SMU SINGAPORE MANAGEMENT UNIVERSITY
School of Computing and Information Systems

Exploring Transaction Anomalies Using Quantum Siamese Neural Networks

Jiawei Yuan
Supervisor: Prof. Paul Griffin | Advisor: Nikolaos Schetakis

Motivation & Objectives

Transaction anomaly is still a problem today. Especially for trade finance, profit and loss can depend on it. To prove the concept that quantum computing would assist on transaction anomaly detection, this project studies on the potential of quantum circuits operated on a Siamese Neural Network (SNN) architecture detecting the anomalies in transaction liked data.

The objective of this research is to begin to explore how quantum machine learning can improve anomaly detection for finance transactions in trade finance.

Data & Pre-processing

Data Source

The data used for the project are from well-known datasets to enable characterization and and benchmarking. Selected datasets are from Outlier Detection Datasets (ODDS). 3 specific point anomaly datasets that have less than 20 attributes are selected, since more than 20 qubits run too slowly on quantum simulators.

Dataset	# of Instance	# of Attributes	% of Outliers
Glass	214	9	4.2%
Breast Cancer	683	9	35%
Lymphography	148	18	4.1%

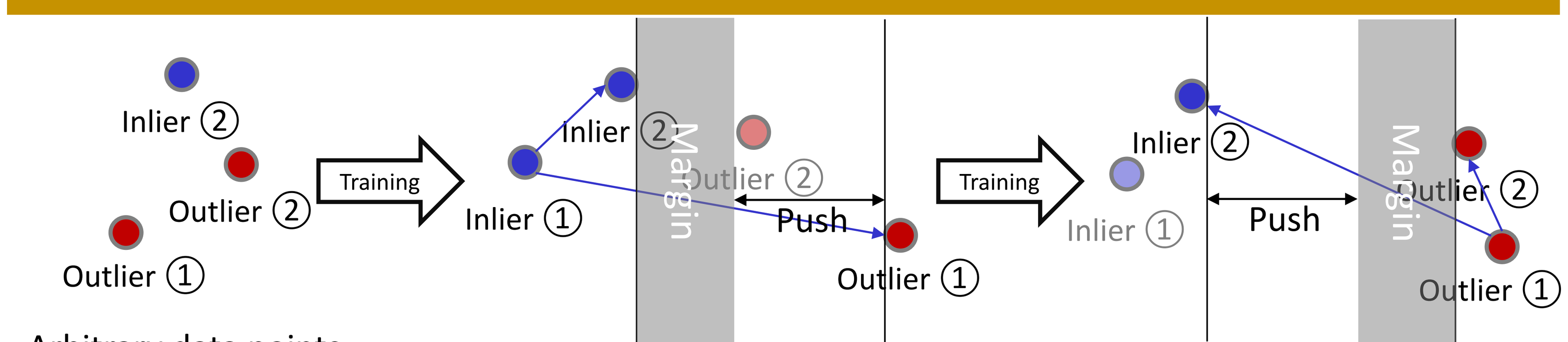
$$Zscore_{xi} = \frac{x_i - \mu}{\sigma}$$

Process to Z-score

None of the 3 dataset's specific information on its attributes is descripted, so the data are processed as numbers without special meaning to their Z-score.

Model Design Methods

Triplet Loss Function



Arbitrary data points before training
Inlier-anchor triplet training
Outlier-anchor triplet training

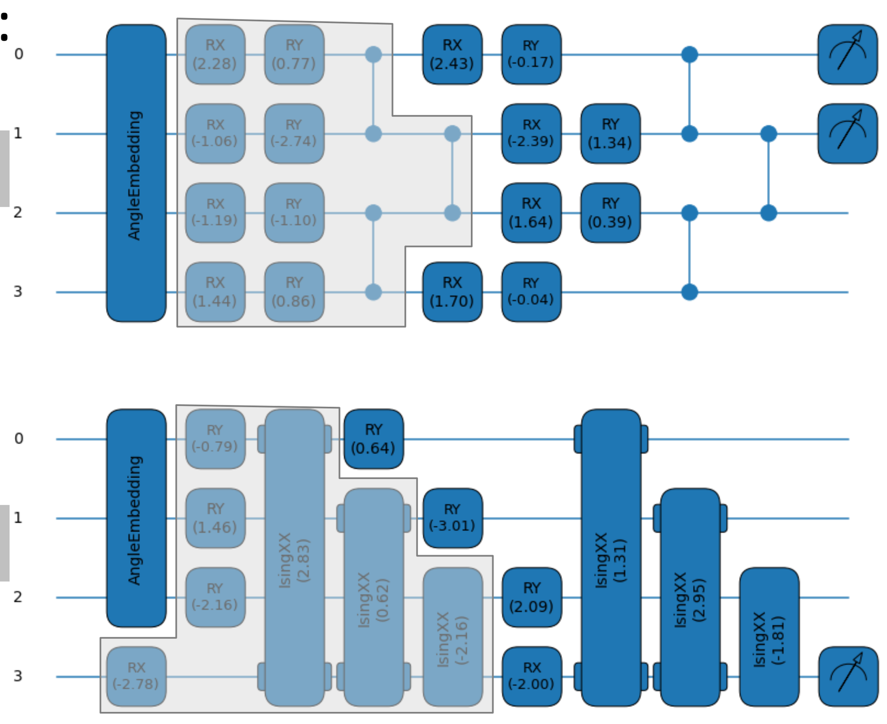
$$L = \max\left(d_{(a,p)} - d_{(a,n)} + margin, 0\right)$$

Triplet loss is chosen to be the loss function due to its performance on classical neural networks. The hope is that triplet loss fits well in the Siamese structure and generates similar outputs in proximity to other feasible functions

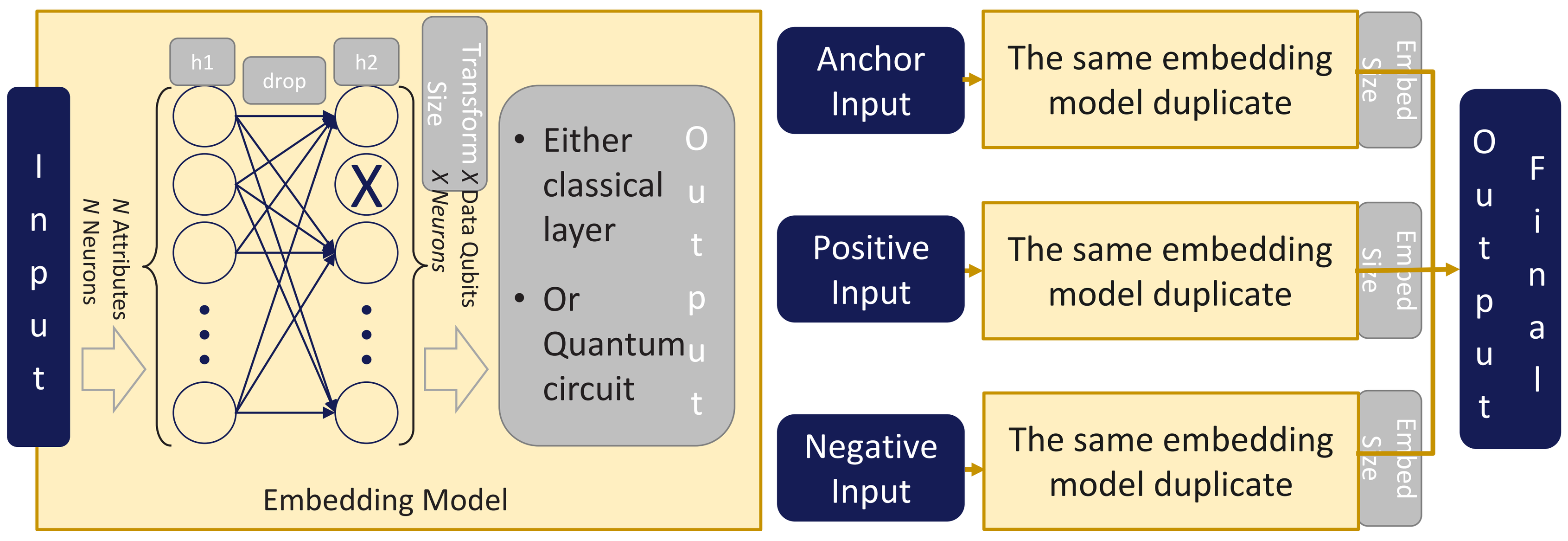
Quantum Circuits

Two types of circuits are tested:

1. Angle embedding all n qubits. Each variational layer has nRX, nRY, and n-1 entangling CZ. Z or ZZ can be measurement.
2. Angle embedding n-1 qubits. Each variational layer has n-1 RY, 1 RX and n-1 RXX. Z measurement.



SNN Model Building



A SNN model (right) consists of 3 duplicates of embedding models (left). The architecture is built for triplet batch data.

Hyperparameters Configuration & Selection

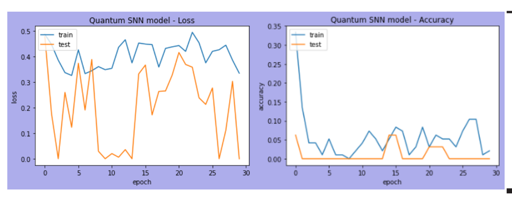
Hyperparameters	Configuration Values	Selection Results
Adam learning rate	$[10^{-4}, 0.5]$ (log)	Main hyperparameter
Batch size	16, 32, 64	Less effective
Depth	{1, 2, ..., 5}	Main hyperparameter
Alpha	{0.1, 0.2, ..., 1.0}	Less effective
Transform size	{3, 5, 9}	Fixed to 3
Embed size	{1, 2, 3}	Fixed to 1
Reuploading	True, False	Fixed to False
Circuit type	Triplet loss type (1.), Classification type (2.)	Fixed to Classification type (2.)
Output circuit	Z, ZZ	Fixed to Z
Use simulator	True, False	Fixed to True

Qubit number limits transform size, embedding size and output circuit. TL circuit or reuploading are by actual results

Performance Benchmarking & Insights

Dataset	Quantum – Variational Layer 1	2	3	4	5	Classical
Glass						
Lr=0.0005	.51(NC)#	.82	.83	.22(NC)#	.07(NC)	.86
Lr=0.005	.03(NC)	.91	.94	.85	.03(NC)	.98
Lr=0.05	.90	.19(NC)#	.85	**.95***	.01(NC)	**.99***
Breast Cancer						
Lr=0.0005	.03(NC)	.80	.78	.02(NC)	.05(NC)	.81
Lr=0.005	.79	**.83***	.81	.03(NC)	.02(NC)	.88
Lr=0.05	.81	.82	.02(NC)	.03(NC)	.77	**.93***
Lymphography						
Lr=0.0005	.57	.83	.85	.56(NC)#	.13(NC)#	.22(NC)#
Lr=0.005	.90	**.99***	.91	.91	.00(NC)	.13(NC)#
Lr=0.05	.68	.05(NC)	.02(NC)	.89	.31(NC)#	**.97***

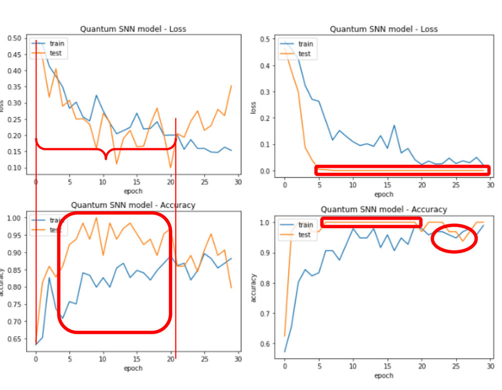
Models(NC) show poor(>10%) score, and their loss has no fixed pattern. These 14 seem to learn nothing.



Models(NC)# mostly show 10-30% score and their loss fluctuates, but there are 2 unique examples whose accuracy is more than 50% and loss is descending. It is hard to decide if it converged at 50.

Rest of the Models all converge within 30 epoch, while **these ones with *** perform the best score of the datasets.

The two examples show an intriguing "foresee" that performs 5-10% better on test sets than on training, while classical SNNs run on same Lr and data behave normally. This may attribute to unknown potential of QSNN.



Conclusion & Future Works

QSNN is a feasible architecture for detecting anomalies, learning rate and the variational layers are crucial to related potential.

Future works: Test models on real transaction data and real quantum device. Refine triplet design. Utilize confusion matrix.