Journal Report 13
1/6/20-1/10/20
Sohom Paul
Computer Systems Research Lab
Period 1, White

Daily Log

Monday, December 16-Thursday, December 19

Ran final tests in preparation for presenting demonstration.

Monday, January 6

Started refactoring code to allow for better code integration and to remove code duplication.

Tuesday, January 7

Continued code refactoring.
Started reading paper on deep relationship networks[1].

Thursday, January 9

Continued with above paper, and started reading paper on deep representation learning using tensor factorization[2].
Started building toy networks using tensor factorization.

[1] https://arxiv.org/abs/1506.02117
[2] https://arxiv.org/abs/1605.06391

Timeline

Date	Goal	Met
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.	Yes.
Jan 13	Start investigating deep relationship networks.	Partial; have not yet integrated these new architectures into our project.
Jan 20	Complete investigation of deep relationship networks. Start building models for chemical property prediction.	
Jan 27	Complete deep relationship networks. Compare to naive models.	

Reflection

Out of a sense of duty towards my partner (and to reduce the amount of work she would need to spend making sense of my messy code as she works on integrating our current progress), I spent Monday and part of class on Tuesday working to refactor my code, which mostly consisted of removing function name collisions and importing functions and objects from sibling files rather than duplicating the code defining each object. For the rest of the week, I began to investigate the applications of deep relationship networks for multitask learning, which had been one of my goals prior to winter break before I committed myself to completing the code demonstration.

Otherwise, I did significantly less coding this week than I usually do because of the mathematical content presented in these papers on tensorial techniques for deep multitask learning. One of the issues that was bothering me prior to winter break was how we were going to extend our networks to deeper learning without having the multitask architecture totally ad hoc. For hard parameter sharing, we arbitrarily set number of layers to be shared, and soft parameter sharing additionally requires the punishment weight to be set. There is no good way for me to determine whether I'm properly leveraging the hidden task relations without introducing negative transfer biases, especially because these parameters would ideally be fine-tuned for each task pair, because certain pairs of tasks are more related than others. The two papers I'm currently reading claim that by stacking the model parameters from between tasks into rank-3 tensors, we can allegedly dynamically learn task relations and leverage these relations at each layer in the network. The first paper, Deep Multi-Task Representation Learning: A Tensor Factorisation Approach by Yang and Hospedales, 2017, takes this tensor and decomposes it in task-specific and shared factors. However, I do not yet have intuition for why we expect this to work, but their reported improvements for gender/age prediction given facial data are intriguing. I've been working on developing my understanding by analyzing the implementation posted on Github and toying with these models. The other paper I attempted to read, Learning Multiple Tasks with Multilinear Relationship Networks by Long et al. 2017, seems to use the same ideas as the first paper and builds upon their work by considering tensor normal priors for inference. I do not yet understand this paper, and I'm not yet certain it will be applicable to our research.