Homework 1 Report

Environment:

macos 10.14 mojave
CSIE workstation

Execution:

$ make
$ ./train <iteration> <model_init.txt> <seq_model_XX.txt> <model_XX.txt>
(Do it for 5 models)
$ ./test modellist.txt <testing_data1.txt> <result1.txt>
$ ./test modellist.txt <testing_data2.txt> <result2.txt>

Experiments:

I did some experiments with different batch size per iteration and different iterations. The batch size is the number of the sequences for the program to update the parameters of the model. And the iteration is the total times that the program run through all the sequences.

Minimum batch size required to train the HMM models

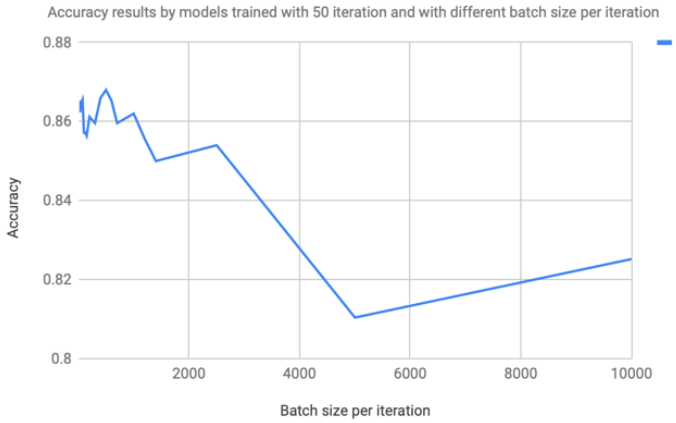
When I tested my program for the first time, I got 'nan' in the parameters of the HMM models. Thus, I checked my program to see if there were any bug, then I found out that the sequences must be trained in batches in order to get the correct results. So I did an experiment on the minimum batch size required to train the HMM models. In my experiment, I trained the models with different batch sizes and with 50 iterations.

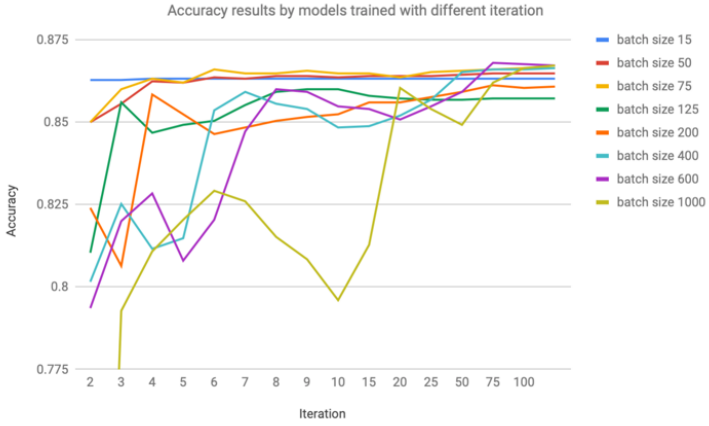
batch		iter		result	time
5		50		x	x
10		50		x	x
12		50		x	x
13		50		x	x
14		50		x	x
15		50		0.8632	3:11
20		50		0.8624	3:16

As you can see, if the batch size if less than 15, 'nan' will occur in the parameters of the models. Thus, the minimum batch size required to train the HMM models is 15 sequences.

The effect of batch size on the accuracy

As you can see in the left figure, with the same amount of iteration, the accuracy is better when the models are trained with smaller batch size. The result of the right figure tells us that the accuracy is bounded by about 0.868. An interesting result is that training with smaller batch size is more efficient than training with larger batch size, since models trained with smaller batch size has great accuracy even with just 1 or 2 iterations.





The training time with different batch size and different iteration

It is pretty obvious that the more iteration it takes to train the models, the longer time it takes to finish the training. The right figure below is coherent with our intuition. On the other hand, as you can see in the left figure below, models are trained faster with smaller batch size. As a result, if we are training HMM models, we should train them with smaller batch size, this will let us both save time and get a great accuracy.