ISSN: 2581-3404 (Online) IF: 5.86 (SJIF)
IJIRTM, Volume-9, Issue-4, August-2025

Network Based Intrusion Detection System using Supervised Machine Learning: Survey & Discussion

[1]Dr. Rizwana Parveen, [2]Prof. Dinkar Likhitkar
[1,2]Assistant Professor
[1]Department of Computer Science,
[1] NRI Group of Institutions, Bhopal, (M.P.), India.
[2]Government College Sarni, (M.P.), India.

Abstract. *As we know that internet is very popular nowadays, everyone is using and doing all the works like education, online shopping, marketing using with the internet, to provide a security to the network is very crucial task, there is devices and tools are available to enhance the security of the computer network and provide security to the network. As now days every organization need to have data and other critical information in a secure mode, therefore it is very important to keep the safe to individual and network based system, As a result, sophisticated new attacks emerge, endangering vital infrastructure. An IDS is essential to detect and counter these attacks. IDS can be hardware devices or software products that monitor abnormal activity or behavior of the system. Pattern-matching systems detect known attack patterns, while statistical anomaly-based systems store typical behavior patterns in the database. In this research article present an overview of intrusion detection system, review and their challenges, also discuss for a near solution to provide for network based system from unwanted user or any suspicious activity. This research work also emphasize on machine learning based deployment for intrusion detection system to detect and analysis of the behaviours of system in a network.*

Keywords: Machine learning, Supervised classification, Attack detection, Intrusion detection system, Confusion matrix, Network based intrusion detection system.

Introduction

Intrusion is an unwanted activity in the network and intrusion detection is an important research and development topic with many applications that influencing confidentiality, integrity, availability. There are two main approach for security management these approaches are prevention-based and detection-based [4]. In any security plan, if intrusion prevention (encryption, authorization, and authentication) named as the first line of security is passed by attackers, as a second line of defence, intrusion detection comes into prominence. Intrusion detection provides deterrence for intruder and serves an alarm mechanism for a computer system or a network to manage security plan successfully. An intrusion-detection system (IDS) can be defined as software or hardware tools that monitoring network to detect internal or external cyber-attacks [1]. An Intrusion Detection System can observe and investigate system

and user activities, recognize patterns of known attacks, identify abnormal network activity. General definition of IDS is about intrusions to network but for WSN it can be added that physical damages to sensor devices [7]. Identifying sensor damage is important in order to serve fault tolerance and reliability. With the high usage of Internet in our day today life, security of network has become the key foundation to all web applications, like online auctions, online retail sales, etc. Detection of Intrusion, attempts to detect the attacks of computer by examining different information records observed in network processes. This can be considered as one of the significant ways to effectively deal with the problems in network security. An intrusion in the internet can compromise the data security through several internet means. Nowadays, the fast rising networks proliferation, data transfer rate, and an unpredictable Internet usage have added more anomaly problems [23, 22]. Thus researchers need to develop more reliable, effective, and self-monitoring systems, which sort troubles and can, carry out operation devoid of human interaction. By undergoing this kind of attempts, catastrophic failures of susceptible systems can be reduced. Detection stability and detection precision are two key indicators used to evaluate IDS (Intrusion Detection System). Many of the IDS research studies have been done in order to improve the detection stability and detection precision. In the beginning stage, the research work focus lies in using statistical approaches and rule-based expert systems [19]. But, the results of statistical approaches and rule-based expert systems were not accurate, when encountering larger datasets. In order to overcome the abovementioned problem, many data mining techniques were developed. An Intrusion Detection System is designed to detect an intrusion while it is in progress, or after it has occurred.

The major functions performed by IDS are monitoring users and systems activity, auditing system configurations, recognizing known attacks, identifying abnormal activities, managing audit data, highlighting normal activities, correcting system configurations and storing information about intruders [2]. Figure 1 showing a general overview of an intrusion detection system is a software application that continuously observes a network or system for abnormal activity. Any abnormal event noticed by IDS that can be reported immediately either to a system administrator or collected centrally using a security information and event management system [17]. The intrusion detection system checks thoroughly the incoming and outgoing traffic of Host or a network.

ML is a popular programming tool used for predicting and classifying large amounts of data. It focuses on grouping the same data or patterns in a single group, a method of data mining that differs from traditional programming languages [8,9]. Machine learning uses pattern analysis or behavioral analysis to categorize data based on properties or behavior. Classification is challenging due to the dynamic nature of data, and there are three classification methods: supervised, unsupervised, and semi-supervised. The network traffic undergoes some pre-processing operations that extract the relevant features from the data, transforming it into samples accepted by the ML model [6]. These samples are then forwarded to the (trained) ML model that will analyze them and determine whether they are legitimate or not [11]. There exist several solutions of ML-NIDS. An organization may adopt and maintain a ML-NIDS on premises, for example by leveraging open-source software. Alternatively, it may rely on third-party products.

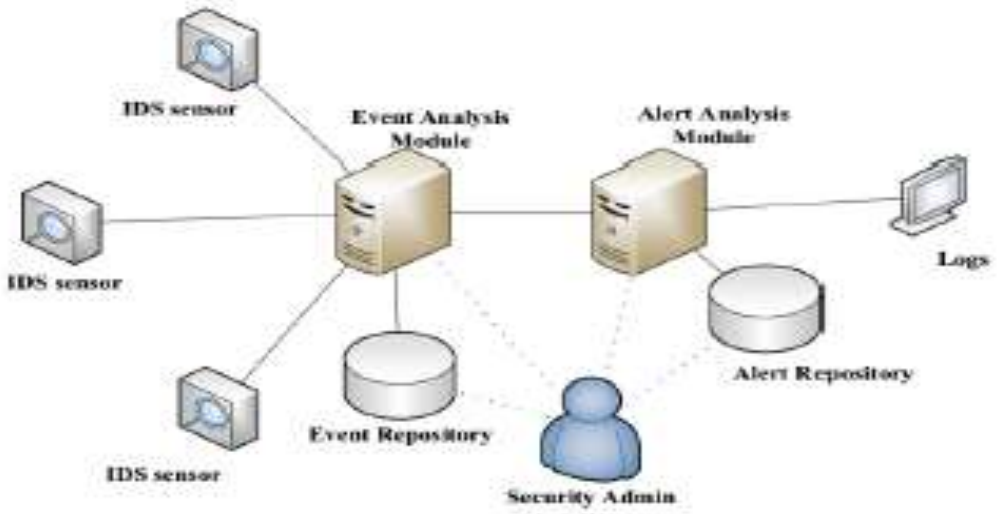

Figure 1: General Architectural diagram of IDS [4].

Intrusion Detection Techniques

The objective of the HIDS is the controlling state and dynamic behaviour of the computer system. This detection system checks all the activities of inspected packets on a network [25]. HIDS recognize what resources are being utilized and which program gets to those resources. If in the network any alternations or adjustment happens, system administer receive some network alerts [5]. HIDS is progressively becoming essential to ensure the host computer frameworks and its network activities. HIDS with host based information is incorporated into the computer frameworks to identify the intruder abnormal activities, noxious Behaviour, application abnormalities and preserve the Information Systems from intruders and report the occasions to the HIDS System Administrator. The figure 2 showing architectures of host based intrusion detection system [10].

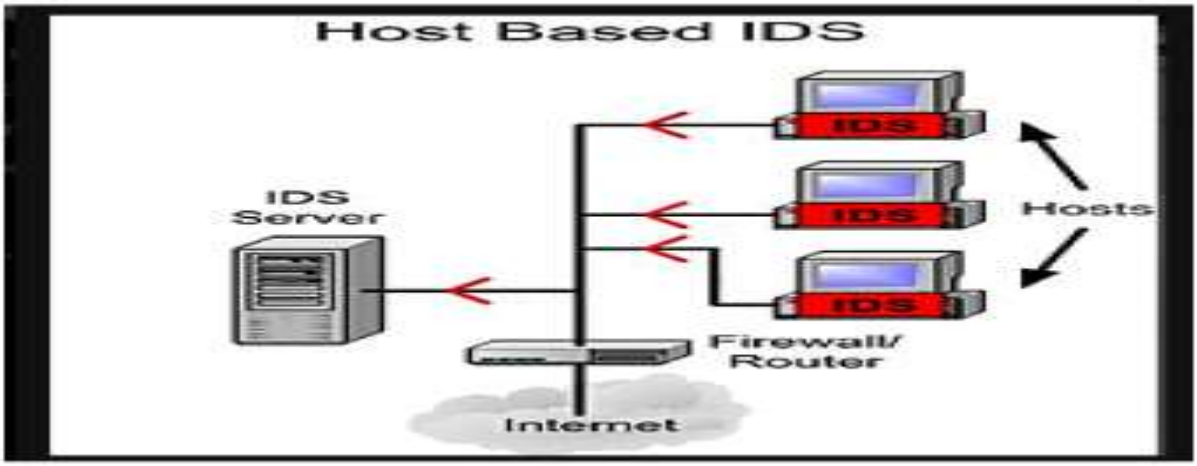

Figure 2: Architecture of host based IDS [8].

NIDS is the attribute function of target system and function modules are observed in network. The investigation of NIDS based on either manually or automatically. The NIDS is significantly used in the security infrastructure of the system. In NIDS to control the incoming and outgoing threads, anti-thread software is installed on the servers [32]. It is very essential to provide security for the several fields such as government application, business, industries and educational institution and so on. The Figure 3 represents the architecture of network based IDS. The NIDS consist of signature based classification, which identify the abnormalities by comparing with previous log files or signatures. Anomaly based technique, identifies the misuse as well as computer abnormalities, then classified as normal or attack depends on heuristics of signatures. The NIDS controls the device's outgoing and incoming packets. NIDs also classified as host based and network based [19].

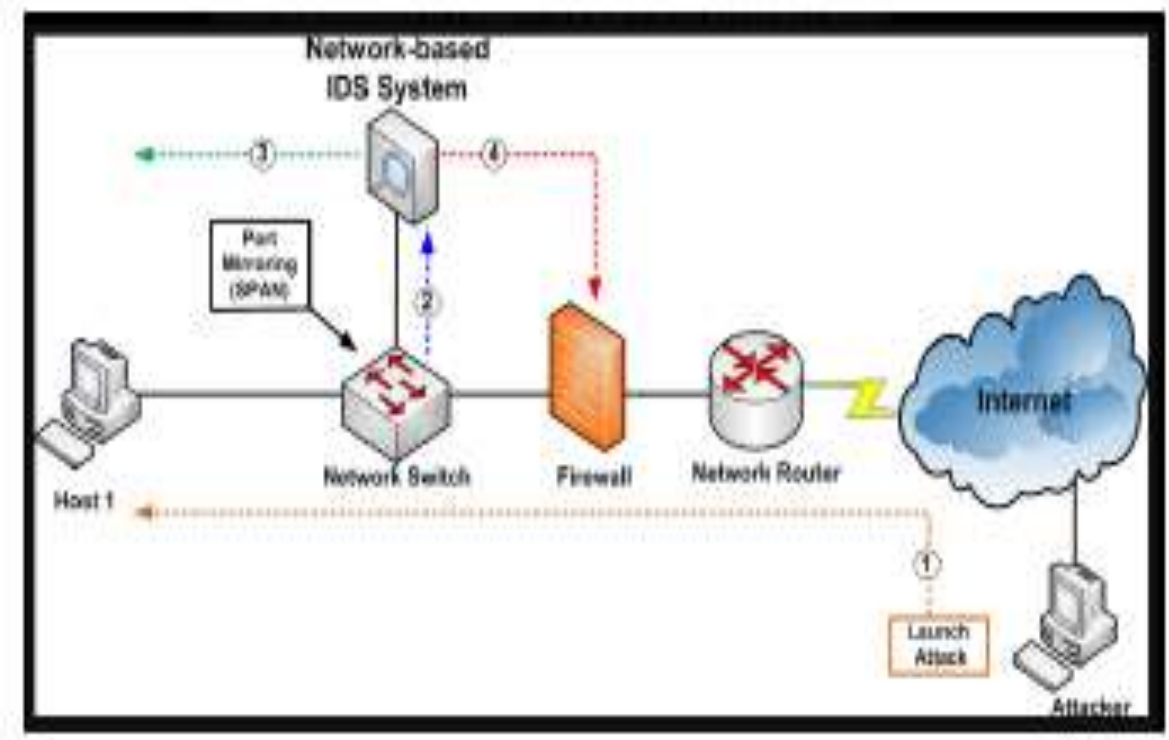

Figure 3: Architecture of network based IDS [11].

Literature Survey

In that research work author proposed a new approach called outlier detection where, the anomaly dataset is measured by the Neighbourhood Outlier Factor (NOF). Here, trained model consists of big datasets with distributed storage environment for improving the performance of Intrusion Detection system [36]. The experimental results proved that the proposed approach identifies the anomalies very effectively than any other approaches [1]. They have presented the details of a new approach called Outlier Detection

approach to detect the intrusion in the computer network. Their training model consists of big datasets with distributed environment that improves the performance of Intrusion detection system. The proposed approach is also been tested with the KDD datasets that are received from real world. The machine learning approaches detect the intrusion in the computer network with huge execution time and storage to predict the when compared to the proposed IDS system which takes less execution time and storage to test the dataset. That research work proposed a new similarity measure, the covering similarity that we formally define for evaluating the similarity between a symbolic sequence and a set of symbolic sequences [2]. A pair-wise similarity can also be directly derived from the covering similarity to compare two symbolic sequences. An efficient implementation to compute the covering similarity is proposed that uses a suffix-tree data-structure, but other implementations, based on suffix-array for instance, are possible and possibly necessary for handling very large-scale problems. They have used this similarity to isolate attack sequences from normal sequences in the scope of Host-based Intrusion Detection. They have assessed the covering similarity on two well-known benchmarks in the field. In view of the results reported on these two datasets for the state of the art methods, and according to the comparative study we have carried out based on three challenging similarity measures commonly used for string processing or in bioinformatics, they show that the covering similarity is particularly relevant to address the detection of anomalies in sequences of system calls. That research works proposed a three layer Intrusion Detection System (IDS) that uses a supervised approach to detect a range of popular network based cyber-attacks on IoT networks [3]. The system consists of three main functions: 1) classify the type and profile the normal behavior of each IoT device connected to the network, 2) identifies malicious packets on the network when an attack is occurring, and 3) classifies the type of the attack that has been deployed. The system is evaluated within a smart home test bed consisting of 8 popular commercially available devices. The effectiveness of the proposed IDS architecture is evaluated by deploying 12 attacks from 4 main network based attack categories such as: Denial of Service (DoS), Man-In-The-Middle (MITM)/Spoofing, Reconnaissance, and Replay. In that research work presented a Gender classification through Support Vector Machine and Scaled Conjugate Gradient Back Propagation Neural Network from face images using Local Binary Patterns [4]. To achieve better classification performance, need to be applied pre-processing technique first and then extracted the features on facial images from Local Binary Pattern Histogram method. These extracted features were stored into a vector called feature vector. Later, the feature vector is inputted to Polynomial SVM and SCG Back Propagation Neural Network classification methods along with labelled target vector. In that research work, author proposed a novel two-stage deep learning model, based on a stacked auto-encoder with a soft-max classifier [5], for efficient network intrusion detection. The model comprises two decision stages: an initial stage responsible for classifying network traffic as normal or abnormal, using a probability score value. This is then used in the final decision stage as an additional feature, for detecting the normal state and other classes of attacks. The proposed model is able to learn useful feature representations from large amounts of unlabeled data and classifies them automatically and efficiently. To evaluate its effectiveness, several experiments are conducted on two public datasets, specifically the benchmark KDD99 and UNSW-NB15 datasets. In that research work, proposed IDS based on deep learning using feed forward deep neural networks coupled with a filter-based feature selection algorithm [6]. The FFDNN-IDS is evaluated using the well-known NSL-knowledge discovery and data mining (NSL-KDD) dataset and it is compared to the following existing machine learning methods: support vectors machines, decision tree, K-Nearest Neighbor, and Naïve Bayes. The experimental results prove that the FFDNN-IDS achieve an increase in accuracy in

comparison to other methods. In that research work,, Big Data and Deep Learning Techniques are integrated to improve the performance of intrusion detection systems [7]. Three classifiers are used to classify network traffic datasets, and these are Deep Feed-Forward Neural Network (DNN) and two ensemble techniques, Random Forest and Gradient Boosting Tree (GBT). To select the most relevant attributes from the datasets, they use homogeneity metric to evaluate features. In that research work proposed new evolutionary multilayer perception neural networks using the recently proposed Bird Swarm Algorithm. The problem of finding the optimal connection weights and neuron biases is first formulated as a minimization problem with mean square error as the objective function [8].

Problem Statement

Intrusion identification is to monitor irregular behavior and misuse in the network. Intrusion recognition was presented in 1980's after the development of internet with surveillance to monitor the risk presents in the network. The reputation and incorporation of security infrastructures are suddenly increased. From that point onward, a few activities in IDS innovation have advanced detection of network interruption in its present state. One of the focused areas to resolve cyber-attacks quickly is to detect the attack process early from the network using NIDS. Network intrusion detection systems (NIDS) are designed to detect malicious activities including virus, worm, DDoS attacks. The critical success factors for NIDS are abnormality detection speed, accuracy and reliability. The current limitations of IDS systems are:

- False Positives: The major problem is IDS predict the false intrusion attacks. If this rate is high, then normal attack predicted as malicious. Reduction of false positive rates in IDS is the complex task.
- False Negatives: If the false negative rates are high then it is a problem because when intrusion occurred, IDS doesn't produce any alert.
- True Positive: An occurs when an actual attack occurs and the IDS responds to it by raising the appropriate alarm.
- True Negative: When no attack happens, the IDS does not raise alarm.
- Low value of accuracy.
- Sometime system is not able to efficiently handle or classify the datasets into normal and abnormal categories.

IDS Deployment with Machine Learning

The detection of malicious events is a prominent issue in the cyber security landscape. As manual inspection is impossible when millions of events per day occur, human defenders are supported by intrusion detection systems (IDS) that analyze data from different sources and, when specific conditions are met, generate alerts for the triage phase [41]. We consider the specific category of Network-IDS, which aim at identifying intrusions at the network traffic level. Several types of NIDS exist, but common differences involve the data type analyzed, and the method used to perform the detection [7]. The first generation of NIDS used to analyze network packets by inspecting their payload. The machine learning techniques can be used in IDSs as integrated, combined, or group in classifying. The classifying can be supervised, semi-supervised, and unsupervised in three operational modes [34]. In general, the supervised mode (or the classification) works better than the other modes; however, due to the high dimensions of the data exchanged in IDS, classifiers' use is very time-consuming, especially when looking to detect

intrusion in real-time. Figure 4 shows typical deployment of a network intrusion detection system. While this approach may be more accurate, it cannot be applied when data is encrypted, and requires high amounts of computational resources to process each network packet [43]. The exponential growth of traffic that often is encrypted raised the interest toward NIDS inspecting metadata, such as network flows.

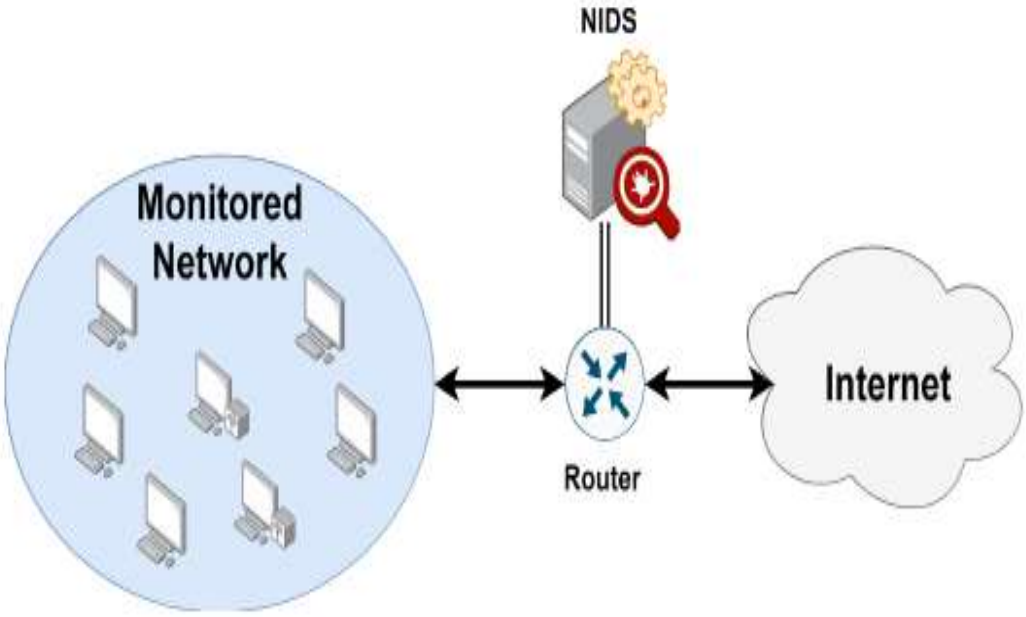

Figure 4: Typical deployment of a network intrusion detection system [28].

Machine Learning generates models that are able to learn specific patterns by providing them with training data. These patterns are then used to make predictions on new and unseen sets of data. The main advantage of these detection schemes is their capability of automatically learning from training data without human intervention, thus simplifying the resource-intensive management procedures required by traditional misuse-based approaches. Furthermore, they are also able of detecting novel attack variants [28], for which no known signature exists and that would be undetectable by NIDS based exclusively on rules.

Conclusion

Today, with the rapid growth and the wide application of the Internet and Intranet, computer networks have brought great convenience to people's life and work. However, at the same time, they also brought a lot of security problems, such as various types of viruses, vulnerabilities and attacks, which cause heavy losses. In recent years, many researchers have introduced more and more innovative techniques to detect intrusions, such as machine learning, data mining and evolutionary techniques. The above mentioned intrusion detection evaluation results are very encouraging, but these classification techniques still have detection defects, low detection rate for unknown attacks and high false positive rate for unbalanced samples, in this article present intrusion detection approach with machine learning.

References

[1] G. Uthradevi, P. Thiruvasagamm "A Semi-Supervised Deep Learning Approach for Intrusion Detection and Classification for the Internet of Things", Springer, 2025, pp. 1-17.

[2] Abdelwahed Berguiga, Ahlem Harcha, "HIDS-RPL: A Hybrid Deep Learning-Based Intrusion Detection System for RPL in Internet of Medical Things Network", IEEE 2025, pp. 38404-38429.

[3] Z.K. Maseer, Q.K. Kadhim, B. Al-Bander, R. Yusof, A. Saif, "Meta-analysis and systematic review for anomaly network intrusion detection systems: Detection methods, dataset, validation methodology, and challenges". IET Netw, 2024.

[4] M.K. Roshan, A. Zafar, "Boosting robustness of network intrusion detection systems: A novel two phase defense strategy against untargeted white-box optimization adversarial attack". Expert Syst. Appl. 2024.

[5] LirimAshiku, CihanDagli, " Network Intrusion Detection System using Deep Learning", Procedia Computer Science, 2021, pp. 239–247.

[6] Alabdulatif, "Network intrusion detection system using an optimized machine learning algorithm", Mehran University Research Journal of Engineering and Technology, 2023, pp. 153-164.

[7] B. Ida Seraphim, E. Poovammal, "Analysis on Intrusion Detection System Using Machine Learning Techniques", 2021, pp. 423-442.

[8] Mohanad Sarhan, Siamak Layeghy, "Towards a Standard Feature Set for Network Intrusion Detection System Datasets", 2021, pp. 1-13.

[9] Muawia A. Elsadig, Abdelrahman Altigani, "Breast cancer detection using machine learning approaches: a comparative study", International Journal of Electrical and Computer Engineering, 2023, pp. 736-745.

[10] Md. Alamin Talukder, Khondokar Fida Hasan, Manowarul Islam, "A Dependable Hybrid Machine Learning Model for Network Intrusion Detection", IEEE, 2022, pp. 1-44.

[11] Rakesh Shrestha, Atefeh Omidkar, "Machine-Learning-Enabled Intrusion Detection System for Cellular Connected UAV Networks", Electronics 2021, pp. 1-28.

[12] Shereen Ismail, "An Ensemble-Based Machine Learning Approach for Cyber-Attacks Detection in Wireless Sensor Networks", Appl. Sci. 2023, pp. 1-15.

[13] Giovanni Apruzzese, Luca Pajola, "The Cross-evaluation of Machine Learning-based Network Intrusion Detection Systems", IEEE Transactions on Network and Service Management, 2022, pp. 1-18.

[14] R Dubey, "An empirical study of intrusion detection system using feature reduction based on evolutionary algorithms and swarm intelligence methods", International Journal of Applied Engineering Research, 2017. pp. 8884-8889.

[15] Tao Wu, Honghui Fan, Hongjin Zhu, "Intrusion detection system combined enhanced random forest with SMOTE algorithm", Journal on Advances in Signal Processing, 2022, pp. 1-20.

[16] Thi-Thu-Huong Le, Haeyoung Kim, "Classification and Explanation for Intrusion Detection System Based on Ensemble Trees and SHAP Method", Sensors 2022, pp. 1-28.

[17] Seok-Jun Bu, Han-Bit Kang, "Ensemble of Deep Convolutional Learning Classifier System Based on Genetic Algorithm for Database Intrusion Detection", Electronics 2022, pp. 1-16.

[18] Rozin Majeed Abdullah, "Machine learning Algorithm of Intrusion Detection System", Asian Journal of Research in Computer Science, 2021, pp. 1-13.

[19] Deepak Rathore, Anurag Jain, "Design Hybrid method for intrusion detection using Ensemble cluster classification and SOM network", International Journal of Advanced Computer Research, 2012, pp. 181-186.

[20] Syed Rizvi, Mark Scanlon, "Deep Learning Based Network Intrusion Detection System for Resource-Constrained Environments", IEEE, 2021, pp. 1-13.

[21] Hooman Alavizadeh, Hootan Alavizadeh, "Deep Q-Learning Based Reinforcement Learning Approach for Network Intrusion Detection", Computers 2022, pp. 1-19.

[22] Abid Salih, Siddeeq Y. Ameen, "Deep Learning Approaches for Intrusion Detection", Asian Journal of Research in Computer Science, 2021, pp. 50-64.

[23] Ripon Patgiri, Udit Varshney, Tanya Akutota, and Rakesh Kunde, "An Investigation on Intrusion Detection System Using Machine Learning", IEEE 2018, pp 1684-1691.

[24] Shone, N, Tran Nguyen, N, Vu Dinh, P and Shi, "A Deep Learning Approach to Network Intrusion Detection", IEEE Transactions on Emerging Topics in Computational Intelligence, 2017, pp 1-11.

[25] Longjie Li , Yang Yu, Shenshen Bai, Jianjun Cheng, Xiaoyun Chen, "Towards Effective Network Intrusion Detection: A Hybrid Model Integrating Gini Index and GBDT with PSO", Journal of Sensors, 2018, Pp 1-10.

[26] Yanqing Yang, Kangfeng Zheng, Chunhua Wu, Xinxin Niu, Yixian Yang, "Building an Effective Intrusion Detection System Using the Modified Density Peak Clustering Algorithm and Deep Belief Networks", Applied Science Journal, 2019, Pp 1-25.

[27] M. Mazhar Rathore, Faisal Saeed, Abdul Rehman, Anand Paul, Alfred Daniel, "Intrusion Detection using Decision Tree Model in High-Speed Environment", International Conference on Soft-computing and Network Security, IEEE 2018, Pp 1-5.

[28] L. Khalvati, M. Keshtgary, N. Rikhtegar, "Intrusion Detection based on a Novel Hybrid Learning Approach", Journal of AI and Data Mining, 2018, Pp 157-162.

[29] Deepak Kumar Rathore, Dr. Praveen Kumar Mannepalli, "A Review of Machine Learning Techniques and Applications for Health Care", International Conference on Advances in Technology, Management & Education, 2021, IEEE proceeding, 978-1-7281-8586-6/21.

[30] Ban Salman Shukur1, Maad M. Mijwil, "Involving machine learning techniques in heart disease diagnosis: a performance analysis", International Journal of Electrical and Computer Engineering, 2023, pp. 2177-2185.

[31] Belal Abuhaija, Aladeen Alloubani, "A comprehensive study of machine learning for predicting cardiovascular disease using Weka and SPSS tools", International Journal of Electrical and Computer Engineering, 2022, pp. 1891-1902.

[32] Huru Hasanova, Muhammad Tufail, Ui-Jun Baek,Jee-Tae Park, Myung-Sup Kim."A novel blockchain-enabled heart disease prediction mechanism using machine learning", Computers and Electrical Engineering, 2022, pp. 1-13.

[33] Deepak Kumar Rathore, Praveen Kumar Mannepalli, "Recent Trends in Machine Learning for Health Care Sector", International Journal of Innovative Research in Technology and Management, Vol-5, Issue-2, 2021.

[34] GhulabNabi Ahmad, Hira Fatima, Shafiullah, "Efficient Medical Diagnosis of Human HeartDiseases Using Machine Learning TechniquesWith and Without Grid SearchCV", IEEE Access, 2022, pp. 80151-80173.

[35] AlaaMenshawi, Mohammad Mehedi Hassan, "A Hybrid Generic Framework for Heart Problem DiagnosisBased on a Machine Learning Paradigm", Sensors 2023, pp. 1-17.

[36] Anna MarkellaAntoniadi, Yuhan Du, "Current Challenges and Future Opportunities for XAI in Machine Learning-Based Clinical Decision Support Systems: A Systematic Review", Appl. Sci. 2021, pp. 1-23.

[37] Baptiste Vasey, MMed; Stephan Ursprung, "Association of Clinician Diagnostic Performance With Machine Learning–Based Decision Support Systems A Systematic Review", JAMA Network, 2021, pp. 1-15.

[38] Muhammad Sajjad, Sana Zahir Amin Ullah, Zahid Akhtar, Kha, n Muhammad, “Human Behavior Understanding in Big Multimedia Data Using CNN based Facial Expression Recognition”, Mobile Networks and Applications, Springer 2019, pp 1-11.

[39] Xiaofeng Liu, B.V.K. Vijaya Kumar, Ping Jia, Jane You, “Hard negative generation for identity-disentangled facial expression recognition”, Pattern Recognition, 2019, pp. 1-12.

[40] Tata Sutabri, Pamungkur, Ade Kurniawan, Raymond Erz Saragih, “Automatic Attendance System for University Student Using Face Recognition Based on Deep Learning”, International Journal of Machine Learning and Computing, 2019, pp. 668-674.

[41] Deepak Rathore, “Diseases Prediction and Classification Using Machine Learning Techniques”, AIP Conference Proceedings 2424, 070001 (2022); https://doi.org/10.1063/5.0076768.

[42] Emmanuel Ileberi, Yanxia Sun, Zenghui Wang, “Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost”, IEEE Access, 2021, pp. 165286-165295.

[43] Ebenezer Esenogho, Ibomoiye Domor Mienye, “A Neural Network Ensemble with Feature Engineering for Improved Credit Card Fraud Detection”, IEEE Access, 2022, pp. 16400-16408.

[44] Wei Zhou, Xiaorui Xue, “Credit card fraud detection based on self-paced ensemble neural Network”, ITCC 2022, pp. 92-99.

[45] Tzu-Hsuan Lin, Jehn-Ruey Jiang, “Credit Card Fraud Detection with Autoencoder and Probabilistic Random Forest”, Mathematics 2021, pp. 1-16.

[46] Gayan K. Kulatilleke, “Credit Card Fraud Detection Classifier selection Strategy”, 2022, pp. 1-17.