15-400 Milestone 5

Prithvi Gudapati (pgudapat)

March 2019

1 Major Changes

There are no major changes. The deadline that is coming up is the SOSP deadline, which is April 24th. If we are not able to meet this, the next deadline would be the NIPS deadline, which is May 23rd.

2 What You Have Accomplished Since Your Last Meeting

I implemented Char-RNN in Caffe using the caffe training process, which will allow it to be ported easily to GeePs, Gaia, FedAvg, and DGC. This turned out to be a lot more challenging that initially expected because the text file needed to be converted to LMDB and then passed to the data layer in caffe. Although Caffe internally represents Dtype, the data type inside blobs, as floats, lmdb works with bytes, so we had to use chars instead of floats for the data. This ended up working out because the vocabulary is relatively small, since we are working with a character-based language model as opposed to a word-based language model. While working with this model we encountered unstable convergence, but eventually were able to see a similar convergence to the original model. The primary difference between our model and the original model is that we randomly shuffle all the possible sequences and then group them into batches as opposed to the original model which just randomly picks batches with replacement from all of the possible sequences.

This model also achieved a similar perplexity as the original.

3 Meeting Your Milestone

It took longer than expected to get Caffe Char-RNN working, but now that it is working I think that we can get back on track by the end of the week.

4 Surprises

There were not really any major surprises besides the one described above.

5 Looking Ahead

With a Caffe Char-RNN implemented, we will now need to see how it fairs in a distributed setting. We will first look at the IID case, where we will just split the text into contiguous chunks. For the non-IID case, we will take the same approach that FedAvg Paper and create a non-IID partition by separating the text by character roles.

In addition, I still need to look into the theoretical analysis. Kevin is also currently working on some solutions, and we are primarily exploring a reinforcement learning solution to dynamically tune the mirror update threshold, which determines how often updates are pushed to the global model.

6 Revisions to Your Future Milestones

There are no revision to future milestones.

7 Resources Needed

I have all the resources that I need to complete my project.