Journal Report 13
1/6/20-1/12/20
Emily Ye
Computer Systems Research Lab
Period 2, White

Daily Log

Monday January 6

Began rereading Ramsundar et al. paper on multitask networks for drug discovery

Tuesday January 7

Finished rereading Ramsundar et al. paper on multitask networks for drug discovery

Began reviewing Sohom's code in QM9GNN2Multitask (making and training the multitask networks)

Thursday January 9

Finished reviewing Sohom's code in QM9GNN2Multitask and compared to my code for making and training GCNs

Reviewed Sohom's code in CalculateTransferCoeffs and Demo

Timeline

December 15	Merge code for reading in molecules with code for model with loaded-in weights and test winter goal demo	Mostly done with demo, with a little debugging left to do
December 22	Finalize demo and have it ready by December 20 (before winter break)	Successfully finished winter goal demo to show time/accuracy comparison of GCN with DFT
January 12	Review multitask papers and Sohom's multitask code in preparation for integrating the multitask and GCN networks	Reread multitask papers and read through Sohom's code
January 19	Implement clustering of single-task properties on GCN	
January 26	Finish combining multitask and GCN networks and begin testing and tweaking to improve accuracy	

Reflection

After discussing with Sohom, we decided that for the next few weeks, I would focus on combining my graph convolutional network code with his multitask code while he looks into modeling multitask representation learning using tensor factorization. Therefore, I spent this week reviewing a paper on multitask learning, as well as Sohom's code from the winter goal. This helped me refamiliarize myself with multitask neural networks after spending the past few months just working on GCNs.

For the upcoming weeks, I will work on integrating the two networks. Hopefully, I will be finished by the end of the month so that we can test the multitask GCN network and tweak features to improve its accuracy. Looking forward to next semester, we are hoping to continue tweaking the multitask GCN network, while working on allowing users to draw the molecule they want to predict properties for (as opposed to digging through QM9 files, which are inconveniently labeled with ID numbers, like we are doing right now).