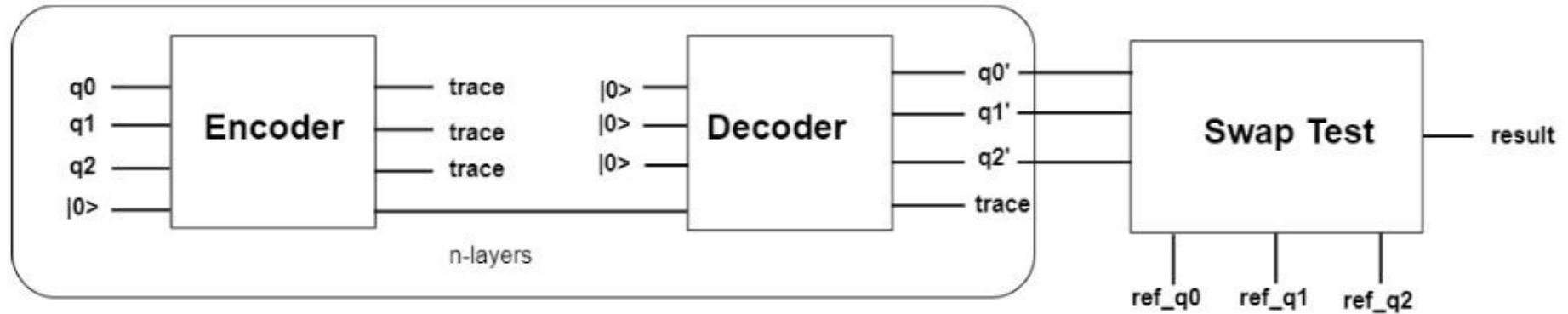


QUANTUM AUTOENCODER STRUCTURE AND TESTS

Content

- Three Qubit QEC General Structure – Page 3
- Implementation – Page 4
- Ansatz 1 – Page 5
- Ansatz 2 – Page 6
- Setup 1 – Page 7
- Results for Setup 1 – Page 9
- Setup 2 – Page 18
- Results for Setup 2 – Page 19
- Setup 3 – Page 24
- Results for Setup 3 – Page 25
- Setup 4 – Page 38
- Results for Setup 4 – Page 39
- Interpretations – Page 44

Three Qubit QEC General Structure

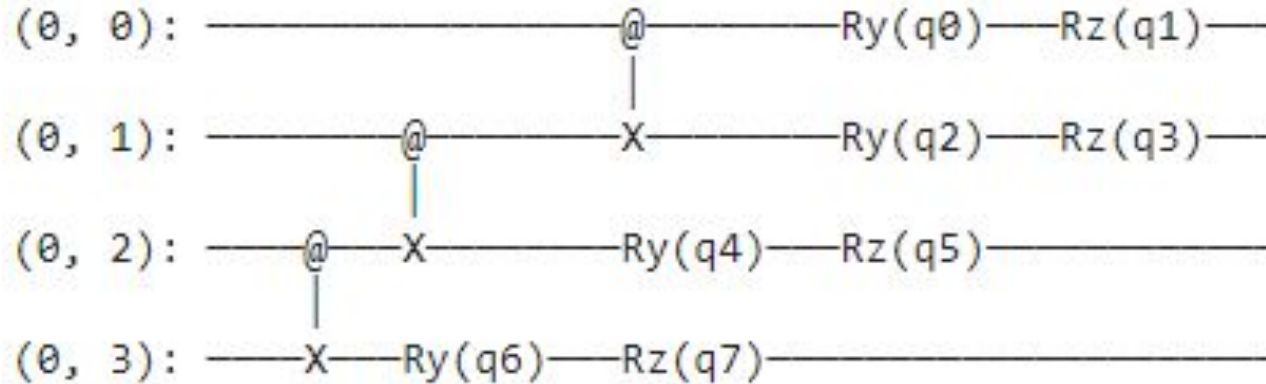


Implementation

- Based from Owen Lockwood's Quantum Autoencoder implementation:
https://github.com/lockwo/quantum_computation/tree/master/TFQ/QAE
<https://www.youtube.com/watch?v=ju4B0t25Ky8>
- Tried different ansatz, number of layers, optimizers, epochs, batch sizes and, training sets.
- Loss function -> Mean Absolute Percentage Error
- Created four different setups and used two different ansatz for our tests.

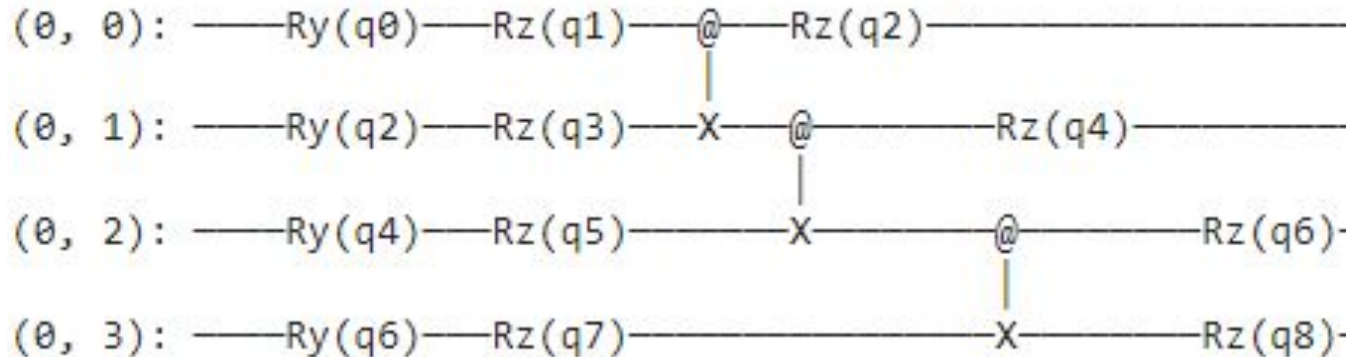
Ansatz 1

- This is the ansatz that Owen Lockwood used



Ansatz 2

- This is the ansatz used by the in the Noise Assisted Quantum Autoencoder paper (Cao, C., & Wang, X. (2021). Noise-assisted quantum autoencoder. Physical Review Applied, 15(5), 054012.).



Setup 1

- Ansatz 1 is used.
- Training and test is done with $|0\rangle$, $|1\rangle$ and $|+\rangle$ (just one state) states for three qubits.
 - $|0\rangle$: Errornous and legit statuses
 - **Legit:** $|000\rangle$
 - **Errornous:** $|001\rangle$, $|010\rangle$, $|100\rangle$
 - $|1\rangle$: Errornous and legit statuses
 - **Legit:** $|111\rangle$
 - **Errornous:** $|110\rangle$, $|101\rangle$, $|011\rangle$
 - $|+\rangle$: Just the legit status
 - $(|000\rangle + |111\rangle)/\text{sqrt}(2)$

- Training and test outputs are done with results of Swap Test.
 - Swap Test is used to compare two quantum states. The probability of getting zero as the result of this test increases as two states get similar.
- We checked training and test loss results for validation.
- For tests with that uses just one layer, epoch size is set to 100 or 200. These runs aborted far more earlier than epoch sizes. Also, the history of them shows that values do not change much. Due to these reasons and seeing gradient calculation takes too much time, number of epochs kept as a small number (like 5) for other number of layers. This value is enough for comparison.
- Size of training and test sets are 9000 and 1000. All of them are randomly selected from $|0\rangle$, $|1\rangle$ and $|+\rangle$.

Results

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss(%)	Final Test Loss(%)
1	1	Adam(learning_rate=0.1)	ForwardDifference	12	10	60.5842	64.2502
2	3	Adam(learning_rate=0.1)	ForwardDifference	5	10	19.2453	20.5269
3	7	Adam(learning_rate=0.1)	ForwardDifference	5	10	4.9924	5.0051

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss(%)	Final Test Loss(%)
4	1	Adam(learning_rate=0.1)	ForwardDifference	13	3	62.0938	59.0006
5	3	Adam(learning_rate=0.1)	ForwardDifference	5	3	20.2496	18.6794
6	7	Adam(learning_rate=0.1)	ForwardDifference	5	3	6.4826	7.7014

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
7	1	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	11	3	67.1428	64.9002
8	3	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	5	3	19.9684	18.6347

Results contd.

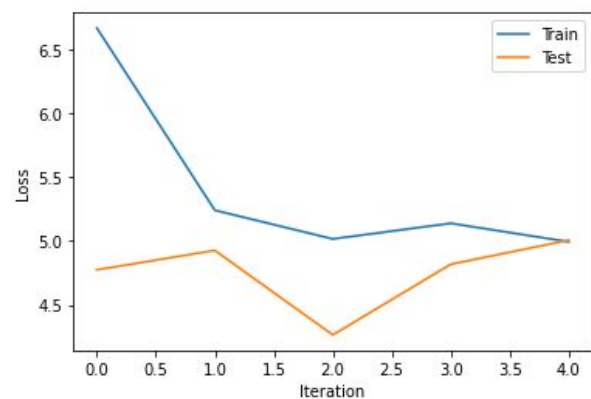
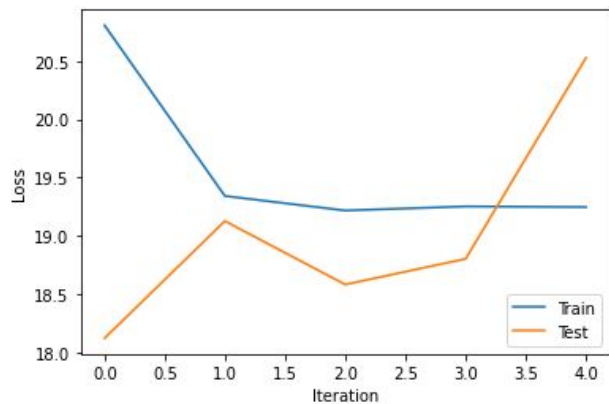
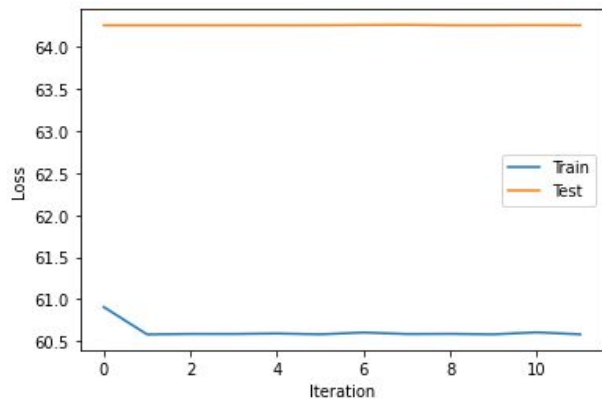
Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
9	7	Nadam(learnin g_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDiffer ence	5	3	5.8952	5.2210
10	1	Nadam(learnin g_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDiffer ence	11	10	61.4717	61.0500

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
11	3	Nadam(learnin g_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDiffer ence	5	10	19.3736	18.8781
12	7	Nadam(learnin g_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDiffer ence	5	10	5.5875	4.9944

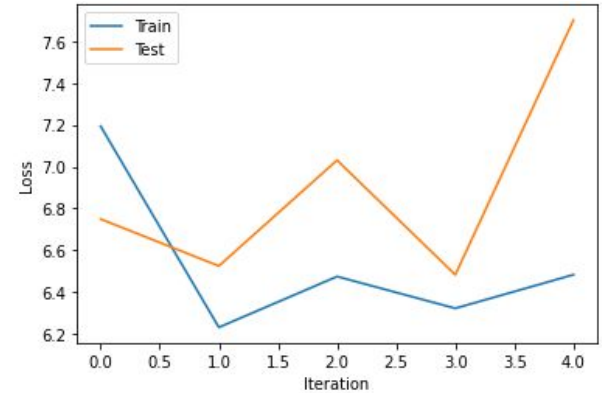
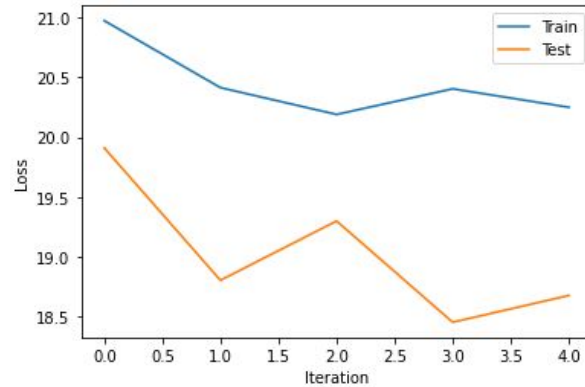
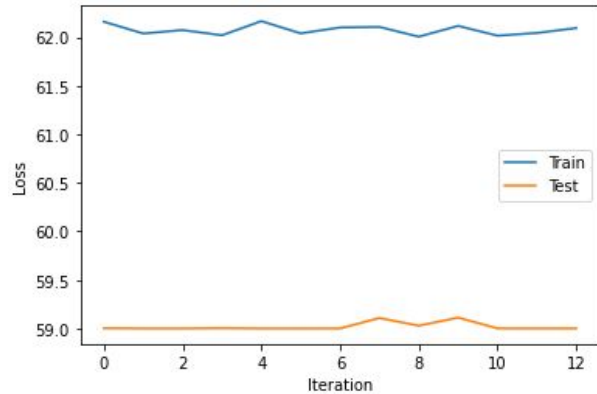
Loss History

- Test Run 1, 2 and 3



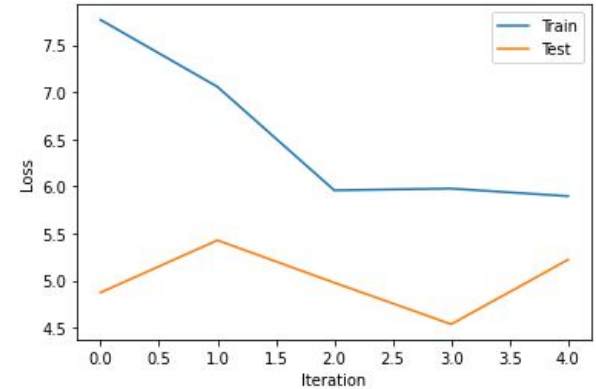
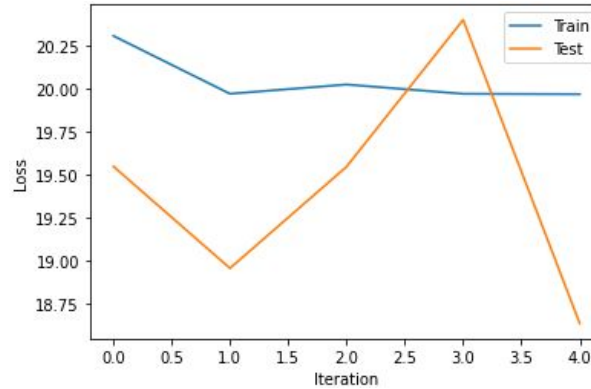
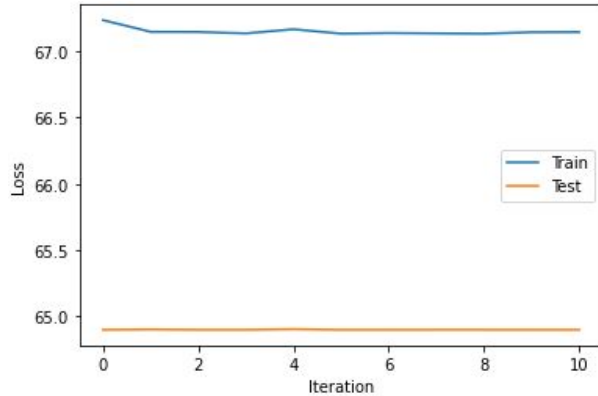
Loss History contd.

- Test Run 4, 5 and 6



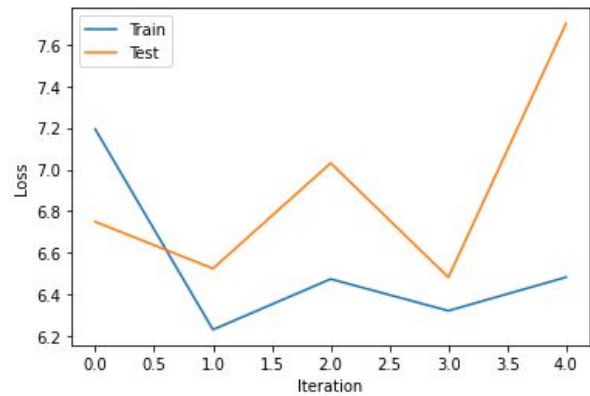
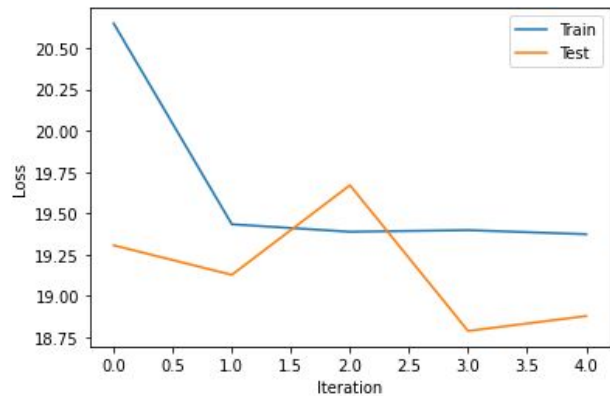
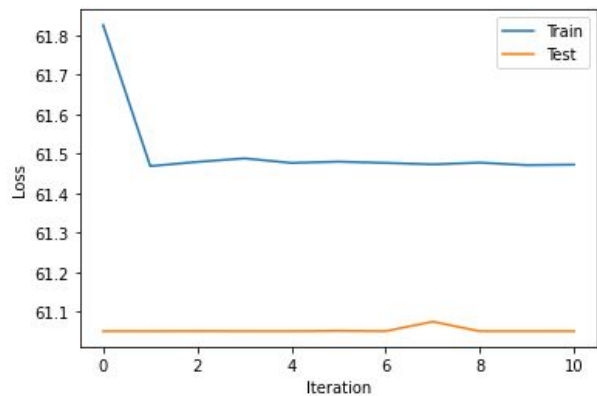
Loss History contd.

- Test Run 7, 8 and 9



Loss History contd.

- Test Run 10, 11 and 12



Setup 2

This setup is similar to the Setup 1 except:

- Ansatz 2 is used.
- Tests with optimizer=Adam, batch=3 and optimizer=Nadam, batch=5 are skipped because from as seen from previous setup, results do not change much.

Results

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
1	1	Adam(learning_rate=0.1)	ForwardDifference	15	10	27.4645	28.3505
2	3	Adam(learning_rate=0.1)	ForwardDifference	5	10	15.5248	14.4815
3	7	Adam(learning_rate=0.1)	ForwardDifference	5	10	5.5472	5.4367

Results contd.

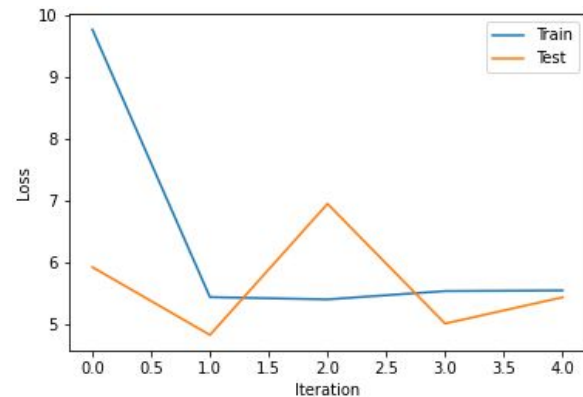
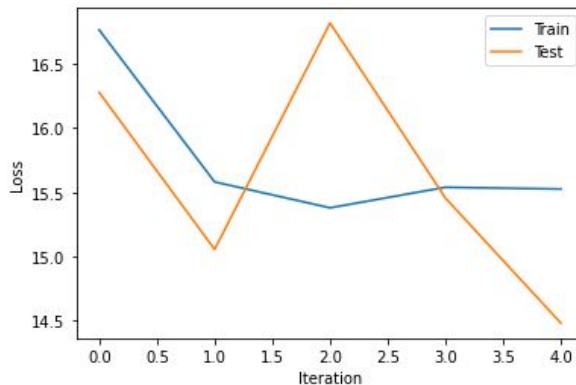
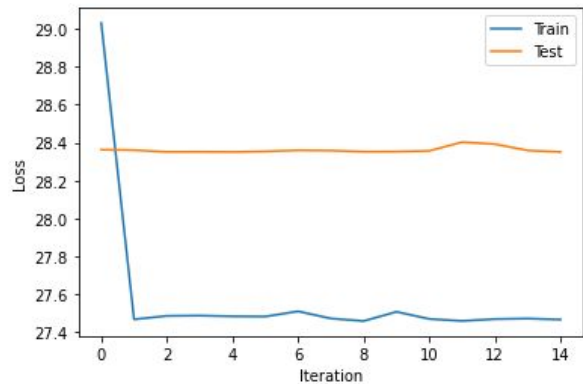
Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
4	1	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	16	3	27.8800	25.7071
5	3	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	5	3	20.0302	20.1113

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
6	7	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	5	3	6.1238	5.6381

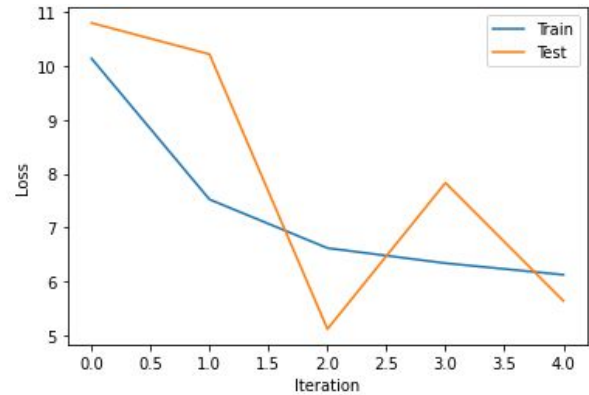
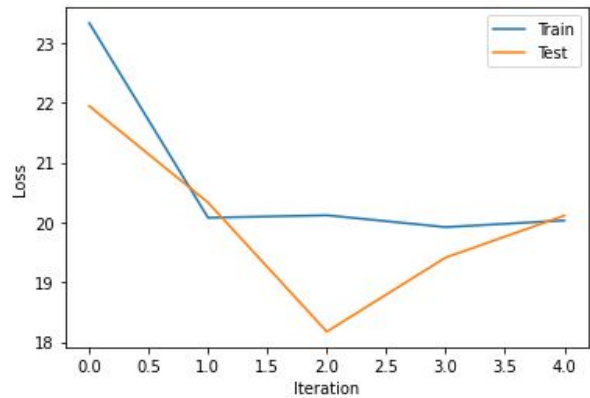
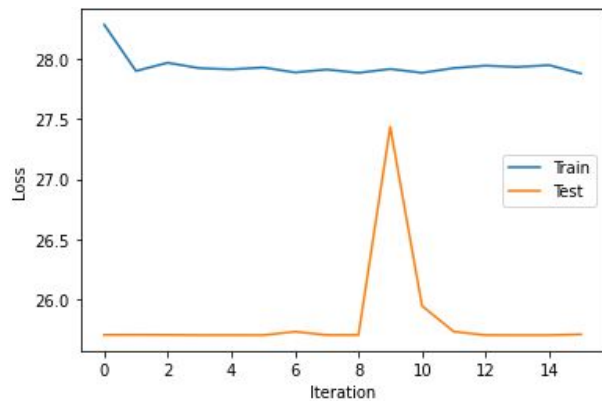
Loss History

- Test Run 1, 2 and 3



Loss History contd.

- Test Run 4, 5 and 6



Setup 3

This setup is similar to the Setup 2 except:

- Training and test is done with $|0\rangle$, $|1\rangle$ and $|+\rangle$ states for three qubits.
 - $|0\rangle$: Errornous and legit statuses
 - **Legit:** $|000\rangle$
 - **Errornous:** $|001\rangle$, $|010\rangle$, $|100\rangle$
 - $|1\rangle$: Errornous and legit statuses
 - **Legit:** $|111\rangle$
 - **Errornous:** $|110\rangle$, $|101\rangle$, $|011\rangle$
 - $|+\rangle$: Just the legit status
 - **Legit:** $(|000\rangle + |111\rangle)/\sqrt{2}$
 - **Errornous:** $(|001\rangle + |110\rangle)/\sqrt{2}$, $(|010\rangle + |101\rangle)/\sqrt{2}$, $(|100\rangle + |011\rangle)/\sqrt{2}$

Results

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
1	1	Adam(learning_rate=0.1)	ForwardDifference	25	10	26.8083	25.9614
2	3	Adam(learning_rate=0.1)	ForwardDifference	5	10	21.3655	20.7298
3	7	Adam(learning_rate=0.1)	ForwardDifference	5	10	10.3257	13.1211

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
4	1	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	22	3	27.0454	26.5334
5	3	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	5	3	21.4473	22.5160

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
6	7	Nadam(learning_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDifference	5	3	11.8276	10.2903
7	1	Adam(learning_rate=0.01)	ForwardDifference	14	10	26.5501	25.5450
8	3	Adam(learning_rate=0.01)	ForwardDifference	10	10	18.9915	19.0568

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
9	7	Adam(learning_rate=0.01)	ForwardDifference	10	10	8.2460	8.5038
10	1	Nadam(learning_rate=0.01, beta_1=0.9, beta_2=0.999)	ForwardDifference	13	3	26.4941	26.4982
11	3	Nadam(learning_rate=0.01, beta_1=0.9, beta_2=0.999)	ForwardDifference	10	3	17.1553	16.9130

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
12	7	Nadam(learning_rate=0.01, beta_1=0.9, beta_2=0.999)	ForwardDifference	10	3	9.3315	9.1887
13	1	Adam(learning_rate=0.001)	ForwardDifference	26	10	25.8677	26.1264
14	3	Adam(learning_rate=0.001)	ForwardDifference	5	10	19.7636	19.9868

Results contd.

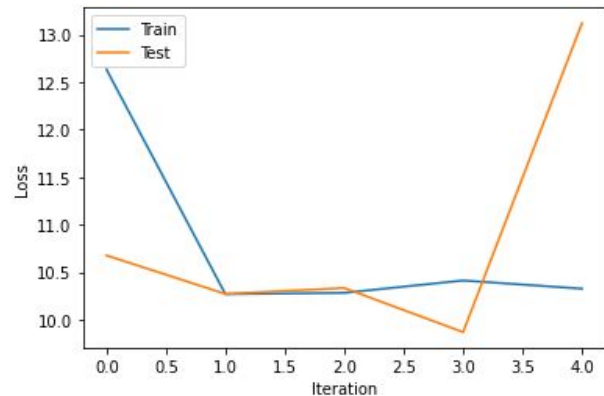
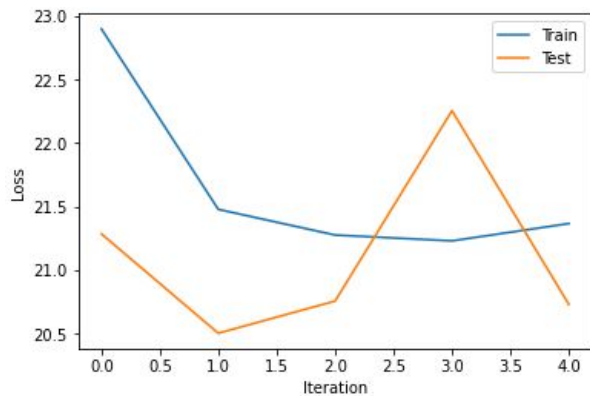
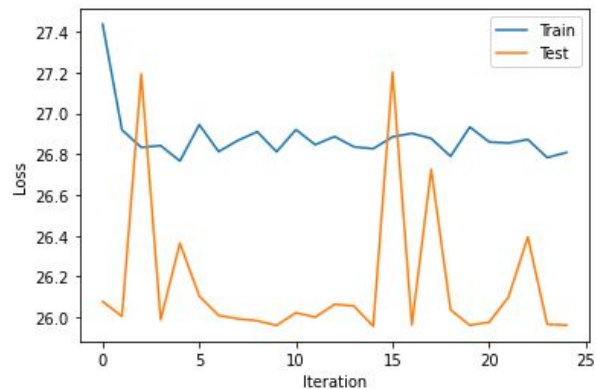
Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
15	7	Adam(learning_rate=0.001)	ForwardDifference	5	10	13.8372	13.4300
16	1	Nadam(learning_rate=0.001, beta_1=0.9, beta_2=0.999)	ForwardDifference	24	3	25.9686	28.1369

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
17	3	Nadam(learning_rate=0.001, beta_1=0.9, beta_2=0.999)	ForwardDifference	5	3	33.4405	33.7549
18	7	Nadam(learning_rate=0.001, beta_1=0.9, beta_2=0.999)	ForwardDifference	5	3	9.1788	8.7624

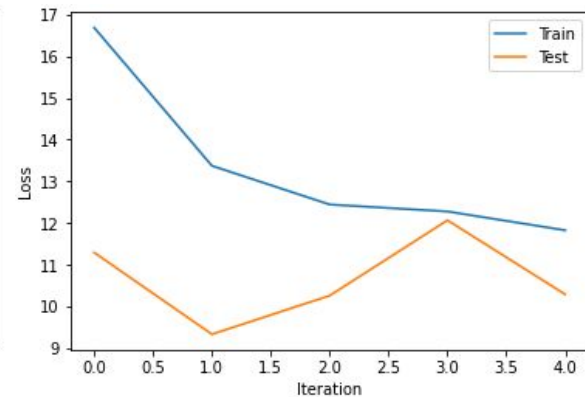
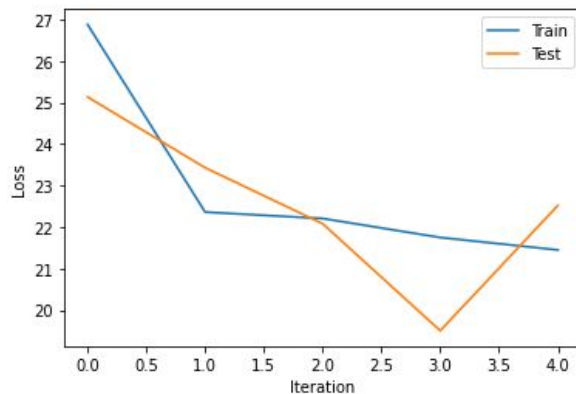
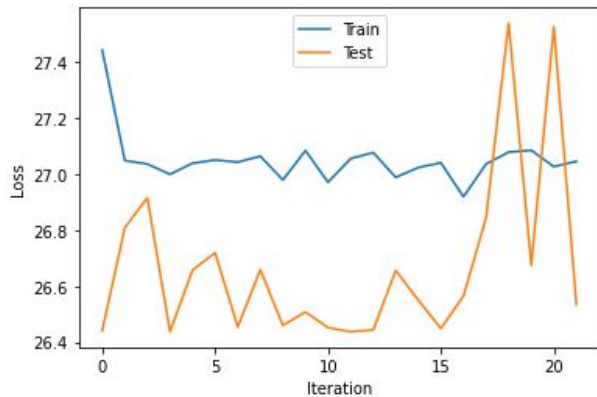
Loss History

- Test Run 1, 2 and 3



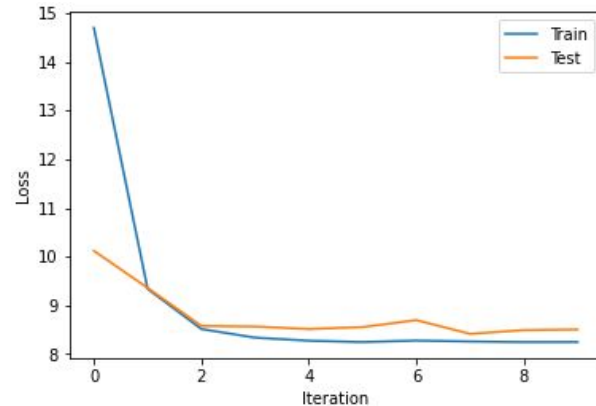
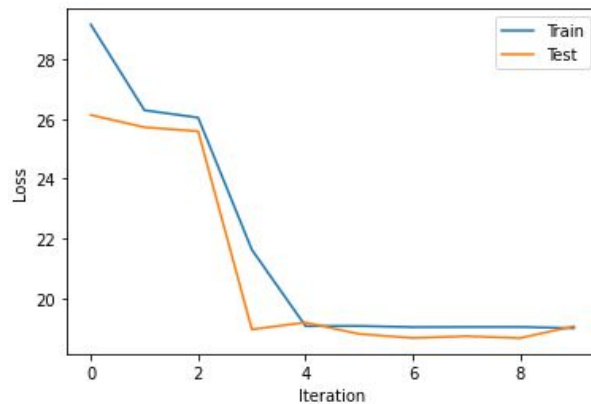
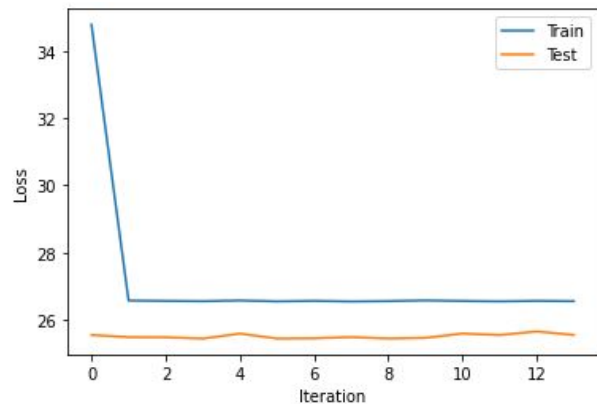
Loss History contd.

- Test Run 4, 5 and 6



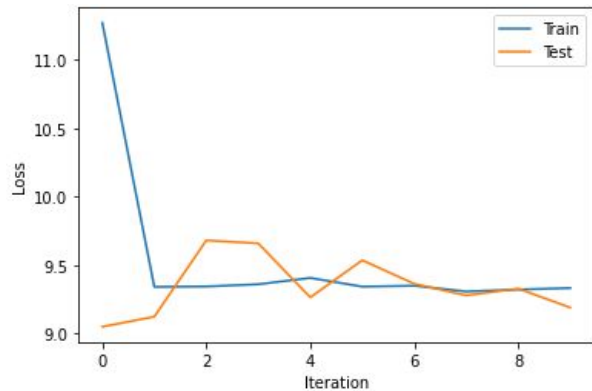
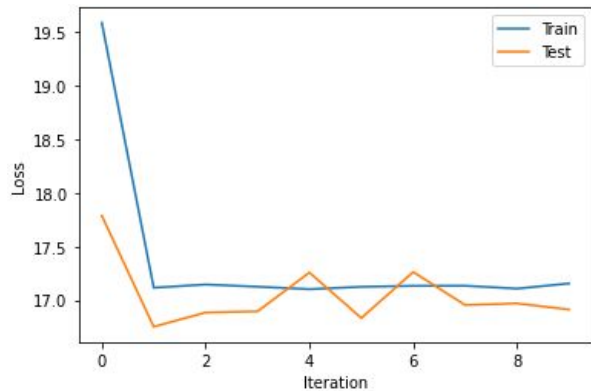
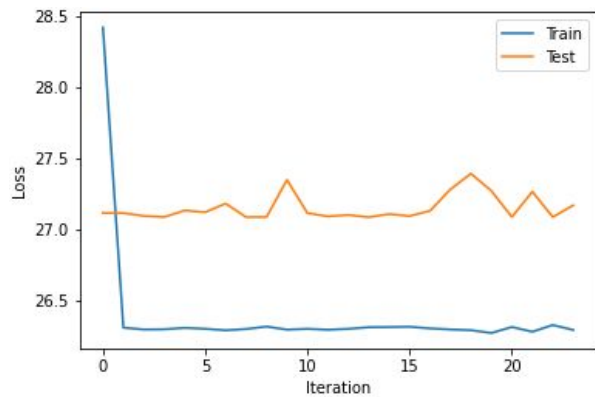
Loss History contd.

- Test Run 7, 8 and 9



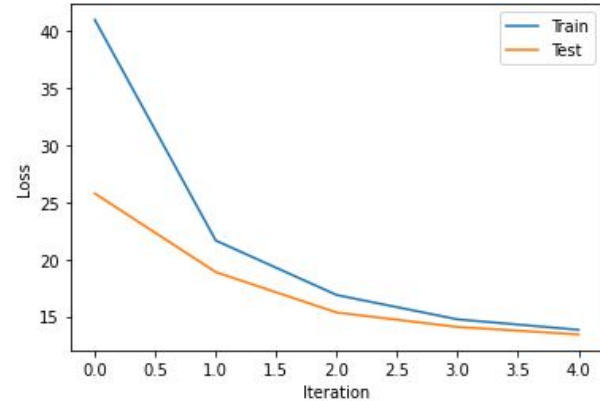
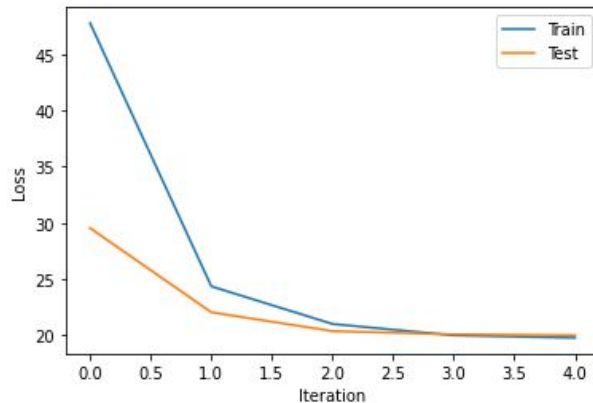
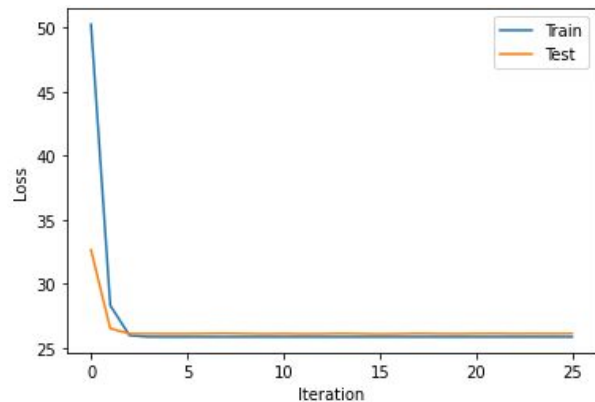
Loss History contd.

- Test Run 10, 11 and 12



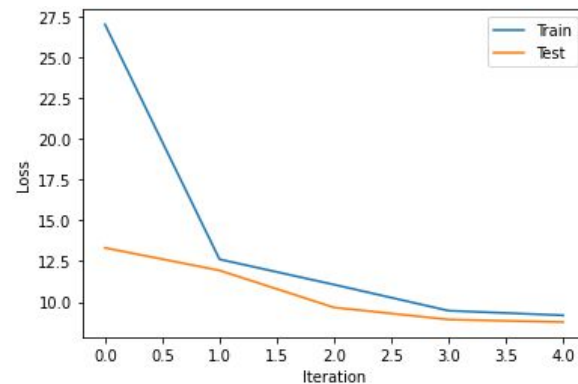
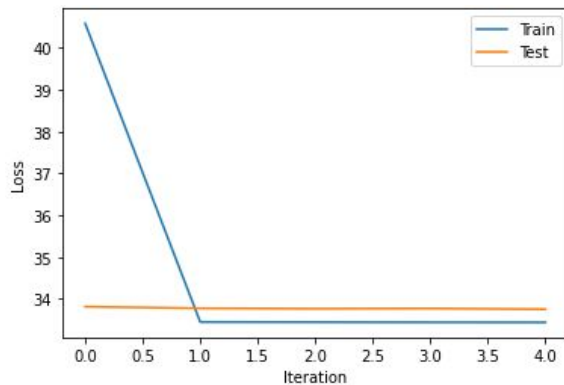
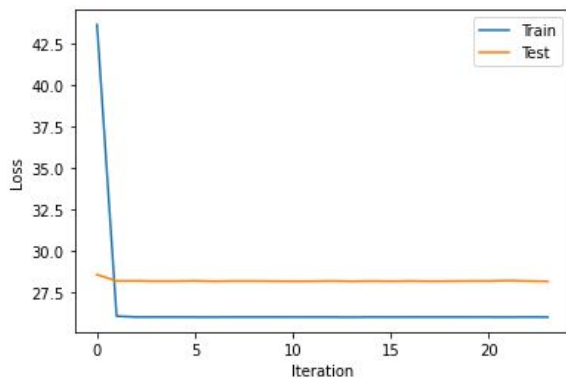
Loss History contd.

- Test Run 13, 14 and 15



Loss History contd.

- Test Run 16, 17 and 18



Setup 4

This setup is similar to the Setup 3 except:

- We passed repetitions=1000 to the PQC in order to simulate measuring results for 1000 times. The result PQC gives the average of the 1000 'measurements'.
- For time purposes, just tests with single layer are performed.

Results

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
1	1	Adam(learning_rate=0.1)	ForwardDifference	26	10	97.3975	98.3172
2	1	Adam(learning_rate=0.01)	ForwardDifference	34	10	43.7091	45.1272
3	1	Adam(learning_rate=0.001)	ForwardDifference	50	10	35.9870	36.2372

Results contd.

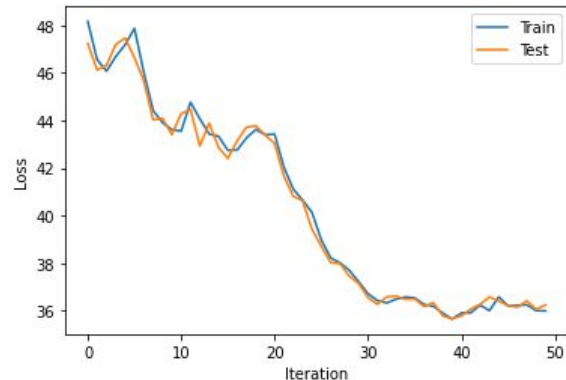
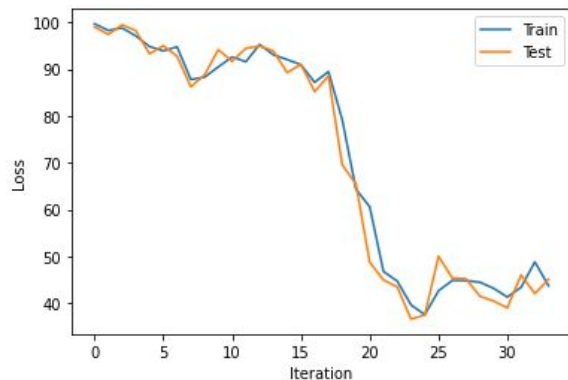
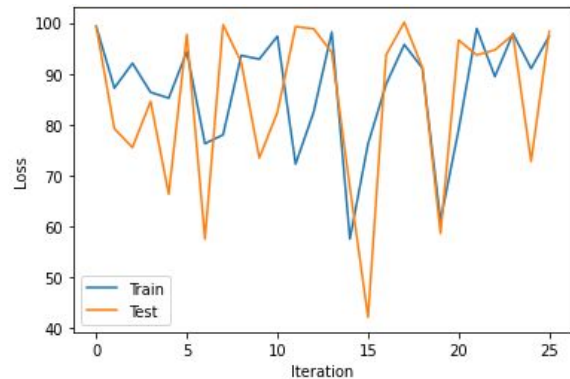
Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
4	1	Nadam(learnin g_rate=0.1, beta_1=0.9, beta_2=0.999)	ForwardDiffer ence	16	3	89.6014	99.8226
5	1	Nadam(learnin g_rate=0.01, beta_1=0.9, beta_2=0.999)	ForwardDiffer ence	27	3	54.0134	66.3436

Results contd.

Run Num	Number of Layers	Optimizers	Gradient Calculator	Epochs	Batch Size	Final Training Loss	Final Test Loss
6	1	Nadam(learning_rate=0.001, beta_1=0.9, beta_2=0.999)	ForwardDifference	95	3	33.5446	33.5388

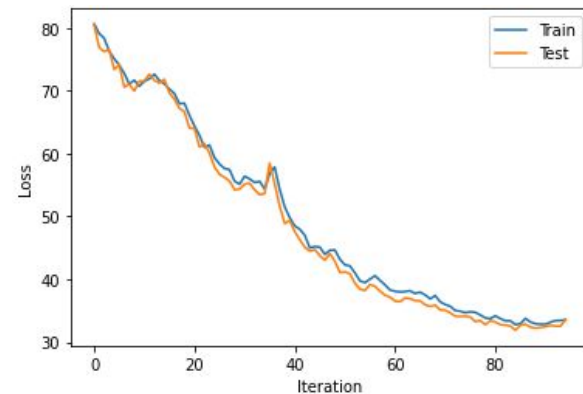
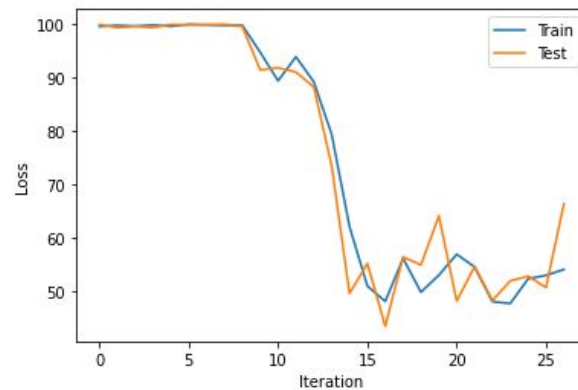
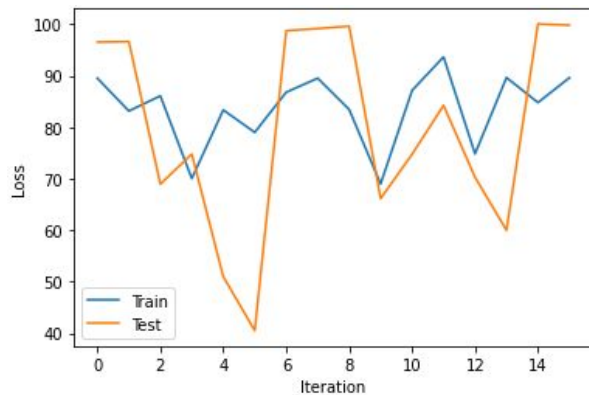
Loss History

- Test Run 1, 2 and 3



Loss History contd.

- Test Run 4, 5 and 6



Interpretations

For Setup 1 and Setup 2:

- For smaller number of layers, there is a huge difference between the first setup and the second one. This gap closes fastly as we increase number of layers. This result may also imply that there is a minimum loss that does not change no matter what setup and parameters are used.
- Results do not change much as other parameters except number of batches change within a setup.
- For each test, loss values do not change much after approximately 5 runs.

Interpretations contd.

For the Setup 3:

- For the Adam optimizer results did not change too much compared to the Setup 2.
- For the Nadam optimizer, effectiveness of the structure with 7 layers fell significantly, though other results did not change much to the Setup 2.
- Different learning rates gave similar final results though shape of histories are smoother for smaller learning rates.
- For each test, loss values do not change much after approximately 5 or 10 runs.

Interpretations contd.

For the Setup 4:

- Results for the Adam and the Nadam optimizers are similar.
- Learning rate makes a really big difference.
 - Bigger rates give better results.