# AI-Based Cyber Security Threat Prediction Model Research

## I. Introduction

Cyber-attacks are evolving in complexity, utilizing sophisticated strategies such as multi-stage infiltration, lateral movement, stealthy data exfiltration, ransomware encryption, and coordinated distributed assaults. Traditional security mechanisms such as firewalls, rule-based Intrusion Detection Systems (IDS), and signature-based antivirus solutions are inadequate in detecting emerging and zero-day attacks. These systems lack the capability to understand behavioral patterns, contextual relationships, and dynamic threat evolution.

To address this challenge, intelligent and autonomous AI-based systems are needed that can analyze diverse cybersecurity data, uncover hidden behavioral and relational patterns, detect anomalies, and predict cyber threats in real time.

This research focuses on the development of an **AI-Based Cyber Security Threat Prediction AI Agent**, capable of performing **autonomous monitoring, real-time anomaly detection, threat prediction, and adaptive threat response** without continuous human supervision.

# II. Comparative Analysis of State-of-the-Art AI Models

Different types of cyber threats exhibit unique characteristics, requiring specialized AI models for different analytical perspectives. These perspectives are:

| Perspective | Description |
|---|---|
| **Statistical Classification** | Detects malicious patterns in structured tabular data such as flow records, protocol data, log statistics, and feature-engineered input. |
| **Sequential Behavior Analysis** | Identifies threats based on time-dependent event sequences such as login attempts, session transitions, API logs, and system calls. |
| **Relational Mapping** | Understands relationships between entities (users, devices, IPs, sessions) to detect coordinated, multi-entity cyber campaigns. |
| **Unsupervised Anomaly Detection** | Discovers unknown or zero-day attacks by learning the baseline of normal behavior and detecting anomalies or deviations. |

## Table 1: Comparison of AI Models for Cyber Threat Prediction

| Model Category | Example Models | Core Function | Why Choose (Best Fit) | Typical Accuracy / F1-Score |
|---|---|---|---|---|
| **Ensemble / Gradient Boosting** | XGBoost, LightGBM | Utilizes optimized gradient boosting to analyze structured cyber data with high explainability (SHAP) | Ideal for **Stage 1 filtering** due to high speed, accuracy, interpretability, and scalability | 98–99.5% Accuracy, up to 99.2% F1-Score |
| **Sequential Deep Learning** | Transformers, LSTM, RNN | Models time-based event sequences and captures long-range dependencies using self-attention | Best for detecting **multi-step attacks**, insider activity, and behavioral anomalies | 98–99%; CNN-LSTM hybrids up to 99.86% |
| **Relational Deep Learning** | Graph Neural Network (GNN) | Models relationships between entities using graph structures (nodes, edges) | Detects **coordinated and campaign-level cyber-attacks** | 96–97.9% Accuracy/F1 |
| **Anomaly Detection (Unsupervised)** | Autoencoder, Isolation Forest | Learns baseline of normal behavior and detects deviations or anomalies | Excellent for detecting **zero-day, rare, and previously unseen threats** | Evaluated by reconstruction error; High sensitivity |
| **Traditional ML (Baseline)** | Random Forest | Ensemble-based decision trees; robust with minimal tuning | Useful for **baseline comparison and quick deployment** | 97–99% Accuracy |

# III. Hybrid Fusion Architecture for Threat Prediction

To achieve more accurate and resilient cyber threat detection, a **Multi-Stage Hybrid Fusion System** is proposed. This system integrates the strengths of all model categories into a single layered architecture, resulting in a **Unified Threat Risk Score**.

## Table 2: Multi-Stage Hybrid AI Architecture

| Stage | Model Type | Key Role in Threat Prediction |
|---|---|---|
| **Stage 1** | XGBoost / LightGBM | Rapid statistical filtering of structured log and telemetry data |
| **Stage 2** | LSTM / Transformer | Sequence-based behavioral threat analysis |
| **Stage 3** | Graph Neural Network (GNN) | Detection of relationship-based, coordinated cyber campaigns |
| **Stage 4** | Autoencoder / Isolation Forest | Unsupervised detection of zero-day and unknown anomalies |
| **Stage 5** | Meta-Learner (Logistic Regression, SVM, Gradient Boosting) | Combines model outputs to generate a **Unified Threat Risk Score** |

## Table 3: Inputs to Meta-Learner

| Feature Source | Score Type |
|---|---|
| XGBoost | Threat Probability |
| LSTM/Transformer | Sequence Risk Score |
| GNN | Entity Correlation Risk |
| Autoencoder | Anomaly Reconstruction Error |
| Final Fusion Output | Unified Threat Risk Score |

# IV. Benefits of the Hybrid Fusion System

| Benefit Area | Description |
|---|---|
| **Detection Performance** | Increased accuracy, fewer false alarms, better detection of rare threats |
| **Zero-Day Threat Detection** | Capable of identifying novel, unseen attacks that bypass signature-based systems |
| **Explainability** | XGBoost with SHAP and GNN visualizations enhance analyst interpretability |
| **Scalability** | Easily integrates with real-world systems like SIEM, SOC, EDR, and SOAR |
| **Resilience to Evasion** | Attackers cannot bypass all system layers due to multi-perspective detection |

# V. Real-World Use Cases

- Insider threat monitoring and unauthorized behavior detection
- Botnet and ransomware command-and-control (C2) analysis
- API and cloud infrastructure security monitoring
- Financial and enterprise fraud detection (multi-step)
- Government cyber defense and national security analytics
- SOC automation and AI-powered SIEM enhancement