Hindawi
Wireless Communications and Mobile Computing
Volume 2021, Article ID 7154587, 17 pages
https://doi.org/10.1155/2021/7154587

Research Article

Intrusion Detection in Industrial Internet of Things Network-Based on Deep Learning Model with Rule-Based Feature Selection

Joseph Bamidele Awotunde,[1] **Chinmay Chakraborty**,[2] **and Abidemi Emmanuel Adeniyi**[3]

[1]*Department of Computer Science, University of Ilorin, Ilorin, Nigeria*
[2]*Department of Electronics and Communication Engineering, Birla Institute of Technology, Mesra, Jharkhand, India*
[3]*Department of Computer Science, Landmark University, Omu-Aran, Nigeria*

Correspondence should be addressed to Joseph Bamidele Awotunde; awotunde.jb@unilorin.edu.ng

Received 19 May 2021; Revised 21 July 2021; Accepted 4 August 2021; Published 3 September 2021

Academic Editor: Alireza Jolfaei

The Industrial Internet of Things (IIoT) is a recent research area that links digital equipment and services to physical systems. The IIoT has been used to generate large quantities of data from multiple sensors, and the device has encountered several issues. The IIoT has faced various forms of cyberattacks that jeopardize its capacity to supply organizations with seamless operations. Such risks result in financial and reputational damages for businesses, as well as the theft of sensitive information. Hence, several Network Intrusion Detection Systems (NIDSs) have been developed to fight and protect IIoT systems, but the collections of information that can be used in the development of an intelligent NIDS are a difficult task; thus, there are serious challenges in detecting existing and new attacks. Therefore, the study provides a deep learning-based intrusion detection paradigm for IIoT with hybrid rule-based feature selection to train and verify information captured from TCP/IP packets. The training process was implemented using a hybrid rule-based feature selection and deep feedforward neural network model. The proposed scheme was tested utilizing two well-known network datasets, NSL-KDD and UNSW-NB15. The suggested method beats other relevant methods in terms of accuracy, detection rate, and FPR by 99.0%, 99.0%, and 1.0%, respectively, for the NSL-KDD dataset, and 98.9%, 99.9%, and 1.1%, respectively, for the UNSW-NB15 dataset, according to the results of the performance comparison. Finally, simulation experiments using various evaluation metrics revealed that the suggested method is appropriate for IIOT intrusion network attack classification.

1. Introduction

A modern industrial revolution brings deep change and human growth, resulting in "Automation of Everything." It uses computer networks to link both digital devices, data mining, and real-world application management [1]. This revolution's opportunity helps everybody to access trillions of data and information that brings new opportunities. Significant increases in efficiency in the physical and digital industries may be felt by humans, resulting in a better quality of life and a more prosperous society. The creation of vast quantities of data from various sensors is popular in the Industrial Internet of Things (IIoT) world. These applications can be felt in various industries like healthcare, retail, automotive, and transport. In many industries, the IIoT can greatly increase efficiency, productivity, and operational efficiency. The IIoT will first develop existing processes and facilities, but the ultimate aim is to create completely new and vastly enhanced goods and services. Many companies recognize how and where IIoT innovations and solutions can lead to organizational changes, new and improved goods and services, and entirely new business models. On the IIoT, machine learning and deep learning algorithms can increase reliability, production, and customer satisfaction by combining technological innovations, sensors, programs, and applications.

Anything necessitates a wide range of technology that must be carefully integrated and orchestrated. These advancements in technology allow intelligent machines,

machinery, appliances, and integrated automation systems [2, 3] to automate routine operations and solve complex problems without human interference. Improvements in the smart workplace, smart data exploration, cognitive automation, and other aspects of business smartness should all be included. A digital twin is a virtual representation of physical assets, systems, and so on. It is commonly known as the Internet of Things (IoT), which is constantly developing all of these appliances and supplying us with an incredibly growing dataset that can be evaluated for efficiency, architecture, maintenance, and a host of other issues. A key feature of any digital twin is that it is constantly updated and "learns" any changes that arise in real-time. The IoT concept and its solutions have made a lot of changes in the physical world.

Since the cloud has altered how individuals and organizations communicate and perform business online, cyberspace plays an important role in today's societies and economies [4, 5]. As a result, the IIoT encompasses a variety of devices, software, and facilities that bridge the gap between the virtual and physical worlds [6]. Due to the connectivity of information technology (IT) and organizational technology (OT), industrial systems that depend on locked and exclusive communication systems are vulnerable to a wide range of interference activities [7, 8].

Machine-to-machine (M2M) and machine-to-person (M2P) connections to the network are used in IIoTs using the TCP/IP interface using various IIoT protocols [9, 10]. The number IIoTs have the number of flaws and bugs that can be abused using a range of advanced attack methods which has increased significantly. The attackers attempt to take advantage of these processes to steal sensitive information, commit financial funds, and corrupt device resources [11]. If the cybersecurity domain does not discover interesting mitigation strategies for stopping cyberthreats to the IIoT, it is estimated that they will cost up to $90 trillion by 2030 [12].

Protecting vital services and infrastructure is becoming a more critical problem in every organization as the volume of IIoT devices and implementations continues to grow [13]. Among the most frequent risks in IIoT networks is malware that abuses zero-day vulnerabilities. The perpetrators infect vulnerable computers to track and change their activities, using a variety of techniques like Progressive Determined Risk (PDR), Denial-of-Service (DoS), and Decentralized DoS. (DDoS). For instance, in 2010, the Stuxnet worm attacked Iran's nuclear program, in 2013, Iranian hackers hacked into the ICS of a dam in New York, and in 2015, the black-energy passive attack was explicitly equivalent to approximately 80.000 power outages in Ukraine [14, 15]. These nefarious practices showed that conventional cyberthreat methods, like security protocols, cryptography, access controls, and biometrics Interruption Discovery Systems (IDSs), are no longer sufficient for delivering successful vital infrastructure protection.

As a network security tonic, the network intrusion detection system (NIDS) is important in detecting and addressing all Internet attacks. The IIoT has become an essential portion of present machinery for data and knowledge transfer, necessitating the need for global network security [16]. To safeguard workstation schemes from multiple grid invasions, network intrusion detection systems (NIDS) are often used to recognize system traffic. In [17], intrusion is a framework that attempts to break information system's security services. Researchers have been inspired to create new IDSs in response to the threats posed by these invasive frameworks. Several intrusion detection systems (IDS) have previously been developed and upgraded, but they are still susceptible to a range of assaults. An increasing interest in anomaly detection research is due to IDS' ability to track and forecast malicious behavior unknown assaults. However, current machine learning-based irregularity discovery methods still have a high false alarm rate [18].

Recently, findings indicate that feature extraction is now at the core of a more accurate IDS [19, 20]. In most detection methods, the feature selection technique is used to pick the fitness values which input attributes for classification models, with the goal of aggregate discovery performance and reducing error rate in NIDS [8]. In particular, classifier feature vectors are massive, and not all of them apply to the groups to be categorized, requiring the use of a feature selection strategy. Conversely, feature selection approaches can be divided into three categories: filter approach, wrapper approach, and embedded approach [21]. The most popular feature selection strategy focused on selecting the best-fitted functionality which relies on dataset measurements lacking seeing classifier's performance. The wrapper method, on the other hand, is superior since the classifier feedback is used to evaluate the quality of the feature subclass, leading to higher prediction performance. The integrated process is analogous to wrapper approaches in that an intrinsic process modeling function in the classifier could be used to improve the learning algorithm's search efficiency.

Until now, several various categories for IDSs have been planned. Depending on the classification algorithm utilized, intrusion detection systems (IDSs) may be categorized as rule-based, misappropriation discovery, or diverse schemes. IDSs may either be classified as real-time if they use persistent system tracking or as sporadic or inactive if the tracking occurs only occasionally taking place at fixed times or even offline using data collected and processed over some time. Furthermore, new classifications have recently been introduced while discussing Industrial Control Systems (ICSs) with unique criteria and characteristics. The authors of [22] suggested a new classification system for IDSs called ICS, which are classified into three types: protocol review, traffic processing, and control process modeling.

Countermeasures are taken based on the information gathered from the detection systems about the identified attacks. The more accurately the type of attack is classified, the more effective the chosen countermeasures will be, and the less they will interfere with the device or network's proper operation. Furthermore, in some situations, countermeasures may have more severe effects than the attack itself if we do not detect the same form of attack. As a result, we aim to develop an intrusion prevention method that has proven expertise in each type of attack Moreover, for both

routine and irregular assaults, our system must have a low false alarm rate and a high detection accuracy, allowing limited processing to correctly classify. The latter function is important because intrusion detection systems are used in industrial control systems that operate critical infrastructures, where reliable and timely warning of cyberthreats is critical [23].

The feature extraction strategy is effective for the design and execution of legitimate security solutions, as well as for improving IDS performance [24–26]. In certain phenomenon detection methods, the need for greater accuracy and a lowered false alarm rate inspired the concept of data preprocessing and identification as the two mutual levels for IDS prototypes [27, 28]. The preprocessing phase removes the identification process which uses the reduction of attributes after removing redundant features from the dataset, retaining a decreased feature set that can be used to generate a high-performance version to predict attack classes using the base classifier.

Therefore, based on [8], this paper integrates the emerging infrastructure for applications of the Industrial Internet of Things. The authors reviewed the proposed scheme, offered the incorporation of work into a three-tier design for IIoT systems, and tested it against the NSL-KDD and UNSW-NB15 datasets. A rule-based model and a genetic search tool were used for the hybrid feature selection; thus, the evaluator subset was used to compute the connection between the class and each feature. The highest correlation from the attribute and class relationship is then chosen for selection. The merits of each attribute were then evaluated; function selection is known as the genetic search method, which produces attributes with the greatest value. If two attribute segments have the same performance score, the rule-based algorithm (rule assessment phase) produces the feature subset with the fewest volume of subset features. Finally, the features that have been selected are loaded into the ANN for template matching and assault selection. The ability of rule-based schemes combined with learning techniques to improve output precision has been demonstrated [29].

This was inspired by the assumption that integrating classifier optimization techniques into the feature representation and driving it with a rule-based algorithm would improve the performance of an IDS. The paper identifies intrusion in the IIoT network using the proposed model. The datasets used has huge features and parameters; an effective feature selection has to be employed to effectively reduce the high dimensionality of the datasets. This was done to reduce the burden this will have on the classifier. Furthermore, a feature extraction technique would make it easier for the classifier to select the most relevant qualities and exclude those that have a detrimental impact on classifier's performance. This motivated the creation of a new model using rule-based feature selection to effectively select the most relevant features from the datasets. The DFFNN classifier is used to train the features selected using this hybrid rule-based feature selection. The suggested model's performance is then assessed using current methodologies.

This paper's contributions are as follows:

(i) A system for intrusion prevention in the Industrial Internet of Things network is suggested

(ii) A hybrid approach focused on hybrid deep learning and rule-based feature selection for in-depth intrusion detection analysis

(iii) A relation of the current approach to prior methods for intrusion detection in the IIoT network is made. The proposed approach is stable, more efficient, and less resource-intensive, according to experimental results

2. Industrial Internet of Things Analytics Overview

Manufacturing, transportation, electricity, and healthcare are all affected by the Industry 4.0 revolution, which necessitates a change in industries that depend heavily on operational technology (OT). Previously, fog and edge computing [30] technologies were needed for Industrial IoT to ensure the required integration across Industry 4.0. However, this uprising introduces a new interrelated aspect that is critical for IIoT Analytics. The DL algorithms improve big data analytics capabilities, while IIoT Technologies enhance the utility of each of these categories. These algorithms can aid in the identification, categorization, and decision-making of each of these data types. The DL in combination with big data technologies generates practical and valuable data for policymaking. DL will be critical in IIoT and data analytics for effective and efficient selection, specifically in the field of streaming data and real-time insights in conjunction with edge computing systems [1].

Several business verticals, such as healthcare, grocery, automobiles, and transportation, are using IIoT applications. In many industries, the IIoT can greatly increase dependability, performance, and service quality. The IIoT will first develop current procedures and facilities, but the eventual aim is to create completely novel and vastly enhanced goods and services. Many companies recognize how and where IoT innovations and solutions can lead to organizational changes, new and improved goods and services, and entirely new business models. On the IIoT, machine learning and deep learning algorithms can increase reliability, production, and customer satisfaction by combining various machinery, procedures, apps, and applications. Anything necessitates a wide range of technology that must be carefully integrated and orchestrated.

These advancements in technology allow intelligent machines, tools, engines, and integrated control systems to execute repetitive duties to solve difficult problems without the need for human involvement [31]. Smart workplace advancements, intelligent data discovery, cognitive automation, and other aspects of business smartness should all be included. A digital twin is a virtual representation of physical assets, systems, and so on. This is generally alluded to as a result of the Internet of Things, which increasingly extends all of these appliances while providing us with a similarly increasing data collection that can be evaluated for

efficiency, architecture, and repair, among other things. Any digital twin's main advantage is that it is continually updated and "learns" any updates that occur in almost real time. The IIoT model and its applications are creating significant disruptions in the market globally.

2.1. The Four Key Components of Industrial IoT Architecture. Intelligent Edge Gateway: An intelligent edge gateway is a computer program closely aligned with sensor nodes that can capture, aggregate, and sanitize light data streaming. It allows one to upload tabulated and relevant data to the Internet of Things network. It acts as a connection between the hardware and the cloud IoT network in general.

IoT Cloud: The main IoT framework that uses data processing, machine knowledge, and artificial intelligence methods to handle massive quantities of data. The processing capabilities including device control, stream analytics, event management, a rules engine, alerts, and updates are all available. It offers components like big data analytics, as well as authorization, virtualization, end-to-end encryption, SDKs, and application APIs.

Business Incorporation and Platform: This is a backend framework that connects many IT schemes to certify that computer data is collected and processed in the full operational loop. ERP, QMS, planning and scheduling, and other systems are examples of such systems. Data analysis can be divided into three groups depending on the form of the result obtained. There are three types of analysis: descriptive, predictive, and prescriptive. Figure 1 displays IIoT architecture with four (4) layers including things, intelligent gateway, IoT cloud, and business application and integrations.

3. Related Work

As Anomaly Detection System (ADS) is an essential security management system that functions as a sniffer and deciding driver for routing traffic and spot suspicious activities [32], it functions as a packet capture and decoding engine for ensuring security and recognizing anomalous behavior. Since it can track both visible and invisible (zero-day) threats, the focus is on creating a pattern from standard data and treating any variance from it as an intrusion [33]. For example, the aim of [34, 35] centered on finding ADS using Particle Swarm Optimization (PSO) techniques for optimizing the performance of the One-Class Support Vector Machine (OCSVM) method by harvesting Modbus/TCP message network streams for testing and verifying the system. In [36], the authors built an IDS/ADS centered on this design, which was learned on offline data from a SCADA setting using network traces.

In [37], the authors constructed an IDS centered on the Modbus/TCP protocol setting using a K-NN classifier. While the aforementioned mechanisms performed admirably in certain cases, they were designed for particular configurations with a strong FPR. Similarly, in [38], the authors proposed an improved intrusion detection system (IDS) for matching the diverse structures of SCADA schemes using diverse OCSVM frameworks to select the right one for efficiently identifying multiple assaults. When operating, nevertheless, this computer used a large amount of computational power and had a high false warning rate for identification. Using SCADA mechanisms to obtain different aspects of contact events and using an SVM algorithm to identify attacks, authors in [39] suggested an ADS for detecting Modbus/TCP protocol-infiltrated assaults. The detection method, on the other hand, was ineffective in detecting irregular behaviors.

To prevent the effects of factors associated with the OCSVM's ability to track network attacks successfully, the authors in [40] merged the OCSVM method, and the recurrent K-means clustering algorithm was used. In another valuable effort, [41] proposed a critical infrastructure intrusion detection system centered on an artificial neural network (ANN) method that trained a multiperceptron ANN to identify anomalous network activity using fault back-propagation and Levenberg-Marquard features. Using a virtual network, in a relevant try, [42] used an ANN to detect DoS/DDoS attacks in IoTs, and in [43], the authors proposed a decentralized IDS based on artificial immunity for IoT devices. In [44], another set of researchers projected a Possibility Risk Identification-centered Intrusion Detection System (PRI-IDS) method for detecting replay attacks by inspecting Modbus TCP/IP protocol network traffic. However, these schemes had a high rate of false alarms and had trouble identifying certain new attacks.

In a related effort, the authors of [45] create a learning firewall that receives tagged samples and automatically configures itself by writing conservative preventive rules to avoid false alerts. We create a novel classifier family called classifiers that, unlike standard classifiers that just focus on accuracy, use zero false positive as the decision-making criterion. The authors first illustrate why naïve modifications of current classifiers, such as SVM, do not produce acceptable results and then present a generic iterative technique to achieve this goal. The proposed classifier, which is based on CART, is used to create a firewall for a Power Grid Monitoring System. We also put the technique to the test on the KDD CUP'99 dataset to see how well it works. The outcomes support the efficacy of our strategy.

IDSs have indeed been analyzed utilizing subsurface networks for identifying irregular findings from host and network-based systems by several researchers [46–48]. An ANN with a shallow network has one or two hidden layers, while a deep network has several hidden states of various architectures [49]. Deep learning is a form of a common machine-learning technique used by academic and industrial researchers because it can learn a detailed computational mechanism that mimics the normal behaviors of the human mind [50].

Several researcher has proved that the swiftness which received system signals is converted into massive datasets which pose a significant obstacle to IDS architectures' ability to analyze the subsequent large amounts of data for actual processing [51–53]. The authors in [54] suggested a new rule-based approach for detecting DoS assaults that relied on domain expert knowledge. For identifying DoS attacks, a rule-based classification algorithm was used, and the final classification was carried out by applying the rules from

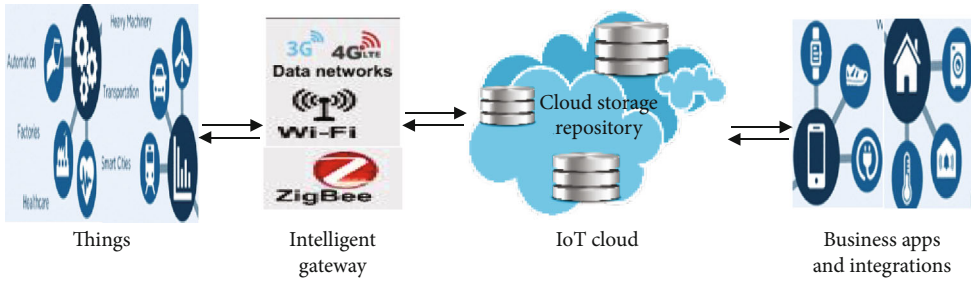

Figure 1: Industrial Internet of Things architecture.

the rule base and was confirmed using a domain expert. Feature selection techniques, also known as spatial removal, can aid in the conversion of databases from an elevated to a lesser spatial domain that better represents the problem space with the same efficacy [55, 56]. Unconnected variables may be eliminated without lowering data's importance to the detection model, which is the foundation of introduced feature collection [57, 58]. Most datasets have several attributes but few examples, according to [59, 60], possibly requiring and using feature selection techniques.

In [61], the authors provide an attack taxonomy based on the several layers of the IoT stack, such as device, infrastructure, communication, and service, as well as the specific characteristics of each layer that can be exploited by adversaries. Furthermore, we explain IoT-related vulnerabilities, exploitation techniques, attacks, impacts, and potential mitigation mechanisms and defense strategies using nine real-world cybersecurity incidents that attacked IoT devices deployed in the consumer, commercial, and industrial sectors. These with various additional examples emphasize the fundamental security vulnerabilities of IoT systems and indicate the possible attack implications of such interconnected ecosystems, while the suggested taxonomy provides a systematic approach for categorizing attacks based on the impacted layer and its impact.

A rule-based classifier-based data reduction strategy has been proposed in [62]. The suggested dimension reduction technique is an innovative data preparation technique that decreases both attributes and occurrences in testing specimens while keeping classifier precision. In [63], the authors suggested a fuzzy-based semisupervised learning method for IDS that constitutes a significant quantity of unlabelled data powered by labeled data to increase classification performance. The authors used the fuzzy measure to produce a trained independent hidden node feedforward neural network used to generate a fuzzy set vector of small, medium, and large specimen classification on unlabelled data. The training set is reused after using each vector of data classification independently in the initial training dataset. In a related work, the authors in [21] suggested a new method wrapper-based NIDS architecture based on Bayesian networks. The feature selection technique is used in this context to extract the appropriate features from the sample so that the Bayesian network classifier can reliably predict attack types.

For intrusion detection, [64] suggested a crossbreed method combining SVM and the ant colony. The aim of integrating the two machine learning techniques is to account for the shortcomings and capabilities of both methods to provide a more precise occurrence grouping. Similarly, in [65], the authors projected a wrapper approach for lightweight malware discovery based on decision trees. The suggested technique has four processes: preprocessing or removal of duplicate attack patterns, feature selection centered on a genetic algorithm (GA), postprocessing for standardized results, and traffic classification techniques centered on a neurotree technique. Similarly, in [66], a wrapping suitability purpose centered on a violation word for a wide amount of attributes with good classification precision and strict enforcement. The suggested wrapping fitness value is effective for feature extraction while maintaining prediction performance, according to the experiments. In [67], the paper suggested a decision tree classifier-based NIDS function collection depending on GA. The researchers used a GA to derive input data for decision trees as a classification algorithm to improve identification and reduce false alarms in cyberthreat detection.

In [53], the authors proposed a smart rule-based identification scheme for detecting Deprivation of Service (DoS) assaults in cloud servers scheme. The study used scoring and rating algorithms to simulate a cloud service, assault identities, and choose the best functionality. To discover assaults, a rule-based grouping procedure grounded on quality expertise was used to the selected features. The key benefit of their proposed model is a lower rate of false alarms and increased protection. But, due to the complex nature of attacks, the risk of confusion was not addressed. A modern feature selection strategy and a more effective KKN classifier were proposed in [68] for intrusion detection. The introduced feature set significantly reduces the existence of irrelevant features, thus improving KNN classifier's classification ability to distinguish kinds of invasion. Furthermore, the suggested feature selection algorithm reduces classifier's error warning rate dramatically. In a related work in [69], the authors suggested a new sophisticated artificial potential field technique for selecting features, as well as the implementation of a phased architecture as a base classifier for assault identification, when the suggested algorithm had a better classification exactness and a low wrong alarm degree as opposed to other approaches.

A machine learning classification algorithm to extract malware photographs with a mix of local and global characteristics was propped in [70, 71]. Their processes had a classification precision of 98.4% on a broad-scale study, using

9339 samples from 25 malware relatives in the Malimg dataset. Their methods achieved 99.21% classification accuracy in small-scale research, with 5288 samples from 8 malware relatives in the Malimg dataset. The authors in [72] created a CNN model that is used to separate threats from a corpus of binary executables. Moreover, this method had a classification accuracy of 98.52% when tested against a dataset of 9339 samples from 25 malware executables. Besides that, this template is used to randomly select 10% of samples in each loop to assess a malware family. In [73], the authors proposed a CNN-based malware classification model. From a dataset of 9339 samples, this model had a 98% accuracy. In each loop, a random method is used to pick 10% of samples to evaluate the malware family in question.

CNN used to create a malware classification model in [74]. The study used a corpus of 9339 samples from 25 diverse malware groups; this method had a 94.5% accuracy rate. In the same vein, in [75], the authors created a deep convolutional neural network that uses color image visualization to discover malware assaults on the Internet. Their findings showed that their classification efficiency for measuring cybersecurity threats had improved. The authors in [76] suggested a system built on Random Coefficient Selection and Mean Adjustment Method (RCSMMA). RCSMMA performs well against a variety of modern cyberattacks. Authors in [77] outlined the most important smart city applications and discussed the major issues of privacy and protection in smart city application architecture as a result of malware attacks. To avoid antagonists in the global sensor network, the authors in [78] proposed a stable steering and watching protocol using multivariant tuples.

To establish a powerful defense system against invaders, authors in [79] recommend developing strong intrusion detection systems that can detect intruders. In this paper, an ensemble classifier based on Crowd-Search is employed to categorize the UNSW-NB15 dataset, which is based on IoT. The most important characteristics from the dataset are first identified using the Crow-Search method and then provided to the ensemble classifier for training using the linear regression, Random Forest, and XGBoost algorithms. The proposed model's performance is then compared to that of state-of-the-art models to ensure that it is effective. The experimental results show that the suggested model outperforms the other models studied.

The widespread use of the internet in all aspects of human existence has raised the possibility of malicious attacks on the network. Intrusion detection systems have emerged as a result of the ease with which activities carried out via the network can spread. The patterns of attacks are also dynamic, necessitating effective cyberattack classification and prediction. To identify intrusion detection system (IDS) datasets, in [80], the authors proposed a hybrid principal component analysis (PCA)-firefly-based machine learning model. The dataset for this study was obtained from Kaggle. For the transformation of the IDS datasets, the model first uses One-Hot encoding. For dimensionality reduction, the hybrid PCA-firefly method is used. For classification, the XGBoost algorithm is used on the reduced dataset. To demonstrate the superiority of our suggested strategy, we undertake a detailed evaluation of the model using state-of-the-art machine learning approaches. The results of the experiments show that the suggested model outperforms the existing machine learning models.

From the existing related work, it can be seen that DL algorithms can considerably be used to increase the efficiency of IDS for IIoT by achieving the highest prediction performance while maintaining a low false alarm rate. Thus, motivating the use of the DL model with a hybrid rule-based technique for the automatic feature selection and sensing anomaly trends in data as suspect vectors using data transmission depth coverage. The proposed DFFNN based with hybrid rule-based feature selection model contains a rule-based model using a genetic search engine to select the relevant features and the DAE-DFFNN algorithm to classify IIoT network by classifying the constraint values of the DAE. It can find a good approximation for communication networks and transform high-dimensional data to low-dimensional data using DAE-DFFNN model's decreased layer, as explained in the following subsections.

4. The Proposed Intrusion Detection for Industrial Internet of Things Network

In this analysis, the deep feedforward neural network (DFFNN) is used to generate an effective ADS for IIoT locations. In the testing stage, a dual feature extraction employs a genetic search system as well as a rule-based algorithm. The subsection assesses or calculates the connection between individual features as well as the category. The class-attribute interaction with the highest similarity is used for filtering. This is referred to as function assessment. The genetic search procedure determines the qualities of each feature based on this function assessment and returns the attributes with the uppermost suitability value. If two attribute subsections have the same performance score, the rule-based algorithm (rule assessment phase) yields the feature vectors with the fewest quantity of subsection attributes. Finally, the chosen attributes are fed into the ANN, which is used to create models and classify attacks. These parameters are used to set up a standard DFFNN for discovering current and new attack instances. The DFFNN is used to detect mischievous vectors during the testing process. By translating the reduced hidden units, various hidden layers in the methodology will properly develop a detailed feature vector and grab the most important features. The subsections go into the specifics of the proposed system methodology.

4.1. Deep Feedforward Neural Network (DFFNN). The fundamental deep learning models are deep feedforward networks, often known as feedforward neural networks or multilayer perceptrons (MLPs). A feedforward network's purpose is to approximate a function f^x. For example, $y = f^x(\mathrm{x})$ transfers an input x to a category y in a classifier. A feedforward network learns the values of the parameters that result in the best function approximation by defining a mapping $y = f(\mathrm{x})$. Because information flows through the function being evaluated from x, the intermediate calculations

necessary to define f, and finally, to the output y, these models are referred to as feedforward models. There are no feedback links; therefore, model's outputs do not feedback into it. Recurrent neural networks are feedforward neural networks that have been extended to incorporate feedback connections. The DFFNN is usually described as an ANN method with input neurons, several hidden nodes, and an output neuron that are all directly connected without the use of a cycle [81].

The secret surface of each node reflects indistinct attributes dependent on the preceding stage's display, which are dynamically computed and processed in multiple layers to produce the outputs. This strategy is trained using a stochastic slope descent back-propagation methodology [82]. You can give a deeper feedforward neural network the ability to capture more complicated representations by creating a deeper feedforward neural network. If the complexity is justifiable, this could be justified. It has the advantage of being able to readily represent more complex functions.

The source data is fed into input nodes before being forwarded on to the hidden units, which generates a nonlinear manipulation of the information before being moved on to the output nodes in this deep-learning technique. To calculate the quality of the result, a feature role or back-propagation fault [83] is calculated, which is the discrepancy between the predicted and real presentation, and its value is transmitted backward across the unknown nodes to change the masses. The loss function is measured utilizing sole or minibatch specimens of the training examples rather than the whole set, with loads calibrated during each test to determine that the model is correctly suited.

This computation training data approach is based on the random chance of neural network variable activation, which results in the template being put in minima solutions with poor normalization [84]. To improve the convergence rate and the results of supervised learning, pretraining unsupervised strategies, specifically an AE, can be used to build the activation specifications [11].

4.2. Deep Autoencoder (DAE). A DAE is a feedforward neural network strategy for fast unsupervised computing execution [85]. It investigates the estimation of a unique task, where the result (x) is equal to the input $(\check{x})$ to construct a definition of a collection of data, that is, $(x \longrightarrow \check{x})$, (x). Its schematic representation consists of vectors $(x^{(\tilde{i})})$ in the input nodes and several concealed units of nonlinear initiation attributes. To learn compact features of the input data, the extracted features employ fewer neurons than the input nodes. As a result, it knows the most significant attributes and lowers spatial size and views the input data as an abstraction. At the end of the method, the output layer $(\check{x}^i)$ is shown as a close depiction of the input layer.

An AE's simplest framework comprises three layers: input, secret, and output. If the training data $(x^{(\tilde{i})})$ has n samples, each $(x^{(\tilde{i})})(i\epsilon(1, \cdots, n))$ has several proportions, as well as a spatial function vector $(d0)$; the Tanhinitiation function [85] is used and calculated using

$$T(t) = \frac{1 - e^{-2t}}{1 + e^{-2t}}. \tag{1}$$

The encoder and decoder are the two key components of the AE algorithm [86, 87]. A deterministic mapping called an encoder method $(f\theta)$ is used [86] to transform the input vector $(x^{(\tilde{i})})$ into a hidden layer representation $(z^{(\tilde{i})})$, and the dimensionality $x^{(\tilde{i})}$ is reduced to provide the right number of codes.

$$f\theta\left(x^{(i)}\right) = T(W_{x^{(i)}} + b), \tag{2}$$

where W is a d^0 x d^h, d^h weight matrix, d^h is the number of neurons in a concealed level $(d^0 < d^h)$, b is the bias vector, T is the Tanhinitiation utility, and θ, $[W, b]$ are the mapping parameters.

The product of the concealed layer's depiction is plotted, and the translator method is calculated by the deterministic plotting $(g\theta')$ as an approximation $(\check{x}^i)$ to restructure the input as an estimate $(\check{x}^i)$.

$$g\theta'\left(x^{(i)}\right) = T\left(W_{z^{(i)}}' + b'\right). \tag{3}$$

W′ is a d^0 x d^h weight matrix, b' is a bias vector, and θ' represents the mapping parameters $[W', b']$.

The information in that compressed representation is then used as inputs to reconstruct the original information after being transformed to fit the secret surface. The reform mistake (i.e., the alteration between the raw document and its low-dimensional reproduction) for a standard or minibatch training set(s) is calculated by the training process.

$$E\left(x, \check{x}\right) = \frac{1}{2}\sum_{i}^{s}\left\|x^{(i)} - \check{x}^{(i)}\right\|^2, \theta = \{W, b\} = \operatorname{argmin}_{\theta} E\left(x, \check{x}\right). \tag{4}$$

Feature selection phase:

Definition 1 (subset). A feature V_i is said to be relevant if there exists some v_i and c for which $p(V_i = v_i) > 0$ such that

$$p(C = c \mid V_i = v_i) \neq p(C - c). \tag{5}$$

Definition 2 (SubsetEval). If the connection between an individual component of when the association between a function and the outside parameter is understood, as well as the intercorrelation across each set of parameters, the connection between a standardized test made up of the combined modules and the outside parameter can be estimated in (6).

$$r_{zc} = \frac{k_{\overline{r_{zi}}}}{\sqrt{k + k(k-l)\overline{r_{ii}}}}, \tag{6}$$

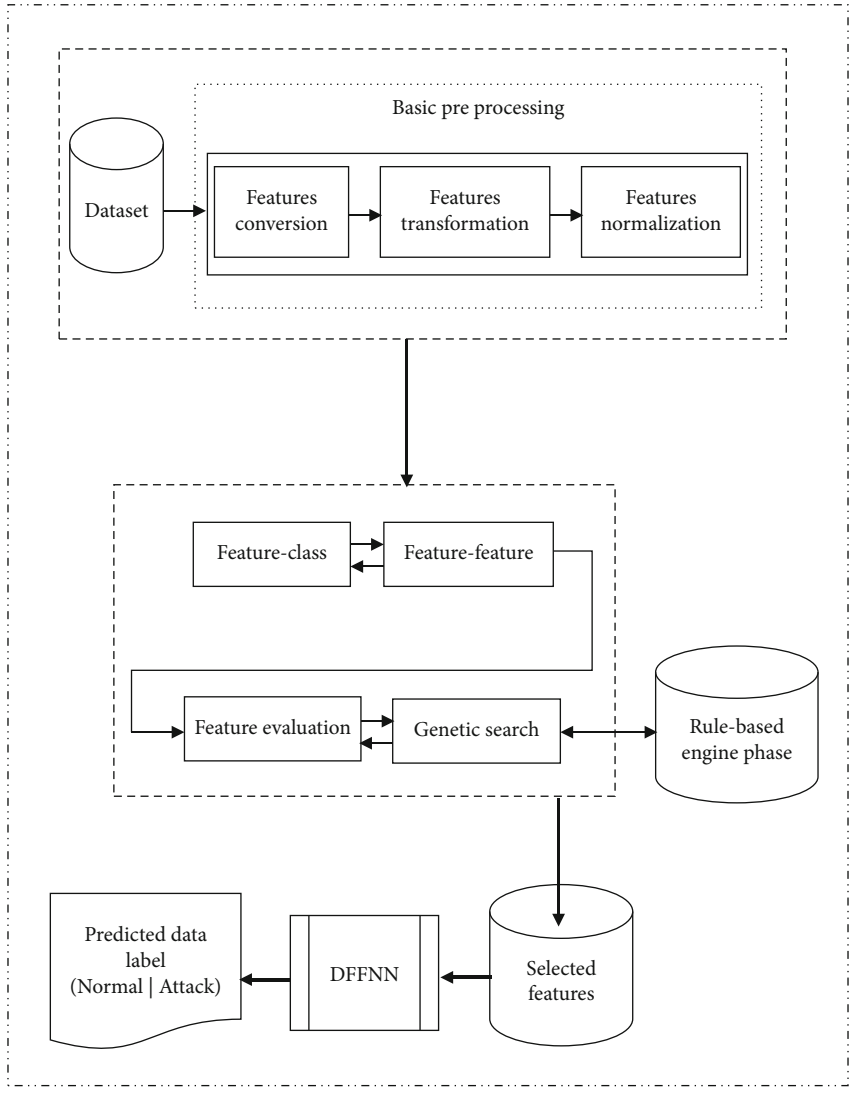

Figure 2: The proposed intrusion detection in industrial internet of things network.

where r_{zc} seems to be the association between the sum of the modules and the external parameter and k is the number of elements, r_{zi} is the average of the component-to-outside-variable correlations, and r_{ii} is the average component-to-outside-variable intercorrelation.

Definition 3 (Genetic search). The term "genetic search" refers to an exploration that is motivated by normal progression. A suitability task that is a lined grouping of an accuracy duration and an effortlessness duration is used in this genetic search.

$$\text{Fitness}(X) = \frac{3}{4}A + \frac{1}{4} = \left(1 - \frac{S+F}{2}\right), \tag{7}$$

where X represents a function subset, A represents DFFNN's average cross-validation precision, S represents the number of instances or training samples, and F represents the number of subset features.

Definition 4 (Rule engine). If there are several feature subsets $(F>)$ with identical fitness values, the rule-based instrument yields a feature subgroup (V_i) with fewer features (X_F), else, it yields the feature subgroup with the uppermost appropriateness value (F_{hi}) to the base classifier as in (8).

$$R = \begin{cases} V_i, & \text{if } V_i \epsilon F > \cap X_f, \\ V_i, & \text{if } F_{hi \cap \varnothing}. \end{cases} \tag{8}$$

This study suggests an effective intrusion discovery model for safeguarding the IIoT system against the mischievous activity. Figure 2 displayed the architecture of the projected model with the training and testing phases.

Figure 2 shows the proposed intrusion detection in IIoT network. In an IIoT setting, the proposed scheme investigates and chooses critical information from large-scale data. The first phase in the suggested method is data preprocessing, which includes function translation and regularization model.

4.2.1. Feature Transformation. Since the suggested framework only embraces mathematical properties, the apiece rhetorical attribute value is transformed into a mathematical formula; for instance, the NSL-KDD dataset contains multiple figurative attributes like procedure natures with reference values like ICMP, TCP, and UDP, which are plotted to 1, 2, and 3, respectively.

4.2.2. Feature Normalization. Since DL relies on different features based on masses. Data could be skewed into various spots due to levels, causing some values to update quicker than others [8, 11]. As a result, it is important to deal with this problem using statistical normalization, in which the Z − score function for each feature value $(v^{(i)})$ is calculated by

$$Z^{(i)} = \frac{v^{(i)} - \mu}{\sigma}, \tag{9}$$

where σ is the standard deviation and is the mean of the μ values for a given function $(v^{(i)}(i\epsilon\, 1, 2, 3, \cdots, n))$.

Since networks have high dimensionality, it is important to minimize it to increase computing resources and develop a compact and flexible ADS strategy [88]. As a result, the suggested DAE-DFFNN method is used to decrease high proportions to low proportions via a main reduced surface. More specifically, the model contains a nonlinear mechanism that encrypts a lot of features into the lesser feature set in the reduced hidden state, requiring dimensionality reduction to be realistic without the necessity for professional acquaintance. The purpose of the rule based with DAE-DFFNN dimension reduction is to identify excellently embodiments from the unclear framework in the probability model in terms of increased learning and also processed and decreased attributes.

4.3. Details of the Datasets Used. The testing, examining, and assessing of the behavior of the discovery scheme depends solely on the dataset, and this plays a vital role in getting a better result. A high-performance one not only yields effective outcomes for an offline device but can be successful in an actual setting. Most authors also used the well-known NSL-KDD datasets, which is a revised variant of the KDD CUP 99 database that solves the KDD CUP 99's main problems by deleting duplicate information and selecting documents concerning their proportions. It comprises 148,517 documents (77,054 standards and 71,460 assaults) after preprocessing, apiece of which includes 41 attributes and a class mark. Probing, DoS, user to root (U2R), remote to local (R2L), and normal are the five classes [89, 90]. However, despite being commonly used in IDSs, it is now obsolete [91]. As a result, a novel dataset called UNSW-NB15 is used to effectively test our proposed work. It includes contemporary synthesized attack activities and represents actual current normal behaviors [92]. It has a total of 257,673 records (93,000 regular and 164,673 attacks), each with 41 features and a classification mark. Fuzzers, examination, backdoors, DoS, vulnerabilities, standard, reconnaissance, shellcode, and worm are all among the ten separate class labels, one standard, and nine attacks.

4.4. Performance Analysis. To assess the performance and comparison of the proposed algorithm using DL and hybrid rule-based model with other existing models, the following performance metrics were used. The amount of correct and incorrect outcomes in a classification problem was summed and compared; the results were with the reference results. Accuracy, precision, recall, specificity, and $F1$-score are just a few of the most common matrices. True-positive (TP), true-negative (TN), false-positive (FP), and false-negative (FN) statistical indices were calculated to solve the confusion matrix, as shown in Equations (10)–(16).

$$\text{Accuracy} : \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}, \tag{10}$$

$$\text{Precision} : \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{11}$$

$$\text{Sensitivity or Recall} : \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{12}$$

$$\text{Specificity} : \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{13}$$

$$\text{F1} - \text{score} : \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}, \tag{14}$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{15}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \tag{16}$$

From Equations (10) to (16), accuracy denotes how often the prediction is correct, whereas precision denotes how often the class will be correct during prediction. However, recall indicates how much of the all-positive class was correctly predicted, whereas specificity assesses how well the negatives were identified. The $F1$-score is a combination of exactness and recall. The quantity of correct negative estimates distributed by the total quantity of negative forecasts is known as specificity. The true-positive rate (TPR) is defined as the proportion of properly recognized attacks over the total quantity of dataset classes, as seen in Equation (15). The TPR stands for discovery rate. The false alarm rate (FAR) is calculated by dividing the number of records wrongly denied by the total number of normal records. Equation (16) defines the FAR evaluation metric. As a result, in the IIoT system, the impetus for intrusion detection prediction is to achieve a higher accuracy and detection rate (DR) with a lower false alarm rate.

5. Results and Discussion

The R programming language platforms were used to implement the proposed model, and the evaluation was done using the explained performance metrics. Both datasets with the relevant DAE-DFFNN with dual rule-based design are used to seamlessly incorporate all characteristics. The NSL-KDD dataset contains 77,054 regular documents and 71,460 assault documents, as well as different samples from the UNSW-NB15 dataset, which contains

Table 1: Proposed method evaluation.

Dataset	TP rate	FP rate	Precision	Recall	*F*-measure	ROC	Class
UNSW-NB15	0.998	0.1	0.967	0.998	0.989	0.989	Attack
	0.999	0.001	0.998	0.996	0.967	0.998	Normal
NSL-KDD	0.996	0.1	0.984	0.999	0.997	0.997	Attack
	0.999	0.001	0.998	0.993	0.969	0.998	Normal

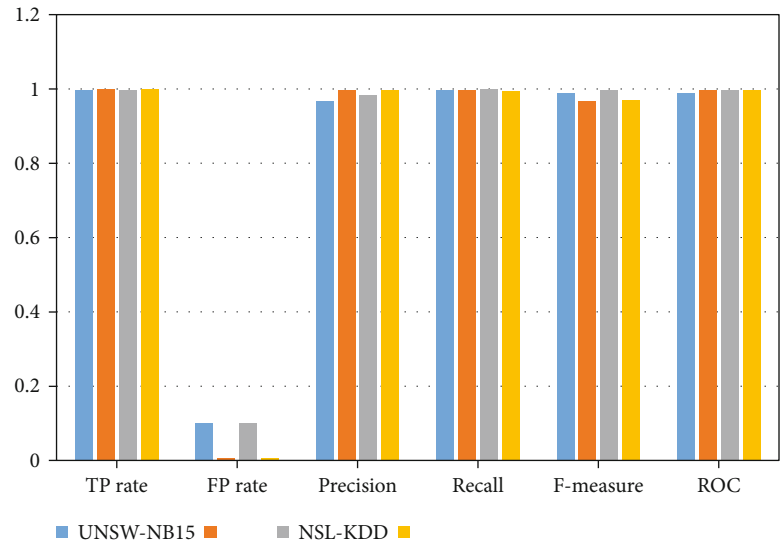

Figure 3: The performance evaluation of the proposed model on both datasets.

93,000 regular documents and 92,000 assault documents, with 20 percent of the normal records represents 40%, 20%, and 60% of the testers which were used for testing, respectively.

The network structures and parameters adopted based on the experiments yield the peak DR and lowermost FPR. The proposed model used the best network structures for both datasets after the best features are selected using the hybrid rule-based genetic search engine in addition to the DAE feature selection model are 41 nodes for one input layer, 10, 3, and 10 nodes for the three hidden layers and 41 nodes for the output layer for the DAE technique, and 2 nodes for the DFFNN model for the output layer with 2 nodes. For the NSL-KDD dataset, 0.0015 is the learning rate and 0.2 momenta start, and for UNSW-NB15 dataset, L1 and L2 regularizations of L1 = L2 = $1e-6$, momentum start of 0.2, momentum stable of 0.4, $1e7$ ramp momentum, annealing rate of $2e$-6, and 100 epochs; 0.002 learning rate for the Tanh activation function was used.

Table 1 and Figure 3 show the performance of the projected model using both NSL-KDD and UNSW-NB15 datasets using numerous metrics. The results obtained using various metrics show that the projected model is very important and relevant in intrusion detection of IIoT network for attack prediction and classification.

Table 2 displays the accuracy, detection rate, and FPR of the proposed model on the datasets. The findings reveal that the model outperforms the UNSW-NB15 dataset on NSL-KDD, with a precision of 99.0 percent, a detection rate of 99.0 percent, and an FPR of 1.0 percent.

Table 2: Evaluation of performances for two datasets.

Dataset	Accuracy	Detection rate	FPR
NSL-KDD	99.0%	99.0%	1.0%
UNSW-NB15	98.9%	99.9%	1.1%

Table 3 shows the discovery rates for the classes in the NSL-KDD and UNSW-NB15 datasets using the projected model. The outcomes for the UNSW-NB15 dataset are displayed in Table 3 and Figure 3 for the discovery rates of the record types classes: analysis (92.3%), backdoor (95.2%), DoS (97.3%), exploits (98.0%), fuzzer (67.1%), generic (99.8%), normal (99.6%), salicode (90.3%), worm (81.7%), reconnaissance (92%), and shellcode (90.2%), respectively. The results for the NSL-KDD dataset using the projected model are displayed in Table 3 and Figures 4 and 5 to determine the records types like DoS, normal, U2R, R2L, and probe with discovery rates of 99.2%, 99.7%, 75.5%, 94.3%, and 99.0%, respectively. The proposed model demonstrated overall better performance for intrusion detection in both used datasets even though some results like U2R, fuzzer, and worms are not too high in both datasets.

5.1. The Comparison of the Proposed Model with Existing Methods. To show how feature selection affects classification algorithm's detection efficiency, Table 4 compares the

Table 3: Detection rates for NSL-KDD and UNSW-NB15 dataset classes.

Dataset	DoS	Normal	Backdoor	Worm	Shellcode	Probe	Exploits	U2R	Salicode	Analysis	R^2	Fuzzer	Generic	Reconnaissance
NSL-KDD	99.2%	99.7%	—	—	—	99.0%	—	75.5%	—	—	94.8%	—	—	—
UNSW-NB15	97.3%	99.6%	95.2%	81.7%	90.2%	—	98.0%	—	90.3%	92.3%		67.1%	99.8%	92.0%

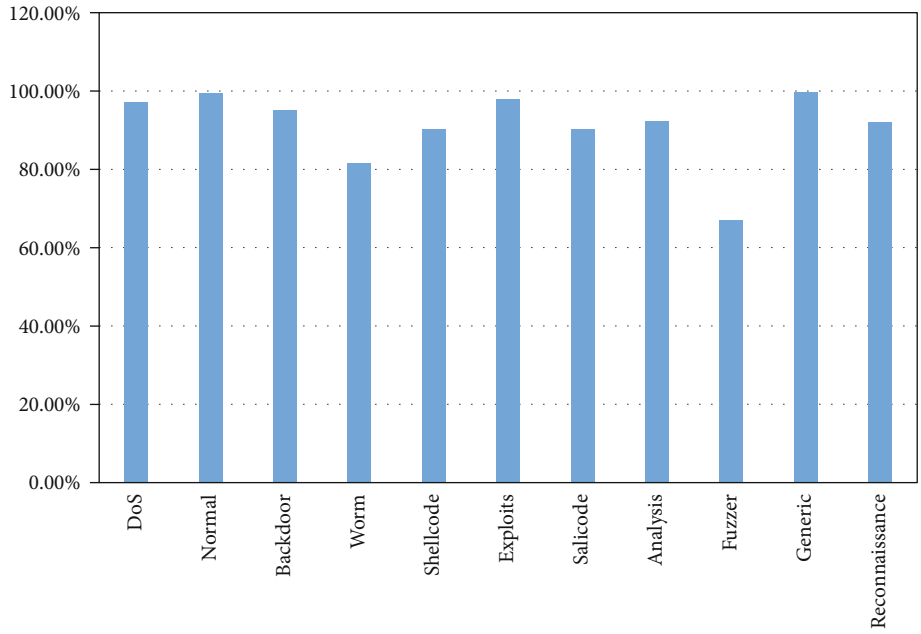

Figure 4: Detection rates for UNSW-NB15 dataset classes.

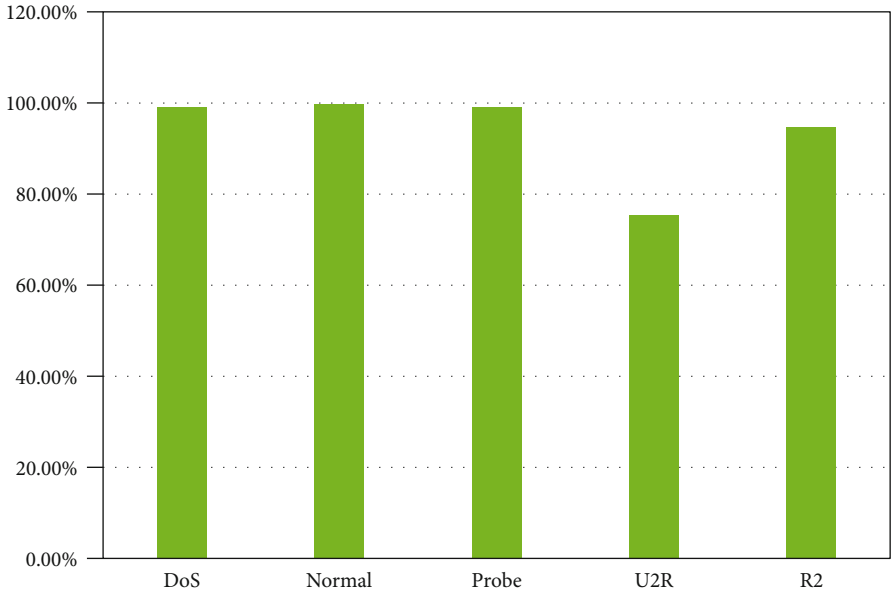

Figure 5: Detection rates for NSL-KDD dataset classes.

proposed approach to several known approaches. Table 4 displays the cumulative performance measures for the proposed system and other models using the decreased UNSW-NB15 dataset. The precision and FPR of the suggested approach are better than those of other approaches. The suggested network intrusion detection method, in general, has a 98.9 percent accuracy, which is 0.1 percent higher than the modified KNN with the second-highest accuracy. Similarly, as compared to other classifiers, the proposed method's FPR has a very low error percentage of 1.1 percent. When equated to other techniques using the reduced UNSW-NB15 dataset, the proposed approach performed better across all evaluation metrics. The proposed method's marginally higher accuracy is due to its robust feature selection and rule-based fitness assessment.

The suggested model's efficiency is contrasted to that of nine recently developed anomaly detection techniques, including the ADS system based on DL, the Filter-based Support Vector Machine (F-SVM) [95], the Computer Vision Method (CVT) [96], the Dirichlet Mixture Model (DMM) [91], the Triangular Area Nearest Neighbors (TANN) [97], DBN [98], RNN [52], DNN [81], and Ensemble-DNN [99]. Table 5 compares the identification rate and false-positive rate of our proposed system to other

Table 4: Summary of performance comparison for UNSW-NB15 dataset.

Model	Performance metrics					
	Accuracy (%)	FPR (%)	*F*-score (%)	Recall (%)	Precision (%)	ROC curve (%)
Wrapper + neurotree [67]	98.38	1.62	0.984	0.980	0.989	0.998
SVM+EML+K-means [58]	95.75	1.87	0.944	0.997	0.897	0.986
GA +SVM [93]	97.3	0.017	0.966	0.997	0.938	0.981
CNN+LSTM [94]	94.12	—	0.956	0.989	0.925	0.984
Modified KNN [70]	98.7	1.3	0.992	0.996	0.988	0.998
CfsSubsetEval + GA+RuleEval+ANN [8]	98.8	1.2	0.989	0.989	0.989	0.998
Proposed model	98.9	1.1	0.989	0.998	0.967	0.989

Table 5: For the NSL-KDD dataset, the results of the proposed model have been compared with nine other classifiers.

Technique	Detection rate	FPR
F-SVM [95]	92.2%	8.7%
CVT [96]	95.3%	5.6%
DMM [91]	97.2%	2.4%
TANN [97]	91.1%	9.4%
DBN [98]	95.1%	4.5%
RNN [52]	73%	3.6%
DNN [81]	76%	15%
Ensemble-DNN [99]	98%	14.7%
ADS-DL [11]	99%	1.8%
Proposed model	98.9	1.1%

models tested on the NSL-KDD dataset. Our developed scheme delivers the desired performance, with 99 percent DR and 1.8 percent FPR. The first four models demonstrated rational results in identifying destructive events after a feature selection process. F-SVM used shared information to solve linear and nonlinear data properties, which was then paired with the SVM for attack detection. Nevertheless, to improve IDS efficiency, this model's search strategy must be refined. CVT and TANN used the PCA technique to reduce the data measurements.

The F-SVM has a detection rate of 92.2% and FPR of 8.7%, CVT with a detection rate of 95.3% and FPR of 5.6%, DMM with a detection rate of 97.2% and FPR of 2.4%, TANN with a detection rate of 91.1% and FPR of 9.4%, DBN with a detection rate of 95.1% and FPR of 4.5%, RNN with a detection rate of 73.0% and FPR of 3.6%, DNN with a detection rate of 76.0% and FPR of 15%, ensemble-DNN with a detection rate of 98.0% and FPR of 14.7%, and ADS with detection rate of 99.0% and FPR of 1.8%. The proposed model differs from previous DL-based IDSs in that it uses a basic mathematical algorithm (DAE) and a hybrid rule-based function selection to estimate parameters that are appropriate DFFNN input to create its classification effectively and efficiently. Moreover, the model knows and examines high-level functionality, automatically decreases data dimensionality, and effectively portrays important features due to the reduced hidden layer. As a consequence, the proposed model is optimal for use in a real-world industrial environment with a vast amount of unlabeled and unstructured data, such as IIoT.

6. Conclusion

This paper proposes an ADS model for identifying destructive activities in IIoT networks utilizing data from TCP/IP packets. It employs unsupervised DL strategies that are hybrid rule-based with automated dimensionality reductions to provide a good description of standard network structures for unsupervised learning. The suggested DAE-DFFNN with hybrid rule-based design is successfully used to develop and remove essential features that improve its overall efficiency. As compared to other strategies developed in recent research, the proposed model achieves the maximum identification rate of 99.0 percent and the fewest false alarms of 1.0 percent when checked on different data samples from the NSL-KDD and NSW-NB15 datasets. Both NSL-KDD and NSW-NB15 were included in the proposed model since they are often used by researchers in intrusion detection and as a benchmark. The use of hybrid rule-based feature collection improves the consistency of the proposed model by using only appropriate features for class classification in the datasets. The future analysis would consider the use of real-world data gathered by the IIoT system to determine the effectiveness of its operation in these settings. In addition, in future work, the proposed model will be extended to accommodate different protocols.

Data Availability

No data available.

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] P. Ambika, "Machine learning and deep learning algorithms on the Industrial Internet of Things (IIoT)," *Advances in Computers*, vol. 117, no. 1, pp. 321–338, 2020.

[2] R. Ashima, A. Haleem, S. Bahl, M. Javaid, S. K. Mahla, and S. Singh, "Automation and manufacturing of smart materials in Additive Manufacturing technologies using the Internet of Things towards the adoption of Industry 4.0," *Materials Today: Proceedings*, vol. 45, pp. 5081–5088, 2021.

[3] L. M. Gladence, V. M. Anu, R. Rathna, and E. Brumancia, "Recommender system for home automation using IoT and artificial intelligence," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–9, 2020.

[4] T. Sherasiya, H. Upadhyay, and H. B. Patel, "A survey: intrusion detection system for internet of things," *International Journal of Computer Science and Engineering (IJCSE)*, vol. 5, no. 2, pp. 91–98, 2016.

[5] J. B. Awotunde, R. G. Jimoh, S. O. Folorunso, E. A. Adeniyi, K. M. Abiodun, and O. O. Banjo, "Privacy and security concerns in IoT-based healthcare systems," *Internet of Things*, pp. 105–134, 2021.

[6] E. A. Adeniyi, R. O. Ogundokun, and J. B. Awotunde, "IoMT-based wearable body sensors network healthcare monitoring system," in *IoT in Healthcare and Ambient Assisted Living*, pp. 103–121, Springer, Singapore, 2021.

[7] K. Amit and C. Chinmay, "Artificial intelligence and Internet of Things based healthcare 4.0 monitoring system," *Wireless Personal Communications*, pp. 1–14, 2021.

[8] F. E. Ayo, S. O. Folorunso, A. A. Abayomi-Alli, A. O. Adekunle, and J. B. Awotunde, "Network intrusion detection based on deep learning model optimized with rule-based hybrid feature selection," *Information Security Journal: A Global Perspective*, vol. 29, no. 6, pp. 267–283, 2020.

[9] M. Abdulraheem, J. B. Awotunde, R. G. Jimoh, and I. D. Oladipo, "An efficient lightweight cryptographic algorithm for IoT security," in *Communications in Computer and Information Science*, pp. 444–456, Springer, 2021.

[10] A. Bakhtawar, R. J. Abdul, C. Chinmay, N. Jamel, R. Saira, and R. Muhammad, "Blockchain and ANFIS empowered IoMT application for privacy preserved contact tracing in COVID-19 pandemic," *Personal and Ubiquitous Computing*, 2021.

[11] A. H. Muna, N. Moustafa, and E. Sitnikova, "Identification of malicious activities in industrial internet of things based on deep learning models," *Journal of information security and applications*, vol. 41, pp. 1–11, 2018.

[12] E. Sitnikova, E. Foo, and R. B. Vaughn, "The power of hands-on exercises in SCADA cybersecurity education," in *Information Assurance and Security Education and Training*, pp. 83–94, Springer, Berlin, Heidelberg, 2013.

[13] S. Dash, C. Chakraborty, S. K. Giri, S. K. Pani, and J. Frnda, "BIFM: big-data driven intelligent forecasting model for COVID-19," *IEEE Access*, vol. 9, pp. 97505–97517, 2021.

[14] G. Tzokatziou, L. A. Maglaras, H. Janicke, and Y. He, "Exploiting SCADA vulnerabilities using a human interface device," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 7, pp. 234–241, 2015.

[15] D. Kushner, "The real story of stuxnet," *IEEE Spectrum*, vol. 50, no. 3, pp. 48–53, 2013.

[16] P. W. Khan and Y. Byun, "A blockchain-based secure image encryption scheme for the industrial Internet of Things," *Entropy*, vol. 22, no. 2, p. 175, 2020.

[17] Q. Yan and F. R. Yu, "Distributed denial of service attacks in software-defined networking with cloud computing," *IEEE Communications Magazine*, vol. 53, no. 4, pp. 52–59, 2015.

[18] A. C. Enache and V. Sgârciu, "Anomaly intrusions detection based on support vector machines with an improved bat algorithm," in *2015 20th International Conference on Control Systems and Computer Science*, pp. 317–321, Bucharest, Romania, May 2015.

[19] O. Folorunso, F. E. Ayo, and Y. E. Babalola, "Ca-NIDS: a network intrusion detection system using combinatorial algorithm approach," *Journal of Information Privacy and Security*, vol. 12, no. 4, pp. 181–196, 2016.

[20] H. Zhang, D. D. Yao, N. Ramakrishnan, and Z. Zhang, "Causality reasoning about network events for detecting stealthy malware activities," *Computers & Security*, vol. 58, pp. 180–198, 2016.

[21] M. R. Kabir, A. R. Onik, and T. Samad, "A network intrusion detection framework based on Bayesian network using a wrapper approach," *International Journal of Computer Applications*, vol. 166, no. 4, pp. 13–17, 2017.

[22] Y. Hu, A. Yang, H. Li, Y. Sun, and L. Sun, "A survey of intrusion detection on industrial control systems," *International Journal of Distributed Sensor Networks*, vol. 14, no. 8, 2018.

[23] T. Cruz, L. Rosa, J. Proenca et al., "A cybersecurity detection framework for supervisory control and data acquisition systems," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2236–2246, 2016.

[24] J. Camacho, A. Pérez-Villegas, P. García-Teodoro, and G. Maciá-Fernández, "PCA-based multivariate statistical network monitoring for anomaly detection," *Computers & Security*, vol. 59, pp. 118–137, 2016.

[25] M. Grill, T. Pevný, and M. Rehak, "Reducing false positives of network anomaly detection by local adaptive multivariate smoothing," *Journal of Computer and System Sciences*, vol. 83, no. 1, pp. 43–57, 2017.

[26] L. A. Maglaras, J. Jiang, and T. J. Cruz, "Combining ensemble methods and social network metrics for improving accuracy of OCSVM on intrusion detection in SCADA systems," *Journal of Information Security and Applications*, vol. 30, pp. 15–26, 2016.

[27] R. O. Ogundokun, J. B. Awotunde, E. A. Adeniyi, and F. E. Ayo, "Crypto-Stegno based model for securing medical information on IOMT platform," *Multimedia tools and applications*, pp. 1–23, 2021.

[28] J. Soto and M. Nogueira, "A framework for resilient and secure spectrum sensing on cognitive radio networks," *Computer Networks*, vol. 115, pp. 130–138, 2017.

[29] M. S. Abadeh, J. Habibi, and C. Lucas, "Intrusion detection using a fuzzy genetics-based learning algorithm," *Journal of Network and Computer Applications*, vol. 30, no. 1, pp. 414–428, 2007.

[30] M. Aazam and E. N. Huh, "Fog computing microdata center-based dynamic resource estimation and pricing model for IoT," in *2015 IEEE 29th International Conference on Advanced Information Networking and Applications*, pp. 687–694, Gwangju, Korea, March 2015.

[31] C. Cecchinel, M. Jimenez, S. Mosser, and M. Riveill, "An architecture to support the collection of big data in the internet of

things," in *2014 IEEE World Congress on Services*, pp. 442–449, Anchorage, AK, USA, June 2014.

[32] N. Moustafa, J. Hu, and J. Slay, "A holistic review of network anomaly detection systems: a comprehensive survey," *Journal of Network and Computer Applications*, vol. 128, pp. 33–55, 2019.

[33] R. Gupta, S. Tanwar, S. Tyagi, and N. Kumar, "Machine learning models for secure data analytics: a taxonomy and threat model," *Computer Communications*, vol. 153, pp. 406–440, 2020.

[34] N. Moustafa and J. Slay, "The evaluation of Network Anomaly Detection Systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Information Security Journal: A Global Perspective*, vol. 25, no. 1-3, pp. 18–31, 2016.

[35] W. Shang, P. Zeng, M. Wan, L. Li, and P. An, "Intrusion detection algorithm based on OCSVM in industrial control system," *Security and Communication Networks*, vol. 9, no. 10, p. 1049, 2016.

[36] L. A. Maglaras and J. Jiang, "Intrusion detection in SCADA systems using machine learning techniques," in *2014 Science and Information Conference*, pp. 626–631, London, UK, August 2014.

[37] P. Silva and M. Schukat, "On the use of k-nn in intrusion detection for industrial control systems," in *Proceedings of The IT&T 13th International Conference on Information Technology and Telecommunication*, pp. 103–106, Dublin, Ireland, August 2014.

[38] B. Stewart, L. Rosa, L. A. Maglaras et al., "A novel intrusion detection mechanism for scada systems which automatically adapts to network topology changes," *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, vol. 4, no. 10, 2017.

[39] W. Shang, J. Cui, M. Wan, P. An, and P. Zeng, "Modbus communication behavior modeling and SVM intrusion detection method," in *Proceedings of the 6th International Conference on Communication and Network Security*pp. 80–85, Singapore, November 2016.

[40] L. A. Maglaras and J. Jiang, "Ocsvm model combined with k-means recursive clustering for intrusion detection in scada systems," in *10th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, pp. 133-134, Rhodes, Greece, August 2014.

[41] O. Linda, T. Vollmer, and M. Manic, "Neural network-based intrusion detection system for critical infrastructures," in *2009 International Joint Conference on Neural Networks*, pp. 1827–1834, Atlanta, GA, USA, June 2009.

[42] E. Hodo, X. Bellekens, A. Hamilton et al., "Threat analysis of IoT networks using artificial neural network intrusion detection system," in *2016 International Symposium on Networks, Computers and Communications (ISNCC)*, pp. 1–6, Yasmine Hammamet, Tunisia, May 2016.

[43] R. Chen, C. M. Liu, and C. Chen, "An artificial immune-based distributed intrusion detection model for the internet of things," in *Advanced materials research*, pp. 165–168, Trans Tech Publications Ltd, 2012.

[44] T. Marsden, N. Moustafa, E. Sitnikova, and G. Creech, "Probability risk identification based intrusion detection system for SCADA systems," in *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, pp. 353–363, Springer, Cham, Switzerland, 2017.

[45] M. S. Haghighi, F. Farivar, and A. Jolfaei, "A machine learning-based approach to build zero false-positive IPSs for industrial IoT and CPS with a case study on power grids security," *IEEE Transactions on Industry Applications*, pp. 1–9, 2020.

[46] N. Gao, L. Gao, Q. Gao, and H. Wang, "An intrusion detection model based on deep belief networks," in *2014 Second International Conference on Advanced Cloud and Big Data*, pp. 247–252, Huangshan, China, November 2014.

[47] B. Abolhasanzadeh, "Nonlinear dimensionality reduction for intrusion detection using auto-encoder bottleneck features," in *2015 7th Conference on Information and Knowledge Technology (IKT)*, pp. 1–5, Urmia, Iran, May 2015.

[48] Y. Li, R. Ma, and R. Jiao, "A hybrid malicious code detection method based on deep learning," *International Journal of Security and Its Applications*, vol. 9, no. 5, pp. 205–216, 2015.

[49] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *IEEE Access*, vol. 5, pp. 21954–21961, 2017.

[50] J. C. van Dijk and P. Williams, "The history of artificial intelligence," in *Expert Systems in Auditing*, pp. 21–26, Palgrave Macmillan, London, 1990.

[51] A. Karkouch, H. Mousannif, H. Al Moatassime, and T. Noel, "Data quality in internet of things: a state-of-the-art survey," *Journal of Network and Computer Applications*, vol. 73, pp. 57–81, 2016.

[52] Y. Qin, Q. Z. Sheng, N. J. Falkner, S. Dustdar, H. Wang, and A. V. Vasilakos, "When things matter: a survey on data-centric Internet of Things," *Journal of Network and Computer Applications*, vol. 64, pp. 137–153, 2016.

[53] F. Zafar, A. Khan, S. U. R. Malik et al., "A survey of cloud computing data integrity schemes: design challenges, taxonomy and future trends," *Computers & Security*, vol. 65, pp. 29–49, 2017.

[54] R. Rajendran, S. V. N. Santhosh Kumar, Y. Palanichamy, and K. Arputharaj, "Detection of DoS attacks in cloud networks using intelligent rule based classification system," *Cluster Computing*, vol. 22, no. S1, pp. 423–434, 2019.

[55] W. L. Al-Yaseen, Z. A. Othman, and M. Z. A. Nazri, "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system," *Expert Systems with Applications*, vol. 67, pp. 296–303, 2017.

[56] L. C. Leonard, "Web-based behavioral modeling for continuous user authentication (CUA)," in *Advances in Computers*, pp. 1–44, Elsevier, 2017.

[57] C. Sammut and G. I. Webb, "Feature selection," in *In Encyclopedia of Machine Learning Edited by Claude Sammut, Geoffrey I. Webb*, pp. 429–433, Springer, 2010.

[58] S. Goswami, A. K. Das, A. Chakrabarti, and B. Chakraborty, "A feature cluster taxonomy based feature selection technique," *Expert Systems with Applications*, vol. 79, pp. 76–89, 2017.

[59] D. Acarali, M. Rajarajan, N. Komninos, and I. Herwono, "Survey of approaches and features for the identification of HTTP-based botnet traffic," *Journal of network and computer applications*, vol. 76, pp. 1–15, 2016.

[60] V. Snášel, J. Nowaková, F. Xhafa, and L. Barolli, "Geometrical and topological approaches to Big Data," *Future Generation Computer Systems*, vol. 67, pp. 286–296, 2017.

[61] C. Xenofontos, I. Zografopoulos, C. Konstantinou, A. Jolfaei, M. K. Khan, and K. K. R. Choo, "Consumer, commercial and

industrial IoT (in) security: attack taxonomy and case studies," *IEEE Internet of Things Journal*, 2021.

[62] V. Herrera-Semenets, O. Andrés Pérez-García, R. Hernández-León, J. van den Berg, and C. Doerr, "A data reduction strategy and its application on scan and backscatter detection using rule-based classifiers," *Expert Systems with Applications*, vol. 95, pp. 272–279, 2018.

[63] R. A. R. Ashfaq, X. Z. Wang, J. Z. Huang, H. Abbas, and Y. L. He, "Fuzziness based semi-supervised learning approach for intrusion detection system," *Information Sciences*, vol. 378, pp. 484–497, 2017.

[64] W. K. Kirnapure and A. R. B. Patil, "Classification, detection and prevention of network attacks using rule based approach," *International Research Journal of Engineering and Technology*, vol. 4, no. 4, pp. 1453–1459, 2017.

[65] S. S. Sivatha Sindhu, S. Geetha, and A. Kannan, "Decision tree based light weight intrusion detection using a wrapper approach," *Expert Systems with Applications*, vol. 39, no. 1, pp. 129–141, 2012.

[66] B. Chakraborty and A. Kawamura, "A new penalty-based wrapper fitness function for feature subset selection with evolutionary algorithms," *Journal of Information and Telecommunication*, vol. 2, no. 2, pp. 1–18, 2018.

[67] G. Stein, B. Chen, A. S. Wu, and K. A. Hua, "Decision tree classifier for network intrusion detection with GA-based feature selection," in *Proceedings of the 43rd annual southeast regional conference on - ACM-SE 43*, pp. 136–141, Kennesaw, Georgia, 2005.

[68] B. Senthilnayaki, K. Venkatalakshmi, and A. Kannan, "Intrusion detection system using fuzzy rough set feature selection and modified KNN classifier," *International Arab Journal of Information Technology*, vol. 16, no. 4, pp. 746–753, 2019.

[69] S. Ganapathy and A. Kannan, "Energy-aware cluster and neuro-fuzzy based routing algorithm for wireless sensor networks in IoT," *Computer Networks*, vol. 151, pp. 211–223, 2019.

[70] H. Naeem, B. Guo, M. R. Naeem, F. Ullah, H. Aldabbas, and M. S. Javed, "Identification of malicious code variants based on image visualization," *Computers & Electrical Engineering*, vol. 76, pp. 225–237, 2019.

[71] H. Naeem, B. Guo, F. Ullah, and M. R. Naeem, "A cross-platform malware variant classification based on image representation," *KSII Transactions on Internet & Information Systems*, vol. 13, no. 7, 2019.

[72] M. Kalash, M. Rochan, N. Mohammed, N. D. Bruce, Y. Wang, and F. Iqbal, "Malware classification with deep convolutional neural networks," in *2018 9th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, pp. 1–5, Paris, France, February 2018.

[73] R. Kumar, Z. Xiaosong, R. U. Khan, I. Ahad, and J. Kumar, "Malicious code detection based on image processing using deep learning," in *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence - ICCAI 2018*, pp. 81–85, Chengdu, China, March 2018.

[74] Z. Cui, F. Xue, X. Cai, Y. Cao, G. G. Wang, and J. Chen, "Detection of malicious code variants based on deep learning," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3187–3196, 2018.

[75] F. Ullah, H. Naeem, S. Jabbar et al., "Cyber security threats detection in internet of things using deep learning approach," *IEEE Access*, vol. 7, pp. 124379–124389, 2019.

[76] N. N. Hurrah, S. A. Parah, J. A. Sheikh, F. Al-Turjman, and K. Muhammad, "Secure data transmission framework for confidentiality in IoTs," *Ad Hoc Networks*, vol. 95, p. 101989, 2019.

[77] F. Al-Turjman, H. Zahmatkesh, and R. Shahroze, "An overview of security and privacy in smart cities' IoT communications," *Transactions on Emerging Telecommunications Technologies*, pp. 1–19, article e3677, 2019.

[78] B. D. Deebak and F. Al-Turjman, "A hybrid secure routing and monitoring mechanism in IoT-based wireless sensor networks," *Ad Hoc Networks*, vol. 97, article 102022, 2020.

[79] G. Srivastava, G. Thippa Reddy, N. Deepa, B. Prabadevi, and M. Praveen Kumar Reddy, "An ensemble model for intrusion detection on the Internet of Softwarized Things," in *Adjunct Proceedings of the 2021 International Conference on Distributed Computing and Networking*, pp. 25–30, Nara, Japan, January 2021.

[80] S. Bhattacharya, S. R. K. S, P. K. R. Maddikunta et al., "A novel PCA-firefly based XGBoost classification model for intrusion detection in networks using GPU," *Electronics*, vol. 9, no. 2, p. 219, 2020.

[81] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho, "Deep learning approach for network intrusion detection in software-defined networking," in *2016 International Conference on Wireless Networks and Mobile Communications (WINCOM)*, pp. 258–263, Fez, Morocco, 2016, October.

[82] M. Li, T. Zhang, Y. Chen, and A. J. Smola, "Efficient mini-batch training for stochastic optimization," *In proceedings of the 20th ACM SIGKDD international conference on knowledge discovery and data mining*, 2014, pp. 661–670, New York, USA, 2014.

[83] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemometrics and Intelligent Laboratory Systems*, vol. 39, no. 1, pp. 43–62, 1997.

[84] D. Erhan, A. Courville, Y. Bengio, and P. Vincent, "Why does unsupervised pre-training help deep learning?," in *In proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 201–208, Sardinia, Italy, March 2010.

[85] X. Tao, D. Kong, Y. Wei, and Y. Wang, "A big network traffic data fusion approach based on fisher and deep auto-encoder," *Information*, vol. 7, no. 2, p. 20, 2016.

[86] W. Hardy, L. Chen, S. Hou, Y. Ye, and X. Li, "DL4MD: a deep learning framework for intelligent malware detection," in *Proceedings of the International Conference on Data Science (ICDATA). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).*, p. 61, Las Vegas, USA, 2016.

[87] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2014.

[88] M. Yousefi-Azar, V. Varadharajan, L. Hamey, and U. Tupakula, "Autoencoder-based feature learning for cyber-security applications," in *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 3854–3861, Anchorage, AK, USA, May 2017.

[89] S. Rathore, A. Saxena, and M. Manoria, "Intrusion detection system on KDDCup99 dataset: a survey," *International Journal of Computer Science and Information Technologies*, vol. 6, no. 4, 2015.

[90] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pp. 1–6, Ottawa, ON, Canada, July 2009.

[91] N. Moustafa, G. Creech, and J. Slay, "Big data analytics for intrusion detection system: statistical decision-making using finite Dirichlet mixture models," in *Data Analytics and Decision Support for Cybersecurity*, pp. 127–156, Springer, Cham, Switzerland, 2017.

[92] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *2015 Military Communications and Information Systems Conference (MilCIS)*, pp. 1–6, Canberra, ACT, Australia, November 2015.

[93] B. M. Aslahi-Shahri, R. Rahmani, M. Chizari et al., "A hybrid method consisting of GA and SVM for intrusion detection system," *Neural Computing and Applications*, vol. 27, no. 6, pp. 1669–1676, 2016.

[94] C. M. Hsu, H. Y. Hsieh, S. W. Prakosa, M. Z. Azhari, and J. S. Leu, "Using long-short-term memory-based convolutional neural networks for network intrusion detection," in *International Wireless Internet Conference*, pp. 86–94, Taipei, Taiwan, 2018.

[95] M. A. Ambusaidi, X. He, P. Nanda, and Z. Tan, "Building an intrusion detection system using a filter-based feature selection algorithm," *IEEE Transactions on Computers*, vol. 65, no. 10, pp. 2986–2998, 2016.

[96] Z. Tan, A. Jamdagni, X. He, P. Nanda, R. P. Liu, and J. Hu, "Detection of denial-of-service attacks based on computer vision techniques," *IEEE Transactions on Computers*, vol. 64, no. 9, pp. 2519–2533, 2014.

[97] C. F. Tsai and C. Y. Lin, "A triangle area based nearest neighbors approach to intrusion detection," *Pattern Recognition*, vol. 43, no. 1, pp. 222–229, 2010.

[98] M. Z. Alom, V. Bontupalli, and T. M. Taha, "Intrusion detection using deep belief networks," in *2015 National Aerospace and Electronics Conference (NAECON)*, pp. 339–344, Dayton, OH, USA, 2015.

[99] S. A. Ludwig, "Intrusion detection of multiple attack classes using a deep neural net ensemble," in *2015 National Aerospace and Electronics Conference (NAECON)*, pp. 1–7, Dayton, OH, USA, November 2017.