Pre-seminar questions - 9/7/20

Dr. Zhang - Studying Age Prediction by DNA Methylation in Neural Network

1. For the datasets used, how large of a difference in output values actually results from different models' MAEs? For instance, 3.7 versus 3.8, or 3.8 versus 2.8.

2. Were there any CpGs that had a more complicated methylation pattern, such as alternating between methylation and demethylation at different life stages still associated with age? How did the model weigh those data to predict age?

3. What on table 5 led you to the conclusion that the CPFNN model more accurately predicted age for Down syndrome? The Horvath and Hannum models have lower p-values and the Hannum model has lower standard deviation relative to the mean for both the control and the patient.

4. What additional work would this model need to predict age-related illnesses before they onset?

5. The model performed best with the dataset with mostly ages below 30, and worst with the ages over 70. Does the dataset being distributed differently from the training datasets have an effect, or in the latter case, is it only because (as you say in the paper) the model hadn't seen many ages over 70?