

Code Challenge

This is a simple programming challenge as an alternative to a code sample designed to give you an opportunity to demonstrate your capabilities to tackle an ad-hoc data analysis problem. This is not expected to be onerous nor is the code expected to be of production quality.

Write a small program to do simple analysis on several MRI images.

Language

- Python3/R is preferred, but any other commonly used language is acceptable.
- List all required third party modules or plugins (`requirements.txt` file or similar).

Data Description

MRI data from one subject at two different test sessions, stored as bzip2 compressed nifti files. Each data file is organized into 182 slices in z, 31 time points, 182 pixels in x and 218 pixels in y. The first time point is a high contrast image prior to testing. We will use Python indexing to refer to the data.

Download:

https://synaptomes2.neurodata.io/data/public/neurodata-job/sub-0025865_dwi_aligned.tar.bz2

Problem

1. Scale the data (window) to allow easy viewing of the data and plot time=5 and z slice=15 from each session (and explain why you chose the levels that you did)
2. Crop out a region of interest (brain area) and plot that region for each session
3. Choose a threshold (intensity value) at slice 15 and return the number of pixels above that threshold for each session.
4. Using the same threshold, perform some simple first order analyses of that region over time
5. Building on the previous steps, train a random forest classifier to differentiate between sessions (we expect performance will be near zero). Please report:
 - out-of-bag error
 - plot the distribution of class conditional posteriors estimated random forest

Send us the code and results in a convenient format, e.g., an exported Jupyter notebook, or standalone code in a zip file with the image output and text results.

Suggestions

- nibabel (<http://nipy.org/nibabel/>) in python will read nifti images into numpy arrays
- [Sklearn](#) for random forest in Python