# A Novel ANN-Based Support Vector Machine for improving Classification Accuracy in Intrusion Detection Systems

**Mallaradhya C[1]*, Dr. G N K Suresh Babu[2]**

[1]Research Scholar, Department of Computer Science, Srishti College of Commerce and Management, University of Mysore, Email: mallaradhyac1985@gmail.com
[2]Professor, Department of Computer Science, Srishti College of Commerce and Management, University of Mysore
*Corresponding Author

Received: 14.07.2024 Revised: 17.08.2024 Accepted: 20.09.2024

**ABSTRACT**

This work recommends a innovative hybrid context for intrusion discovery, specifically tailored to address the challenges posed by the composite and energetic landscape of cyber threats, utilizing the CICIDS2017 dataset, that affords a complete and realistic illustration of modern web transportation. This approach integrates the strengths of Artificial Neural Networks (ANN) and Support Vector Machines (SVM) to develop classification accurateness and toughness in identifying malicious activities within network traffic. The use of ANN in this framework serves a crucial role in feature learning and representation. By employing a multi-layered architecture, ANN can autonomously extract intricate features from raw network data, capturing both spatial and temporal dependencies that are indicative of various types of network intrusions. This adaptive feature extraction capability enables the system to adapt to evolving attack strategies and effectively differentiate between normal and anomalous network behavior. Complementing the feature learning aspect, Support Vector Machines (SVM) are employed for classification tasks, leveraging their ability to construct optimal hyper planes in high-dimensional feature spaces. SVMs excel in binary classification tasks, making them well-suited for distinguishing between normal and intrusive network traffic. By integrating SVM into the framework, the researcher aims to exploit its robustness and generalization capability to improve the general performing and reliability of this intrusion detecting method. Extensive experimentation and evaluation are handled using the CICIDS2017 dataset to assess the effectiveness of the recommended approach. Reasonable analyses are performed against traditional SVM and ANN classifiers, as well as other contemporary intrusion detecting methods, to benchmark the execution in terms of classification accuracy. The results demonstrate consistent improvements achieved by hybrid approach, highlighting its efficacy in detecting various types of network intrusions while minimizing false positives. This research work presents a comprehensive and effective framework for intrusion detection, leveraging the synergies between ANN and SVM to enhance classification accuracy to 97.20%. By advancing the high-tech in intrusion detecting techniques, this study contributes to strengthening network defense measures and mollifying the threats posed by progressing cyber coercions in modern computing environments.

**Keywords:** Intrusion Detection, Artificial Neural Networks (ANN), Support Vector Machines (SVM), CICIDS2017 Dataset, Classification Accuracy

## 1. INTRODUCTION

In today's interconnected world, the security of computer networks is paramount to safeguarding sensitive information and maintaining the integrity of digital assets. However, with the proliferation of cyber threats and the increasing sophistication of malicious actors, the task of protecting systems versus not permitted approach and malevolent behavior has become increasingly challenging [Thakkar A et al 2021]. Intrusion Detection Systems (IDS) show a critical task in this work by dynamically examining network stream of traffic and recognizing anomalous comportment that may designate probables anctuary breaks or unauthorized access attempts [Sultana N et al 2019].

An Intrusion Detection System (IDS) is a critical element of a comprehensive cyber security strategy, providing an additional layer of defense against cyber threats beyond traditional firewalls and antivirus software [Gautam RKS et al 2018]. Unlike preventive security measures that aim to block known threats, IDS operates on the principle of anomaly detection, identifying deviations from normal network behavior that may signify suspicious or malicious activities [Md. Nayer et al 2022]. By uninterruptedly observing system traffic and investigating outlines of communication, IDS can perceive and observant overseers to possible safety breaks in real-time, enabling prompt response and mitigation efforts to minimize the impact of cyber-attacks [Moshref M et al 2022].

The evolution of cyber threats has necessitated the development of increasingly sophisticated IDS technologies capable of detecting a extensive assortment of attacks, comprising network interference attempts, malware infections, denial-of-service attacks, and insider threats [Fan L et al 2021]. Traditional IDS approaches, such as signature-based detection and rule-based filtering, are effective against known threats but often struggle to detect novel or previously unseen attacks [Gulab Sah et al 2022]. To address this limitation, modern IDS systems employ advanced techniques such as machine learning, anomaly detection, and behavioral analysis to improve detection accuracy and resilience against evolving threats [Mendonca R V et al 2021].

In recent years, the availability of comprehensive datasets, such as the CICIDS2017 dataset [Engelen G et al 2021], has facilitated the development and evaluation of IDS systems in realistic network environments. These datasets provide researchers with access to diverse and representative samples of network traffic, enabling rigorous testing and benchmarking of intrusion detection techniques [Vijayakumar D S et al 2022]. By conducting extensive experimentation and evaluation using real-world datasets, researchers can assess the effectiveness and scalability of different IDS approaches and identify opportunities for improvement [Lohiya R et al 2021]. IDS serve as a frontline defense mechanism alongside the ever-developing site of cyber intimidations, playing a crucial part in classifying and uncomfortable unofficial admittance efforts, malware contagions, denial-of-service occurrences, and various other forms of malicious activities targeting computer networks [Maseer Z K et al 2021]. As systems remain to grow in complexity and scale, the requirement for efficient and accurate detection methods becomes increasingly imperative. In this context, classification methods portray a pivotal task in developing the efficacy and reliability of IDS by enabling automated decision-making processes based on the analysis of network traffic patterns and behaviors [Hosseini S et al 2021].

The primary objective of classification methods in IDS is to accurately differentiate between normal network traffic and potentially malicious activity, thereby minimizing false positives and false negatives [Anurag Chhetri et al 2022]. False positives occur when legitimate network activities are incorrectly flagged as suspicious or malicious, leading to unnecessary alarms and resource wastage. Conversely, false negatives occur when actual security threats go undetected, leaving networks vulnerable to exploitation and compromise [Kumara A et al 2018]. Classification methods aim to uncover a equilibrium among detection accurateness and effectiveness, ensuring that genuine threats are promptly identified while minimizing the occurrence of false alarms [Mohamad Faiz Ahmad et al 2022].

One of the key challenges in intrusion detection is the inherent complexity and variability of network traffic patterns, which can exhibit significant fluctuations over time due to factors such as user behavior, software updates, and network topology changes [Thakkar A et al 2022]. Classification methods address this challenge by leveraging machine learning procedures, arithmeticexamination techniques, and pattern recognition approaches to recognize distinctive features and characteristics associated with different types of network activity [Anwer HM et al 2018]. By extracting relevant features from network data and mapping them to predefined classes or categories, classification methods enable IDS to make informed decisions about the nature and severity of detected events.

Machine learning-based cataloguing approaches, such as verdict trees, neural networks, support vector machines, and ensemble learning algorithms, have gained prominence in IDS due to their capability to absorb from historical information and adapt to changing threat landscapes [Thakkar A et al 2021]. These methods can automatically identify complex patterns and relationships within network traffic, allowing IDS to detect both known and previously unseen forms of malicious activity. Furthermore, machine learning techniques enable IDS to continuously evolve and recover over time, as they can incorporate new information and adapt their decision-making criteria based on emerging trends and attack vectors.

Another advantage of classification methods in IDS is their scalability and efficiency in handling large volumes of network traffic. Traditional signature-based detection approachestrust on predefined systems and models to recognize known threats, which can be cumbersome to maintain and update in dynamic network environments. In contrast, classification methods can examine vast volumes of information in real-time, leveraging comparable managing and dispersed computation methods to accomplish high quantity and low latency. This scalability enables IDS to effectively monitor high-speed networks and detect sophisticated attacks without compromising performance [Zhang J et al 2019]. Classification methods play a critical part in developing the effectiveness and efficiency of Intrusion Detection Systems (IDS) by enabling automated decision-making processes based on the analysis of network traffic patterns and behaviors. By leveraging machine learning algorithms, statistical analysis techniques, and pattern recognition approaches, classification methods empower IDS to accurately differentiate between normal and malicious activity, minimizing false positives and false negatives [Harush S et al 2021]. Moreover, the scalability and adaptability of classification methods make them well-suited for focusing the confronts modeled by the evolving risk model and the increasing complexity of modern computer networks.

## 2. LITERATURE SURVEY

In the scope of cyber security, Intrusion Detection Systems (IDS) serve as critical tools for identifying and thwarting intrusion attacks, safeguarding digital environments from malicious threats. With the exponential growth in data generation, the spectrum of potential intrusion attacks expands correspondingly, necessitating advanced methodologies for effective detection. Feature selection emerges as a pivotal aspect in enhancing IDS performance, allowing for the identification of relevant patterns amidst vast datasets. Moreover, the organization of the dataset itself profoundly impacts the efficacy of machine learning models employed in IDS, highlighting the importance of dataset carnation and preprocessing. Addressing data imbalance represents another significant challenge in IDS, where certain classes of data may be underrepresented, potentially leading to biased model outcomes. However, innovative sampling approaches offer promising avenues for mitigating such imbalances, thereby enhancing the robustness of intrusion detection systems.

A comprehensive literature review reveals a myriad of research endeavors aimed at exploring diverse machine learning approaches for IDS, each offering unique insights into dataset selection, algorithmic methodologies, and performance evaluation metrics. Notably, studies such as those conducted by Dini P et al (2023) and E L Asry C et al (2024) delve into the intricate interplay between machine learning algorithms and IDS performance, utilizing datasets such as KDD 99, UNSW-NB15, and CSE-CIC-IDS 2018 to evaluate classification accuracy and model efficacy. These investigations underscore the critical role of dataset choice in ensuring the relevance and applicability of IDS models to real-world scenarios. Moreover, Qazi Emad-ul-Haq et al (2022) showcase the potential of deep learning techniques in enhancing IDS capabilities, leveraging non-symmetric deep auto-encoders to achieve remarkable accuracy rates in network intrusion detection. Furthermore, A. Ugendhar et al (2022) advocate for a paradigm shift towards deep multilayer classification approaches in IDS, highlighting the ought for adaptable and scalable defense systems in the challenge of advancing cyber threats. Their proposed methodology, comprising preprocessing, autoencoding, database, classification, and feedback modules, demonstrates confirming results in times of accurateness and detection rates across benchmark datasets such as NSL-KDD. Additionally, Abbas Q et al (2023) present a original amalgam collaborative prototypical utilizing arbitrary forest recursive feature elimination (RF-RFE) to optimize IDS performance, offering an efficient alternative to deep learning paradigms with reduced computational costs and training times. Moreover, Emad E. Abdallah et al (2022) provide a comprehensive taxonomy of supervised machine learning procedures for intrusion detection, highlighting the high classification performance exhibited across popular datasets such as KDD'99, NSL-KDD, CICIDS2017, and UNSW-NB15. Their review underscores the importance of feature selection techniques and data sampling methodologies in mitigating challenges associated with dataset imbalance and enhancing overall IDS performance.

**Table 1.** Review of Classification Methods for Different Datasets

| S.No | Author Names | Dataset Used | Methodology | Performance Metrics |
|---|---|---|---|---|
| 1 | Dini P et al 2023 | KDD 99, UNSW-NB15, CSE-CIC-IDS | Machine Learning Approaches | Accuracy (100% for binary classification) |
| 2 | E L Asry C et al 2024 | NSL-KDD, UNSW-NB15 | Feature Selection, PV-DM, | Accuracy (98.92% in NSL-KDD, 82.86% in UNSW-NB15) |
| 3 | Qazi Emad-ul-Haq et al 2022 | KDD CUP'99 | Deep Learning Approach | Accuracy (99.65%) |
| 4 | A. Ugendhar et al 2022 | NSL-KDD | Ensemble Learning Method | Accuracy (96.7%) |
| 5 | Abbas Q et al 2023 | NSL-KDD, UNSW-NB15, CSE-CIC-IDS2018 | Random Forest-RFE | Overall Accuracy (99%, 98.53%, 99.9%) |
| 6 | Emad E. Abdallah et al 2022 | KDD'99, NSL-KDD, CICIDS2017, UNSW-NB15 | Taxonomy Approach | Classification Performance (High and Promising) |

| | | | | |
|---|---|---|---|---|
| 7 | Kumar N et al 2023 | NSL-KDD, KDD-CUP99, UNSW NB15 | Hybrid Intelligent System | Accuracy (99.967%, 99.567%, 99.726%) |
| 8 | M. Reji et al 2023 | CIC-IDS-2018 | SOA and ELM Classifier | Accuracy (94.22%) |
| 9 | Jiyuan Cui et al 2023 | NSL-KDD, UNSW-NB15 | GMM-WGAN-IDS | Outclasses State-ofHigh-tech Methods |
| 10 | Tao Wu et al 2022 | NSL-KDD | Enhanced Random Forest | Accuracy (99.72% on training set, 78.47% on test set) |

The table 1 gives the consolidated performance information in the review of related works. for This literature review elucidates the multifaceted landscape of intrusion detection systems, emphasizing the intricate interplay between dataset selection, algorithmic methodologies, and performance evaluation metrics. By synthesizing insights from diverse research endeavors, this review provides valuable perspectives on the current contemporary in IDS, paving the way for future advancements in cyber security research and practice.

## 3. Challenges Of The Existing Ids

By the extensive review of the related works in the previous section, the challenges which exist in the available IDS is described below.

### 3.1 Complexity of Cyber Threats

Existing classification systems face significant encountersfitting to the intricate and ever-developingtype of cyber threats. Intruders continuously devise new tactics and attack vectors, necessitating classification algorithms capable of adapting and learning in real-time to accurately differentiate between normal network behavior and malicious activities.

### 3.2 Data Imbalance

A prevalent challenge in classification tasks, data imbalance occurs when certain classes of data are underrepresented compared to others. In intrusion detection systems (IDS), this unevenness can precede to prejudicedpattern outcomes, where the classifier may prioritize the common class and exhibit poor performance in detecting minority class intrusions. Addressing data imbalance requires sophisticated sampling approaches and algorithmic techniques to ensure fair representation of all classes and improve overall model robustness.

### 3.3. Scalability

The scalability of classification algorithms is crucial, especially in large-scale network environments where IDS must handle vast volumes of data in real-time. Traditional methods may struggle to cope with the computational demands of processing extensive datasets, leading to scalability issues and degraded performance. Thus, there is a pressing need for scalable classification algorithms capable of efficiently handling the immense data volumes encountered in modern network infrastructures.

Addressing these challenges necessitates innovative approaches that leverage advanced machine learning techniques, mitigate data imbalance, ensure scalability, and optimize feature selection processes. By overcoming these hurdles, the classification systems can improve their accuracy and robustness, thereby enhancing network security against evolving cyber threats.

## 4. Feature Learning And Representation Solution

Artificial Neural Networks (ANNs) are effective machine learning replicaencouraged by the organization and utility of genetic neural networks. In the context of intrusion detection systems (IDS), ANNs play a vitalfunction in feature learning and representation, enabling the system to independently extract complexfeatures from direct network data.

### 4.1 Process of ANN in IDS

#### 4.1.1 Multi-Layered Architecture

ANNs consist of numeroussheets of intersected neurons, including an input sheet, one or supplementary hidden sheets, and an output sheet. Individual neuron in a sheet is linked to each neuron in the successivesheet, forming a network of interconnected nodes. Precisely, the end product of a neuron in a hidden sheet can be represented as in equation 1

$z_j = \sigma \left( \sum_{i=1}^{n} w_{ij} x_i + b_j \right)$ (1)

As the sigmoid $\sigma(x) = \frac{1}{1+e^{-x}}$(2)
Or ReLU (Rectified Linear Unit) $\sigma(x) = \max(0, x))$ (3)
Where$z_j$ is the product of the j-th neuron in the secreted level. σ is the initiation function, which introduces non-linearity into the network by using equations 2 or 3. $w_{ij}$ is the weightiness linking the i-th neuron in the contribution sheet to the j-th neuron in the hidden layer. $x_i$ is the input to the i-th neuron in the input layer. $b_j$ is the bias term for the j-th neuron in the unseenlevel.

### 4.1.2 Feature Extraction

Through onward transmission, the involvement information is fed into the input layer, and calculations are completed layer by layer until the product layer produces a prediction. The yield of each layer is computed as in equation 4

$$a^l = \sigma\left(W^l a^{l-1} + b^l\right) \qquad (4)$$

$$a^l = \sigma(z^l) \qquad (5)$$

Where$a^l$ is the yield of layer l. $W^l$is the weightiness template for layer l. $a^{l-1}$ is the input to layer l (output of the previous layer). $b^l$ is the prejudicedirection for layer l. σ is the triggeringevent applied component-wise to the subjective sum.

### 4.1.3 Capturing Dependencies

In a neural network with multiple hidden layers, each layer learns increasingly abstract representations of the input data. The weightiness and preferences are corrected during training to decrease the error concerninganticipated and genuine outputs. This procedure enables the complex to capture both spatial and temporal dependencies in the key data.In a neural network with multiple hidden layers, each layer learns increasingly abstract representations of the input data. This is achieved through a sequence of mathematical transformations applied to the input data. Let's denote $z_j^l$ as the outturn of the j-th neuron in the l-thsecreted layer, where $l = 1,2,\ldots,L$ represents the layer index. The output $z_j^l$ is calculated by concerning an openingoperateσ to the biased sum of inputs to the neuron in equation 6

$$z_j^l = \sigma\left(\sum_{i=1}^{n^{l-1}} w_{ij}^l z_i^{l-1} + b_j^l\right) \qquad (6)$$

Here, $w_{ij}^l$ represents the weightiness attaching the i-th neuron in the $(l-1)$-th layer to the j-th neuron in the l-th layer, $z_i^{l-1}$ is the outturn of the i-th neuron in the $(l-1)$-th sheet, $b_j^l$ is the bias term for the j-th neuron in the l-th sheet, and $n^{l-1}$ is the number of neurons in the $(l-1)$-th layer. The foundationoperate σ establishes non-one-dimensionality into the complex, acknowledging it to attaincompound patterns in the data.

### 4.1.4 Adaptive Feature Extraction

During training, the neural network adapts its internal representations to different attack strategies and evolving network conditions. This adaptability is achieved through the optimization of the network's parameters (weightiness and preferences). The parameters are modified iteratively using techniques such as inclinedecline, which aims to decrease a loss operate measuring the disparity between anticipated and definiteyields. Mathematically, the parameters are updated in the direction of steepest descent of the loss function in equations 7 and 8

$$w_{ij}^l = w_{ij}^l - \eta \frac{\partial L}{\partial w_{ij}^l} \qquad (7)$$

$$b_j^l = b_j^l - \eta \frac{\partial L}{\partial b_j^l} \quad (8)$$

Here, η represents the realizing rate, controlling the dimensions of the parameter updates, and $\frac{\partial^L}{\partial w_{ij}^l}$ and $\frac{\partial L}{\partial b_j^l}$ denote the slopes of the loss operate L with recognize to the weightiness$w_{ij}^l$ and preferences$b_j^l$, respectively. By iteratively adjusting the parameters based on the observed discrepancies between predicted and actual outputs, the neural network gradually improves its ability to differentiate between normal and anomalous network behavior, even in the presence of previously unseen attack patterns.

## 5. Optimization And Classification Solution

Support Vector Machines (SVM) are powerful and widely used supervised learning algorithms for classification tasks. They are particularly effective in scenarios where the data is high-dimensional and may not be linearly separable. SVMs work by constructing hyper planes in the characteristicdistance that best separate different sets of data spots.

### 5.1 Process of SVM in IDS

Let's consider a binary classification problem where we have a dataset $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, where $x_i$denotes the stored attributes and $y_i$signifies the matching classification markers. In SVM, the goal is to find the optimum hyperplane that splits the data into two divisions while maximizing the boundary, which is the separateinvolving the hyperplane and the nearby data point from both classes.

To formulate this mathematically, let $x_i$ be a p-dimensional input feature vector $x_i = \left(x_{i1}, x_{i2}, \ldots, x_{ip}\right)^T$, and $y_i$be the class label where $y_i \in \{-1, +1\}$. The decision function of SVM is defined as in equation 9
$f(x) = \mathrm{sign}(\sum_{i=1}^{n} \alpha_i y_i K(x, x_i) + b)$ (9)
Here, $\alpha_i$ is the Lagrange multipliers obtained through the optimization process, b is the bias term, and $K(x, x_i)$ is the kernel function, which handles the inward consequence relating the input vectors x and $x_i$ in the transformed feature space.

### 5.1.1 Sequential Minimal Optimization (SMO)

The optimization problem in SVM impliesobtaining the finesthyper plane that separates the data whilst maximizing the margin. This can be devised as a hampered optimization difficulty as in equation 10
Minimize$\frac{1}{2} || w ||^2$ Subject to $y_i (w \cdot x_i + b) \geq 1$ for all i (11)
Where w is the weightiness vector perpendicular to the hyperplane. The objective function represents the margin, and the constraints ensure that each data point is correctly classified with a margin of at least 1.To solve this optimization problem, we use Lagrange duality method. By introducing Lagrange multipliers$\alpha_i$, the constrained optimization problem can be transformed into an unconstrained one as in equation 12
Maximize $\sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \cdot \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K\left(x_i, x_j\right)$ Subject to $\sum_{i=1}^{n} \alpha_i y_i = 0$ and $\alpha_i \geq 0$ for all i (12)
This optimization problem is resolved using Sequential Minimal Optimization (SMO) procedure, which iteratively updates the Lagrange multipliers to find the optimal solution.

### 5.1.2 Margin Maximization

The concept of maximizing the margin is fundamental to SVM. The boundary is defined as the spaceconcerning the hyper plane and the closest data viewpoint from bothclasses. In SVM, the optimum hyper plane is unique that expands this margin. The space concerning a data spot x and the hyper plane can be computed as in equation 13.
Distance $= \frac{|w \cdot x + b|}{||w||}$ (13)
Maximizing the margin ensures better generalization performance and improves the classifier's ability to classify new, unseen data points correctly.

### 5.1.3 Kernel Tricks

In many real-world scenarios, the informationcould not be linearly discrete in the uniqueelementarea. SVM addresses this limitation by employing the kernel trick, which implicitly maps the entered features into aadvanced-dimensional feature area where the information may become linearly distinguishable. The choosing of kernel role determines the transformation applied to the input features.The kernel functions used are in equations 14, 15 and 16
Linear Kernel $K(x, x_i) = x^T x_i$ (14)
Polynomial Kernel $K(x, x_i) = (x^T x_i + c)^d$ (15)
Gaussian Radial Basis Function (RBF)Kernel$K(x, x_i) = \exp(-\frac{||x - x_i||^2}{2\sigma^2})$ (16)
The kernel trick allows SVM to handle complex, nonlinear decision boundaries effectively.Support Vector Machines (SVM) stand as a pivotal tool in the realm of intrusion detection classification, particularly adept at discerning patterns in high-dimensional feature spaces. In this context, SVMs operate by crafting hyper planes that optimally segregate various types of network activity, with the goal of maximizing the boundary concerning assigns. This space, the space involving the hyper plane and the closest data spot from each classification, is critical for robust classification. Mathematically, SVMs seek to find a decision function f(x) that maps input features x to class labels as in equation 9. The optimization process involves minimizing$\frac{1}{2} || w ||^2$ subject to constraints that enforce correct classification of data points. This is accomplished through Sequential Minimal Optimization (SMO) algorithm, iteratively updating Lagrange multipliers to converge towards an optimal solution. The essence of SVMs in intrusion detection lies in their ability to handle high-dimensional, nonlinear data by maximizing the margin between classes and effectively separating different types of network activity.

## 6. Ann-Based Support Vector Machine

The below are the key benefits of fusing ANN and SVM as in Figure 1 to upgrade the precision of the Intrusion detection systems.

### 6.1 Enhanced Feature Learning

This advantage refers to the capability of the ANN to get intricate attributes from raw data. ANN, through its multi-layered architecture, can autonomously learn and represent complex features present in the input data. These learned features are then utilized by the SVM for classification tasks.

### 6.2 Improved Classification Performance

By integrating ANN's feature learning capabilities with SVM's robust classification, the combined model achieves higher accuracy in classification tasks compared to using ANN or SVM individually. ANN helps in extracting meaningful features from the data, which are then used by SVM to make accurate predictions.

### 6.3 Adaptability to Complex Data Structures

ANN's ability to portray both spatial and temporal territories in the data complements SVM's effectiveness in handling complex data structures. This adaptability allows the model to handle intricate relationships and patterns present in the data, making it suitable for tasks involving complex data.

### 6.4 Scalability and Generalization

The combined approach of ANN and SVM provides scalability and generalization capabilities. ANN can handle high-dimensional data efficiently, while SVM's ability to construct optimal hyper planes makes it suitable for an expansive selection of categorization tasks. This combination grants the representation to dimension efficiently to large datasets and generalize effectively to unseen data.

### 6.5 Effective Anomaly Detection

Leveraging ANN's adaptability to evolving attack strategies and SVM's ability to separate classes effectively, the combined model is particularly effective in anomaly detection tasks. By utilizing features learned by ANN and the classification power of SVM, the pattern can efficiently differentiate between regular and irregular behavior in the data, making it suitable for intrusion detection and other anomaly detection applications.

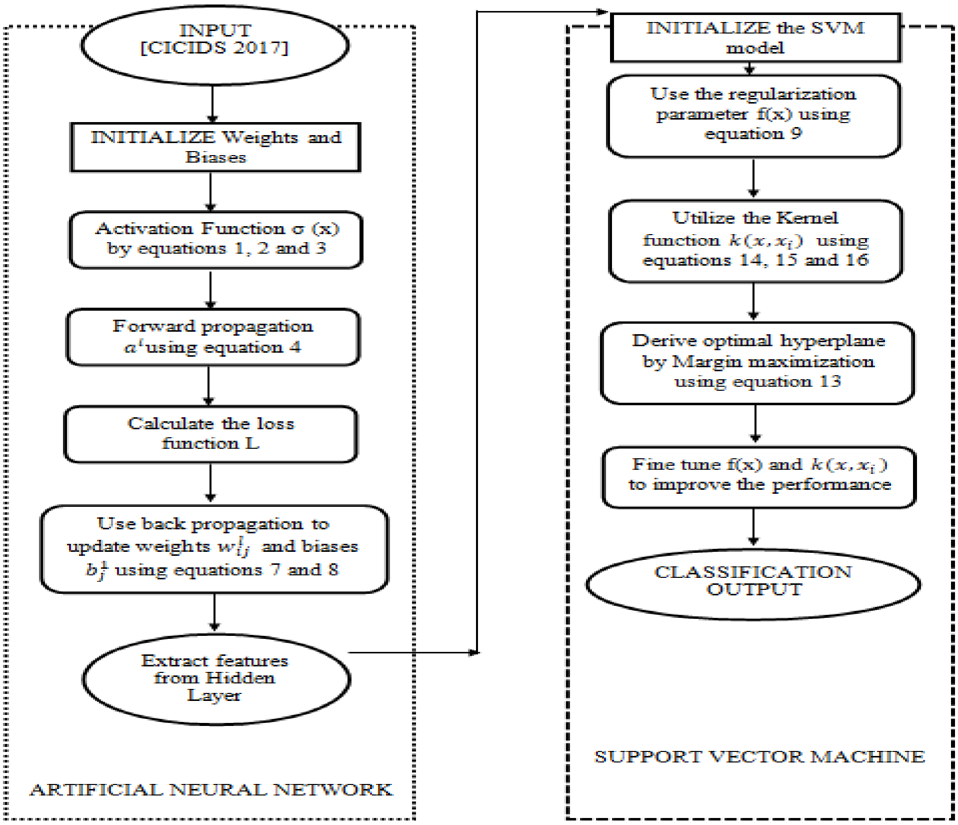
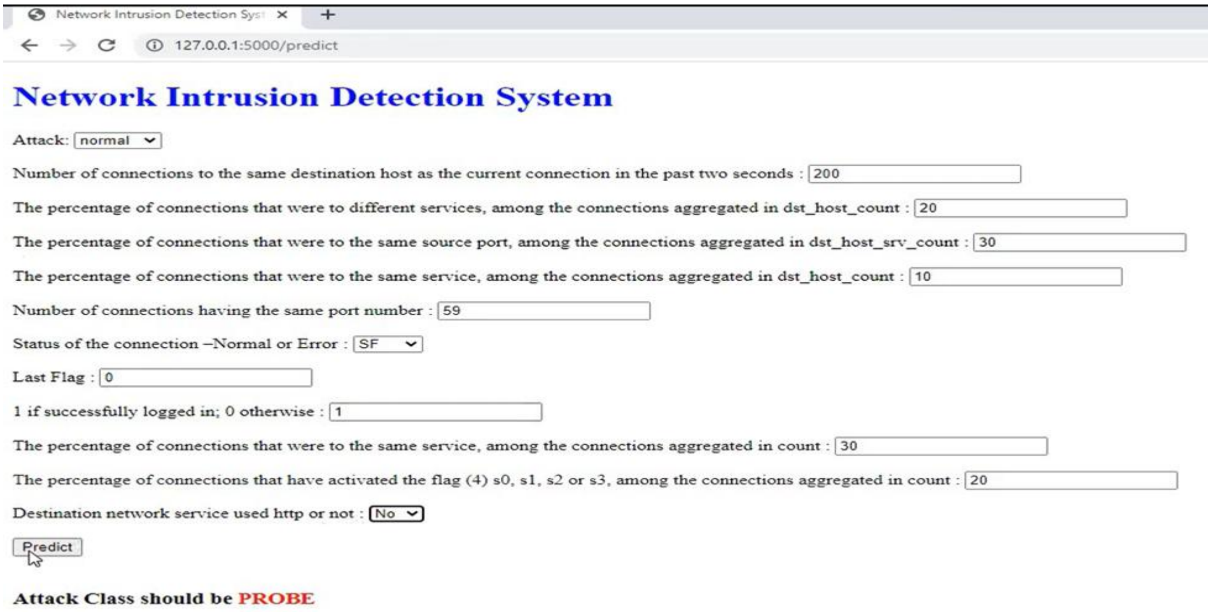

INPUT [CICIDS 2017]
INITIALIZE Weights and Biases
Activation Function σ (x) by equations 1, 2 and 3
Forward propagation $a^l$ using equation 4
Calculate the loss function L
Use back propagation to update weights $w_{ij}^l$ and biases $b_j^l$ using equations 7 and 8
Extract features from Hidden Layer
ARTIFICIAL NEURAL NETWORK
INITIALIZE the SVM model
Use the regularization parameter f(x) using equation 9
Utilize the Kernel function $k(x, x_i)$ using equations 14, 15 and 16
Derive optimal hyperplane by Margin maximization using equation 13
Fine tune f(x) and $k(x, x_i)$ to improve the performance
CLASSIFICATION OUTPUT
SUPPORT VECTOR MACHINE

**Figure 1.** Architecture of ANN Based Support Vector Machine IDS

## 7. Implementation And Results Of The Proposed System

The recommended system is employed in Python within a Linux RHEL 8 environment using the CICIDS2017 dataset, which is started by setting up the necessary environment. This involves installing Python along with essential libraries of scikit-learn, pandas, numpy, and matplotlib. These libraries will be instrumental in data manipulation, model training, and evaluation. Additionally, the CICIDS2017 dataset is downloaded and extracted to a local directory for further processing. With the environment set up, we proceeded to preprocess the dataset. Use pandas DataFrame to load the dataset and perform data cleaning tasks such as handling missing values and removing irrelevant columns. Then the dataset is splitted into features (X) and target labels (y), and applied normalization or scaling techniques to ensure uniform contributions from all features. Once the dataset is preprocessed, split it into training and testing sets to facilitate model assessment. Next, train an Artificial Neural Network (ANN) on the preprocessed data. Initialize the ANN model using TensorFlow, defining its architecture and compilation parameters as explained in previous sections. Fitting the version to the training data, stating the quantity of times and group size. Extract features from the hidden layers of the trained ANN model, which driveoblige as input for Support Vector Machine (SVM). Initialize the SVM model using scikit-learn's SVC class, providing the extracted features as input. Fine-tune the SVM model using hyper parameter tuning techniques and evaluate its performance on the testing dataset. The Figure 2 provides the implementation outputs for intrusion classifications.

The CICIDS2017 (Canadian Institute for Cyber security Intrusion Detection Systems 2017) dataset stands out as a crucial resource in the realm of IDS research and development. Crafted by the Canadian Institute for Cyber security, this dataset offers a completeassemblage of data for network traffic, covering both nonthreatening and nasty activities. Its significance lies in the diverse range of attack scenarios it covers, counting but not limited to Denial of Service (DoS), Distributed Denial of Service (DDoS), brute force attacks, Probe and botnet activities. This diversity enables researchers and practitioners to explore various intrusion detection techniques effectively. Featuring over 2.5 million records, the CICIDS2017 dataset provides scalability and robustness, catering to the demanding requirements of training and evaluating IDS models. Each traffic sample within the dataset is meticulously annotated with its corresponding attack category or benign class, facilitating supervised learning approaches for intrusion detection. This annotation ensures that the dataset is well-suited for the development and evaluation of machine learning algorithms, enabling researchers to benchmark the performance of their IDS systems accurately. The availability of such a comprehensive and well-annotated dataset empowers cyber securityscholars and experts to improve the effectiveness and reliability of IDS. By leveraging the CICIDS2017 dataset, they can explore novel methodologies, refine existing techniques, and ultimately contribute to bolstering the security posture of modern network infrastructures. Additionally, the scalability and realism of the dataset make it an invaluable asset in addressing the evolving landscape of cyber threats, thereby fostering advancements in intrusion detection technology.



Network Intrusion Detection System

Attack: normal

Number of connections to the same destination host as the current connection in the past two seconds : 200

The percentage of connections that were to different services, among the connections aggregated in dst_host_count : 20

The percentage of connections that were to the same source port, among the connections aggregated in dst_host_srv_count : 30

The percentage of connections that were to the same service, among the connections aggregated in dst_host_count : 10

Number of connections having the same port number : 59

Status of the connection –Normal or Error : SF

Last Flag : 0

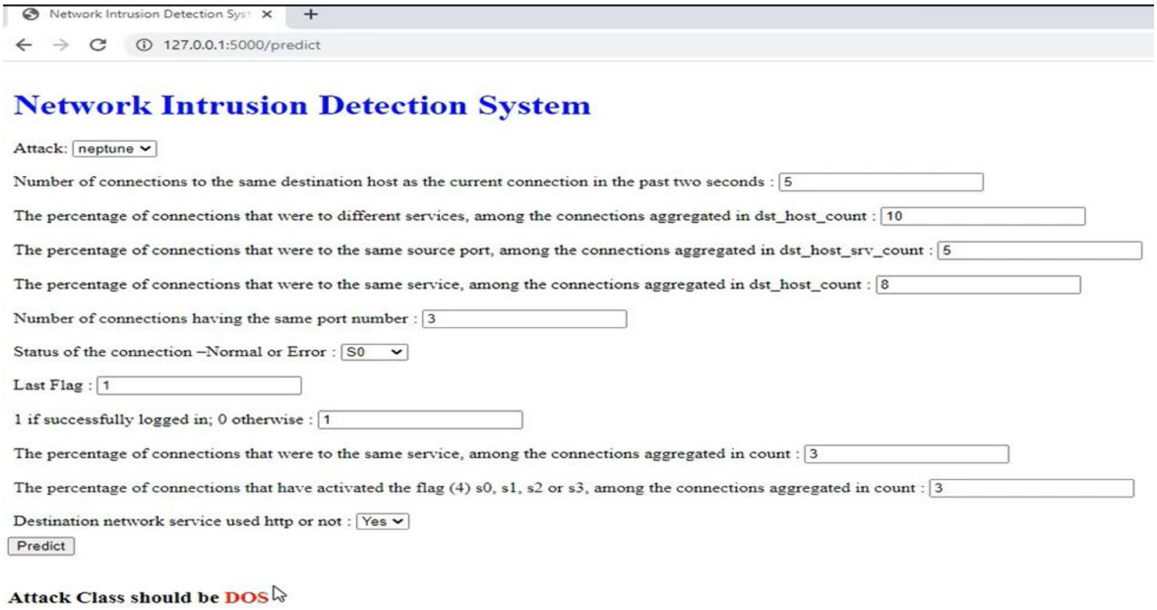
1 if successfully logged in; 0 otherwise : 1

The percentage of connections that were to the same service, among the connections aggregated in count : 30

The percentage of connections that have activated the flag (4) s0, s1, s2 or s3, among the connections aggregated in count : 20

Destination network service used http or not : No

Attack Class should be PROBE



Network Intrusion Detection System

Attack: neptune

Number of connections to the same destination host as the current connection in the past two seconds : 5

The percentage of connections that were to different services, among the connections aggregated in dst_host_count : 10

The percentage of connections that were to the same source port, among the connections aggregated in dst_host_srv_count : 5

The percentage of connections that were to the same service, among the connections aggregated in dst_host_count : 8

Number of connections having the same port number : 3

Status of the connection –Normal or Error : S0

Last Flag : 1

1 if successfully logged in; 0 otherwise : 1

The percentage of connections that were to the same service, among the connections aggregated in count : 3

The percentage of connections that have activated the flag (4) s0, s1, s2 or s3, among the connections aggregated in count : 3

Destination network service used http or not : Yes

Predict

Attack Class should be DOS

**Figure 2.** Implementation of ANN+SVM IDS classifying Probe and DOS attacks

**Table 1.** Classification Accuracy of the proposed method

| S.No | Attack Type | Number of Samples | Samples Classified Correctly | Classification Accuracy (%) |
|---|---|---|---|---|
| 1 | Brute Force Attack | 16,678 | 16,528 | 99.10% |
| 2 | DoS Attack | 56,171 | 55,329 | 98.50% |
| 3 | Probe Attack | 36,127 | 35,509 | 98.29% |
| 4 | DDoS Attack | 466,625 | 453,560 | 97.20% |
| 5 | Botnet Attack | 286,191 | 277,033 | 96.80% |

The table 2 presents a breakdown of diverse kinds of cyber-attacks, including Brute Force Attack, DoS (Denial of Service) Attack, Probe Attack, DDoS (Distributed Denial of Service) Attack, and Botnet Attack, along with corresponding statistics regarding the number of samples, samples classified correctly, and classification accuracy percentages. Each row corresponds to a specific attack type, with the "Number of Samples" column indicating the total instances of each attack type present in the dataset. The "Samples Classified Correctly" column represents the count of samples accurately classified by the detection system for each attack type, while the "Classification Accuracy (%)" column displays the share of appropriately categorized samples out of the entire samples for each attack type. These system of measurement provide understandings into the efficacy of the recognition system in precisely identifying and categorizing different cyber threats, with higher accuracy percentages indicating better performance in distinctive between type of network attacks.

## 8. Performance Comparision

The performance metrics used for the comparative analysis are explained below.

- **Precision** measures the accurateness of optimistic estimates made by the standard. A greater correctness indicates less in correct positives, meaning that when the representation forecasts a case as optimistic, it is more probable to be accurate.
- **Recall**, also known as sensitivity, measures the capacity of the simulation to fittingly detect all optimistic occurrences. A prominent recall reveals limited mistaken rejections, implication that the representation can capture more of the convinced occurrences present in the records.
- The **F1 score** is the harmonized suggest of exactness and recall, presenting a evaluate among the two system of measurement. It is principally helpful while buying with extreme datasets. A superior F1 score implies superior overall performing of the form in terms of together precision and recall.
- **Specificity** trials the capability of the representation to perfectly recognise all negative cases. It is essentially the real negative rate. A higher specificity indicates fewer false positives among the negative instances.

The figure 3 presents a comprehensive comparison of performance metrics across various machine learning models, including Enhanced Random Forest, SOA + ELM Classifier, Artificial Neural Network (ANN), Support Vector Machine (SVM), and a fused approach combining ANN and SVM. Each model is appraised based on key metrics such as precision, recall, F1 score, and specificity. Notably, the proposed fused ANN + SVM demonstrate compelling performing across most system of measurement, with high accuracy, recall, and specificity values. The SOA + ELM Classifier also performs well, particularly in terms of recall and specificity. However, the ANN model exhibits the highest F1 score, indicating a stable trade-off amongst precision and recall. The SVM model achieves competitive performance, with relatively high precision and recall values.

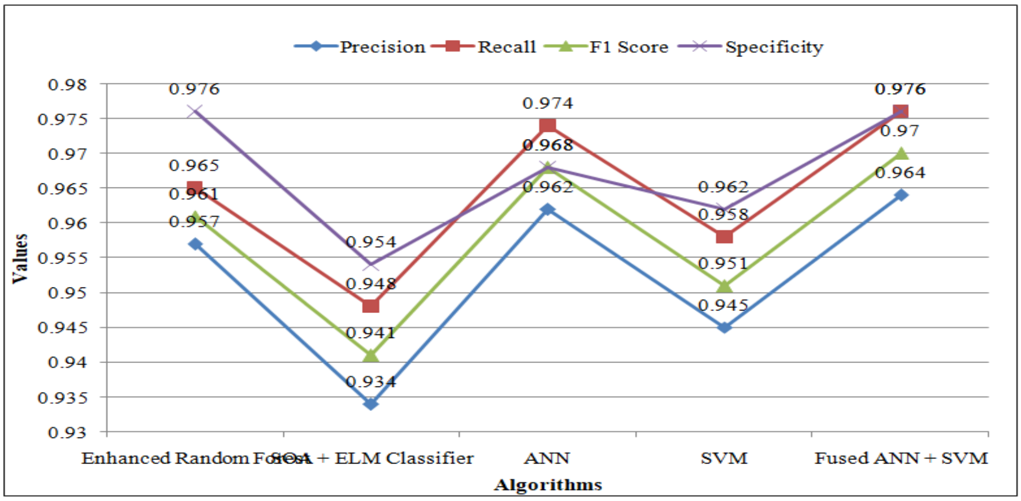


**Figure 3.** Precision, Recall, F1 Score and Specificity comparison

The figure 4 presents the accuracy values for different machine learning models employed in IDS. Accuracy is a fundamental performance metric that procedures the general rightness of expectations presented by a model. In this context, the Fused ANN + SVM approach demonstrates the highest accuracy of 97.63%, suggesting that the combined model yields slightly better overall performance compared to Enhanced Random Forest with 97.20%. On the other hand, the SOA + ELM Classifier and ANN models achieve accuracies of 94.20% and 96.80%, respectively. While these accuracies are relatively high, they are slightly lower compared to Enhanced Random Forest and Fused ANN + SVM. Additionally, the SVM model achieves an accuracy of 95.43%, which is also respectable but falls short of the top-performing models. These accuracy values provide effective perceptions into the efficacy of diverse machine learning algorithms in accurately identifying and classifying intrusions in network traffic data. Higher accuracy values indicate better performance in correctly predicting the presence or absence of intrusions, thereby enhancing the overall security posture of the system.

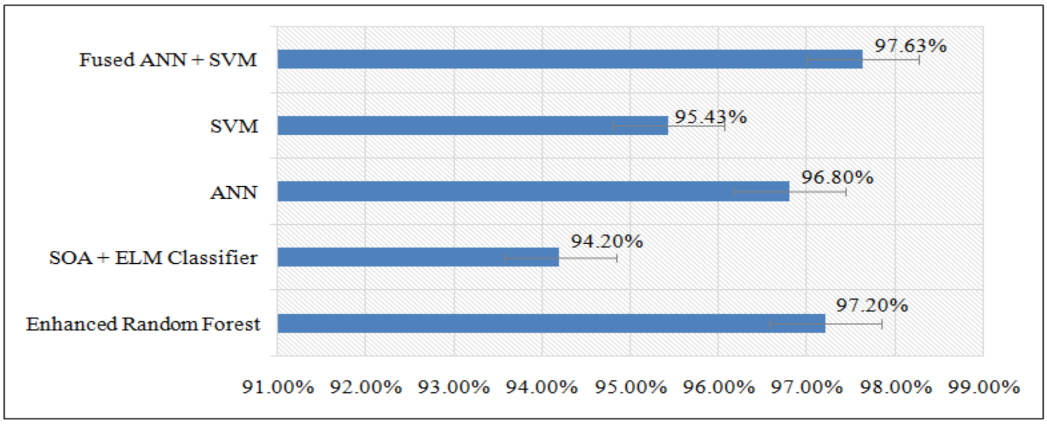


**Figure 4.** Classification Accuracy comparison

The comparison in Figure 3 and 4 highlights the strong suit and limitations of each machine learning model in the perspective of intrusion detection. While some models excel in specific metrics, such as precision or recall, others demonstrate a more balanced performance across multiple metrics. The findings underscore the significance of believing various performance indicators when evaluating the effectiveness of intrusion detection systems. Additionally, the fused approach combining ANN and SVM shows promise in achieving superior classification accuracy, suggesting potential avenues for further research and development in enhancing cyber security measures.

## 9. CONCLUSION

The research findings highlight the competence of various machine learning algorithms, including Enhanced Random Forest, SOA + ELM Classifier, ANN, SVM, and Fused ANN + SVM, in the context of IDS. The study demonstrates that these algorithms exhibit varying levels of performance in accurately classifying different types of network intrusions. Fused ANN + SVM emerge as the top-performing models, achieving high accuracy rates above 97%, indicating its robustness in identifying and mitigating potential threats. Moreover, the results underscore the importance of utilizing a combination of machine learning techniques, such as feature extraction with ANN and classification with SVM, to improve the overall presentation of intrusion detection systems. By leveraging the strengths of multiple algorithms, the fused approach demonstrates improved accuracy compared to individual models. These findings contribute to expanding the sphere of cyber security by grantingunderstandings into the efficacy of different machine learning methodologies for intrusion detection. The research underscores the significance of employing sophisticated algorithms and integrated approaches to effectively detect and mitigate network intrusions, ultimately improving the defense posture of modern digital infrastructures.

## REFERENCES

[1] Ugendhar, Babu Illuri, Sridhar Reddy Vulapula, Marepalli Radha, Sukanya K, Fayadh Alenezi, Sara A. Althubiti, Kemal Polat, "A Novel Intelligent-Based Intrusion Detection System Approach Using Deep Multilayer Classification", Mathematical Problems in Engineering, vol. 2022, Article ID 8030510, 10 pages, 2022. https://doi.org/10.1155/2022/8030510

[2] Abbas Q, Hina S, Sajjad H, Zaidi KS, Akbar R. Optimization of predictive performance of intrusion detection system using hybrid ensemble model for secure systems. PeerJComput Sci. 2023 Sep 4;9:e1552. doi: 10.7717/peerj-cs.1552. PMID: 37705624; PMCID: PMC10496009.

[3] Anurag Chhetri , Sanjay Kumar , Arya Nanda , Priyanshu Panwar, 2022. Applications of machine learning and rule induction. International Journal of Innovative Science and Research Technology, Vol.7, Issue 5, pp.1-4.

[4] Anwer, HM, Farouk, M., Hamid, AA, “A Framework for Efficient Network Anomaly Intrusion Detection with Features Selection”, IEEE International Conference on Information and Communication Systems (ICICS), 2018, pp. 157 - 162.

[5] Dini P, Elhanashi A, Begni A, Saponara S, Zheng Q, Gasmi K. Overview on Intrusion Detection Systems Design Exploiting Machine Learning for Networking Cybersecurity. Applied Sciences. 2023; 13(13):7507. https://doi.org/10.3390/app13137507

[6] Kankam, Kunrada, Prasit Cholamjiak, and Watcharaporn Cholamjiak. "A modified parallel monotone hybrid algorithm for a finite family of $\mathcal {G} $-nonexpansive mappings apply to a novel signal recovery." Results in Nonlinear Analysis 5.3 (2022): 393-411.

[7] E L Asry C, Benchaji I, Douzi S, E L Ouahidi B. A robust intrusion detection system based on a shallow learning model and feature extraction techniques. PLoS One. 2024 Jan 24;19(1):e0295801. doi: 10.1371/journal.pone.0295801. PMID: 38266011; PMCID: PMC10807775.

[8] Emad E. Abdallah, WafaEleisah, Ahmed Fawzi Otoom, Intrusion Detection Systems using Supervised Machine Learning Techniques: A survey, Procedia Computer Science, Volume 201, 2022, Pages 205-212, https://doi.org/10.1016/j.procs.2022.03.029.

[9] Engelen, G., Rimmer, V., & Joosen, W. (2021, May). Troubleshooting an intrusion detection dataset: the CICIDS2017 case study. In 2021 IEEE Security and Privacy Workshops (SPW) (pp. 7-12). IEEE.

[10] Fan, L., Liu, L., Gao, H., Ma, Z. and Wu, Y., 2021. Secure K-Nearest neighbor queries in two-tiered mobile wireless sensor networks. Digital Communications and Networks, 7(2), pp.247-256.

[11] Gautam, RKS, Doegar, A., “An Ensemble Approach for Intrusion Detection System Using Machine Learning Algorithms”, IEEE International Conference on Cloud Computing, Data Science and Engineering, 2018, pp. 1-8.

[12] Gulab Sah, Subhasish Banerjee, and Manash Pratim Dutta. "Ensemble learning algorithms with feature reduction mechanism for intrusion detection system." International Journal of Information and Computer Security 19, no. 1-2 (2022): 88-117.

[13] Harush, S., Meidan, Y., & Shabtai, A. (2021). DeepStream: autoencoder-based stream temporal clustering and anomaly detection. Computers & Security, 106, 102276.

[14] Hosseini, S., &Seilani, H. (2021). Anomaly process detection using negative selection algorithm and classification techniques. Evolving Systems, 12(3), 769-778.

[15] Jiyuan Cui, Liansong Zong, Jianhua Xie & Mingwei Tang, A novel multi-module integrated intrusion detection system for high-dimensional imbalanced data. Appl Intell 53, 272–288 (2023). https://doi.org/10.1007/s10489-022-03361-2

[16] Kumar N, Sharma S. A Hybrid Modified Deep Learning Architecture for Intrusion Detection System with Optimal Feature Selection. Electronics. 2023; 12(19):4050. https://doi.org/10.3390/electronics12194050

[17] Kumara, A., Jaidhar C.D, "Automated multi-level malware detection system based on reconstructed semantic view of executables using machine learning techniques at VMM", Future Generation Computer Systems, vol. 79, issue 1, 2018, pp. 431-446.

[18] Lohiya, R., & Thakkar, A. (2021). Intrusion detection using deep neural network with antirectifier layer. In Applied Soft Computing and Communication Networks: Proceedings of ACN 2020 (pp. 89-105). Springer Singapore.

[19] Muralidharan, J. "Advancements in 5G Technology: Challenges and Opportunities in Communication Networks." Progress in Electronics and Communication Engineering 1.1 (2024): 1-6.

[20] M. Reji, Christeena Joseph, P. Nancy, and A. Lourdes Mary. 2023. An intrusion detection system based on hybrid machine learning classifier. J. Intell. Fuzzy Syst. 44, 3 (2023), 4245–4255. https://doi.org/10.3233/JIFS-222427

[21] Maseer, Z. K., Yusof, R., Bahaman, N., Mostafa, S. A., &Foozy, C. F. M. (2021). Benchmarking of machine learning for anomaly based intrusion detection systems in the CICIDS2017 dataset. IEEE access, 9, 22351-22370.

[22] Md. Nayer & Subhash Chandra Pandey, Chromosomes identification based differential evolution (CIDE): a new bio-inspired variant for network intrusion detection, Cluster Computing volume 25, pages 3459–3480 (2022)

[23] Mendonca, R. V., Teodoro, A. A., Rosa, R. L., Saadi, M., Melgarejo, D. C., Nardelli, P. H., & Rodríguez, D. Z. (2021). Intrusion detection system based on fast hierarchical deep convolutional neural network. IEEE Access, 9, 61024-61034.

[24] Mohamad Faiz Ahmad, Nor Ashidi Mat Isa, Wei Hong Lim, Koon Meng Ang, 2022, Differential evolution: A recent review based on state-of-the-art works, Alexandria Engineering Journal, Volume 61, Issue 5, pp. 3831-3872.

[25] Moshref, M., Al-Sayyed, R. and Al-Sharaeh, S., 2022. Improving the quality of service in wireless sensor networks using an enhanced routing genetic protocol for four objectives. Indonesian Journal of Electrical Engineering and Computer Science, 26(2),pp.1182-1196.

[26] Qazi Emad-ul-Haq, Imran Muhammad, Haider Noman Shoaib Muhammad and Razzak Imran (2022). An intelligent and efficient network intrusion detection system using deep learning. Computers & Electrical Engineering. 99. 107764. DOI: 10.1016/j.compeleceng.2022.107764.

[27] Sultana, N., Chilamkurti, N., Peng, W., Alhadad, R., "Survey on SDN based network intrusion detection system using machine learning approaches", Peer to Peer Networking and Applications, vol. 12, issue 2, 2019, pp. 493–501.

[28] Tao Wu, Honghui Fan, Hongjin Zhu, Congzhe You, Hongyan Zhou &Xianzhen Huang, Intrusion detection system combined enhanced random forest with SMOTE algorithm. EURASIP J. Adv. Signal Process. 2022, 39 (2022). https://doi.org/10.1186/s13634-022-00871-6

[29] Thakkar, A., &Lohiya, R. (2021). A review on machine learning and deep learning perspectives of IDS for IoT: recent updates, security issues, and challenges. Archives of Computational Methods in Engineering, 28, 3211-3243.

[30] Thakkar, A., &Lohiya, R. (2021). Attack classification using feature selection techniques: a comparative study. Journal of Ambient Intelligence and Humanized Computing, 12, 1249-1266.

[31] Thakkar, A., &Lohiya, R. (2022). A survey on intrusion detection system: feature selection, model, performance measures, application perspective, challenges, and future research directions. Artificial Intelligence Review, 55(1), 453-563.

[32] Vijayakumar, D. S., & Ganapathy, S. (2022). Multistage ensembled classifier for wireless intrusion detection system. Wireless Personal Communications, 122(1), 645-668.

[33] Wu, T., Fan, H., Zhu, H. et al. Intrusion detection system combined enhanced random forest with SMOTE algorithm. EURASIP J. Adv. Signal Process. 2022, 39 (2022). https://doi.org/10.1186/s13634-022-00871-6

[34] Zhang, J., Vukotic, I., Gardner, R., "Anomaly detection in wide area network mesh using two machine learning anomaly detection algorithms," Networking and Internet Architecture, vol. 93, 2019, pp. 418-426.