Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

# Detection of Intrusion Data using Deep Learning for Multi Class and Binary Class Data

Bhoomi Sharma , Muskan , Jyoti Devi
Department of Computer Science and Engineering
Manav Rachna International Institute of Research and Studies (MRIIRS), Faridabad, India

*Abstract* - **In today's world, Wi-Fi networks are almost everywhere connecting our phones, laptops and smart devices, which make our life easier but also exposes to a variety of cyber threats. Malicious users can exploit vulnerabilities in Wi-Fi systems to perform attacks such as flooding networks, impersonating legitimate users, or injecting harmful data packets. Traditional intrusion detection system are often unable to attack these attacks efficiently and tend to generate many false alarms, which reduces trust in the system. This research proposes a new approach that combines artificial neurons with a bioinspired optimization technique called Harris Hawk optimization (HHO). We used the AWID dataset to train and test the model carefully preprocessing the data and selecting the important features. The result shows that the proposed system can detect attack with an accuracy of 99.16%, a detection rate of 99.49, ana a very low false alarm rate of only 1.27%. Compared to other machine learning and optimization methods, this model performs better, making it a practical and effective solution for Wi-Fi security.**

***Index Terms – Intrusion Detection Systems (IDS), Artificial Neurons, Harris Hawk Optimization Machine Learning, Bio-inspired Algorithm. Deep Learning, AWID.***

## I. INTRODUCTION

Smart homes are becoming very common today. People use devices like CCTV cameras, smart locks, smart speakers, and Wi-Fi lights in their houses. But these devices are not fully safe. Hackers can control them, steal data, or shut them down. For example, Mirai malware attacked thousands of smart devices and made them unusable. This paper reviews various deep learning models used for intrusion detection, the datasets commonly applied (like **NSL-KDD** and **CICIDS2017**), and compare their performance. The goal is to understand how deep learning improves detection accuracy and what challenges still remain in building effective IDS systems. Most smart home devices have low memory and weak processors, so we cannot install normal anti virus inside them. That is why we need a smart system that can detect attack from outside by checking the network traffic.

Machine learning helps to find abnormal behavior without writing manual rules. In this research, we use an ensemble machine learning model called XGBoost to detect different attacks in smart homes.

## II. LITERATURE REVIEW

Several researchers have explored the use of deep learning techniques for building efficient Intrusion Detection Systems (IDS). Earlier studies mainly focused on **machine learning algorithms** like Decision Trees, Support Vector Machines (SVM), and k-Nearest Neighbors (KNN). These models gave good results for small datasets but failed to perform well on large and complex network traffic.

To improve performance, researchers started using **deep learning models** because of their ability to automatically extract important features from data. For example, **Convolutional Neural Networks (CNNs)** have been used to detect network intrusion by learning spatial patterns in network traffic. CNN-based IDS models achieved higher accuracy and reduced false alarms compared to traditional methods.

Similarly, **Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM)** models are effective in analyzing time-dependent data such as network traffic sequences. They can detect both short-term and long-term attack patterns. Many studies using the **NSL-KDD** and **CICIDS2017** datasets have shown that LSTM-based IDS gives better results for both **binary** and **multi-class intrusion detection.**

Some researchers also combined multiple deep learning models, such as
**CNN-LSTM hybrid models,** to improve
detection speed accuracy. These hybrid models can handle feature learning and sequences analysis together. However, studies also note challenges like **high computational cost, imbalanced datasets,** and **difficulty in explaining model decisions.**

Overall, the literature shows deep learning method outperform traditional techniques but further improvement are needed to make IDS more accurate, efficient, and explainable.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

## III. RESEARCH METHODOLOGY

### A. Dataset

Intrusion Detection system (IDS) help us find unwanted and harmful activities in a computer network. Deep learning is used in IDS because it can learn patterns from large amount of data on its own. When using deep learning, the data is usually divided into two types **binary class**, where the system is only checks if the traffic is normal or attack, and **multi-class,** where the system also identifies what type of attack it is, such DoS, DDoS, port scan, or brute force. To build these systems, firstly researcher have to collect network dtaa from the popular datasets like NSL-KDD, CICIDS2017, or UNSW-NB15. And then they clean the data by removing errors, converting text into numbers, and scaling values so the model can learn better . Deep learning model like CNN, LSTM, Autoencoders are used to detect attacks. CNN is good at finding patterns, LSTM is good for time based attacks, and Autoencoders help find new or unknown attacks. Sometimes, two models are combined like CNN+LSTM to get better results. Binary-class detection is simple because it only decides "normal" or "attack" while multi-class detection is more difficult because it must correctly identifies different attack types. Researchers measure the performances of these models using accuracy , precision, recall, and F1- score. Overall , deep learning helps IDS become more accurate and reliable by automatically learning how both normal and attack traffic look.

### B. Data Preprocessing

Before using the data to train a deep learning model, it must be cleaned or prepared. This process is called pre-processing. During this step, missing values are fixed, duplicate entries are removed, and text values like 'TCP" or "UDP" are converted into numbers because the model can only understand numeric data. The data is also scaled so that large numbers do not dominate small ones. Another important step is balancing the dataset, because in many cases, some attack type are just a samples. SMOTE are used to create more sample attacks.

### C. Convolutional neural networks

CNNs are more advanced models that can find deeper and more more complex patterns in data. Although CNNs are mainly used for images, they also work very well with network traffic because the feature can be arranged in a grid-like structure. CNNs can automatically pick out important details from the data and often perform very well in detecting different types of attacks. They are very effective for multi class intrusion detection.

### D. Recurrent Neural Networks (RNN), LSTM, GRU

RNNs and their improved versions, LSTM and GRU, are used when the data has a time-based sequence. Network traffic flows over time, so these models are good at understanding pattern that occur step by step. For Example they can detect attack like DDoS or brute force attempts that follows a repeated pattern. These model remember the previous inputs and use this information to make a batter prediction.

### E. Autoencoder

Autoencoder are special dep learning models used in mainly for finding unusual behavior. They learn what normal traffic look like and then try to reconstruct it. When the autoencoder sees something different from normal traffic, it produces a large error, which indicates a possible attack. Autoencoders are useful for detecting new or unknown attacks, especially when there is very little labeled attack data.

### F. Binary Class Intrusion Detection

Binary class intrusion detection means the model only need to decide whether the a traffic is normal or attack. This is simpler than multi class detection because there are only two categories to choose from. Binary classification is fast and more accurate. It is very useful in real time monitoring systems and devices with limited computing power- such as IoT devices.

### G. Multi class Intrusion Detection

Multi-class intrusion detection is more challenging because the model identify exactly which type of attack is happening. Examples include DoS, DDoS, brute force , port scanning. This types of detection give more detailed information to the security teams, but it is more difficult because many attacks look similar.

## IV. RESULT AND DISCUSSION

The experiments carried out on the AWID dataset provided clear insights into how the proposed IDS behaves under both binary and multi-class classification settings. The goal throughout the evaluation was to check whether the system could reliably separate normal Wi-Fi traffic from different types of attacks, while keeping the training process stable and efficient. The use of the Harris Hawk Optimization Algorithm played an important role in improving model consistency and tuning.

### 4.1 Binary Classification Results

In the binary classification task, the model was required to distinguish between only two classes: normal traffic and attack traffic.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

The overall performance was highly encouraging, and the system was able to learn the main characteristics of both categories very effectively

The results are summarized below:

• Accuracy : ~99.16%

• Detection rate: ~99.49%

• False alarm rate: Extremely low

Such high accuracy and detection rate indicate that the model understood the underlying data patterns well. The very low false alarm rate is especially important for real-time intrusion detection, because unnecessary alerts overload security teams and reduce the usefulness of the system.

The improvements brought by HHO were quite noticeable. Traditional hyper parameter tuning methods sometimes cause the model to settle in poor local optima, but HHO's adaptive exploration helped the network reach better configurations. During training, the loss curve remained smooth, and the model showed very limited over fitting.

In summary, the binary classification results confirm that the proposed IDS can accurately and reliably separate benign and malicious Wi-Fi traffic, making it suitable for real operating environments.

### 4.2 Multi – Class classification Results

The multi-class classification scenario was more challenging, as the dataset includes several different types of attacks such as injection, impersonation, flooding, and authentication misuses. Many of these attack categories share similar traffic patterns, which makes correct classification more difficult.

Even with this complexity, the model delivered strong and balanced performances.

Key observations include:

• High accuracy across multiple attack types

• Very little confusion between similar categories such as impersonation and injection.

• Strong generalization capability on the test set

These results suggest that the model was able to learn both the structural and temporal features of the AWID traffic. The role of HHO was even more evident in the multi-class task, as it helped the model find cleaner decision boundaries between closely related attack types.

Overall, the multi-class results demonstrate that the combination of deep learning and HHO can handle diverse intrusion scenarios and capture subtle variations in Wi-Fi attack behavior.

## V. CONCLUSION

The main objective of this research was to design an Intrusion Detection System that performs well in Wi-Fi environments and remains stable during training. The proposed system combines a deep learning model with the Harris Hawk Optimization Algorithm, and the experimental results show that this approach is highly effective.

Both binary and multi-class evaluations produced strong performance. The model provided high accuracy, maintained a very low false alarm rate, and successfully detected various types of attacks. The good separation between similar attack types– especially in the multi-class task–shows that the model captured detailed patterns present in the dataset.

Based on the overall findings, it can be concluded that using bio-inspired optimization techniques such as HHO together with deep learning can significantly improve IDS performance. This combination not only enhances accuracy but also contributes to better model stability and faster convergence.

## VI. FUTURE SCOPE

Although the proposed system shows promising results, there are still several ways in which this work can be extended:

1. Real-Time Deployment

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

Future work can focus on deploying the model in real-time monitoring tools. This will help evaluate how well the system performs when faced with live, unpredictable network traffic.

2. Integrating with Edge or Fog Computing

As IOT continues to grow, lightweight versions of the IDS can be developed for edge or fog devices. This will reduce delay and help in quick, localized detection.

3. CNN-LSTM Hybrid Models

Future studies can explore hybrid architectures where CNNs handle spatial feature extraction and LSTMs capture time- based patterns. This can further improve the detection of evolving attacks.

4. Zero-Day Attack Detection

Using transfer learning techniques may help the system recognize previously unseen attack types, making it more adaptable to real-world threats

5. Explainable AI

Adding explainable AI tools can help security analysts understand why the model classified a specific packet or flow as malicious, increasing trust in the system.

6. Evaluation in Multiple Datasets

Extending the experiments to other datasets such as NSL-KDD, CIC-IDS2018, or UNSW-NB15 will strengthen the reliability and general applicability of the proposed IDS.

## VII. REFERENCES

[1] Aegean Wi-Fi intrusion dataset (AWID).
"AWID: A New Generation of Wireless Intrusion Dataset." University of the Aegean, 2015. Available: https:icsdweb.aegean.gr/awid/

[2] T.T Nguyen and G. Armitage, "A Survey of techniques for Internet Traffic Classification Using Machine Learning," IEEE Communication Surveys & Tutorials, vol.10, no. 4,pp. 56,2008.

[3] H. Heidari, H. Faris, I.Aljarah and S. Mirjalili, "Harris Hawks Optimization: Algorithm and Applications," Future Generation Computer Systems, vol. 97,pp. 849-872,2019.

[4] J. Kim, J. Kim, H. Shim and E. Choi, "Deep Learning-Based Intrusion Detection System for Wireless Networks," IEEE Access, vol.7, pp. 100,104-100,112,2019.

[5] A. Javaid, Q. Niyaz, W. Sun and M. Alam, "A Deep Learning Approach for Network Intrusion Detection Using NSL-KDD Dataset," in Proceedings of the 9th EAI International Conference on Bio-inspired Information and Communication Technologies, 2016,pp.21-26.

[6] B. Bhuyan, D. Bhattacharyya and J.K. Kalita, "Network Anomaly Detection: Methods, Systems and Tools," IEEE Communication Surveys & Tutorials,vol.16, no. 1,pp. 303-336,2014

[7] W. Wang, Y. Sheng J. Wang et al., "HAST-IDS: Learning Hierarchical Spatial-Temporal Features Using Deep Neural Networks to Improve Intrusion Detection," IEEE Access, vol.6,pp.1792-1806,2018.

[8] N. Mustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems," in Proc. Military Communications and information Systems Conference (MilCIS), 2015,pp. 1-6.

[9] Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning," Nature, vol. 521,pp. 436-444,2015.

[10] S. Revathi and A. Malathi, "A Detailed Analysis on NSL-KDD Dataset Using Various Machine Learning Techniques," International Journal of Engineering Research & Technology, vol.2, no. 12, pp. 1848-1853, 2013.

[11] K. Kim, "Deep Learning-Based Intrusion Detection System Using Autoencoders," Journal of Information Security and Applications, vol. 7, no. 3-4, pp. 197-387, 2014.