CS 4782 Coding Assignment 1 Written Responses

Due: 2/12/26 11:59 PM on Gradescope

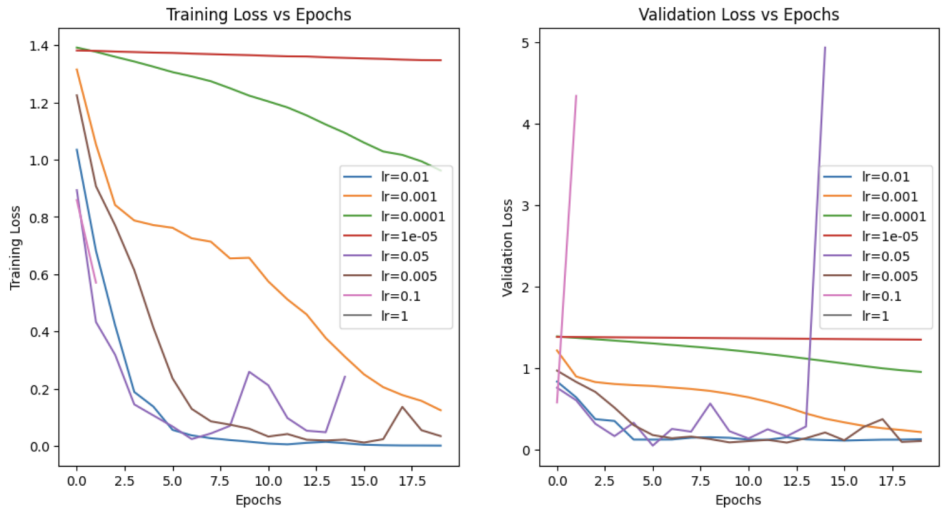
Late submissions accepted until 2/14/26 11:59 PM

Note: For homework, you can work in teams of up to 2 people. Please include your teammates' NetIDs and names on the front page and form a group on Gradescope.

Problem 1:

Note down anything you have observed from your experiment with learning rates (one or two sentences will suffice). Include an image of the plots being referenced in this question.

Answer: From the experiment, very high learning rates (i.e. 1) led to unstable training and divergence in validation loss, while very low learning rates (e.g., 1e-5) resulted in slow or ineffective learning. Moderate learning rates around 1e-3 to 1e-2 produced the most stable convergence and lowest validation loss for the ConvNet during training.



Problem 2:

What do you notice about the loss graphs and the differences between the five models you have implemented? Write 2-3 sentences. Include an image of the plots referenced in this question.

Answer: From the loss graphs, the baseline ConvNetDropout shows relatively higher training and validation loss compared to the other architectures. This indicates that it has the worst generalizatoin. Models incorporating Batch Normalization and residual connections (ConvNetBN and ResNet) converge much faster and achieve similar validation loss, though ResNet seems to have the most stable and lowest loss overall.