NSL_KDD Dataset: Anomaly Detection Model Evaluation

Evaluation Protocol

Strict Train-Test Separation
No Data Leakage Confirmed | **Overfitting Analysis Below**

1. One-Class SVM

- Abandoned due to high compute cost + inferior performance vs. other models.

2. Isolation Forest (contamination=0.2)

Performance Metrics

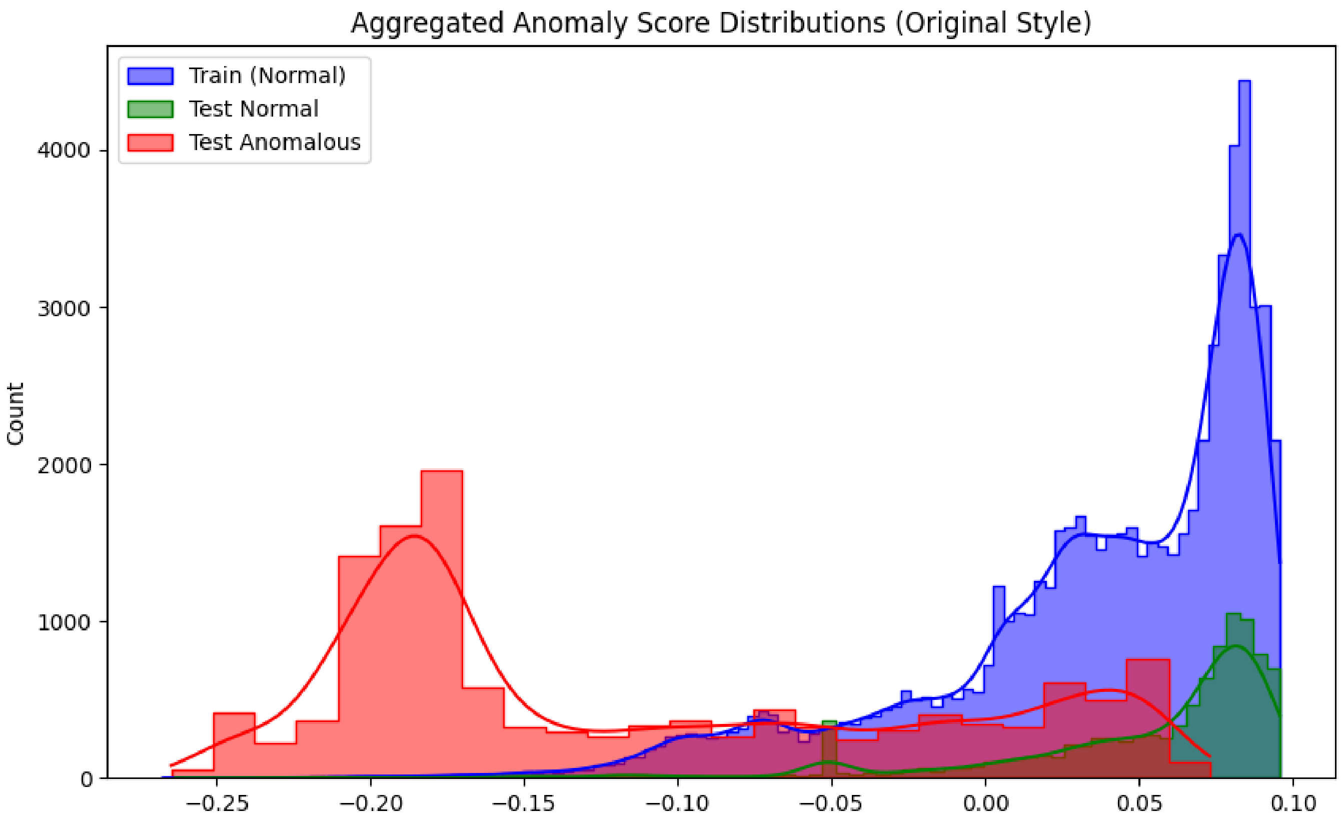
Test Set:

Class	Precision	Recall	F1-Score	Support
Normal (0)	0.78	0.88	0.83	9711
Anomaly (1)	0.90	0.81	0.85	12833
Accuracy: 84%	**Macro Avg F1**: 85%			

Training Set:

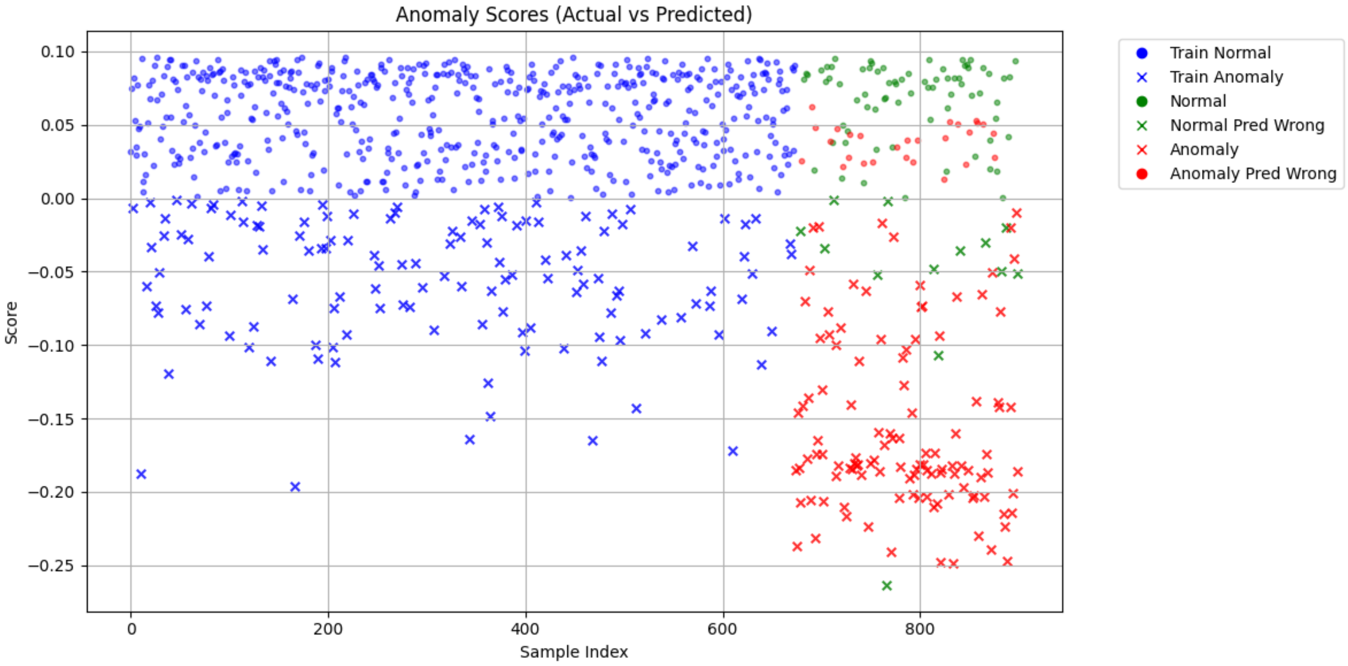
Metric	Value
Normal Recall	80%
Training Accuracy	80%

Visual Analysis



Color Code:

- Blue: Training normal scores
- Green: Test normal scores
- Red: Test anomaly scores



Markers:

- ○: Predicted normal (score > threshold)
- ✕: Predicted anomaly
  Color:
- Green: True normal

- Red: True anomaly
- Blue: Training normal

Analysis

- As Train accuracy is 80%, contaminaiton is 0.2 and Test accuracy is 84%, the model is not overfitting.
- From the visualization, we can see that the model is able to separate the normal and anomaly classes well upto a certain extent as the overlapping is not much.

3. Autoencoder (contamination=0.1)

Performance Metrics

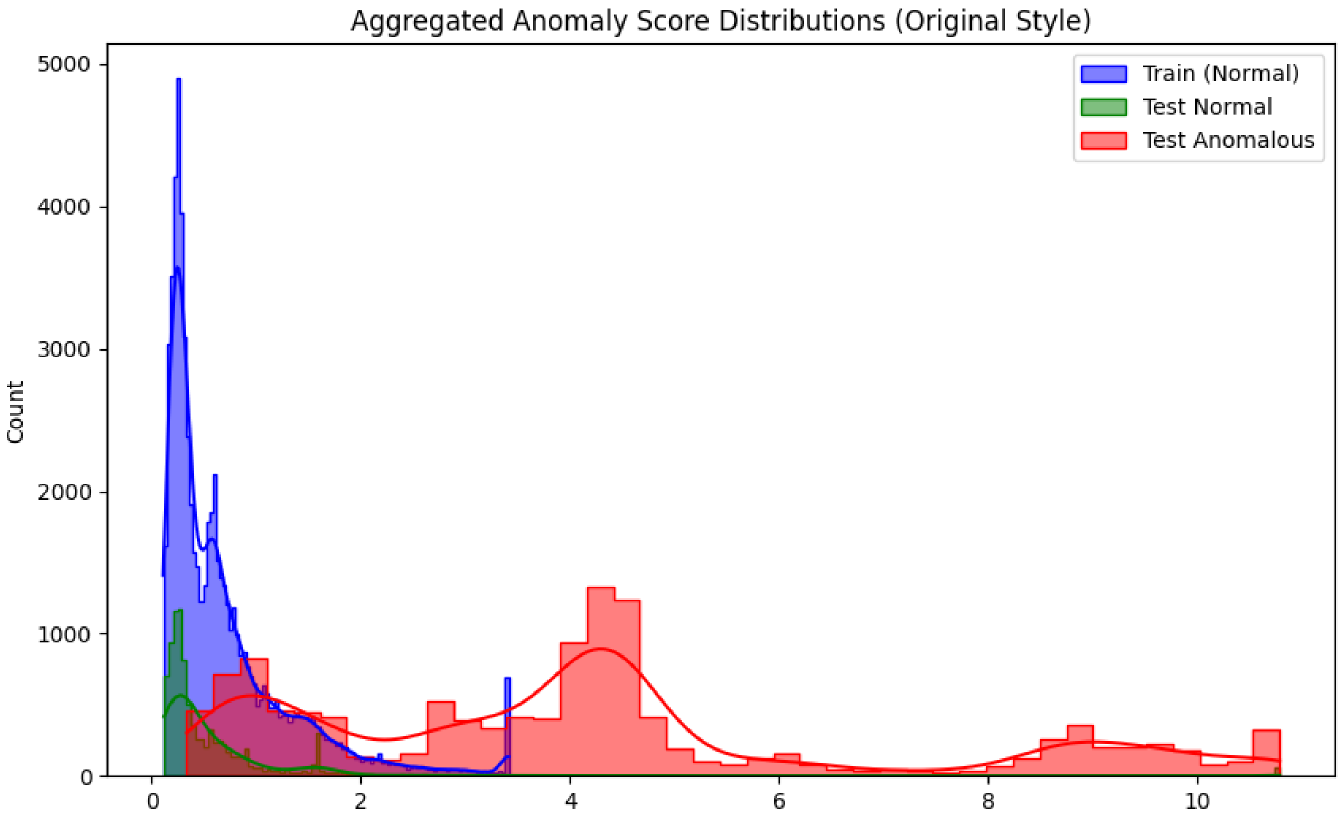
Test Set:

Class	Precision	Recall	F1-Score	Support
Normal (0)	0.76	0.91	0.82	9711
Anomaly (1)	0.92	0.78	0.84	12833
Accuracy: 83%	**Macro Avg F1**: 83%			

Training Set:

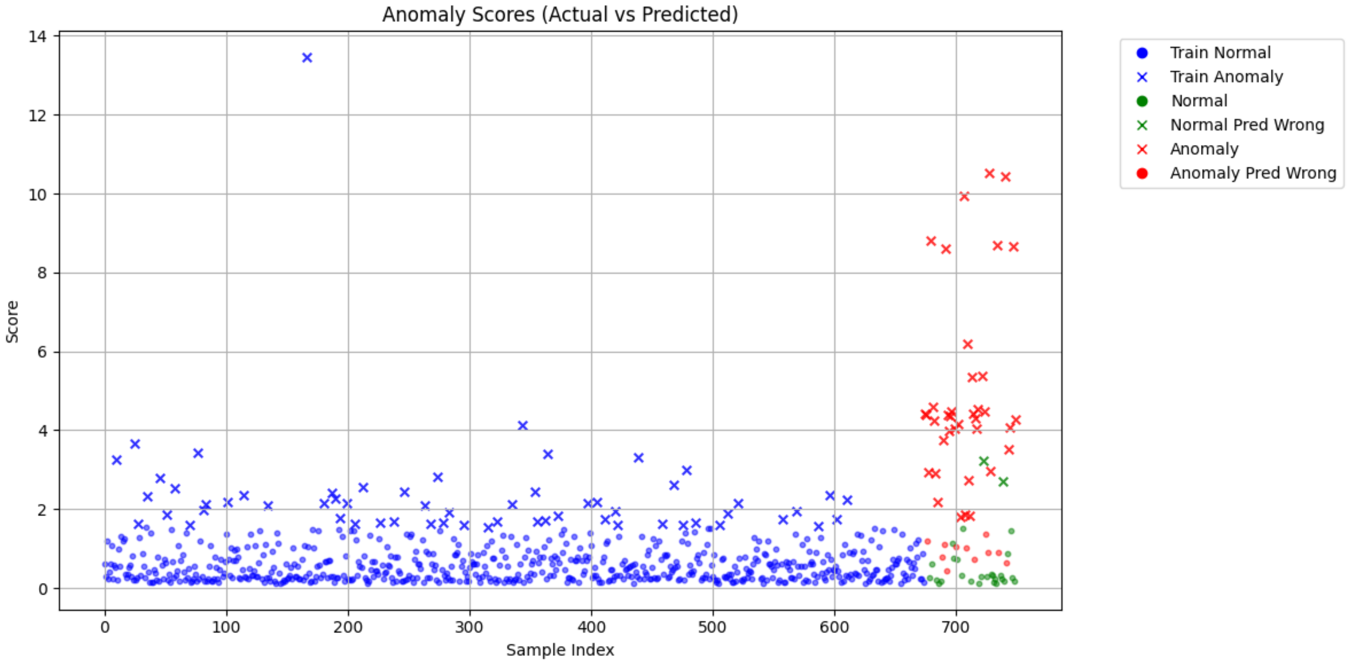
Metric	Value
Training Accuracy	90%

Visual Analysis



Color Code:

- Blue: Training normal scores
- Green: Test normal scores
- Red: Test anomaly scores



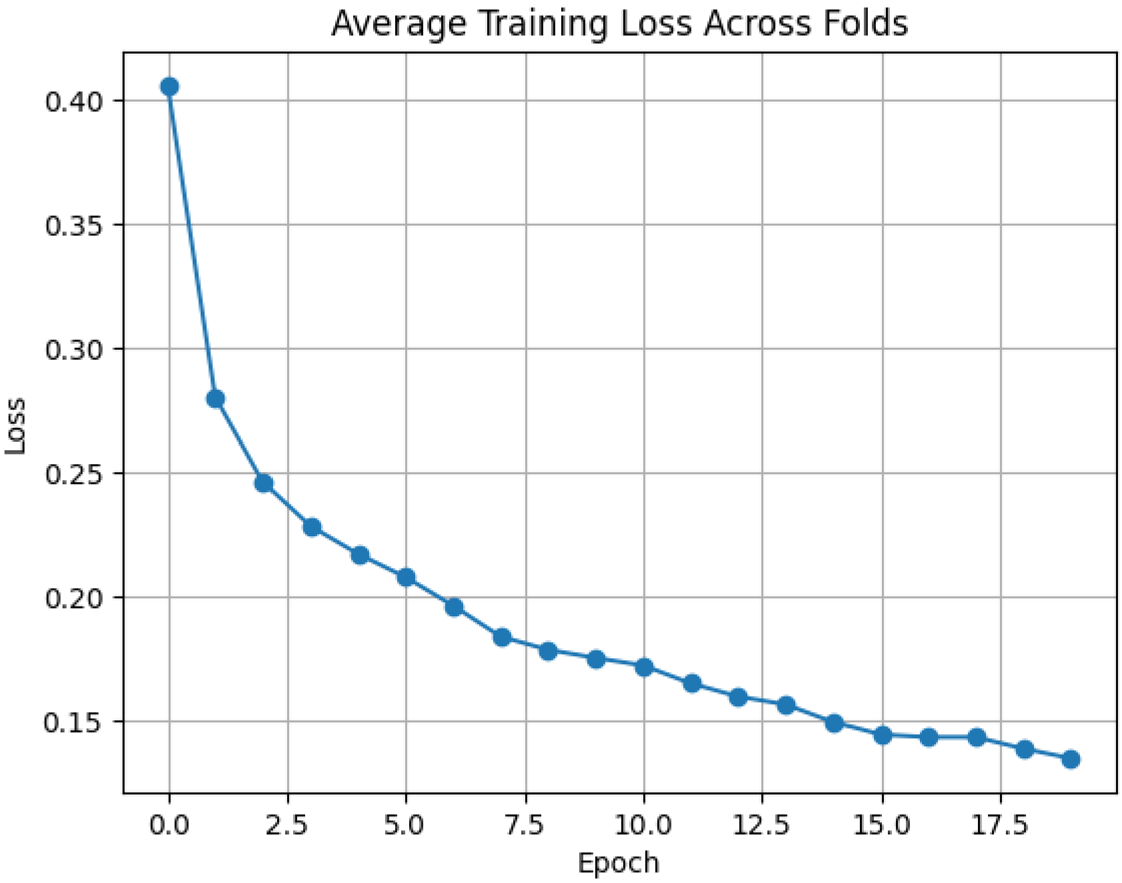
Markers:

- ○: Predicted normal (score > threshold)
- ✕: Predicted anomaly

  Color:
- Green: True normal

- Red: True anomaly
- Blue: Training normal

Training Loss



Analysis

- As Train accuracy is 80%, contaminaiton is 0.2 and Test accuracy is 84%, the model is not overfitting.
- From the visualization, we can see that the model is able to separate the normal and anomaly classes well upto a certain extent as the overlapping is not much.