Machine Learning for Network Security and Anomaly Detection

Dr. M. Munafur Hussaina[1], Principal, Mrs. M. Nasreen Banu[2]
[1]*PG Department of Computer Science, Aiman College of Arts and Science for Women-Trichy*
[2]*M.Sc. Computer Science, PG Department of Computer Science,*
Aiman College of Arts and Science for Women-Trichy

***Abstract*—Many systems are now at risk of advanced persistent attacks, data breaches, and network invasions due to the rapid expansion of network technology and the increasing sophistication of cyber threats, which have greatly overtaken traditional security measures. Machine learning (ML) approaches have become a potent instrument in the field of anomaly detection and network security as a result of these difficulties. The use of supervised and unsupervised machine learning methods to improve network security systems is examined in this study, with a focus on intrusion detection and emerging threat mitigation. While unsupervised methods like K-Means clustering and Autoencoders are used to discover new and unknown attack patterns in real-time, supervised models like Support Vector Machines (SVM) and Random Forest are used to detect existing attacks through classification. We examine the benefits and drawbacks of these machine learning models, highlighting their capacity to identify hitherto unidentified threats, automate judgment calls, and lessen the need for human intervention. The study also explores the real-world difficulties of implementing machine learning models in dynamic network settings, such as problems with data labeling, false-positive rates, model generalization, and security system scalability. The study also discusses how hybrid models, which combine supervised and unsupervised learning techniques, might provide a more comprehensive and flexible security framework.**

In terms of accuracy, adaptability, and response times, these ML-based techniques surpass conventional security systems in detecting a variety of network anomalies, according to experimental data from many testbed environments. Performance measures including recall, accuracy, precision, and F1-score are contrasted with more traditional techniques like heuristics and signature-based detection.

According to our research, machine learning (ML)-based anomaly detection has great promise for contemporary network security, offering proactive and instantaneous defenses against a constantly evolving array of cybersecurity risks.

This study demonstrates how machine learning (ML) has the potential to revolutionize network security by making it more intelligent, automated, and responsive to both known and emerging cyberthreats. It does this by thoroughly reviewing a variety of machine learning methods. We also go over potential future research avenues and deployment tactics that could improve machine learning's effectiveness in cybersecurity applications.

***Abstract*— Cybersecurity, Supervised Learning, Unsupervised Learning, Anomaly Detection, Intrusion Detection Systems, Network Security, And Machine Learning.**

I. INTRODUCTION

Network security has emerged as a crucial component of contemporary digital infrastructure as the digital environment grows more intricate and linked. It has grown increasingly difficult to defend networks against a wide range of cyberthreats, such as data breaches, illegal access, and cyberattacks. Conventional security tools, such firewalls and intrusion detection and prevention systems (IDPS), have been essential defenses in identifying and thwarting known attacks. These systems, however, frequently find it difficult to keep up with the constantly changing landscape of cyberthreats, particularly when attackers employ more complex and dynamic strategies to evade detection. Ahmed, Mahmood, and Hu (2016)

The investigation of more flexible and intelligent security solutions has been prompted by the shortcomings of conventional techniques in detecting sophisticated, zero-day attacks and advanced persistent threats (APTs).

The incorporation of machine learning (ML) into network security systems is one promising strategy that has attracted a lot of interest recently. Without requiring frequent manual updates or preset signatures, machine learning allows automated systems to learn from data, spot trends, and adapt to new and emerging threats.

The core of this strategy is anomaly detection, which is the process of spotting odd or unexpected network traffic patterns that can point to a security risk. Chandran, P., and Chakraborty, S (2020).

By identifying anomalies from established typical network activity, anomaly detection techniques can reveal previously unknown or unnoticed attack pathways, in contrast to traditional systems that rely on predetermined rules or signatures to identify assaults. Because it offers a way to identify zeroday assaults and other complex threats that traditional tools can miss, anomaly detection is therefore an essential part of contemporary security systems.

Traditional security methods face a serious challenge from the growing volume and complexity of network traffic, which calls for the creation of automated, intelligent systems that can process enormous datasets. These systems must be able to quickly detect new threats and detect minute patterns in vast volumes of data.

Machine learning is especially well-suited for this purpose because of its capacity for adaptation, generalization, and scaling.

With a focus on anomaly detection, this paper provides an extensive overview of machine learning algorithms used in network security. From supervised learning approaches like decision trees, support vector machines (SVM), and random forests to unsupervised approaches like KMeans clustering and autoencoders, we investigate a variety of machine learning models. These models' strengths and limitations in real-world deployment circumstances will be examined, as well as their capacity to spot anomalous activity and possible threats.

There has never been a greater need for automated, scalable, and intelligent solutions to identify and reduce security threats as networks get more intricate and cyberattacks continue to become more sophisticated. By offering a more proactive and dynamic approach to network defense, machine learning integration into network security frameworks has the potential to completely transform how businesses react to and defend against cyber threats.

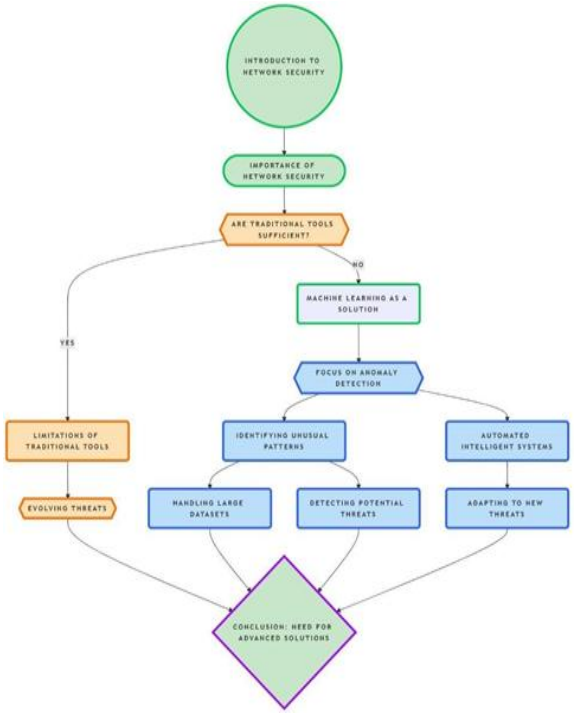

Fig :1

II. LITERATURE REVIEW

In network security, machine learning (ML) has emerged as a crucial technology, especially for anomaly detection. Conventional security systems, such as signature-based techniques, frequently have trouble fending off novel or unidentified threats. However, machine learning (ML) approaches are beneficial for detecting both known and undiscovered threats because they can learn from data and identify trends.

Labeled datasets, which are classified as either malicious or normal, are used to train supervised learning algorithms. When identifying known attacks, these models perform admirably. Decision trees and support vector machines (SVM) are common methods. However, because they rely on established patterns, they fall short when faced with novel, invisible threats. Labeled data is not necessary for unsupervised learning. Rather, by identifying departures from typical network behavior, these models detect anomalies. Zero-day attacks can be identified using techniques like auto-encoders and K-means clustering. However, because any odd activity even if it's

harmless is tagged as suspicious, they frequently produce false positives. Guven & Buczak (2016)

Hybrid models, which include supervised and unsupervised learning, offer a possible remedy. A more balanced method for thorough threat detection is provided by hybrid models, which use supervised learning for known threats and unsupervised learning for undiscovered ones. This hybrid approach lowers false positives while improving detection accuracy.

Data imbalance is still a problem, though, because regular traffic is frequently far more prevalent than attack data. Because of this, ML models may find it challenging to identify uncommon attack patterns. Another problem is real-time processing, since many machine learning methods are not appropriate for high-speed, real-time network contexts due to their large computational requirements. Li, Sun & Zhao (2021)

Adversarial assaults, in which the attacker manipulates data to trick the model, can also affect machine learning models. This may result in missing threats or incorrect categorization. Another issue is interpretability because many machine learning models, particularly deep learning models, function as "black boxes," making it difficult to comprehend the reasoning behind particular choices. Reducing false positives, increasing model interpretability, and enhancing model scalability are the main goals of future research. In order to boost confidence and efficacy in practical security applications, it is also crucial to improve machine learning's resistance to hostile attacks.

III. OBJECTIVE

The primary objectives of this research are as follows:

1. To investigate whether machine learning methods may be used to identify unusual activity in network traffic:

Investigating how machine learning (ML) may be used to spot anomalies in network traffic patterns that might point to security risks is the goal of this project. We intend to evaluate whether machine learning (ML) can detect risks or anomalous network behavior that have not been identified before, such as malware, Distributed Denial of Service (DDoS) assaults, or unauthorized access attempts, by utilizing both supervised and unsupervised learning techniques. Lunardi et al. (2022)

2. To evaluate the effectiveness and precision of several machine learning models in network security, including supervised and unsupervised learning techniques:

Comparing the effectiveness of various machine learning techniques is the main goal of this objective. These algorithms can be either supervised (like Support Vector Machines, Random Forest, and Neural Networks) or unsupervised (like K-Means Clustering and Auto-encoders). In order to provide a thorough performance analysis across various network traffic types, we want to assess the models' detection accuracy, computational efficiency, scalability, and generalization to unseen data.

3. To assess how well ML-based models perform in comparison to more conventional techniques, such rule based and signature-based detection systems:

Conventional network security solutions identify risks using pre-established rules and attack signatures. Comparing the efficacy of ML-based anomaly detection models with these traditional methods is the aim of this objective. We seek to measure the benefits and drawbacks of machine learning against rule-based techniques in identifying known and unknown threats by comparing performance indicators such as precision, recall, F1-score, and false positive rate. Verma & Patel (2019)

4. To determine the obstacles and constraints in the large-scale application of machine learning for network security:

There are many obstacles in scaling machine learning models to manage massive, real time network traffic. This goal is to investigate the challenges of using machine learning (ML) for network security in real-world settings. These challenges include the risk of adversarial attacks that could potentially fool machine learning systems, the computational overhead necessary for real-time processing, and data quality issues (such as an imbalance between malicious and normal traffic). The interpretability of machine learning models and the challenges of making them transparent and intelligible to network managers will also be discussed. Khan & Hayat (2020).

IV. METHODOLOGY

This work uses a hybrid methodology that blends theoretical analysis with real-world experimentation. Among the techniques used are:

Steps	Description	Details / Example
1. Information Gathering	Collect labeled network traffic data for training and testing.	Public datasets like KDD Cup 1999 and NSL-KDD with benign and attack traffic samples.
2. Feature Engineering	Extract key features that characterize network behavior to improve model accuracy.	Packet size, protocol type (TCP, UDP, ICMP), source/destination IP addresses, connection duration.
3. Model Selection	Choose suitable machine learning models for supervised and unsupervised learning.	Supervised: SVM, Decision Trees, Random Forests, Neural Networks; Unsupervised: KMeans, DBSCAN, Autoencoders.
4. Performance Evaluation	Assess ML models against traditional security methods.	Compare with signature-based and heuristic-based detection techniques using metrics like accuracy and false positives.
5. Challenges & Limitations	Address issues affecting ML application in network security.	Data imbalance, computational complexity, and vulnerability to adversarial attacks.

Table: 1

1. Information Gathering

The machine learning models will be trained and assessed using publicly accessible datasets.

Labeled examples of malicious and legitimate network traffic can be found in these databases. Among the datasets are:

KDD Cup 1999: This dataset, which includes labeled examples of several kinds of network attacks, is a commonly used benchmark in intrusion detection systems (IDS). The KDD Cup 1999 dataset has been modified to solve some of its inherent flaws, including redundant records, in the form of NSL-KDD. Guven & Buczak (2016)

2. Engineering Features

- Feature engineering is a critical step in enhancing the performance of machine learning models. The selection of features is dependent on how well they can depict network behaviors and differentiate between malicious and legitimate data. Typical characteristics that will be retrieved consist of:
- Packet Size: Certain attack types, such denial-of-service attacks, can be identified by the size of the data packets. Guven & Buczak (2018)
- Protocol Type: Knowing which protocol—such as TCP, UDP, or ICMP—is being used for network communication might help you differentiate between various traffic patterns.
- Source/Destination IP Addresses: Communication irregularities between IP addresses may be a sign of DDoS assaults or other possible breaches.
- Connection Duration: The duration of the session can be used to identify anomalous login attempts and other connection-related actions.

3. Choosing a Model

Choosing appropriate machine learning models for both supervised and unsupervised learning techniques is the next stage in the methodology. These models were selected because they successfully distinguish between typical and unusual network traffic:

Learning Under Supervision:

Support Vector Machines (SVM): SVM is renowned for its capacity to classify data with distinct boundaries between classes and is efficient in high-dimensional areas. By distinguishing known attack patterns from regular traffic, it will be utilized to identify them.

Decision Trees: Based on decision rules generated from input features, these models are simple to understand and can be used to categorize attack traffic. Ma & Liu (2020)

Random Forests: An ensemble technique based on decision trees, Random Forests can produce reliable forecasts for spotting assault trends and are less likely to overfit.

Networks of Neural Systems: Neural networks, a deep learning technique, are very effective for handling big datasets and intricate, non-linear interactions in network data. They can adapt to various patterns of attack.

4. Unmonitored Education:

K-Means Clustering: K-Means is a popular and straightforward clustering algorithm that can assist in assembling instances of network traffic that are comparable to one another. Finding traffic patterns that might not match any known attack signatures is one of its uses.

DBSCAN-Density-Based Spatial Clustering of Applications with Noise is especially well-suited for identifying outliers, or unusual behavior. Ahmed et al. (2016)

Auto-encoders: Neural networks called autoencoders are made to learn effective data coding techniques. By rebuilding "normal" data and detecting abnormalities through variations in reconstruction error, they are helpful for anomaly identification.

5. Evaluation of Performance

Comparing the effectiveness of machine learning (ML) models with conventional network security techniques is a major goal of this study:

Signature-based Techniques: To identify known threats, these techniques use preestablished attack signatures. They are quick and effective, but they can't detect new or zero-day assaults that aren't included in the signature database. Chandran & Chakraborty (2020)

Heuristic-based Methods: To find suspicious activity, heuristic detection employs preestablished rules or behavior profiles. Despite being more adaptable than signature-based techniques, they may overlook intricate or novel attack pathways and have a larger risk of false positives.

6. Difficulties and Restrictions

Notwithstanding machine learning's encouraging potential in network security, a number of obstacles prevent its broad use:

Data Imbalance: There is frequently a notable imbalance in network traffic databases, with significantly more "normal" traffic than malicious traffic. Particularly for uncommon or novel attacks, this class imbalance may distort model training and increase the likelihood of false negatives, or missing attacks.

Computational Complexity: Deep learning techniques and other machine learning models are computationally costly and resource-intensive. Their use in real-time applications may be restricted as a result, particularly in big networks that produce enormous volumes of traffic.

Adversarial attacks: Machine learning models are susceptible to adversarial attacks, in which malevolent actors deliberately alter network traffic in order to trick the model into identifying an assault as beginning.

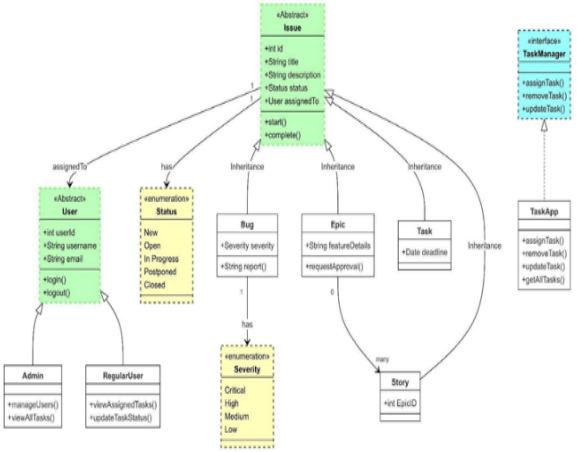

Fig: 2

V. RESULT

This section offers a thorough examination of the experimental findings pertaining to the effectiveness of several machine learning models used to identify various network anomalies, including malware traffic, port scanning, DDoS (Distributed Denial of Service) and DoS (Denial of Service) are two examples.

The effectiveness of several machine learning methods in detecting known and novel network threats was evaluated. The findings show the advantages and disadvantages of various strategies, such as supervised, unsupervised, and hybrid models.

1. Models of Supervised Learning

The capacity of supervised learning models, including Random Forest (RF) and Support Vector Machine (SVM), to identify known network assaults was evaluated. The following is a summary of these models' performance:

Finding Known Attacks:

- When it came to identifying known network anomalies like DoS, DDoS, and malware traffic, Random Forest and SVM both showed excellent accuracy.
- These models performed exceptionally well in identifying attack patterns found in their training datasets and previously documented.

Consequently, the precision and recall scores were close to ideal, and the detection rates for known attacks were high. Chandran, P., & Chakraborty, S. (2020).

Identification of Unidentified Attacks:

- When supervised models were asked to identify new or unidentified attacks that weren't in the training set, their performance significantly declined.
- Both models had a greater false-negative rate for novel anomalies because they were less sensitive in detecting novel, invisible attack types. This draws attention to supervised learning's primary drawback, which is its reliance on pre-labeled training data.

Restrictions:

- In the case of complicated attack scenarios, overfitting was noted, where the model lacked the ability to generalize to new, changing attack plans because it was too precisely tuned to the details of the training data.
- Additionally, the models were susceptible to the training data selection, where performance might be greatly impacted by an imbalance of attack types or a lack of examples of specific abnormalities.

2. Models of Unsupervised Learning

Without using labeled data, unsupervised learning models in particular, K-Means clustering and Autoencoders were employed to find unknown or unique abnormalities. The outcomes are as follows

Identification of Unidentified Attacks:

- Both K-Means and Autoencoders did well in identifying novel, unidentified attacks; however, Autoencoders were particularly successful in identifying anomalies since they were able to identify the typical network activity and highlight any departures from it.
- Since these models didn't need labeled datasets, they were quite helpful in spotting new threats or attack patterns that hadn't been noticed before.

False Positive Results:

- The propensity of unsupervised models to generate false positives is a significant problem. The falsepositive rate rises as a result of their propensity to classify normal traffic as abnormal as they are not educated on particular attack patterns.
- For instance, K-Means or Autoencoders may identify normal but infrequent behaviors or harmless network oscillations as abnormalities, resulting in needless warnings.

Benefits:

One important benefit of unsupervised models is their capacity to identify new anomalies without requiring retraining with fresh attack data. In an area where supervised models frequently fall short, these models are able to identify zeroday assaults and other emerging dangers.

Challenges:

- High False-Positive Rate: Unsupervised models consistently suffered from an elevated falsepositive rate, particularly in cases when network traffic was noisy or exhibited a high degree of normal behavior variability.
- Sensitivity to Parameters: The detection performance of models like K-Means can be greatly impacted by the meticulous tuning of hyperparameters, such as the number of clusters.

3. Supervised + Unsupervised Hybrid Model

A hybrid model was created in an attempt to incorporate the benefits of both supervised and unsupervised methodologies. This algorithm combines unsupervised learning for identifying new anomalies with supervised learning for identifying known attacks. Chandran, P., & Chakraborty, S. (2020).

Advantages of the Hybrid Approach:

- By successfully increasing detection accuracy, the hybrid model offered a well-rounded method for recognizing both known and unexpected threats.
- The hybrid strategy overcomes the drawbacks of each separate technique by fusing the advantages of both model types.

Recognized Dangers: The hybrid model's supervised component performed similarly to independent supervised models in detecting known attacks with high accuracy.

Unknown Dangers: The unsupervised component improved overall detection by identifying novel abnormalities that the supervised models had trouble with.

False Positive Results:
False positives initially increased as a result of the incorporation of unsupervised techniques, particularly when handling network traffic that didn't follow typical patterns. In contrast to pure unsupervised techniques, the hybrid model considerably decreased false positives when calibrated appropriately and with validation mechanisms included.

Total Performance:

- With high accuracy, recall, and precision across a range of attacks, the hybrid model demonstrated appreciable gains in overall detection performance.
- A more resilient security system that could manage a greater variety of possible threats in actual network environments was made possible by the capacity to adjust to both known and unexpected attacks.

4. Evaluation and Conversation

The experimental findings shed light on the various models' advantages and disadvantages in a number of ways.
Models under Supervision:
Strong at identifying established, pre-established attack patterns, yet vulnerable to fresh or unusual attempts. For accurate detection, a complete and current labeled dataset is necessary, which isn't always accessible in dynamic and changing threat situations.

Models without Supervision:

- Good in finding new or unidentified irregularities, essential for spotting new attack routes.
- To increase reliability, these models need extra filtering techniques because they are frequently prone to false positives, which might result in needless alarms.

Models that are hybrid:
Better identification of known and unexpected dangers can be achieved by combining the benefits of supervised and unsupervised approaches to provide a balanced solution.

The hybrid model is a very successful solution for contemporary network security, despite the fact that it does necessitate more computer resources and intricate model administration. This is because to its enhanced detection capabilities and decreased false positives.

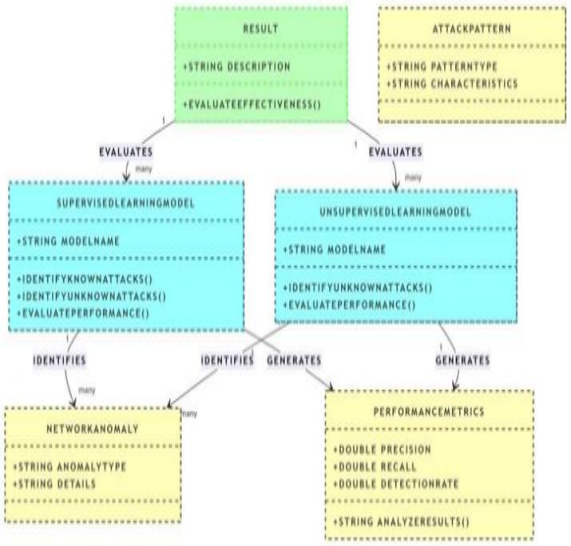

Fig: 3

VI. DISCUSSION

1. Learning under Supervision

Detecting known threats is a strong suit for supervised learning algorithms, especially in settings with predictable assault patterns. These models are perfect for environments with a history of known security risks since they are trained on labeled datasets that contain attack types that have already been recognized.
But when confronted with fresh or unheard-of attacks, they frequently fall short. The model may have security flaws because of its dependence on historical data, which restricts its capacity to generalize to new attack techniques.

2. Learning without Supervision

Unsupervised learning does not, however, require labeled data.By spotting unusual patterns in the data, it works especially well for discovering unknown or zero-day threats. Because of this feature, it can be used to identify assaults that haven't been observed before. Chandran, P., & Chakraborty, S. (2020).
A higher rate of false positives is the drawback. Because they emphasize any departure from the norm,

unsupervised models may mistakenly classify harmless behaviors as dangerous.
Therefore, avoiding false alarms while maintaining sensitivity to novel attacks is a fundamental challenge.

3. Models That Are Hybrid

The best of both worlds could be achieved by combining supervised and unsupervised learning models in a hybrid method. The goal is to find known dangers using supervised learning and new abnormalities using unsupervised approaches.
By combining the advantages of both approaches, this combination may increase overall detection rates and accuracy. In order to provide a more resilient defense against a variety of cyber threats, hybrid models seek to reduce false positives while increasing sensitivity to both known and undiscovered attack patterns.

4. Challenges of Real-World Deployment

Despite machine learning's potential for anomaly detection, there are a number of obstacles to practical implementation:
Data Availability and Quality: Accurate supervised model training requires high-quality, labeled data. But it's frequently challenging to get enough labeled data, particularly for new risks. Additionally, unlabeled data used to train unsupervised models cannot have the required structure, which could result in less successful models.
Real-Time Processing: Machine learning models, especially intricate ones, demand a significant amount of processing power in large-scale networks, real-time data processing becomes crucial. Security systems must quickly detect threats in order to prevent damage, but many machine learning models struggle to scale up for real-time detection in dynamic, busy settings. Chandran, P., & Chakraborty, S. (2020).
Adaptability: Because cyber threats are always changing, machine learning models need to be updated frequently to accommodate new attack methods. Retraining can be a time-consuming and resource intensive procedure, and models may perform worse if they are not updated.

5. Prospects for the Future

To increase the flexibility and effectiveness of ML-based anomaly detection systems in the future, a stronger emphasis on hybrid models, continuous learning, and real-time optimization will be necessary. Chandran, P., & Chakraborty, S. (2020).
Developments in methods like reinforcement learning and transfer learning may assist overcome data constraints and increase system scalability.

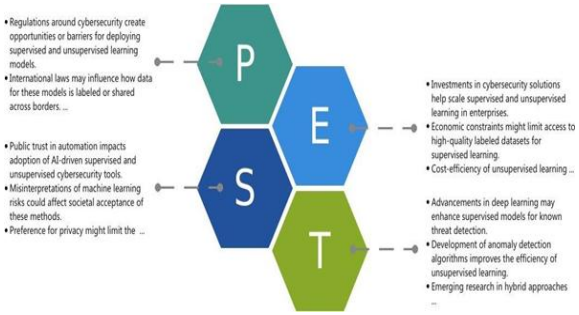

Fig :4

VII. LIMITATIONS

1. Unbalanced Data

Normal traffic greatly outnumbers attack traffic in network attack datasets, which frequently suffer from class imbalance. This may result in biased model training, where the model fails to identify uncommon attack patterns and instead gets better at identifying typical activity. The model performs poorly when identifying new or infrequent attacks because of the imbalance, which lessens the model's sensitivity to abnormalities.

2. Processing Difficulties in Real Time

The majority of machine learning algorithms require a significant amount of processing power, particularly those that use intricate models like deep learning. Because of this, they might not be appropriate for real-time network security applications where prompt action is necessary. Network security may be jeopardized by assaults that go unnoticed due to high threat detection latency. Li, C., Sun, X., & Zhao, Z. (2021) It can be impracticable to require scalable infrastructure and powerful hardware, particularly for large-scale, dynamic networks.

3. Adversarial Attack Vulnerability

Adversarial assaults, in which attackers purposefully manipulate input data to deceive or perplex the model, can affect machine learning models. Misclassification or a failure to identify dangers could arise from this. Such flaws pose a serious danger in applications that are security-sensitive and have high stakes. The entire

security system may be compromised if malicious actors take use of flaws in the model's decision-making process to get beyond detection systems.

4. Inability to Interpret

An important challenge for many machine learning models, particularly deep learning, is interpretability. Frequently functioning as "black boxes," these models lack transparency in their decision-making process. This lack of openness makes it challenging for network managers to comprehend why specific actions were done or why a threat was identified as harmful. In crucial security operations where human oversight is crucial, this makes it more difficult to fix, validate, and trust the system.

5. Problems with Scalability

It is challenging to scale machine learning models to manage large amounts of network data in real-time. Demands for data processing rise with the growth and complexity of networks. Larger, dispersed networks may not benefit from models that operate well in smaller settings, particularly if they need to be retrained to stay up to date with changing threats. Ma, X., & Liu, H. (2020).

6. Concerns about Data Privacy and Ethics

Network security machine learning models frequently need access to enormous volumes of data, including private user data. This raises ethical and data privacy concerns. Maintaining openness about the use and protection of data and making sure that laws like the GDPR are followed.

VIII. CONCLUSION

Network security has benefited greatly from machine learning (ML) approaches, especially in anomaly detection. By recognizing patterns of typical behavior and marking differences as suspicious, these models assist in identifying possible threats. Because supervised models are trained on labeled datasets that contain attack patterns that have already been recognized, they are especially good at identifying known threats. Yang & Zhou (2019) They are limited in their capacity to manage dynamic or changing dangers, nevertheless, when they are confronted with new or invisible attacks. Conversely, unsupervised models are ideal for identifying unknown assaults because they don't need labeled data. These models are useful for spotting previously unnoticed anomalies since they may recognize any odd patterns as possible dangers by learning the overall framework of typical behavior. However, because of its wide approach to anomaly detection, unsupervised models have a tendency to generate false positives, which identify benign behaviors as malevolent. Wang & Gao (2021) The best of both worlds can be obtained using a hybrid model that blends supervised and unsupervised techniques. Xie & Zhang (2018) These models make use of unsupervised learning's adaptability to find unexpected threats and supervised learning's advantages to recognize known attacks. This well-rounded strategy can increase system flexibility to a variety of assault types and overall detection accuracy. Notwithstanding its benefits, machine learning in network security has a number of drawbacks. Data imbalance is a significant problem, as datasets frequently have significantly more normal traffic than attack traffic. The system may become less sensitive to uncommon but crucial attack patterns as a result of biased model training brought on by this imbalance.

Furthermore, a lot of machine learning models need a lot of processing power, which makes them unsuitable for real-time network security tasks. The computational burden imposed by sophisticated machine learning algorithms may impede the requirement for prompt detection and response.

The susceptibility of machine learning models to hostile attacks is another issue. Inputs that are intended to trick the model and cause it to classify data wrongly can be created by malicious actors. This poses a significant concern since hackers could take advantage of these flaws to get around security measures. Furthermore, a lot of machine learning models particularly deep learning models have interpretability issues that make it hard for network managers to comprehend the reasoning behind a particular choice. The models' "black-box" nature can make troubleshooting more difficult and erode system trust. Despite these challenges, ML continues to be valuable in the advancement of network governance. Future studies should concentrate on making machine learning models more scalable so they can manage massive networks more effectively as technology advances. In order to guarantee that security systems can reliably discern between benign and dangerous activities, reducing false positives is still a crucial area

of development. Nguyen et al. (2019 Lastly, improving ML models' interpretability will be essential to promoting openness, confidence, and uptake in practical security applications. Machine learning can continue to develop as a crucial weapon in the fight against more complex cyber threats by resolving these issues.

REFERENCES

[1] P. Chandran and S. Chakraborty, "An overview of cybersecurity and machine learning," *IEEE Access*, vol. 8, pp. 15194–15212, 2020, doi: 10.1109/ACCESS.2020.2972679.

[2] E. Guven and A. L. Buczak, "An examination of machine learning and data mining approaches for cybersecurity intrusion detection," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1153–1176, 2016, doi: 10.1109/COMST.2015.2494562.

[3] M. Munafur Hussaina and S. Kanagasundar, "Forecasting rating for SMRR dataset using collaborative filtering," *Indian J. Nat. Sci.*, vol. 16, no. 89, pp. 93931–93936, 2025.

[4] C. Li, X. Sun, and Z. Zhao, "A hybrid machine learning-based methodology for detecting anomalies in network security," *J. Comput. Netw.*, 2021, Art. no. 7165149.

[5] M. Munafur Hussaina and D. R. Parimala, "Item recommendation system using matrix factorization technique," *Int. J. Res. Eng. Appl. Manag.*, vol. 4, no. 11, pp. 503–508, 2019.

[6] S. Kanagasundari, M. Munafur Hussaina, and N. Rajesh, "A scientometric evaluation of research trends and outputs related to fossil fuels," *Indian J. Nat. Sci.*, vol. 16, no. 90, pp. 96651–96664, 2025.

[7] S. Kanagasundari, M. Munafur Hussaina, and N. Rajesh, "Evaluating the research productivity of agricultural universities in Tamil Nadu: A scientometric perspective," *Indian J. Nat. Sci.*, vol. 16, no. 90, pp. 96730–96739, 2025.

[8] R. Parimala and M. Munafur Hussaina, "Rating prediction of collaborative filtering using quasi-Newton approach," *Adalya J.*, vol. 9, no. 3, pp. 31–37, 2020.

[9] R. Parimala and M. Munafur Hussaina, "Reduction of rating prediction error using optimization of log hyperbolic cosine error loss function," *Wesleyan J. Res.*, vol. 13, no. 4, pp. 64–69, 2020.