Journal Report 12
12/9/19-12/13/19
Odin Woitek
Computer Systems Research Lab
Period 1, White

Daily Log

Monday December 9

In order to solve my problem of the accuracy staying the same, I made a small version of the network and made my batch size smaller, which didn't seem to effect much. I suspect that the learning rate is the problem here, so I will work on that.

Tuesday December 10

I tried to look into the learning rate further, setting it manually. Just by looking at it over the course of training, it appeared to stay stagnant after a quick drop in the beginning, which suggests that the decay was too great. I also decided to test my network for many epochs on 50 images from the training set. Although this should have overfit the data, it ended up getting stuck at 26 percent accuracy, which definitely should not be happening. I decided that it must be an extremely low learning rate that is causing this issue, so I tried to set a fixed learning rate of 0.1 to see what happens. Unfortunately, I kept getting an error any time I tried to change the optimizer.

Thursday December 12

I noticed that this entire time, I was running tensorflow for CPU on zoidberg, so I uninstalled that and installed the GPU version. Solved optimizer issue by importing from tensorflow.keras rather than tensorflow.python.keras. Now that I was able to do this, I started messing around with the learning rate a bit, but it was always either too high or there would be no improvement at all, again maxing out at exactly 26 percent accuracy for the 50 images. This made me suspect that perhaps learning rate is not the problem and that there is something else wrong here. After looking at the output data, I found that author 0 came up 13 times, so if the model always predicted 0, it would be correct 26 percent of the time.

Timeline

Date	Goal	Met
November 18-22	Run and evaluate three models on zoidberg	Yes, looked at model summary, took note of accuracy as it trained
December 2-6	Improve accuracy to 0.3	No, tried implementing validation set, modifying other things but accuracy did not improve.
December 9-13	Successfully overfit model on small dataset of 50 images	No, but identified problem being guessing the same author each time
December 16-20	Train a model that can predict author identity with at least 0.7 accuracy on the test set	

Reflection

I suspected that my main issue was learning rate being too low, as my model would quickly hit a certain accuracy and just sit there, not improving. Unfortunately, after trying to find ways to modify the learning rate, I kept getting an issue where I couldn't specify the optimizer as a function, and it only worked if I specified it as a string, which doesn't allow you to alter the learning rate. In looking for solutions to this problem, I found that one solution that might work would be to change the import statements, but I needed my import statements to be written the way they were to fix another issue I had earlier in the project, so I was stuck. After some research, I found that a good metric for debugging the issue I was having with my model was to attempt to overfit on a small dataset. When trying that out, it didn't work, which brought me back to the learning rate problem, as that is a common cause of this issue. I had hit what seemed like a wall previously, but I decided to try running the code on the tensorflow gpu version that I installed on zoidberg. Although the import statements did not work on the tensorflow cpu version I have on my computer, they did work on zoidberg, which allowed me to alter the learning rate. However, upon further inspection it appeared that learning rate was not the problem. After looking at my data, I suspect that the issue is that the model outputs author '000' for every input picture, which would account for the issues I am having.