# Efficient Intrusion Detection and Classification Using Enhanced MLP Deep Learning Model

G. Gowthami(✉) and S. Silvia Priscila

Department of Computer Science, Bharath Institute of Higher Education and Research, Chennai, India
Gowthami.ramya@gmail.com,
Silviaprisila.cbcs.cs@bharathuniv.ac.in

**Abstract.** Everyone has entered a new stage of the digital world during this era. The digital world has created numerous opportunities and facilities, but it has also become a threat to the data that is kept there. Internet security is seen by many enterprises as a major challenge. Organizations use a variety of methods, including firewalls, virtual private networks (VPNs), authentication, and encryption, to protect credential data. The primary goals are to protect network infrastructure security and internet communication security. The arsenal of technologies for securing security has expanded. One of the most recent advancements in security technology is intrusion detection. In this study, EMLP (MLP + PSO) is used to identify and categorise paper incursions which is been compared over ANN and MLP (MLP + PSO) deep learning models. The dataset used for analysis is KDD CUP99 dataset. Among these models, Enhanced MLP (MLP + PSO) produces better outcomes in terms of accuracy of about 93%, precision of about 0.88, and recall of about 0.84 respectively. The tool used for analysis is python.

**Keywords:** Intrusion · Attacks · Anomaly · Accuracy · Networking · Hacker · Security · Digital services

## 1 Introduction

As more networked digital devices adopt various technologies on a daily basis, the risk of various forms of attacks by large numbers of attackers also rises. It becomes necessary to bridge the gap between the attack and the data's safety.

The network administrator may identify suspicious online activity with the use of IDS, and it also alerts the administrator to secure the data by launching the necessary defences against those threats. In networking, the phrase intrusion refers to any hostile application of information or unauthorised access of any type. Attackers or invaders are those who intend to obtain unauthorised access to guarded data. They cause damage to the information stored by their malicious activities [3].

The most important requirements for effective IDS have been identified by researchers as ML (Machine Learning) and DL approaches (Deep Learning). Both of the IDS methodologies are branches of AI (Artificial Intelligence), which tries to analyse


S. Rajagopal et al. (Eds.): ASCIS 2022, CCIS 1759, pp. 208–219, 2022.
https://doi.org/10.1007/978-3-031-23092-9_17

the critical facts concealed within huge data. Due to the development of powerful GPUs over the past ten years, the aforementioned methodologies have gained popularity in the field of network security (Graphics Processor Units). With the help of robust technologies like ML and DL, it is possible to discover the best characteristics of network traffic. Additionally, it is employed to estimate the typical and exceptional works pertaining to the ingrained patterns. The ML-based IDS technique relies on feature engineering, which is used to study the useful information present in the network traffic whereas the IDS based on DL doesn't depend on feature engineering and automatically studies the complex attributes received from raw data because of its strong structure [7].

## 2 Literature Review

The intrusion detection system based on DL proposed by Akhil Krishna et al., 2020, effectively detects malicious attacks including Probe, DOS, U2R, and R2L coupled with attack prevention. The MLP, which is trained using high accuracy dataset, is used to detect intrusions using the DL model. A CSV file containing the relevant data is downloaded from the network and added to the used DL model in real-time intrusion detection, obtaining the detection as a result. The second stage involves running a script in the background to prevent the incursion. In the script, a precise choice is made to protect the data from numerous assaults during the prevention phase. The Multi-Layer Perceptron model is used to make judgments using the data that was gathered from the classification part. The proposed combination approach, which combines IDS and IPS, accomplishes the objective of intrusion prevention and detection rapidly and effectively [1].

The 2017 study by Jin Kim et al. focuses on the application of AI-based IDS with the DNN model to effectively detect attacks. The given data were extracted into samples for the investigation. The entire dataset, which contains around 4.9 million records, was used for the process of confirming the testing set, while the training data used for the study only includes 10% of the corrected data. The findings from the suggested model demonstrate an astonishingly high rate of 99% accuracy in detecting the attack. Additionally, the rate of false alarms was found to be roughly 0.08%, showing how infrequently routine data was mistaken for a danger. The investigations that have been done so far have only examined a single traffic data point. To combat Distributed Denial of Service (DDoS) assaults, time series data analysis using the recurrent neural network (RNN) model and the LSTM (Long Short Term Memory) model is required [2].

S. Santosh Kumar et al., 2021 organised a successful study to achieve the goal of detecting intrusion, different methods used for intrusion detection, a significant category of hacker attacks, a variety of tools and techniques used to protect, the area of research to improve the efficiency of identifying intrusion, challenges faced by the user, and finally the evolution of new IDS tools in terms of research. The instruments that were created have the ability to recognise and guard against incursion brought on by intruders. IDS has been integrated into many businesses that want to safeguard their digital data within the network perimeter since the introduction of firewall technology. IDS has been recognized as a competent technology to protect valuable information from internal attackers and external hackers where the traffic doesn't move ahead of the firewall itself [3].

The implementation of the DL technique using binary classification to separate the regular, legitimate message packets from the suspicious message packets has been the focus of Vipparthy Praneeth et al., 2021. The method begins with the creation of a training dataset made up of 1,20,223 network packets with 41 attributes that were taken from the open-source CICIDS 2018 and KDD 99 datasets. The chosen one-dimensional network dataset is first pre-processed by using an autoencoder to remove any extraneous data. Out of forty-one qualities, 23 were deemed deserving. The structured DNN is used to assess the suggested model. Additionally, it is combined with Softmax classifier and Relu activation algorithms. The proposed intrusion deterrent approach has been researched and tested using Google Colab, a non-proprietary tensor flow and an open forum for cloud services. Using a simulation dataset produced via network simulation, the recommended intrusion prevention classifier model was validated. The experimental findings show an accuracy of 99.57%, which is the highest level among the current CNN and RNN-based models. In a short time, the method can be studied with various datasets that result in the process of improving the precision and effectiveness of the suggested model [5].

Vijayakumar R et al., 2019, offer their opinion on how to choose the best method for effectively identifying upcoming cyberattacks. The use of existing DNNs and other traditional ML classifiers on a variety of benchmarked malware datasets that are publically accessible has been thoroughly evaluated. The network topologies specifically for DNNs and the most important network parameters were determined using the hyper-parameter methods of selection in relation to the KDD Cup 99 dataset. The trials used DNNs with a training rate range of 0.01 to 0.5 over a period of 1000 epochs. To create a benchmark, the well-performing DNN model from the KDD Cup 99 dataset was applied to the UNSW-NB15, NSL-KDD, WSN-DS, CICIDS 2017, and Kyoto datasets. The recommended DNN model studies the high-dimensional attribute recognition of the IDS data by transferring them into various hidden layers. This method is recommended for real-time usage ineffective monitoring of the network traffic and it also assists the host-level events by sending early alerts of the cyberattacks that are possible [6].

M. Azhagiri et al., 2015 render an outlook of IDPS techniques. The research on IDPS summarizes all the key functions utilized in performing the IDPS techniques and it also evaluates the detection methodologies utilized by it. In addition, the technique embosses the significant properties of every chief class of the IDPS system. The paper has elaborated on the different kinds of IDPS security competencies, the challenges identified, and the limitations of the technique. Securing the sensitive information stored in a computer has become a legal concern by organizations because of the growing trust in electronic transactions and operations. Many methods are utilized to support the organizations intending to protect their information against attacks [8].

Now a day, the blending of methods has become very popular which leads to confusion in choosing the apt methodology that deploys to secure the system. David Mudzingwa et al., 2012 demonstrate and provide a clear clarification of every methodology along with the way to compare the performance and efficiency of each methodology. The research has considered four important methodologies which are used to detect and prevent systems from intrusion. Even though the anomaly-based methodology performs well than the other methods in the process of detecting new attacks without any updates

in the software or without any change fed by the user, the current network world moves for IDPS makes a mixture of four chief methodologies. The research also concentrates on the comparison and estimation procedure of the IDPS methodologies which are utilized by the IDPS products in the market [9].

Due to the nature of the network, security has become a threat to utilizing a wireless network. With this concern, many researchers are thinking to find a suitable and reliable solution for the issue. Jafar Abo Nada et al., 2018 propose a novel IDPS, particularly for a wireless system. The study discusses the framing of a new system, especially for a wireless network, and named it WIDPAS. The framework is designed to perform three important tasks such as monitoring, analyzing, and defending. The system initially monitors the network for any type of false networks or service attacks, then it analyses the characteristics of the attack and finds out the intruder. Finally, the system safeguards the network users. The research paper demonstrates the method to improve the effectiveness of the prevailing system. The proposed system monitors all the attacks and it is equipped to defend against any type of attack from counterfeit networks. It cut the path of the attacker and safeguards the user from being scammed. The future of network systems is developed with AI and ANN. A large amount of data is involved in every transaction. Therefore, it is essential to monitor, analyze and study the data and secure it safely from intruder attack [10].

## 3 Proposed Methodology

IDS has become a highly propounded approach that aids in securing digital data. The major change occurs in the collected data set that includes different samples of intrusion algorithms similar to denial of service, brute force, or at least an infiltration taking place within a network. Figure 1 depicts a passive deployment of NIDS which is associated with a network switch that is configured with the technology of port mirroring.

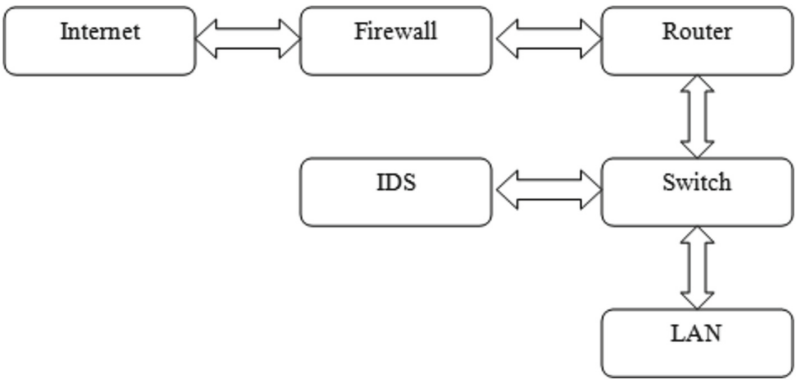


**Fig. 1.** Network-based IDS

**Artificial Neural Network (ANN)**

The single layer of ANN consists of multiple nodes that perform the actual computation. The computation process in ANN is designed in such a way that it reflects the actions of neurons present in the human neural network. The nodes in the neural network react to the stimulus at a particular degree or extensively. The magnitude of the reaction is related proportionally to the feed value multiplied by the node weight. The nodes commonly consist of weights achieved from multiple inputs. Different weights are extracted by modifying many inputs. Finally, each multiplied value was summed up and the added value is fed into the activation function. At last, its outcome is implemented into the regression or classification analysis. The application of the ANN approach has increased in recent times and it is used for various categories, reasoning, prediction, and demonstration of positive results for endorsement [2].

The familiar quality of the objective kind method contains two components, training type loss, and the regularization component.

$$obj(\theta) = L(\theta) + \Omega(\theta) \tag{1}$$

From the above formula (1) $L$ describes the training loss mode and $\Omega$ represents the regularization component. The loss value computes the system's capability to project information. The common option of the value $L$ is MSE, which is described by

$$L(\theta) = \sum\nolimits_i \left(y_i - \hat{y}_i\right)^2 \tag{2}$$

Another type of common loss function is logistic kind loss. It is expressed as

$$L(\theta) = \sum_i \left[ y_i \ln\left(1 + e^{-\hat{y}_i}\right) + (1 - y_i)\ln\left(1 + e^{\hat{y}_i}\right) \right] \tag{3}$$

ANN model is composed mathematically as portrayed below:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in \mathcal{F} \tag{4}$$

In formula (4) K denotes the total quantity of layers, function space represents the symbol, and a cluster of possible classifications is depicted by $\mathcal{F}$. Layer complexity of the value $\Omega$(f), and the description of the layer f(x) is expressed by

$$f_t(x) = w_{q(x)}, w \in R^T, q : R^d \to \{1, 2, \ldots, T\} \tag{5}$$

Here T denotes the definite number of layers, and q portrays the method to assign each data point equivalent to the leaf information. The MLP model complexity level is usually described as

$$\Omega(\mathrm{f}) = \gamma T + \frac{1}{2}\lambda \sum_{\mathrm{j}=1}^{\mathrm{T}} \mathrm{w}_{\mathrm{j}}^2 \tag{6}$$

**Multi Layer Perceptron (MLP)**
MLP is based on the concept of evaluating the values of neurons present in the current layer in the form of activated summation of the weighted outcomes of the neurons from the previous layer that is connected to each neuron. The term activation infers to the sum of weighted inputs that are used as inputs for the above-termed activation function for mapping the input to the output performed either directly by identity activation or with certain restrictions like tanh or sigmoid or mapping it during the removal of unwanted values. The process is illustrated with an example of ReLU that removes the negative values along with the direct mapping with positive ones. At the initial stage, the weights of the neuron links are random. Some adjustments are done using the backward propagation process where the error found for the forward propagation of the MLP outcomes receives back-propagated through, and the weights are altered proportionally according to the error.

Commonly MLP model is trained in a supervised manner, with a backpropagation technique to measure the weight derivatives. Here the error function $E$ is illustrated as:

$$E = \sum_{k=1}^{n} d^{(k)} - y^{(k)} \tag{7}$$

In the above equation, d describes the target value and y denotes the MLP output-based vector. After measuring the error value E, the formulas are used to keep posted the bias value θ and the value of the weight w.

$$w_{new} = w_{prev} - \eta \frac{\partial E}{\partial w_{prev}} \tag{8}$$

$$\theta_{new} = \theta_{prev} - \eta \frac{\partial E}{\partial \theta_{prev}} \tag{9}$$

From the Eq. (8) and (9) η shows the learning value, and $d^{(k)}$ describes the position of the target vector. θ denotes the weight value used in the learning process, the identifier w manages the weight, and y indicates the output vector information.

**Enhanced Multi Layer Perceptron (EMLP)**
The main reason for the selection of MLP as the technique for utilizing in research was to ease the implementation process of such techniques. MLP technique, which is known for rendering high-quality models it keeps the time needed for training relatively lower than the other compared complex methods.

**PSO (Particle Swarm Optimization)**
In continuation to it, the position and velocity updates are done, as the values of every particle are as discrete numbers with the utilization of the sigmoid equation for updating the position concerning the necessary velocity based on the following Eq. 10.

$$v_n = w * v_n + c_1\, rand_1 * (Pbest - x_n) + c_2\, rand_2 * (Gbest - x_n) \tag{10}$$

From the Eq. 10 $v_n$ indicates the n[th] particle velocity, $w$ describes the inertial weight, $rand_1$ and $rand_2$ describes the random value between 1 and 0, and $x_n$ indicates the present position.

$$sigmoid = \frac{1}{1 + e^{-v_n}} \tag{11}$$

$$x_{i,j}(t+1) = \{x_{i,j}(t+1) = rand(B_i), rand(t) < sigmoid\big(v_{i,j}(t+1)\big)x_{i,j}(t+1) = x_{i,j}(t+1), rand(t) < sigmoid\big(v_{i,j}(t+1)\big) \tag{12}$$

To move the particle into a new position, a sigmoid equation is needed, when a velocity is added to the earlier position which is a discrete number that results in a continuous number that doesn't work in the prevailing case. Therefore, to search for a new position concerning the velocity, the positions are mapped into the sigmoid equation. After that, the outcomes are compared with random numbers ranging from zero to one.

It provides more extra weight values of the weak observation. The DS method can be illustrated using the following equation.

$$f(x) = s(x_k > c) \tag{13}$$

$f(x)$ will create a forecasting value 1 when the component of K of the x vector is advanced the threshold value c and the value will be set as $-1$. The value of s is $-1$ or 1 which will create the two methods named $x_k > c$ and $x_k \leq c$. Then all the predicted values are integrated, and the higher votes are producing the final forecasting result. The iterative process consists of relating weighting value $t = 1 \ldots T$ for every learning sample in the Eqs. 14, 15, 16 and 17.

Given that: $(x_1, y_1), \ldots, (x_m, y_m)$ *where* $x_i \in X,\ y_i \in Y = \{-1, +1\}$.

Initialization of $D_1(i) = \frac{1}{n}$, here n denotes the quantity of data.

*For* $t = 1$ *to* $T$ : *train the base learner with* $D_t$ *distribution*

Find the week type hypothesis $h_t$: $X \to \{-1, +1\}$ *with error*

$$\in_t = Pr_{i \sim D_t}\big[h_t(x_i) \neq y_i\big] \tag{14}$$

$$a_t = \frac{1}{2}\lambda v\left(\frac{1 - \in_t}{\in_t}\right) \tag{15}$$

Update:

$$D_{t+1}(i) = \frac{D_t(i)}{Z_t} \times \begin{cases} e^{-at}, jika\, h_t(x_i) = y_i \\ e^{at}, jika\, h_t(x_i) \neq y_i \end{cases} = \frac{D_t(i)\exp(-a_t y_i h_t(x_i))}{Z_t} \tag{16}$$

From the above equation $Z_t$ method is called the normalization factor, so the outcome in the following equation:

$$H(x) = sign\left(\sum\nolimits_{t=0}^{T} a_t h_t(x)\right) \tag{17}$$

## 4 Results and Discussion

When the word IDS is mentioned, it indicates two concepts namely intrusion and detection system. The word intrusion has become familiar among many network hosts, as they are confronted with the problem of unauthorized access of attackers and hackers to the valuable information in a network or computer system compromising their confidentiality, availability, or integrity.

The performance of the model was evaluated by tabulating and calculating the features like accuracy, false alarms, and detection rate. TP represents true positive that means the data of real attacks which are segregated as attacks, FP represents false positives that provide the real attack data which are separated as normal from the dataset, FN means false negative which indicates the real attack data that are categorized as normal and TN stands for true negative which are the normal data that are assorted as normal.

The term accuracy infers to the rate at which data are classified rightly. In other words, the cases for which actual attack data are classified as attacks and normal data as normal. Accuracy can be mathematically represented as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{18}$$

Recall and precision are additional measures that we take into account in this study and are defined as follows:

$$Precision = \frac{TP}{TP + FP} \tag{19}$$

$$Recall = \frac{TP}{TP + FN} \tag{20}$$

**KDD CUP99 Dataset Analysis**

For our research work we have downloaded cyber hacking dataset from a online database provider. It consists of 10244 Rows and 12 Columns. The attributes used in the dataset are Time, Date, Delivery ratio, Packet Length etc. We have measured output parameters such as Accuracy, Precision and Recall respectively.

**Accuracy Analysis**

Now we are going to apply our proposed EMLP algorithm along with existing algorithms such as ANN and MLP over KDDCUP99 Dataset. The following Table 1 and Fig. 2 represents Accuracy Analysis of proposed EMLP compared over ANN and MLP. From the results its proved that proposed EMLP produces Accuracy of about 93% which is higher than ANN Accuracy which is 84% and MLP Accuracy which is 89% respectively.

**Table 1.** Accuracy comparison of proposed EMLP with other existing algorithms

| No of Iterations | ANN Accuracy (%) | MLP Accuracy (%) | EMLP Accuracy (%) |
|---|---|---|---|
| 10 | 82.1 | 87.3 | 91.3 |
| 20 | 82.6 | 88.5 | 92.5 |
| 30 | 83.2 | 89.3 | 92.6 |
| 40 | 84.4 | 89.5 | 93.1 |
| 50 | 84.5 | 89.8 | 93.3 |

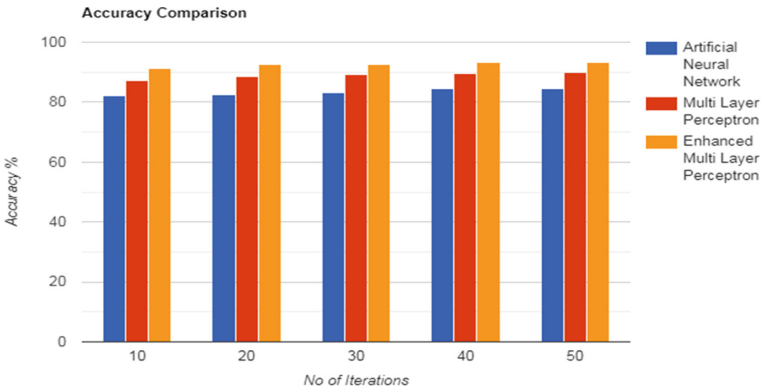


**Fig. 2.** Accuracy comparison of proposed EMLP with other existing algorithms graph

**Precision Analysis**

Now we are going to apply our proposed EMLP algorithm along with existing algorithms such as ANN and MLP over KDDCUP99 Dataset. The following Table 2 and Fig. 3 represents Precision Analysis of proposed EMLP compared over ANN and MLP. From the results its proved that proposed EMLP produces Precision of about 0.88 which is higher than ANN Precision which is 0.80 and MLP Precision which is 0.85 respectively.

**Table 2.** Precision comparison of proposed EMLP with other existing algorithms

| No of Iterations | ANN Precision (%) | MLP Precision (%) | EMLP Precision (%) |
|---|---|---|---|
| 10 | 80.3 | 84.6 | 87.6 |
| 20 | 80.4 | 85.2 | 88 |
| 30 | 80.7 | 85.3 | 88.2 |
| 40 | 81.1 | 85.5 | 88.6 |
| 50 | 81.2 | 85.7 | 88.8 |

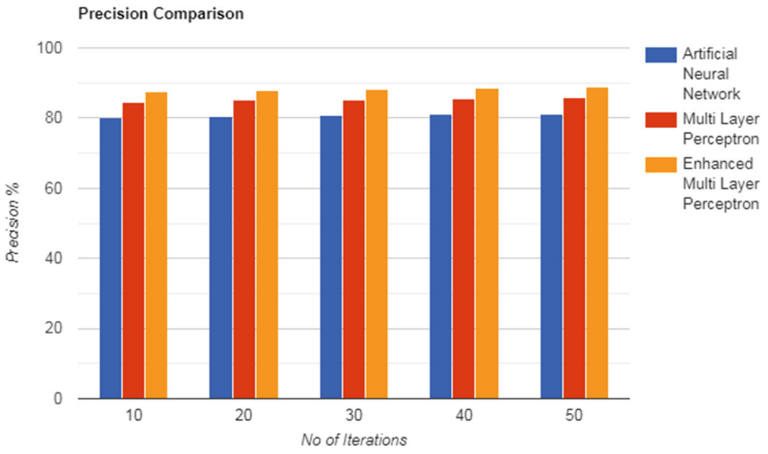


**Fig. 3.** Precision comparison of proposed EMLP with other existing algorithms graph

**Recall Analysis**

Now we are going to apply our proposed EMLP algorithm along with existing algorithms such as ANN and MLP over KDDCUP99 Dataset. The following Table 3 and Fig. 4 represents Recall Analysis of proposed EMLP compared over ANN and MLP. From the results its proved that proposed EMLP produces Recall of about 0.84 which is higher than ANN Recall which is 0.79 and MLP Recall which is 0.81 respectively.

**Table 3.** Recall comparison of proposed EMLP with other existing algorithms

| No of Iterations | ANN Accuracy (%) | MLP Accuracy (%) | EMLP Accuracy (%) |
|---|---|---|---|
| 10 | 78.5 | 81.4 | 82.1 |
| 20 | 78.1 | 81.5 | 83.3 |
| 30 | 79.3 | 81.7 | 83.6 |
| 40 | 79.5 | 81.9 | 84.2 |
| 50 | 79.7 | 82.4 | 84.7 |

In terms of Accuracy we have evaluated three algorithms with 50 iterations on KDD-CUP99 dataset. From the Accuracy Table and graph we can analyze that the average Accuracy of ANN is 84%, MLP is 89% and EMLP is 93%. From the results we can prove that EMLP outperforms other algorithms in terms of Accuracy.

In terms of Precision we have evaluated three algorithms with 50 iterations on KDD-CUP99 dataset. From the Precision Table and graph we can analyze that the average Precision of ANN is 0.80, MLP is 0.85 and EMLP is 0.88. From the results we can prove that EMLP outperforms other algorithms in terms of Precision.

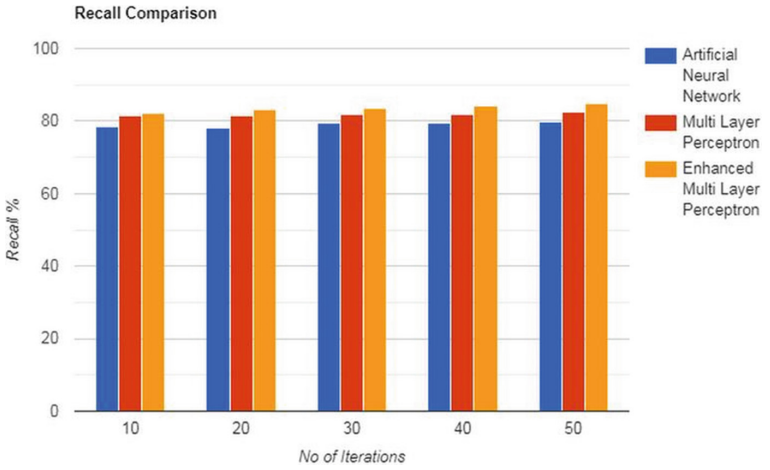


**Fig. 4.** Recall comparison of proposed EMLP with other existing algorithms graph

In terms of Recall we have evaluated three algorithms with 50 iterations on KDD-CUP99 dataset. From the Recall Table and graph we can analyze that the average Recall of ANN is 0.79, MLP is 0.81 and EMLP is 0.84. From the results we can prove that EMLP outperforms other algorithms in terms of Recall.

## 5 Conclusion

In today's world of the internet, data security has become a dare for many organizations. Protecting authorized data from the attack of intruders is very much essential. Intrusion detection plays a crucial role in achieving the best accuracy in securing data and filling the gaps of the existing techniques. IDS is the new evolution in the digital world that enhances network security by safely preserving the credential data of the organization. The IDS technology assists the network administrator and the host of the network in identity any type of malicious functions performed on the network. The system alerts the host to take appropriate corrective actions against the threat identified. The research proves that IDS is the best solution for conserving the critical data from intruders who are real-world entities with the help of deep learning techniques.

## References

1. Krishna, A., Lal M.A., A., Mathewkutty, A.J., Jacob, D.S., Hari, M.: Intrusion detection and prevention system using deep learning. In: IEEE International Conference on Electronics and Sustainable Communication Systems (ICESC) (2020)
2. Kim, J., Shin, N., Jo, S.Y., Kim, S.H.: Method of intrusion detection using deep neural network. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp) (2017). https://doi.org/10.1109/BIGCOMP.2017.7881684
3. Santosh Kumar, S., Kannan, M., Vignesh, B., Rajarajan, S.: Intrusion detection system using deep learning. Int. J. Eng. Res. Technol. (IJERT) **9**(5), 8–13 (2021). ISSN: 2278-0181 Published by, www.ijert.org ICRADL - 2021 Conference Proceedings.

4. Anuradha, K., Nirmala Sugirtha Rajini, S., Bhuvaneswari, T., Vinod, V.: TCP/SYN Flood of Denial of Service (DOS) Attack Using Simulation. Test Eng. Manage., 14553–14558 (2020). ISSN 0193-4120
5. Praneeth, V., Kumar, K.R., Karyemsetty, N.: Security: intrusion prevention system using deep learning on the internet of vehicles. Int. J. Saf. Secur. Eng. **11**(3), 231–237 (2021)
6. Vijayakumar, R, Alazab, M., Soman, K.P., Poornachandran, P., Al-Nemrat, A., Venkatraman, S.: Deep Learning Approach for Intelligent Intrusion Detection System. IEEE Access. 7, pp. 41525-41550 (2019).https://doi.org/10.1109/ACCESS.2019.2895334
7. Ahmad, Z.: Network intrusion detection system: a systematic study of machine learning. Emerg. Telecommun. Technol. **32**, e4150 (2020)
8. Azhagiri, M., Rajesh, A., Karthik, S.: Intrusion detection and prevention system : technologies and challenges. Int. J. Appl. Eng. Res. **10**(87), 1–13 (2015). ISSN 0973-4562
9. Mudzingwa, D., Agrawal, R.: A study of methodologies used in intrusion detection and prevention systems (IDPS). IEEE Access, pp. 1–6 (2012). 978-1-4673-1375-9/12
10. Nada, J.A., Al-Mosa, M.R.: A proposed wireless intrusion detection prevention and attack system. In: IEEE International Arab Conference on Information Technology (ACIT), Werdanye, Lebanon, 28 November 2018–30 November 2018, pp. 1–5 (2018). https://doi.org/10.1109/ACIT.2018.8672722