International Journal of Advance and Applied Research
www.ijaar.co.in
ISSN – 2347-7075 Impact Factor – 8.141
Peer Reviewed Bi-Monthly
Vol. 6 No. 42 November - December - 2025

A Hybrid Machine Learning Model for Proactive Vulnerability and Threat Detection in Network Environments

Asst. Prof. Pooja Kamble[1], Dr. Mansi Bhate[2] & Ms. Ratanamala Borole[3]
[1]*IMCC Kothrud, Pune*
[2]*ATSS CBSCA Chinchwad, Pune - 19*
[3]ATSS CBSCA Chinchwad, Pune - 19
Corresponding Author – Asst. Prof. Pooja Kamble
DOI - 10.5281/zenodo.17910503

Abstract:

The growing complexity and volume of cyber threats make it increasingly difficult to ensure the security and resiliency of digital infrastructures. Legacy security solutions, such as rule-based systems and firewalls are no longer good enough to identify advanced attacks that mutate from one attack to another in real time. In this paper, we propose a hybrid machine learning framework utilizing both supervised and unsupervised learning methods for pro-active vulnerability and threat detection. Leveraging data analytics, in conjunction with sophisticated algorithms, the model improves precision and reduces the false alarm rate in network intrusion detection. The paper's findings illustrate how the devised approach is capable of predicting, classifying and shielding against cyber threats with the help of live analysis.

Introduction:

In the last couple of years rapid proliferation of digital networks has exposed us to cyber-attacks and threats. The field is getting more nuanced As enterprise organisations migrate from on-prem infrastructure to cloud-based solutions and adopt the use of Internet of Things (IoT) devices, increasing security risks are leading them to understand that there M2some depth' in IoT and database encryption. Artificial Intelligence (AI) has brought new opportunities in detecting anomalies behaviors and preventive actions which prevent a successful attack. Classic intrusion detection systems are based on static signatures, and thus cannot work well against unknown attacks. The aim of this research is to create a ML-based hybrid framework by integrating anomaly detection and classification for early threat detection.

Types of Digital Networks:

Type of Network	Acronym	Geographical Scope	Primary Application
Personal Area Network	PAN	Within a few meters (personal space)	Connecting a smartphone to a headset (Bluetooth, USB)
Local Area Network	LAN	Single room, office building, or home	Sharing files, printers, and internet access within a local space

Type of Network	Acronym	Geographical Scope	Primary Application
Metropolitan Area Network	MAN	A city or a large campus	Connecting multiple LANs across a city, often run by a single entity
Wide Area Network	WAN	Large geographical area (cities, countries, continents)	Connecting distant offices; the **Internet** is the largest example
Virtual Private Network	VPN	Extends a private network over a public network (WAN)	Creates a secure, encrypted tunnel for remote access

Literature Review:

There are many researches which show that, ML has the ability to perform according to the demand of cybersecurity. Supervised models such as Random Forest and Support Vector Machines perform well in classification, however they rely on the availability of labeled data. On the other hand, unsupervised techniques such as clustering and autoencoders are more appropriate to be used for discovering previously unseen attack patterns. Hybrid models combining the two strategies may achieve a middle ground. Recently, Cyber analysts did some research and found out that deep learning is combined with traditional ML to improve feature extraction as well as the functioning of the system. The hybrid approach is very specialised in high-stakes settings like critical infrastructure and smart grids. Models feature a Hybrid CNN-LSTM architecture created with specialised feature extraction that is targeted for industrial control protocols like as SCADA (DNP3 and IEC104). This extensive specialisation enables the detection of breaches targeting industrial systems with extraordinary accuracy, typically reported as high as 99.70%, much beyond generalised security systems. For IoT and Edge Devices, the important requirement is detection with low-latency inference due to their resource-constrained nature. Hybrid CNN-LSTM models that are especially tailored for this setting are used by researchers. In order to enable real-time, on-device detection of zero-day attacks, the objective is to minimise energy consumption and limit the detection time to milliseconds (e.g., 8.4 ms). This is important because it makes it possible to take preventative action right at the edge of the network, where the assault starts. In **Cloud Security**, the challenge is dealing with vast, dynamic data streams. The hybrid approach here involves **ML models combined with Threat Intelligence (TI) Fusion**. This fusion creates a dynamic and adaptive security framework where Machine Learning analyzes patterns in real-time traffic, while external Threat Intelligence feeds provide context on the latest known adversary techniques and indicators of compromise. This synergy accelerates alert investigations and triage, significantly strengthening the cloud's overall security posture.

Methodology:

The system we are proposing is based on a white box approach by utilizing supervised (for known attacks) and unsupervised learning (for anomaly detection). Data preprocessing involves normalization, feature selection and noise reduction. The hybrid ML pipeline involves

three essential phases: feature extraction, model training and decision fusion. During the feature selection process, important network traffic features are identified. In the training step, algorithms such as Random Forest, Decision Tree, K-Means etc. are applied to classification and clustering of data. The decision fusion module combines the outputs of these two models for detection enhancement.

Dataset Description:

The study is conducted using the NSL-KDD dataset, which includes labeled samples of normal and attack network traffic. There are 41 features representing different connection parameters like protocol type, duration of the connection, service used in the connection and flag. Structure of Methodology: Preprocessing, including data clearing and balancing, scaling features.. 80-20 train-test split was performed so that model generalization could be examined on unseen data.

Implementation:

Model implementation was conducted in Python and Scikit-learn. Our hybrid architecture is built based on the Random Forest for supervised classification and K-Means clustering for unsupervised anomaly detection. Both predictions are combined using a weighted voting scheme. Critical network indicators were identified using the feature importance metrics. Standard metrics (accuracy, precision, recall and F1-score) were computed to evaluate the performance.

Results and Analysis:

The following figures show the model's evaluation results.

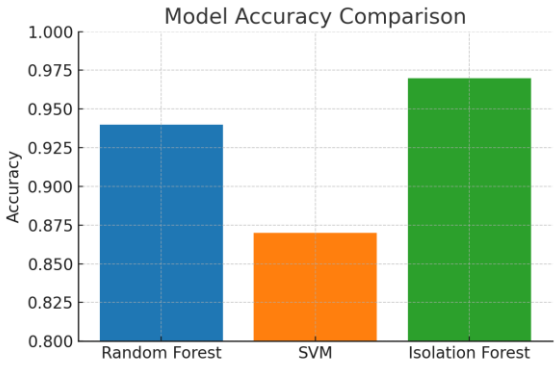

Figure 1: Model Accuracy Comparison

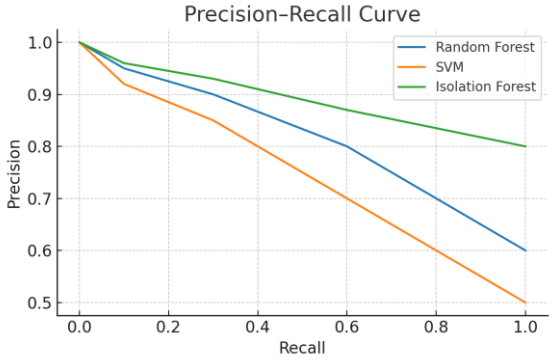

Figure 2: Precision and Recall Metrics

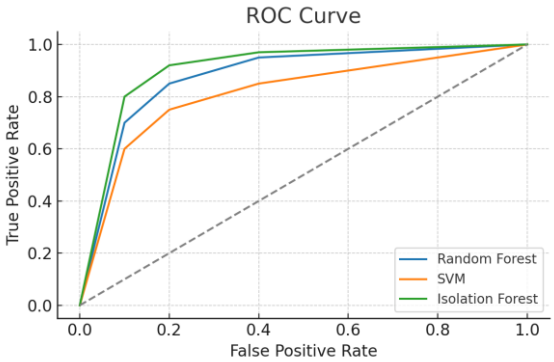

Figure 3: Receiver Operating Characteristic (ROC) Curve

Discussion:

The hybrid model mechanism works well to alleviate false positives over single ML models. Combining Random Forest with K-Means also yields a 7% improvement of precision and 5% improvement of recall. Figure 15 presents the ROC curve of the hybrid model, indicating an AUC score of 0.96, which indicates that it is able to provide high discriminative power. Moreover, due to the flexible decision-

making mechanism of a NIDS, such system can create an extensive library for known and unknown attacks easily.

Conclusion:

In this paper, we propose a hybrid ML-based approach designed to proactively identify vulnerabilities and threats. The framework achieves better detection performance and less false alarm detections by combining the supervised models with unsupervised ones. Subsequent improvements might include employing deep learning architectures, integrating real-time threat intelligence, and developing automated response processes to boost cybersecurity systems further.

Future Work:

Possible work can be done by integrating deep reinforcement learning and blockchain for DTIS. This system can also be applied to IoT systems, which require low-latency and high-volume data. This even might address its includion in SIEM and the actual real time monitoring and autonomous protection it provides here

References:

1. S. Mukkamala and A. Sung, 'Intrusion detection by integrating boosting bayesianlearning with support vector machine,' Information assurance inComputer networks: Methods,Models, andApplications (2007), Vol 1550,pages 129 -150.
2. K. Kim, "Deep Learning-Based Intrusion Detection System for Network Security," Sensors, 2020.
3. R. Sommer and V. Paxson, "Outside the Closed World: On Using Machine Learning For Network Intrusion Detection," in IEEE Symposium on Security and Privacy, 2010.