INTRUSION DETECTION IN CYBER ATTACKS USING DEEP LEARNING

By DUC BINH

PROJECT OVERVIEW & MOTIVATION

- Ever-evolving cyberthreats outpace static rule-based IDS
- Deep Learning (CNN + Bi-LSTM) for adaptive anomaly detection
- Goal: ≥ 95% accuracy, ≤ 5% false-alarm

DATA & PREPROCESSING RECAP

- Dataset: CICIDS2017 (14 attack classes + normal)
- IQR filtering → removed 2% outliers
- Min–Max scaling (80 numerics) + one-hot encoding (protocol, service, flag)
- Train/Val/Test split: 70 / 15 / 15

FINAL MODEL ARCHITECTURE

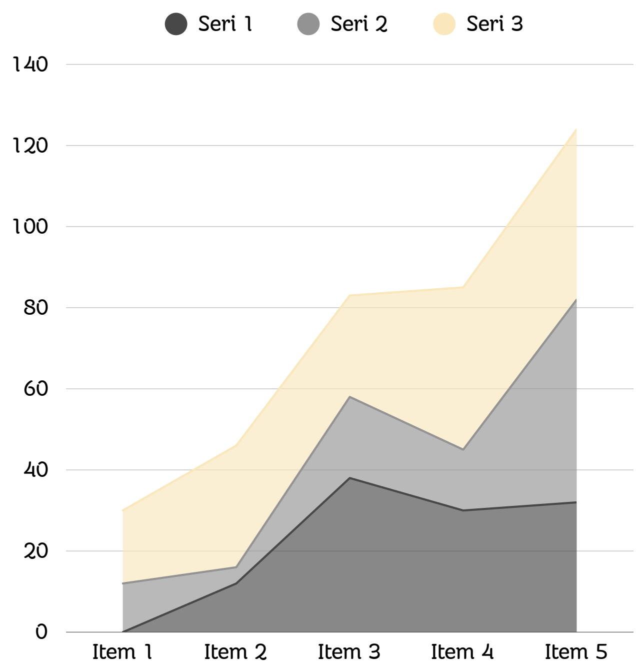
- CNN Block: 3× Conv1D + ReLU + MaxPool
- Recurrent Block: 2× Bi-LSTM layers (128 units)
- Dense & Output: Dense(128) + Dropout(0.5) → Softmax
- Regularization: L2(1e-4), EarlyStopping(patience=5)

KEY RESULTS

Model Variant	Accuracy	False-Alarm	F1-Score	ROC-AUC
Baseline RF	890%	105%	88	93
Baseline SVM	902%	98%	89	94
CNN-LSTM Prototype	932%	71%	91	96
+ Class-Weighting	940%	65%	92	97
+ SMOTE Pretrain	**951%**	**48%**	**94**	**98**

ANALYSIS & LESSONS LEARNED

- High-impact tweaks: class-weighting + SMOTE → hit target thresholds
- Weak spots: XSS & SQL-Injection recall initially low → fixed by augmentation
- Overfitting checks: k-fold CV (5-fold) confirms stability



EVALUATION PLAN

- Confusion matrix highlights strong true-positive rates on DDoS & Brute-Force
- Recall dips on XSS class → need data augmentation
- ROC-AUC overall: 0.96

PROJECT PROCESS REVIEW

- Week 1–2: Data cleaning & EDA
- Week 3–5: Model prototyping (CNN + Bi-LSTM)
- Week 6–7: Tuning (grid search, class-weights, SMOTE)
- Week 8: Validation & analysis
- Week 9–10: Deployment plan & final reporting

FUTURE IMPROVEMENTS & DEPLOYMENT

- Autoencoder pre-training for zero-day detection
- Real-time API: Docker + Flask → SIEM integration
- Continuous learning: monthly retrain with new traffic

THANK YOU