International Transactions on Electrical Engineering and Computer Science
M. A. Daneswar, Vol: 3, No: 4, pp: 207-214, December 2024

Anomaly Detection in Data Streams Using Machine Learning and Deep Learning

Bakhtiyar Mahmood Abdullah, Muhammed Amin Daneswar

Abstract: Data stream mining for movement has emerged as an important area of machine learning because of the huge amount of changing and continuous data coming from diverse sources such as social media, business sensors, and mobile communications. The goal of this anomaly identification in the data streams is to find patterns that deviate substantially from the way things usually work. This will be valuable information for making decisions in a large number of areas, including healthcare, management of financial risk, keeping communities safe, and operating the power grid. This research discusses the intractable ways of finding oddities in a stream of data with the corresponding predicaments of always having a new inflow of data, creating information fast, and also dynamics of information changing. We also observe how distinct deep learning and machine learning approaches are being used in different fields to rapidly detect anomalies. Some examples of the way these techniques have been used to discover network intrusions, malware, IoT outliers, healthcare anomalies, and credit card frauds are a demonstration of the techniques work.

Keywords: Deep learning, Data stream mining, Anomaly detection, Network intrusion detection systems.

1. Introduction

In the beyond few years, statistics circulation mining has turned out to be one of the maximum critical and speedy growing areas of system mastering. Streaming facts resources that are extensively used and continually growing have created new aggressive opportunities for thorough and green statistics use, fast-growing huge quantities of multimodal data. Our networks and structures are usually getting new bits of facts from such things as cellular phones, websites, emails, social media, films, and business sensors. These streaming websites might be beneficial due to the fact they're new or up-to-date.

Article History
Received: 25-10-2024; Revised: 29-11-2024;
Accepted: 05-12-2024

M. A. Daneswar
mdaneshwar@yahoo.com

[1]Department of Computer Science, Faculty of Science, Soran University, Soran, Iraq

Many companies already realize that they want to apply online processing models and keeping their links with information resources up to date allows you to provide better models and offerings [1]. A critical part of records mining is locating patterns in data streams that don't behave usually. This is referred to as anomaly identity. This form of odd however essential records is frequently hidden in styles, and this information is regularly used to help humans make selections. Therefore, statistics circulate anomaly identification is becoming an increasingly number of critical in facts analysis and security. It is used a lot in the following situations [2]

Financial risk management: In finding, analyzing, and predicting the frauds, which occur in credit card transactions, insurance scams, and other fraud activities involved in banking cards, the data stream anomaly detection is used. Data stream anomaly detection is also involved in the commercial banks to conduct the investigation in the real-time exchange rate abnormalities to avoid facing substantial financial loss, which might be experienced by them and their clients.

https://doi.org/10.62760/iteecs.3.4.2024.118

***Power grid operations*:** Identifying power schedule data abnormalities is utilized to ensure safe power system operation.

Healthcare applications: Finding uncommon patterns in a medical image is a popular use for data stream anomaly detection, blood pressure, pulse, and other streaming healthcare data. Such irregularities could serve as crucial markers for identifying anomalous human situations.

Security of computer networks: Anomaly detection in data streams is often utilized for identifying intrusions that are not in accordance with the security rules.

Compared to static facts, statistics stream anomaly detection is notably extra difficult because of the following features

- ***Constant inflow of facts***: Storing all the information or doing a couple of scans is high priced because of the steady intake of facts. As a result, algorithms and fashions are required if you want to operate in environments with restricted sources.
- **High records era prices**: As a result, applicable algorithms must be capable of manner and evaluate statistics immediately.
- ***The dynamics and evolution of data flow***: algorithms that can adapt to keep their accuracy must detect static anomaly detection techniques are flawed because aberrant behaviors in statistics streams may evolve Anomalies in data flow. This means that the algorithms must be able to detect changes in the flow of data and then adjust accordingly [3].

1.1. Intrusion Detection in Network Traffic

Network safety is now an essential part of all web-based totally apps, like shopping, transactions, and other corporations. This is due to the fact such a lot of human beings use the net for everyday chores. Intrusions into the net can manifest via a number of exclusive ways to connect with the net. Malicious or hacks are words which can be sometimes used to explain individuals who break into network visitors. Changes within the velocity at which statistics is despatched, unpredictable internet use, sudden changes in get entry to instances, and different intrusions are all examples of intrusions. For instance, let's consider the plan is primarily based at the internet and that regular tracking of the way the net is used has taken area. An unusually high quantity of network usage as compared to how data has been used inside the beyond method that the record is an intrusion. It wishes to be located and fixed proper away to hold the network secure. This makes it clean that finding abnormal matters is a way to forestall people from moving into your network. The predominant goal is to find abnormal things occurring in signaling records in cell networks. Look at one extra example of a cellular community wherein the ambiguity detection approach is wanted to locate surprising changes in signal go with the flow. Things like the quantity of information in TBs, the wide variety of more than one record events in keeping with 2nd, and the speed of facts activities consistent with 2nd are used in this utility to discover things that do not appear proper. For locating outliers, the above examination assignment tried both batch processing and real-time analytics [4].

1.2. Malware identification in computing systems

Malware detection is the system of locating any sort of illegal behavior that does a lot of harm to computers. In turn, this makes the computer structures forestall operating properly. For example, computer systems that have dangerous software often don't work nicely, lose facts, and cannot protect their resources. The writers communicate approximately a way to use anomaly detection to find styles in facts that do not in shape what needs to be visible. The authors of the above-mentioned paper used an incremental clustering set of rules to technique the VMware movement data and stored the song of its success via measuring CPU load, reminiscence usage, and different factors. This helped them find extraordinary things within the statistics. The writers extensively utilized Apache Spark and Apache Kafka to find abnormal matters right away [5].

1.3. IoT outlier detection

Internet of Things made the availability of Internet Protocol (IP) services more and more possible due to the introduction of sensors and computers that can talk over the internet. The Internet of Things application needs proper detection of outlier problems.

For example, let's look at telephony, which is an Internet of Things application. Someone might receive a call from an attacker out of nowhere. Also, the attackers usually change their phone numbers and make fake calls. In this case, finding strange things helps find fake calls and stops other threats like data loss. It talk about how important it is to deal with real-time problems in streaming data from devices, mobile phones, the Internet of Things (IoT), machines, network data flow, application logs, and other sources. The authors tried using the machine learning algorithms Naive Bayesand Random Forest to find outliers in large datasets quickly in the Spark environment. The writers also suggested the method of Random Forest as it works better with fast data that gets big quickly [6].

1.4. Anomalies in Healthcare

Healthcare is the maximum touchy when it comes to actual-time anomaly spotting due to the fact reducing the death rate is so crucial. There are plenty of sensors and health tracking gadgets that may be used to preserve a watch on ill human beings all of the time. Real-time statistics are very critical in healthcare because it's miles one of the maximum crucial regions of lifestyles. For instance, let's assume that the coronary heart fee of an affected person is constantly being watched. Let's say that the heart price being tracked is 120 beats per minute, that is better than usual. The person in this example desires to get help right away so they don't die. Because of this, it's far very essential for the healthcare business to find troubles properly away. A lot of the paintings are targeted at the way to use machine-gaining knowledge to discover extraordinary matters in healthcare tracking structures. For locating outliers, the individuals who wrote this text have attempted both supervised and unsupervised systems gaining knowledge of algorithms. To find outliers, managed system studying algorithms like Support Vector Machine and Random Forest are used. Unsupervised devices gaining knowledge of algorithms like Local Outlier Factor, Isolation Forest, and K-Nearest Neighbor are used to discover outliers. The writers determined that the supervised device gaining knowledge of the algorithm is better at locating anomalies than the unsupervised technique [7].

1.5. Credit Card Fraud Detection

Credit cards are indispensable in today's marketplace and are used for all personal and professional shopping. It is easy to fraud a credit card and, hence, misuse the same. Credit card fraud involves unauthorized use of credit card information. The transaction, amount, place of transaction, etc., are analyzed to identify fraud in a credit card. For example, an amount above the average transaction amount is treated as fraud. The authors of [8] have tried to detect fraud in credit card transactions using several ML algorithms, viz., MLP, RF, NB, and LR. The Kaggle dataset is used, which contains 492 fraudulent transactions out of a total of 2, 84, 807 transactions. As per their results, RF is effective in identifying fraudulent transactions. In addition, the Table. 1 provided below gives a snapshot of the usage of several anomaly detection techniques in different fields.

Table. 1: Techniques for detecting anomalies in a range of applications

APP.	Instruments and Techniques for Anomaly Detection	Results	Reference
detection of intrusions in network traffic	Tools used: Apache Spark; Pearson Correlation and Relative Entropy	For real-time data, relative entropy works best.	[5]
malicious computer security attacks	Decision Tree, K-Nearest Neighbors; Genetic Algorithms; Naive Bayes, Artificial Neural Networks, Fuzzy Logic; Support Vector Machines; Hidden Markov Model.	The methods have been evaluated on a number of datasets, including NSL-KDD, CAIDA, DARPA, and others. Out of all the methods, SVM is proven to yield superior results.	[22]
(IoT)	NRDD-DBSCAN, DBSCAN, Tools used: Apache Spark RDD	Scalability is not a good fit for DBSCAN. As a result, anomalies are detected in nature using NRDD-DBSCAN, which operates in parallel.	[21]
Healthcare	Random Forest; Local Outlier Factor; Support Vector Machines; Isolation Forest	The optimal hit rate and accurate rejection of anomalies are provided by Random Forest.	[23]

	K-Nearest Neighbors		
Credit card fraud	Logistic Regression Random Forest; Naive Bayes Multilayer Perceptron	Random Forest accurately categorizes the fraudulent transaction.	[8]

2. Network Intrusion Detection Systems (NIDS) Using AI Techniques.

This section details the most common ML and DL algorithms used in making an effective NIDS. This section also discusses the functioning of AI-based NIDS as a whole. In a broad sense, ML and DL are both types of controlled and unsupervised algorithms. Supervised algorithms look through tagged data to gather important data, while unsupervised algorithms use unlabeled data to pull out useful information and traits.

2.1. Machine Learning Algorithms

Machine learning is a subfield of artificial intelligence that includes all methods and programs to let computers learn on their own, using mathematical models to find useful data in very large datasets. The most frequently used machine learning methods in IDS are k-nearest neighbors (KNN), decision trees, artificial neural networks (ANN), support vector machines (SVM), k-mean clustering, fast learning networks, and ensemble methods. Machine learning is sometimes referred to as "shallow learning" [9].

2.1.1. K-Nearest Neighbor

Among the simplest of supervised machine learning algorithms, KNN uses the idea of \"feature similarity\" to forecast the classification of the sample data. By determining how much of a sample is away from its neighbors, it can recognize it. In The model's performance depends on the parameter k in the KNN algorithm. In situations where k is too small, the model may over-fit. On the other hand, it may badly misclassify the sample instance when we take an overly large range of k. [10] evaluated the performance of various machine learning techniques on a new benchmark dataset CSE-CIC-IDS-2018. By applying the Synthetic Minority Oversampling Technique to decrease the imbalance ratio, the issue of the imbalance of the dataset has been eliminated, and the detection rate of minority class attacks has been enhanced.

2.1.2 Decision tree

One of the best approaches for guided device learning uses a decision tree (DT). It structures and makes predictions about a collection based on a set of decisions, or rules. Like an actual tree, the model has nodes, branches, and leaves. Every node is a characteristic or way of being. Each leaf is a potential outcome or a decision for a class, and each branch is a decision or a rule. So that it doesn't overfit, the DT algorithm chooses the best characteristics to build a tree by itself. It then prunes away any branches that are not needed. Many people purchase DT models like the ID3, C4.5, and CART. Many complex learning algorithms such as XGBoost60 and Random Forest (RF) are built from numerous decision trees [11].

2.1.3 Support vector machine

SVM is a supervised machine learning algorithm based on the concept of maximum margin hyperplane in an n-dimensional feature space. It can work on linear as well as nonlinear situations. Kernel functions will be applied to solve problems that do not follow a straight line. In the first place, the kernel function transforms input vectors with a small number of dimensions into a feature space with a large number of dimensions. Later, the support vectors are used to define an ideal maximum marginal hyper-plane that is used as a boundary for making a choice. NIDS can be improved and more appropriately used with the support vector machine method, where threats are classified into good and bad groups [12].

2.1.4 K-mean clustering

The goal of clustering is to put together pieces of data that are very similar so that the data can be grouped in useful ways. Some people like K-Mean clustering, which is a centroid-based iterative ML method that learns on its own. The number K in a collection stands for the number of centroids, or cluster centers. Most of the time, distance is used to put data points into groups. The main goal is to make the gaps

between each data point and its center as small as possible within a cluster [13].

To deal with IDS, [9] proposed a multilevel intrusion detection model framework called multilevel semi-supervised ML (MSML). The RF model is implemented along with the clustering concept. The four modules that constituted the proposed approach were pure cluster extraction, pattern recognition, fine-grained classification, and model updating. The aim is to forward an attack to the next module for detection if it is not labeled in the previous one. We evaluated the proposed methodology on the KDD Cup'99 dataset. The model was better at detecting attacks, even with few occurrences in the dataset, based on the experimental results.

2.1.5. Artificial neural network

ANN is a supervised machine learning technique inspired by the working of the human nervous system. It comprises connections between neurons, which are units of processing. Those nodes are organized with an input layer, some hidden layers, and an output layer. The backpropagation algorithm is used in learning an ANN. The most vital advantage of the artificial neural network (ANN) is that it can carry out nonlinear modeling due to learning from a massive dataset. A weakness of implementing the ANN model in training is that it is time-consuming because of its complexity, which inhibits learning and gives suboptimal results.. ELM, which is a novel type of ANN that overcomes the limitations of ANN, is adopted in this study, and it is, in fact, a single hidden layer feed-forward neural network used to solve the IDS problem. The paper proposes a model based on FLN and particle swarm optimization, abbreviated as PSO-FLN and tests the model with KDD Cup'99 dataset. PSO-FLN applies the input weights and hidden layer weights randomly, not requiring any form of tuning, and decides the output weights analytically. To test the model, FLN was tested with various optimization techniques because the results indicated that PSOFLN—when compared to other FLN models that apply various optimization techniques, namely Genetic Algorithm, Harmony Search optimization, and Optimization Based on Ameliorated Teaching and Learning, is superior in terms of its working approach, on the grounds that the accuracy increases when the number of neurons in the buried layer increases. The main disadvantage is that the accuracy in the detection rate drops for the smaller class of attacks [12].

2.1.6 Ensemble methods

The basic idea behind ensemble methods is that you extract the best from the different algorithms by learning together. Given that every classification has its strengths and weaknesses. Some systems may be better at detecting one type of attack versus another. An ensemble method trains many classifiers and then combines the weakest ones. The strongest classifier is then chosen by a vote. [14] proposed an ensemble-based IDS that selected ELM as the base classifier. In the ensemble pruning stage, the BAT optimization algorithm is adopted to optimize the suggested methodology. Kyoto, NSL-KDD, and KDD Cup'99 datasets are used to test the model. The experiment results claim that many ELMs working together in an ensemble are better than individual ELMs. With several base classifiers that include DT, RF, KNN, and Deep Neural Network (DNN), and finally the best one by an adaptive voting algorithm, [15] constructed an adaptive ensemble model. In the experiment, the NSL-KDD dataset was used to confirm the suggested methodology. Comparison of the experimental results between various models proved how well the performance was. For the less potent attack classes, the suggested methodology did not give results that were acceptable.

2.2. Deep learning algorithms

DL is a form of ML that makes use of many hidden layers to get features from deep networks. These strategies paintings higher than gadget mastering because they have a deeper shape and may pull out the crucial capabilities from a dataset on their very own and supply an output. This element suggests the DL strategies that were used inside the research that have been checked out to suggest DL-based totally NIDS solutions [11].

2.2.1 Recurrent neural networks

The motive of recurrent neural networks (RNNs) is to model collection data and to increase the talents of feed-forward neural networks. The hidden gadgets in an RNN are the notion because of the memory

additives. The RNN consists of enter, hidden, and output units. Each RNN unit uses both the output of the preceding enter and its modern-day input to make your mind up. RNN is notably employed in many domain names, consisting of handwriting prediction, audio processing, semantic knowledge, and human hobby reputation, to say some. The motive of recurrent neural networks (RNNs) is to version sequence information and to increase the abilities of feed-ahead neural networks. The enter, hidden, and output devices that make up an RNN are concepts of the reminiscence elements. Each RNN unit makes use of the output of the previous input and its modern-day input to determine.

RNN is significantly hired in lots of domains, which include handwriting prediction, audio processing, semantic knowledge, and human interest popularity, to mention some. RNNs can be applied for supervised type and characteristic extraction in an IDS. RNNs commonly have constrained duration series handling capacity and revel in brief-time period memory troubles with prolonged sequences87. In the context of binary and multi-magnificence classification of the NSL-KDD dataset, [16] offered RNN-based totally IDS. Several getting-to-know quotes and hidden node counts have been used to test the version. The quantity of hidden nodes and varying studying fees have been found to have an effect on the version's accuracy. Eighty hidden nodes and mastering fees of zero. 1 and 0.5 had been proven to yield nice accuracy in binary and multi-elegance settings. Xu et al. Cautioned an RNN-primarily based IDS in Reference 93, utilizing a multilayer perceptron, a softmax classifier, and GRU because of the number one reminiscence. The KDD Cup'ninety nine and NSL-KDD datasets had been used to test the recommended technique. The experimental findings tested high detection charges whilst in comparison to alternative methods. Lower detection charges for minority attack instructions like U2R and R2L are a fundamental flaw of their technique [17].

2.2.2 Deep neural network

DNN is a fundamental DL shape that we could the version examine at unique degrees. It has many hidden stages, as well as an enter layer and an output layer. Complex functions that are not linear are modeled through DNN. Adding extra secret layers to a model makes it a greater summary and increases its potential. [17] offered a CNN-based totally community IDS with 4 hidden layers to kind the KDD cup'99 and NSL-KDD datasets into corporations. In the output layer, a fully connected layer and a softmax classifier were brought to help with the class. This layer's activation function becomes a corrected linear unit [18]. Except for U2R, which had fewer facts than the alternative assault instructions, the statistics showed that the model became an excellent one. According to the authors, including more nodes and layers makes the machine greater complex, which requires extra sources and takes longer to procedure. The solutions to those problems are the optimization approach and automatic correction. [18] supplied a CNN-primarily based network IDS with four hidden layers to kind the KDD up ninety-nine and NSL-KDD datasets into businesses. In the output layer, a totally related layer and a softmax classifier have been delivered to assist with classification. This layer's activation characteristic turned into a corrected linear unit [18] Except for U2R, which had less information than the opposite attack lessons, the records showed that the version become a great one. According to the authors, including extra nodes and layers makes the system greater complicated, which requires more assets and takes longer to technique. The answers to these issues are the optimization approach and automated correction.

2.2.3. Convolutional neural network

Convolutional Neural Networks (CNNs) are a type of Deep Learning architecture used to handle data stored in arrays. Composed of an input and fully connected layer, this architecture also comprises a feature extraction stack with convolutional and pooling layers that feed into a classification layer whose classifier is a Softmax classifier. CNNs have been very successful in computer vision. In IDS, they are used for supervised feature extraction and classification. An effective use of CNNs was proposed by [19] for an IDS. The technique begins with feature extraction via Auto Encoders (AE) and Principal Component Analysis (PCA). The resultant feature set, one-dimensional in nature, is then shaped into a two-dimensional matrix that is finally fed into CNN. Experiments using the KDD Cup'99 dataset demonstrated the system's efficiency in lowering the time required to train and test the algorithms. The primary query, however, is that the

U2R and R2L classes have lower detection rates than other attack types [20].

3. Conclusion

Anomaly detection on data streams is an important part of present-day data analytics and security systems in use in many fields, such as banking, healthcare, network security, and the Internet of Things. Data streams are ongoing and move quickly, so they pose unique problems that must have resource-efficient and flexible algorithms for dealing with the data in real-time. This paper critically reviews how data stream anomaly detection techniques are used to find malware, protect the settings of IoT, find financial scams, monitor patient health, keep computer networks safe, and ensure the grid stays stable.

It demonstrates how different machine learning methods can be used and how effective they are, such as K-Nearest Neighbour, Decision Trees, Support Vector Machines, K-Mean Clustering, Ensemble Methods and Artificial Neural Networks. This fact is particularly important, as ensemble methods apply a lot of promise in a lot of areas where it needs many classifiers to combine and make detection more accurate. Anomaly detection is enhanced much better by using deep learning methods, especially Recurrent Neural Networks, to model sequential data and to find complex patterns. Bringing cutting-edge technologies like blockchain and edge computing together with systems looking for strange behavior could also make data more secure and processing faster. For this area to gain rapid development, it would be important for business and education to cooperate more to push forward new ideas and problem-solving in new ways.

Conflict of interest

The authors declared 'No conflict of interest'.

References

[1] Ł. Korycki, A. Cano and B. Krawczyk, "Active Learning with Abstaining Classifiers for Imbalanced Drifting Data Streams," *2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, CA, USA, pp. 2334-2343, 2019. https://doi.org/10.1109/BigData47090.2019.9006453

[2] R. A. A. Habeeb, F. Nasaruddin, A. Gani, I. A. T. Hashem, E. Ahmed, & M. Imran, "Real-time big data processing for anomaly detection: A survey", *International Journal of Information Management*, Vol. 45, pp. 289-307, 2019. https://doi.org/10.1016/j.ijinfomgt.2018.08.006

[3] M. A. Daneshwar & N. M. Noh, "Detection of stiction in flow control loops based on fuzzy clustering. *Control Engineering Practice*, Vol. 39, pp. 23-34, 2015. https://doi.org/10.1016/j.conengprac.2015.02.002

[4] S. Kunasekaran & C. Suriyanarayanan "Anomaly detection techniques for streaming data–an overview", *Malaya Journal of Matematik*, pp. 703-710, 2020. https://doi.org/10.26637/MJM0S20/0133

[5] L. Rettig, M. Khayati, P. Cudré-Mauroux and M. Piórkowski, "Online anomaly detection over Big Data streams", *2015 IEEE International Conference on Big Data (Big Data)*, Santa Clara, CA, USA, pp. 1113-1122, 2015. https://doi.org/10.1109/BigData.2015.7363865

[6] M. Sughasiny "Zero event anomaly detection in big data using spark for fast and streaming applications", *International Journal of Pure and Applied Mathematics*, Vol. 119, No. 15, pp. 3407-3412, 2018. [Cross Ref].

[7] M. A. Daneshwar, & N. M. Noh "Identification of a process with control valve stiction using a fuzzy system: A data-driven approach. *Journal of Process Control*, Vol. 24, No. 4, pp. 249-260, 2014. https://doi.org/10.1016/j.jprocont.2014.01.013

[8] D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic and A. Anderla, "Credit Card Fraud Detection - Machine Learning methods," *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, East Sarajevo, Bosnia and Herzegovina, pp. 1-5, 2019. https://doi.org/10.1109/INFOTEH.2019.8717766

[9] H. Yao, D. Fu, P. Zhang, M. Li and Y. Liu "MSML: A Novel Multilevel Semi-Supervised Machine Learning Framework for Intrusion Detection System", in *IEEE Internet of Things Journal*, Vol. 6, No. 2, pp. 1949-1959, April 2019. https://doi.org/10.1109/JIOT.2018.2873125

[10] G. Karatas, O. Demir and O. K. Sahingoz, "Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date

Dataset," in *IEEE Access*, Vol. 8, pp. 32150-32162, 2020.
https://doi.org/10.1109/ACCESS.2020.2973219

[11] S. S. Dhaliwal, A. A. Nahid & R. Abbas "Effective intrusion detection system using XGBoost", *Information*, Vol. 9, No. 7, pp. 149, 2018.
https://doi.org/10.3390/info9070149

[12] E. M. Roopa Devi, & R. C. Suganthe "Enhanced transductive support vector machine classification with grey wolf optimizer cuckoo search optimization for intrusion detection system", *Concurrency and Computation: Practice and Experience*, Vol. 32, No. 4, art. No. e4999, 2020.
https://doi.org/10.1002/cpe.4999

[13] A. Alalousi, R. Razif, M. AbuAlhaj, M. Anbar, & S. Nizam "A preliminary performance evaluation of K-means, KNN and EM unsupervised machine learning methods for network flow classification", *International Journal of Electrical and Computer Engineering*, Vol. 6, No. 2, pp. 778 - 784, 2016.
http://doi.org/10.11591/ijece.v6i2.pp778-784

[14] Y. Shen, K. Zheng, C. Wu, M. Zhang, X. Niu and Y. Yang, "An Ensemble Method based on Selection Using Bat Algorithm for Intrusion Detection", *The Computer Journal*, Vol. 61, No. 4, pp. 526-538, April 2018.
https://doi.org/10.1093/comjnl/bxx101

[15] X. Gao, C. Shan, C. Hu, Z. Niu and Z. Liu, "An Adaptive Ensemble Machine Learning Model for Intrusion Detection," in *IEEE Access*, Vol. 7, pp. 82512-82521, 2019.
https://doi.org/10.1109/ACCESS.2019.2923640

[16] 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.
https://doi.org/10.1109/ACCESS.2017.2762418

[17] Y. Jia, M. Wang, & Y. Wang "Network intrusion detection algorithm based on deep neural network", *IET Information Security*, Vol. 13, No. 1, pp. 48-53, 2019.
https://doi.org/10.1049/iet-ifs.2018.5258

[18] G. E. Dahl, T. N. Sainath and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout", *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, BC, Canada, pp. 8609-8613, 2013.
https://doi.org/10.1109/ICASSP.2013.6639346

[19] Y. Xiao, C. Xing, T. Zhang and Z. Zhao, "An Intrusion Detection Model Based on Feature Reduction and Convolutional Neural Networks," in *IEEE Access*, Vol. 7, pp. 42210-42219, 2019.
https://doi.org/10.1109/ACCESS.2019.2904620

[20] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac and P. Faruki, "Network Intrusion Detection for IoT Security Based on Learning Techniques," in *IEEE Communications Surveys & Tutorials*, Vol. 21, No. 3, pp. 2671-2701, 2019.
https://doi.org/10.1109/COMST.2019.2896380

[21] H. Ghallab, H. Fahmy, & M. Nasr, "Detection outliers on internet of things using big data technology. *Egyptian Informatics Journal*, Vol. 21, No. 3, pp. 131-138, 2020.
https://doi.org/10.1016/j.eij.2019.12.001

[22] A. Khraisat, I. Gondal, P. Vamplew, & J. Kamruzzaman "Survey of intrusion detection systems: techniques, datasets and challenges", *Cybersecurity*, Vol. 2, No. 1, pp. 1-22, 2019.
https://doi.org/10.1186/s42400-019-0038-7

[23] Z. Zojaji, R. E. Atani, & A. H. Monadjemi "A survey of credit card fraud detection techniques: data and technique oriented perspective", *arXiv preprint arXiv:1611.06439*, 2016.
https://doi.org/10.48550/arXiv.1611.06439

***