National College of Ireland

Configuration Manual

MSc Research Project
Msc in Data Analytics

Sai Priya Unnava
Student ID: x23329173

School of Computing
National College of Ireland

Supervisor: John Kelly

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name: Sai Priya Unnava

Student ID: X23329173

Programme: Msc in Data Analytics **Year:** 2024-2025

Module: Msc Research Praticum

Lecturer: John Kelly

Submission Due Date: 15-09-2025

Project Title: NT-MAMBA- Network Traffic Anomaly Detection

Word Count: 839 **PageCount:** 4

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Sai Priya Unnava

Date: 15-09-2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Sai Priya Unnava

Student ID: x23329173

1 NT-MAMBA: Code Review & Google Colab Setup Manual

What This Code Does

NT-MAMBA (Network Traffic Multi-scale Anomaly detection with MAMBA) is an advanced cybersecurity system that detects malicious network traffic using state-of-the-art deep learning techniques.

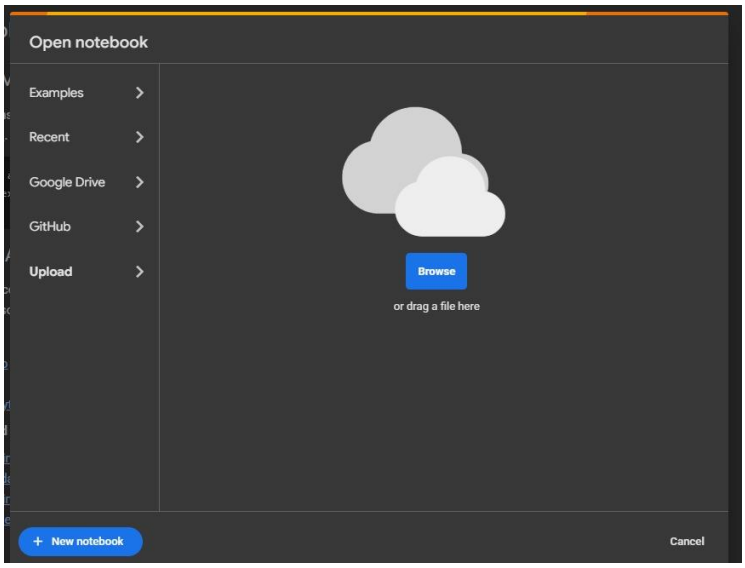
Core Functionality:

1. **Detects 10 types of network attacks** including:
    - DoS attacks (Hulk, SlowHTTPTest, GoldenEye, Slowloris)
    - Brute Force attacks (FTP, SSH)
    - Botnet activities
    - DDoS attacks
    - Infiltration attempts
    - Normal (Benign) traffic
2. **Uses cutting-edge AI models**:
    - **Mamba State Space Model**: A novel architecture for sequence modeling
    - **Graph Neural Networks**: Captures relationships between network flows
    - **Ensemble approach**: Combines multiple models for robust detection
3. **Processes large-scale network data**:
    - Handles 400,000+ network traffic samples
    - 84 features per traffic flow
    - Real-time anomaly detection capability

2 5-Step Google Colab Setup Guide

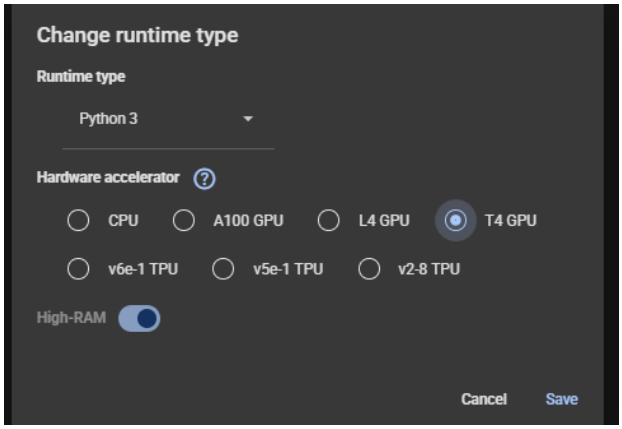
2.1 Step 1: Upload to Google Colab

1. Open Google Colab
2. Click **File → Upload notebook**
3. Upload your ***Priya_NT_MamBA_Aug11_WorkingOne.ipynb*** file
4. Wait for upload to complete



2.2 Step 2: Configure Runtime

1. Click **Runtime → Change runtime type**
2. Select:
   - **Runtime type**: Python 3
   - **Hardware accelerator**: T4 GPU
3. Click **Save**
4. The notebook will restart with GPU enabled



2.3 Step 3: Upload Dataset

1. Click the **folder icon** on the left sidebar
2. Click **Upload** button
3. Select your network_traffic_schema_normalized.csv file
4. Wait for upload (this may take 2-3 minutes for large files)
5. Verify file appears in /content/ directory

2.4 Step 4: Run Installation Cell

Run the first cell to install dependencies:
 # This will install PyTorch, PyTorch Geometric, and other requirements
 # Expected time: 3-5 minutes

2.5 Step 5: Execute Cells Sequentially

Run cells in order:

1. **Cell 1**: Setup & imports (30 seconds)
2. **Cell 2**: Data loading & preprocessing (2-3 minutes)
3. **Cell 3**: Model training & evaluation (10-15 minutes)

3 Configuration Parameters

3.1 Data Processing Settings

```
# Class balancing configuration
classes_to_keep = {
   'Benign': 40000,            # Normal traffic
   'Infilteration': 35000,      # Reduced from original
   'DoS attacks-Hulk': 50000,    # Major attack type
   'DoS attacks-SlowHTTPTest': 50000,
   'FTP-BruteForce': 50000,
   'SSH-Bruteforce': 50000,
   'Bot': 50000,
   'DDOS attack-HOIC': 50000,
   'DoS attacks-GoldenEye': 41508,  # Keep all
   'DoS attacks-Slowloris': 10990   # Keep all
}
```

3.2 Model Configurations

3.2.1 Mamba Model

```
mamba_config = {
   'input_dim': 91,       # 84 original + 7 engineered features
   'd_model': 128,        # Model dimension
   'n_layers': 3,         # Number of Mamba blocks
   'd_state': 16,         # State space dimension
   'd_conv': 4,           # Convolution kernel size
   'expand': 2            # Expansion factor
}
```

3.2.2 Graph Neural Networks

```
gnn_config = {
   'hidden_dim': 128,     # Hidden layer dimension
   'num_layers': 3,       # Number of GNN layers
   'num_heads': 4,        # For GAT model
   'dropout': 0.2,        # Dropout rate
   'max_nodes': 500       # Max nodes per graph
}
```

3.2.3 Training Settings

```
training_config = {
   'batch_size': 512,     # Batch size for training
   'learning_rate': 0.001, # Initial learning rate
   'epochs': 50,          # Training epochs
   'validation_split': 0.2,# Validation set size
   'test_split': 0.2      # Test set size
```

}

4 Common Issues & Solutions

4.1 Issue 1: CUDA Out of Memory

Solution: Reduce batch size to 256 or 128
python
batch_size = 256 *# Instead of 512*

4.2 Issue 2: Dataset Upload Timeout

Solution:

1. Compress CSV file to ZIP
2. Upload ZIP and extract in Colab:

python
!unzip network_traffic_schema_normalized.zip

4.3 Issue 3: PyTorch Geometric Installation Fails

Solution: Use specific CUDA version:
python
!pip install torch-scatter torch-sparse -f https://data.pyg.org/whl/torch-2.0.0+cu118.html

4.4 Issue 4: Slow Training

Solution: Enable mixed precision training:
python
from torch.cuda.amp import autocast, GradScaler
scaler = GradScaler()

5 Monitoring Progress

5.1 Training Progress Indicators

1. **Loss decreasing**: Should drop from ~2.3 to <0.5
2. **Accuracy increasing**: Should reach >90% by epoch 10
3. **GPU utilization**: Should be >80% during training
4. **No error messages**: Red error boxes indicate issu

REFERENCES

Google Colab. (n.d.). *Colab.google*. colab.google. https://colab.google/