Submission:

You need to submit the code of your model and fill in the following information:

Model Information:

★ Number of layers: 2 layers of lstm model

★ Number of units in each layer: 5

★ Activation functions used:

input:sigmoid/leaky relu(model4/model5)

hidden:sigmoid

output:relu

★ Loss function: mean square error

Training Epochs: 250/350(model6)

Best Training mae: 6.9%

Best **Testing mae**: 7.3 %

Optimization techniques employed: Adam,early stopping

Difference in accuracies after each optimization technique that you applied: (compare with model 1)

(1) Optimization technique name: batch normalization and drop out(model2)

Before optimization: Training/Testing mae = 6.9 /7.3

After optimization: Training/Testing mae = 8/ 8.8

Any other changes: _____

(2) Optimization technique name: SGD stochastic gradient decent

Before optimization: Training/Testing Accuracies =6.9/7.3

After optimization: Training/Testing Accuracies = 7.28/7.75

Any other changes: _____

(3) Optimization technique name: weight initializers

Before optimization: Training/Testing Accuracies = 6.9/7.3

After optimization: Training/Testing Accuracies =6.9/7.4

Any other changes: _____

Anything special about your model:

When using model 1 and model 6 .Although they have lower validation loss and testing loss.They also cause vanishing gradient problem sometimes.So I use leaky relu function to solve this problem.The model with weight initializers often has lower loss than normal one .But during the last time I train those model.The normal one is better than the one which implements weight initializers.

Comments on the course:

Teacher makes a lot of effort teaching us and try to design a lot of lab class and homework to let us better know about the implementation in deep learning and also invite person who has a lot of experience in the industry relation to deep learning to have a speech. As a student which is not major in CSIE. I learn a lot in this course.