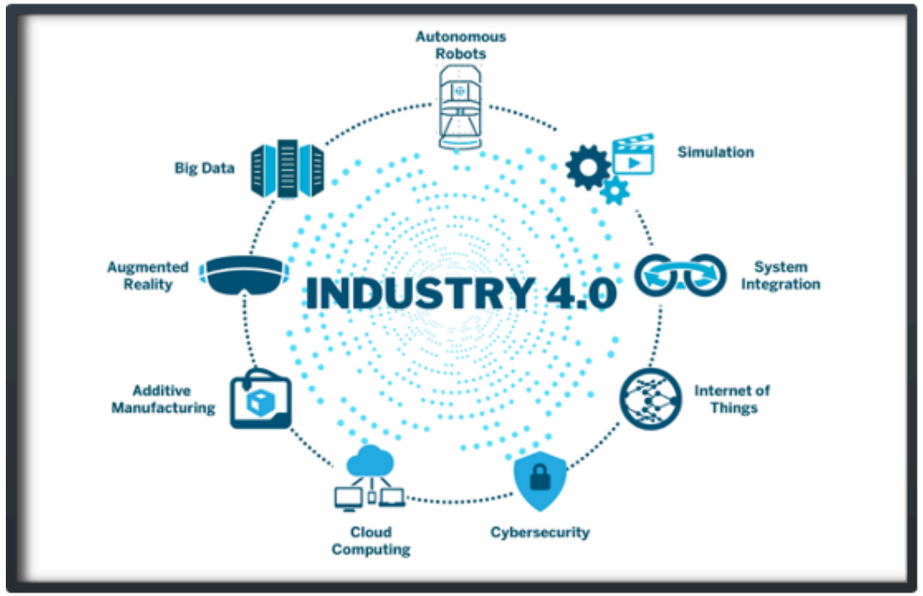
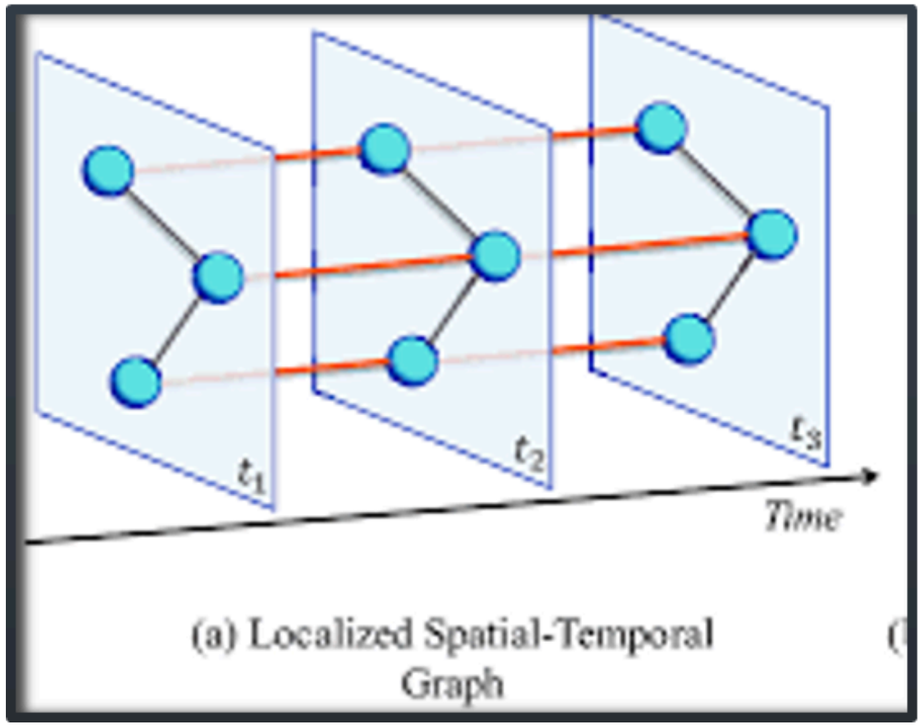
# ANOMALY DETECTION ON TIME SERIES GRAPHS

Roberto Giordano
A.A 2024/2025

Supervisor:
Karina Chichifoi

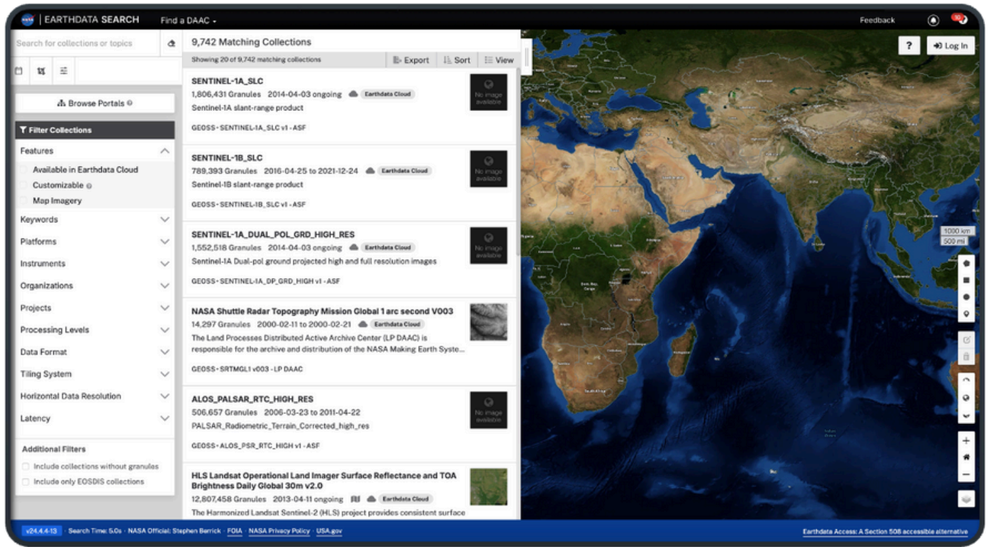



(a) Localized Spatial-Temporal Graph

# Introduction

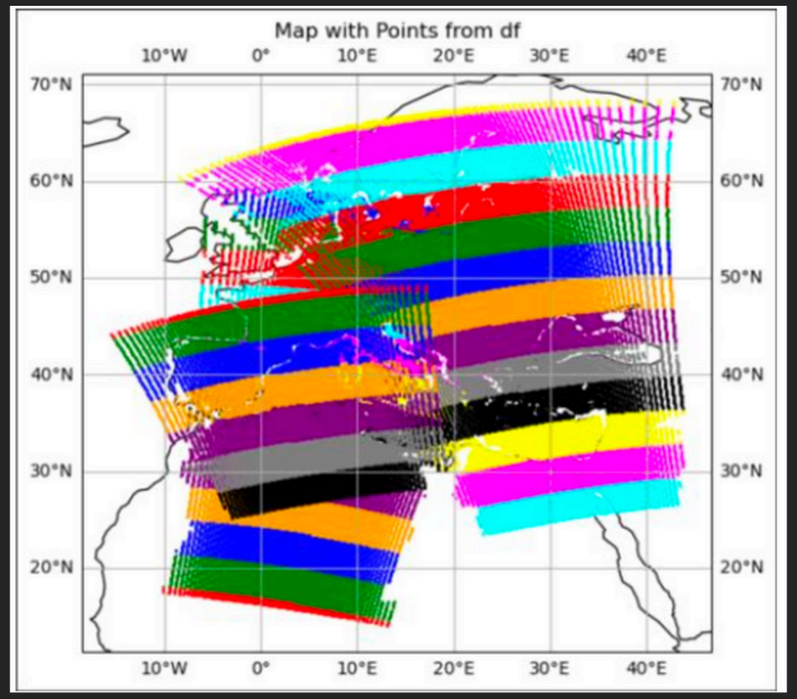
- Industry 4.0
- Cybersecurity threats
- Anomaly Detection
- Spatial and Temporal Correlation
- Novel approach applied in this field

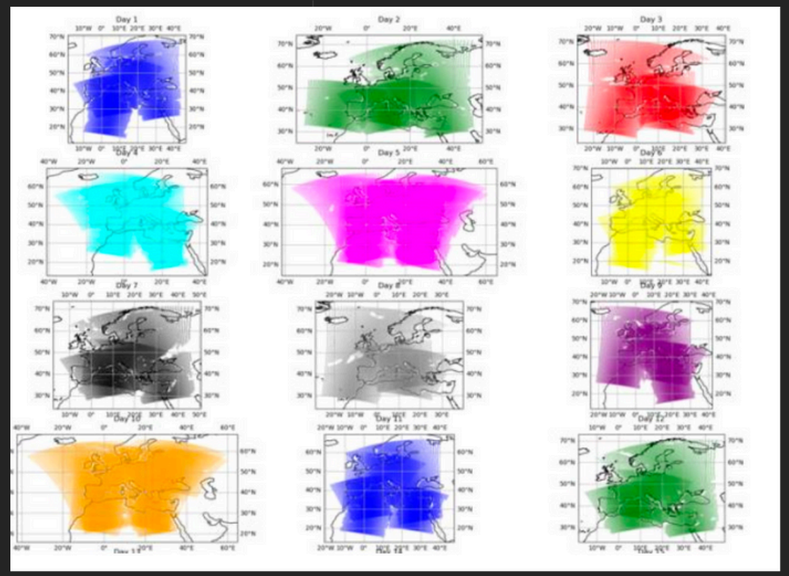



# DATASET: SNDRSNML2RMS_1

- The EarthData tool
- Data Organization

| | |
|---|---|
| **1410** Granules | **64.1 GB** out of 2.5 TB |
| **8 months** across 2013 and 2015 | **Europe** window selected |

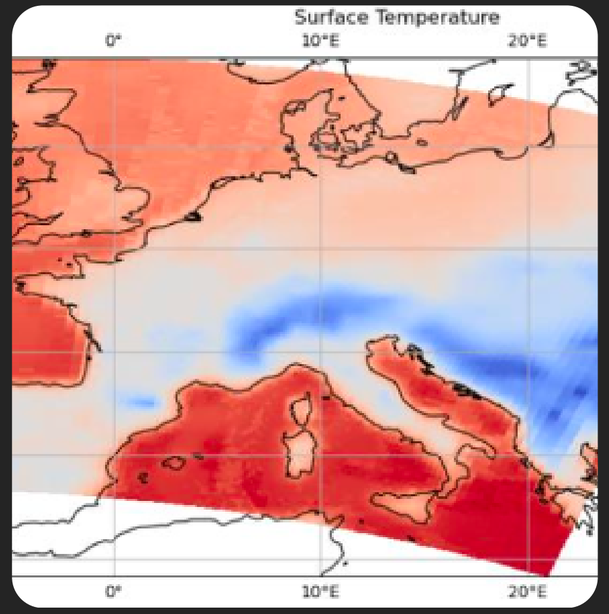


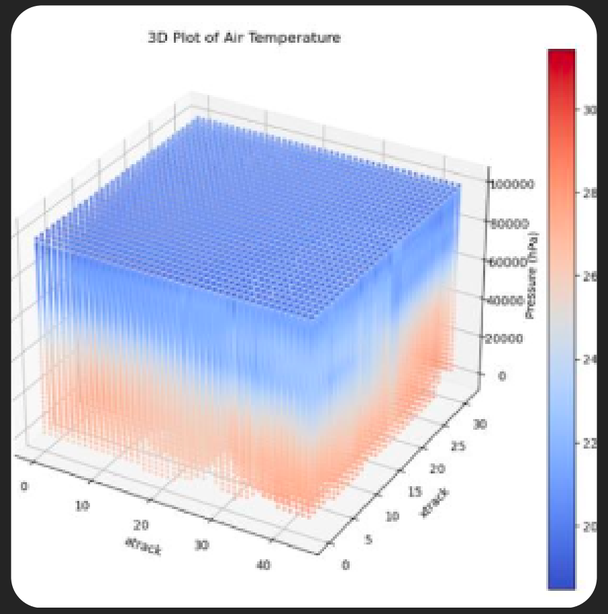
# DATASET: SNDRSNML2RMS_1

- Analysis and interpretation of data

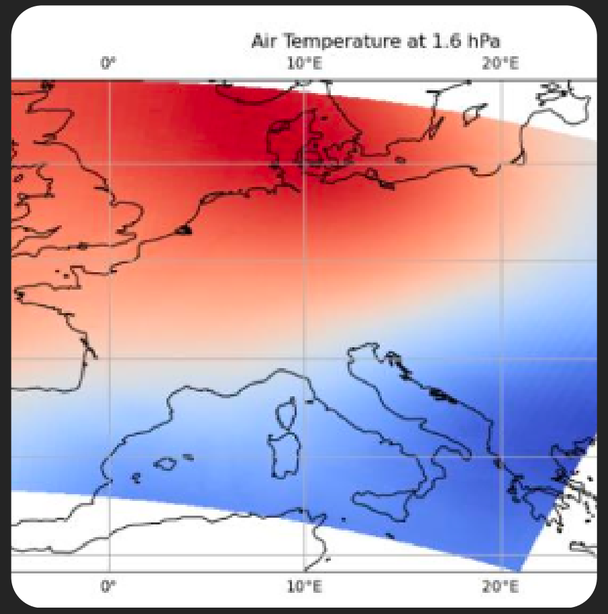
# Variables




Surface temperature




Air temperature




Air temperature

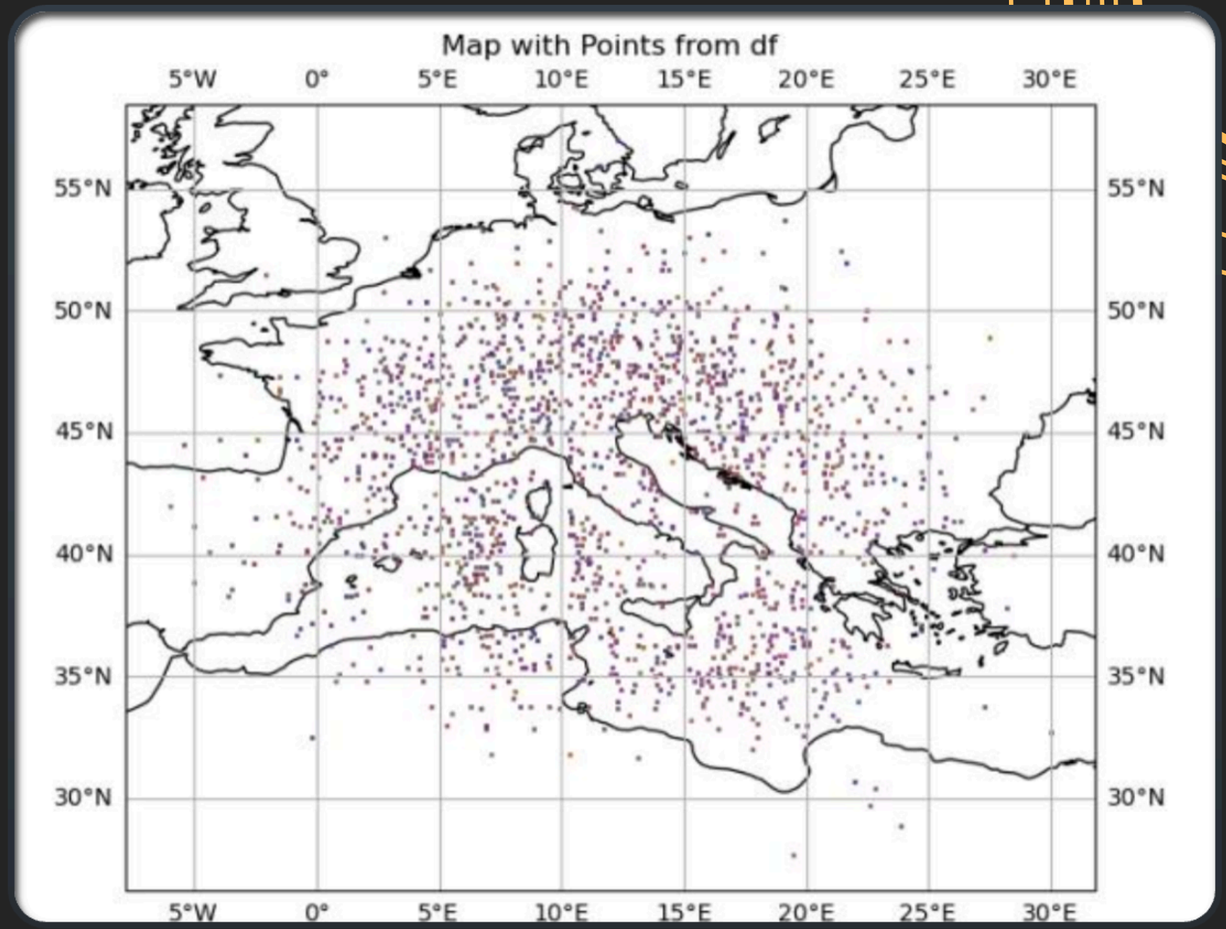
at pression 1.6hPa

# Variables Selection

**Table 1**

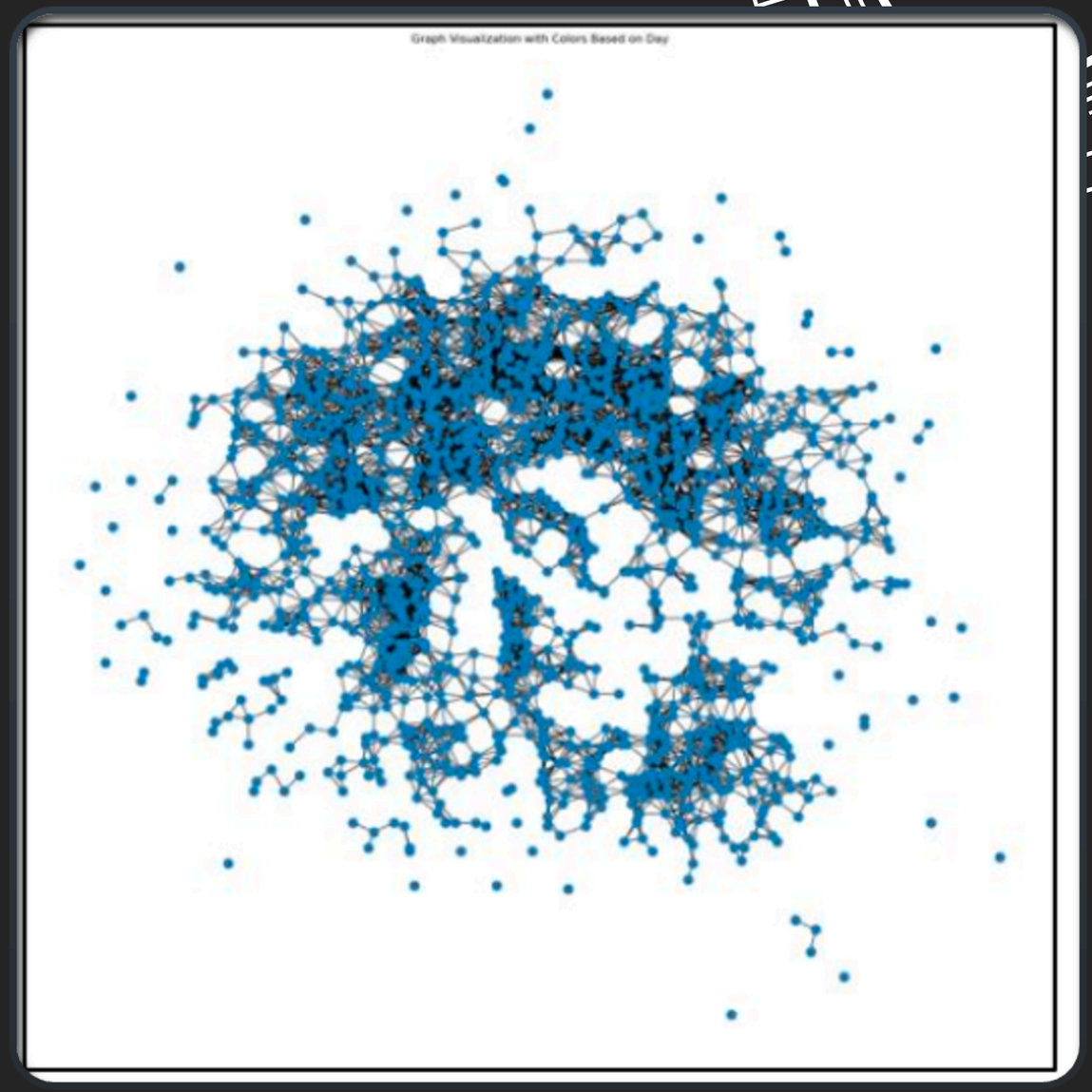
| Variable Name | Description | Type | Units | Dimensions |
|---|---|---|---|---|
| **solar_zenith_angle** | The angle between the sun and the zenith (directly overhead) at a specific observation point. | float32 | degrees | (atrack, xtrack) |
| **solar_azimuth_angle** | The angle of the sun's position relative to true north at a specific observation point. | float32 | degrees | (atrack, xtrack) |
| **surf_air_temp_masked** | The near-surface air temperature (~2 meters above ground), with invalid data masked. | float32 | Kelvin | (atrack, xtrack) |
| **surf_temp_masked** | The radiative temperature of the Earth's surface, with invalid data masked. | float32 | Kelvin | (atrack, xtrack) |
| **surf_spec_hum_masked** | The specific humidity (mass fraction of water vapor) near the surface, with invalid data masked. | float32 | unitless | (atrack, xtrack) |
| **h2o_vap_tot_masked** | The total column amount of water vapor in the atmosphere, with invalid data masked. | float32 | $kg/m^2$ | (atrack, xtrack) |
| **cloud_liquid_ water_masked** | The total liquid water content in clouds, with invalid data masked. | float32 | $kg/m^2$ | (atrack, xtrack) |
| **atmosphere_mass_ content_of_cloud _ice_masked** | The total mass of ice within clouds in the atmospheric column, with invalid data masked. | float32 | $kg/m^2$ | (atrack, xtrack) |

| Decimal Places | Approximate Error (km) |
|---|---|
| 0 decimals | 111 km |
| 1 decimal | 11.1 km |
| 2 decimals | 1.11 km |
| 3 decimals | 111 m |
| 4 decimals | 11.1 m |
| 5 decimals | 1.11 m |
| 6 decimals | 0.11 m |




# Latitude and Longitude Approximation

# Static Graph Creation



# Model Architecture

# Model Architecture: overview

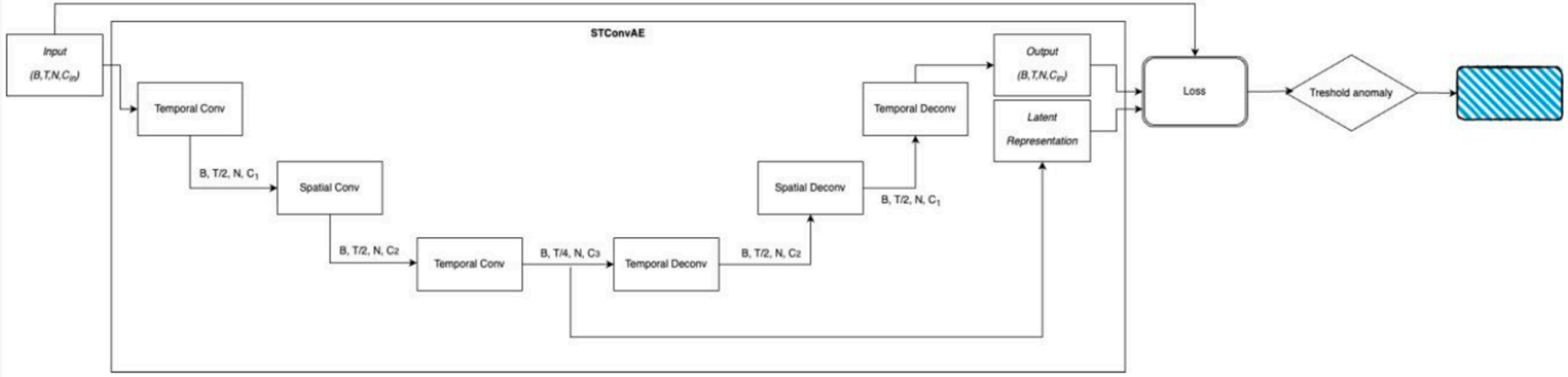
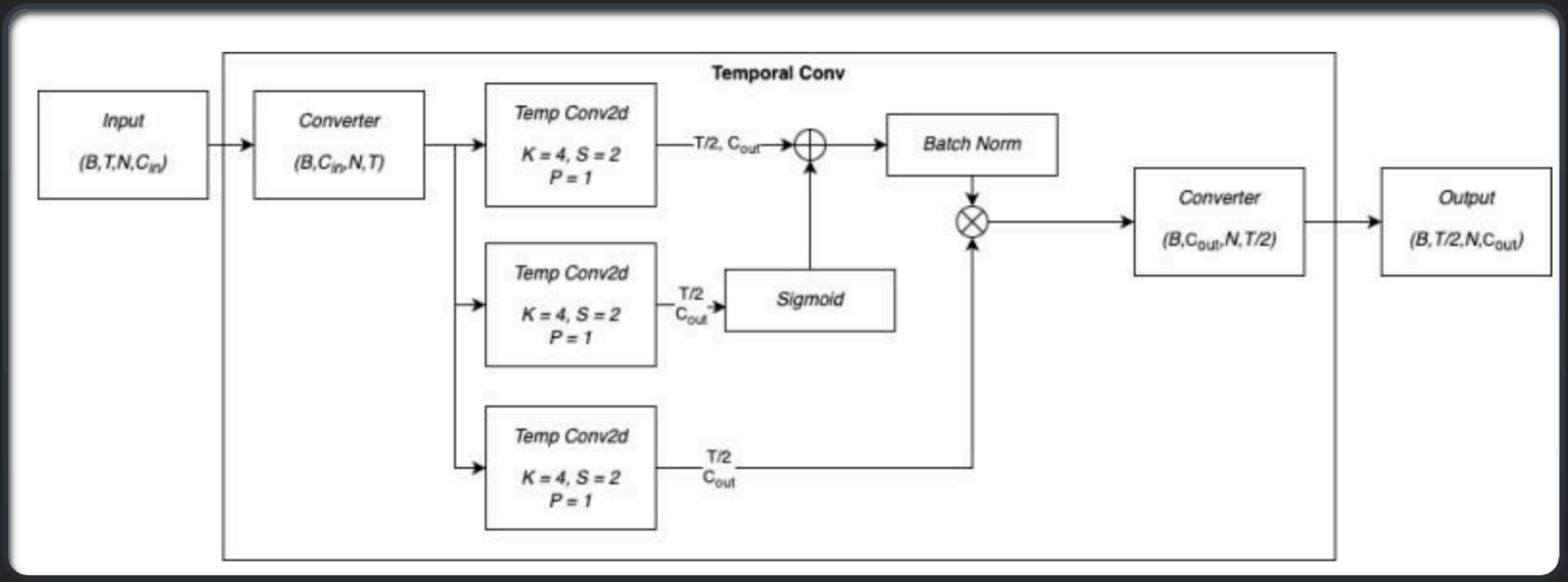


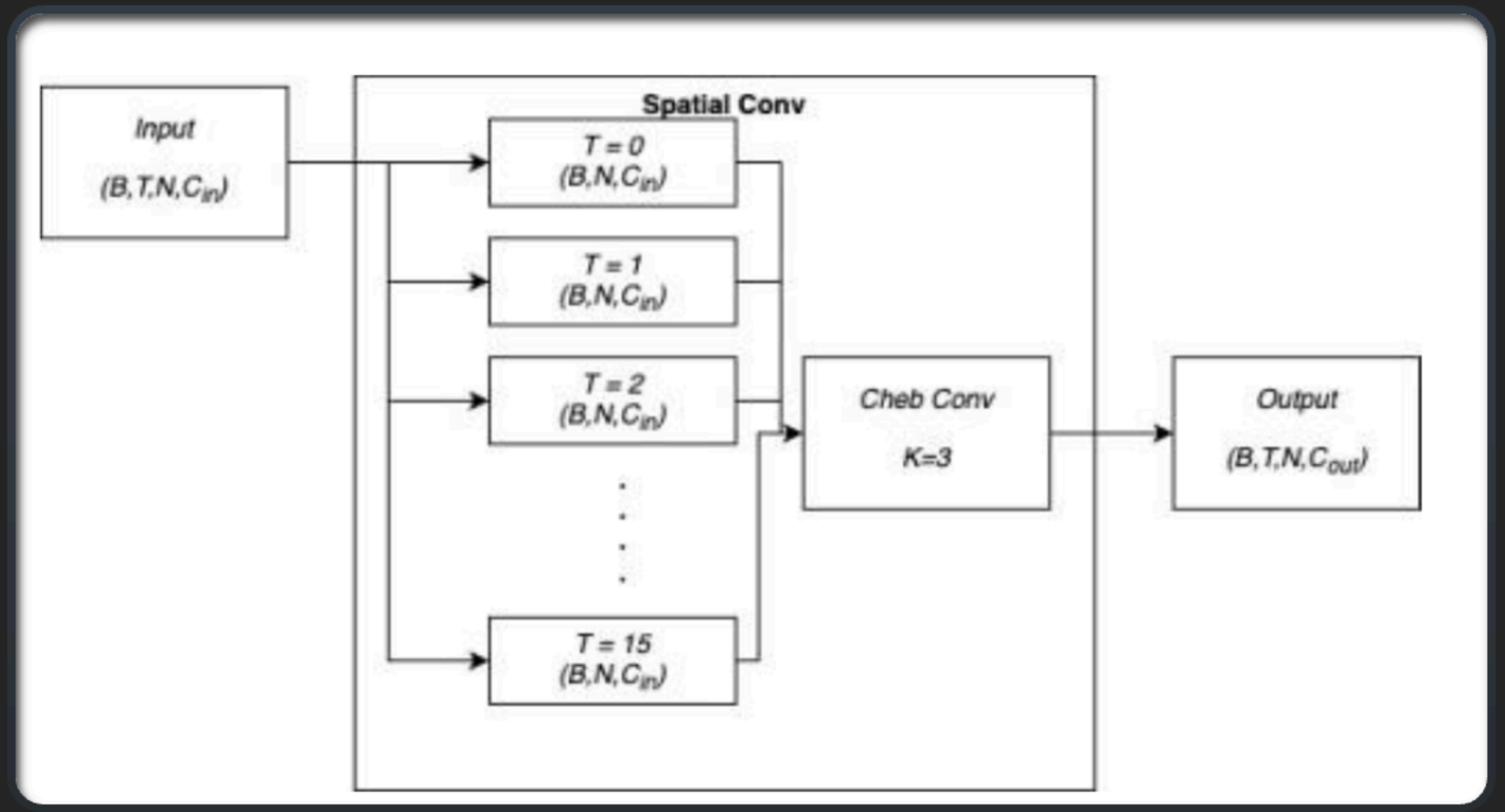

***Figure 8***

*General architecture of the STConvAE model. Here, $C_1=C_{in}$. In the final model configuration, $[C_1,C_2,C_3]=[8,16,32]$.*

# Model Architecture: temporal block

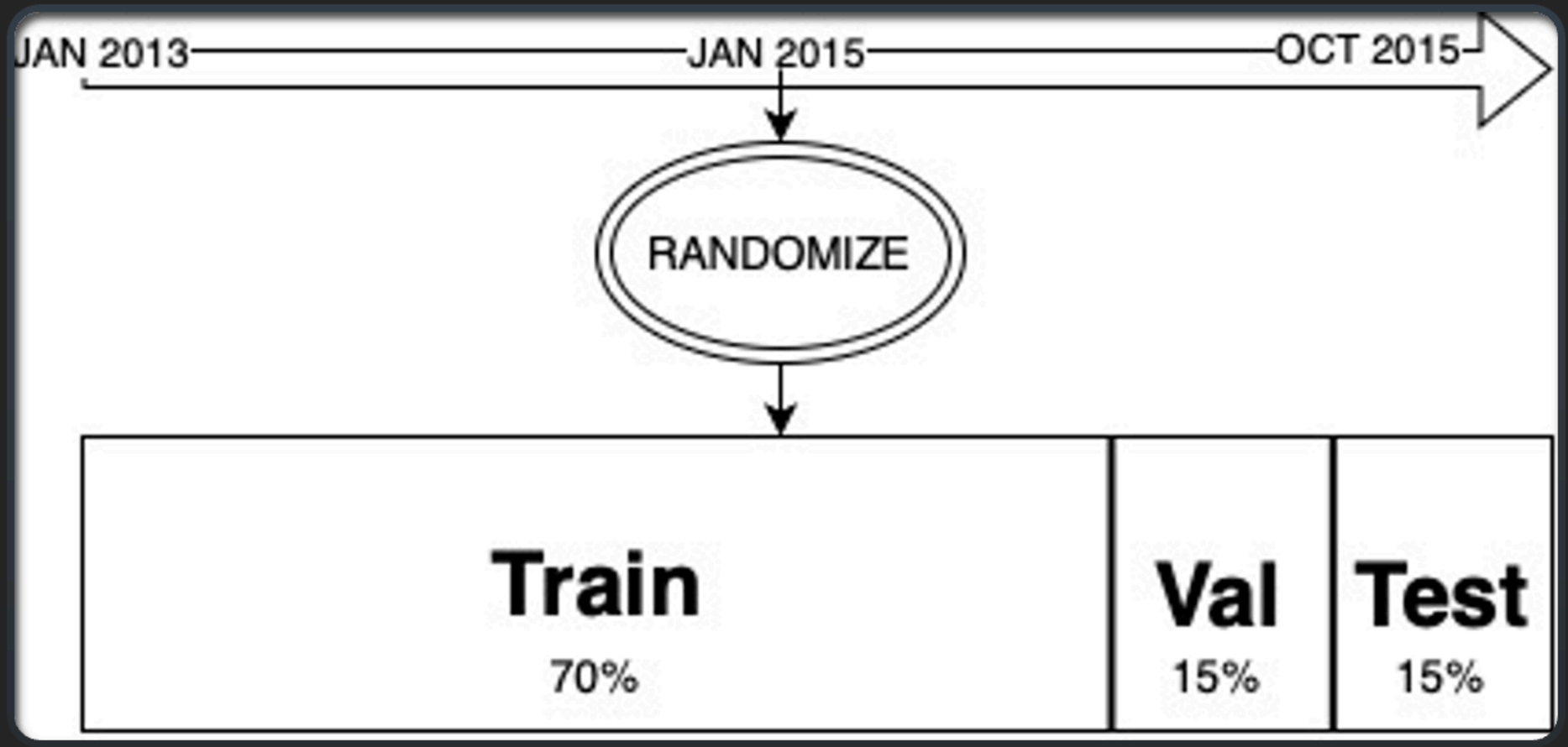


# Model Architecture: spatial block

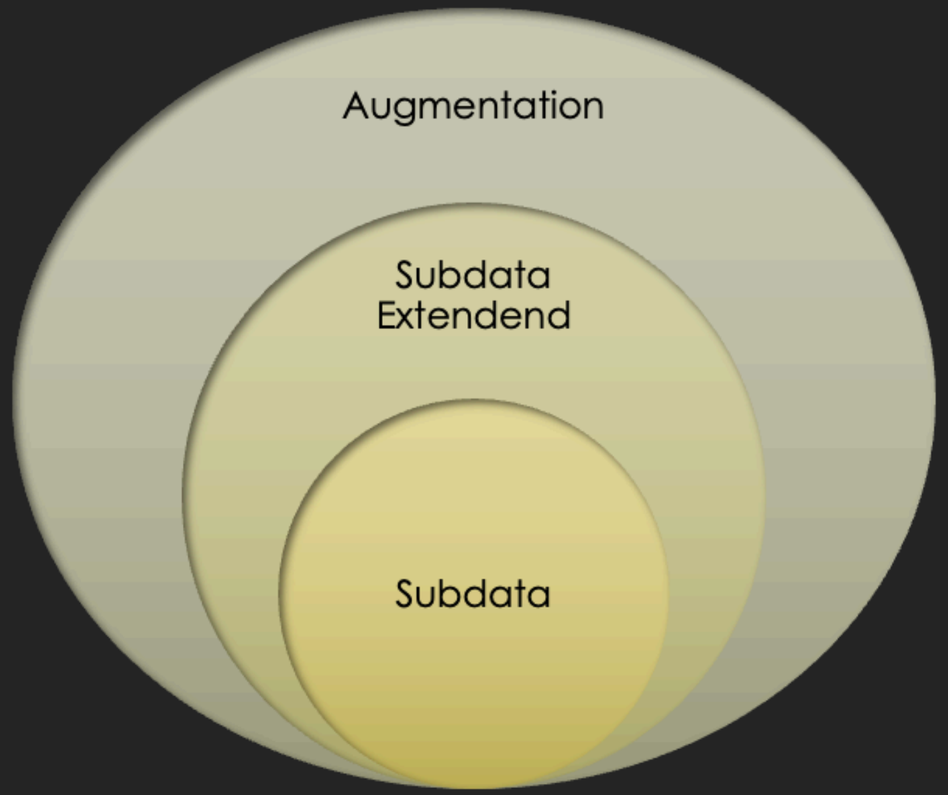


# Experiments

# Experiments: Data Splitting Strategy



# Experiments: Dataset variation



# Experiments: Architectures

Temporal Sandwich

Spatial Sandwich

Double Spatial and Temporal Sandwich

# Experiments: Activation and Gate Functions

**Activation:**

Relu

Leaky Relu

**Gate:**

Sigmoid

TanH

# Experiments: Loss functions

## MSE

Mean square error between input and output.

## Edge Weight Preservation

The distances of the nodes in the input must be preserved in the latent space.

## Pairwise similarity

Linked nodes must be similar

## L1

Absolute error between input and output

## Latent space

MSE and/or L1 applied to latent space

## Weighted Function

Different weights for a combination of them

# Experiments: Data input for training

01 Temporal Window

8-16-24 days

02 Batch Size

2-4-16

# Final Architecture and Parameters

# Best model configuration

| | |
|---|---|
| 01 | Kernel Size: 4 |
| 02 | Stride: 2 |
| 03 | Padding: 1 |
| 04 | Chebyshev Polynomial Order (K): 3 |
| 05 | Gated Function: Sigmoid |
| 06 | Activation Function: Leaky ReLU (Slope 0.01) |
| 07 | Architecture: Encoder-decoder |
| 08 | Training Epochs: 150 |
| 09 | Loss: Edge Weight Preservation + Pairwise Similarity + MSE |
| 10 | Dataset: subdata_extended (2013 and 2015), without augmentation. |

# Architectural Details

| | |
|---|---|
| 01 | Parameters: 22976 |
| 02 | Number of channels evoved through the layers: 8->16->32 |
| 03 | Temporal layers halved the temporal dimension at each step. |
| 04 | Training and evaluation were conducted using a batch size of 2 for training and validation,and 1 for testing. |

# Data statistics and model Inputs:

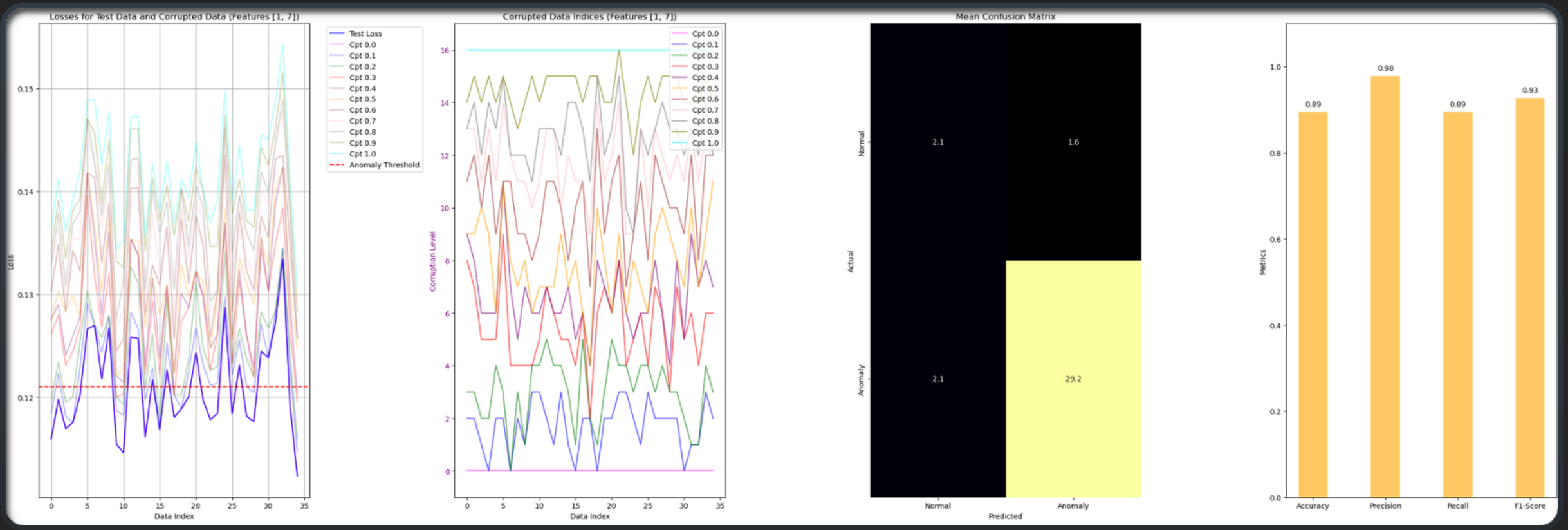
| | |
|---|---|
| 01 | Number of nodes: 1657 |
| 02 | Number of features: 8 |
| 03 | Temporal steps: 16 |
| 04 | Train/Val/Test split: 70%-15%-15% |

# Edge Structure

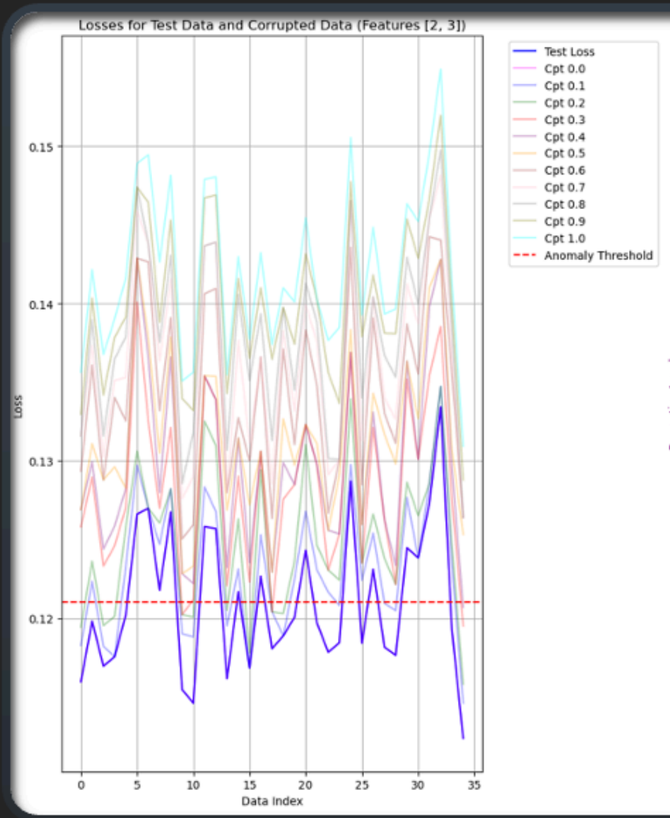
01 Edge index shape: [2, 12564]

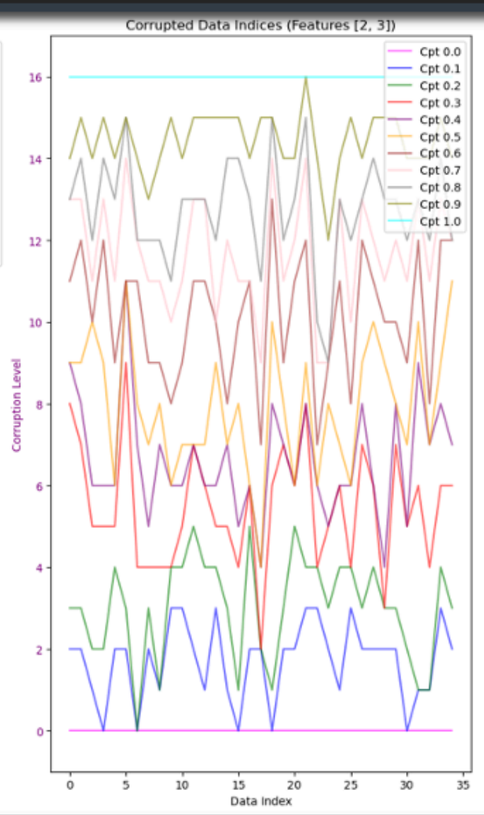
02 Edge weights shape: [12564]

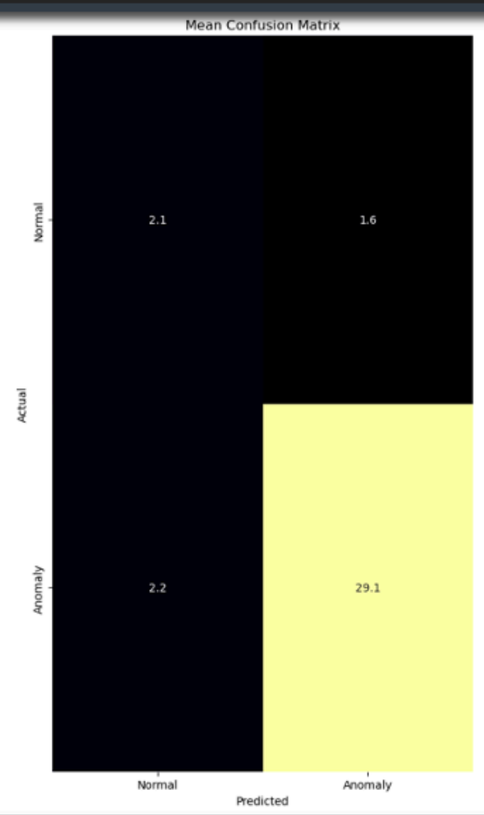
# Results

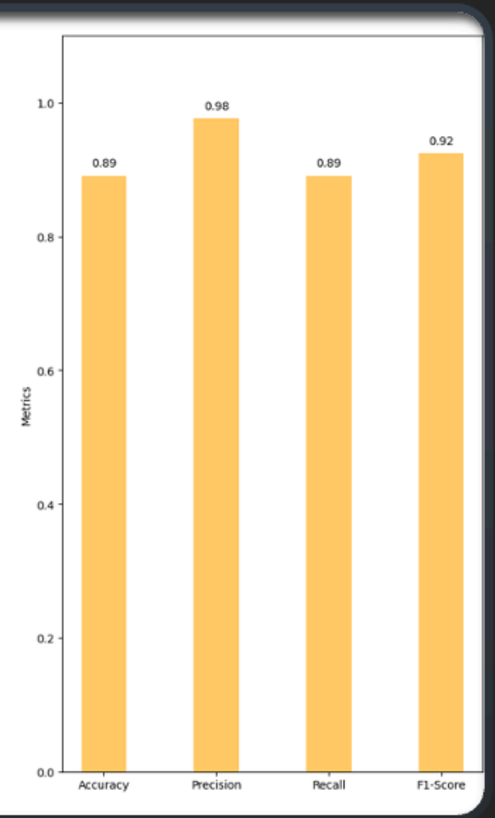

Losses for Test Data and Corrupted Data (Features [1, 7])
Loss
Data Index
Test Loss
Cpt 0.0
Cpt 0.1
Cpt 0.2
Cpt 0.3
Cpt 0.4
Cpt 0.5
Cpt 0.6
Cpt 0.7
Cpt 0.8
Cpt 0.9
Cpt 1.0
Anomaly Threshold
Corrupted Data Indices (Features [1, 7])
Corruption Level
Data Index
Mean Confusion Matrix
Normal
Anomaly
Actual
Predicted
2.1
1.6
2.1
29.2
Metrics
Accuracy
Precision
Recall
F1-Score
0.89
0.98
0.89
0.93

# Results

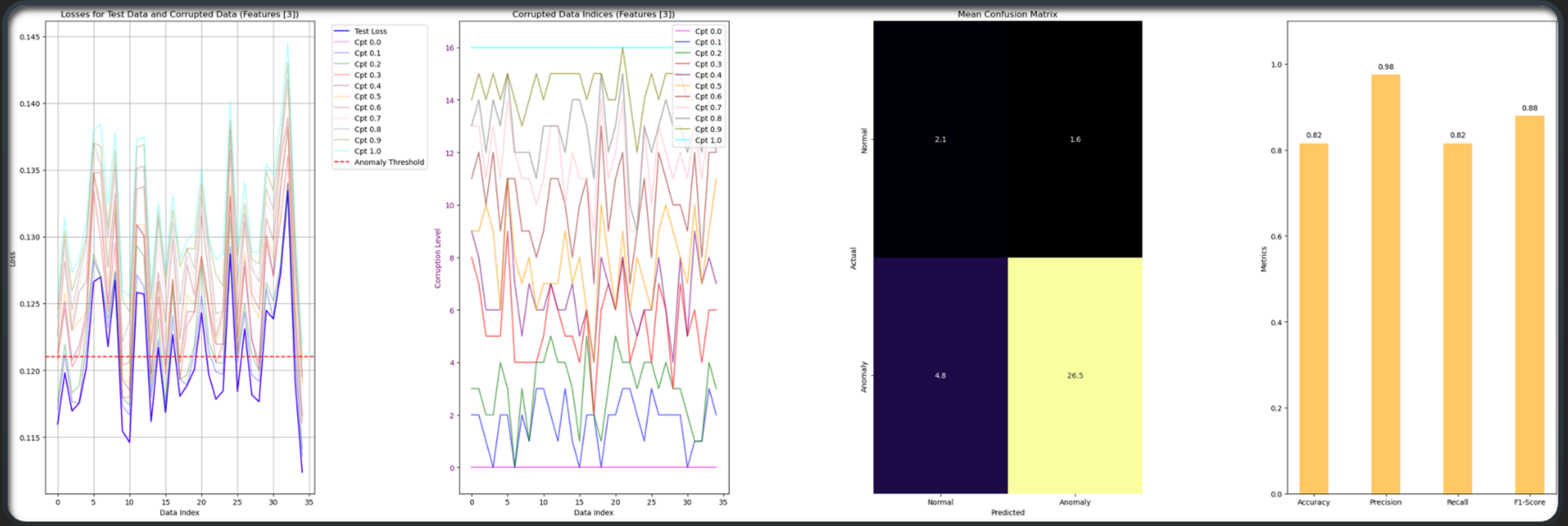

Losses for Test Data and Corrupted Data (Features [2, 3])
Test Loss
Cpt 0.0
Cpt 0.1
Cpt 0.2
Cpt 0.3
Cpt 0.4
Cpt 0.5
Cpt 0.6
Cpt 0.7
Cpt 0.8
Cpt 0.9
Cpt 1.0
Anomaly Threshold
Loss
Data Index


Corrupted Data Indices (Features [2, 3])
Cpt 0.0
Cpt 0.1
Cpt 0.2
Cpt 0.3
Cpt 0.4
Cpt 0.5
Cpt 0.6
Cpt 0.7
Cpt 0.8
Cpt 0.9
Cpt 1.0
Corruption Level
Data Index

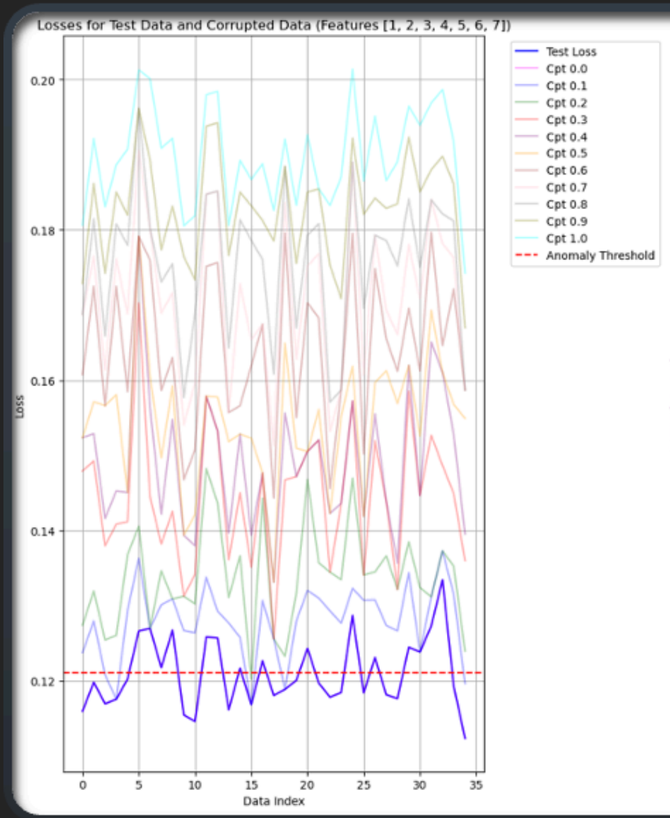

Mean Confusion Matrix
Normal
Anomaly
2.1
1.6
2.2
29.1
Actual
Normal
Anomaly
Predicted

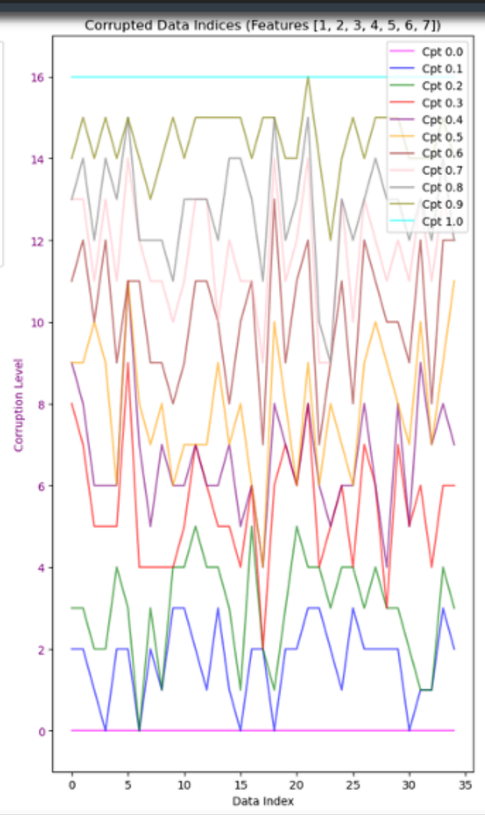

0.89
0.98
0.89
0.92
Metrics
Accuracy
Precision
Recall
F1-Score

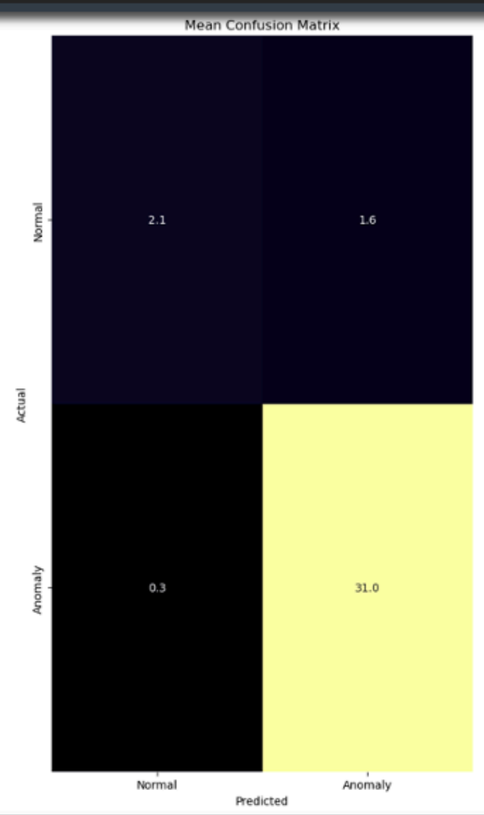
# Results

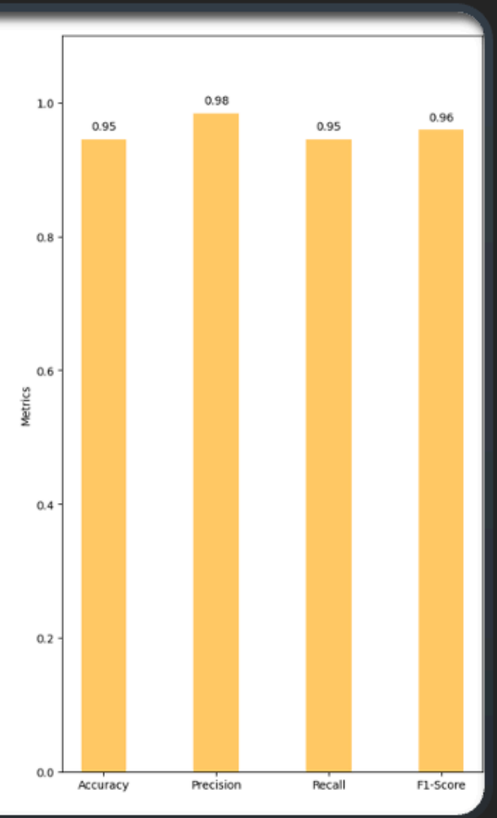

Losses for Test Data and Corrupted Data (Features [3])
Loss
Data Index
Test Loss
Cpt 0.0
Cpt 0.1
Cpt 0.2
Cpt 0.3
Cpt 0.4
Cpt 0.5
Cpt 0.6
Cpt 0.7
Cpt 0.8
Cpt 0.9
Cpt 1.0
Anomaly Threshold
Corrupted Data Indices (Features [3])
Corruption Level
Data Index
Mean Confusion Matrix
Actual
Predicted
Normal
Anomaly
2.1
1.6
4.8
26.5
Metrics
0.82
0.98
0.82
0.88
Accuracy
Precision
Recall
F1-Score

# Results

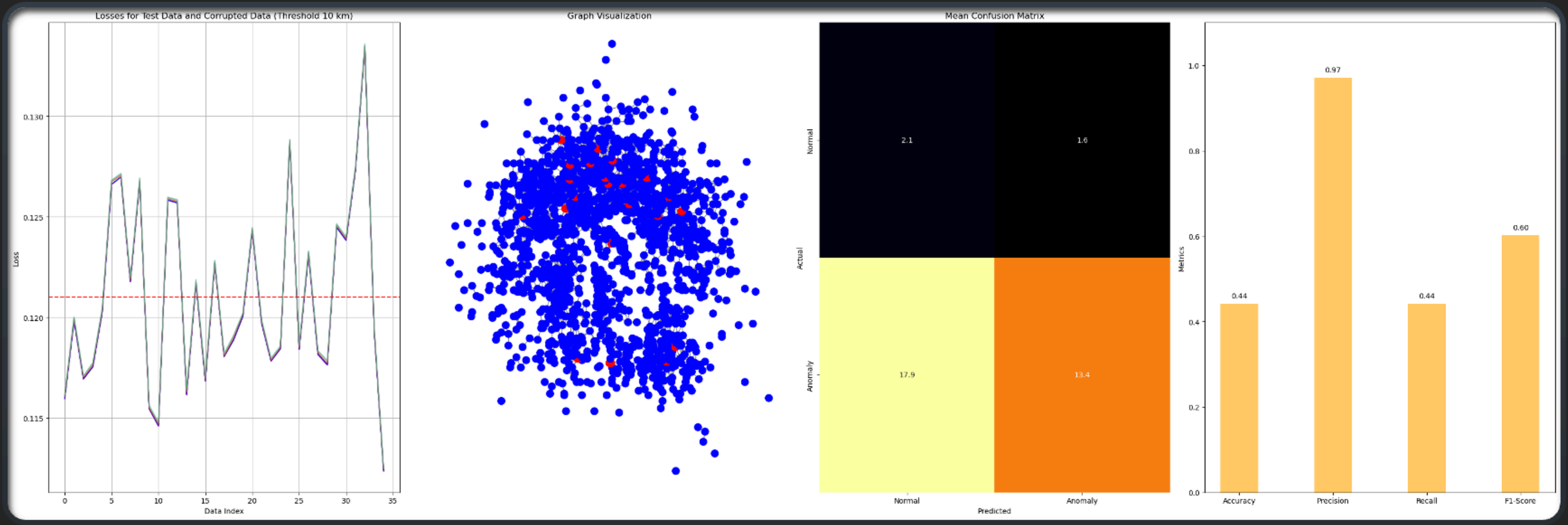

Losses for Test Data and Corrupted Data (Features [1, 2, 3, 4, 5, 6, 7])
Test Loss
Cpt 0.0
Cpt 0.1
Cpt 0.2
Cpt 0.3
Cpt 0.4
Cpt 0.5
Cpt 0.6
Cpt 0.7
Cpt 0.8
Cpt 0.9
Cpt 1.0
Anomaly Threshold
Loss
Data Index


Corrupted Data Indices (Features [1, 2, 3, 4, 5, 6, 7])
Cpt 0.0
Cpt 0.1
Cpt 0.2
Cpt 0.3
Cpt 0.4
Cpt 0.5
Cpt 0.6
Cpt 0.7
Cpt 0.8
Cpt 0.9
Cpt 1.0
Corruption Level
Data Index

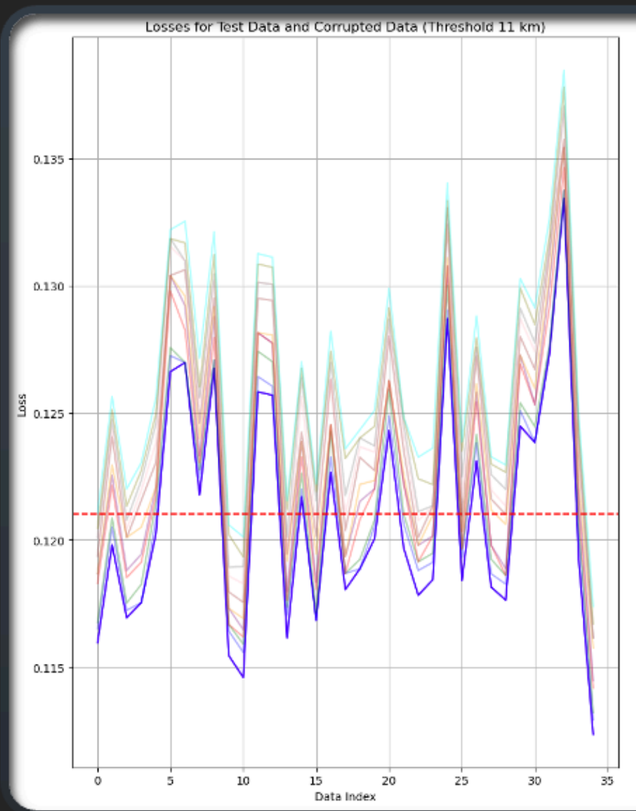

Mean Confusion Matrix
Normal
Anomaly
2.1
1.6
0.3
31.0
Actual
Normal
Anomaly
Predicted

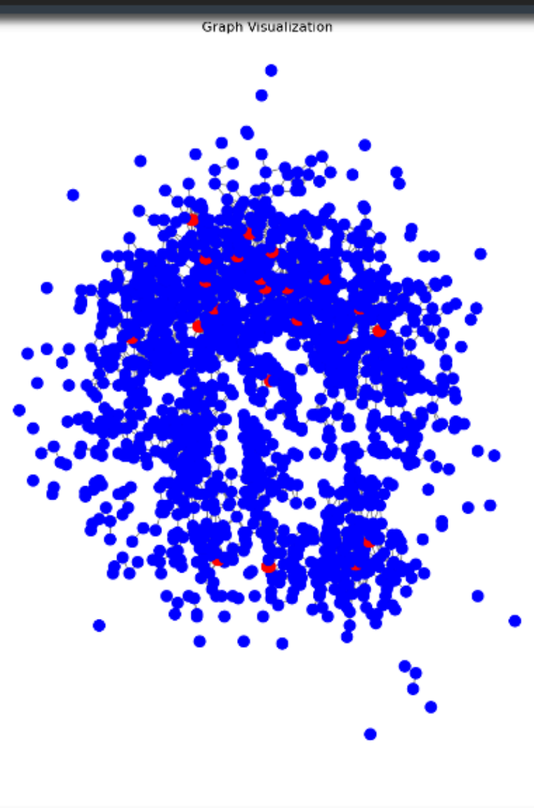

0.95
0.98
0.95
0.96
Metrics
Accuracy
Precision
Recall
F1-Score

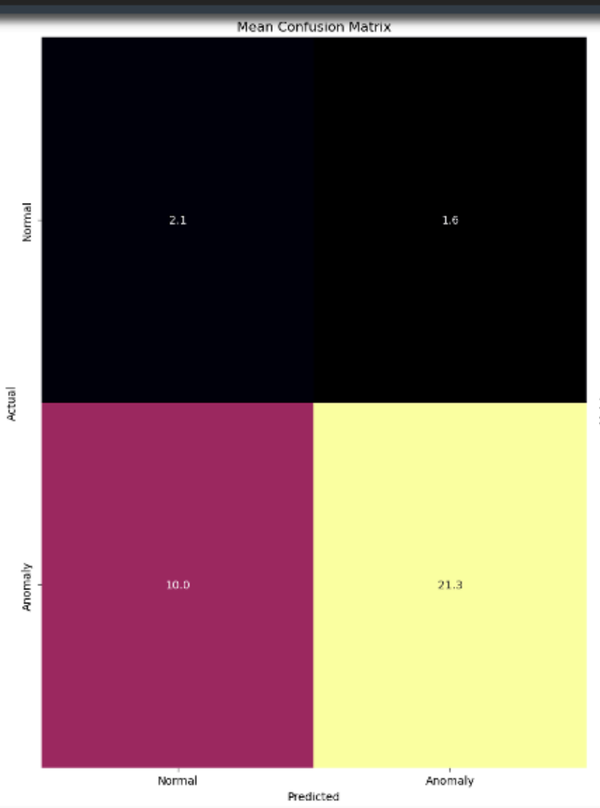
# Results

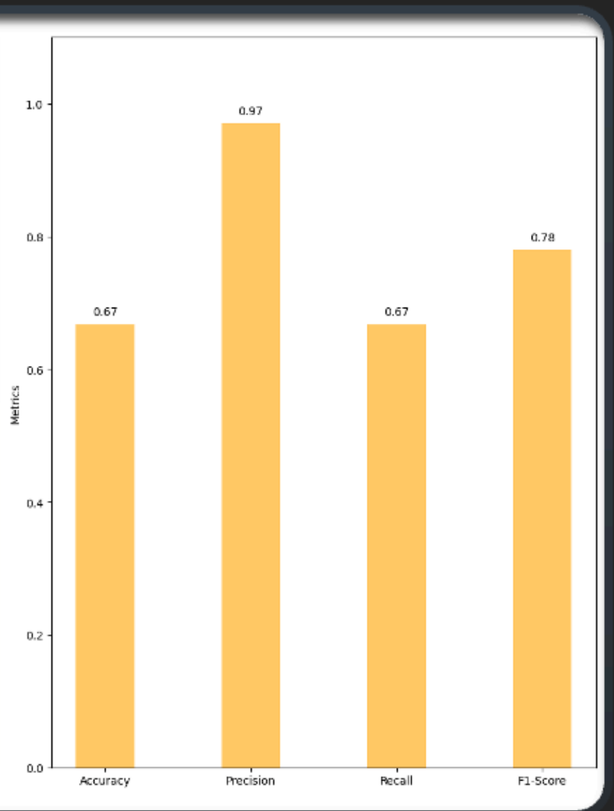

Losses for Test Data and Corrupted Data (Threshold 10 km)
Loss
Data Index
0.130
0.125
0.120
0.115
0
5
10
15
20
25
30
35
Graph Visualization
Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
17.9
13.4
Normal
Anomaly
Predicted
Metrics
1.0
0.8
0.6
0.4
0.2
0.0
0.44
0.97
0.44
0.60
Accuracy
Precision
Recall
F1-Score

# Results

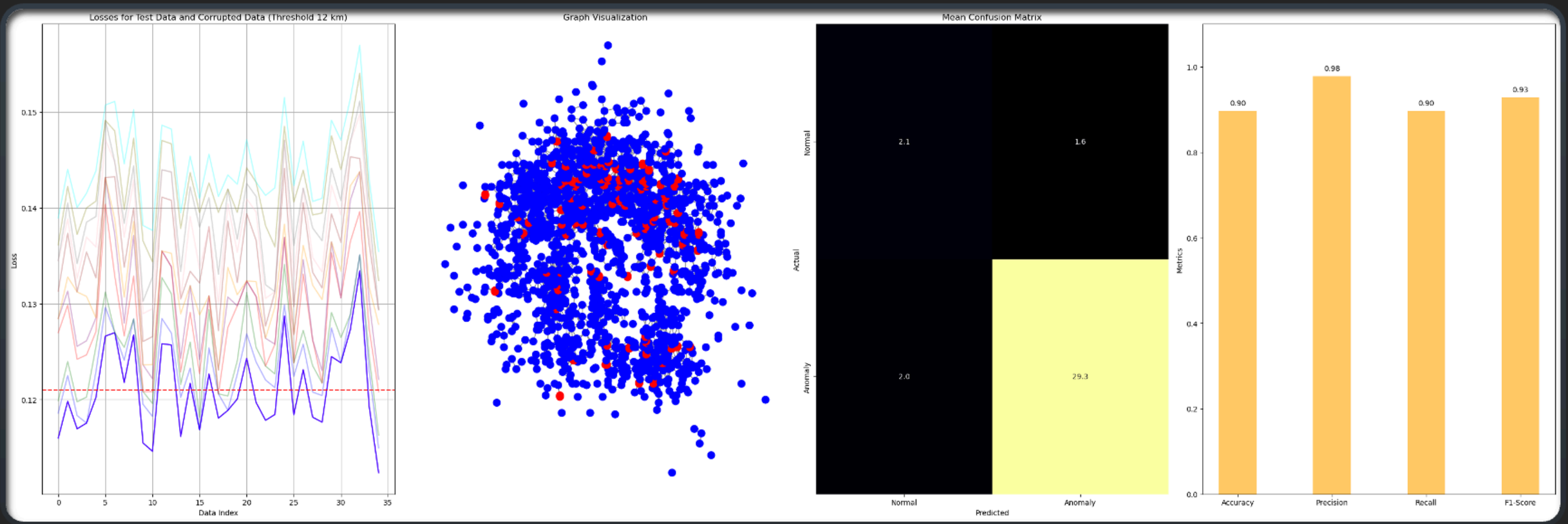

Losses for Test Data and Corrupted Data (Threshold 11 km)
Loss
Data Index
0.135
0.130
0.125
0.120
0.115
0
5
10
15
20
25
30
35


Graph Visualization

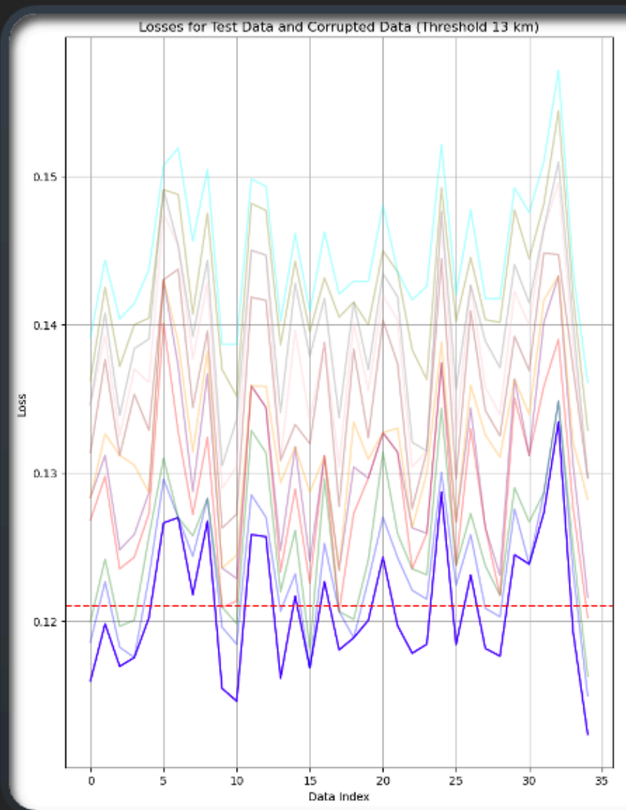

Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
10.0
21.3
Normal
Anomaly
Predicted

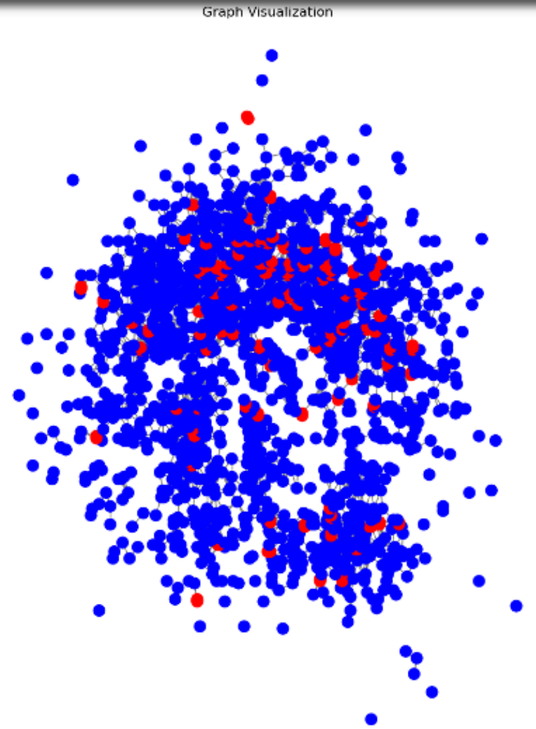

Metrics
1.0
0.8
0.6
0.4
0.2
0.0
0.67
0.97
0.67
0.78
Accuracy
Precision
Recall
F1-Score

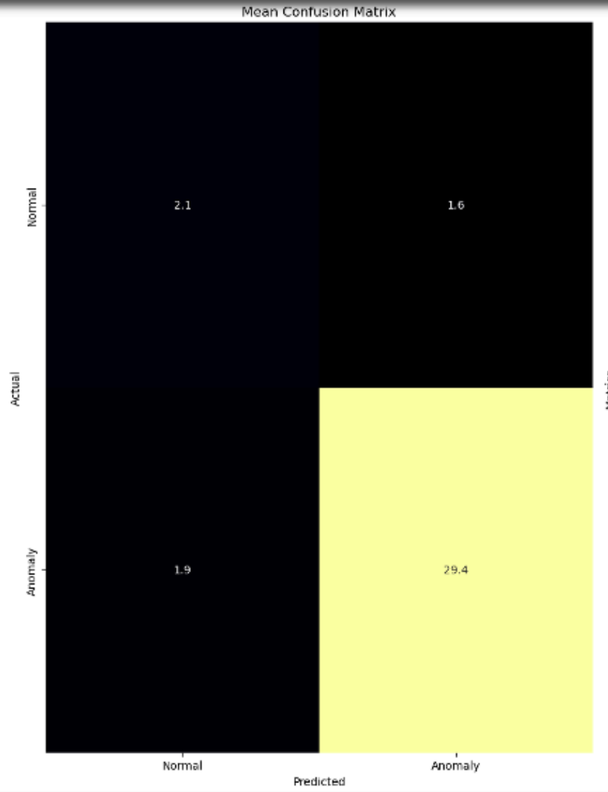
# Results

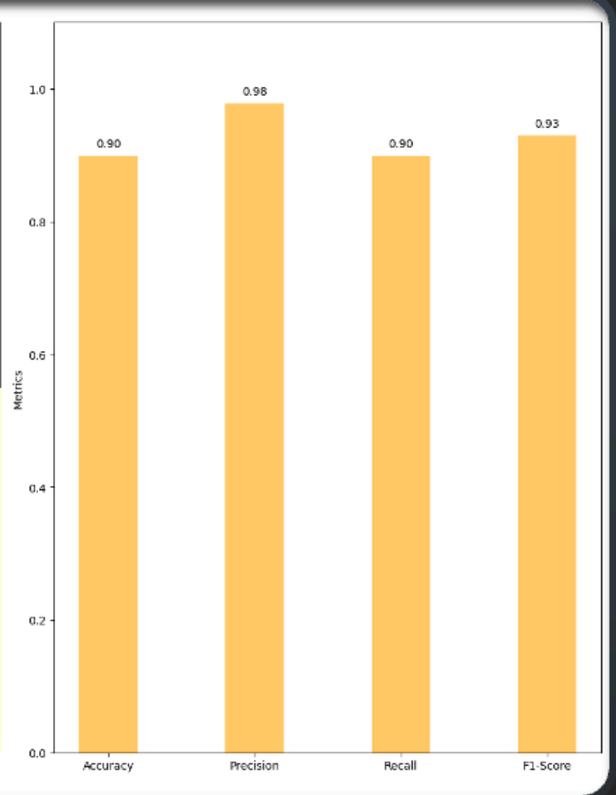

Losses for Test Data and Corrupted Data (Threshold 12 km)
Loss
Data Index
0.15
0.14
0.13
0.12
0
5
10
15
20
25
30
35
Graph Visualization
Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
2.0
29.3
Normal
Anomaly
Predicted
Metrics
1.0
0.8
0.6
0.4
0.2
0.0
0.90
0.98
0.90
0.93
Accuracy
Precision
Recall
F1-Score

# Results

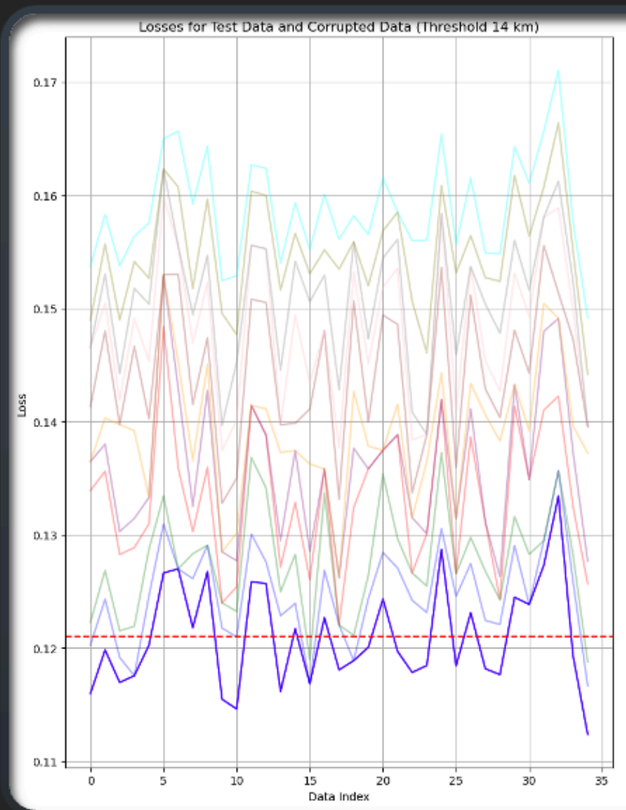

Losses for Test Data and Corrupted Data (Threshold 13 km)
Loss
Data Index
0.15
0.14
0.13
0.12
0
5
10
15
20
25
30
35

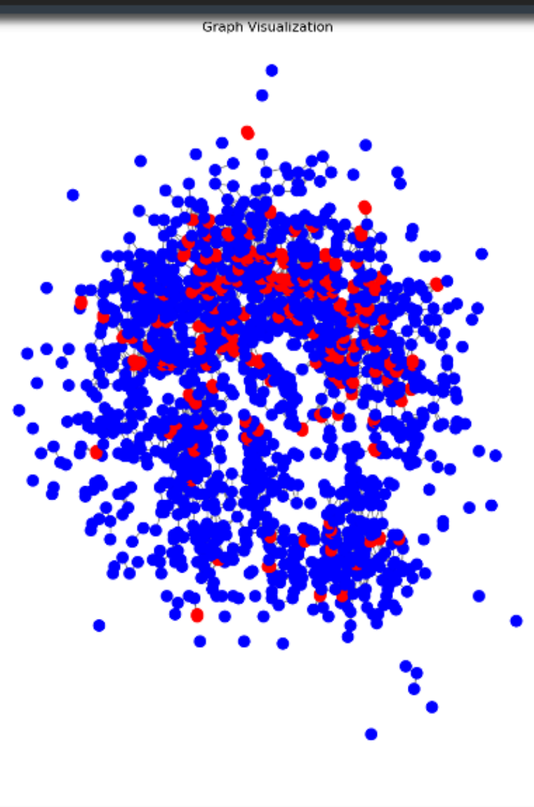

Graph Visualization

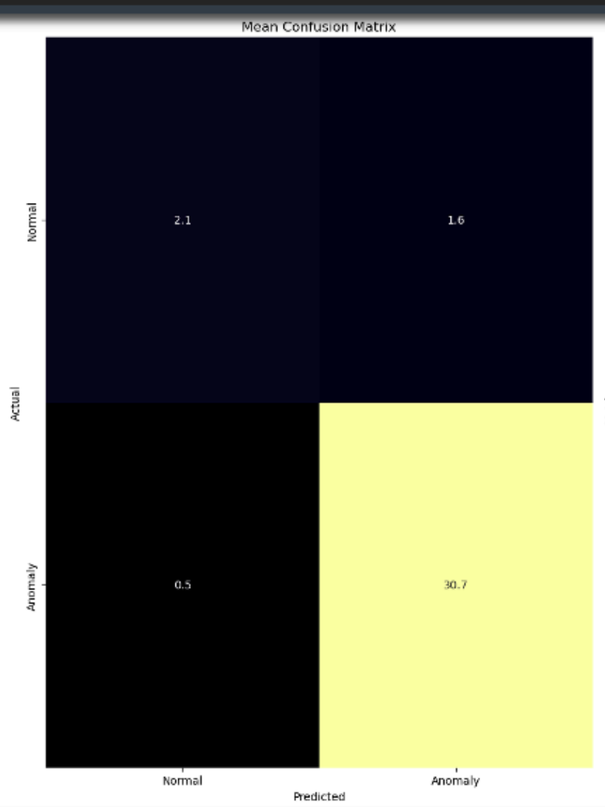

Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
1.9
29.4
Normal
Anomaly
Predicted

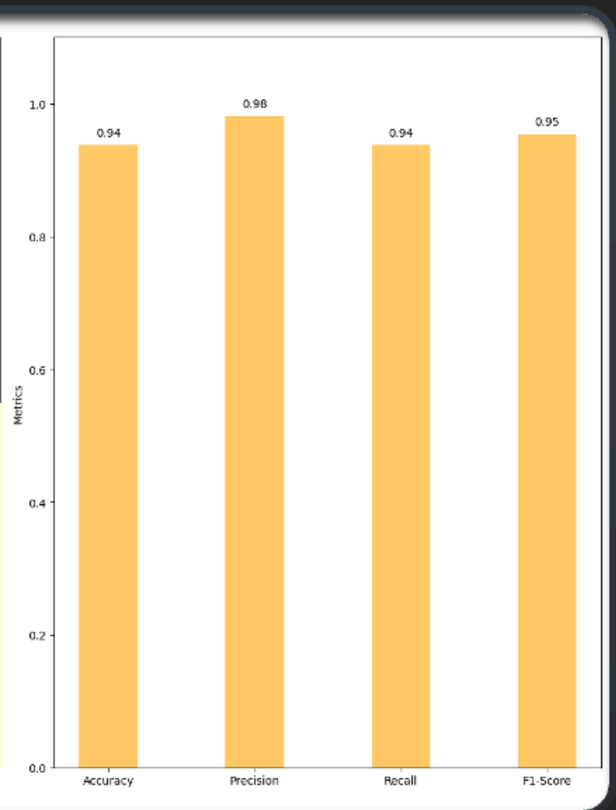

Metrics
1.0
0.8
0.6
0.4
0.2
0.0
0.90
0.98
0.90
0.93
Accuracy
Precision
Recall
F1-Score

# Results

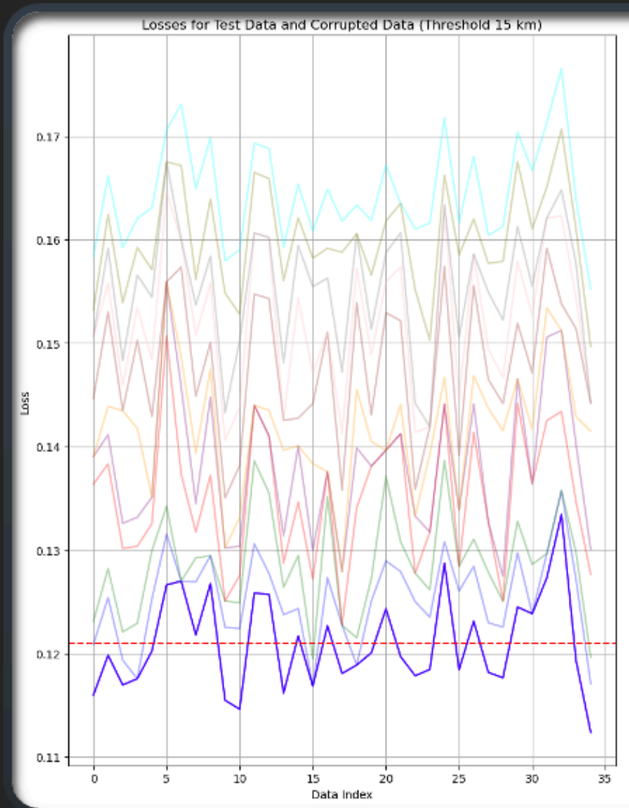

Losses for Test Data and Corrupted Data (Threshold 14 km)
Loss
Data Index
0.17
0.16
0.15
0.14
0.13
0.12
0.11
0
5
10
15
20
25
30
35

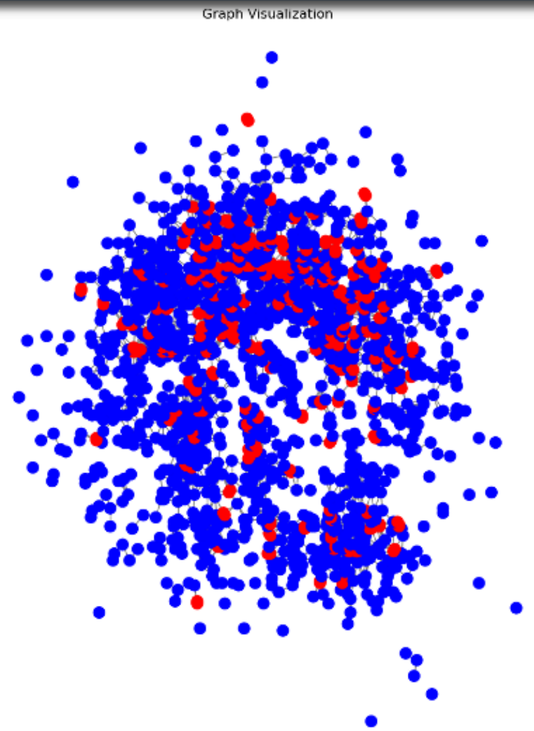

Graph Visualization

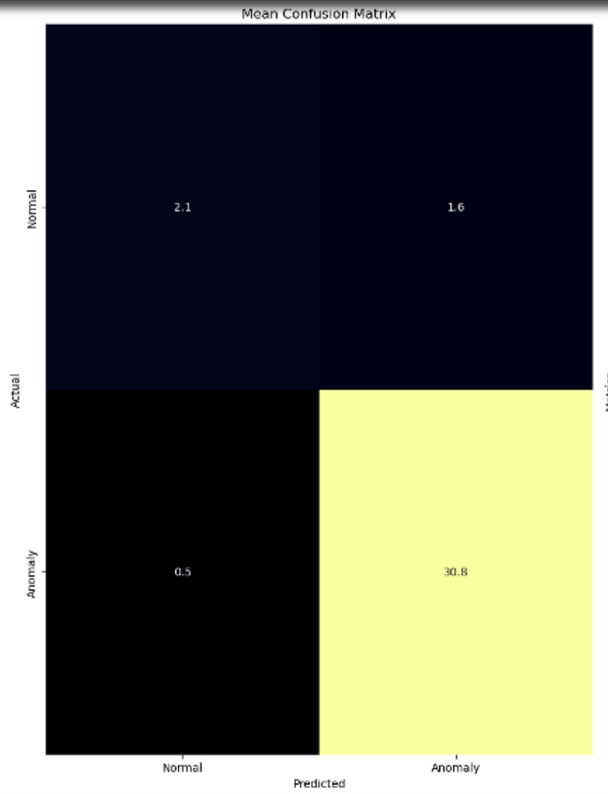

Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
0.5
30.7
Normal
Anomaly
Predicted

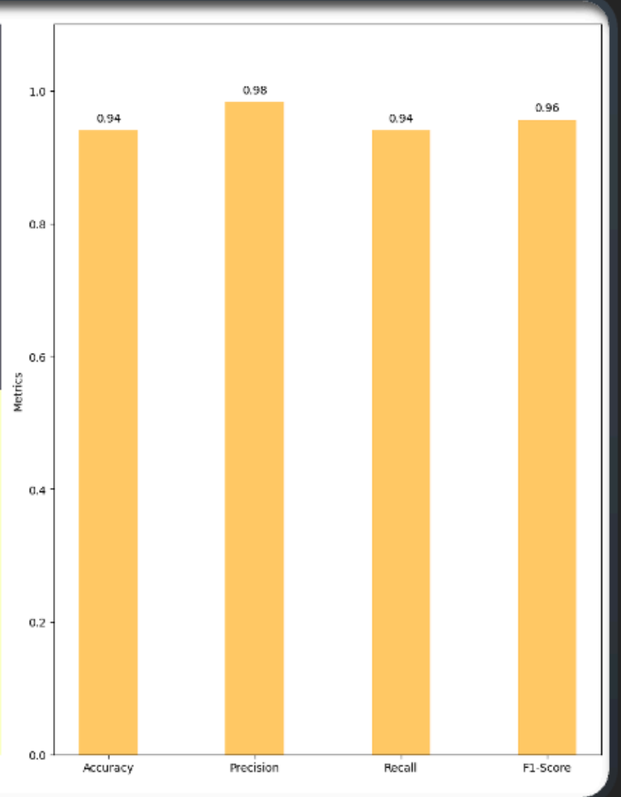

Metrics
1.0
0.8
0.6
0.4
0.2
0.0
0.94
0.98
0.94
0.95
Accuracy
Precision
Recall
F1-Score

# Results

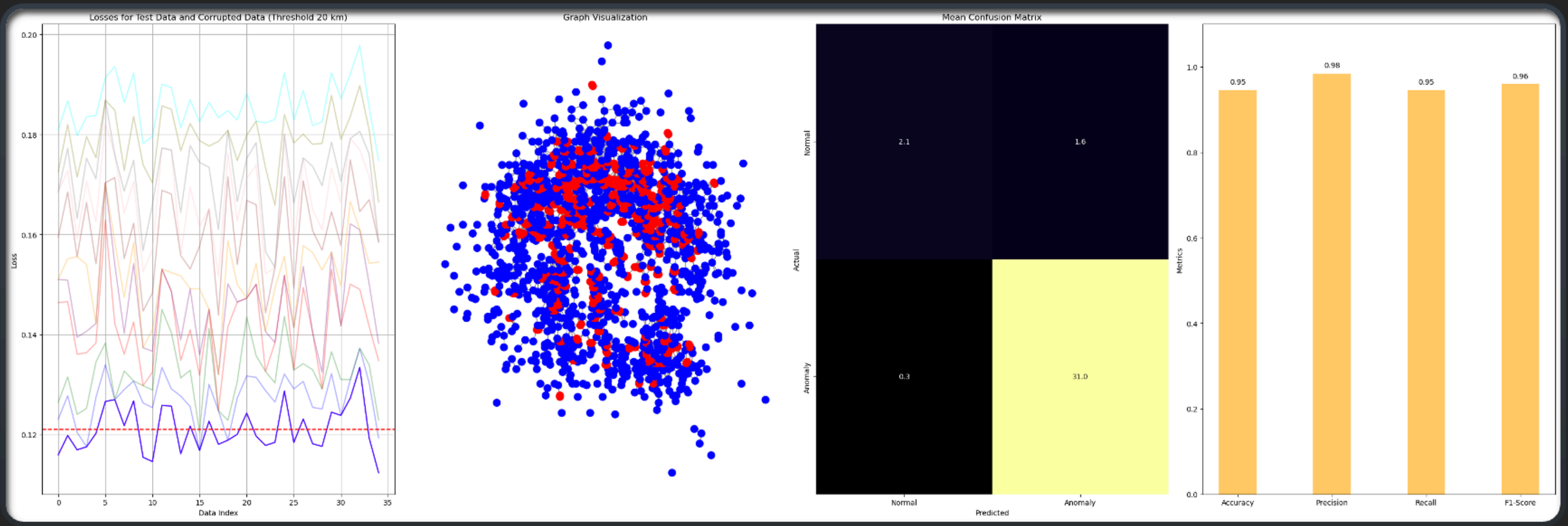

Losses for Test Data and Corrupted Data (Threshold 15 km)
Loss
Data Index
0.17
0.16
0.15
0.14
0.13
0.12
0.11
0
5
10
15
20
25
30
35


Graph Visualization

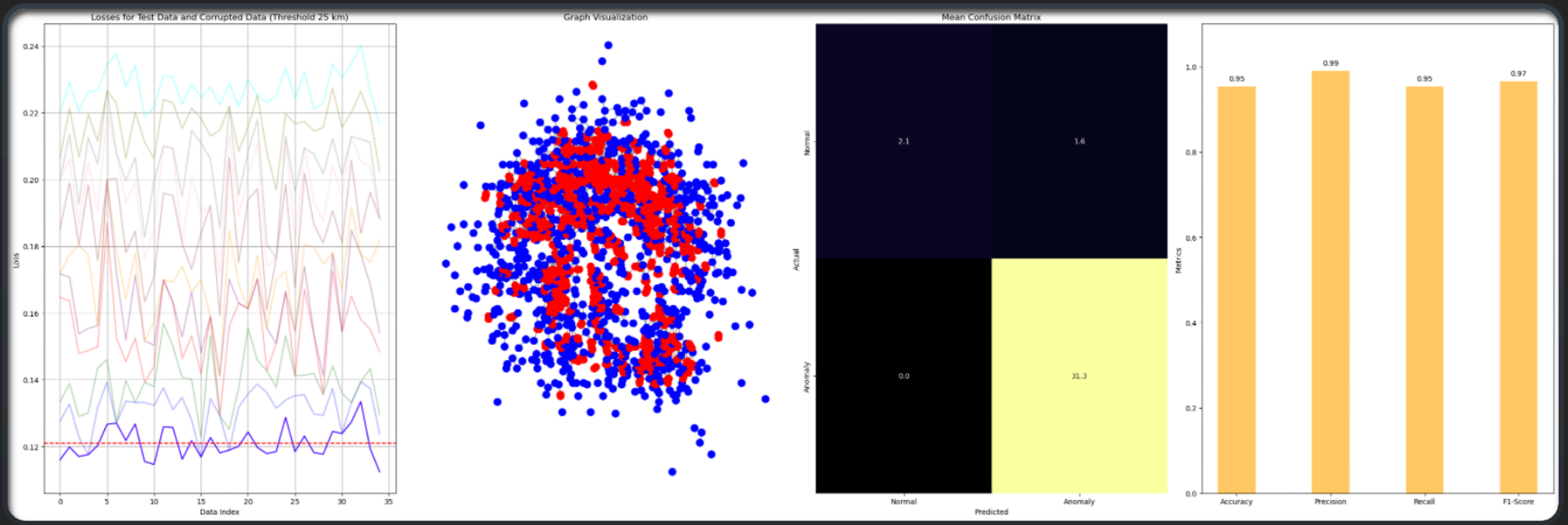

Mean Confusion Matrix
Actual
Normal
Anomaly
2.1
1.6
0.5
30.8
Normal
Anomaly
Predicted


Metrics
1.0
0.8
0.6
0.4
0.2
0.0
0.94
0.98
0.94
0.96
Accuracy
Precision
Recall
F1-Score

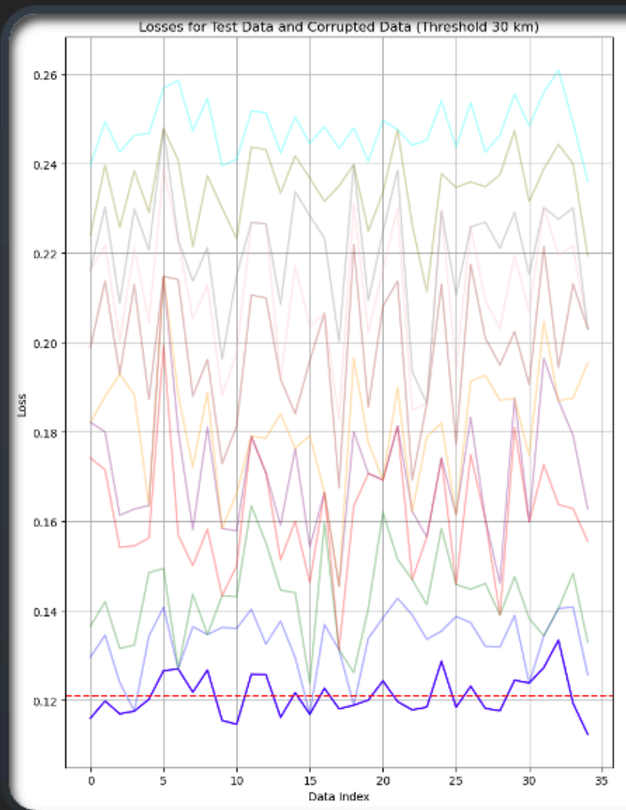
# Results

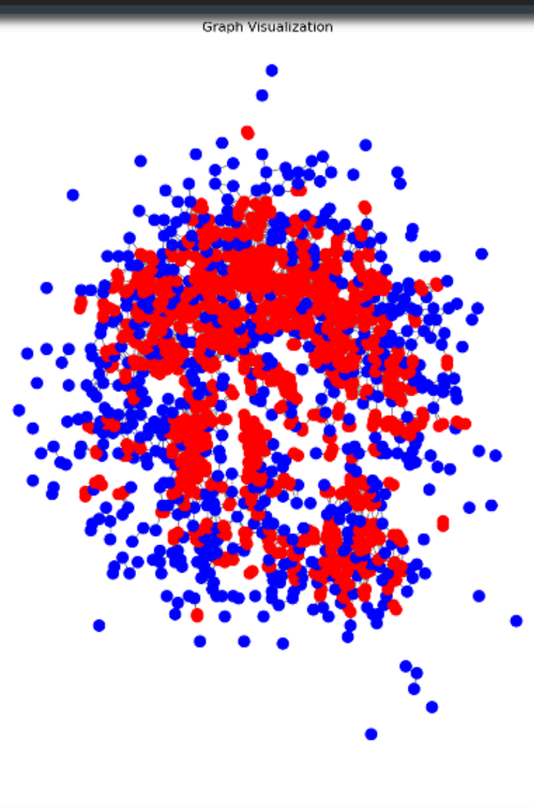

Losses for Test Data and Corrupted Data (Threshold 20 km)
Loss
Data Index
Graph Visualization
Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
0.3
31.0
Predicted
0.95
0.98
0.95
0.96
Metrics
Accuracy
Precision
Recall
F1-Score

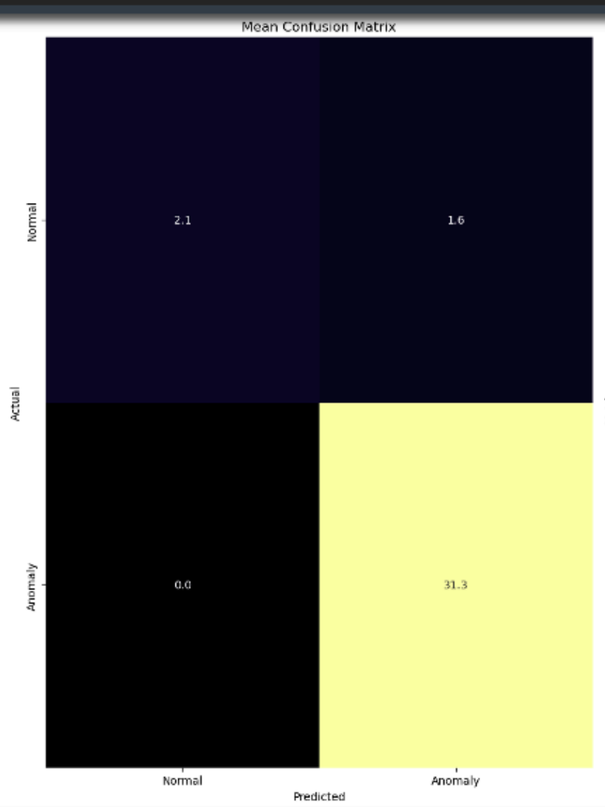
# Results

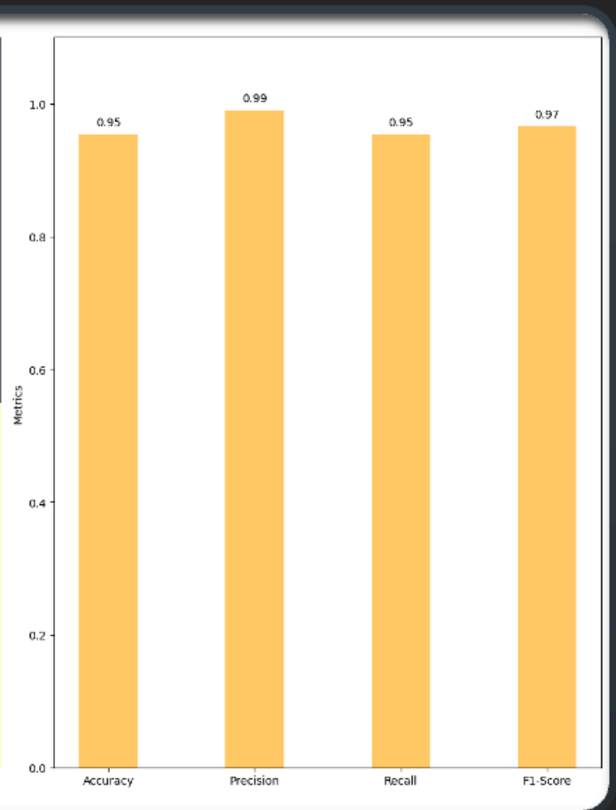

Losses for Test Data and Corrupted Data (Threshold 25 km)
Loss
Data Index
Graph Visualization
Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
0.0
31.3
Predicted
Normal
Anomaly
Metrics
0.95
0.99
0.95
0.97
Accuracy
Precision
Recall
F1-Score

# Results

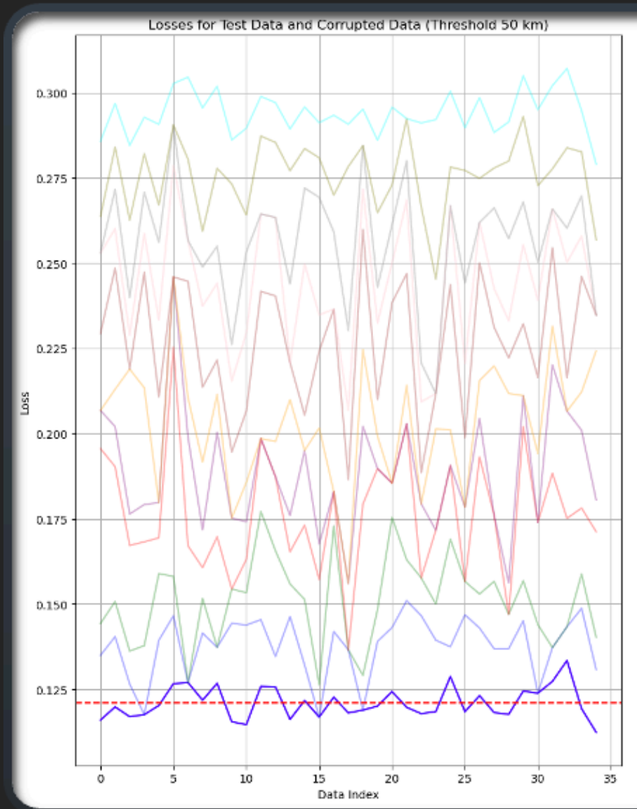

Losses for Test Data and Corrupted Data (Threshold 30 km)
Loss
Data Index
0.26
0.24
0.22
0.20
0.18
0.16
0.14
0.12
0
5
10
15
20
25
30
35

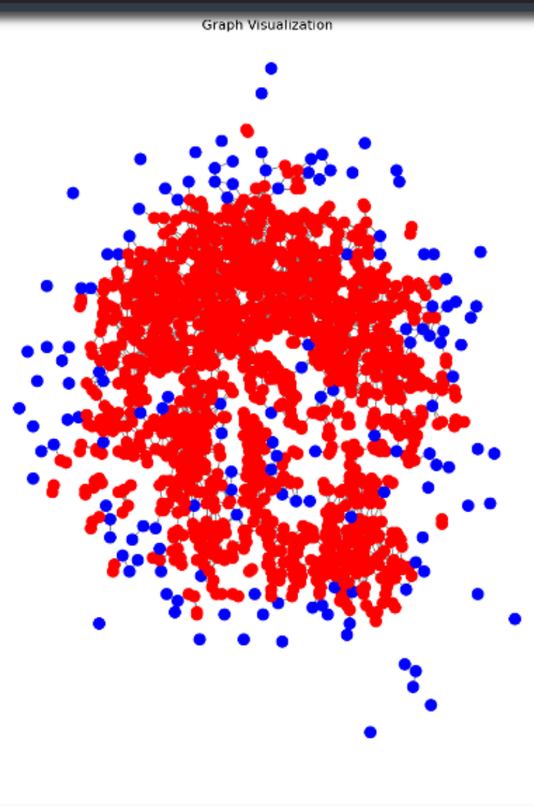

Graph Visualization

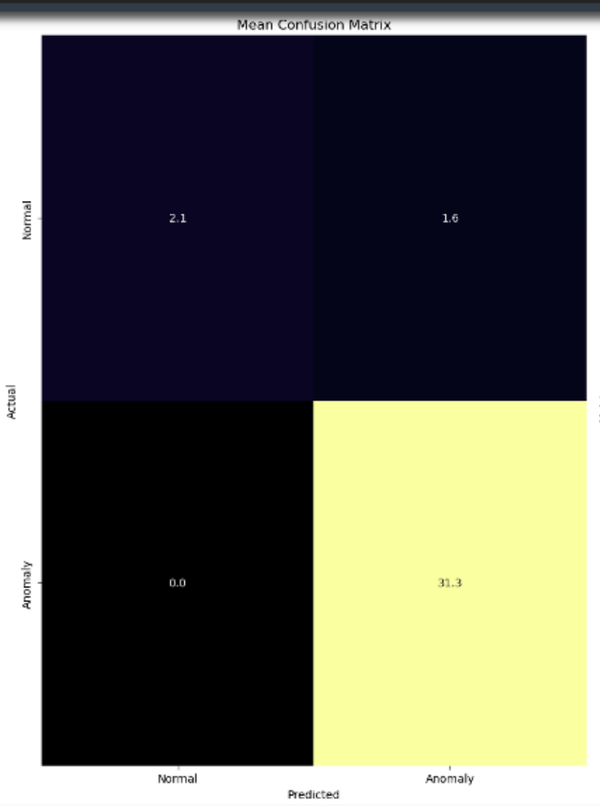

Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
0.0
31.3
Normal
Anomaly
Predicted

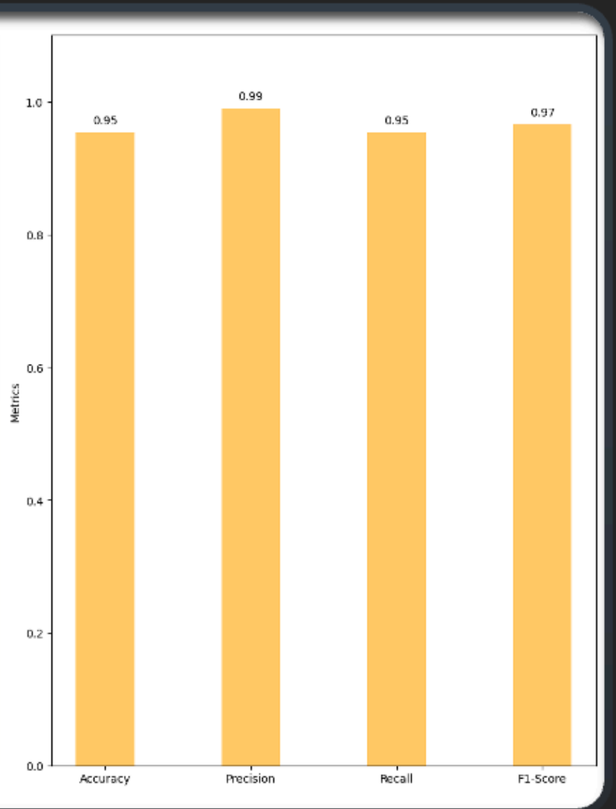

Metrics
1.0
0.8
0.6
0.4
0.2
0.0
0.95
0.99
0.95
0.97
Accuracy
Precision
Recall
F1-Score

# Results

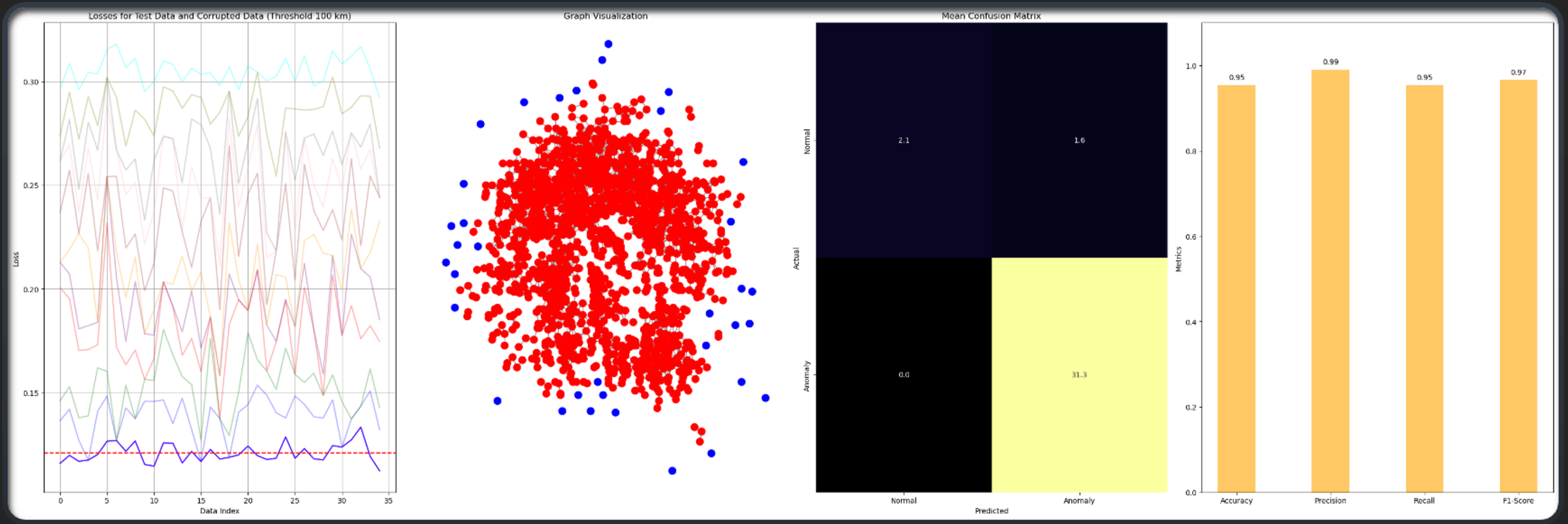

Losses for Test Data and Corrupted Data (Threshold 50 km)
Loss
Data Index
0.300
0.275
0.250
0.225
0.200
0.175
0.150
0.125
0
5
10
15
20
25
30
35


Graph Visualization


Mean Confusion Matrix
Actual
Predicted
Normal
Anomaly
2.1
1.6
0.0
31.3


Metrics
0.95
0.99
0.95
0.97
Accuracy
Precision
Recall
F1-Score
0.0
0.2
0.4
0.6
0.8
1.0

# Results


Losses for Test Data and Corrupted Data (Threshold 100 km)
Loss
Data Index
Graph Visualization
Mean Confusion Matrix
Normal
Anomaly
Actual
2.1
1.6
0.0
31.3
Predicted
Metrics
0.95
0.99
0.95
0.97
Accuracy
Precision
Recall
F1-Score

# Thank You for Attention

Roberto Giordano

roberto.giordano2@studio.unibo.it