Document: Model Research for Cyber Threat Dashboard

Title: Artificial Intelligence and Machine Learning Models-Comparative Study for Cyber Threat Visualization Dashboard

Author: Shireen Naaz
Date: 11-02-2026
Project: Development of Interactive Cyber Threat Visualization Dashboard

1. Introduction

In today's digital world, the importance of cybersecurity cannot be emphasized enough. The goal of the project titled “Development of an Interactive Cyber Threat Visualization Dashboard” is to enable security analysts and stakeholders to comprehend the threat landscape in real-time through data-driven evidence. In order to leverage the concepts of Advanced Threat Detection and Visualization, the choice of appropriate “AI & Machine Learning Models” needs to be evaluated.

This document provides a comparative analysis of the most popular AI/ML models, focusing on their advantages, disadvantages, and potential application, alongside their appropriateness for the project at hand. It aims to determine which model can efficiently detect, classify, and predict cyber threats while providing interactive visualization support.

2. Overview of AI/ML Models

Model Name	Type	Strengths	Weaknesses	Use Cases	Suitability for Cyber Threat Dashboard
Linear Regression	Supervised (Regression)	Simple, interpretable, fast	Assumes linear relationships only	Trend prediction, numeric	Low – Threat detection often requires non-

				forecasting	linear modeling
Logistic Regression	Supervised (Classification)	Easy to implement, interpretable	Limited to simple patterns, binary outcomes	Binary classification (malicious vs safe)	Medium – Can be used for initial threat classification
Decision Tree	Supervised (Classification/Regression)	Easy to visualize, handles non-linear relationships	Prone to overfitting	Classification of alerts, risk scoring	High – Useful for interpretable threat detection rules
Random Forest	Supervised (Classification/Regression)	Reduces overfitting, accurate, handles large datasets	Computationally intensive	Intrusion detection, anomaly detection	Very High – Ideal for real-time detection of complex cyber threats
Support Vector Machine (SVM)	Supervised (Classification)	Works well with small datasets, finds clear boundaries	Slow on large datasets, sensitive to parameters	Malware classification, anomaly detection	Medium – Effective for specific pattern recognition in threats
K-Nearest Neighbors (KNN)	Supervised (Classification/Regression)	Simple, no training required	Computationally slow for large datasets	Attack type classification,	Low-Medium – May not scale

				anomaly scoring	well for real-time dashboards
Naive Bayes	Supervised (Classification)	Fast, effective for text, probabilistic approach	Assumes feature independence	Phishing detection, spam alerts	Medium-High – Good for textual threat intelligence
Neural Network (MLP)	Supervised (Classification/Regression)	Can model complex patterns, scalable	Needs large datasets, less interpretable	Predicting attacks, behavior analysis	High – Suitable for predicting complex threat patterns
Convolutional Neural Network (CNN)	Supervised (Mostly Image)	Excellent for spatial and visual data	Not ideal for tabular/temporal data	Malware visualization, network traffic visualization	Medium – Useful if visual patterns in threat logs are analyzed
Transformer	Supervised/Unsupervised (NLP)	Excellent for sequential data, NLP	Computationally heavy	Threat report analysis, log parsing, sequence prediction	Very High – Useful for analyzing large cyber threat datasets and logs

3. Detailed Model Analysis

3.1 Linear & Logistic Regression

- Linear Regression predicts numeric trends. It is simple and interpretable but cannot model complex cyber threat patterns, which are often non-linear.
- Logistic Regression is suitable for basic threat classification (e.g., malicious vs safe). It is interpretable but not effective for multi-class threat scenarios.

3.2 Decision Tree & Random Forest

- Decision Trees provide clear rules that analysts can interpret directly. Useful for explaining why an alert was flagged.
- Random Forest aggregates multiple trees to improve accuracy and reduce overfitting. It is ideal for detecting complex and subtle threats in large datasets.

3.3 Support Vector Machine (SVM)

- SVM can separate normal from anomalous traffic with high precision. It works best with small-to-medium datasets but may struggle with very large, real-time threat data streams.

3.4 K-Nearest Neighbors (KNN)

- KNN classifies threats based on similarity to known attacks. Simple but computationally heavy for dashboards requiring real-time response.

3.5 Naive Bayes

- Particularly effective for text-based threat intelligence, such as phishing email classification. Fast and lightweight, making it suitable for some modules of the dashboard.

3.6 Neural Networks (MLP)

- Multi-layer perceptrons can capture complex attack patterns across multiple features. Requires more computational power but offers high detection accuracy.

3.7 Convolutional Neural Networks (CNN)

- If visual patterns exist in network traffic or malware signatures, CNNs can process these efficiently. Useful for threat visualization modules.

3.8 Transformers

- Transformers excel in sequential and text data. Can analyze logs, threat reports, and sequences of attack events. Very powerful for dashboards that include NLP-based analysis.
-

4. Model Suitability for the Cyber Threat Dashboard

After considering the details, it came out that the recommended model for this project is Random Forest, with support from Neural Networks and Transformers in specialized modules.

Justification:

Random Forest: Accurate, scalable, and interpretable anomaly detection. It supports multi-feature network data and cyber events.

Neural Networks: Useful for predicting complex attack patterns over time.

Transformers: They can analyze textual logs and also threat intelligence to ensure actionable insights.

Coupled together, these models grant the capability to the dashboard for cyber threats detection, classification, and visualization in an interactive manner, hence very responsive and reliable.

5. Conclusion

The selection of a proper AI/ML model plays a fundamental role in developing an efficient Cyber Threat Visualization Dashboard. The current study assessed 10 different models for the selection of the best one based on utmost efficiency, scalability, interpretability, and applicability in real-time threat detection. The study recommends Random Forest, Neural Networks, and Transformers, which could be implemented to make the response of security analysts efficient and gain insights from complex datasets.

6. References

- Bishop, C. M. *Pattern Recognition and Machine Learning.* Springer, 2006.
- Goodfellow, I., Bengio, Y., & Courville, A. *Deep Learning.* MIT Press, 2016.
- Kaggle tutorials on Cybersecurity ML applications
- Research papers on Random Forests in intrusion detection systems