/ELECSYM.2017.8240373.

[40] K. Akkaya, M. Younis, and W. Youssef, "Positioning of base stations in wireless sensor networks," *IEEE Commun. Mag.*, vol. 45, no. 4, pp. 96–102, Apr. 2007, doi: 10.1109/MCOM.2007.343618.

[41] L. Yunhong and Q. Meini, "The design of building fire monitoring system based on ZigBee-WiFi networks," in *Proc. 8th Int. Conf. Measuring Technol. Mechatronics Autom. (ICMTMA)*, Mar. 2016, pp. 733–735, doi: 10.1109/ICMTMA.2016.180.

[42] A. Alkhatib, "Sub-network coverage method as an efficient method of wireless sensor networks for forest fire detection," in *Proc. ACM Int. Conf.*, vols. 22–23, Mar. 2016, pp. 1–7, doi: 10.1145/2896387.2896450.

[43] H. Dabiri, V. Farhangi, M. J. Moradi, M. Zadehmohamad, and M. Karakouzian, "Applications of decision tree and random forest as tree-based machine learning techniques for analyzing the ultimate strain of spliced and non-spliced reinforcement bars," *Appl. Sci.*, vol. 12, no. 10, p. 4851, May 2022, doi: 10.3390/app12104851.

[44] G. S. Fischer, R. D. R. Righi, G. D. O. Ramos, C. A. D. Costa, and J. J. P. C. Rodrigues, "ElHealth: Using Internet of Things and data prediction for elastic management of human resources in smart hospitals," *Eng. Appl. Artif. Intell.*, vol. 87, Jan. 2020, Art. no. 103285, doi: 10.1016/j.engappai.2019.103285.

[45] N. Mozaffaree Pour and T. Oja, "Prediction power of logistic regression (LR) and multi-layer perceptron (MLP) models in exploring driving forces of urban expansion to be sustainable in Estonia," *Sustainability*, vol. 14, no. 1, p. 160, Dec. 2021, doi: 10.3390/su14010160.

[46] A. Kumari, R. K. Patel, U. C. Sukharamwala, S. Tanwar, M. S. Raboaca, A. Saad, and A. Tolba, "AI-empowered attack detection and prevention scheme for smart grid system," *Mathematics*, vol. 10, no. 16, p. 2852, Aug. 2022, doi: 10.3390/math10162852.

[47] Y. Song, S. Hyun, and Y.-G. Cheong, "Analysis of autoencoders for network intrusion detection," *Sensors*, vol. 21, no. 13, p. 4294, Jun. 2021, doi: 10.3390/s21134294.

[48] I. Almomani, B. Al-Kasasbeh, and M. Al-Akhras, "WSN-DS: A dataset for intrusion detection systems in wireless sensor networks," *J. Sensors*, vol. 2016, pp. 1–16, Jan. 2016, doi: 10.1155/2016/4731953.

[49] N. Farnaaz and M. A. Jabbar, "Random forest modeling for network intrusion detection system," *Proc. Comput. Sci.*, vol. 89, pp. 213–217, Jan. 2016, doi: 10.1016/j.procs.2016.06.047.

[50] T. G. Nick and K. M. Campbell, "Logistic regression," in *Topics in Biostatistics* (Methods in Molecular Biology), vol. 404. Springer, 2007, pp. 273–301, doi: 10.1007/978-1-59745-530-5_14.

FATIMA AL-QUAYED is currently an Assistant Professor with the College of Computer and Information Sciences, Jouf University, Saudi Arabia. She has multiple publications in WoS/ISI/SCI/Scopus. She has vast experience in academic qualifications. Her research interests include cyber security, wireless sensor networks (WSN), the Internet of Things (IoT), and knowledge management.

ZULFIQAR AHMAD received the M.Sc. degree (Hons.) in computer science from Hazara University, Mansehra, Pakistan, in 2012, the M.S. degree in computer science from COMSATS University Islamabad, Abbottabad, Pakistan, in 2016, and the Ph.D. degree in computer science from the Department of Computer Science and Information Technology, Hazara University, in 2022. He is the author of several publications in the fields of fog computing, cloud computing, artificial intelligence, high-performance computing, and scientific workflow scheduling and management. His current research interests include scientific workflow management in cloud computing, the Internet of Things, fog computing, edge computing, artificial intelligence, cybersecurity, and wireless sensor networks (WSNs).

MAMOONA HUMAYUN is currently an Assistant Professor with the College of Computer and Information Sciences, Jouf University, Saudi Arabia. She has highly indexed publications in WoS/ISI/SCI/Scopus and her collective research impact factor is more than 200 plus points. Her Google Scholar H-index is 28 and I-10 Index is close to 78, with more than 150 publications on her credit. She has several international patents on her account, including U.K. and Japanese. She has edited/authored over five research books published by World-Class Publishers. She has excellent experience in supervising and co-supervising postgraduate students and more than 13 postgraduate scholars graduated under her supervision. She has completed more than 15 funded research grants successfully. She has vast experience in academic qualifications, including ABET and NCAAA. Her research interests include cyber security, wireless sensor networks (WSN), the Internet of Things (IoT), requirement engineering, global software development, and knowledge management. She has served as a keynote/invited speaker for many international conferences and workshops. She serves as a reviewer for several reputable journals.