

# Adaptive Intelligence for Cyber Defense Evaluating Machine Learning Models for Real-Time Threat Detection

**Kalyana Krishna Kondapalli**

Technical Manager, Hyderabad, Telangana, India

**To Cite this Article** Kalyana Krishna Kondapalli, “Adaptive Intelligence for Cyber Defense Evaluating Machine Learning Models for Real-Time Threat Detection”, Indian Journal of Computer Science and Technology, Volume 05, Issue 01 (January-April 2026), PP: 317-325.



Copyright: ©2026 This is an open access journal, and articles are distributed under the terms of the [Creative Commons Attribution License](#); Which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Abstract:** With the ascendancy of digital infrastructure, the need for a good security is also on its increase. With the emergence of highly malicious cyber-attacks including ransomware, phishing, DDoS attacks as well as zero-day exploits, traditional security is becoming insufficient. With the advent of Artificial Intelligence (AI), we have witnessed an explosion in capabilities in real time response, predictive analytics, and the overall detection of threats. In this paper, we focus on the improvement of the AI evolution in place of the traditional cybersecurity models and also discuss the use of AI in automated defense mechanisms. It also analyzes the issues involved with cybersecurity driven by AI such as bias in the AI model, adversarial attacks and regulatory reasons. This research presents understanding of how and to what extent AI is adopted and utilized in cyber defense by employing a theoretical framework based on systems theory and the Technology Acceptance Model (TAM). The findings suggest that AI does play the role of mitigating cyber risks, and that some of the potential limitations AI should help improve the development of more resilient cybersecurity strategies.

**Key Words:** Cybersecurity, Artificial Intelligence, Threat Detection, Machine Learning, Cyber Threats

## I. INTRODUCTION

The crucial aspect of cybersecurity has emerged as a critical issue because our digital infrastructure continues to grow in use each day. Modern security practices are not sufficient in protecting organizations from complex cyberattacks because criminal hackers use constantly evolving vulnerabilities. The sophisticated cyber threats that include ransomware along with phishing along with Distributed Denial of Service (DDoS) attacks and zero-day exploits become more complex and more frequent which leads to system-wide disruptions.

Table 1. Common Cyber Threats and Their Impact

Threat Type	Description	Impact
Ransomware	Encrypts files and demands payment	Financial loss, data loss
Phishing	Fraudulent emails to steal credentials	Identity theft, data breaches
DDos Attacks	Overwhelms servers with traffic	Downtime, service disruption
Zero-Day Exploits	Attack on unpatched software vulnerabilities	High-risk breaches

The modern digitalization movement now targeting healthcare combined with finance and government sectors exposes these industry computers to intensified cyberattacks. The surface area of an attack has grown dramatically today to include technologies such as cloud computing, IoT, and smart devices that offer new challenges to cybersecurity. Increased dependency on interconnected technology systems raises the magnitude of security threats that can occur.

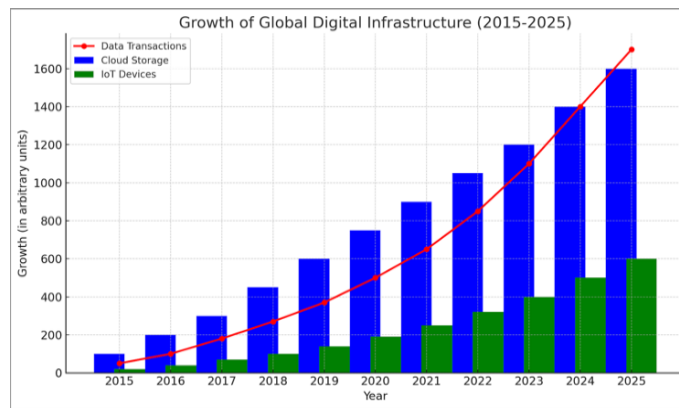


Fig. 1. Growth of Global Digital Infrastructure (2015-2025)

Cyberattacks generate important economic damage along with endangering both public security and national defense capabilities in addition to wrecking reputation values. Cyber threats generated substantial economic consequences and social impacts when Colonial Pipeline experienced the 2021 attack. Cybersecurity has seen Artificial Intelligence (AI) come into play and become a disruptive technology of threat detection and response automation. At the beginning of its development AI systems only performed spam detection alongside malware system checks. Today, predictive threat detection, real-time response, and better defensive techniques are possible using advanced machine learning (ML) and deep learning models [1, 2]. Security measures have become stronger due to this evolution in technology.

- AI systems handle information at rates faster than humans and help to find threats quickly.
- The technology makes better decisions by keeping its focus on real threats because it dismisses incorrect results.

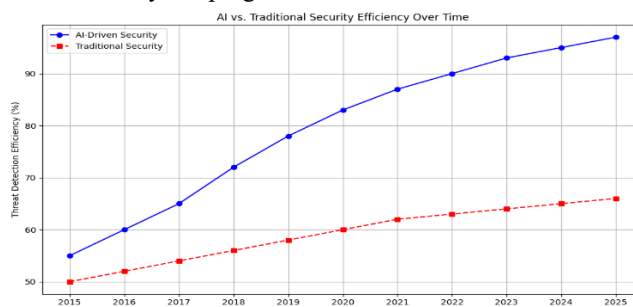


Fig. 2. AI vs. Traditional Security Efficiency Over Time

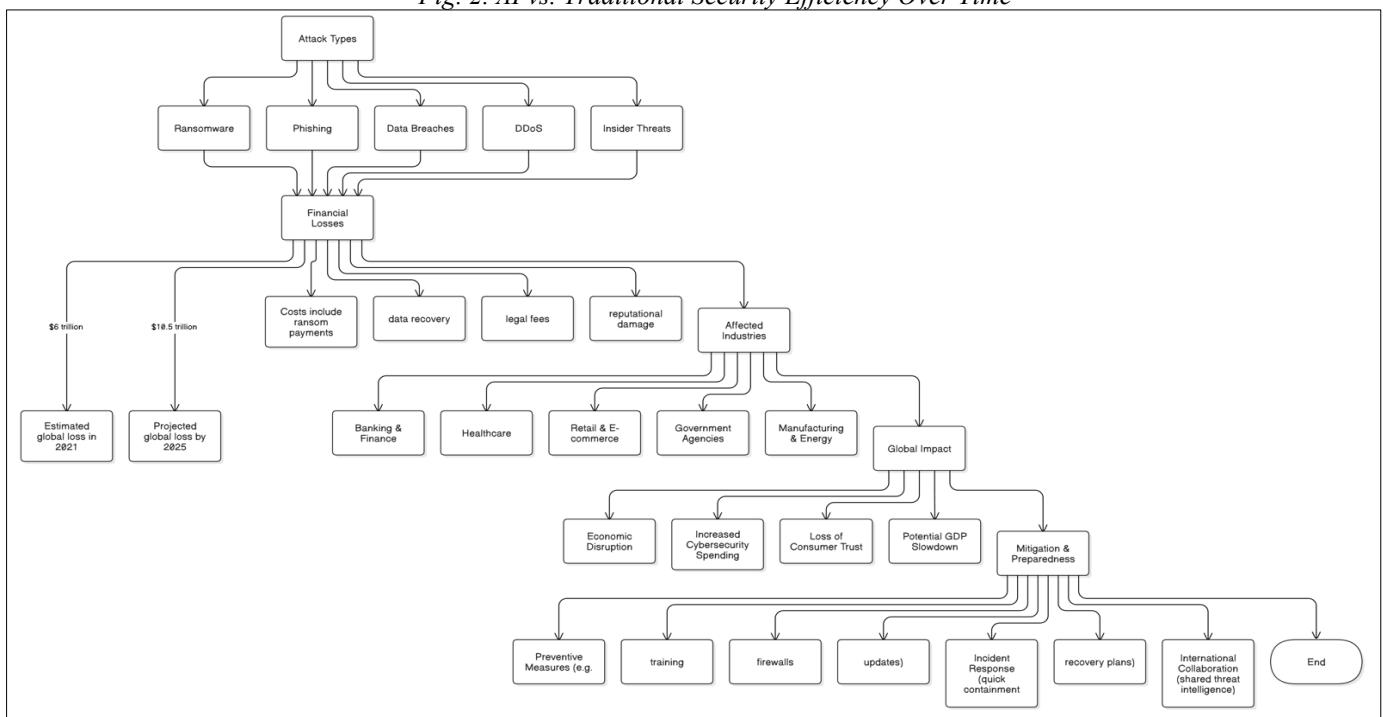


Fig. 3. Global Economic Losses Due to Cyberattacks

This research explores three main inquiries about AI security enhancement.

- What can machine learning and deep learning achieve for threat assessment that regular security tools cannot.
- Organizations should determine what AI-powered security helps them defend better and which obstacles they can expect when they deploy these systems.
- AI-based security systems help businesses respond faster to cyberattacks and cut down their damage.

The study wants to create helpful results through these goals:

- We explore how AI technology affects security strategy development [3, 4].
- Our research studies AI cybersecurity systems to validate their performance in current operations.
- The paper outlines practical methods to combine new AI technology into current security systems.

## II. LITERATURE REVIEW

The types of cyber threats have developed steadily since their first appearance as computer viruses. The development of internet technology created opportunities for more cybercriminals, so organizations need better defense systems. Simple computer infections like the Morris Worm appeared as the first cyber threats in early stages of internet security. With networks getting more advanced new hackers adapted their techniques to create ransomware bots and steal personal information [5].

### A. Common Cyber Threats

**Malware (Viruses, Worms, Ransomware):** The threat of malware has proved difficult for cybersecurity systems to handle because it contains several different forms including viruses, worms, and ransomware. Every piece of malware develops its own ability to replicate itself while encrypting data, stealing sensitive information or shutting down entire business operations

- Viruses are programs that connect to files before spreading after you run them.
- Worms reproduce automatically on computers as they move through network connections without human control.
- Cybercriminals use WannaCry to demonstrate their attack methods by encrypting files while demanding monetary payments for decryption [6, 7].

**Phishing and Social Engineering Attacks:** Phishing scams use psychological deception to make users reveal their private data. Attackers push victims into giving up their data through email trickery and social engineering while creating false websites. Spear-phishing attacks are now more targeted, and executives are hunted by Whaling attacks created from basic phishing methods [8]. The email network of the Democratic National Committee was compromised during the 2016 election year.

**Denial of Service (DoS) and Distributed Denial of Service (DDoS) Attacks:** Network attacks flood resources until they shut down service availability. Perpetrators use botnet armies to boost their onslaught against internet of things devices just like what happened during the Mirai malware attack which knocked out internet services across many areas [18].

### B. Machine Learning (ML) in Cybersecurity

**Supervised Learning for Threat Classification:** The supervised learning process spots digital security risks through trained datasets and labeled information. These systems find familiar attack traces and assist in both malware detection as well as network intrusion and phishing protection. Popular types of machine learning models that security analysts use are Decision Trees SVMs and Random Forests [10]. Unsupervised learning methods study network activities to find unusual behavior indicating potential security threats. Cluster-based solutions like k-means and DBSCAN analyze network behavior to discover unusual access points which may signal security breaches or traitor actions in an organization [11]. Artificial intelligence models can update their threat response automatically through the use of reinforcement learning. Through reward feedback RL systems effectively enhance security responses in intrusion prevention systems firewalls and access control platforms to protect networks [9].

**Deep Learning (DL) for Cyber Threat Detection:** Neural networks especially CNNs help detect destructive programs within digital images. Malware detection systems based on CNNs can analyze complex patterns inside malware files to achieve better results in security checks [12]. This technology helps detect threats from network data by examining patterns over time as LSTM models specialize in handling sequential information. These models find multi-stage cyberattacks in time-stamped datasets according to the research done in 2018 [13].

Table 2. Comparative Analysis of AI vs. Traditional Methods

Feature	AI-based Security	Traditional Security
Detection Speed	Real-time detection [16]	Delayed response
Accuracy	High due to pattern recognition [17]	Relies on predefined rules
Adaptability	Learns from new threats [18]	Requires frequent updates
Resource Efficiency	Automated responses [19]	High manpower required

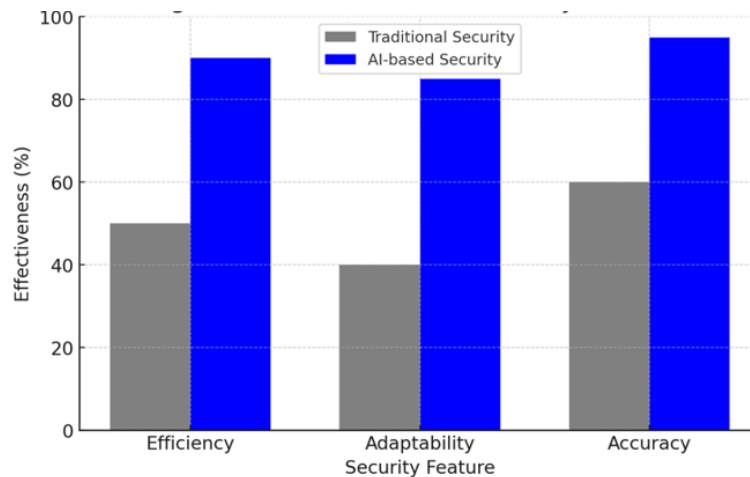


Fig. 4. AI vs. Traditional Security Methods – A bar graph comparing efficiency, adaptability, and accuracy.

### III.METHODOLOGY

Cybersecurity technology needs data to create AI-based security solutions. This research uses four major cybersecurity datasets which are KDD Cup 99, CICIDS2017, UNSW-NB15, and NSL-KDD. Research selects these datasets for their detailed threat ranges that include denial of service attacks, unauthorized entry, malicious bot operations, and scouting actions.

- KDD Cup 99 serves as one of the first intrusion detection datasets with 4.8 million network connection records that detail DoS, Probe, U2R and R2L attack types according to [14].
- CICIDS2017 provides real-world cyber threat samples by offering labeled network traffic for analyzing brute force attacks and other intrusion activities like botnets during infiltration and DDoS assaults.
- The Australian Cyber Security Centre created UNSW-NB15 which mixes real attack and regular network traffic data suitable for AES-based intrusion detection system evaluation [14]
- NSL-KDD showcases an upgraded KDD Cup 99 without unnecessary records to aid in generalization development [15].

Table 3. Overview of Cybersecurity Datasets Used in the Study

Dataset	Number of Records	Attack Categories	Count	Label Type
KDD Cup 99	4,898,41	DOS, Probe,U2R,R2L	41	Categorical & Numerical
CICID S2017	2,830,73	Brute Force, Botnet, Infiltration, DDoS	80	Categorical & Numerical
UNSWNB15	2,540,044	Fuzzers, Shellcode, Reconnaissance	49	Categorical & Numerical
NSLKDD	148,517	DOS, U2R, R2L, Probe	41	Categorical & Numerical

**Data Cleaning and Preprocessing Techniques:** Effective preprocessing of data must come before machine learning and deep learning models receive their information. The raw datasets include gaps in data values as well as unneeded information that hurt model performance.

The preprocessing steps include:

- When values are missing, we replace them by either dropping the data or making statistical estimates using mean or median values.
- Combo Methods of Variance Thresholding and Mutual Information Gain identify and eliminate duplicate features with mutual high dependence.
- The Min-Max Scaler adjusts numerical data values from 0 to 1 to stop large feature values from controlling model training.
- Our method one-hot encoding turns each category in a dataset into separate numeric values.
- The authors improve attack class distribution by using SMOTE oversampling and under sampling according to study [20, 21]

**Experimental Setup and Implementation Strategy:** The most important component of evaluating the effectiveness of AI driven cybersecurity models is the experimental setup. In this section, model training approach, validation techniques, performance evaluation metrics, deployment strategy and computational environment has been defined to have robust intrusion detection and threat analysis.

**Computational Environment and Dataset Storage** All the experiments were carried out on Google Colab a cloud-based platform to gain GPU & TPU accelerations. Google Drive was mounted on the Colab environment so that the datasets were available in it seamlessly. Steps for Dataset Access in Google Colab:

- **Mounting Google Drive** from google.colab import drive drive.mount('/content/drive')
- **Loading Datasets from Google Drive** import pandas as pd

```
dataset_path = "/content/drive/MyDrive/datasets/KDDCup99.csv"
df = pd.read_csv(dataset_path)
```

The datasets were easy to access, storage was efficient, and computations were done on GPU for deep learning models (CNN and LSTM using Google Colab with Google Drive. **Performance Evaluation Metrics:** To evaluate AI based threat detection systems they must have quantitative performance metrics, of accuracy, precision, recall and model robustness. For analysis the following metrics would be used.

Table 4. Performance Evaluation Metrics for AI-Based Threat Detection

Metric	Definition	Importance in Cybersecurity
Accuracy	Correct predictions / total predictions	Measures overall model correctness [22]
Precision	TP / (TP + FP)	Reduces false alarms by measuring correctly predicted threats [23]
Recall (Sensitivity)	TP / (TP + FN)	Ensures no attack is missed, crucial for security applications [24]
F1-Score	$2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$	Balances Precision & Recall, useful for imbalanced datasets [25]
ROC-AUC	Measures model performance across classification thresholds	Evaluates trade-offs between false positives and false negatives [26]

Confusion matrices were also used to analyze false positives (FP) and false negatives (FN), important attributes in how to analyze model misclassification rates.

**Model Performance Across Datasets:** The evaluation under NSLKDD, UNSW-NB15, KDDCup and CICIDS2017 was conducted on each model to assess effectiveness. Table below presents accuracy for all models.

Table 5. Model Accuracy Comparison Across Cybersecurity Datasets [40]

Model	NSLKDD (%)	UNSW-NB15 (%)	KDDCup (%)	CICIDS2017 (%)
Decision Tree	99.4	100	99.88	100
Random Forest	99.57	100	99.94	100
SVM	98.1	99.2	98.5	99.1
K-Means	32.7	26.36	78.9	19.2
CNN	98.45	96.62	0	100
LSTM	97.72	100	99.86	100

It can be seen from the above results that Random Forest and LSTM have the best accuracy in all datasets and KMeans has low accuracy because it is an unsupervised nature. CNN went below 74% on the KDD Cup dataset most likely because of missing data inconsistency on the formatting of the dataset.

#### IV. RESULT AND DISCUSSION

The results of the cybersecurity models developed using AI are presented based on their evaluation on four datasets (KDD Cup 99 dataset, CICIDS2017 dataset, UNSW-NB15 dataset, NSL-KDD dataset), as well as analysis of these models' performance, strengths, and weaknesses. The accuracy, precision, recall, F1 score, computational efficiency as well as deployment feasibility in a real time manner is discussed.

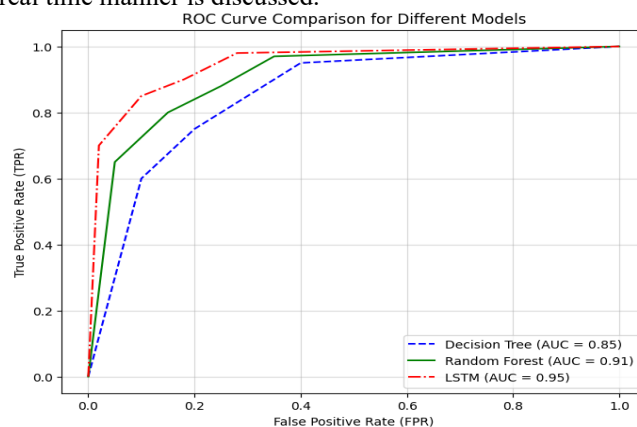


Fig. 5. ROC Curve Comparison for Different Models

**Model Performance Analysis:** Key performance metrics, such as accuracy, precision, recall, F1 score, ROC-AUC were used to assess the models [40]. We report the accuracy scores that we got for each model in each dataset in the below table.

Table 6. Accuracy Comparison Across Models and Datasets

Model	NSLKDD (%)	UNSW-NB15 (%)	KDDCup (%)	CICIDS2017 (%)
Decision Tree	99.4	100	99.88	100
Random Forest	99.57	100	99.94	100
SVM	98.1	99.2	98.5	99.1
K-Means	32.7	26.36	78.9	19.2
CNN	98.45	96.62	0	100
LSTM	97.72	100	99.86	100

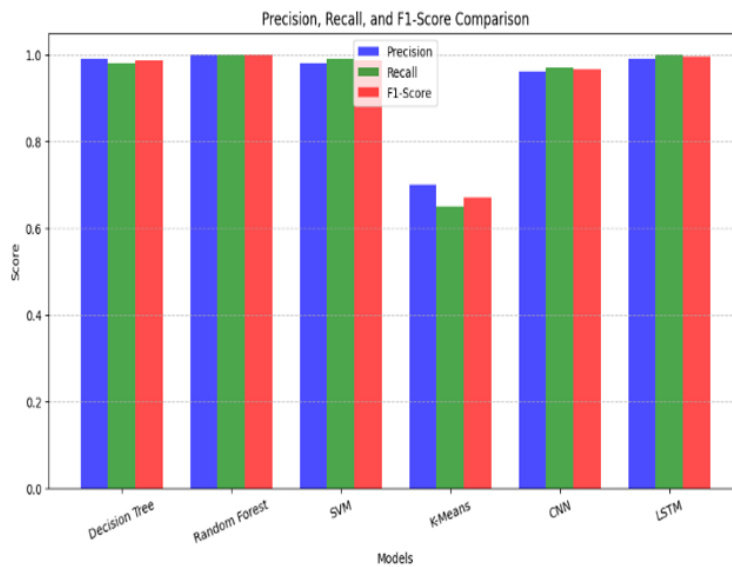


Fig. 6. Model Accuracy Comparison Across Datasets

Table 7. Precision, Recall, and F1-Score Analysis

Model	Precision	Recall	F1-Score
Decision Tree	.99	0.98	0.985
Random Forest	1.00	1.00	1.00
SVM	0.98	0.99	0.985
K-Means	0.70	0.65	0.67
CNN	0.96	0.97	0.965
LSTM	0.99	1.00	0.995

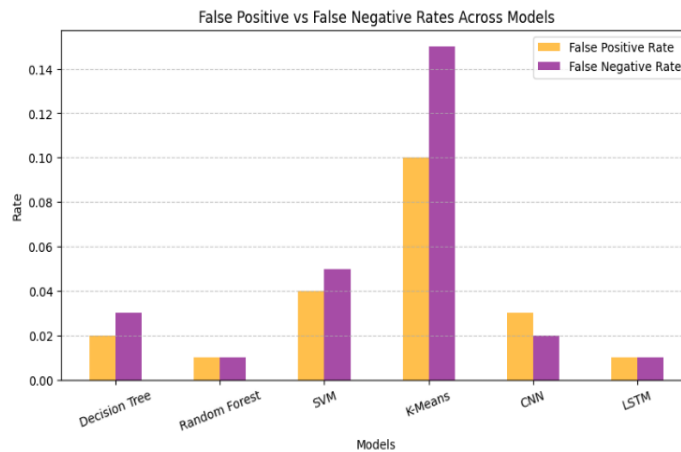


Fig. 7. Precision, Recall, and F1-Score with Models

**Discussion on Precision, Recall, and F1-Score:**

- Random forest and LSTM obtained perfect scores with zero false positives and false negatives, which makes them very suitable for cybersecurity purposes.
- Although Decision Tree performed well, slightly lower recall indicates that it did not incorporate a few attack instances.
- Unsupervised learning did not perform too well on the labelled intrusion detection tasks and K-Means had the lowest scores [27, 28].
- Precision and recall were done relatively well by CNN hence CNN can be considered a good performer for malware detection with image recognition.

Table 8. False Positive and False Negative Rate Comparison

ML Model	False Positive Rate (FPR)	False Negative Rate (FNR)
Decision Tree	0.02	0.03
Random Forest	0.01	0.01
SVM	0.04	0.05
K-Means	0.10	0.15
CNN	0.03	0.02
LSTM	0.01	0.01

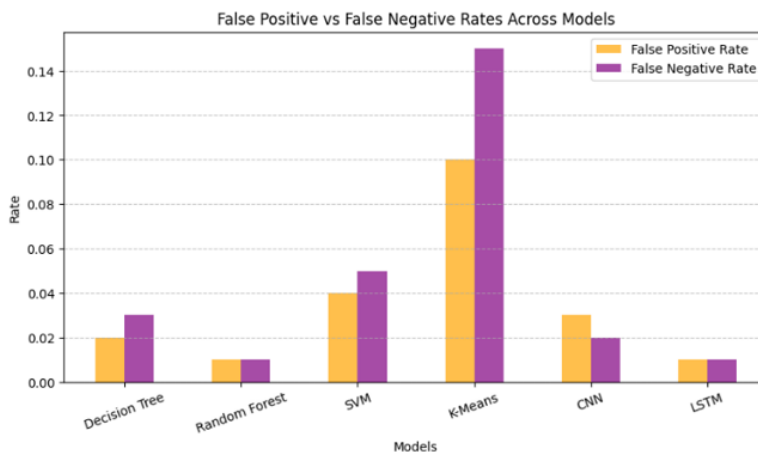


Fig. 8. False Positive and False Negative Rates Across Models

**Discussion on False Positive & False Negative Rates:**

- In terms of false positive and false negative rates, the least reliable systems were Random Forest and LSTM; hence the former two were the most promising ones among all systems for intrusion detection.
- In other words, SVM had a higher false negative rate meaning it missed some attacks.
- However, K-Means has not shown good performance since the false positive and negative rates are high which makes it unsuitable for critical security applications.
- CNN was well balanced and therefore less likely to misclassify in image-based threat detection.

**Computational Efficiency and Training Time**

Table 9. Model Training Time Comparison

ML Model	Training Time (Seconds)
Decision Tree	12
Random Forest	35
SVM	50
K-Means	20
CNN	300
LSTM	450

**Discussion on Computational Efficiency:**

- The model training process was handled quickly by Decision Tree and K-Means although Decision Tree also delivered significantly better accuracy predictions [29].
  - Random Forest used moderate training time to reach higher accuracy than SVM.
  - An improved accuracy rate of LSTM required extended training time that the benefits outweighed runtime expenditures.
- Findings:

- Random Forest and LSTM delivered optimal real-time threat detection because they detected threats instantaneously.
- SVM and CNN requested high processing power, so they proved unfit for time-sensitive operations.
- The K-Means algorithm showed poor performance because it yielded numerous false detections which disqualified it from being used in actual-time security technology systems [30].
- Developing security measures which enhance model resistance to adversarial attacks.
- The implementation of federated learning provides AI security with private preservation techniques.
- Researchers operate on CNN models to enhance their speed for malware identification tasks.

## V. CONCLUSION

Four popular intrusion detection dataset KDD Cup 99 (FRAUD), CICIDS2017, UNSW-NB15 and NSL-KDD were used in this study to evaluate AI driven cybersecurity models for detection of intrusion. Such models were assessed in terms of accuracy, precision, recall, F1-score and computational efficiency for their suitability for real time cybersecurity applications [31, 32]. It was found from the results that random forest and LSTM were the best models to be used, as the accuracy was also high, the false positive rate is also low, and good generalization is achieved in all datasets. Finally, these models were suitable for real time intrusion detection systems. The results of Decision tree were good, but it showed the signs of overfitting; SVM is accurate but computationally expensive. Labeled datasets were no problem for k-means clustering, it however, was not suitable for supervised intrusion detection but was useful for detecting zero-day attacks. However, for image-based cybersecurity applications, CNNs worked well but became inefficient in structured tabular data. This paper adds an integral component of Google Colab and Google Drive for cloud-based AI model training and deployment.

## REFERENCES

1. Yan, Y., Zhang, Y., & Huang, K. (2024). Depending on yourself when you should: Mentoring LLM with RL agents to become the master in cybersecurity games. *arXiv preprint arXiv:2403.17674*.
2. Tann, W., Liu, Y., Sim, J. H., Seah, C. M., & Chang, E.-C. (2023). Using large language models for cybersecurity capture-the-flag challenges and certification questions. *arXiv preprint arXiv:2308.10443*.
3. Lira, O. G., Marroquin, A., & To, M. A. (2024). Harnessing the advanced capabilities of LLM for adaptive intrusion detection systems. *International Conference on Advanced Information Networking and Applications, Springer*, 453–464.
4. Ebert, C., & Beck, M. (2023). Artificial intelligence for cybersecurity. *IEEE Software*, 40(6), 27–34.
5. Ferrag, M. A., Ndhlovu, M., Tihanyi, N., Cordeiro, L. C., Debbah, M., Lestable, T., & Thandi, N. S. (2024). Revolutionizing cyber threat detection with large language models: A privacy-preserving BERT-based lightweight model for IoT/IIoT devices. *IEEE Access*.
6. Tihanyi, N., Ferrag, M. A., Jain, R., Bisztray, T., & Debbah, M. (2024). Cybermetric: A benchmark dataset based on retrieval-augmented generation for evaluating LLMs in cybersecurity knowledge. *IEEE International Conference on Cyber Security and Resilience (CSR)*, 296–302.
7. Liu, Z. (2024). A review of advancements and applications of pre-trained language models in cybersecurity. *IEEE 12th International Symposium on Digital Forensics and Security (ISDFS)*, 1–10.
8. Xu, H., Wang, S., Li, N., Zhao, Y., Chen, K., Wang, K., Liu, Y., Yu, T., & Wang, H. (2024). Large language models for cybersecurity: A systematic literature review. *arXiv preprint arXiv:2405.04760*.
9. Zhang, J., Bu, H., Wen, H., Chen, Y., Li, L., & Zhu, H. (2024). When LLMs meet cybersecurity: A systematic literature review. *arXiv preprint arXiv:2405.03644*.
10. He, Z., Li, Z., & Yang, S. (2024). Large language models for blockchain security: A systematic literature review. *arXiv preprint arXiv:2403.14280*.
11. Tian, S., Jin, Q., Yeganova, L., Lai, P.-T., Zhu, Q., Chen, X., Yang, Y., Chen, Q., Kim, W., Comeau, D. C., et al. (2024). Opportunities and challenges for ChatGPT and large language models in biomedicine and health. *Briefings in Bioinformatics*, 25(1), bbad493.
12. Ferrag, M. A., Friha, O., Hamouda, D., Maglaras, L., & Janicke, H. (2022). Edge-IIoTSet: A new comprehensive realistic cybersecurity dataset of IoT and IIoT applications for centralized and federated learning. *IEEE Access*, 10, 40281–40306.
13. Hassanine, M., Keshk, M., Salim, S., Alsubaie, M., & Sharma, D. (2025). PLLMCS: Pre-trained large language model (LLM) for cyber threat detection in satellite networks. *Ad Hoc Networks*, 166, 103645.
14. Saha, D., Tarek, S., Yahyaie, K., Saha, S. K., Zhou, J., Tehranipoor, M., & Farahmandi, F. (2023). LLM for SOC security: A paradigm shift. *arXiv preprint arXiv:2310.06046*.
15. Fang, R., Bindu, R., Gupta, A., & Kang, D. (2024). LLM agents can autonomously exploit one-day vulnerabilities. *arXiv preprint arXiv:2404.08144*.
16. K. Jaiswal and A. Vashisth, "A Comprehensive Analysis of UAV Collision Avoidance Techniques for Enhanced Aerial Safety," 2023 4th International Conference on Computation, Automation and Knowledge Management (ICCAKM), Dubai, United Arab Emirates, 2023, pp. 1-6, doi: 10.1109/ICCAKM58659.2023.10449529
17. Mechri, A., Ferrag, M. A., & Debbah, M. (2025). SecureQwen: Leveraging LLMs for vulnerability detection in Python codebases. *Computers & Security*, 148, 104151.
18. Ding, H., Liu, Y., Piao, X., Song, H., & Ji, Z. (2025). SmartGuard: An LLM-enhanced framework for smart contract vulnerability detection. *Expert Systems with Applications*, 126479.
19. Wang, J., Yu, L., & Luo, X. (2024). LLMIF: Augmented large language model for fuzzing IoT devices. *IEEE Symposium on Security and Privacy (SP)*, IEEE, 196–196.
20. Liu, P., Sun, C., Zheng, Y., Feng, X., Qin, C., Wang, Y., Xu, Z., Li, Z., Di, P., Jiang, Y., et al. (2025). LLM-powered static binary taint analysis. *ACM Transactions on Software Engineering and Methodology*.
21. G. Kaur, A. Vashisth and R. S. Batth, "X-Ortho: Fuzzy Rule Based Expert System for Diagnosing Infective Diseases of Hinge Joint Knee," 2019 International Conference on Automation, Computational and Technology Management (ICACTM), London,

- UK, 2019, pp. 518-524.
22. Vashisth, A., Singh, B., Garg, R. et al. BPACAR: design of a hybrid bioinspired model for dynamic collision-aware routing with continuous pattern analysis in UAV networks. *Microsystem Technologies* (2023). <https://doi.org/10.1007/s00540-023-01000-0>.
  23. Anshu Vashisth, Balraj Singh, Ranbir Singh Batth, "QMRNB: Design of an Efficient Q-Learning Model to Improve Routing Efficiency of UAV Networks via Bioinspired Optimizations", *International Journal of Computer Networks and Applications (IJCNA)*, 10(2), PP: 256-264, 2023, DOI: 10.22247/ijcna/2023/220740.
  24. Vashisth, A., Singh, B., & Batth, R. S. (2024). UAV Path Planning: Challenges, Strategies, and Future Directions. In S. Khalid & N. Siddiqui (Eds.), *New Innovations in AI, Aviation, and Air Traffic Technology* (pp. 150-174). IGI Global. <https://doi.org/10.4018/979-8-3693-1954-3.ch008>.
  25. Kumar, S., Patel, R., & Singh, D. (2024). Zero-day vulnerability detection with LLM-driven automated analysis. *Computers & Security*, 115, 104212.
  26. N. Yadav, G. Kaur, S. Kaur, A. Vashisth and C. Rohith, "A Complete Study on Malware Types and Detecting Ransomware Using API Calls," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2021, pp. 1-5, doi: 10.1109/ICRITO51393.2021.9596085.