UNIVERSITEIT GENT

FACULTEIT PSYCHOLOGIE EN PEDAGOGISCHE WETENSCHAPPEN

Classification of fMRI data

Roma Siugzdaite & Ruth Seurinck

Department of Data-Analysis

April 17th, 2014

Outline

1. Toolboxes
2. Tutorial: Pereira, Neuroimage 2009
   - Creating and selecting features
   - Pre-processing features
   - Choosing a (classification) algorithm
   - Cross-validation
   - Assessing the results
3. Scenarios
   - Haxby Data (Pronto manual)
   - IXI Data (Pronto manual)
   - Ishai Data (Beta images)
   - Motion After Effect (Beta images)

- Multi Voxel Pattern Analysis: MVPA
  - requires Matlab (OctaveMVPA available)
  - imports images from SPM, AFNI and Brainvoyager
  - https://code.google.com/p/princeton-mvpa-toolbox/
- pyMVPA
  - runs on Python
  - includes a searchlight (moves across the brain to calculate local multivariate information content)
  - http://www.pymvpa.org/
- Functional Real-time Interactive Endogenous Neuromodulation and Decoding: FRIEND
  - FSL toolbox
  - reference paper: Sato et al., PLOS one 2013
  - http://www.nitrc.org/projects/friend/

- Pattern Recognition for Neuroimaging Toolbox: PRoNTo (what we will use)
  - requires Matlab and SPM8
  - reference paper: Schrouff et al., Neuroinformatics 2013
  - http://www.mlnl.cs.ucl.ac.uk/pronto/

- Two specials:
  1. Representational Similarity Analysis: RSA http://www.mrc-cbu.cam.ac.uk/methods-and-resources/toolboxes/
  2. Cortical surface-based searchlight decoding (Hayneslab, Berlin) https://sites.google.com/site/hayneslab/links/

General principle

Infer variable from multivariate brain information

1. categorical variable: CLASSIFICATION
   (Scenario 1: Did the subject see houses or faces?)
2. scale variable: REGRESSION
   (Scenario 2: Does the brain contain information that can predict age?)

Feature Selection

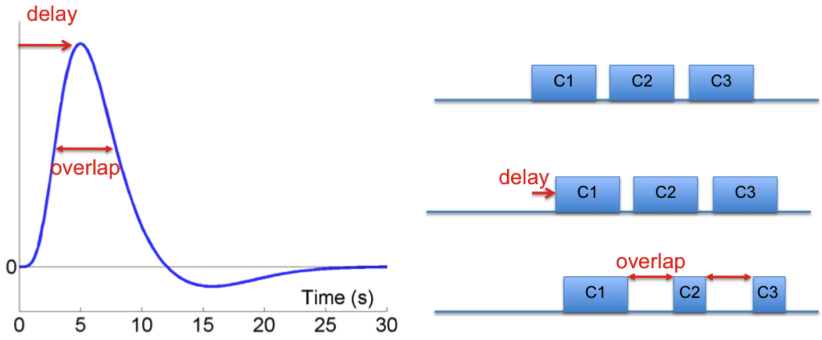
Selecting informative voxels and isolating relevant information:

- use a brain mask to isolate only voxels within the brain
- further feature selection in PRoNTo: use a priori mask image (e.g. only fusiform gyrus)
- classification: association with class labels? (e.g. faces or houses)

Associating scans with classes

1. use the entire timeseries of the experiment: what scans should be associated with what event?
   - HRF-correction (delay and overlap)
   - block design works well, fast event-related design does not
   - choose scans that coincide with event of interest
   - overlapping scans should be omitted, given sluggish HRF

Creating and selecting features



1. use the entire timeseries of the experiment: what scans should be associated with what event?
2. use Beta-images from a regular GLM analysis you perform prior to classification
   - recommended procedure for fast-event related design
   - results depend on the goodness-of-fit of the univariate GLM

Pre-processing features

- not the same as pre-processing images prior to GLM!
- adapting features/voxels to make them comparable across the brain
- independently for training and test data (see cross-validation)
- different possibilities in PRoNTo:
  minimum requirement is to mean centre the data

Algorithms

PRoNTo has several algorithms for both regression and classification

1. regression: for the moment only if you have one image per subject and do a group analysis
2. classification: can be performed both at the individual and the group level

Cross-validation

- Dividing the data in n folds (e.g. runs, subjects, ...).
- Use $n-1$ folds as training data, the remaining fold is the test data
- Rotate which fold is the test data.
- Result for each fold and an overall result

Is my result statistically signficant?

We focus on the classification:

- we get a classification accuracy: % events that has been classified correctly
- PRoNTo:
  - regular accuracy (each category has equal frequencies)
  - balanced accuracy (corrected accuracy if one category occurs more)
  - an accuracy for each class/category
  - a permutation test for the accuracy with a p-value
- classification accuracies for each subject can be combined at the group-level in a regular test (e.g. t-test, non-parametric test, ...) compared to chance level (e.g. 50 % for two-class problem)

Locating informative voxels: the weights map

If your algorithm uses a linear model, you can display the weights, or the relative amount of information a voxel conveys, in a map.

You can never focus on certain peaks and coordinates, as it is a relative contribution within a multivariate analysis!

Single-subject Block Design

- one subject
- viewing 8 categories: faces, houses, cats, chairs, bottles, scissors, shoes and scrambled pictures
  limited to two-class problem of houses and faces
- block design, entire timeseries available
- 12 runs, runs used as folds
- Leave-One-Block-Out cross-validation (run = block)

Regression example

- one group of 102 subjects
- one anatomical image and the age per subject
- subjects used as folds
- Leave-One-Subject-Out cross-validation
- correlation and MeanSquaredError (MSE) as output measure

What to do with a classification problem in a fast-event related design?

- Ishai et al., Journal of Cognitive Neuroscience 2000 (could be analysed as a block design as well)
- one subject, three classes: faces, houses, chairs and their scrambled pictures
- 12 runs, only a Beta-image for each category and each run
- for a two-class problem: faces versus houses
  - model two groups, one with the 12 Beta-images of the face category, one with the 12 Beta-images of the house category
  - make sure the images are in the same order across groups!
  - Leave-One-Subject-per-Group-Out cross-validation (run = subject)

What if you have only one run per subject?

- Seurinck et al., Journal of Cognitive Neuroscience 2011 (could be analysed as a block design as well)
- motion after effect localizer: rotating disk with continuous direction (MAE) and rotating disk with alternating direction (noMAE)
- 13 subjects, only a Beta-image for each category (MAE and noMAE)
- for a two-class problem: MAE versus noMAE
  - model two groups, one with the 13 Beta-images of the MAE category, one with the 13 Beta-images of the noMAE category
  - make sure the images are in the same order across groups!
  - Leave-One-Subject-per-Group-Out cross-validation (subject = subject)

Let's try...