Journal Report 12

12/9/19-12/13/19
Sohom Paul
Computer Systems Research Lab
Period 1, White

Daily Log

Monday, December 9

Started clustering by thresholding cross-task transfer coefficients.

Tuesday, December 10

Used a greedy approach to select which sets of 3 tasks had the most similarity (taking the L2 norm of each of the vector of transfer coefficients).

Thursday, December 12

Started training multitask networks in anticipation of demonstration.
Wrote code to predict properties using saved models.

Timeline

Date	Goal	Met
Nov 25	Finish performing neural clustering. Generate 19x19 table of task relatedness. Start building demonstration program with DFT and single-task results.	Successful builds of DFT, partial completion of code to cluster tasks.
Dec 2	Continue building demonstration with preliminary multitask results.	Yes.
Dec 9	Cluster tasks. Train multitask network. Finish building demonstration with multitask and DFT methods.	Yes.
Dec 16	Consolidate demonstration code into single file. Final testing of my demonstration.	
Winter Goal	Have a program capable of taking an input molecule and predicting the relevant molecular properties, using a choice of density functional theory or multitask-learned networks. The demo will let the use see firsthand the relative accuracies and speeds of the methods.	

Reflection

My decision to cluster my tasks into groups of 3 was more or less ad hoc, based on empirical evidence from earlier in the year. When looking at the cross transfer coefficients, I noticed that some tasks tended to transfer better than others; for example, calculation of the largest rotational constant (A) seemed to transfer onto virtually every other task. That is why I chose to use a greedy approach to selecting clusters, by putting the 3 tasks that were most correlated into one cluster, then calculating which 3 tasks of the remaining were most correlated and so on. Even though A and any other task can be learned simultaneously well, that doesn't mean we should train all the tasks together, because the other tasks will have less correlation. I also calculated the "goodness" of each set of 3 by taking an L2 norm as to prefer clusters were all 3 tasks transferred well to each other, rather than have one pair transfer really poorly while the other pairs transfer well. My final clusters are as follows:

```
A B alpha
C r2 u0
zpve g298 cv
lumo u298 h298
mu homo
```

I chose to leave out gap because it's just lumo - homo, and I left out the atomization energies because those can't be calculated using our DFT demo (and, to be honest, they're not that useful to actual chemists). For next week, I just need to integrate the code I've written into a single runnable demonstration file, and test it to see if the results are reasonable.