International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

Comparison of Machine Learning Algorithms for Cyberattack Detection

Sumedh B. Dongre[1], Nikita P. Ramteke[2], Prof. A. J. Pimparikar[3]

[1,2,3]*MCA II yr Sem IV P.G. Dept of Computer Applications, PRMITR Badnera City-Amravati country India*

***Abstract*- Today, internet services are used everywhere, and because of this, cyberattacks have also increased. Cyberattacks are dangerous for network security and personal data. Some common attacks are denial of service (dos), phishing, brute force, and dictionary attacks. These attacks can stop services or steal important information. Traditional security systems sometimes fail to detect new and advanced attacks. Because of this, machine learning is becoming very useful in modern intrusion detection system. This research compares three types of machine learning methods these are supervised, unsupervised, and reinforcement learning for detecting cyberattacks. The supervised learning algorithms used in this study are Logistic Regression, Decision Tree, Support Vector Machine (SVM), Random forest and Naïve Bayes. For unsupervised learning, K-Means clustering, Isolation Forest are used. Reinforcement learning algorithms such as Q-Learning and Deep Q Network are studied. Experiments were performed using cybersecurity datasets. The results shows that supervised learning algorithms give high accuracy when classifying attacks. Unsupervised learning algorithms are helpful for detecting new and unknown attacks. Reinforcement learning algorithms can learn and improve their defense method over time. The results show that using different machine learning methods together can help create a stronger intrusion detection system.**

***Keywords*-- Cybersecurity, intrusion detection system, machine learning, supervised learning, unsupervised learning, reinforcement learning, cyberattack detection.**

I. Introduction

Cybersecurity is very important today because many people and organizations use computer networks. The use of cloud computing, online payments, and digital communication has increased a lot. Because of this, cybercriminals try to attack networks and steal information. Some common cyberattacks are denial-of-service attacks which sends too much traffic to a system so that real users cannot access the service/ make server unavailable to authorized users, phishing attacks try to trick users into giving their personal information, such as passwords or credit card numbers, brute force attacks try many password combinations, and dictionary attacks uses predefined password lists to break into accounts.

Intrusion Detection System (IDS) is used to watch network activity and find suspicious behavior. Traditional IDS systems used fixed rules or known attack patterns. Because of this, they cannot always detect new types of cyberattacks. Machine learning is a better solution because it can study large amounts of network data and find patterns related to cyberattacks. Machine learning models learn from past data and help detect unusual behavior that may show an attack.

Machine learning methods used for cyberattack detection can be divided into three main groups:

1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning

This research compares different algorithms from these three groups to see how well they detect cyberattacks.

II. Literature Review

Many researchers have studied how machine learning can help intrusion detection systems improve cybersecurity. Supervised learning algorithms are often used because they work well when labelled datasets are available. Logistic regression and naïve bayes are simple algorithms that are often used to classify network traffic as normal or malicious. Decision tree algorithms are also popular because they create models that are easy to understand. SVM has shown good results in cyberattack detection because it can work well with large and complex datasets. Random forest is another strong algorithm that combines many decision trees to improve accuracy. Unsupervised learning algorithms are also used for detecting unusual patterns in network traffic. K-means clustering algorithm group similar data points together and help find abnormal behavior that may indicate attacks. Isolation forest especially useful for detecting anomalies in large cybersecurity datasets. Recently, reinforcement learning has also been studied in cybersecurity. Algorithms like Q-learning and deep Q network allow systems to learn better defense strategies by interacting with the network environment and receiving rewards.Even though many machine learning techniques have been developed, choosing the best algorithms for cyberattack detection is still difficult. Therefore, comparing different algorithms is important to understand their advantages and limitations.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

III. METHODOLOGY

A. Dataset

Cybersecurity datasets such as NSL-KDD are used to test machine learning algorithms. These datasets contain network traffic records that represent normal activity and different types of cyberattacks.

The dataset includes many features that describe network behavior, such as: Connection duration, Protocol type, Number of packets, Service type, Error rate.

The dataset is divided into two parts: training data is used to train the algorithms and testing data is used to evaluate their performance.

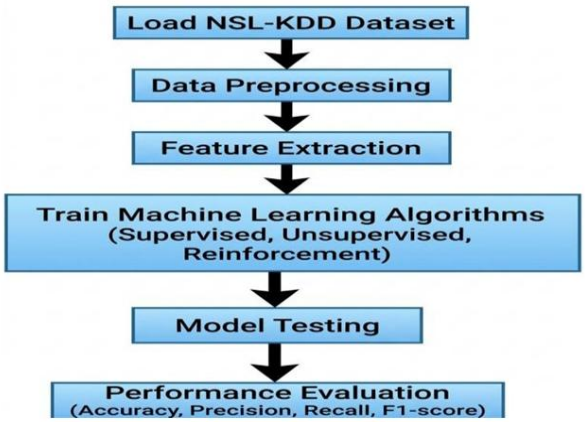


Fig.1: Workflow of machine learning for cyberattack detection

B. Data preprocessing

Data preprocessing is an important step before training machine learning models. The following steps are performed:

Data cleaning to remove duplicate and incomplete records

Feature selection to choose important network attributes

Normalization to scale numerical values

Encoding categorical data into numerical form These steps help improve the accuracy of the models and make the computation faster.

C. Machine learning algorithms 1.Supervised learning algorithms

The supervised learning algorithms are used:

Logistic regression, Decision tree, Naïve bayes, Support vector machine (SVM), Random forest,

These algorithms learn from labeled data and classify network traffic as normal or malicious.

D. Unsupervised learning algorithms

Unsupervised algorithms analyze data that are unlabeled. They try to find patterns or unusual behavior in the data. The algorithms used are:

1. k-means clustering
2. isolation forest

These algorithms are helpful for detecting new types of cyberattacks that are not present in the training data.

Reinforcement learning algorithms

Reinforcement learning algorithms learn through a system of trial or error, receiving rewards for good action and penalties for bad action. The algorithms are used:

1. Q-learning
2. Deep Q Network (DQN)

These algorithms help systems learn better strategies for defending against cyber threats.

E. Evaluation metrics

Evaluation metrics used to measure algorithms performance:

Sr. No	Evaluation metrics	Description
1.	Accuracy	Overall Performance of the system
2.	Precision	Percentage of classified cyberattack which are actually cyberattack
3.	Recall	Effectiveness of system in detecting cyberattacks
4.	F1-score	Blend of precision and recall

These metrics are calculated using a confusion matrix.

System architecture for cyberattack detection

The proposed IDS system uses different machine learning techniques to detect cyberattacks in network traffic. The system processes data through several stages /layers such as here, cleans and normalizes network data and duplicate or incomplete records.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

Feature extraction:- in this layer, extracts important features such as:

1. Packet size
2. Protocol type
3. Connection duration
4. Login attempts
5. Traffic frequency

Machine learning :- in this layer apply machine learning algorithms to detect attacks.

Detection and response :- here, system classifies traffic as normal or malicious and sends alerts when an attack is detected.

F. Experimental setup

Experimental environment

For machine learning algorithms, python programming language is efficient to used, scikit-learn used as libraries and NSL-KDD dataset used. TCP, UDP protocol type used.

IV. Results And Discussion

The results show that supervised learning algorithms give the highest accuracy. Support vector machine and random forest perform especially well in detecting attacks. Unsupervised learning algorithms are useful for detecting unusual patterns in network traffic. Isolation forest algorithm can find new type of attacks that were not seen during training. Reinforcement learning algorithms allow system to learn and improve over time. Deep Q network shows good results in changing network environments where new attacks appear. Using machine learning algorithms together can make cyberattack detection systems stronger and more effective.

Result table:

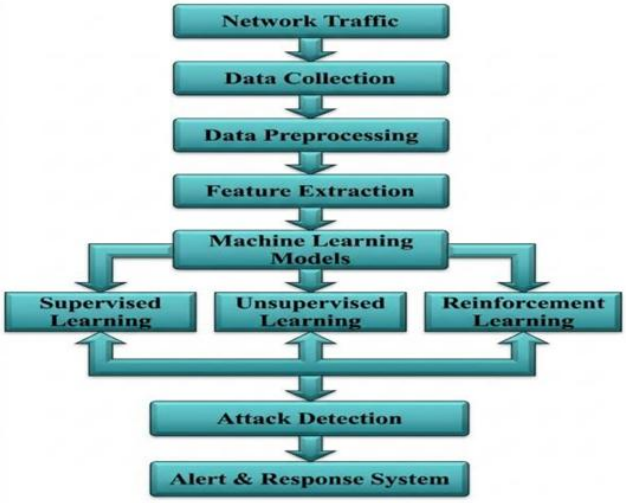


Fig 2: IDS System Architecture of cyber attack

Data collection:- In this layer, collects network traffic from routers, servers, and firewall and captures packet-level information.

Preprocessing:-

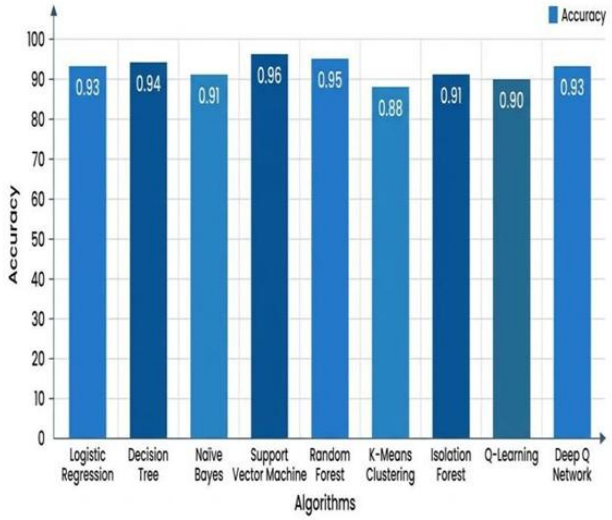
Algorithms	Learning type	Accuracy
1. Logistic regression	Supervised	93%
2. Decision tree	Supervised	94%
3. Naïve bayes	Supervised	91%
4. Support vector machine	Supervised	96%
5. Random forest	Supervised	95%
6. K-means clustering	Unsupervised	88%
7. Isolation forest	Unsupervised	91%
8. Q-learning	Reinforcement	90%
9. Deep Q Network	Reinforcement	93%

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)

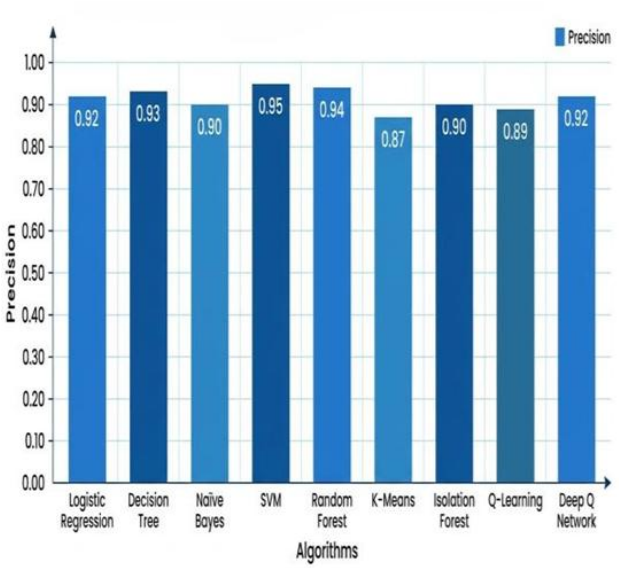
Comparison table of machine learning algorithms for cyberattack detection:

Sr. no	algorithm	accuracy	precision	recall	Fscore
1	Logistic Regression	0.93	0.92	0.91	0.915
2	Decision tree	0.94	0.93	0.92	0.925
3	Naïve bayes	0.91	0.90	0.89	0.895
4	Support vector machine	0.96	0.95	0.95	0.950
5	Random forest	0.95	0.94	0.94	0.945
6	K-means clustering	0.88	0.87	0.86	0.865
7	Isolation forest	0.91	0.90	0.90	0.905
8	Q-learning	0.90	0.89	0.88	0.885
9	Deep Q Network	0.93	0.92	0.91	0.915

Graphical representation of evaluation metrics of machine learning algorithms for cyberattack detection:

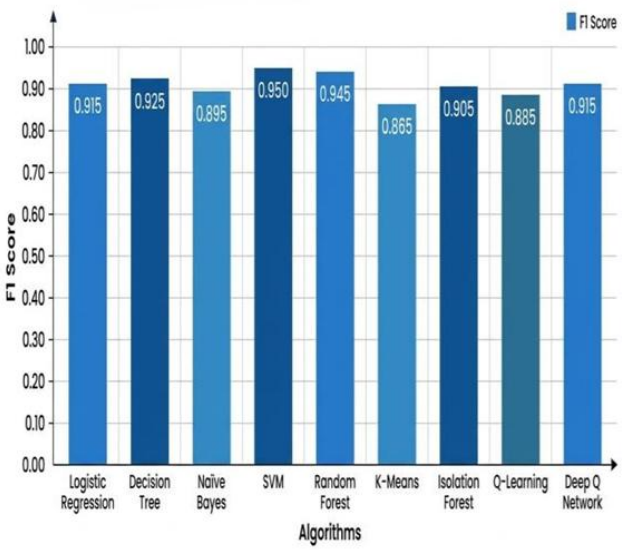


Graph 1 - Accuracy Comparison of Algorithms



Graph 2 - Precision Comparison

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 03, March 2026)



Graph 4 – F1 Score Comparison

V. Conclusion

Cyberattacks are becoming more advanced as network system grows. Traditional security systems alone cannot detect all types of attacks. This research compared some machine learning algorithms for cyberattack detection. The results show that these algorithms provide high accuracy and allow systems to adapt and improve their defense strategies.

REFERENCES

[1] H. Liu and B. Lang, "Machine Learning and Deep Learning Methods for Intrusion Detection Systems: A Survey," Applied Sciences, vol. 9, no. 20, p. 4396, 2019. https://doi.org/10.3390/app9204396

[2] W. Wang, M. Zhu, X. Zeng, X. Ye, and Y. Sheng, "Malware Traffic Classification Using Convolutional Neural Network for Representation Learning," 2017. https://ieeexplore.ieee.org/document/7946998

[3] M. Soltani, M. Jafari Siavoshani, and A. H. Jahangir, "A ContentBased Deep Intrusion Detection System," 2020. https://arxiv.org/abs/2001.05009

[4] S. Axelsson, "The Base-Rate Fallacy and the Difficulty of Intrusion Detection," ACM Transactions on Information and System Security, 2000. https://dl.acm.org/doi/10.1145/357830.357849

[5] J. McHugh, "Testing Intrusion Detection Systems: A Critique of the 1998 and 1999 DARPA Intrusion Detection System Evaluations," 2000. https://dl.acm.org/doi/10.1145/357830.357833

[6] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Rahim, "Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection," IEEE Access, 2018. https://ieeexplore.ieee.org/document/8361332

[7] N. Moustafa and J. Slay, "UNSW-NB15: A Comprehensive Data Set for Network Intrusion Detection Systems," 2015 Military Communications and Information Systems Conference, 2015. https://ieeexplore.ieee.org/document/7348942

[8] M. Almseidin, M. Alzubi, S. Kovacs, and M. Alkasassbeh, "Evaluation of Machine Learning Algorithms for Intrusion Detection System," 2017. https://arxiv.org/abs/1801.02330

[9] R. Chalapathy and S. Chawla, "Deep Learning for Anomaly Detection: A Survey," arXiv, 2019. https://arxiv.org/abs/1901.03407

[10] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A Detailed Analysis of the KDD Cup 99 Data Set," 2009. https://ieeexplore.ieee.org/document/5137283