WCRG Global Research WORLD COMET RESEARCH GROUP

International Journal of Modern Innovations and Emerging Trends
Vol. 5, No. 1, 2022
Doi: 10.67228/30715628/IJMIET-2022PI3W7Z
PP. 01-14

Original Article

AI-Enabled Threat Detection in Network Security

Dr. Rajesh Kumar Sharma

Professor, University of Delhi, India.

Received: 03-03-2022 Revised: 03-25-2022 Accepted: 04-03-2022 Published: 04-05-2022

ABSTRACT

The expansion in the domains of digital communication, cloud computing, Internet of Things (IoT), and 5G technologies has greatly increased the attack area of the contemporary network infrastructure. The conventional network security measures, which rely mostly on fixed rules and signature-based responses, are becoming incapable of responding to more advanced, changing and zero-day cyber-attacks. Artificial Intelligence (AI) has become a disruptive technology that can be used to improve network security through enhanced and real-time and adaptive threat issues. This paper consists of an in-depth analysis of AI-powered threat detection when it is applied to guarantee network security, its principles, techniques, methodology, and the performance results. The paper is started with the discussion of the limitations of traditional intrusion detection and prevental systems and the necessity of intelligent automation in security problems. It has been thoroughly analyzed with the literature review of the modern progress in the fields of machine learning, deep learning, and hybrid AI applications to network threat detection. The suggested methodology is an AI-based design that includes the process of data gathering, feature engineering, model training, and classifying the threats. Several types of machine learning algorithms are analyzed with reference to their usefulness in identifying anomalous and malicious network behavior, such as supervised, unsupervised and deep learning models. The experimental analysis shows that AI threatened detection systems can greatly increase the accuracy of detection, a decrease in false positives and an increase in the response time in comparison to the usual methods. This discussion examines the performance metrics, scalability, and deployment issues in the actual environments. Lastly, the paper has come to an end by summarizing the major findings and stating the direction of future research moves, which will focus on the role of explainable AI, federated learning, and adaptive defense mechanisms in the next-generation network security systems.

KEYWORDS

Artificial Intelligence, Network Security, Intrusion Detection System, Machine Learning, Deep Learning, Cyber Threat Detection.

ISSN: 3071-5628
PP: 01-14.

1. INTRODUCTION

1.1. Background

The high rate of interconnected device proliferation, cloud computing, and massive interchange of data over public and private networks has set forth network security as a more critical issue in the contemporary age of the digital world. Digital infrastructures are essential to organizations in the financial, health, education, and government sectors, and are used on a daily basis to facilitate day-to-day operation, service delivery, and decision-making processes in organizations. Such an increasing reliance makes network environments superior destinations of cyber attackers aiming at exploiting a vulnerability in order to generate financial benefits, espionage, or disrupt services. Cyberattacks like Distributed Denial of Service (DDoS), malware spread, ransomware, phishing, advanced persistent threat (APT), may cause a major threat to the network resource confidentiality, integrity and availability, leading to data breach, loss of money, and reputation damage. Firewalls, antivirus software, rule-based intrusion detect systems (IDS) and other traditional network security agents have been used many years as the first-line defense against cyber threats. Although these methods are prove helpful in identifying the signature of an attack and applying the previously designed security policy, they also show significant weaknesses in the contemporary dynamic world of threats. These systems are very dependent onstatic rules and signature databases and they need to be regularly updated manually to be effective. Since cyberattacks are constantly becoming quite sophisticated and complex to approach, older security solutions tend to fail in the ability to discover zero-day attacks and pattern threats that have never been witnessed before. The inflexibility and failure to adapt to new data shows a necessity of smarter, automated, and adaptable security solutions that will serve the new menacing networks. Artificial Intelligence and Network Security.

1.2. Role of Artificial Intelligence in Network Security

Artificial Intelligence (AI) is a revolution in contemporary network security since it brings information-driven intelligence and automation to the process of threat detection and response. AI based security systems can examine large quantities of network traffic and identify concealed patterns as well as identify anomalies real-time by utilizing machine learning (ML) and deep learning (DL) algorithms to analyze the network traffic patterns in detail. In contrast to traditional rule-based systems, AI-enabled systems can also learn more continuously out of data, allowing them to respond to changes in attack techniques as well as in dynamic network conditions. This is what renders AI one of the main facilitators of active and robust protection of the network.

1.2.1. Adaptive Learning Capabilities

Security systems based on AI have the capacity to develop adaptive learning which enables them to adapt with the changing behavior of the network, which is extended to new cyber threats. Machine learning algorithms can also use constant training and updating of the model to detect new attack patterns without the need to adapt rules manually. This flexibility improves the system with regard to the detection of zero day attacks and also it minimizes the time gap between the introduction and detection of a threat.

ISSN: 3071-5628
PP: 01-14.

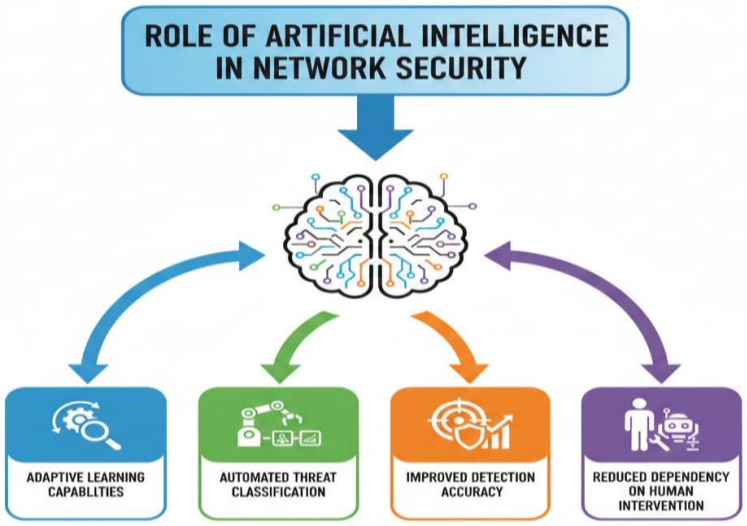

Fig 1 - Role of Artificial Intelligence in Network Security

1.2.2. Automated Threat Classification

The high-value feature of AI-based security systems is automated threat classification. Through the extraction of features out of traffic on a network, AI systems are capable of identifying the abnormal and normal processes automatically, and go further to classify the identified threats into a particular type of attack. This automation will eliminate the necessity of manual analysis, quicken reaction to incidences, and empower security personnel to concentrate on valuable and risky attacks.

1.2.3. Increased Detection Accuracy

AI-based models are very helpful in police-detecting accuracy because they learn to be used in complex and non-linear associations between the network traffic that consists of high dimensions. Deep learning methods, especially, are able to detect minute irregularities and advanced patterns of attacks that other systems might miss. The increased precision allows cutting down on the false positives and false negatives thus making security monitoring better and more trustworthy.

1.2.4. Less Reliance on human Intervention

The systems using AI cut down on the need and engagement of humans in data analysis, threat identification, and classification, as it is automated. This will reduce operational costs besides reducing human error and delays in responding. Consequently, intelligent systems will help security administrators to ensure continuous monitoring and quick detection of threats even in a large scale and high speed network.

1.3. Limitations of Traditional Threat Detection Systems

Conventional network threat detection mechanism has been commonly applied in the security of organizational network, but it has rampant weaknesses in dealing with the complexity and dynamism of the cyber threat in the present day. Traditional methods are mostly based on signature-based methods of detection, static sets of rules and manual analysis and response tools.

ISSN: 3071-5628
PP: 01-14.

Using signature-based detection, the potential threat is defined by the pattern of the traffic and compared with a set of known attack signatures in the database. Although effective in identifying the existing attacks, this approach cannot identify new, altered, or zero day threats that are yet to have the predetermined signatures. Since cyber attackers constantly develop their methods to get around detection, signature based systems are usually slower to adapt to new attack vectors. Traditional threat detection is also limited by the use of the non-adaptive rule-based systems. Such systems rely on the set rules designed by security professionals to detect suspicious activity. Rule sets are useful to express known malicious patterns but these are not flexible and have to be updated many times to be relevant. In a rapidly changing network environment where traffic patterns keep changing, the use of static rules commonly leads to high false positive or loss. Also, it is time consuming and requires extensive knowledge to upkeep and tune these rule sets and, therefore, large and complex networks face considerable challenges in terms of scalability. Another important weakness of the traditional threat detection systems is manual analysis and response. Interpretation of alerts, incident investigation and ultimately response actions can be time consuming and inaccurate, particularly when volumes of alerts are very high, thus security analysts are expected to be involved. Such dependency on human effort raises the time of response and can enable attacks to spread before measures can be put in place to counter the attack. Additionally, the increasing lack of qualified cybersecurity specialists builds on such obstacles. Taken together, these restrictions illustrate how traditional threat detection solutions have not been able to deliver response-to-threat benefits timely, change to meet demands, and scale to address the needs of modern intelligent, automated and learning-based security responses.

2. LITERATURE SURVEY

2.1. Machine Learning-Based Threat Detection

The detection of network intrusion activity has been widely analyzed based on machine learning (ML) techniques owing to their ability to be able to learn patterns and relationships automatically based on past network traffic data. Monitored learning systems like Support Vector Machine (SVM), Decision Trees, Random Forests, and k-Nearest Neighbors (k-NN) have performed well in classification of network traffic into normal and malicious. All of these are based on labeled datasets in order to construct predictive models that are able to recognize known attack patterns. Random Forests, such as others, are useful with high-dimensional data and are also more efficient in minimizing overfitting, whereas SVMs are able to separate complex decision boundaries. The effectiveness of supervised methods is, however, plagued by resources of high-quality labeled datasets, which are prohibitively expensive and time-consuming to produce. Moreover, supervised models can hardly learn new or zero-day attacks and generalize them to previously unnoticed behavioural patterns since they are sensitive to the patterns they have observed during training and cannot predict new behaviour patterns.

2.2. Unsupervised and Semi-Supervised Approaches

In a bid to break the constraints posed by labeled data, scholars have turned to unsupervised and semi-supervised learning fields with increasing frequency to carry out intrusion detection

ISSN: 3071-5628
PP: 01-14.

analysis. The unsupervised classifiers, including k-means clustering, Isolation Forests and Autoencoders, do not require any labeled data: Because the normal network traffic is understood and a deviation detected as a possible anomaly. These are the techniques that are very useful in identifying the zero day type of attack because they are a technique that is not based on the attack signatures. Autoencoders, such as, generate normal traffic modes with minimum error whereas anomalous traffic generate more reconstruction errors which is an indicator of possible intrusions. Semi-supervised methods also utilize larger amounts of unlabeled data to achieve better detection tasks through a limited number of labeled instances which allows the models to learn more robust representations. This point is because this hybrid approach provides a sensible compromise in terms of accuracy in detection and the cost of data labelling and is therefore appropriate in real-world network conditions where labelled information is limited.

2.3. Deep Learning Techniques

The past years have seen deep learning methods receiving much hype as these methods can capture non-linear relationships in high-dimensional network traffic data. Convolutional Neural Networks (CNNs) are typically employed to extract spatial features of traffic representations, e.g. sequence of packets or flow based features, whereas Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks prove useful in learning temporal dependencies and sequence effects of network behavior. These models are also able to acquire hierarchical feature representations automatically, with less manual feature engineering being required. Existing empirical research evidences that deep learning-based intrusion detection systems are frequently more accurate and robust in their detection than more traditional machine learning based methods. These advantages, however, are accompanied by costs of higher computational complexity, longer training times and a high requirement of large scale datasets that may limit their usability in resource constrained or real-time applications.

2.4. Challenges Identified in Existing Research

Although machine learning and deep learning-based intrusion detector systems have made significant progress, the existing studies still have some challenges that have not been solved. Imbalance of datasets is one, in the common case, normal traffic greatly exceeds malicious cases, resulting in biased models and decreased detection rates of those rare types of attacks. Also, the unintelligibility of complex models, especially due to deep architecture, renders it hard to comprehend detection results of the system and rely on automatic systems by security analysts. Computational overhead is also an issue because most of the updated models only run on powerful systems and might not be applicable in the real time in large-scale networks. Lastly, few studies cover real world real time deployment applications, and majority of the tests are done on benchmark data sets in a controlled setting. The challenges will have to be considered to create feasible, extensible, and reliable intrusion detection solutions.

ISSN: 3071-5628
PP: 01-14.

3. METHODOLOGY

3.1. System Architecture

The suggested AI-driven threat detection system can be defined as a multi-phase pipeline, which operates on network traffic data in a systematic manner to determine the existence of the security threats as accurately as possible. The importance of every step followed by the architecture is to have a reliable, efficient, and scalable intrusion detection.

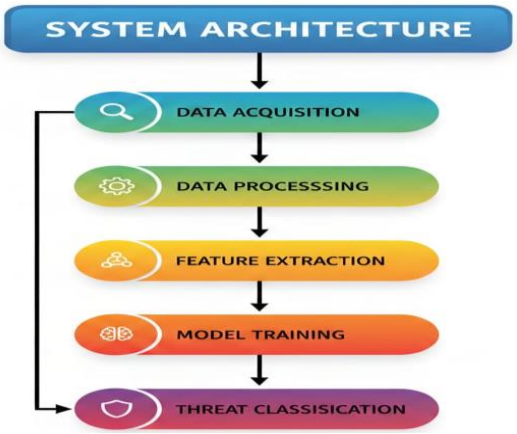

Fig 2 - System Architecture

3.1.1. Data Acquisition

The first step of the framework is data acquisition whereby raw network traffic data are obtained in different channels through different sources, including network sensors, packet sniffers, firewalls, or publicly accessible intrusion detection data. The obtained data can be packet-level data, flow statistics, protocol data, and connection data. This step would guarantee the system is able to get all-inclusive and representative network traffic demanded to perform effective threat analysis.

3.1.2. Data Preprocessing

During the data preprocessing phase, the obtained raw data undergoes the process of cleaning and conversion so that the data quality and consistency could be enhanced. These involve the management of missing or noisy data, deleting irrelevant and redundant data, normalizing numerical data, and encoding nominal data. The preprocessing stage of the process minimizes the dimension of data and eradicates errors making machine learning models to learn significant patterns with high efficiency.

3.1.3. Feature Extraction

The aspect of feature extraction is meant to obtain informative features on network traffic data once it has been preprocessed and which provide the best representation of both normal and malignant behavior. Commonly extracted statistical, temporal and protocol based features include the packet rate, duration of the flows, number of bytes and the flag values. Effective feature extraction is a process that increases the performance models by focusing on discriminatory features at the expense of minimizing the computational complexity.

ISSN: 3071-5628
PP: 01-14.

3.1.4. Model Training

During the model training stage, the features extracted are trained on machine learning or deep material to identify patterns that relate to the various forms of network threats. Supervised, unsupervised, or semi-supervised algorithms are used depending on the method of learning applied. The training procedure consists in the optimalization of the model parameters, based upon historical data, in order to have high detection accuracy and high-level robustness.

3.1.5. Threat Classification

The last step of the system is the threat classification where the trained model examines the incoming network traffic and either classifies it as normal or malicious. In the case of malicious traffic, even more specific types of attacks can be identified by the system. These classification outcomes allow raising alerts the applicable response to security concerns in a timely manner and thus make the framework applicable to real time or close to real time intrusion detection.

3.2. Data Collection and Preprocessing

The success of the proposed threat detection system greatly depends on the quality of the input data. The process of gathering network traffic data and the methodology of preprocessing it so that it can be used in machine learning is outlined in this section.

Fig 3 - Data Collection and Preprocessing

3.2.1. Data Collection

The packet capture tools and network sensors on the monitored network environment are used to capture network traffic data. These are packet-level and flow-level tools that contain the source and destination addresses as well as protocols, timestamps, and statistics related to the payloads. The obtained data give the real picture of network behavior during the normal operation and attack modes, and this is the basis of detection of the threat.

3.2.2. Noise Removal

Noise elimination is undertaken to clear irrelevant, redundant or corrupted data that can have adverse effects on the performance of the model. These involve eliminating redundant records, incomplete visits and outliers such as loss of packets or inaccurate reads. The system enhances reliability of extracted features by reduction of false detections that arise due to the noise that can be filtered.

ISSN: 3071-5628
PP: 01-14.

3.2.3. Data Normalization

Data normalization is used to bring the numerical data into a similar range so that attributes whose numerical values are high do not control the learning process. Minimum and maximum normalization (or z-score standardization techniques) are often employed in order to make features equal. Normalization increases the convergence rate of the model and also increases the overall accuracy of classification.

3.2.4. Handling Missing Values

Missing values are a critical preprocessing technique because such incomplete data would result in biased or flawed predictions. Losses of values can occur as a result of sensor failures or fragmented packet captures. Such values are filled in with the mean or median imputation, forward fill, or deletion of contaminated records depending on the volume of the lost data. Maintaining the consistency of data and model sensitivity are guaranteed by proper treatment of missing data.

3.3. Feature Engineering

The data provided on feature engineering is very important to the efficiency of the suggested AI-based threat detection system because the quality of extracted features is a direct determining factor on how the model will be able to identify normal and malicious network behavior. The raw and preprocessed network traffic data are converted in this stage to a structured set of meaningful features that represent the statistical as well as behavioral characteristic of network flows. Effective feature engineering maximizes accuracy of detection and minimizes the complexity of models and compensates computational costs. Such features as protocol type, connection length, and flow-level statistics are extracted to allow representing a holistic view of the network activity. Packet size characteristics, such as average packet length, maximum and minimum packet size, and packet size variance are used to detect evidence of abnormality in the traffic pattern like flooding or fragmentation attacks.

Features based on protocol types record the details on the use of the communication protocols (e.g. TCP, UDP, ICMP) and the system can distinguish between the normal use of the protocols and protocol-based attacks. Connection duration features show the duration of communication sessions and can be used especially well to detect both suspicious long lived and suspicious short lived connections that may be due to denial-of-service attacks or scanning. Besides those simple features, flow statistics to model traffic intensity and behavior at times include: packet rate, byter rate, inter-arrival Time and number of packets per flow. The features enable the detection system to identify abrupt spikes, abnormal transmission habits and departure of the normal traffic baselines.

Encoding feature encoding is used to transform categorical attributes into numerical forms, and feature scaling is used to bring about similarity across the range of various features. In general, feature engineering process processes raw network data to form a discriminative state of features that increase the efficiency of model learning, generalization to previously unseen attacks, and effective and reliable threat classification.

ISSN: 3071-5628
PP: 01-14.

3.4. AI Model Training

The step of AI model training is aimed at developing and testing various learning models to find the most successful one to detect the threats on the network. Various types of models are trained to get different types of attack patterns and enhance the general detection performance.

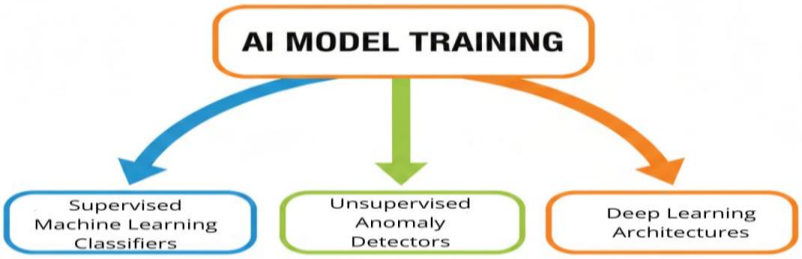

Fig 4 - AI Model Training

3.4.1. Supervised Machine Learning Classifiers

Training of supervised machine learning classifiers is done with picture network traffic data that is labeled to obtain the difference between regular and malicious activities. Support Vector Machines, Decision Trees, Random Forests, k-Nearest Neighbors, etc. are the algorithms that are generally used because they have been proved effective in terms of intrusion detection jobs. These models train on learning decision limits using extracted features and known attack labels. The metrics applied to assess the performance are the accuracy, precision, recall, and F1-score to make sure that there is a reliable classification of the known types of threats.

3.4.2. Unsupervised Anomaly Detectors

Unsupervised anomaly detection models are using the unlabeled or normal traffic data to learn normal behavior of networks only. Several methods used to detect a deviation of normal patterns that may be considered intrusion include k-means clustering, Isolation Forests, and Autoencoders. Such models can be especially useful in identifying the threats of zero-day attacks and unknown threats in the past. Their performance is gauged by the use of anomalies scores and detection rates, and they are focused on dynamism in changing network surroundings.

3.4.3. Deep Learning Architectures

Data in the form of complex and high-dimensional network traffic is automatically trained deep learning architectures to learn high-level representations. Convolutional Neural Networks and Long Short-Term Memory networks are models that do not require much feature engineering to capture spatial and temporal patterns of traffic. Even though such models generally have a greater detection accuracy, they have increased data size, duration of training and demand more computing power. The appropriate hyperparameter tuning and regularization methods are used to improve the model generalization and to avoid overfitting.

ISSN: 3071-5628
PP: 01-14.

3.5. Threat Detection and Alert Generation

The last operation phase within the proposed AI-based intrusion detection framework is threat detection and alert generation, where trained models have been placed to scan through the live or close-to-real-time network traffic. The data in the incoming network is constantly captured, preprocessed and converted to feature vectors aligned with the training stage. These feature vectors are sent to the trained machine learning or deep learning models, which take into consideration the traffic behavior and each network flow or packet is considered as normal or malicious. This automatic classification can be used to detect suspicious activities with minimum human intervention and at the same time, the level of detection is high. Upon detecting malicious activity, the system triggers an alert generation process to be in a position to generate alert in time. Alerts include necessary data, like the nature and origin of the identified threat, origin and destination IP addresses, protocol data, time, and intensity. The level of severity will be determined according to the confidence scores, the magnitude of the anomaly, and its possible effects on the network resources. This priority allows security administrators to attend to high-risk incidents, whereas it minimizes fatigue due to the number of false positives. The alerts are sent to security administrators as either centralized monitoring dashboards, log management systems or notification mechanisms e.g. email or messaging services. The integration allows analyzing an incident quickly and helping to take corresponding prevention measures, such as blocking traffic, ending the session, or changing the policy. Also, events identified could be recorded to be analyzed and reported on a forensic basis. The proposed system positively affects situational awareness, the speed of response to the threat, and contributes to the overall security position of the network to both established and new cyber threats by ensuring the accurate threat classification in conjunction with the effective alert generation and dissemination.

4. RESULTS AND DISCUSSION

4.1. Performance Metrics

The effectiveness of the proposed AI-based threat detection system is measured against standard classification measures that will give a holistic analysis of detection effectiveness and reliability. The metrics are based on the confusion matrix, which will consist of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). All the metrics bring out a different facet of system performance and play a crucial role in testing intrusion detection systems working in security sensitive environments.

The accuracy is used to determine whether the model is generally correct, it is a percentage of the model predictions that are correct. It is defined as

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

Although accuracy gives a rough level of performance of the models, it is misleading in the case of intrusion detection since in most cases datasets are very imbalanced with normal traffic greatly exceeding malicious traffic. Precision is the number of falsely identified malicious instances of all malicious instances predicted. The high level of precision means that the false positive rate is low,

ISSN: 3071-5628
PP: 01-14.

which is crucial to decrease the number of unnecessary messages and alert fatigue among security administrators. In operational settings, accuracy is found to be of great essence because in the instance where there are too many false alarms, this will disparage the reliability of automated devices. As previously mentioned, recall, or detection rate, sensitivity is the technical expression of how the system correctly recognizes actual malicious traffic. It is a ratio of the actual attacks that are successfully identified by the model. Recall is vital in reducing false negatives since attack detection through which attacks have not been detected may lead to the occurrence of critical security breaches and compromise of the system. F1-Score is the harmonic mean of the precision and recall, and it gives a balanced score on the detection performance. It comes in exceptionally handy when one is working with disbalanced data sets since it takes into consideration both false positive and false negatives. Collectively, these measures will provide a sound assessment framework that will guarantee the proposed system will deliver sound, credible, and workable performance in regard to threat detection.

4.2. Comparative Analysis

Performance of the various AI models in the proposed threat detection system is evaluated by conducting a comparative analysis. Comparison is done based on the key performance metrics, such as accuracy, precision, and recall, to compare the efficacy of each model in the detection of malicious network traffic.

Table 1: Comparative Analysis

Model	Accuracy (%)	Precision (%)	Recall (%)
SVM	91.2	89.4	88.7
Random Forest	94.8	93.1	92.6
Deep Learning	97.3	96.8	96.1

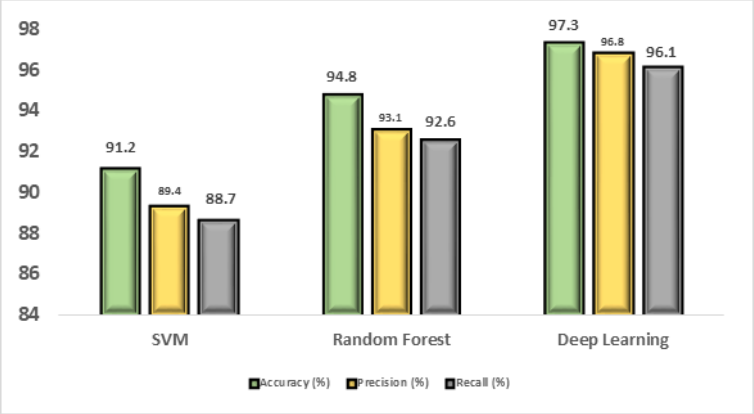

Fig 5 - Comparative Analysis

4.2.1. Support Vector Machine (SVM)

Support Vector machine model gives an accuracy of 91.2, and its precision and recall are 89.4% and 88.7 percent respectively. These findings show that the performance of SVM is relatively good in terms of differentiating normal and malicious traffic, especially in the case of familiar

ISSN: 3071-5628
PP: 01-14.

patterns of attacks. Nevertheless, its comparatively low recall implies the existence of some evil cases that are not noticed. This is a limitation that could be as a result of the sensitivity of SVM to features selection as well as parameter tuning particularly in cases of complex and high dimensional network traffic data.

4.2.2. Random Forest

Random Forest model shows better results than SVM, and its accuracy is 94.8, precision 93.1 and recall is 92.6. Random Forest has the ensemble feature that allows the algorithm to better manage variability in features as well as data imbalance. Its greater accuracy in situates that there is a decrease in false positives whereas the enhanced recall signifies the increased malicious traffic detection. These features make random forest a strong and valid consideration in dealing with intrusion detection activities in a real network environment.

4.2.3. Deep Learning Model

The deep learning model performs better as compared to both traditional machine learning models, and has the highest accuracy of 97.3, precision of 96.8 and recall of 96.1. This outperformance indicates that the model is capable of capturing non-linear patterns and time based dependencies in network traffic which is complicated. Even though deep learning models consume more computation resources and take more time to train, their high detection ability renders them adequate in high degree and extensive threat detection frameworks.

4.3. Discussion

The findings of the experiments clearly show that the approaches based on deep learning are superior in identifying the presence of the complex and sophisticated network attacks as compared to the traditional machine learning approaches. The deep learning model exhibited the best accuracy, precision, and recall which means it has high learning power with complex patterns and non-linear associations that exist in high-dimensional network traffic data. This is especially relevant to detecting highly evolved and multi-phase attacks, which might not be represented by signatures that can be easily distinguished. Deep learning models learn hierarchical representations of features automatically, so that they are not required to engineer features manually and can generalize to a broader range of attack variants. Along with these benefits, significant trade-offs related to the deep learning-based intrusion detection systems are also presented. The main issues are that training and deployment are expensive in terms of computational cost. Deep learning models are typically very CPU-intensive, require a large amount of memory and take longer to train, potentially compromising their practicality on resource-limited hardware or real-time detection. Conversely, conventional machine learning models that include SVM and Random Forest have shorter training time, and reduced stage of computation, which render the tools more adaptable to lightweight and edge-based implementation. The other major challenge is interpretability. Whereas traditional machine learning models can theoretically give better understanding of the feature importance and decision making, deep learning models are complex black-box models. This non-disclosure can lead to lack of trust and adoption especially in critical security applications where it is necessary to know the reasoning behind the decisions taken during detection in order to analyze them forensically and other

ISSN: 3071-5628
PP: 01-14.

compliance measures. As such, despite the better detection accuracy of deep learning models, they should be cautiously adopted on large systems taking into consideration system resources and explainability. Future studies ought to dwell on streamlining the architecture of deep learning and enhancing the interpretability of models and hybrid models that embrace accuracy, performance and explainability.

5. CONCLUSION

The paper has offered a detailed investigation on AI-enabled threat detection in network security, in which intelligent approaches have increasingly become a significant mode of managing the shortcoming of traditional security mechanisms. The proposed framework can be used to enhance the accuracy and scalability of detection and adhere to the current needs of emerging cyber threats by employing machine learning and deep learning techniques. Machine learning models are useful at classifying familiar patterns of attacks, whilst deep learning designs are useful at capturing non-linear, complex and temporal patterns of network traffic, allowing the identification of highly-advanced and novel attacks. The system architecture as a modular system which includes the data acquisition and preprocessing, features engineering, model training, and alert generation provide a systematic and effective threat detection system that is viable in contemporary and high speed network conditions. The findings of the experiment prove the usefulness of AI-based intrusion detecting systems, and deep learning models are more vigorous as compared to traditional machine learning algorithms regarding accuracy, precision, and recall. These results verify the fact that intelligent models are more suitable to process high dimensional and dynamic network data. Nonetheless, the research also finds significant problems, including the inability to handle computational overhead, the ability to interpret a model and the possession of constraints in real-time. As such issues are important, it would be necessary to address them so that AI-enabled security solutions can be adopted in practice in operational networks. The future study has the potential to extend the current knowledge in AI-based threat detection systems by introducing the Explainable AI (XAI) methods to increase the level of transparency and trust. The XAI techniques can help people comprehend the decision of a model and thus security analyst can learn why certain traffic is identified as malicious and they can assist in responding to an incident and conducting a forensic investigation. The next perspective is a promising option which is the adoption of federated learning which facilitates collaborative model training in distributed network settings without exchanging raw data. This method maintains the privacy of data and Asset security whilst enabling the models to connect with varied traffic patterns of various organizations. Also, it is essential to create lightweight AI models that optimize their resources and edges as well as to create them in resource-constrained settings so that they could detect threats in real-time even closer to the data sources. These models have the capability to lower the latency, bandwidth consumption, and dependence on the centralized infrastructure. In general, these avenues of the future are opening up opportunities to having a strong, transparent and scaleable system of AI-based network security, capable of providing resistance to new cyber threat forms.

ISSN: 3071-5628
PP: 01-14.

REFERENCES

[1] D. E. Denning, "An intrusion-detection model," *IEEE* Transactions *on Software Engineering*, vol. SE-13, no. 2, pp. 222–232, 1987.

[2] C. Modi, D. Patel, B. Borisaniya, A. Patel, M. Rajarajan, and A. Patel, "A survey of intrusion detection techniques in cloud," *Journal of Network and Computer Applications*, vol. 36, no. 1, pp. 42–57, 2013.

[3] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81–106, 1986.

[4] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[5] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.

[6] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[7] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *Proc. IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 2009.

[8] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[9] F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *Proc. IEEE International Conference on Data Mining*, 2008, pp. 413–422.

[10] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[12] R. Vinayakumar, K. P. Soman, P. Poornachandran, and S. Sachin Kumar, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525–41550, 2019.

[13] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[14] N. Moustafa and J. Slay, "UNSW-NB15: A comprehensive data set for network intrusion detection systems," in *Proc. Military Communications and Information Systems Conference (MilCIS)*, 2015.

[15] Z. Lin, Y. Shi, and Z. Xue, "IDSGAN: Generative adversarial networks for attack generation against intrusion detection," *IEEE Access*, vol. 6, pp. 2315–2325, 2018.