**From:** **Tiago Azevedo** tiago.azevedo@cst.cam.ac.uk
**Subject:** Re: Follow-up from the skype talk
**Date:** 8 August 2019 at 20:18
**To:** Kamile Stankeviciute ks830@cam.ac.uk
**Cc:** Pietro Lio' pl219@cam.ac.uk

Hi Kamile,

Ok, let's wait for when you are back then, it's just a few days indeed. Shall we try to skype monday or tuesday? I will be in Portugal and can talk at 2.30pm or 9:15pm uk time. Would this be a possibility for you?

Regarding the ideas, what do you think about the paper I sent to you some time ago? https://www.sciencedirect.com/science/article/pii/S1361841518303554 They do quite a nice job comparing different approaches for the creation of the graph.

From HCP you could get different imaging modalities (not genetic I'm afraid). So you could have node features from structural images, and the edges constructed based on the functional data (eg. the graph kernels I mentioned which are not used in the paper). A project in which definitely you need to implement many parts so the software development part that it is required for Part II projects would still be there. The problem with HCP is that you'd need to be regressing out a variable instead of classification, which usually are more difficult to have a good evaluation metric.

Maybe applying a similar method in PPMI (Parkinson's) could be better. In PPMI we could use both structural images and genetic data for creating a graph, and maybe a semi-supervised approach to account for the people who don't have functional data. Advantage here is that, similarly to AD, we know the classification will have at least a reasonably evaluation metric and classifications is usually easier to analyse and explain (and in a way is a safer option for you).

An interesting thing from a CS perspective is to try to use other methods (eg a list with a lot of those implemented: https://github.com/rusty1s/pytorch_geometric), and maybe do something the paper didn't do: try to interpret what the model is learning. Just throwing ideas out loud here, but the attention mechanisms have been quite nice for interpreting image classifications, and maybe can be successfully used in graphs.

All these options are interesting from a research point of view, and also allow for a nice part II project where you need to code a lot, and make nice evaluation metrics between different possibilities of graph construction. It can also be nice for you as a computer scientist as it's an opportunity for you to learn quite a recent paradigm in ML/DL.

Have a thought, I will try to think more about reasonable "multi modal projects" as well.

Cheers,
Tiago.

On 2019-08-07 18:27, Kamile Stankeviciute wrote:

> Hi Tiago,
> I could still look at multimodality in other datasets (e.g. the human connectome or the autism-related one) or we could discuss the gene expression dataset which seems quite vast and interesting.
> It makes it a bit harder to tell whether I want to go with multimodality or just training training as well because I haven't explored the datasets themselves enough (and don't really have the experience) to reasonably know what is feasible with them them—which could easily result in some infeasible ideas. A further practicality is that the success of the model most of the time depends on a good dataset, and one of the aims, of course, is a successful model (for a more positive evaluation section)—I'd favour a better dataset/question combination over a particular dataset or a particular question (especially without the ability to collect my own data :)).
> Tl;dr I'm still open to most of the ideas, but more leaning more towards those applied (to medical datasets) rather than theoretical (only using the dataset as a benchmark). And maybe more towards multimodality rather than training as the former seems to be a less popular idea and—as you mentioned—more extensible to an applied paper.
> I'm in Pittsburgh, PA which is 5 hours behind London time. I will be back in Zurich (1 hour ahead) on Saturday afternoon. I should be available on Friday morning/afternoon (which would be the afternoon/evening in Cambridge) though.
> With best wishes,
> Kamile
>
> > On 7 08 2019, at 11:36, Tiago Azevedo <tiago.azevedo@cst.cam.ac.uk> wrote:
> > Hi Kamile,
> > Yes, it's important you focus the objective of the project into the implementation/coding of a methodology, with the correct evaluation chapter as well.
> > Regarding the ideas in my homepage, they are from last year, and both were developed into a part II project. One thing though: the project on Parkinson's multimodal data was changed because the amount of fMRI data that they have in PPMI is