Enhancing Network Intrusion Detection: An Investigation of Hybrid Deep Learning Approaches

Prateek Srivastava
Associate Professor, Department of Comp. Sc. & Info. Tech., Graphic Era Hill University, Dehradun, Uttarakhand India 248002

ABSTRACT

Network intrusion detection uses deep learning (DL) techniques to find hidden patterns and recognise assaults. DL's ability to simultaneously conduct feature extraction and classification tasks is another key aspect. The network's security has become more important as hostile assaults occur more often, protecting user information. A variety of preventive and detection methods are used to protect network services. A potential that may be recognised and identify any new form of assaults presents an even bigger gap for intrusion detection systems (IDS). In addition, a new challenge has been introduced by the rapid growth of information and communications technology applications. With the advent of the new technology age, passing the vast quantity of data from many sources created in a short period of time on the network is another issue since it is difficult to identify invasive behaviours in these big amounts of data and quick network speeds. A suspicious assault may be found using a variety of techniques, including artificial intelligence and machine learning. Instead of using conventional machine learning approaches, researchers are moving ahead to apply and study the DL method. With the help of DL, which is a contemporary method for working with enormous volumes of data, models for inference, decision-making, and prediction may be created from huge data by extracting important characteristics.

DOI Number: 10.48047/nq.2022.20.5.NQ22802 **NeuroQuantology 2022; 20(5): 5219-5227**

1. INTRODUCTION

A Deep Neural Network (DNN) has multilayers and additional deep layers. DL requires more inputs and complex neural layers. Three types of DL exist. Feature extraction begins with supervised feature learning. Basic machine learning algorithms are utilised to obtain these features for classification and detection. The second type of unsupervised feature learning uses the best features from the complete model. Generative feature learning is used to train hybrid deep neural networks, the third alternative. DL solves complex problems, save money, reduce data labelling, and train a large number of parameters in intelligent applications. However, it is difficult to understand, expensive to compute, requires a lot of clean data, and has more complicated algorithms. DL is used for prediction, classification, and decision-making. DL relies on feature gathering and automated extraction. Unsatisfactory outcomes were obtained today when machine learning methods were used in conjunction with expanding internet area and various assault aspects. DL methods have shown their effectiveness at choosing features.

Machine learning, growing internet space, and different attack factors yielded unsatisfactory results today. DL solves dimensionality reduction, classification, and automated feature selection challenges. DL powers speech recognition, NLP, computer vision, image processing, intrusion detection, and others. DL creates multi-layer ANN techniques. DL algorithms focus on learning and categorization. Finding correlations in massive data sets is another matter. DL has

three architectures and methods: discriminative, generative, and hybrid. Names help discriminative or supervised learning systems forecast. CNN, one of the most popular discriminative DL algorithms, is ideal for feature selection and image identification due to its intriguing design. Unsupervised learning, also known as generative learning, uses unlabeled data and learns each lower layer individually. AE, BM, and RNN are unsupervised methods. Deep hybrid approaches like DNN and GAN combine generative and discriminative techniques to benefit from both.

Information and communication technology (ICT) systems' capabilities are essential to every facet of business and daily life. A ground-breaking IDS was created as a consequence of the recent increased vulnerability of several organisations to sophisticated cyber-attacks. An IDS is a network security technique that is underutilised for spotting different types of hostile intrusions. In 1980, John Anderson was the first individual to make a substantial contribution to the identification field. All cyberattacks have financial expenses, reputational damage, and legal repercussions; as a result, the development of IDSs affects both the academic community and the commercial sector on a worldwide scale. In addition to identifying new security vulnerabilities, networks must be secured against unauthorised access and user involvement while also protecting user data. An IDS is a mechanism that can effectively enhance network or system security by detecting and preventing cyberattacks on computer networks or computer systems. IDSs are tasked with identifying potentially malicious behaviour, protecting an entire network infrastructure from the onslaught of cyberattacks, and reducing the amount of money and time lost as a result. According to the available research, an IDS is classified into one of the following three types based on its network architecture:

- Network-based IDS that look at specific packet components to identify hazardous patterns of network traffic.
- Server signature IDS, which scrutinises the system logs of a variety of hosts' activities in order to identify malicious attacks and hybrid identification methods.
- The security procedures used by systems that use anomaly- and signature-based IDS are of a higher calibre.

To more accurately evaluate harmful attacks, the signature detection approach takes use of established patterns and classifiers. Since it makes use of available knowledge to identify dangerous threats, it is known as a knowledge-based strategy. The method yields improved accuracy and a low false positive (FP), but it is unable to identify fresh network assaults. Heuristic techniques are used in the anomaly detection strategy to find hostile threats that are unknown. As a consequence, while having a high false-positive rate, this anomaly detection technique is excellent in finding abnormalities. To get around this issue, several firms have started using protocol analysis, which combines anomaly and signature-based methods. ID systems is divided into two primary categories, distributed and non-distributed, based on the deployment pattern. General basis architecture for Intrusion detection using DL technique as follow:

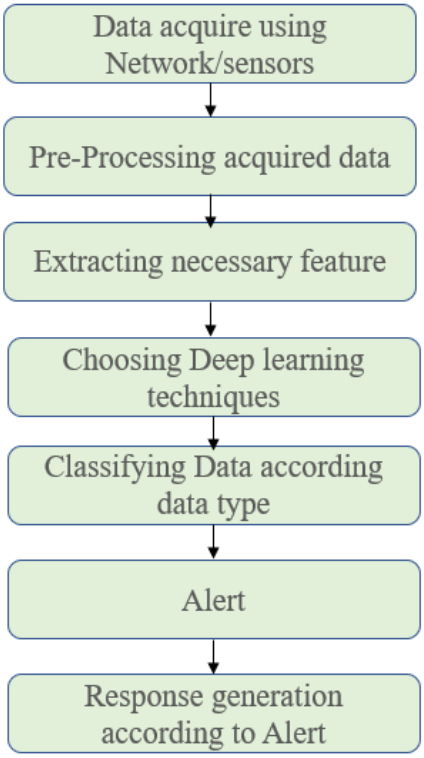

Figure 1. Basic flowchart for IDS using Deep learning

In the present day, statistical testing and threshold calculation approaches are being used in businesses to identify network intrusions. Depending on the network traffic of the model in a given period of time, the ID system based on statistical testing depends on a number of traffic limits including, the timing of packet arrival, packet length, and traffic flow volume. It's likely that these tactics won't work since harmful assaults in today's contemporary world are so complicated. It is necessary to find a solution that is highly optimised and effective to replace these statistically based methods. In order to cope with a variety of destructive assaults and avoid them, network managers have often used machine learning (ML)-based technologies.

By integrating many models into one, ensemble learning (EL) improves machine learning results. These algorithms categorise the processed data into normal and problematic categories in order to assess the network's health. These algorithms precisely test and simulate assaults to assess capabilities using a range of datasets. The vast majority of these datasets, on the other hand, are extremely imbalanced. The remaining 2% of these datasets are classified as attacks, leaving a total of 98% of them as normal. For analysing non-stationary datasets like IDS logs, a novel DL model based on ensemble learning has been developed. Better detection systems should be possible to build, particularly when using ensemble classifiers. Choosing appropriate classifiers and combiners are two crucial decisions when building an ensemble. However, the majority of conventional ML methods fall into the superficial learning category and give little consideration to the design and selection of features; as a result, they are unable to handle the substantial classification task posed by attack data in a practical network application. The accuracy of multiclassification attack detection will decline as the number of datasets increases. Because of this, machine learning and the projection needs of higher-dimensional learning with huge volumes of data are not compatible with intelligent evaluation. There are four components for both IDS and DL. The background of DL in the first section. The second part provides a description of the DL architecture and methods. Proposed system for intrusion detection are covered in the third section. The fourth segment discusses result

and analysis, and the last portion discusses conclusion.

2. LITERATURE SURVEY

The study by A. L. Buczak et al [1] offered a summary of several data mining and machine learning algorithms for cybersecurity intrusion detection. Both the basic ideas of intrusion detection and the most recent advancements in IDS are covered. The article examines several machine learning and data mining approaches used for intrusion detection, as well as numerous kinds of IDS and their categorization. The writers contrast different methods and point out the advantages and disadvantages of each strategy.

J. A. Khan et al [2] focused on the categorization methods used by IDS. It provides a thorough explanation of the various categorization methods and how they are used in intrusion detection. The authors provide a thorough evaluation of each categorization method, stressing its advantages and disadvantages. The different IDS and their categorization are also covered in the study.

A network IDS based on support vector machines (SVMs) and the Ant Colony Optimisation (ACO) method for feature selection was proposed by T. Mehmood and H. B. M. Rais in [3]. The suggested solution uses SVMs to categorise network traffic as benign or harmful and ACO to choose the best attributes to include in the SVM model. The results demonstrate how well the suggested approach detects network issues.

A feature augmentation strategy was put out by H. Wang et al [4] to enhance the performance of SVM-based IDS. To improve the performance of feature vectors used in SVM models, the suggested technique makes use of a brand-new feature augmentation algorithm. The study presents experimental findings that show how the suggested strategy might increase the precision of SVM-based IDS.

Using a deep auto-encoder-based methodology, F. Farahnakian et al [5] suggested a method for detecting intrusions. In order to identify anomalous traffic, the suggested technique employs deep neural networks to understand the fundamental characteristics of network traffic data. The experiment findings in the research show how well the suggested technique works at spotting different kinds of network assaults.

Specifically, Salur MU et al. [6] created a hybrid DL model for emotion classification using LSTM models and convolutional neural networks (CNN). The proposed model improves classification accuracy over existing common machine learning methods and DL models.

A hybrid anomaly classification method for IDS that combines DL and binary algorithms (BA) was suggested by Atefi K, et al [7]. The suggested approach uses supervised and unsupervised learning strategies to increase IDS accuracy. The results show that the suggested model works better at identifying network abnormalities than other approaches already in use.

In order to identify network attack, Xiao Y et al [8]. devised an intrusion detection model that makes use of feature reduction and CNN. Principal component analysis (PCA) is used in the proposed methodology to decrease the dimensionality of network traffic data before a CNN-based classification model is used to identify intrusion attempts. According to the experimental findings, the suggested model obtains a high level of detection accuracy.

A cloud-based network IDS that uses DL methods to evaluate network traffic in real-time was suggested by Parampottupadam S et al [9]. The suggested approach employs a multilayer perceptron (MLP) for classification and a stacked autoencoder (SAE) to learn the underlying properties of network data. The results of the experiments show that the suggested model can successfully identify network intrusions in real-time.

A deep belief network (DBN) is used to learn the characteristics of network traffic data in a DL method for NIDS developed by Niyaz et al [10]. When compared to conventional machine learning algorithms like decision trees and support vector machines, the proposed model is more accurate. The results show that the suggested approach is capable of efficiently identifying network intrusions.

3. PROPOSED MODEL

Deep learning algorithms (DL) are a subset of machine learning (ML) that use several hidden layers to simulate a deep neural network. Due to their deep structure and capacity to independently learn the key features from the dataset and produce an output, these techniques are more effective than ML. Following are several steps which are carried out for executing the proposed system;

1. Gather instances of both regular traffic and other kinds of network assaults in a collection of network traffic data.
2. Prepare the data for the DL model by transforming it into an appropriate format. Techniques for data augmentation, feature extraction, and normalisation may be included.
3. Divide the data into sets for testing, validation, and training.
4. Create a hybrid deep model that integrates recurrent neural networks (RNNs) and autoencoders with CNNs and other DL architectures. The Keras or PyTorch libraries may be used to do this.
5. Use an appropriate optimisation technique, such as stochastic gradient descent (SGD) or Adam, to train the hybrid model on the training data. A appropriate loss function, such as binary cross-entropy or categorical cross-entropy, should be trained into the model to minimise it.
6. Assess the model's performance on the validation set and adjust the model hyperparameters as needed.
7. Run the finished model through the testing set to see how well it performs with omitted data.
8. Use the model to identify intrusions in a real-world setting.
9. Continue to monitor the model's performance and adjust or retrain it as appropriate to increase accuracy and make it more flexible to changes in the network environment.

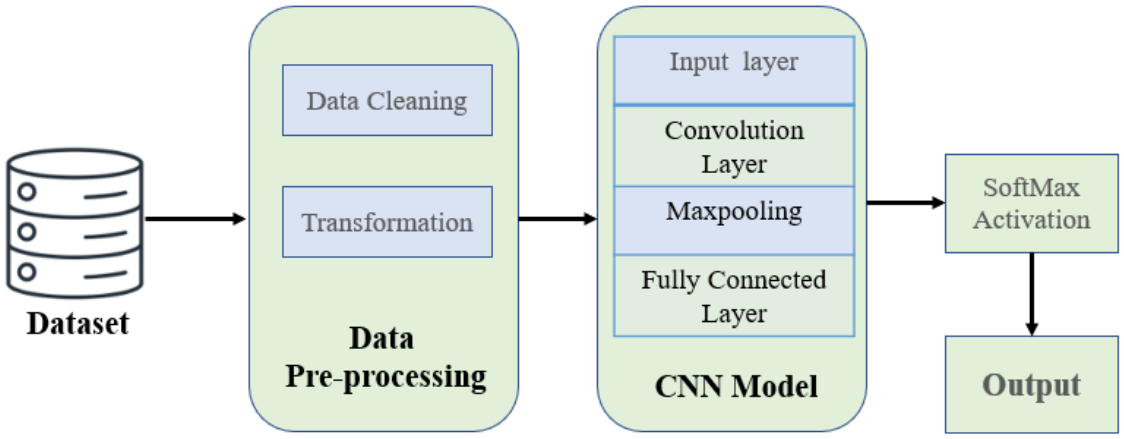

Figure 2. Architectural Flow diagram of proposed model

Algorithm 1: Convolutional Neural Network

When dealing with data that is organised in arrays, the structure of DL known as the CNN works the best. The classification layer is comprised of an input layer, a feature extraction stack that includes convolutional and pooling layers, a fully connected layer, and a softmax classifier. In the field of computer vision, CNN has made significant strides in recent years. In order to do supervised feature extraction and classification, the IDS makes use of them. They can be broken down into three primary sorts of layers:

-Fully-connected (FC) layer,
-pooling layer, and
-convolutional layer

The first layer of a CNN is referred to as the convolutional layer. The fully-connected layer is always the final layer of a CNN, regardless of the number of convolutional layers, further convolutional layers, or pooling layers that come after it. In the initial levels, there is a focus placed on the fundamentals such as colour and borders.

Convolutional Layer

CNNs process most data at the convolutional layer. Start with data, a filter, and a feature map. A three-dimensional pixel matrix representing a colourful image is the input. Like an image, the input will have height, width, and depth. Height, width, and length. Feature detectors, sometimes called kernels or filters, iteratively scan the image's receptive fields to detect features. "Convolution" describes this process. 2-D weighted array to represent and detect visual characteristics. Usually a 3x3 matrix, the filter's size and receptive field can vary. After filtering an image area, the output dot product of the input pixels and the filter is calculated. The output array receives this dot product. After the kernel scans the image, the filter advances one step and starts over. Dot products on the input and filter produce a feature map, activation map, or convolved feature.

CNNs apply Rectified Linear Units (ReLUs) to the feature map after each convolution to account for nonlinearity. After the initial convolution layer, more can be added. The CNN's structure can become hierarchical when later layers can perceive pixels in earlier levels' receptive fields. Let's say we're looking for a bicycle in an image. Bicycles are disassembled. Bicycle parts include the frame, handlebars, wheels, and pedals. CNN features are sorted by bicycle parts. The bicycle symbolises a more complicated neural network pattern than these portions.

Pooling Layer

Dimensionality reduction is achieved through down sampling, which can also be referred to as pooling layers. This technique reduces the amount of factors that are included in the input. In a manner analogous to that of the convolutional layer, the pooling operation applies a filter to the entire input; however, in contrast to that layer, this filter does not make use of weights. Instead, the kernel populates the output array by applying an aggregation function to the values in the receptive field in order to accomplish this task. Pooling can often be broken down into one of two categories:

• Maximum pooling: As it moves across the input, the filter selects the input pixel that has the highest value to send to the output array. This occurs as it applies the filter. In passing, I might mention that this form of pooling is utilised more frequently than the standard one.

• Calculating the average value inside the receptive field as it travels across the input and sending that value to the output array is what's referred to as "average pooling," and it's performed by the filter.

Although the pooling layer causes the loss of a great deal of information, it does provide the CNN with a number of benefits. They simplify the process, make it more effective, and lessen the risk of overfitting at the same time.

Fully Connected Layer

The full-connected layer is precisely what one would expect from its name. As was mentioned previously, partially linked layers do not have a relationship that is direct between the pixel values of the input image and the values of the output layer. In contrast, every node in the output layer of the fully-connected layer is directly connected to a node in the layer above it. This is the case with the fully-connected layer.

The classification process is carried out by this layer using the attributes that were gathered from the layers that came before it as well as the different filters that were applied to those attributes. In order to correctly classify inputs, FC layers frequently make use of a softmax activation function, which results in a probability that can range anywhere from 0 to 1. ReLu functions are usually utilised in the convolutional and pooling layers.

Algorithm 2. Recurrent neural networks

Recurrent neural networks (RNN) augment feed-forward neural networks to simulate sequence data. RNNs have input, hidden, and

output units. Hidden units store memory. Each RNN unit decides using its current input and a preceding input. Voice processing, handwriting analysis, semantic comprehension, and human activity recognition use RNN. IDS feature extraction and supervised classification can use RNN. RNNs have weak short-term memory and struggle with long sequences. RNN versions like LSTM and GRU have been proposed to solve these issues. RNN-based IDS classifies binary and multiclass NSL-KDD datasets. Hidden node counts and learning rates were tested on the model. The model's accuracy depended on hidden nodes and learning rates. 80 hidden nodes and 0.1 and 0.5 learning rates gave the greatest accuracy for binary and multi-class instances. The suggested model outperformed ML approaches and a Reference condensed RNN model. This research's main issue is increased computational processing, which slows model training and lowers R2L and U2R detection rates. The study does not compare the suggested model to other DL techniques.

Deep neural network

A basic DL structure called DNN makes it possible for the model to learn in layers. It has many hidden levels and an input layer, an output layer, and an input layer. DNN is used to show complicated functions that don't follow a straight line. By adding more hidden layers, which raises the model's abstraction level, the model's powers get better. As the function that turned on the hidden layer, a rectified linear unit was used. Results showed that the suggested model is strong because it increased detection rates for almost all types of attacks, except for U2R, for which there were fewer data. The writers say that adding nodes and layers makes a structure that is more complicated, uses more resources, and takes longer to compute. These problems can be fixed with the tuning method and automatic optimisation.

4. RESULT ANALYSIS

In below line graph chart, we can linear line is Users and dots are presenting number of packets in intrusion detection analysis done using open-source dataset.

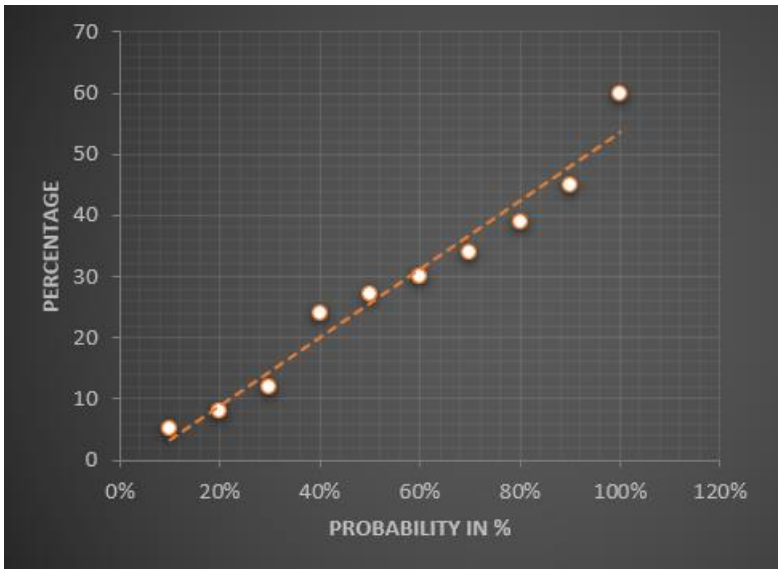

Figure 3. Probability check for Intrusions Packet

Figure 4 shows the proposed hybrid DL model accuracy and loss comparison graph. The hybrid DL algorithm achieved 99.02 % of accuracy.

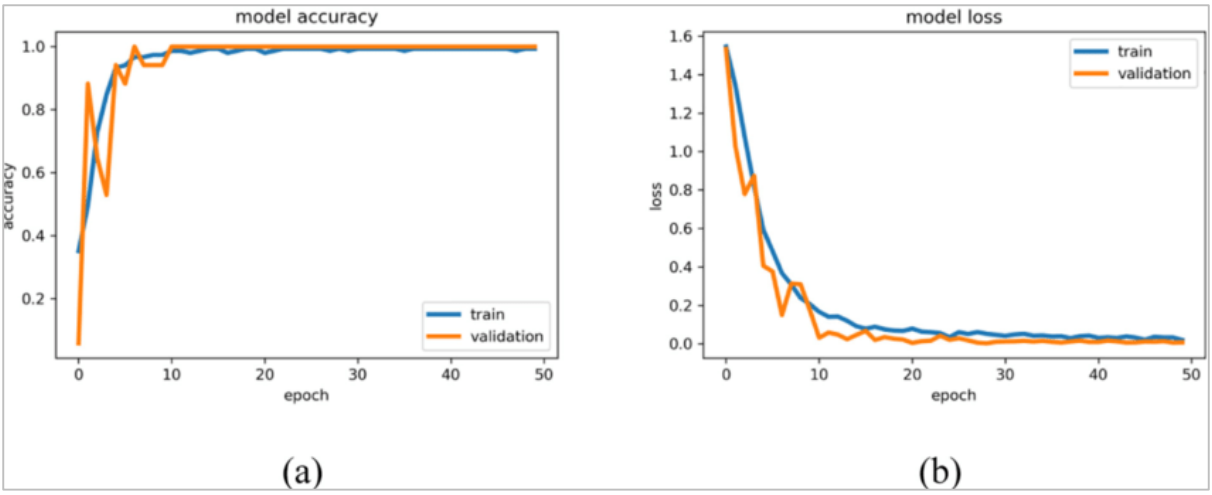

Figure 4. Accuracy and Loss Comparison Graph of Hybrid Deep Learning Model

5. CONCLUSION

Despite its benefits, the proposed CNN deep learning model requires more training time than common DL strategies like a single model because of its complexity. In the future, this research may be enhanced by deploying the model in real-time to classify network traffic and utilising a different optimisation and feature selection technique. We want to implement the distinction between regular and attack types of incoming traffic. Network assaults are now the most important problem confronting contemporary civilization. Regardless of scale, all networks are prone to network dangers. An intrusion detection system (ID) is a critical component of a network's defence against malicious assaults. A very sophisticated security solution is needed for the network since hostile threats are always evolving and emerging. When hybrid DL used to identify fraudulent traffic, different changes in the traffic may be recognised, improving performance and letting only legitimate traffic into the system. Using a variety of DL approaches, we wish to build more effective algorithms for other harmful network traffic in order to extend our model with additional factors that contribute to improved performance and detection. As our first area for development, we'll examine and increase our model's resistance against zero-day attacks. Additionally, we will try to improve upon our current evaluations by using actual backbone network traffic to demonstrate the usefulness of the expanded model.

REFERENCES

[1] A. L. Buczak and E. Guven “A survey of data mining and machine learning methods for cybersecurity intrusion detection,” IEEE Communications Surveys and Tutorials, vol. 18, no. 2, pp. 1153–1176, 2016.

[2] J. A. Khan and N. Jain, “A survey on intrusion detection systems and classication techniques,” International Journal of Scientific Research in Science, Engineering and Technology, vol. 2, no. 5, pp. 202– 208, 2016.

[3] T. Mehmood and H. B. M. Rais, ‘‘SVM for network anomaly detection using ACO feature subset,’’ in Proc. Int. Symp. Math. Sci. Comput. Res. (iSMSC), May 2015, pp. 121–126.

[4] H. Wang, J. Gu, and S. Wang, ‘‘An effective intrusion detection framework based on SVM with feature augmentation,’’ Knowl.-Based Syst., vol. 136, pp. 130–139, Nov. 2017.

[5] F. Farahnakian and J. Heikkonen, ‘‘A deep auto-encoder based approach for intrusion detection system,’’ in Proc. 20th Int. Conf. Adv. Commun. Technol. (ICACT), Feb. 2018, p. 1.

[6] Salur MU, Aydin I. "A novel hybrid deep learning model for sentiment classification," IEEE Access. 2020; 8:58080-58093

[7] Atefi K, Hashim H, Khodadadi T. "A Hybrid Anomaly Classification with Deep Learning (DL) and Binary Algorithms (BA) as Optimizer in the Intrusion Detection System (IDS)," in 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA). 2020;29-34.

[8] Xiao Y, Xing C, Zhang T, Zhao Z. "An intrusion detection model based on feature

reduction and convolutional neural networks," IEEE Access. 2019; 7:42210- 42219.
[9] Parampottupadam S, Moldovann A-N. "Cloud-based real-time network intrusion detection using deep learning," in 2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security). 2018;1-8.
[10] Q. Niyaz, W. Sun, A. Y. Javid and M. Alam, “A deep learning approach for network intrusion detection system,” in Proc. 9th EAI Int. Conf. on Bio-Inspired Information and Communications Technologies, New York City, USA, pp. 21–26, 20