Programming Assignment 3 - EE5179
EE21S049 Report

1. MNIST Classification using RNN:

Models:

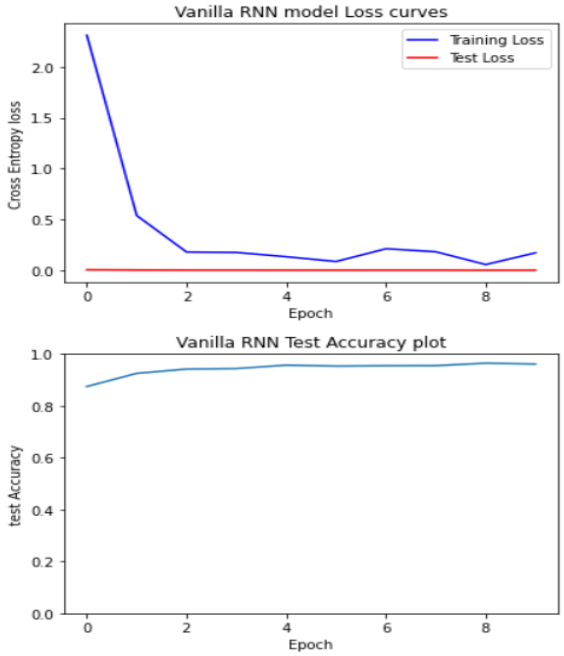
a. Vanilla RNN
b. Vanilla LSTM
c. Vanilla GRU
d. Bidirectional RNN
e. Bidirectional LSTM

Hyperparameters:

a. Learning rate = 0.01
b. Number of epochs = 10
c. Optimizer - Adam
d. Batch Size = 64
e. Hidden state size = 128

Outputs:

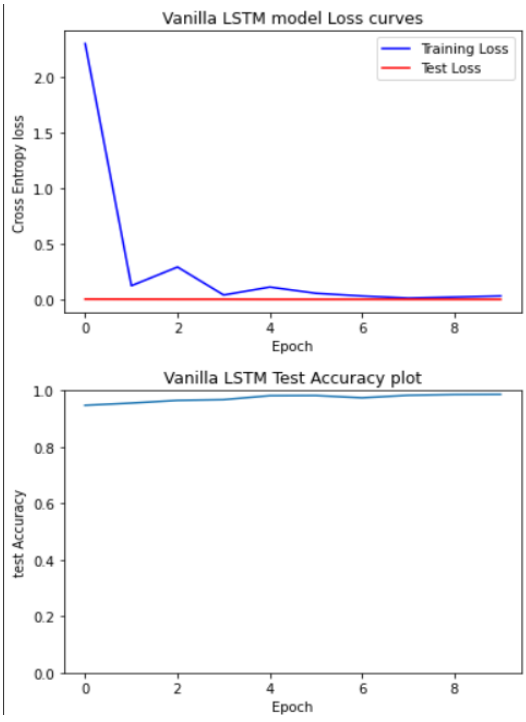
1. Vanilla RNN:
    a. Loss and Accuracy plots:



b. Average Prediction accuracy of Vanilla RNN model = 94.329
c. Random Predictions of Vanilla RNN model:

True label:4, predicted as 4
True label:2, predicted as 2
True label:4, predicted as 4
True label:1, predicted as 1
True label:4, predicted as 4
True label:7, predicted as 7
True label:8, predicted as 8
True label:8, predicted as 8
True label:7, predicted as 7
True label:7, predicted as 7

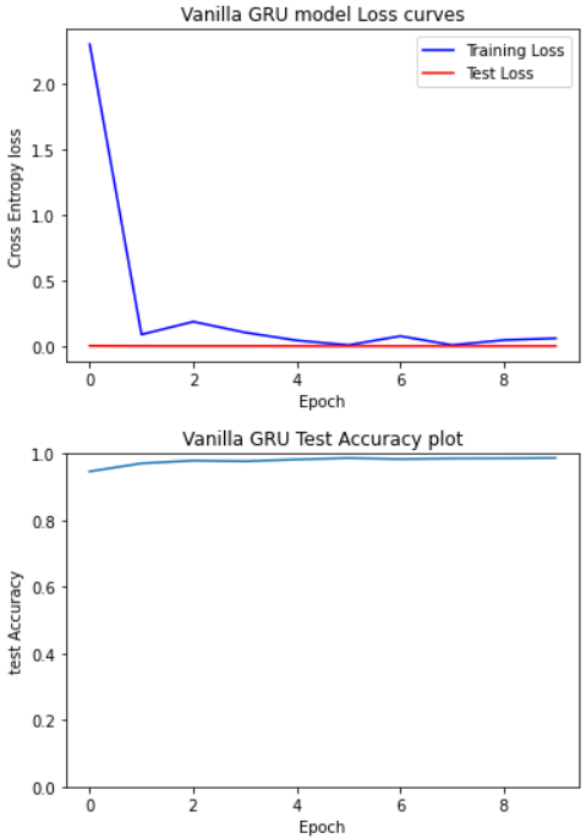
2. Vanilla LSTM :
   a. Loss and Accuracy plots:



   b. Average Prediction accuracy of Vanilla LSTM model is 97.3109%
   c. Random Predictions of Vanilla LSTM model:
      True label:1, predicted as 1
      True label:2, predicted as 2

True label:9, predicted as 9
True label:6, predicted as 6
True label:4, predicted as 4
True label:0, predicted as 0
True label:9, predicted as 9
True label:0, predicted as 0
True label:8, predicted as 8
True label:3, predicted as 3

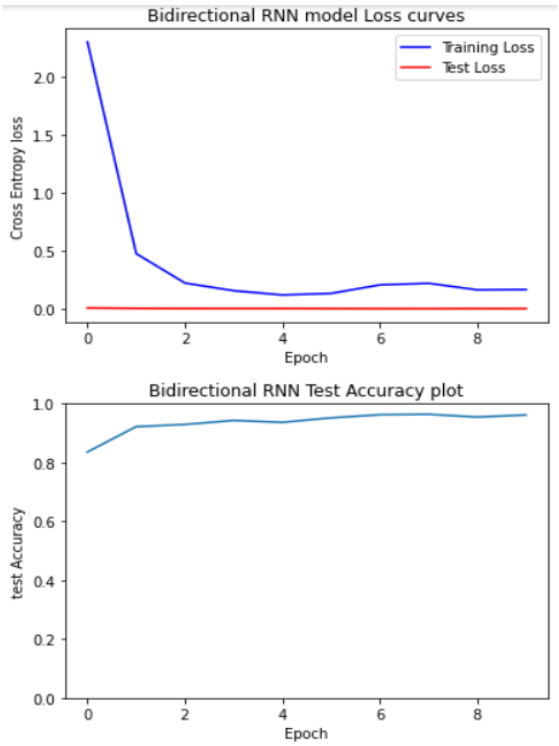
3. Vanilla GRU:
    a. Loss and Accuracy plots:



    b. Average Prediction accuracy of Vanilla GRU model = 97.870999%
    c. Random Predictions of Vanilla GRU model
        True label:8, predicted as 8
        True label:9, predicted as 9
        True label:8, predicted as 8
        True label:1, predicted as 1
        True label:9, predicted as 9

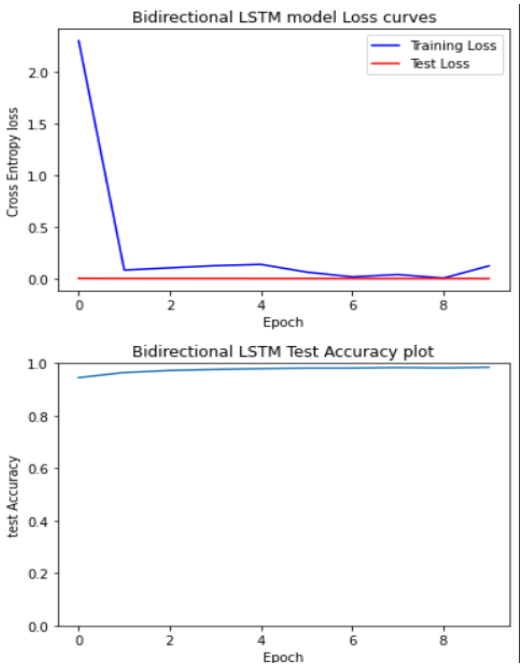
True label:3, predicted as 3
True label:3, predicted as 3
True label:1, predicted as 1
True label:8, predicted as 8
True label:7, predicted as 7

4. Bidirectional RNN:
    a. Loss and Accuracy plots:



    b. Average Prediction accuracy of Bidirectional RNN model = 93.63
    c. Random Predictions of Bidirectional RNN model
        True label:4, predicted as 4
        True label:6, predicted as 6
        True label:9, predicted as 9
        True label:1, predicted as 1
        True label:9, predicted as 9
        True label:0, predicted as 0
        True label:9, predicted as 9
        True label:3, predicted as 3
        True label:8, predicted as 8
        True label:7, predicted as 7

5. Bidirectional LSTM :
    a. Loss and Accuracy plots;



    b. Average Prediction accuracy of Bidirectional LSTM model = 97.54899%
    c. Random Predictions of Bidirectional LSTM model

        True label:6, predicted as 6
        True label:1, predicted as 1
        True label:7, predicted as 7
        True label:2, predicted as 2
        True label:3, predicted as 3
        True label:4, predicted as 4
        True label:7, predicted as 7
        True label:6, predicted as 6
        True label:0, predicted as 0
        True label:2, predicted as 2

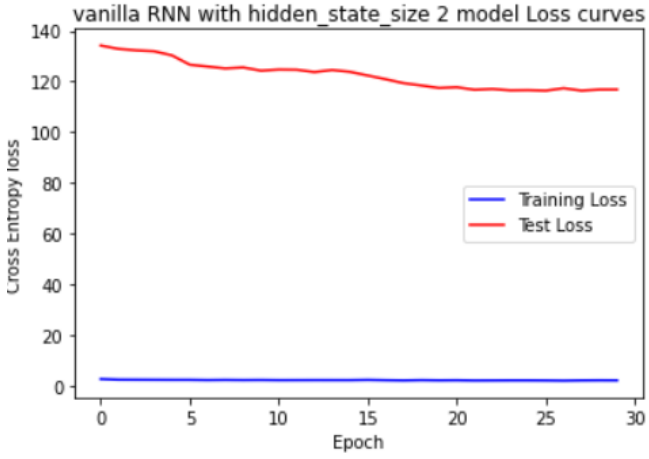
2. Remembering the number at a particular index in a given sequence

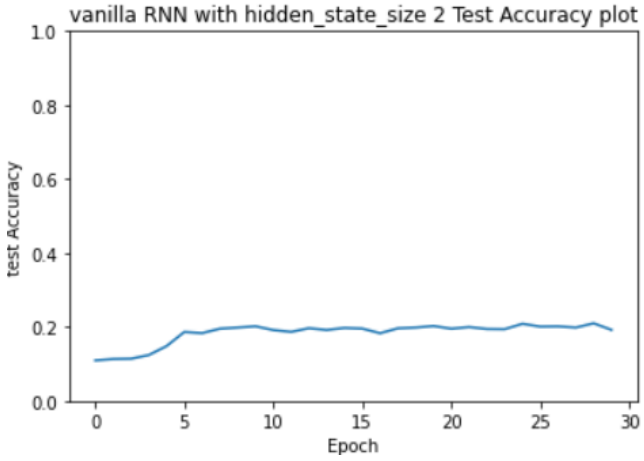
Hyperparameters:

a. Number of epochs = 30
b. Train iterations = 300
c. Test iterations = 60
d. Prediction iterations = 5

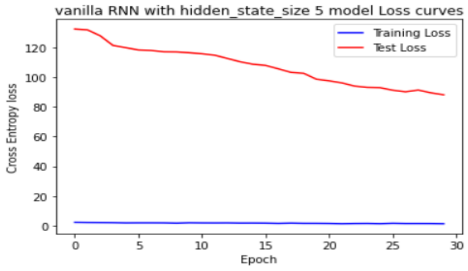
Loss and Accuracy plots:

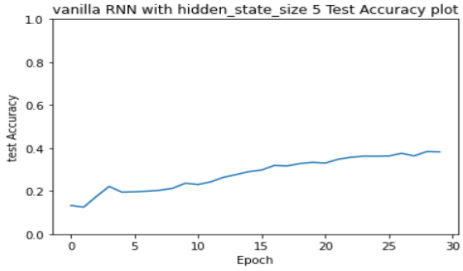
1. Vanilla RNN :
    a. Hidden state size = 2



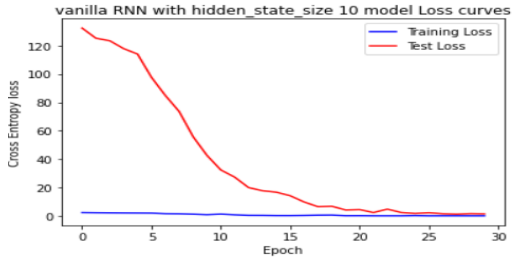


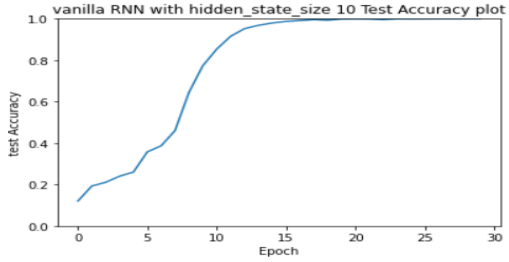
b. Hidden state size = 5



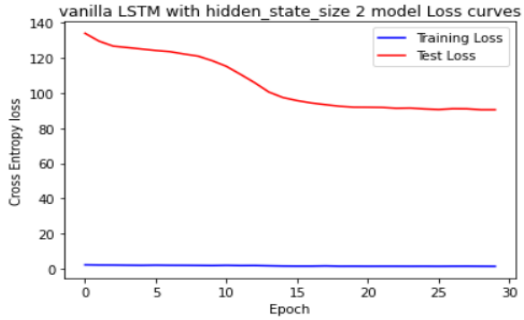


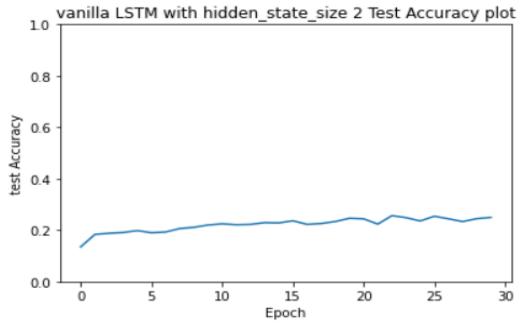
c. Hidden state size = 10



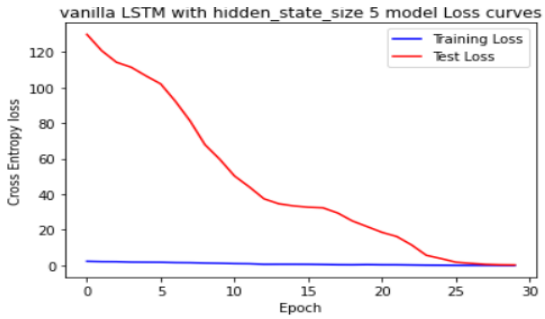


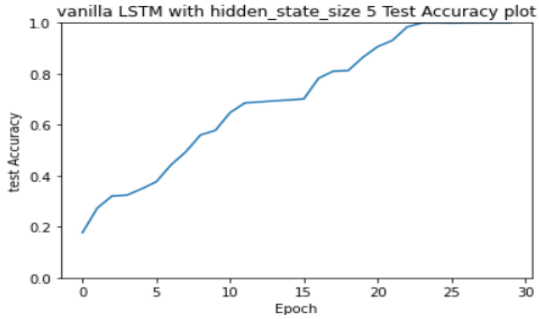
2. Vanilla LSTM :
    a. Hidden state size = 2



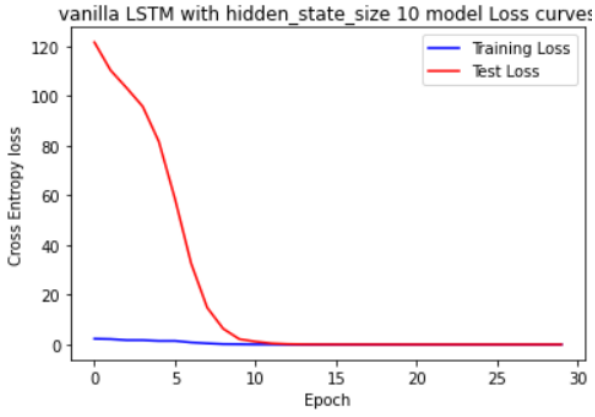


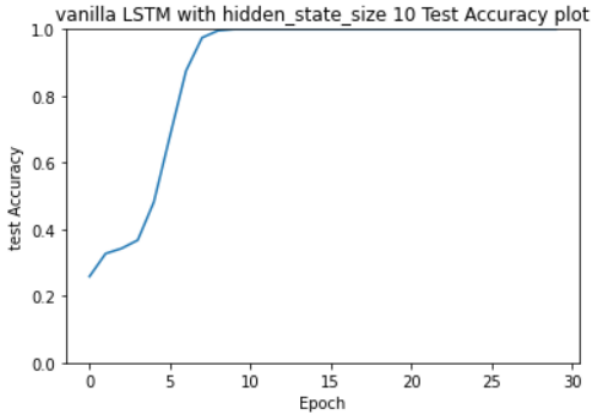
    b. Hidden state size = 5



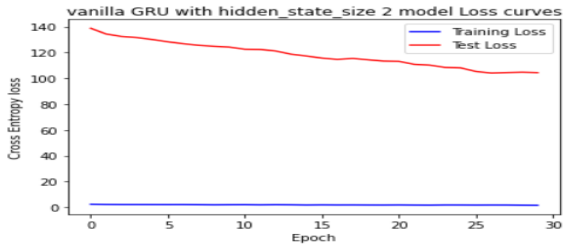


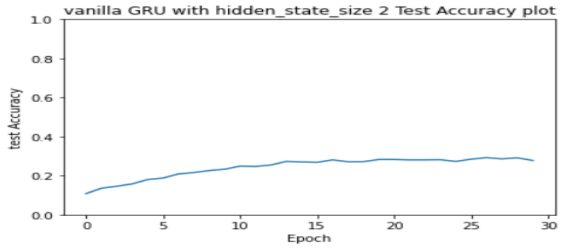
c. Hidden state size = 10



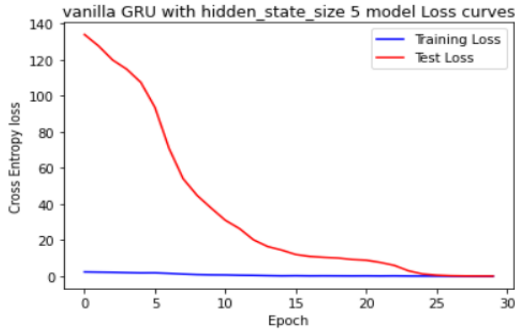


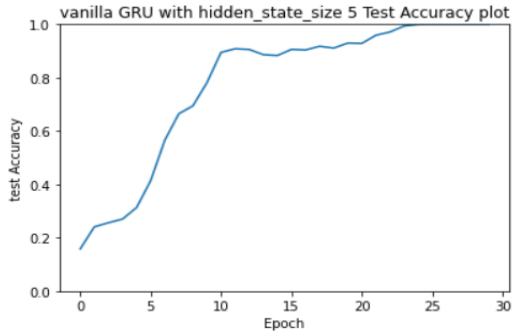
3. Vanilla GRU :
   a. Hidden state size = 2



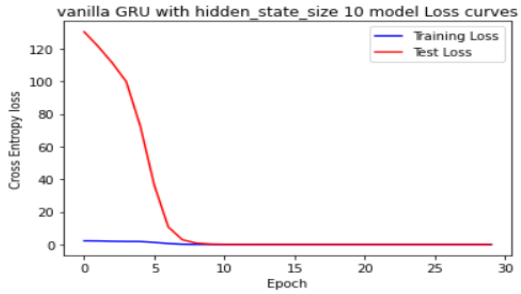


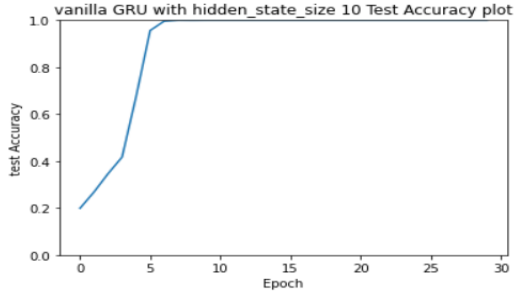
b. Hidden state size = 5





c. Hidden state size = 10





4. Random Predictions of best accuracy models:

From the above accuracy plots of vanilla RNN, LSTM and GRU models with different hidden state sizes, it can be observed that vanilla RNN with hidden state size as 10 , vanilla LSTM with hidden state size 5 and 10 and vanilla GRU with hidden state size 5 and 10 gave better results. As hidden state size increases, performance of the model increases.

a. Random Predictions of vanilla RNN with hidden_state_size 10 model for K=2

Generated Sequence:tensor([[2, 4, 5, 3, 0, 4, 5, 8, 2]], dtype=torch.int32)
Predicted Output:tensor([4])
Generated Sequence:tensor([[5, 1, 7, 7, 6, 4, 3, 3]], dtype=torch.int32)
Predicted Output:tensor([1])
Generated Sequence:tensor([[1, 1, 4, 5, 3]], dtype=torch.int32)
Predicted Output:tensor([1])
Generated Sequence:tensor([[3, 0, 7, 4, 7]], dtype=torch.int32)
Predicted Output:tensor([0])
Generated Sequence:tensor([[2, 7, 2, 7, 0, 7, 4]], dtype=torch.int32)
Predicted Output:tensor([7])

b. Random Predictions of vanilla LSTM with hidden_state_size 10 model with K=2

Generated Sequence:tensor([[6, 4, 4, 7, 4]], dtype=torch.int32)
Predicted Output:tensor([4])
Generated Sequence:tensor([[2, 7, 4, 7, 0]], dtype=torch.int32)
Predicted Output:tensor([7])
Generated Sequence:tensor([[5, 4, 5, 3, 5, 1]], dtype=torch.int32)
Predicted Output:tensor([4])
Generated Sequence:tensor([[6, 5, 3, 1, 4, 2]], dtype=torch.int32)
Predicted Output:tensor([5])
Generated Sequence:tensor([[0, 5, 6, 5, 5, 1, 0, 1, 0]], dtype=torch.int32)
Predicted Output:tensor([5])

c. Random Predictions of vanilla GRU with hidden_state_size 10 model with K=2

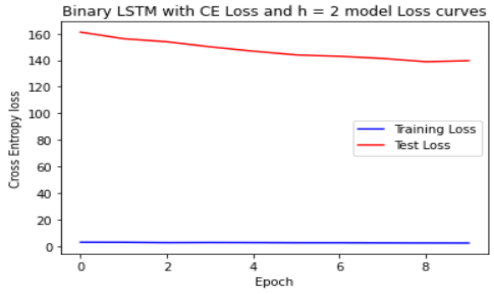
Generated Sequence:tensor([[7, 6, 6]], dtype=torch.int32)
Predicted Output:tensor([6])
Generated Sequence:tensor([[7, 1, 0, 6, 1, 6, 7, 2]], dtype=torch.int32)
Predicted Output:tensor([1])
Generated Sequence:tensor([[6, 0, 6]], dtype=torch.int32)

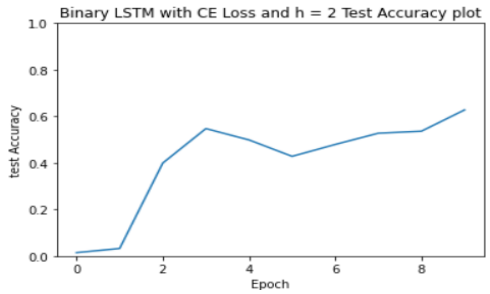
Predicted Output:tensor([0])
Generated Sequence:tensor([[4, 1, 1, 4, 3, 1, 7, 6]], dtype=torch.int32)
Predicted Output:tensor([1])
Generated Sequence:tensor([[8, 5, 8, 5, 7, 7, 3, 0]], dtype=torch.int32)
Predicted Output:tensor([5])

3. Adding two binary strings:

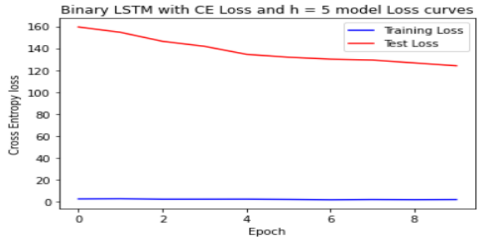
a. Cross Entropy loss:

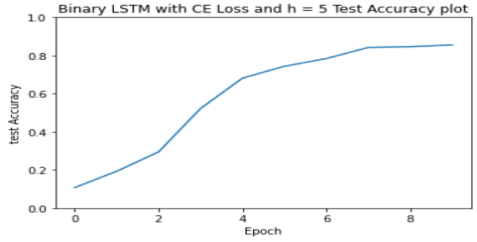
i. Hidden State Size = 2



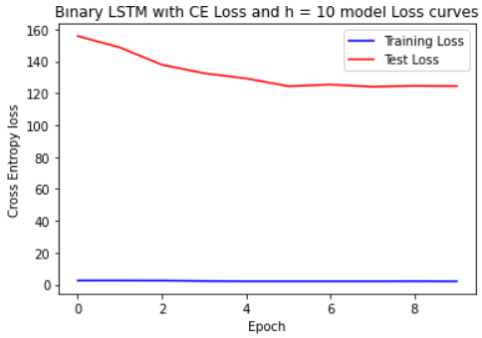


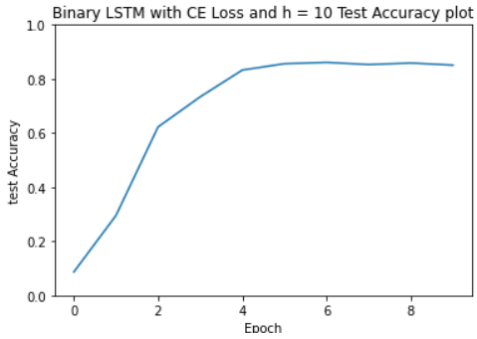
ii. Hidden State Size = 5





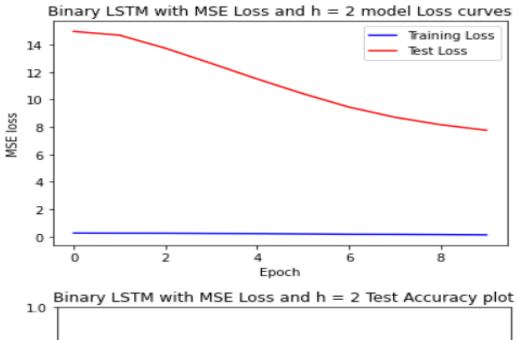
iii. Hidden State Size = 10

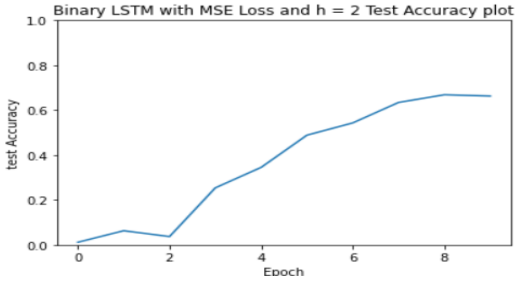




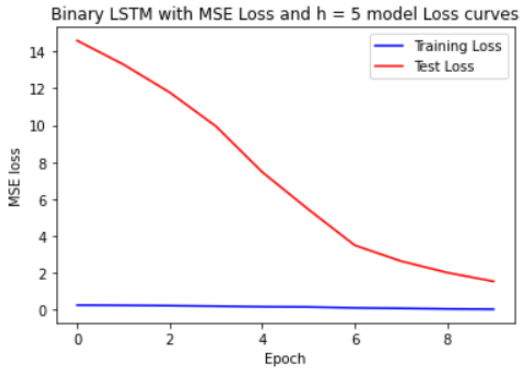
b. MSE Loss:

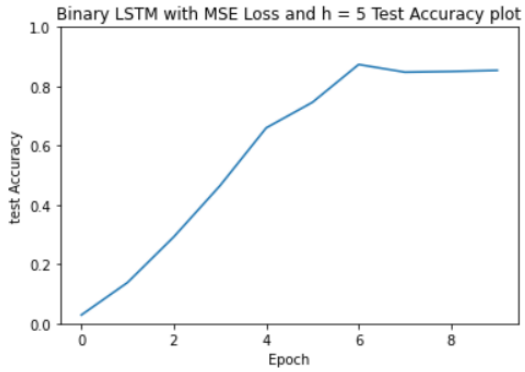
i. Hidden State Size = 2





ii. Hidden State Size = 5





iii. Hidden State Size = 10