# Comparative Report: QAE vs QSVM vs QNN on IEEE-CIS, SECOM, and NSL-KDD Datasets

---

## Introduction

Quantum Machine Learning (QML) is emerging as a powerful approach to learning from complex, high-dimensional data by leveraging quantum information processing. In this study, three different QML architectures are evaluated:

- **QAE** – Quantum Autoencoder for representation learning and compression
- **QSVM** – Quantum-enhanced Support Vector Machine using quantum kernels
- **QNN** – Quantum Neural Network leveraging variational quantum circuits

Each model is tested on three real-world datasets representing distinct application domains:

- **IEEE-CIS** – Fraud detection in online transactions
- **SECOM** – Quality prediction in semiconductor manufacturing
- **NSL-KDD** – Intrusion detection in network traffic data

The goal is to understand how different quantum models perform on varying data structures and learning tasks.

---

## Dataset Overview

### 1. IEEE-CIS Fraud Detection

- **Domain**: Finance / Cybersecurity
- **Objective**: Binary classification of transactions as fraudulent or legitimate
- **Samples used**: 1,000 to 10,000
- **Features**: Transaction amount, card data, derived numerical stats
- **Challenge**: Imbalanced dataset with subtle patterns

### 2. SECOM Manufacturing

- **Domain**: Industrial Engineering
- **Objective**: Predict pass/fail status based on 590+ sensor readings
- **Samples used**: 1,536 (filtered for completeness)
- **Features**: Sensor voltages, flows, and temporal signals
- **Challenge**: Noisy, high-dimensional data with weak signals

### 3. NSL-KDD Intrusion Detection

- **Domain**: Network Security
- **Objective**: Classify connections as normal or an attack
- **Samples used**: 100–10,000
- **Features**: Rates of errors, host metrics, TCP-level flags

- **Challenge**: Structured, well-separated classes

---

# Quantum Model Summary

| Model | Core Idea | Strengths | Typical Usage |
|---|---|---|---|
| **QAE** | Compresses input data to latent space and reconstructs | Captures structure, useful for dimensionality reduction | Preprocessing, anomaly detection |
| **QSVM** | Replaces kernel function with quantum feature space | Good at non-linear separations | Binary classification |
| **QNN** | Hybrid model combining classical + quantum layers | Flexible, expressive | Classification, multi-task learning |

---

# Performance Comparison

## 1. IEEE-CIS (Fraud Detection)

| Model | Accuracy | Notes |
|---|---|---|
| QAE (Latent SVM) | 83.50% | Learned meaningful compressed representations |
| QSVM | **96.00%** | Best performing; quantum kernel effective in detecting fraud patterns |
| QNN | **96.49%** | Strong performance, nearly matching QSVM |

## 2. SECOM (Manufacturing Quality)

| Model | Accuracy | Notes |
|---|---|---|
| QAE (Latent SVM) | 82.00% | Effective reconstruction and separation |
| QSVM | 94.00% | Well-suited for noisy industrial data |
| QNN | **93.51%** | Robust and stable over training, competitive with QSVM |

## 3. NSL-KDD (Intrusion Detection)

| Model | Accuracy | Notes |
|---|---|---|
| QAE (Latent SVM) | 80.00% | Performed reasonably despite small latent space |
| QSVM | **87.50%** | Outperformed other quantum models on structured data |
| QNN | 74.27% | Struggled with rigid patterns, lower generalization |

---

# Comparative Insights

## Representation & Learning Ability

- **QAE** excels in uncovering low-dimensional structures, particularly useful in unsupervised settings or when interpretability of latent features is crucial.
- **QSVM** stands out in binary classification tasks, especially on complex, noisy, or imbalanced datasets. It is especially effective when the class boundary is non-linear.
- **QNN** offers a flexible architecture capable of adapting to a wide range of data types. Its performance is close to or better than QSVM in many scenarios, though it may require more fine-tuning.

### Computational Considerations

- **QAE** is the most resource-intensive due to the encoding/decoding of full samples and repeated quantum circuit evaluations.
- **QSVM** is relatively efficient with fewer parameters and faster convergence on small datasets.
- **QNN** benefits from batch training via PyTorch and integrates seamlessly with existing ML pipelines, though training is slower due to variational quantum circuit evaluations per sample.

### Strengths vs. Weaknesses

| Model | Strengths | Weaknesses |
|---|---|---|
| QAE | Excellent for compression, captures structure | Reconstruction-focused, not optimized for classification |
| QSVM | Best classification accuracy, robust to noise | Less scalable to multi-class problems |
| QNN | Flexible, end-to-end, adaptive | Training complexity, performance sensitive to architecture choices |

## Conclusion

Each quantum model exhibits distinct strengths:

- **QSVM** is the top performer for **classification**, showing superior results in fraud detection and network intrusion tasks.
- **QAE** is most suitable for **data compression and feature extraction**, useful when interpretability of data representations is important.
- **QNN** is an excellent **general-purpose model**, delivering strong results in diverse domains, especially with deeper learning capacity.

In conclusion, the choice of quantum model should be guided by the task:

- **QSVM** for high-accuracy classification
- **QAE** for unsupervised feature learning
- **QNN** for hybrid applications requiring flexibility