Model Research for AI-Based Cybersecurity Threat Prediction

1. Objective

The main goal of this research is to design and evaluate AI-based models capable of predicting, detecting, and classifying cybersecurity threats in real-time. The platform aims to:

1. Detect known and unknown (zero-day) attacks.
2. Predict potential vulnerabilities and attack patterns.
3. Provide explainable and interpretable results for security analysts.
4. Operate efficiently on high-dimensional and heterogeneous data sources (network traffic, logs, system events).

.

---

2. Data Sources & Features

- **Static Data:**
    - **CICIDS2017 dataset** – widely used benchmark dataset with labeled attacks (DoS, DDoS, Brute Force, Botnets, etc.).
    - Features: flow statistics, packet sizes, durations, protocols, flags, source/destination IPs.
- **Real-Time Data:**
    - Network traffic logs, system/application logs, user behavior.
    - Threat intelligence feeds (IPs, CVEs, malware signatures).

Feature Engineering: Combine statistical, temporal, and session-level features to capture attack patterns effectively.

.

---

3. Model Selection

Machine Learning Models: Random Forest, XGBoost, SVM, KNN – effective for classification and anomaly detection.

Deep Learning Models:

- **LSTM/RNN** – for sequential attack patterns.
- **CNN** – malware detection from binaries or traffic matrices.
- **Autoencoders** – anomaly detection.
- **Graph Neural Networks** – lateral movement in networks.

Hybrid Models: LSTM + Random Forest, Autoencoder + XGBoost, to handle heterogeneous data.

---

4. Workflow

1. **Data collection** → logs, network, and threat feeds.
2. **Preprocessing** → cleaning, normalization, encoding.
3. **Feature extraction** → statistical and temporal features.
4. **Model inference** → predict attacks or anomalies.
5. **Alert generation** → actionable warnings with confidence scores.

6. **Response** → optional automated remediation.

---

5. Evaluation Metrics

- Accuracy, Precision, Recall, F1-Score
- ROC-AUC
- Time-to-Detect (critical for real-time systems)

---

6. Explainable AI

- **SHAP / LIME** – interpret model predictions.
- **Attention mechanisms** – highlight key patterns in sequential or network data.

---

7. Challenges

- Data imbalance, scarcity of zero-day attack examples
- Real-time processing constraints
- Adversarial attacks on AI models
- Balancing interpretability with model complexity

---

8. Emerging Directions

- Zero-day prediction using graph-based methods
- Federated learning for privacy-preserving threat prediction
- Autonomous AI-driven threat hunting
- Multi-modal threat intelligence combining logs, network, IoT

Conclusion

AI-based threat prediction provides **proactive cybersecurity** by learning patterns, detecting anomalies, and predicting attacks. Hybrid models and explainable AI improve accuracy, usability, and trust for real-world deployment.