IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 29 August 2024, accepted 30 September 2024, date of publication 7 October 2024, date of current version 25 October 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3474861

RESEARCH ARTICLE

Exploring the Potential Network Vulnerabilities in the Smart Manufacturing Process of Industry 5.0 via the Use of Machine Learning Methods

VADYM SHKARUPYLO[1,2], JAMIL ABEDALRAHIM JAMIL ALSAYAYDEH[3], (Member, IEEE), MOHD FAIZAL BIN YUSOF[4], ANDRII OLIINYK[5], VOLODYMYR ARTEMCHUK[6,7,8,9], AND SAFARUDIN GAZALI HERAWAN[10]

[1]Department of Computer Systems, Networks and Cybersecurity, National University of Life and Environmental Sciences of Ukraine, 03041 Kyiv, Ukraine
[2]Department of Mathematical and Computer Modelling, G. E. Pukhov Institute for Modelling in Energy Engineering, National Academy of Sciences of Ukraine, 03164 Kyiv, Ukraine
[3]Department of Engineering Technology, Fakulti Teknologi & Kejuruteraan Elektronik & Komputer (FTKEK), Universiti Teknikal Malaysia Melaka (UTeM), Melaka 76100, Malaysia
[4]Research Section, Faculty of Resilience, Rabdan Academy, Abu Dhabi, United Arab Emirates
[5]Department of Software Tools, Faculty of Computer Science and Technology, Zaporizhzhia Polytechnic National University, 69063 Zaporizhzhia, Ukraine
[6]Department of Mathematical and Econometric Modelling, G. E. Pukhov Institute for Modelling in Energy Engineering, NAS of Ukraine, 03164 Kyiv, Ukraine
[7]Department of Environmental Protection Technologies and Radiation Safety, Center for Information-Analytical and Technical Support of Nuclear Power Facilities Monitoring, NAS of Ukraine, 03164 Kyiv, Ukraine
[8]Department of Information Systems in Economics, Kyiv National Economic University named after Vadym Hetman, 03057 Kyiv, Ukraine
[9]Department of Computerized Control System, National Aviation University, 03058 Kyiv, Ukraine
[10]Industrial Engineering Department, Faculty of Engineering, Bina Nusantara University, Jakarta 11480, Indonesia

Corresponding author: Jamil Abedalrahim Jamil Alsayaydeh (jamil@utem.edu.my)

This work was supported in part by the Development of Specialized Computer Technologies for Modeling and Processing of Operational Information in Energy Problems under Grant 0120U102683, and in part by the Development of a Hardware-Software Complex and Methodology for Rapid Detection of Damages in Heating and Water Supply Systems Considering Their Wear and Tear and Military Impacts funded by the National Research Foundation of Ukraine (NRFU) under Grant 2023.04/0022.

ABSTRACT The Industry 5.0 revolution has launched a new age of intelligent manufacturing equipment, which is crucial to our society and economy. Industry 5.0 aims to enhance human potential by integrating cutting-edge IT technology, Artificial Intelligence (AI), the Internet of Things (IoT), robots, and augmented reality into everyday living, especially in smart industrial settings, to maximize human potential. Smart Production Processes (SP2), accessible standards, and resource sharing on the web have increased network vulnerabilities as enterprises adopt them. These vulnerabilities primarily target industry IoT systems, threatening private data. Existing system issues include network vulnerabilities, production capability, and data management. Traditional applications struggle with industrial data's bulk and complexity, causing data analysis, security, and privacy challenges. Modern industrial systems use AI and ML for big data analysis and data processing to overcome these issues. Optimizing human-machine synergy minimizes costs and equipment maintenance and boosts efficiency. An ML-assisted Smart Production Process (ML-SP2) and Intrusion Detection System (ML-IDS) have been proposed to assess and address Industrial 5.0 concerns. Data capture, predictive maintenance and optimization, transparent decision-making, and proactive maintenance are integrated into the manufacturing process with the ML-SP2. The ML-IDS detects network vulnerabilities using ensemble methods and manages distribution efficiently. The ML-IDS uses the variety of ML techniques such as random forest, decision tree and support vector machines that identifies the intruder with maximum prediction accuracy. In addition, the intruder activities are observed with the help of the convolution networks that improve the overall intruder activities recognition. The smart industrial production phase's effectiveness

The associate editor coordinating the review of this manuscript and approving it for publication was Wei-Yen Hsu.

is measured by the RPL (0.9), maintenance cost (2500$), production performance (1.0), VDR (97.6%), FAR, accuracy (99%), precision (94%)-recall (96%), and f1-measure (98%) compared to the existing methods such as ML-RF-FLID, NRDD-DBSCAN, and OC-SVM.

INDEX TERMS Industry 5.0, network vulnerabilities, smart production process, machine learning, Internet of Things, predictive maintenance.

I. INTRODUCTION

The notion ''Industry 5.0'' describes employees seeking the usage of robots and sophisticated machines to complete tasks and improve smart production. Industry 5.0 aims to produce smart production solutions with effective network safety measures and user-friendly resources by combining the innovation of expert systems with accurate, intelligent, and efficient machinery. Previous Industry 4.0 emphasizes automated equipment efficiency, while more recent concept of Industry 5.0 puts an accent on the effectiveness of optimizing manual input for cutting-edge technological innovation. Based on the IoT infrastructure, data can now be analyzed in real-time across the industry, uploaded in the cloud for improved analysis, and used to upgrade ML models for SP2 and network vulnerabilities. A focus is on generating SP2 combining the AI with human expertise to make them more flexible and collaborative. Numerous worries about cyberattacks and data protection have arisen due to the proliferation of smart gadgets and communication through the IoT. ML algorithms are highly effective in dealing with threats because of their exceptional anomaly recognition and categorization abilities.

Industry 5.0 is intended for SP2 enhancement in addition to network efficiency in the work area and communication with consumers from a remote, emphasizing mostly the coordination between geographically distributed machines and human beings [1]. Identifying routine, consistent duties for robots and machines while giving humans more complex, thinking-intensive activities will help Industry 5.0 improve performance [2]. The dire necessity to boost smart production alienates employees, causing significant difficulties for the global financial system [3]. As people become more interconnected, Industry 5.0 focuses on combining their creativity and intellect with cognitive computing capabilities using intelligent techniques in a coordinated manner [4]. Artificial Intelligence (AI) techniques have developed due to the widespread production and accumulated the IoT sensed information and online platforms, the internet, and smart gadgets of industries. The information is then used for training the ML model and data analysis. It relies on a centralized server for training purposes, which causes data leakage in industrial network scenarios [5].

The effects of Industry 5.0 on social and economic aspects, concluding how moving to ''mass automation'' at the smart production level, handled through the creation of human resources in conjunction with the AI, would enable successful production re-emergence [6]. The ML algorithms and smart, intelligent robots are very useful for remotely managing production systems in the industrial sector [7]. Human-cyber-physical systems bring cognition intelligent-based CPSs built by sensors with human intelligence and learned ML systems [8]. Network vulnerabilities include RFID impersonation, dynamic routing attacks, and blackmail attacks. These attacks can be stopped by using a secure hash function and authentication though [9]. With the use of wireless sensor networks for goal-controlling, the facilitation of the incorporation of massive volumes of data, and the supervision and coordinated supervision of actual operations, the cyber-physical production systems (CPPSs) are essential in improving the sustainable production IoT and smart manufacturing [10], [11].

Smart-tailored goods are manufactured using the ML-SP2 in knowledge-intensive industrial unsupervised circumstances structured by the CPS-based production [12]. Smart production and intelligence aid in reducing network traffic, facilitating transactions, and protecting privacy, allowing businesses to employ digital resources to share information about industrial sectors [13]. A unique ML technique called Dual Isolation Forest (DIF), which trains multiple IFs, was used in the semi-supervised approach. Since the test dataset only included information about cyberattacks of industrial scenarios [14]. The Intrusion Detection System (IDS) aids in spotting network vulnerabilities and takes the appropriate safety precautions by using migration learning to guarantee the IoT's secure and dependable functioning in Industry 5.0, which still shows the difficulty in limited communication capability [15]. Even though, several learning approaches are utilized for managing the intruder activities, the security and privacy should be managed while analyzing the industrial data. The system should manage the vulnerable activities to ensure the network safety. The traditional approaches are fails to process the complex industrial patterns which increases the false acceptance rate.

The motivation of this work involves the integration of the ML-SP2 and big data-driven continuous production planning automatically to maintain the vulnerabilities in the network safely. During this process, several machine learning algorithms such as random forest, decision tree, support vector machine and convolution networks are incorporated to identify the intruder activities. This process ensures the trust and minimize the security related issues.

The paper's primary contribution can be summarized as follows:

- Designing ML-SP2 for improving industrial production efficiency using machine lifespan, performance,

TABLE 1. List of abbreviation.

CPPS	Cyber-physical production systems
IDS	Intrusion Detection System
ML	Machine Learning
DIF	Dual Isolation Forest
CAD	Computer-Aided Design
ARF	Adaptive Random Forest
AI	Artificial Intelligence
VDR	Vulnerability Detection Rate
FAR	False Alarm Rate
ROC	Receiver Operating Characteristics
AUC	Area Under Curve
LR	Learning Rate
ML-RF-FLID	multi-view ensemble called Random Forest (RF) based Federated Learning Intrusion Detection Method
NRDD-DBSCAN	Noise Resilient Distributed Datasets (RDDs)-Density-Based Spatial Clustering (DBSCAN)
OC-SVM	One Class-Support Vector Machine

and maintenance cost reduction using predictive maintenance and machine optimization model to reduce maintenance cost.

- Implement ensembling techniques such as Auto Encoder (AE), Decision Tree (DT), and Random Forest (RF) in ML-IDS to detect invasions and vulnerability patterns in the industry network to get continuous SP2.
- For efficient analysis, evaluate necessary metrics or conditions like RPL, maintenance cost, performance, accuracy, precision-recall, f1-measure, VDR, and FAR. Then the efficiency of the system is compared with existing methodologies such as ML-RF-FLID, NRDD-DBSCAN, and OC-SVM.

The remainder of the work is arranged as follows: Section II discusses the existing algorithms relevant to smart production processes (SP2) and network vulnerabilities of Industry 5.0. Section III covers the implementation analysis of the ML-SP2-IDS algorithm based on AI and cloud technologies in the current and emerging smart Industry 5.0. Section IV identifies the baseline experimental analysis for network safety issues and necessary solutions for smart production requirements in Industry 5.0. Section V discusses the suggested algorithms work and future directions of this work and is finally concluded.

II. RELATED WORK

The most recent analysis of ML-based cutting-edge algorithms related to smart Industry 5.0 is examined in this section, and their key points of interest are discussed.

Sharma et al. established a smart method for collaborative design and clothing manufacturing by combining several data-driven smart services, an Illustration of interactive 3D clothing measurements, and a depth of knowledge for development. The associated computations are carried out for each consumer profile using ML algorithms such as Adaptive Random Forest [ARF] algorithms and mining rules for associating the results. By immediately integrating the consumer's perspective and expert designers' expertise into the outfit's Computer-Aided Design (CAD) platform, the complexity of the production process can be significantly decreased and improved production efficiency. However the effective and optimization techniques are requires to understand the rules and relationship [16], [17]. Caiazzo et al. presented a new design for smart manufacturing system monitoring that makes use of the internet of things (IoT) and the cloud. In order to detect and categorize abnormalities, the architecture employs artificial intelligence. Using control charts, autoencoders, LSTM, and FIS, the five-layer platform can identify flaws and their causalities, according to Industry

5.0 principles. Operators and company administrators are provided with real-time status updates and risk levels via the design, which has been experimentally proven on a solar thermal high-vacuum flat panel. In order to avoid product waste and identify causalities in complicated systems, this method is essential [18], [19]. During the abnormality identification, the system fails to explore the complex patterns. Ding et al. proposed that AI-powered manufacturing improves several areas, generating confined network production from standards to supplies. In particular, industrial AI, including contextual information, greatly improves production tracking. Deep Neural networks, adversarial training, and Transfer Learning (TL) are widely employed to assist in the inspection and predictive maintenance of manufacturing units using DNN-TL. Production surveillance metrics include defect identification, remaining usable life forecasts, and industry acceptance testing [20], [21]. The deviation between the outputs requires the fine-tune procedures which takes the computation complexity. Ouda et al. suggested a framework for predicting machine faults and optimizing predictive/corrective maintenance schedules based on sensor data. ML models are trained using previous data to forecast five-day failure risks. The output of the ML models is supplied into a model of optimizing, which then proposes an optimum maintenance standard. It gives the highest prediction accuracy and enhances system reliability, keeping costs down. However, the ML model limited for data availability [22], [23]. Menon et al. presented Novelty detection finds novel information in text datasets. To increase interpretability, spectral graph-based approaches have been researched, however there is a literature gap on merging them with visualization tools. A innovative strategy that combines spectral graph-based approaches with visualization tools to discover novel documents in text data gathering yields interpretable results. This technique may be used in many text data novelty detection areas [24], [25]. The introduced visualization process fails to maintain system scalability and flexibility.

Frankó et al. offered a complete review of ML techniques used in Industrial IoT and Smart Manufacturing for diverse objectives. Privacy and secured asset location, proactive management, and quality management are covered. Resource placement is a subset of smart manufacturing in which machine learning has been actively used. Maintenance applications include defect detection, monitoring, predictive, and production optimization applications [26], [27]. Bagaa et al. implemented an ML-based security architecture that adapts Software Defined Networking(SDN)/Network Function Virtualization (NFV) standards with One Class-Support Vector Machine(OC-SVM) classification as a monitoring agent with an AI-based reaction agent that analyses network patterns of anomaly-based intrusion detection for Internet of Things (IoT) devices. SVM was used in the framework to incorporate an IDS for anomaly detection in IoT sensor data, and this technique increased detection accuracy to over 98% [28]. Samara et al. provided a survey covering fault detection implementation in IoT networks, highlighting anomaly detection samples in industrial IoT.ML systems can provide a broader level of anomaly detection by employing statistical, hybrid, unsupervised, and supervised approaches to discover Exceptions or irregular activity in the network [29], [30]. Ghallab et al. suggested Resilient Distributed Datasets (RDDs) are used in addition to the Density-Based Spatial Clustering of Application with Noise(DBSCAN) method and N-dimensional RDD-DBSCAN to identify outliers that compromise the network performance of IoT technologies in industrial big data analytics. However, it handles only linear data reduction techniques unsupported by non-linear to detect outliers. The ground truth labels are necessary for the V-measure, Homogeneity, Completeness, and Adjusted Mutual Information [31]. Rodríguez et al. presented Industrial Internet of Things (IIoT) enables real-time, secure, and autonomous production in Industry 4.0. However, diverse technologies needed for IIoT systems provide incomplete, unstructured, redundant, and loud data. Anomaly detection systems identify bad data and protect IIoT systems. These systems are difficult to implement owing to method selection, information processing, monitoring devices, and algorithm execution. Automatic anomaly classification in IIoT is an open research field, with minimal application context information employed for anomaly detection, according to a 99-article state-of-the-art review [32]. Attota et al. developed a multi-view ensemble called Random Forest (RF) based Federated Learning Intrusion Detection Method (MV-RF-FLID) perspectives decentralized storage of IoT application data to identify, categorize, and protect from threats. The top-k selection-based gradient compression scheme improves performance efficiency with minimum communication overhead [33], [34], [35].

Mutaz Ryalat et al. [36] suggested Cyber-Physical Systems and Internet of Things for Industry 4.0. An innovative smart factory architecture for Industry 4.0 is outlined in this study, together with the essential industrial, computer, information, and communication technologies that make up a smart factory. It explains how to build an intelligent production system by integrating a smart factory's various elements (pillars). A smart manufacturing case study, including a drilling process, is used to show the simplified smart factory model, and the practicality of the suggested technique is proved and confirmed by tests.

Wenhao Yan et al. [37] proposed the Real-Time Fault Diagnosis Methods (RTFD) for Industrial Smart Manufacturing. This study thoroughly examines recent RTFD developments in the machine condition monitoring and industrial process monitoring fields. The RTFD procedure is described in depth, beginning with the first data collection stage. Techniques based on "end-to-end" neural networks, techniques based on qualitative knowledge reasoning from a fresh viewpoint, and methods based on independent feature extraction make up the current RTFD approaches. In addition, the study

provides a reference for scholars in this area by discussing the difficulties and possible future developments of RTFD.

Salam et al. [38] presented Industry 5.0's use of IoT and cyber-physical systems in production presents a cybersecurity risk. This study suggests detecting web-based assaults utilizing CNNs, RNNs, and transformer models using deep learning. The transformer-based approach surpasses classic machine learning techniques in accuracy, precision, and recall, proving deep learning can identify intrusions in Industry 5.0 contexts.

Adel et al. [39] displayed with an emphasis on technologies like cyber-physical systems, renewable energy, cyber-forensics, machine learning, deep learning, fog computing, unmanned aerial vehicles, and cyber-physical systems, this article investigates the ways in which Industry 5.0 might influence the development of smart cities in the future. Data management, intelligent automation [40], human-machine cooperation, and improved cybersecurity are some of the areas it focuses on. Additionally, the article assesses the ways in which Industry 5.0 technologies could improve upon current frameworks that are influencing smart city applications. Its stated goal is to help smart cities and their environs overcome the technical obstacles they encounter.

The summary of the literature review work includes various related algorithms applied for ML-based smart industrial production and network vulnerabilities issues. The algorithms ARF and GA-ANN using the dynamic concept, DNN-TL, and DAE-DBC are reviewed for SP2. The network vulnerabilities issues are resolved with earlier work like SDN/NFV-based OC-SVM, NRDD-DBSCAN, and MV-RF-FLID techniques. Still, there is a lack of solutions for smart production and security issues. Compared to these methods, the proposed approach uses the Machine learning approaches that identifies the intruder activities in the network with minimum false acceptance rate and high detection accuracy. The method uses the ML that ensures the adaptability, quality and efficiency it also mitigate the security threats in the production process. Thus the system ensures the robust framework in the industry 5.0.

III. PROPOSED SOLUTION

Machines must be effective at doing intellectual labor and even creating innovations for this type of manufacturing to free humans from physically demanding tasks. There is a critical need to understand if Cyber-Physical Production Systems (CPPSs) are appropriate for flexibility in configuring smart industries, given the emerging evidence of ML-assisted SP2(ML-SP2). This study uses the machine learning algorithm to cyber physical production system because it can handle the imbalance data, identifying the complex vulnerable patterns and provide the robust solutions. Especially this study uses the Convolution Neural Networks (CNN) that identifies the complex patterns from large volume of manufacturing data. The extracted patterns retrieve the relationship between the data that minimize the vulnerability issues effectively. The traditional methods like autoencoder approach utilized in the unsupervised learning process to detect the anomaly and unusual pattern which represents that the network vulnerabilities. The other methods such as random forest, decision tree approaches also able to analyze the numerical and categorical data for making decisions regarding the intrusions. However, the neural approach like Convolution Neural Networks(CNN) provides the effective results while analyzing the large volume and complex data in the industry 5.0 platform. This paper uses the Convolutional Neural Networks (CNNs) in the context of the ML-assisted Smart Production Process (ML-SP2) and the Intrusion Detection System (ML-IDS). Even though, the CNN network effectively utilized in the image processing applications, it able to process the sequential data and identify the complex network patterns to identify the network vulnerabilities. The integration of the ML-SP2 with CNN that focuses on the deep learning concept to improve the overall smart production environment. The CNN approach uses the different layers that provides the transparency while making decision-making in the industry 5.0. The effective data analysis process ensures robustness, security and provides the effective SP2. The interoperability of IoT and CPP surveillance standards was considered in this study to show that it can determine how operations advance restoring a device to its initial configuration with CPPSs. Whenever SP2 relates to industrial automation, the goal of evaluating the obtained information for improved overall performance by boosting output while minimizing waste and consuming less energy. Industry 5.0 gives the welfare of the workforce at the core of SP2 to overcome the human difficulties of Industry 4.0. Fully automated SP2 to change conditions and needs in the production system, the delivery model, and consumer demands are involved in the smart production process in Industry 5.0. Fully automated, intelligent machines will take over the arduous task of managing every step of the production process in Industry 5.0. The SP2 ends up being tailored to fit the needs of client-focused industries. Each good from all those industries will be personalized following the clients' preferences, increasing consumer satisfaction and ease of use, decreasing cancellations, and increasing the value of produced goods. Robotic WSN and big data analytics in Industrial 5.0 for cyber-physical monitoring networks will be integrated into smart networked industries to improve them. IoT security issues have been incorporated into these systems because of the rapid rise in IoT device usage and deployment across all industries.

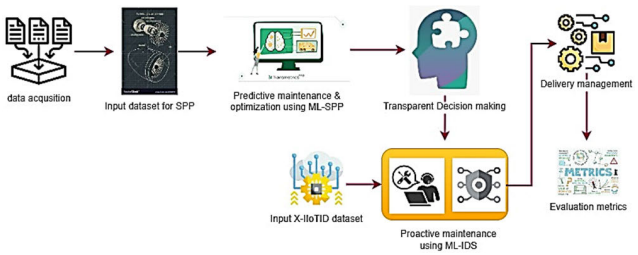


FIGURE 1. Proposed ML-SP2-IDS model for effective Industrial 5.0.

From Figure 1, the implementation of Industry 5.0 for SP2 is given in a detailed manner. Two datasets are used in this manuscript for SP2 to maintain a smooth production process and IDS to identify network vulnerabilities. Initially, data is gathered from necessary devices for pre-processing, followed by predictive maintenance and optimization control for identifying the machine status and working state of the equipment in the production phase, thereby reducing the overall cost for repair maintenance and improving the production quality. Implementing ML-SP2 for training the model to calculate the RPL of a machine and forward it to make decisions clearly to the industrial process. Then proactive maintenance is used with the help of ML-IDS by ensembling three ML algorithms, and results are optimized to classify the malicious and regular activity. The final production phase involves intelligent delivery management using supply chain logistics. Various key performance indicators are analyzed to enhance the smart production phase performance efficiency.

A. DATA ACQUISITION FOR ANALYZING SMART PRODUCTION PROCESSES (SP2)

Data is generated from several intelligent sensors to monitor important production equipment due to Instrumentation, data management, networking transmission, and other technological innovations that are rapidly evolving. With the human touch information, maintaining SP2 is the objective of Industry 5.0. By storing data documents on a cloud with strict access controls and distributing the design flow of produced things. Initially, information has been collected from machinery-sensed data where the devices implemented with smart IoT sensors enabled for production phases. Both user and product data are collected from IoT sensors implemented in machines that collect machinery arm handles data, acceleration sensor data, Data from vibration sensors, video images from greater cameras, and moisture and temperature data at the product level. Online machine and environment monitoring is done using the sensor data. Real-time data enabling collaboration mingled with the big data analytics storage access for smart production capacities. The system production paradigm enables the integration of production capabilities with services to deliver appropriate solutions to customers. By incorporating service components into the production process, the industrial sector hopes to increase production efficiency, financial returns, and share price through technology innovation. Cloud is a networked system used for production resources and a distributed system for machinery analysis. The virtual setting, known as the ''industrial cloud storage,'' offers a supportive atmosphere for smart production in industrial applications like IoT monitoring tools produced by service providers and used for online access. With the development of ML techniques, they can locate and remove quality defects from the production process by using slightly elevated webcams mounted on the robotic arms to track the movement of things in real-time.

The smart production method in Industry 5.0 offers the technological advantage of allowing a small team of highly skilled people to quickly generate high-quality smart products in the market. This paradigm offers a contemporary framework for producing intelligent products through smart production techniques based on ML algorithms. Initially, it aspires to produce smart items more rapidly and with higher quality using an SP2. Second, the cost of the product and how much energy it uses will determine how popular and in demand it is in the current market. According to this viewpoint, demand is influenced by retail cost and resource usage. As a result, the highest revenue and product demand are attained by reflecting the lowest selling price and energy consumption. Increased innovation, process efficiency, capacity management, and production quality improvement are among the most important advantages of adopting ML in smart industrial sectors.

First, new technologies such as Cloud computing, internet big data analytics, and cybersecurity technologies make this possible for collecting, transmitting, storing, and managing, hence expediting the development of huge amounts of data for smart industrial maintenance. As a result, big data analytics serves as the cornerstone of industrial intelligence. IoT technology uses integrated sensors, algorithms, and physical devices to gather and transfer data produced by their use and environment. Using an IoT platform or gateway, it exchanges this data. Usually, data is transmitted to the cloud for analysis and storage. Filing data into the Big Data database; data analysis using powerful and sophisticated analytical platforms, like Hadoop. Data sets are produced by IoT very quickly, frequently in real-time analytics, and require machine-driven tools like deep learning, machine learning, and artificial intelligence. Second, AI methods like machine learning, reinforcement learning, and learning techniques have emerged. Experienced significant progress in recent decades. Production surveillance is essential in the manufacturing cycle, including defect detection, remaining usable life forecasts, and quality assurance.

B. PREDICTIVE MAINTENANCE ALGORITHM AND OPTIMIZATION MODEL FOR MAINTAINING SP2

Because of the fast evolution of big data [41], cloud platforms, and ML innovations, information-related models have become widely used in predictive maintenance, prediction algorithms, and performance outcomes, which aids in reducing costs, includes production effectiveness, and enhances the standards and security of industrial production, and anticipates cumulative system performance. Safety is a key necessity in industrial production since anomalous actions of machinery or production processes can significantly decline product quality, accidents, and casualties. Based on equipment and manufacturing process monitoring data [42], the collaboration Throughout industrial processes and the targeted metrics for the whole process of production are improved.

Pausing a production process for repairs would cost time and impact the reliant assembly lines. So, by minimizing

failures during production, which leads to savings, recognizing the fault beforehand using the proper sensors would support industrial production. Prediction is critical for increasing industrial output.

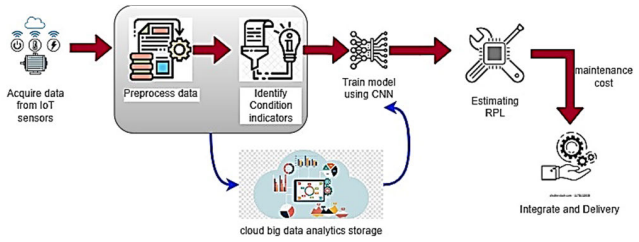


FIGURE 2. Predictive maintenance algorithm and optimization model.

As a result, IoT devices are utilized frequently to capture surveillance stored information for SP2, end product in photos, videos, and time series input in Figure 2. The Convolutional Neural Network (CNN) is used to train the model, and its implementation requires considerable pre-processing before it can use raw and processed input. It can be applied, for instance, to solve the objects detection problem [43]. Data quality is a key factor in a machine learning algorithm's performance. Pre-processing is necessary to properly use any ML algorithm and get acceptable accuracy performance if the data contain outliers and unnecessary information. The gathered data is pre-processed to remove noise, and necessary features, called Condition Indicators (CI), are extracted. The normalization of the feature is the initial phase, which is used to downscale the modification of the characteristics to prevent an extreme disparity. The extracted features are stored and analyzed in a cloud big data analytics and then passed to the ML model for implementation.

A Principal Component Analysis (PCA) was also performed on collected attributes to standardize the data number of attributes. There are several numbers of components that can be supplied to each method, ranging from 0 to 10. The value has now been changed in the parameter estimation of each method to understand better whether applying PCA to the dataset can alter the algorithm's accuracy. The features are processed in this section so they may be sent to the machine learning algorithms. Then, identified features are trained using an ML algorithm for production process maintenance. ML, DL, and other AI approaches are employed with massive data to enable smart diagnosis of abnormalities SP2 in the industry. These problems are commonly tackled using grouping and categorization techniques. Problems, either supervised or unsupervised. CNN technique is used for developing a model for making predictions by analyzing greyscale images. The input vector is designated as a convolution one to filter frames of 2D data. Data are sent from this layer to a pooling and Rectified Linear Unit (ReLU)activation function that applies a non-linear concept. This layer's output is a sequential process of the input. The data is then flattened using a Flatten layer before input to a Dense layer, where each input is linked to each outcome by a weighted score. Ultimately, the output is returned following another Activation layer.

Both monitoring data and empirical deterioration information are utilized in predictive maintenance to anticipate the RPL of industrial equipment and increase production capacity by enabling efficient maintenance plans for optimized supply chains with the best staff selection. Both surveillance data and quantitative degradation information are utilized in predictive maintenance to anticipate the Remaining Productive Lifespan (RPL) of machines, which drives the development of efficient maintenance plans. It is anticipated to be replaced by human shift maintenance, maximize system performance, and reduce cost by RPL. As a result, RPL estimation is a key focus in predictive maintenance programs. The manufacturer estimates demand based on previous monitoring data from the manufacturing line to manage defects and eliminate wastage. Finally, qualified prediction is frequently used. The quality of the product is anticipated by analyzing measurements and the production process output and operational status. The manufacturing then improved to prevent the manufacture of damaged products.

Optimization, divided into machine-level optimization and general system, is the main strategy for enhancing productivity in Industrial 5.0. The use of AI algorithms for online process parameter enhancement is critical for enhancing the performance and effectiveness of industrial operations. A production procedure includes a variety of industrial machinery, whereas an assembly line comprises many production systems. From this, the maintenance cost of equipment can be reduced or optimized. Costs are divided into various components, including maintenance, renewal, or replacement personnel expenses. The expenses vary depending on the machine's state.

C. TRANSPARENT DECISION-MAKING FOR SP2

A significant quantity of data is likely generated every second due to the development of AI and ML, sensor-based input, IoT-connected systems, virtual communities, and digital tools, necessitating real-time analysis to forecast outcomes and make quicker decisions. Big data analytics is a sophisticated analytical technique used consistently to find hidden patterns in data and link them to specific actions that aid decision-making. It supports real-time linkage to customer demand and gives transparent smart production phase order tracking. The key to ending the sequence of industrial production is decision-making, which is related to optimizing industrial processes and equipment maintenance.

Decision-making considers production factors like real economic information, production environments, implementation variables, processing guidelines, monitoring specifications, and operating machinery using robots to achieve enterprise objectives through efficiency and planning.

Algorithm 1 Reducing Maintenance Cost Using an Optimization Model

Input : repair costs (RC), replacement costs (ReC), maintenance schedules (MS), equipment lifetimes (EL), etc.
Output : optimal maintenance plan
Step 1: Define objective function to minimize total maintenance costs (TotalCost):
TotalCost = Σ [RC * Number_of_Repairs + ReC * Replacement_Decision]
Step 2: Initialize optimization variables:
Maintenance intervals for each equipment type (MI)
Number of repairs = type (NumRepairs)
Step 3: Define constraints:
(i)Maintenance intervals must be within specified ranges:
$MI\_min <= MI <= MI\_max$
(ii)Number of repairs cannot exceed maximum for each equipment type:
$NumRepairs <= MaxRepairs$
Step 4: Run optimization to find values of decision variables that minimize objective function while satisfying constraints:
$OptimalValues = minimize(TotalCost, subject\ to\ MI\_min <= MI <= MI\_max, NumRepairs <= MaxRepairs, \ldots)$
Step 5: Output optimal maintenance plan with intervals and repair/replace decisions for each equipment type:
$OptimalMaintenancePlan = \{ EquipmentType1 : \{MaintenanceInterval : OptimalValues [MI\_EquipmentType1], NumRepairs : OptimalValues [NumRepairs\_EquipmentType1]\}, EquipmentType2 : \{MaintenanceInterval : OptimalValues [MI\_EquipmentType2], NumRepairs : OptimalValues [NumRepairs\_EquipmentType2]\}, ..\}$
Step 6: Integrate optimization model with condition monitoring data to update maintenance decisions in real time:
$MI_{updated} = f(MI, ConditionMonitoringData)$
$NumRepairs\_updated = g(NumRepairs, RepairHistoryData)$
Step 7: Deliver optimized maintenance plan to maintenance team for execution
Step 8: Continuously monitor costs and equipment performance to refine optimization model and improve results over time

D. PROACTIVE MAINTENANCE FOR IDENTIFYING NETWORK VULNERABILITIES IN THE PRODUCTION PHASE

It involves real-time safety monitoring of machinery products in the smart industry using an ML-based predictive anomaly detection system. The most dangerous element of a network vulnerabilities are malicious activity identified by hackers for detecting spyware, espionage, keyloggers, Rootkits, worms, and other harmful programs on industrial endpoints and other IoT devices linked to the system. Figure 3 illustrates ML-IDS implementation using ensembling techniques for resolving network vulnerabilities issues.

The network assault can be carried out by an attacker even if they are far away from the network. While good system

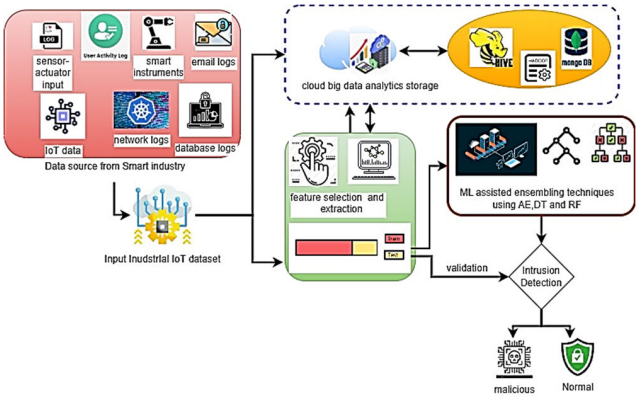


FIGURE 3. Identification of network vulnerabilities using ML-assisted IDS.

design, implementation, and deployment are required for a secure and safe industrial IoT system, ML-based solutions are extensively employed to provide extra security and safety.

The resilience of smart industrial systems also raises threats to sensitive information. The long-term challenge is to address the security threats associated with IoT devices in smart industries by implementing the following security architecture. The best approach to using this technology to acquire intelligence is using sensors, machinery devices, and IoT systems as information sources in an SP2. In the industrial realm, sensors are integrated into machinery to track and manage resources. Other factors, such as heat, moisture, and stress, can be used to identify occurrences and set off the appropriate signals.

The Manufacturing Internet of Things (MIoT), which analyses big industrial data in real time, is another scenario to maintain data security. Data screening and feature engineering are introduced in the first part. Optimizing ML algorithms' model parameters modifies the anomaly detection system. The analysis of anomalies found is the primary focus of the final segment.

E. INITIAL PROCESSING OF DATA

Thus, data are often recorded in diverse forms and have missing values. Thus, a crucial step in developing an ML model is data processing. Integrated data analysis collects, preserves, and analyzes consumer insights to provide smart production with analyzed data. Data is gathered in the first stage from various cybersecurity sensors in various smart industrial sectors for monitoring network vulnerabilities among numerous locations. Each industry uses a different source; data is gathered from various sources and stored in cloud big data storage. Data is collected from smart IoT devices, email logs, network logs, sensors, database access logs, and activity logs of users, followed by feature selection and extraction, deletion of duplicates and erroneous records, data analysis, and standardization. The pre-processed data is divided into training and testing sets for each class, 80% and 20%, respectively. After this procedure, the test dataset is saved on cloud big data storage, while the training data is delivered to the model training layer. After pre-processing, the data

is delivered via the MapReduce process implemented in a cloud big data analytics storage. The cloud server's robust computation and storage capabilities make it the best choice for model training.

The infrastructure of a cloud server is well-equipped, with a wealth of data processing tools, ML algorithms, and visualization libraries, enabling AI-enabled data processing services. MapReduce in the Hadoop framework serves as the processing model for data analysis. ML-assisted ensembling techniques for model training and classification algorithms have been used by combining three algorithms. Ensemble learning refers to techniques that produce many models and integrate them to make predictions, whether for classification or regression. The i) Auto Encoder (AE), ii) Decision Tree (DT), and iii) Random Forest (RF) analyze the intrusion in the network using bagging, boosting, and stacking approaches. Tracking time-series monitoring data on the smart industrial environment that various sensor kinds and businesses have gathered is vital yet challenging. In addition, the operating environment's data privacy must be safeguarded. AE decreases the dimensionality by classifying the model parameters and data through a weighted average distribution. It has identical proportions to the source vector from this tight space. This procedure's reconstruction error identifies a potential anomaly. The Decision Tree (DT) classifier uses a model of decisions and their potential consequences that resembles a tree. This work's target is a binary variable encompassing attack and normal scenarios. Typically, a DT classifier is employed for discrete category targets. Many tree classifiers are combined in the RF algorithm. By adding randomization to the data selection process, this classifier aims to increase variance while minimizing bias by deepening the trees to their maximum extent. The ensemble classifiers receive instruction on the train set using k-fold Cross-Validation (CV). The k-value, in this case, is 10. The test set is also subjected to these ensemble classifiers to forecast the class's malicious or normal. A digitally efficient logistics network in Industry 5.0 is built on real-time decision-making and excellent visualization.

F. DELIVERY MANAGEMENT

Delivery management enables inventory management and logistics optimization scenarios of production capacities. Intelligent supply chain management and smart transportation are two new technical advancements enabling smart logistics operations. The delivery of the product mainly depends on the production performance. The equipment's performance efficiency needs to be calculated to identify production efficiency. A measurement of how successfully a production process uses its resources, installations, workforce, and equipment throughout the planned operating times is identified in equations (1) and (2).

$$\textit{Equipment Efficiency} = \left(\frac{A}{B}\right) * \frac{\textit{total pdt wt} - \textit{scrap} - \textit{rework wt}}{\textit{total pdt wt}} * (\textit{performance}), \quad (1)$$

$$\textit{performance} = (\textit{total pdt wt}/A)/R, \quad (2)$$

where "*total pdt wt*" is the period that the equipment is operational at the fixed time given by A, t the number of hours available less any "not designated time" equals the number of hours scheduled is given as B. The successful running rate of the machine during production is given as R.

G. EVALUATION METRICS

The efficiency of the system is evaluated using the https://www.nasa.gov/intelligent-systems division turbofan dataset information. The dataset information is collected in the real-time, the system designed according to the inputs. The inputs are processed by ML techniques that are trained by passing various inputs. The frequent training procedure improves the overall system efficiency that used to implement the system in the real-time applications.

Parameters like accuracy, precision, recall, False Alarm Rate (FAR), the Receiver Operating Characteristics (ROC) for Area Under Curve (AUC), Learning Rate (LR), Vulnerability Detection Rate (VDR), and F1-score were used for performance evaluation for detecting anomalies in IoT applications of SP2. Finding True Positive prediction (TP), False Positive prediction (FP), True Negative prediction (TN), and False Negative prediction (FN) is necessary before computing. The True Positive Rate (TPR) is the count of relevant occurrences the model correctly classified as relevant. The False Positive Rate (FPR) is the count of events the model mistakenly thought were significant but irrelevant. The quantity of cases that were relevant but were mistakenly labeled as irrelevant by the model is known as the false negative rate.

Learning Timeout (LT).

Indicates how long it took to train the entire dataset and find the best-fitting training algorithm for learning timeout is calculated in equation (3).

$$LT = \textit{Final learning time} - \textit{Initial training}. \quad (3)$$

Vulnerability Detection Rate.

Equation (4) measures the Vulnerability Detection Rate (VDR) measures the proportion of abnormal data points accurately identified as anomalies using the learning algorithm

$$VDR = \textit{Final testing time} - \textit{Initial testing time}. \quad (4)$$

False Alarm Rate (FAR).

The False Alarm Rate (FAR) is the proportion of standard data wrongly classified as anomalies. Using the Receiver Operating Characteristic (ROC) curve, the 2D graph that illustrates the comparison between the VDR and the FAR is another way to assess the performance of an anomaly identification. The most effective method is one with a wide Area Under the Curve (AUC) based on ROC and a relatively minimum rate of false alarms. A high ROC value implies a model's ability to identify attacks, but a low number shows inefficiency. IDS may also be assessed based on how well they manage their time. The time performance reflects the

time the IDS requires to find an incursion. This period is made up of the dissemination and processing times. The processing period refers to the period the IDS needs to process the data to find an attack. Real-time intrusion processing requires the IDS to process data as quickly as possible; otherwise, it is impractical. The dissemination duration is the duration needed to spread information.

Accuracy prediction.

The positive class represents the attacks that have been identified. The negative class has no attacks in the network. Accuracy is one of the effectively organized SP2 in an industrial environment, and intrusion occurrence is analyzed as the original value to determine the right output, which is how detection accuracy is defined in equation (5). False Positive (FP) is given as an incorrect prediction of the positive class (attack), False Negative (FN) is given as an incorrect prediction of the negative class (no attack), True Negative gives the correctly predicted negative class (no attack), True Positive gives correctly predicts the positive class (attack).

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP} \times 100. \tag{5}$$

Recall or sensitivity calculation.

Recall, often known as ''sensitivity,'' is the percentage of retrieved instances across all pertinent instances calculated in equation (6). Precision and recall of a perfect classifier are both equal to one.

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})}. \tag{6}$$

Precision.

The percentage of genuine positives to the overall quantity of good patterns can be used to define precision. It can be calculated by equation (7).

$$Precision = \frac{TP}{(TP\ prediction + FP\ prediction)}. \tag{7}$$

F1-score.

An ML evaluation metric called the F1 score combines precision and recall ratings. To properly test model correctness, learning when and how to apply it is given in equation (8).

$$F1 - score = \frac{Two * precision * recall}{precision + recall}. \tag{8}$$

The summary of the proposed ML-SP2-IDS model enables the Industry 5.0 development into a new phase. Applying predictive maintenance and optimization for SP2 enhances the performance efficiency of the production phase compared to previously implemented algorithms. The ML-IDS implemented for network vulnerabilities issue is resolved using an ensembling concept by combining three ML algorithms, AE, RF, and DT, to accurately classify the intrusion as malicious.

IV. EXPERIMENTAL RESULTS

During the training and validation phases, most ML techniques for industrial IoT, like anomaly diagnosis, require some data. Moreover, sensor data is typically collected over an extended period and at various sampling frequencies in industrial IoT systems, which results in high dimensional dataset analysis.

The data source for SP2 is gathered from [27], in which the turbojet drum production using a smart production phase with various attributes like rotator arm, machine lifetime, and others are taken as input sources for the production phase. Network vulnerabilities in the production phase might affect the smart production performance efficiency. Suricata can serve as the foundation for building the ML-IDS, integrating with the machine learning models to enhance threat detection and response. The collected data is spilt into 80% testing and 20% training that used to measure the efficiency of the system. Hence, the production phase in the network domain needs to be monitored for vulnerabilities. Hence, an ML-IDS approach has been implemented with the data source mentioned in [28], comprising the advanced IIoT structure, the behavior of modern machines, and numerous attack kinds of malicious procedures. It includes multi-view features, network activity, host assets, logs, and warnings and provides an attack classification.

The proposed IDS in the network is implemented by utilizing the X-IIoTID, which represents the Extended Industrial IoT Intrusion Detection and gives data for identifying network vulnerabilities using three ensembling ML algorithms. X-IIoT is a comprehensive Industrial IoT-based attack data source encompassing the variability of network activity and system operations generated by diversified IoT devices. It includes nine intrusions: surveillance, weaponization, lateral movement, Command Control, Transfer information, virtual currency, Ransom Denial of Service (RDoS), and altering. It contains 820,834 cases with 68 features and two categories (421,417 as regular and 399,417 as malicious). Various data sources are used to collect and produce the attributes in the X-IIoTID dataset. It has 68 attributes, including network congestion-related features.

Eighty percent of the resource is used for training, while twenty percent is used for testing. Each smart industrial module receives an equal share of the training component, which they utilize to prepare the learning algorithm. Due to the great flexibility of smart manufacturing and data privacy concerns, it may not be possible to train these models using diverse, impartial, and comprehensive datasets in distributed and remote facilities.

Figure 4 shows that the interpolation procedure depicts the turbo deterioration's trend line for predicting machine life using ML techniques in the predictive maintenance algorithm phase. With minimal training, the matching trend line accurately predicts the RPL.

The input attributes taken from [27] give the decision variable information based on several cycles in terms of the remaining working days of the identified smart machine to replace the machine. In contrast, the vertical axis represents the Conditional indicator of the machine's health, whether faulty or normal. The machine is at the normal range when

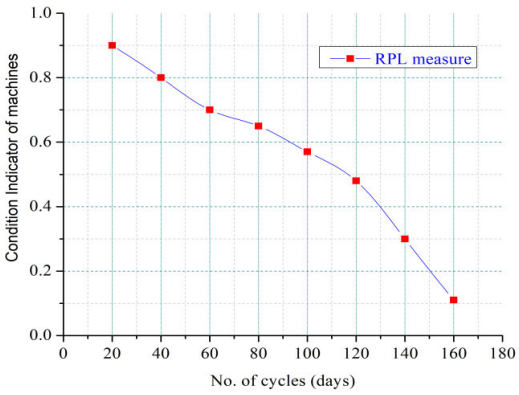


FIGURE 4. RPL prediction using predictive maintenance algorithm for calculating the life of industrial machines.

identified at index value one and finds faulty or needs repair when it reaches nearly 0.1.

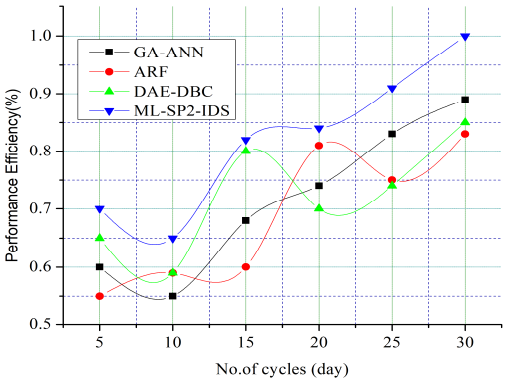


FIGURE 5. Performance efficiency of the production process in an industry.

The performance efficiency of the production line is analyzed from above in Figure 5, where the proposed ML-SP2-IDS model is compared with other existing models, GA-ANN, ARF, and DAE-DBC.

The input attributes of several cycles representing the machine's working condition in a day are taken from [27]. The performance analysis is identified from the implemented CNN algorithm for training the model with ReLu and necessary parameters. The information features like machine uptime, downtime, and scheduled period are extracted using the PCA technique. The performance efficiency of algorithms increased when the number of the working capacity of the equipment improved.

In Figure 6 above, the graph shows that with SP2's predictive maintenance algorithm, maintenance costs are reduced with the help of RPL measures calculated for the machine's longevity. The probability of working days of machinery with necessary CI conditional features is taken from $p\,(CI)$. The minimum cost maintenance of equipment is calculated from the optimized model parameter $minimum(MnC)$. The number of machines presented in a given CI concerning the total serving days compared to the total lifetime of other machines

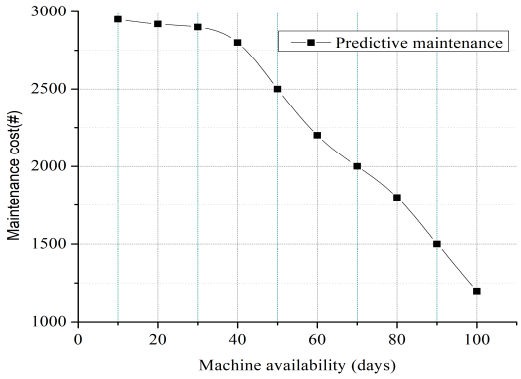


FIGURE 6. Maintenance cost reduction using predictive maintenance algorithm and optimization model in SP2.

with the same production line is given as Z_{opt}. The amount of equipment maintenance must be maintained from various CI values and calculated.

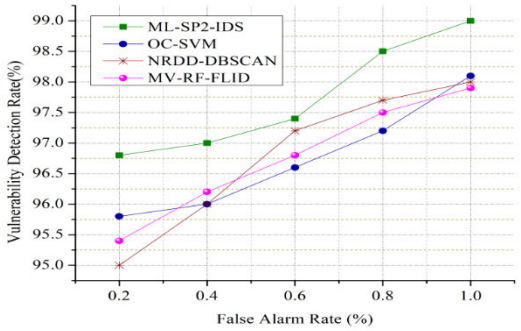


FIGURE 7. The trade-off between VDR and FAR.

Figure 7 shows that the trade-off between the network vulnerabilities detection rate using equation 3 and the FAR rate is represented using an ROC curve. Since the proportions of attack detection and false alarms are sometimes at odds with one another for evaluation comparison, Receiver Operating Characteristics (ROC) analysis is also used to evaluate IDS.

The proposed ML-SP2-IDS model outperforms all existing models, ML-RF-FLID, NRDD-DBSCAN, and OC-SVM. From the above Figure 8, the input attributes categorized as reconnaissance, weaponization, and Ransom Denial of Service (RDoS) are taken for performance analysis from [28] for identifying vulnerabilities in the industrial network for safe and better production process with metrics such as accurateness, precision-recall, and F-measurement score in all scenarios, by using the suggested model for varying numbers of ensembled ML techniques like AE, RF, and DT with bagging, boosting and stacking concepts to analyze network vulnerabilities. The detection accuracy performed well and classified the outcome as malicious or normal behavior based on the implemented techniques.

Figure 9 depicts the implemented ML-IDS model for network vulnerabilities for different epoch sizes for incremental iterations. The train test split function validates the implemented ensembling ML algorithms for various production phases in the network of Industry 5.0. The input attributes

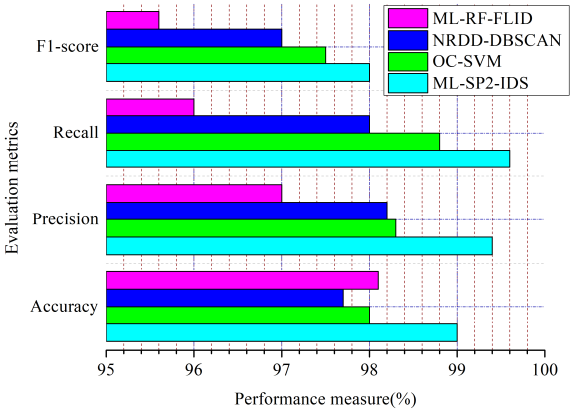


FIGURE 8. Comparison of ML-SP2-IDS model with other ML-based IDS techniques.

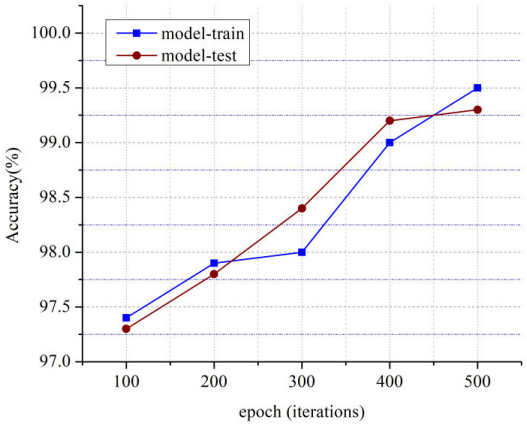


FIGURE 9. Accuracy prediction of ML-IDS based on various iterations for train-test split.

TABLE 2. Computation time analysis.

Methods	Computation Time (s)
ML-SP2-IDS	9.87
SDN/NFV	13.49
NRDD-DBSCAN	15.97
MV-RF-FLID	18.23

are taken from [28] for various attribute types of attack types with model-train and model-test to solve the bias-variance trade-off. In addition, the computation time for the proposed ML-SP2-IDS is computed to measure how much time the proposed approach takes to run the industrial information. The computation time measures the starting and running time for entire task. Then the obtained computation time is shown in the table 2.

From the table 2 it clearly shows that introduced ML-SP2-IDS approach attains minimum computation time (9.87s) compared to other methods such as SND/NFV (13.49s), NRDD-DBSCAN (15.97s) and MV-RF-FLID (18.23s). from the analysis it clearly shows that proposed algorithm effectively utilizes the hardware component that influences the computation time. In addition, the proposed method

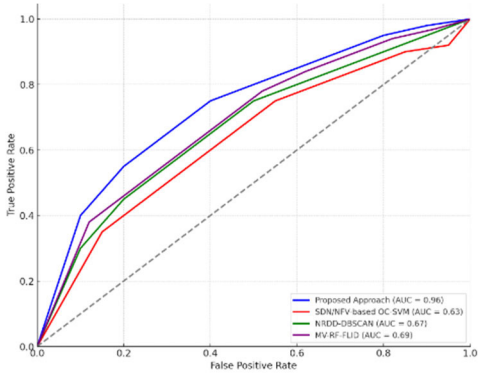


FIGURE 10. ROC- AUC analysis.

TABLE 3. Comparative analysis.

Criteria	Proposed Model	SDN/NFV-based OC-SVM	NRDD-DBSCAN	MV-RF-FLID
Performance Metrics	High accuracy, low false positives	Moderate accuracy, potential false positives	High accuracy, effective in anomaly detection	High accuracy, strong ensemble performance
Architecture	Deep learning-based CNN with feature extraction	One-Class SVM in SDN/NFV environment	Density-based clustering with redundancy removal	Federated learning with Random Forest ensemble
Computational Efficiency	Moderate to high (dependent on hardware)	High efficiency, but lower complexity	Lower efficiency due to clustering and redundancy processing	Moderate efficiency, computationally intensive due to ensemble and distributed nature
Scalability	High scalability with potential parallel processing	High scalability in SDN/NFV environments	Moderate scalability, limited by clustering complexity	High scalability with federated learning, but requires robust infrastructure
Data Handling	Capable of handling large, high-dimensional data	Effective in network traffic analysis	Best for clustering spatial data, may struggle with large-scale data	Handles distributed data effectively, good for federated setups
Model Complexity	High complexity due to CNN architecture	Low to moderate complexity	Moderate complexity with DBSCAN enhancements	High complexity due to ensemble and federated learning
Adaptability to New Threats	High adaptability through retraining CNNs	Limited adaptability, requires model retraining	Moderate adaptability with clustering adjustments	High adaptability, retraining across nodes improves model generalization
Real-time Application	Suitable for real-time detection with optimized implementation	Highly suitable for real-time due to SDN/NFV integration	May struggle in real-time due to clustering overhead	Suitable, but may face latency in federated learning communication
Infrastructure Requirements	Requires GPUs/TPUs for optimal performance	Moderate, leverages existing SDN/NFV infrastructure	Higher computational resources for clustering	High, needs a robust federated learning infrastructure
Strengths	Strong pattern recognition, robust against complex threats	Efficient in network-based environments	Effective in identifying non-redundant patterns	High accuracy, privacy-preserving with federated learning
Weaknesses	Higher computational cost, complex architecture	May miss subtle anomalies, requires precise configuration	Computationally intensive, scalability challenges	Complexity and infrastructure demands may limit deployment in resource-constrained

effectively works on the large dataset by effectively utilizes the machine learning techniques. The convolution neural networks and training process reduces the complexity while

exploring the large data. Finally, the method attains the high scalability in the industrial applications.

From the above analysis, network vulnerability of industry 5.0 is detected with the help of the Convolutional Neural Networks (CNNs) in the context of the ML-assisted Smart Production Process (ML-SP2). During the analysis, the ROC-AUC curve analysis is incorporated with the other metrics such as accuracy, F1-score etc. The ROC-AUC metrics identifies the model identifies the difference between the class with different thresholds. The metric effectively identifies the imbalanced dataset network security concerns. The described metrics provides the effective evaluation and used to justify the system robustness while analyzing the imbalanced data in the industry 5.0 real-time applications. The results summarized all the implemented graphs performed well for the proposed ML-SP2-IDS model and outperformed other existing algorithms using intelligent ML algorithms in the smart industrial production phase. In addition, the excellence of the ML-SP2-IDS system is compared with existing approaches and the obtained comparison is shown in Table 3.

V. CONCLUSION

For decades, industrial smart production process enhancement goals have been improving production efficiency, boosting worker productivity, reducing scrap waste, and prolonging network lifetime by ensuring security and integrity. The fact that the new ML solution makes it easier to increase industrial operators' knowledge and simplify the production process environment using the platform's knowledge is a major point. The ML-SP2-IDS model achieves it by implementing ML algorithms in smart industries using predictive maintenance and optimization and ensembling techniques sorted by bagging and boosting the production performance. During the analysis, different machine learning algorithms are incorporated that reduce the intermediate activities and identifies the abnormal activities with maximum recognition accuracy. In addition, the system uses the convolution network that fine-tune and train the networks that reduce the false acceptance rate.

The ML-SP2-IDS algorithm has added advantages such as equipment safety, increased reliability, and cost reduction. The smart industrial production phase's effectiveness is measured by the RPL (0.9), maintenance cost (2500$), production performance (1.0), VDR (97.6%), FAR, accuracy (99%), precision (94%)-recall (96%), and f1-measure (98%) compared to the existing methods such as ML-RF-FLID, NRDD-DBSCAN, and OC-SVM. Yes, it has limitations related to data protection, cost intensity, scalability, operational and the complexity of the smart transformation for robots in Industry 5.0. Limitations of the ML models used, such as overfitting or sensitivity to data quality. Possible socio-economic impacts, such as job displacement due to increased automation. Further work will include presenting novel ML-assisted ensembling techniques for IDS to prevent Industry 5.0 from adversarial and classic intrusion diagnosis with minimal computational overhead. In addition, encryption methodologies like AES, DES and block chain approaches are utilized for ensuring the data security and privacy.

AUTHOR CONTRIBUTIONS

The authors' contributions are as follows: ''Conceptualization, Vadym Shkarupylo and Andrii Oliinyk; methodology, Jamil Abedalrahim Jamil Alsayaydeh; software, Vadym Shkarupylo; validation, Volodymyr Artemchuk; formal analysis, Andrii Oliinyk; investigation, Vadym Shkarupylo and Jamil Abedalrahim Jamil Alsayaydeh; resources, Volodymyr Artemchuk; writing—original draft preparation, Vadym Shkarupylo and Safarudin Gazali Herawan; writing—review and editing, Jamil Abedalrahim Jamil Alsayaydeh and Mohd Faizal bin Yusof; funding acquisition, Mohd Faizal bin Yusof and Safarudin Gazali Herawan.

ACKNOWLEDGMENT

The authors would like to thank the Centre for Research and Innovation Management (CRIM), Universiti Teknikal Malaysia Melaka (UTeM), for their valuable support in this research.

REFERENCES

[1] I. Kardush, S. Kim, and E. Wong, "A techno-economic study of Industry 5.0 enterprise deployments for human-to-machine communications," *IEEE Commun. Mag.*, vol. 60, no. 12, pp. 74–80, Dec. 2022, doi: 10.1109/MCOM.001.2101068.

[2] C. N. N. Tran, T. T. H. Tat, V. W. Y. Tam, and D. H. Tran, "Factors affecting intelligent transport systems towards a smart city: A critical review," *Int. J. Construction Manage.*, vol. 23, no. 12, pp. 1982–1998, Sep. 2023, doi: 10.1080/15623599.2022.2029680.

[3] J. M. Rožanec, I. Novalija, P. Zajec, K. Kenda, H. T. Ghinani, S. Suh, E. Veliou, D. Papamartzivanos, T. Giannetsos, S. A. Menesidou, R. Alonso, N. Cauli, A. Meloni, D. R. Recupero, D. Kyriazis, G. Sofianidis, S. Theodoropoulos, B. Fortuna, D. Mladenić, and J. Soldatos, "Human-centric artificial intelligence architecture for Industry 5.0 applications," *Int. J. Prod. Res.*, vol. 61, no. 20, pp. 6847–6872, Oct. 2023, doi: 10.1080/00207543.2022.2138611.

[4] N. Prakash, S. Sharma, M. Bhardwaj, and R. K. Mukherji, "Industry 5.0: A paradigm shift towards human-centric industrial revolution," *EEO*, vol. 20, no. 1, pp. 6912–6922, 2021.

[5] N. Deepa, Q.-V. Pham, D. C. Nguyen, S. Bhattacharya, B. Prabadevi, T. R. Gadekallu, P. K. R. Maddikunta, F. Fang, and P. N. Pathirana, "A survey on blockchain for big data: Approaches, opportunities, and future directions," *Future Gener. Comput. Syst.*, vol. 131, pp. 209–226, Jun. 2022, doi: 10.1016/j.future.2022.01.017.

[6] A. Knap-Stefaniuk, "The skills members of multicultural teams need to succeed in Industry 5.0—The opinion of managers from Portugal, France, and Greece," *Proc. Comput. Sci.*, vol. 225, pp. 1478–1485, Jan. 2023, doi: 10.1016/j.procs.2023.10.136.

[7] B. Chander, S. Pal, D. De, and R. Buyya, "Artificial intelligence-based Internet of Things for Industry 5.0," in *Artificial Intelligence-Based Internet of Things Systems*. (Internet of Things), S. Pal, D. De, and R. Buyya, Eds. Cham, Switzerland: Springer, 2022, pp. 3–45.

[8] X. Chen, M. A. Eder, A. Shihavuddin, and D. Zheng, "A human-cyber-physical system toward intelligent wind turbine operation and maintenance," *Sustainability*, vol. 13, no. 2, p. 561, Jan. 2021, doi: 10.3390/su13020561.

[9] M. Aljanabi, "Safeguarding connected health: Leveraging trustworthy AI techniques to harden intrusion detection systems against data poisoning threats in IoMT environments," *Babylonian J. Internet Things*, vol. 2023, pp. 31–37, May 2023, doi: 10.58496/bjiot/2023/005.

[10] M. Biró, A. Mashkoor, and J. Sametinger, "Safe and secure cyber-physical systems," *J. Softw., Evol. Process*, vol. 33, no. 9, 2021, Art. no. e2340, doi: 10.1002/smr.2340.

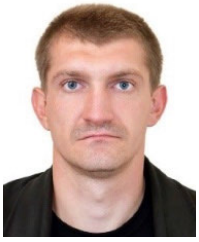
[11] A. Leiden, C. Herrmann, and S. Thiede, "Cyber-physical production system approach for energy and resource efficient planning and operation of plating process chains," *J. Cleaner Prod.*, vol. 280, Jan. 2021, Art. no. 125160, doi: 10.1016/j.jclepro.2020.125160.
[12] M. Soori, B. Arezoo, and R. Dastres, "Internet of Things for smart factories in Industry 4.0, a review," *Internet Things Cyber-Phys. Syst.*, vol. 3, pp. 192–204, Jan. 2023, doi: 10.1016/j.iotcps.2023.04.006.
[13] S. Bag, M. S. Rahman, S. Gupta, and L. C. Wood, "Understanding and predicting the determinants of blockchain technology adoption and SMEs' performance," *Int. J. Logistics Manage.*, vol. 34, no. 6, pp. 1781–1807, Dec. 2023, doi: 10.1108/ijlm-01-2022-0017.
[14] A. Alsajri and A. Steiti, "Intrusion detection system based on machine learning algorithms:(SVM and genetic algorithm)," *Babylonian J. Mach. Learn.*, vol. 2024, pp. 15–29, Jan. 2023, doi: 10.58496/bjml/2024/002.
[15] S. Shitharth, A. M. Alshareef, A. O. Khadidos, K. H. Alyoubi, A. O. Khadidos, and M. Uddin, "A conjugate self-organizing migration (CSOM) and reconciliate multi-agent Markov learning (RMML) based cyborg intelligence mechanism for smart city security," *Sci. Rep.*, vol. 13, no. 1, Sep. 2023, Art. no. 15681, doi: 10.1038/s41598-023-42257-0.
[16] S. Sharma, L. Koehl, P. Bruniaux, X. Zeng, and Z. Wang, "Development of an intelligent data-driven system to recommend personalized fashion design solutions," *Sensors*, vol. 21, no. 12, p. 4239, Jun. 2021, doi: 10.3390/s21124239.
[17] J. A. J. Alsayaydeh, "Development of programmable home security using GSM system for early prevention," *ARPN J. Eng. Appl. Sci.*, vol. 16, no. 1, pp. 88–97, 2021.
[18] B. Caiazzo, T. Murino, A. Petrillo, G. Piccirillo, and S. Santini, "An IoT-based and cloud-assisted AI-driven monitoring platform for smart manufacturing: Design architecture and experimental validation," *J. Manuf. Technol. Manage.*, vol. 34, no. 4, pp. 507–534, May 2023, doi: 10.1108/jmtm-02-2022-0092.
[19] V. Kovtun, K. Grochla, S. Aldosary, and M. Al-Maitah, "Analysis of direct traffic at the transport protocol level in the WiMax-1/2 cluster oriented to offload the smart city's wireless ecosystem," *Roy. Soc. Open Sci.*, vol. 11, no. 7, Jul. 2024, Art. no. 240206, doi: 10.1098/rsos.240206.
[20] H. Ding, R. X. Gao, A. J. Isaksson, R. G. Landers, T. Parisini, and Y. Yuan, "State of AI-based monitoring in smart manufacturing and introduction to focused section," *IEEE/ASME Trans. Mechatronics*, vol. 25, no. 5, pp. 2143–2154, Oct. 2020, doi: 10.1109/TMECH.2020.3022983.
[21] N. F. B. A. Rahim, A. W. Y. Khang, A. Hassan, S. J. Elias, J. A. M. Gani, J. Jasmis, and J. Abedalrahim, "Channel congestion control in VANET for safety and non-safety communication: A review," in *Proc. 6th IEEE Int. Conf. Recent Adv. Innov. Eng. (ICRAIE)*, vol. 6, Dec. 2021, pp. 1–6, doi: 10.1109/ICRAIE52900.2021.9704017.
[22] E. Ouda, M. Maalouf, and A. Sleptchenko, "Machine learning and optimization for predictive maintenance based on predicting failure in the next five days," in *Proc. 10th Int. Conf. Oper. Res. Enterprise Syst.*, 2021, pp. 192–199.
[23] V. Kovtun, K. Grochla, T. Altameem, and M. Al-Maitah, "Evaluation of the QoS policy model of an ordinary 5G smart city cluster with predominant URLLC and eMBB traffic," *PLoS ONE*, vol. 18, no. 12, Dec. 2023, Art. no. e0295252, doi: 10.1371/journal.pone.0295252.
[24] R. R. K. Menon, Fawaz-Al-Faizi, and U. L. Sooraj, "Novelty detection of text using spectral graphs and visualization," in *Proc. 4th IEEE Global Conf. Advancement Technol. (GCAT)*, Oct. 2023, pp. 1–8.
[25] W. A. Indra, A. W. Y. Khang, Y. T. Yung, and J. A. J. Alsayaydeh, "Radio-frequency identification (RFID) item finder using radio frequency energy harvesting," *ARPN J. Eng. Appl. Sci.*, vol. 14, no. 20, pp. 3554–3560, 2019.
[26] A. Frankó, G. Hollósi, D. Ficzere, and P. Varga, "Applied machine learning for IIoT and smart production—Methods to improve production quality, safety and sustainability," *Sensors*, vol. 22, no. 23, p. 9148, Nov. 2022, doi: 10.3390/s22239148.
[27] Z. A. Shaikh, A. Kraikin, A. Mikhaylov, and G. Pinter, "Forecasting stock prices of companies producing solar panels using machine learning methods," *Complexity*, vol. 2022, no. 1, Jan. 2022, Art. no. 9186265, doi: 10.1155/2022/9186265.
[28] M. Bagaa, T. Taleb, J. B. Bernabe, and A. Skarmeta, "A machine learning security framework for IoT systems," *IEEE Access*, vol. 8, pp. 114066–114077, 2020, doi: 10.1109/ACCESS.2020.2996214.
[29] M. A. Samara, I. Bennis, A. Abouaissa, and P. Lorenz, "A survey of outlier detection techniques in IoT: Review and classification," *J. Sensor Actuator Netw.*, vol. 11, no. 1, p. 4, Jan. 2022, doi: 10.3390/jsan11010004.
[30] V. Kovtun, K. Grochla, and K. Połys, "The concept of network resource control of a 5G cluster focused on the smart city's critical infrastructure needs," *Alexandria Eng. J.*, vol. 94, pp. 248–256, May 2024, doi: 10.1016/j.aej.2024.03.038.
[31] H. Ghallab, H. Fahmy, and M. Nasr, "Detection outliers on Internet of Things using big data technology," *Egyptian Informat. J.*, vol. 21, no. 3, pp. 131–138, Sep. 2020, doi: 10.1016/j.eij.2019.12.001.
[32] M. Rodríguez, D. P. Tobón, and D. Múnera, "Anomaly classification in industrial Internet of Things: A review," *Intell. Syst. with Appl.*, vol. 18, May 2023, Art. no. 200232, doi: 10.1016/j.iswa.2023.200232.
[33] D. C. Attota, V. Mothukuri, R. M. Parizi, and S. Pouriyeh, "An ensemble multi-view federated learning intrusion detection for IoT," *IEEE Access*, vol. 9, pp. 117734–117745, 2021, doi: 10.1109/ACCESS.2021.3107337.
[34] *Intelligent Systems Division*. Accessed: Aug. 29, 2024. [Online]. Available: https://www.nasa.gov/intelligent-systemsdivision#turbofan
[35] M. Al-Hawawreh, E. Sitnikova, and N. Aboutorab, "X-IIoTID: A connectivity-agnostic and device-agnostic intrusion data set for industrial Internet of Things," *IEEE Internet Things J.*, vol. 9, no. 5, pp. 3962–3977, Mar. 2022, doi: 10.1109/JIOT.2021.3102056.
[36] M. Ryalat, H. ElMoaqet, and M. AlFaouri, "Design of a smart factory based on cyber-physical systems and Internet of Things towards Industry 4.0," *Appl. Sci.*, vol. 13, no. 4, p. 2156, Feb. 2023, doi: 10.3390/app13042156.
[37] W. Yan, J. Wang, S. Lu, M. Zhou, and X. Peng, "A review of real-time fault diagnosis methods for industrial smart manufacturing," *Processes*, vol. 11, no. 2, p. 369, Jan. 2023, doi: 10.3390/pr11020369.
[38] A. Salam, F. Ullah, F. Amin, and M. Abrar, "Deep learning techniques for Web-based attack detection in industry 5.0: a novel approach," *Technologies*, vol. 11, no. 4, p. 107, Aug. 2023, doi: 10.3390/technologies11040107.
[39] A. Adel, "Unlocking the future: Fostering human–machine collaboration and driving intelligent automation through Industry 5.0 in smart cities," *Smart Cities*, vol. 6, no. 5, pp. 2742–2782, Oct. 2023, doi: 10.3390/smartcities6050124.
[40] J. A. J. Alsayaydeh, M. F. B. Yusof, C. K. Hern, M. R. Ahmad, V. Shkarupylo, and S. G. Herawan, "Greenhouse horticulture automation with crops protection by using Arduino," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 10, pp. 114–123, 2023, doi: 10.14569/ijacsa.2023.0141013.
[41] S. Leoshchenko, A. Oliinyk, S. Subbotin, M. Ilyashenko, and T. Kolpakova, "Neuroevolution methods for organizing the search for anomalies in time series," in *Proc. CEUR Workshop*, vol. 3392, 2023, pp. 164–176.
[42] N. Kulykovska, A. V. Timenko, S. S. Hrushko, and V. V. Shkarupylo, "Automated Internet of Things system for monitoring indoor air quality," in *Proc. CEUR Workshop*, vol. 3666, 2024, pp. 97–104.
[43] D. Borovyk, R. Fedoniuk, A. Oliinyk, S. Subbotin, and T. Kolpakova, "Detection of vehicles in aerial photographs using convolutional neural networks," in *Proc. CEUR Workshop*, vol. 3699, 2024, pp. 161–179.

VADYM SHKARUPYLO received the M.S. degree in computer systems and networks from Zaporizhzhia National Technical University, Zaporizhzhia, Ukraine, in 2010, and the Ph.D. degree in computer systems and components from the G. E. Pukhov Institute for Modelling in Energy Engineering, National Academy of Sciences of Ukraine, Kyiv, Ukraine, in 2014. From 2013 to 2015, he was a Senior Lecturer with the Department of Computer Systems and Networks, Zaporizhzhia National Technical University, where he was an Associate Professor with the Department of Computer Systems and Networks, from 2015 to 2018. Since 2018, he has been an Associate Professor with the Department of Computer Systems, Networks and Cybersecurity, National University of Life and Environmental Sciences of Ukraine, Kyiv. Since 2019, he has also been a Senior Research Fellow of the Department of Mathematical and Computer Modelling, G. E. Pukhov Institute for Modelling in Energy Engineering, National Academy of Sciences of Ukraine, Kyiv. In 2024, he has defended his Dr.Sc. degree in computer systems and components in the G. E. Pukhov Institute for Modelling in Energy Engineering, National Academy of Sciences of Ukraine. His research interests include formal methods application in safety-critical scenarios, techniques of verification and validation, web services, and the Internet of Things.

JAMIL ABEDALRAHIM JAMIL ALSAYAYDEH (Member, IEEE) received the degree in computer engineering and the M.S. degree in computer systems and networks from Zaporizhzhya National Technical University, Ukraine, in 2009 and 2010, respectively, and the Ph.D. degree in engineering sciences with a specialization in automation of control processes from National Mining University, Ukraine, in 2014. He has been a Senior Lecturer with the Department of Engineering Technology, Faculty of Electronic and Computer Engineering and Technology, Universiti Teknikal Malaysia Melaka (UTeM), since 2015. His teaching portfolio includes a range of courses, such as computer network and security, internet technology and multimedia, software engineering, computer system engineering, data communications and computer networks, computer network and systems, real time systems, programming fundamental, digital signal processing, and advanced programming. His research interests include formal methods, simulation, the Internet of Things, computing technology, artificial intelligence and machine learning: computer architecture, algorithms, and applications; he has more than 65 research publications to his credit that are indexed in SSCI, SCIE, and Scopus, which cited by over 300 documents. He supervised undergraduate and postgraduate students and he is a reviewing member of various reputed journals. Currently, he actively publishes research articles, received grants from the government and private sectors, universities, and international collaboration. He is also a Research Member of the Center for Advanced Computing Technology. He is also a member of Board of Engineers Malaysia (BEM).

MOHD FAIZAL BIN YUSOF received the Bachelor of Science degree in electrical engineering from Northwestern University, in 1998, and the M.B.A. degree in technology entrepreneurship from Universiti Teknologi Malaysia, in 2008. He is currently an Associate Researcher and a Lecturer with Rabdan Academy, United Arab Emirates. He is an Experienced Blockchain Researcher, an University Technology Transfers Officer, a Software Developer, and a Former Start-Ups Entrepreneur. His research interests include design science research, blockchain, cryptocurrency, artificial intelligence, and social entrepreneurship.

ANDRII OLIINYK received the master's degree in software engineering from Zaporizhzhia National Technical University, Ukraine, in 2007, and the Ph.D. degree in artificial intelligence, in 2009. He is currently pursuing the Doctor of Science degree in artificial intelligence. He was a Professor with the Software Tools Department, National University "Zaporizhzhia Polytechnic," in 2021. He is the author/co-author of more than 100 research publications cited in over 300 documents. His research interests include artificial intelligence, big data, neural networks, computer vision, and soft computing. He actively supervises Ph.D. students, reviews for reputable journals, and secures grants.



VOLODYMYR ARTEMCHUK received the master's degree in software engineering from Zhytomyr State Technological University, Ukraine, in 2008, the Ph.D. degree in mathematical modeling and computational methods, in 2012, and the Doctor of Science degree in ecological safety, in 2021. He is currently the Deputy Director of the G. E. Pukhov Institute for Modelling in Energy Engineering, NAS of Ukraine. He is the author/co-author of more than 100 research publications cited in over more than 700 documents. His research interests include mathematical modeling, energy resilience, sustainable development, artificial intelligence, and soft computing. He actively supervises Ph.D. students, reviews for reputable journals, and secures grants.

SAFARUDIN GAZALI HERAWAN is currently a Senior Lecturer with Bina Nusantara University, Jakarta, Indonesia. His current research interests include automotive engineering, renewable energy, and heat recovery technologies, where he is the author/co-author of over 80 research publications which cited by over 290 documents.