Graph Analysis Project Write Up

Alex Owen, Kavya Tumkur, Bronte Wen

scienceinthe.cloud demo

The demo worked for the the most part and produced eight graphs - betweenness, clustering coefficient, degree (log scale), edge weight (log scale), dimension (vs. eigenvalue), number of non-zeroes, locality statistic-1 (log scale), and dimension (vs. portion of total variance). These appear to work as expected. However, this isn't very reproducible - the only code that we could really use as reference was for visualizing the connectome as a matrix, though this piece of code was essentially given in the notebook. The other pieces of code for the pipeline were not given and were not very reproducible. This demo is useful for seeing what the results might look like for a given connectome, but isn't very useful in case we want to use it as an example to make our own connectomes and analyze them.

Covariate interpretations

We used the covariates age at scan, session, season, resting state, and time of day to build our connectome graphs. The age at scan was the age of the subjects when their brains were scanned. We know that adults have more developed brains, and so we hypothesize that their connectomes will be more intricate (more connections or clusters of connections). This is true at least until they get older. These subjects range from the age of 17-27, which we split into groups of < 20 years old, 20-23 years old, and older than 23 years old. Since these subjects are mostly young, we found that the connectomes don't vary too much. The season is the season at which the subjects' brains are scanned. During the colder seasons (fall, winter), the brain will be less active. During the warmer seasons (spring, summer) the brain will be more active and the

connectome will consequently be more intricate. The session variable described whether the session was a first test or a retest. The retests did not have data for it, however, and so only the baseline results were graphed. We hypothesize that if the retest data was given, the connectome graphs would be very similar, since they are of the same brain. The resting state variable measures whether or not the subject is resting. When the subjects are resting the brains will be less active, and when the subjects are not resting the brains will be more active, making the connectomes more intricate. The time of day variable is the time of the day at which the subjects' brains were scanned. We hypothesize that people's brains would be most active during the middle of the day and less active when they've just woken up or slept. Therefore we predict that the connectomes for the morning and night will be more similar to each other than they will be to the connectome in the middle of the day

What we learned

We learned that there is a notion of a generalized connectome for humans. The data we were given exemplified this based on the mean connectome that we calculated. There may be more complicated and robust ways to perform this, though. We also learned that for the data set given, there were noticeable differences between the mean connectome for men and women. This is one piece of evidence that there can be an algorithm to determine the sex of a subject based upon their brain. Finally, we learned that age does have an impact on sex classification, as adding it as a feature increased our classification score by about 5%.