# Model Research Report

## Project Title: AI-Driven Cybersecurity Threat Prediction Platform

### 1. Objective

The primary objective of this research is to identify and analyze suitable machine learning and deep learning models that can accurately detect, classify, and predict cybersecurity threats. The models selected must handle complex, high-dimensional network data efficiently while maintaining interpretability and scalability for real-time threat prediction.

### 2. Overview

Cybersecurity data often involves large-scale, imbalanced, and heterogeneous network traffic logs. The goal is to select models capable of learning both known attack patterns and anomalous behaviors representing potential new threats.

For this project, both traditional machine learning algorithms and deep learning architectures are explored to determine the optimal set of models for implementation.

### 3. Models Considered

#### A. Classical Machine Learning Models

1. **Logistic Regression**: A linear model for binary classification such as predicting whether a network connection is benign or malicious. It is simple, interpretable, and fast to train but struggles with nonlinear relationships.
2. **Random Forest**: An ensemble model using multiple decision trees. It is robust to overfitting and handles noisy and imbalanced data well but can be slower for large datasets and less interpretable.
3. **XGBoost**: A gradient boosting algorithm that builds decision trees sequentially for higher accuracy. It provides high predictive performance and efficiency but requires careful parameter tuning.
4. **Support Vector Machine (SVM)**: Finds an optimal hyperplane to separate classes. It performs well in high-dimensional spaces but is computationally expensive on large datasets.
5. **Naïve Bayes**: A probabilistic model based on Bayes’ theorem. It is simple and effective for text-based threat detection (e.g., phishing, spam) but assumes feature independence.

#### B. Deep Learning Models

1. **LSTM (Long Short-Term Memory)**: Captures temporal dependencies in sequential data such as network logs. Excellent for sequential threat pattern detection but requires significant data and training time.
2. **CNN (Convolutional Neural Network)**: Learns local feature patterns and detects spatial correlations in traffic behavior. It needs proper preprocessing and reshaping of input data.
3. **Autoencoder**: An unsupervised neural network for learning data representations. Useful for anomaly detection but may require tuning to reduce false positives.
4. **Deep Neural Network (DNN)**: A fully connected neural network for general threat classification. It has high learning capacity for complex patterns but can overfit without regularization.

#### C. Hybrid / Ensemble Models

* **Stacked Ensemble**: Combines Random Forest and XGBoost for higher accuracy.
* **Autoencoder + Classifier**: Autoencoder detects anomalies, then a classifier labels them.
* **CNN + LSTM Model**: Captures both spatial and temporal features in network data.

These hybrid approaches combine the strengths of multiple algorithms for improved detection precision and generalization.

### 4. Model Selection Strategy

The project adopts a progressive evaluation approach: 1. Baseline evaluation using Logistic Regression and Random Forest. 2. Performance enhancement using XGBoost. 3. Advanced deep learning phase using LSTM and Autoencoder. 4. Final ensemble model integrating classical and deep learning models.

### 5. Dataset Compatibility

The following datasets are suitable for the models above: - **NSL-KDD Dataset** – Standard dataset for network intrusion detection. - **CICIDS2017 / CSE-CIC-IDS2018** – Modern and realistic intrusion datasets. - **UNSW-NB15** – Covers various attack categories (DoS, Exploits, Reconnaissance, etc.). - **PhishTank / UCI Phishing Dataset** – Ideal for phishing URL detection.

All datasets will undergo preprocessing (cleaning, encoding, normalization) before training.

### 6. Evaluation Metrics

The evaluation of the models will be based on: - Accuracy - Precision - Recall (Sensitivity) - F1 Score - ROC-AUC Score

### 7. Conclusion

The models selected for implementation in the AI-Driven Cybersecurity Threat Prediction Platform are: 1. Random Forest – Reliable and interpretable baseline. 2. XGBoost – High accuracy and scalability. 3. LSTM / Autoencoder – Deep learning for sequential and anomaly detection. 4. Hybrid Ensemble – Final integrated model for superior prediction.

These models enable the system to learn from past attack data, detect ongoing threats, and predict future malicious activities with high accuracy.