IEEE CIS Dataset: Anomaly Detection Model Evaluation

Evaluation Protocol

10-Fold Cross Validation | Strict Train-Test Separation
No Data Leakage Confirmed | Overfitting Analysis Below | No patterns are getting detected by the models

---

1. One-Class SVM

The OneClass SVM models are not working well for higer dimensional data. Its taking lot of time nearly 3-4 hours to train them and they are just giving comparable or worse results than the Isolation Forest and AutoEncoder models. Which are just taking at max few minutes to train. So the study on OneClass SVM is not done thoroughly.

2. Isolation Forest (contamination=0.1)

Performance Metrics

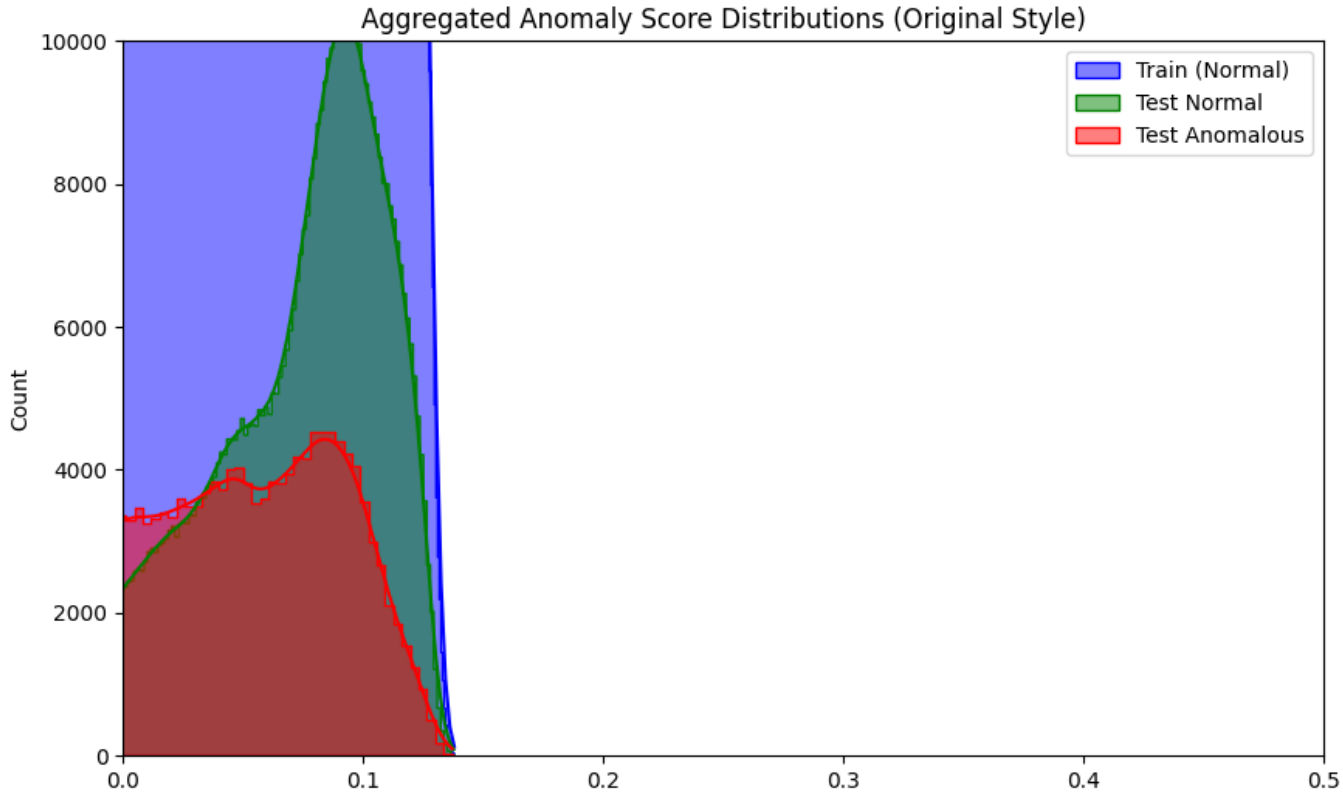
Test Set:

Class	Precision	Recall	F1-Score	Support
Normal (0)	0.80	0.90	0.85	569877
Anomaly (1)	0.57	0.38	0.45	206630
Accuracy: 76%	**Macro Avg F1**: 0.65			

Training Set:

Metric	Value
Normal Recall	90%
Training Accuracy	90%

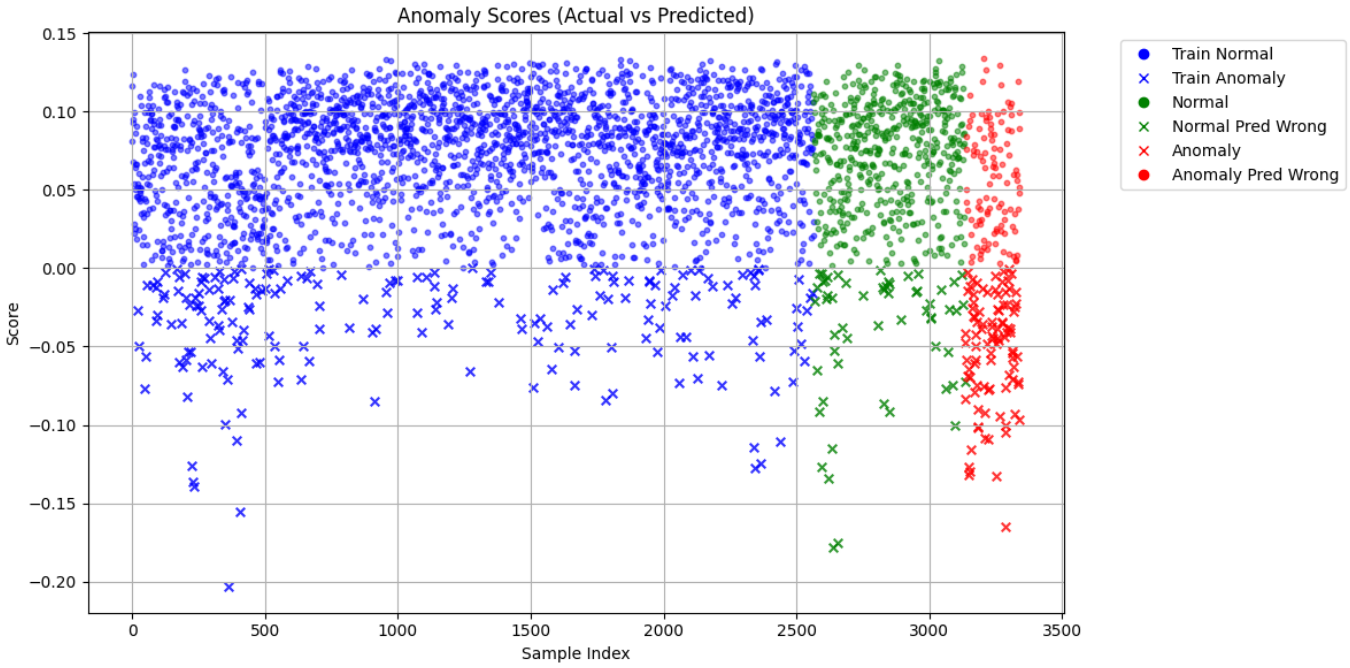
Visual Analysis



Color Code:

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

  Pattern: Overlapping distributions indicate limited separability

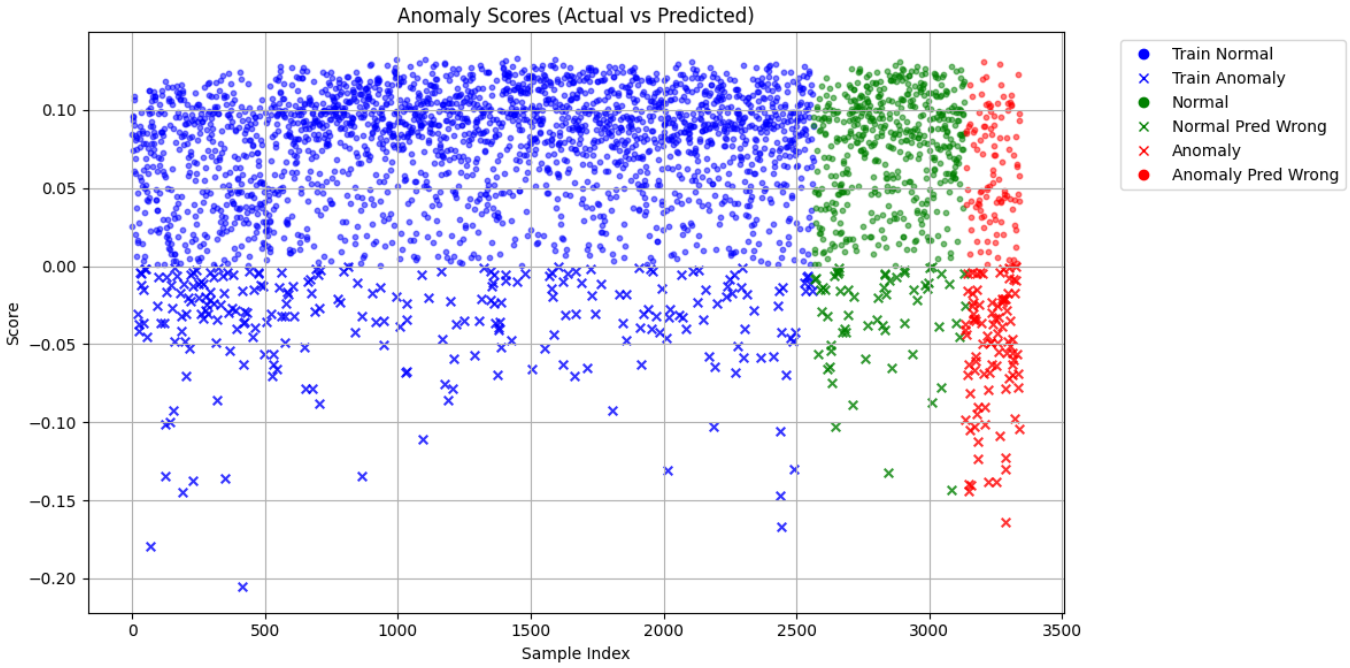


Markers:

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

  Color:
- Green: True normal

- Red: True anomaly
- Blue: Training normal



Key Insight: No clear separation between classes in latent space.

---

3. Autoencoder (contamination=0.1)

Performance Metrics

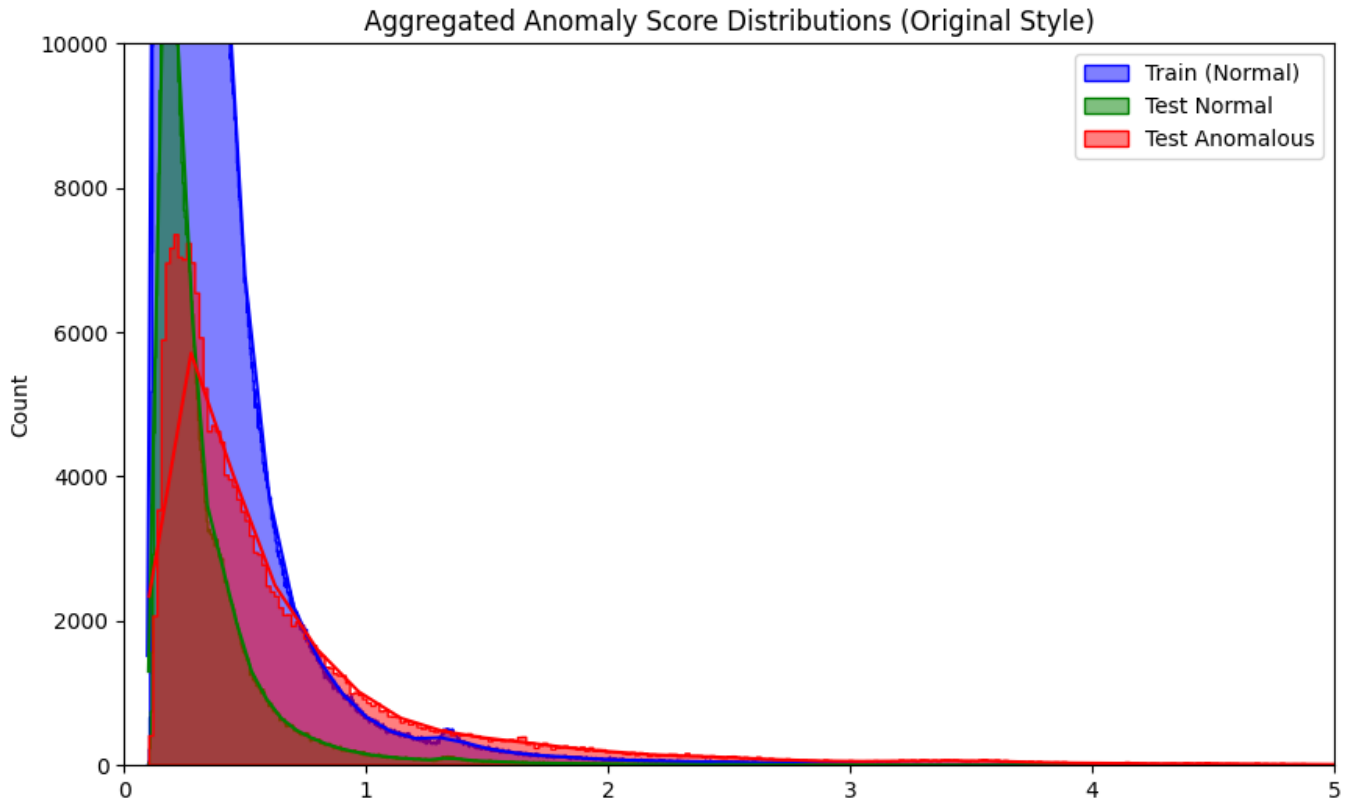
Test Set:

Class	Precision	Recall	F1-Score	Support
Normal (0)	0.80	0.90	0.85	569877
Anomaly (1)	0.57	0.37	0.46	206630
Accuracy: 76%	**Macro Avg F1**: 0.65			

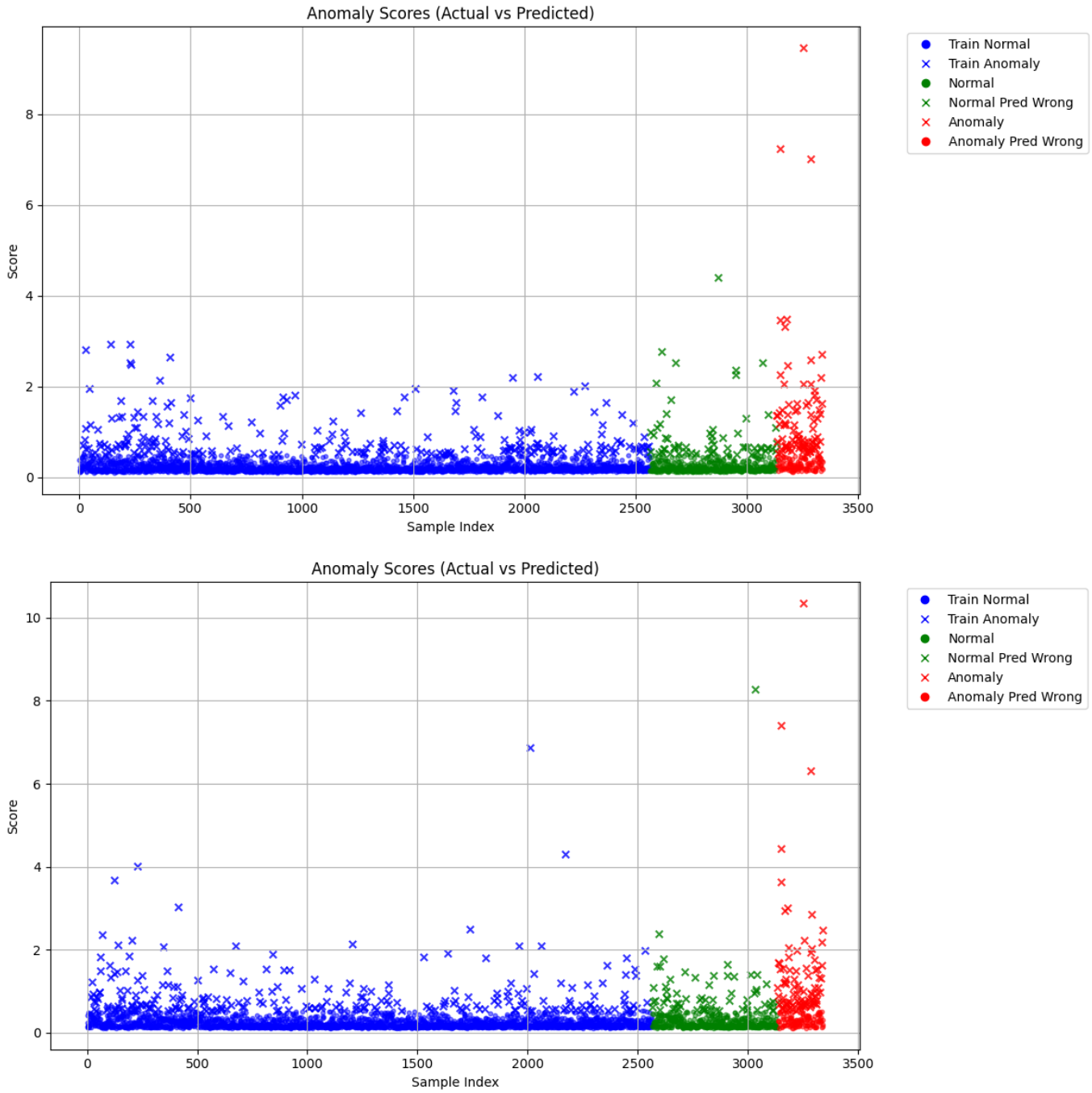
Training Set:

Metric	Value
Reconstruction Error	0 to 7
Training Accuracy	90%

Visual Analysis



Critical Overlap: Most of the samples in test normal and test anomaly overlap in the score distribution.



Key Insight: No clear cluster separation despite reconstruction errors.

Final Assessment

"All models demonstrate limited discriminative power due to fundamental feature relationships in the SECOM dataset. While achieving expected contamination-aligned accuracy, true anomaly detection capability remains constrained by overlapping score distributions. Priority should be given to feature-space transformations before model optimization."