Predictive Analytics Competition (PAC) 2021 Challenge

Using structural neuroimaging data, improving brain age prediction

This semester, we invite students from SWCON253 Machine Learning class to develop a model predicting brain age from healthy individuals based on structural Magnetic Resonance Imaging (sMRI) data (n=481).

All students are asked to engineer a machine learning model using their expertise and fully processed data as described in the paper. Each student has the opportunity to upload predictions for a given test dataset and we will evaluate model performance by comparing the uploaded predictions to actual chronological age for each individual. This semester, there are two distinct objectives:

1. The student submitting the model with the smallest Mean Absolute Error (MAE) for the test datasets (IXI test dataset; n=100 & COBRE test dataset; n=72) while keeping the highest Pearson correlation between predicted age (brain age) and chronological (actual) age will win this semester's PAC Best Model Award.

2. The student submitting the model with the smallest Mean Absolute Error for the test dataset while keeping the Spearman correlation between brain predicted age difference (brain-PAD) and chronological age below r = 0.10 will win this semester's PAC Bias Reduction Award.

As we find the resulting performances to be an interesting benchmark as well as the chosen approaches to be exciting opportunities for inspiration, we would like to finish the Predictive Analytics Competition with a presentation session at **the date of the final exam (June 22nd)**. In addition, both the number of students (right now being > 46 students) as well as the number of training data are sizable so that we expect the outcome to be of major interest for the general machine learning community using neuroimaging data.

Main organizer
Won Hee Lee

Deadline
13:29 pm, June 17th

Presentation session
Top 3 (or 5) winners will be selected and then these students will be invited to present their PAC Best Model at the date of the final exam (June 22nd), 13:30pm
- Top 3 (or 5) winners will be contacted to get ready for presentation (slides)

Additional Context

The brain changes as we age, and these changes are associated with cognitive decline, neurodegenerative disease and dementia. Although brain aging is universal, rates of brain aging differ markedly; some people suffer cognitive decline in later middle-adulthood, while others remains cognitively normal into their tenth decade. The process of brain aging includes morphological and functional changes to the brain, which can be assessed using neuroimaging. This raises the possibility that the variability in brain aging can be measured, and research has focused on developing such a neuroimaging biomarker of brain aging; the so-called 'brain-age' paradigm.

The idea with brain-age is that if statistical models can be developed to accurately predict chronological age in healthy people (using neuroimaging data), then the apparent age of a new individual's brain can be calculated. Where someone's brain-age is older than their real age, this is thought to reflect poorer brain health, relative to their age. Older-appearing brains have been associated with psychiatric and neurological diseases, with greater risk of developing dementia and a shorter lifespan. Younger-appearing brains have been found in people who exercise more, have greater years of education, meditate or play musical instruments.

The hope is that brain-age can provide a sensitive, if unspecific, global measure of brain health, that could be used in many contexts. These include clinical trials of neuroprotective therapies, screening groups of people at-risk of poorer cognitive aging, and providing mechanistic insights into the downstream consequences of different diseases.

Critical to the success of brain-age models, is the accuracy of the healthy training model. Hence, the goal of this semester's PAC is to build the most accurate model, using the training data supplied. Specifically, we would like to minimize brain predicted age difference (brain-PAD) which is calculated as brain-predicted age minus chronological age.