AI Based Network Intrusion Detection System (NIDS)

by Seran Gemechu

Introduction

CYBERSECURITY THREATS ARE INCREASING, REQUIRING EFFECTIVE IDS.

TRADITIONAL IDS SOLUTIONS OFTEN FAIL AGAINST MODERN, EVOLVING ATTACKS.

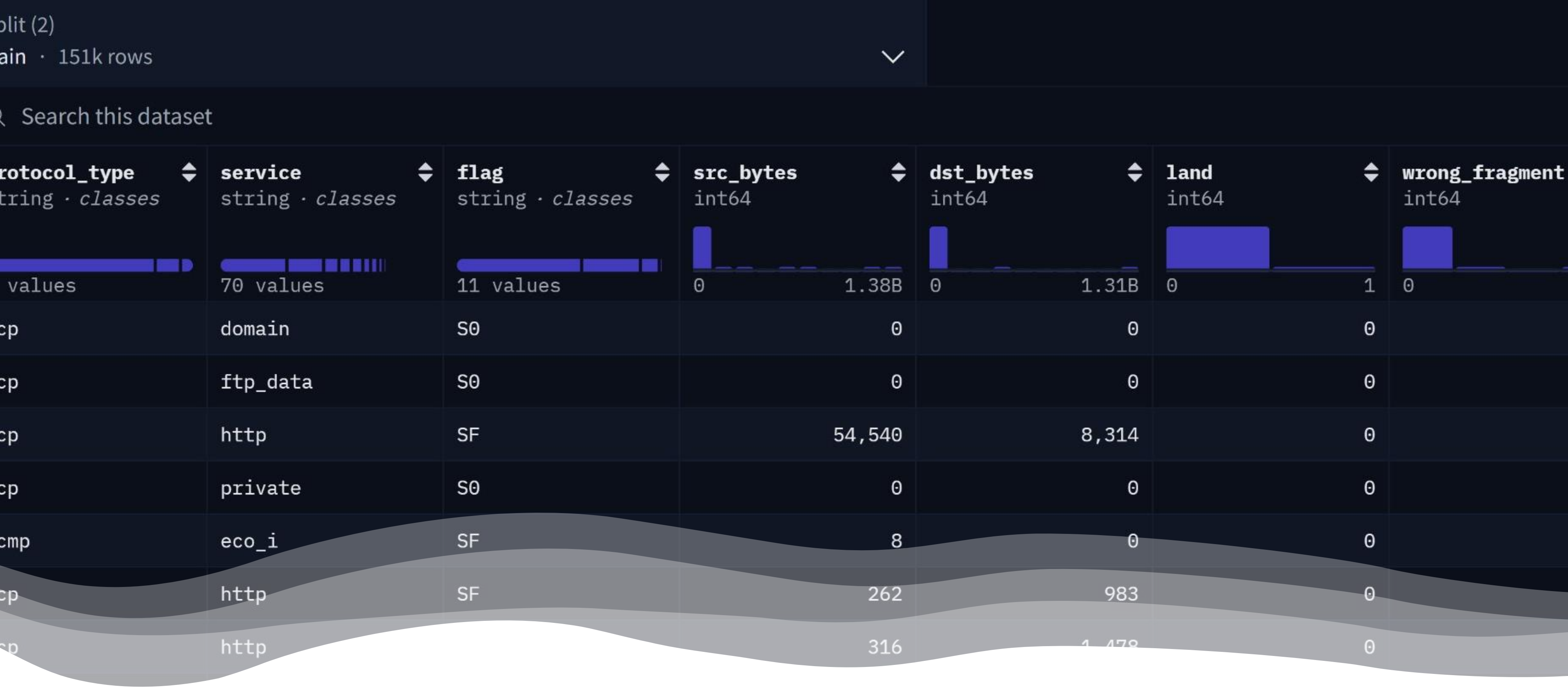
THIS PROJECT DEVELOPS AN IDS USING ANN TO CLASSIFY ANOMALOUS NETWORK TRAFFIC.

Problem Statement

- Cyberattacks such as DoS, R2L, U2R, and Probe are difficult to detect with traditional IDS.
- High false positives make manual investigation difficult. [redacted] **FNs are critical in cybersecurity because it means the IDS missed attacks**, allowing potential intrusions.
- The goal is to develop an AI-powered IDS that improves accuracy and reduces FP & FN alarms.

Project Objectives

- Train an ANN/MLP model to classify network traffic as normal or anomalous.
- Use the NSL-KDD dataset for training and evaluation.
- Measure and optimize precision, recall, CM, and accuracy.
- [**Optional**] Deploy a real-time IDS for monitoring network traffic.



lit (2)
ain · 151k rows

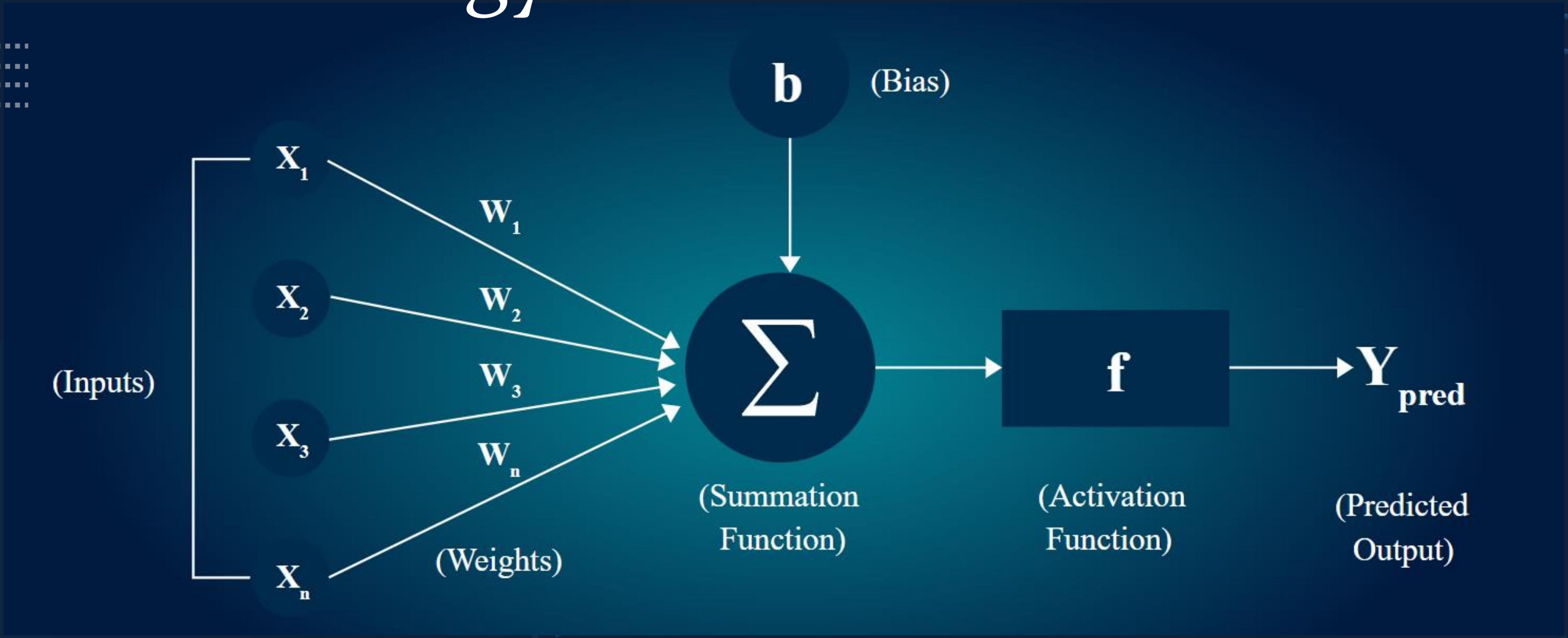
Search this dataset

rotocol_type tring · classes values	service string · classes 70 values	flag string · classes 11 values	src_bytes int64 0 1.38B	dst_bytes int64 0 1.31B	land int64 0 1	wrong_fragment int64 0
cp	domain	S0	0	0	0	
cp	ftp_data	S0	0	0	0	
cp	http	SF	54,540	8,314	0	
cp	private	S0	0	0	0	
cmp	eco_i	SF	8	0	0	
cp	http	SF	262	983	0	
cp	http		316	1,478	0	

NSL-KDD & Features

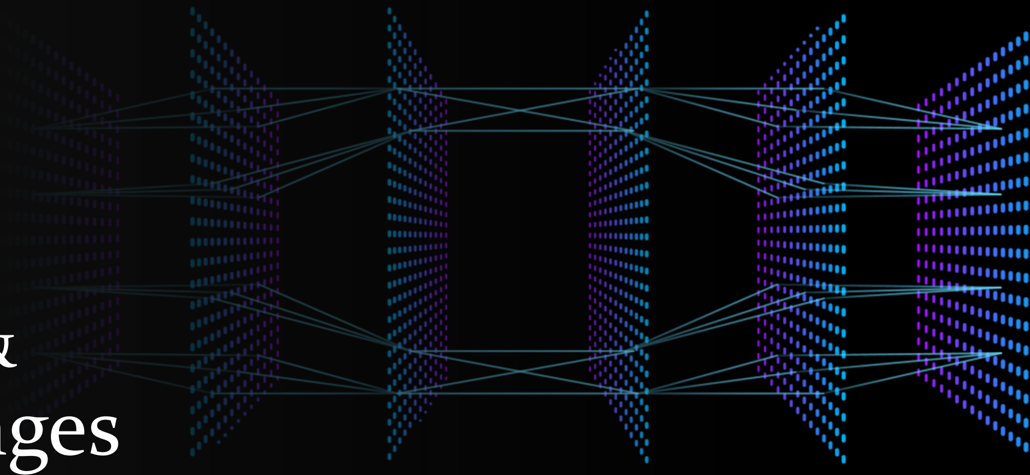
- Contains labeled network traffic data classified as 'Normal' or 'Anomalous'.
- Feature categories:
  - Basic features: protocol, duration, src_bytes, dst_bytes
  - User/Content-based features: failed logins, root access, file creation
  - Traffic-based features: connection rates, packet anomalies

Methodology



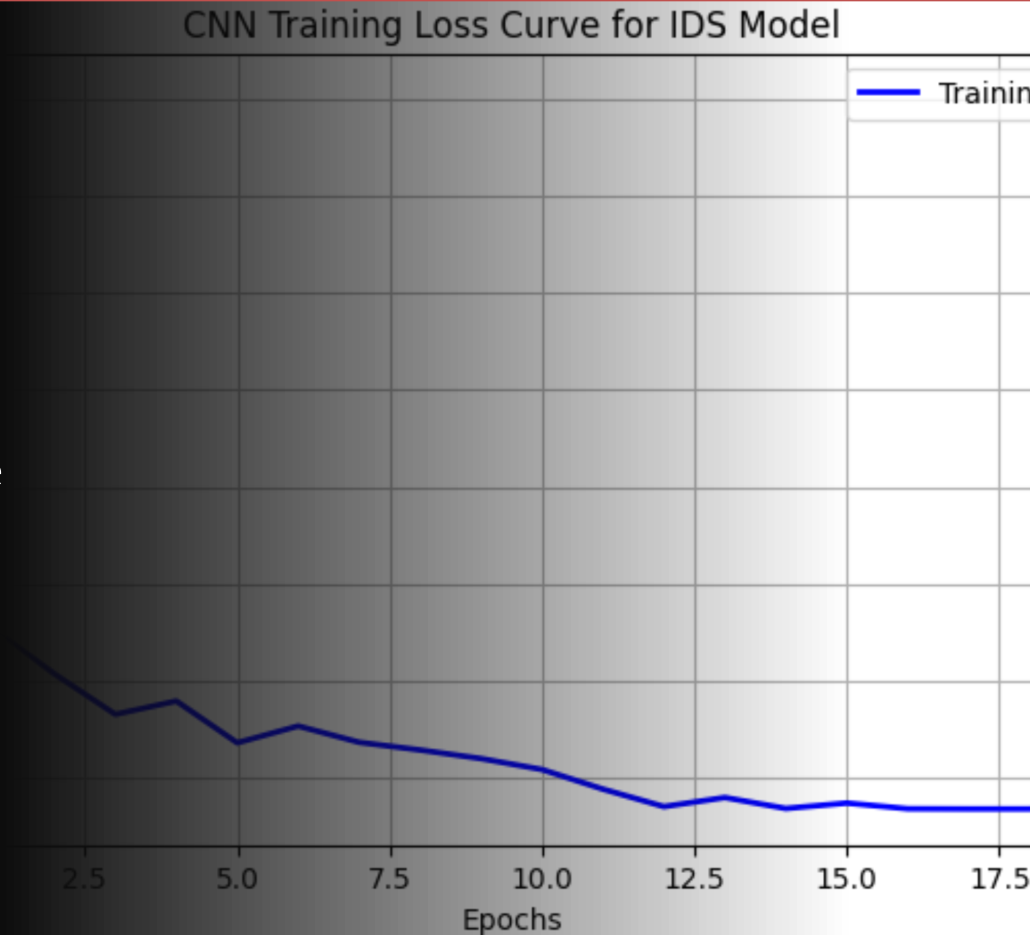
High Level Model Architecture

- Input Layer: 41+ features from NSL-KDD dataset.
- Hidden Layers:
  - 128 neurons --> 64 neurons --> 32 neurons
  - FP + BP + ReLU activation
- Output Layer:
  - 1 neuron
  - Sigmoid for binary classification
- Model performance evaluated using: Accuracy, Precision, Recall, F1-Score, CM



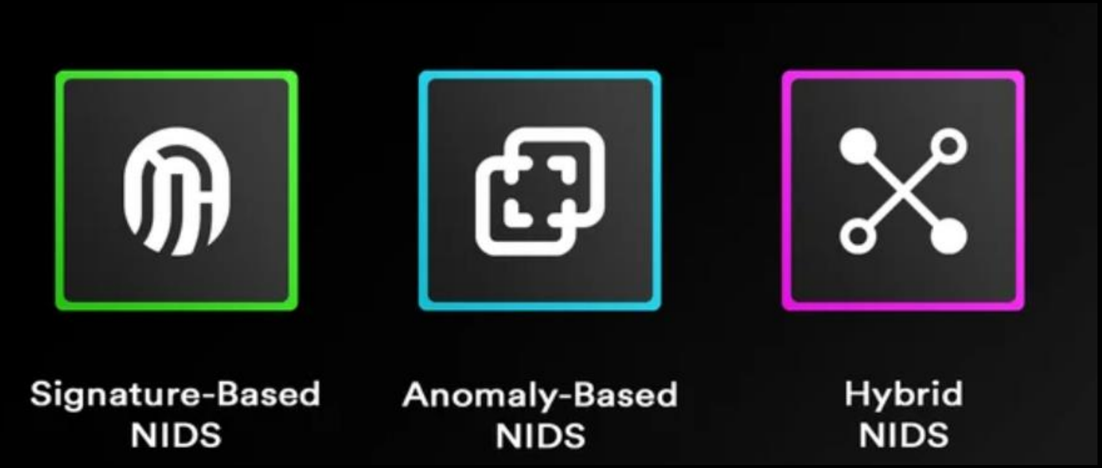
Future Work & Expected Challenges

- Handling imbalanced data within the dataset.
- CNNs or RNNs for better feature extraction.
- Implement unsupervised learning (Autoencoders) for anomaly detection.
- Deploy the model for real-time traffic analysis.



Expected Outcome

- AI-ANN based IDS provides high accuracy and adaptability.
- Can be improved further with hybrid models and real-world deployment.



Timeline

- **Week 1**: EDA
- **Week 2-3**: Model architecture design and initial training.
- **Week 4-5**: Performance tuning and evaluation.
- **Week 6**: Deployment and real-time testing (**optional**).

Thank you! [redacted] Any [redacted] questions [redacted] ? [redacted]