Model Research

Model	What it does	When to use	Dataset(s) Used	Accuracy	Robustness
Random Forest (RF)	Uses decision trees that vote on outcomes for classification or detection.	Good baseline model for tabular and structured cybersecurity data.	CICIDS2017, NSL-KDD	85–95%	High
XGBoost / LightGBM	Efficient gradient boosting models for fast, high-performance learning.	When higher accuracy and speed are required than Random Forest.	UNSW-NB15	88–96%	High
LSTM / BiLSTM	Learns sequential or temporal patterns within data.	For time-series or session-based traffic logs where order matters.	CICIDS2017, TON_IoT	85–90%	Moderate
1D-CNN	Detects spatial or temporal relationships in sequential data.	For packet or flow-level intrusion detection.	MAWILab, CICIDS2017	85–95%	Moderate
Autoencoder (AE)	Learns normal data distribution and flags deviations as anomalies.	When labeled attack data is limited or unavailable (unsupervised).	TON_IoT, UNSW-NB15	80–92%	Moderate
Isolation Forest	Identifies anomalies by isolating outliers in feature space.	For unsupervised, lightweight anomaly detection.	NSL-KDD, MAWILab	78–88%	Moderate
SVM	Creates class boundaries using support vectors for classification.	For small, balanced, and feature-rich datasets.	NSL-KDD	80–90%	Low–Moderate
NBNS	Neural-based scoring system that learns custom detection boundaries.	Experimental or advanced systems requiring custom neural configurations.	CICIDS2017	95–99%	High