Poster prize winner at the annual meeting of the German Society for Psychophysiology and its Application Magdeburg, 2008

PyMVPA: A Python toolbox for classifier-based data analysis

Michael Hanke[1], Yaroslav O. Halchenko[2], Per B. Sederberg[3], Stephen José Hanson[2], James V. Haxby[4], Stefan Pollmann[1]

1) Department of Psychology, University of Magdeburg, Germany
2) Psychology Department, Rutgers Newark, New Jersey, USA
3) Department of Psychology, Princeton University, New Jersey, USA
4) Department of Psychological and Brain Sciences, Dartmouth College, New Hampshire, USA

Classifier-based analysis: Where is the software?

Advantages of classifier-based analysis:

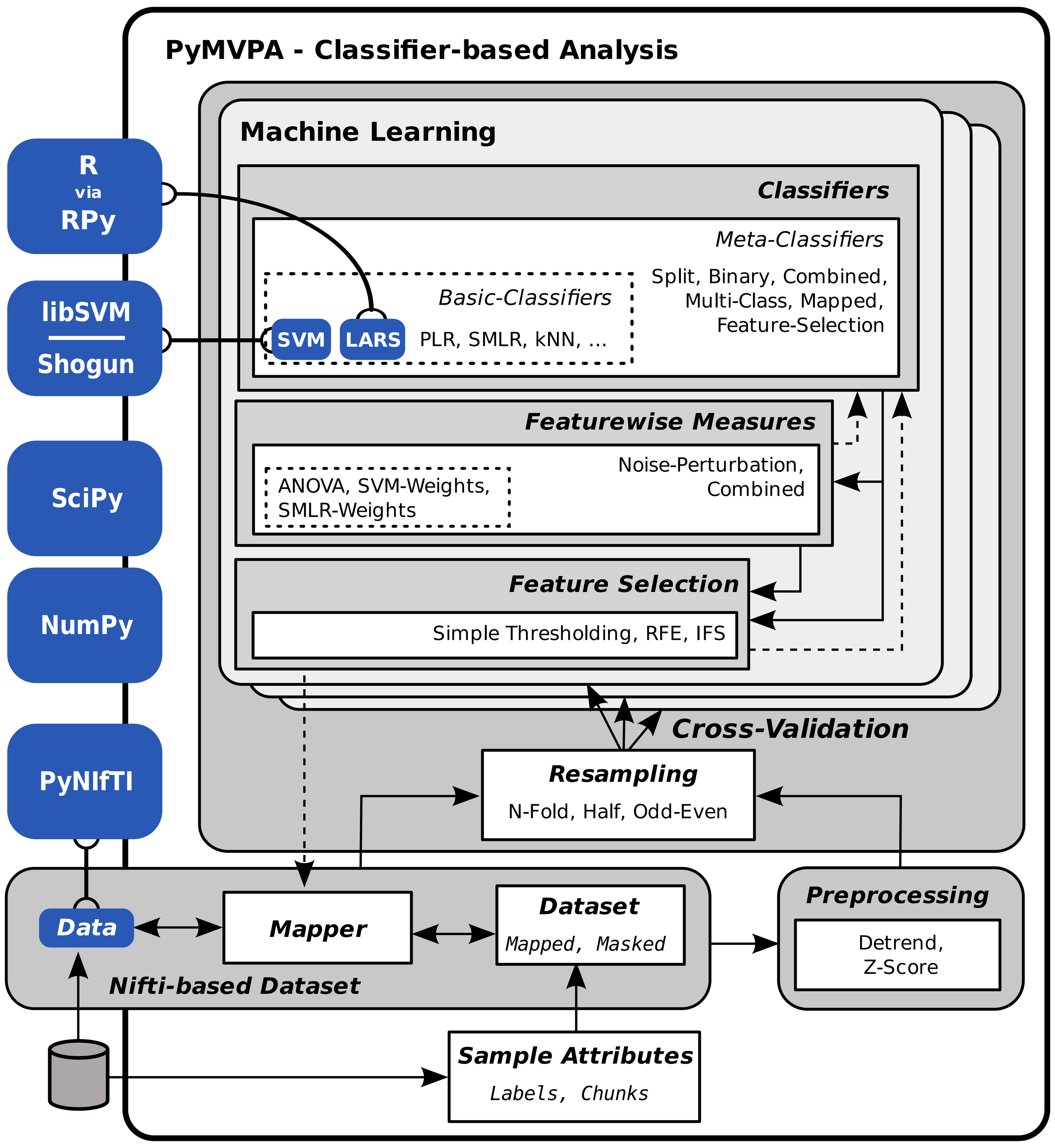
- Multivariate method
- Can focus on distributed patterns instead activation foci
- Direct quantifyable link between neuroimaging data and experimental manipulation (O'Toole et al, 2007)
- Suitable for high-resolution fMRI imaging (Kriegeskorte et. al, 2007)

Increasing number of publications applying classifier-based analysis to neuroimaging data (e.g. Haxby et al., 2001; Kamitani & Tong, 2005; Hanson et al., 2004; Haynes & Rees, 2007; Hanson & Halchenko, 2008), but only few software packages to facilitate this type of analysis, which are available to a broad audience (*3dsvm*, LaConte et al., 2005; *MVPA toolbox*, Detre et al., 2006).
But: wealth of machine learning software (www.mloss.org)

There is a need for a unifying framework to bridge between establish neuroimaging and machine learning software.

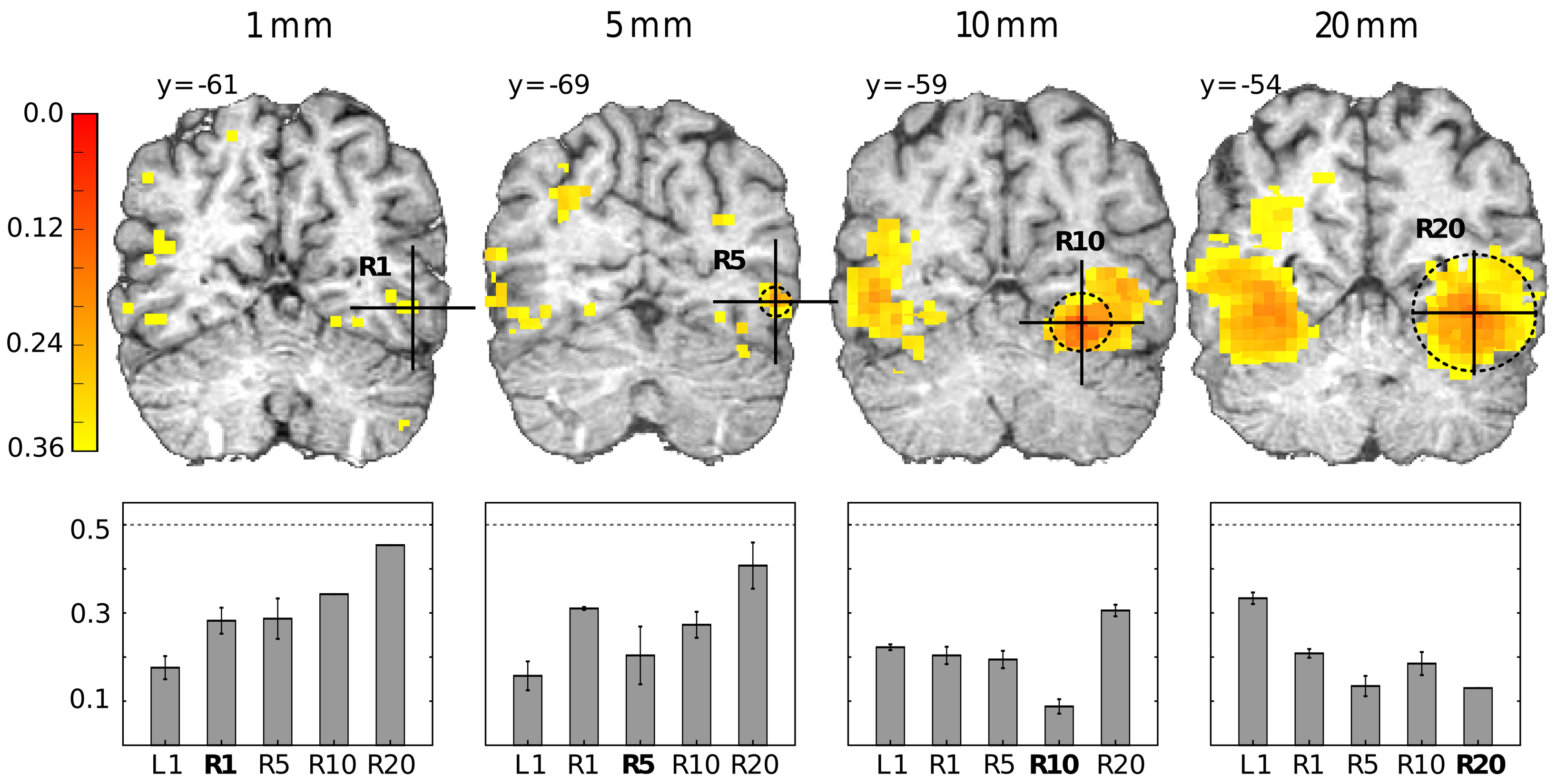
PyMVPA Features

- User-centered programmability with a intuitive user interface: Object-oriented toolbox design leading to readable and verifyable code.
- Extensibility: Modular interface to connect extensions in multiple programming languages.
- Transparent reading and writing of datasets: NIfTI support for input and output and additional generic support of various binary and plain text format.
- Portability: Should run on anything supported by Python.
- Open source software: MIT-licensed free software.



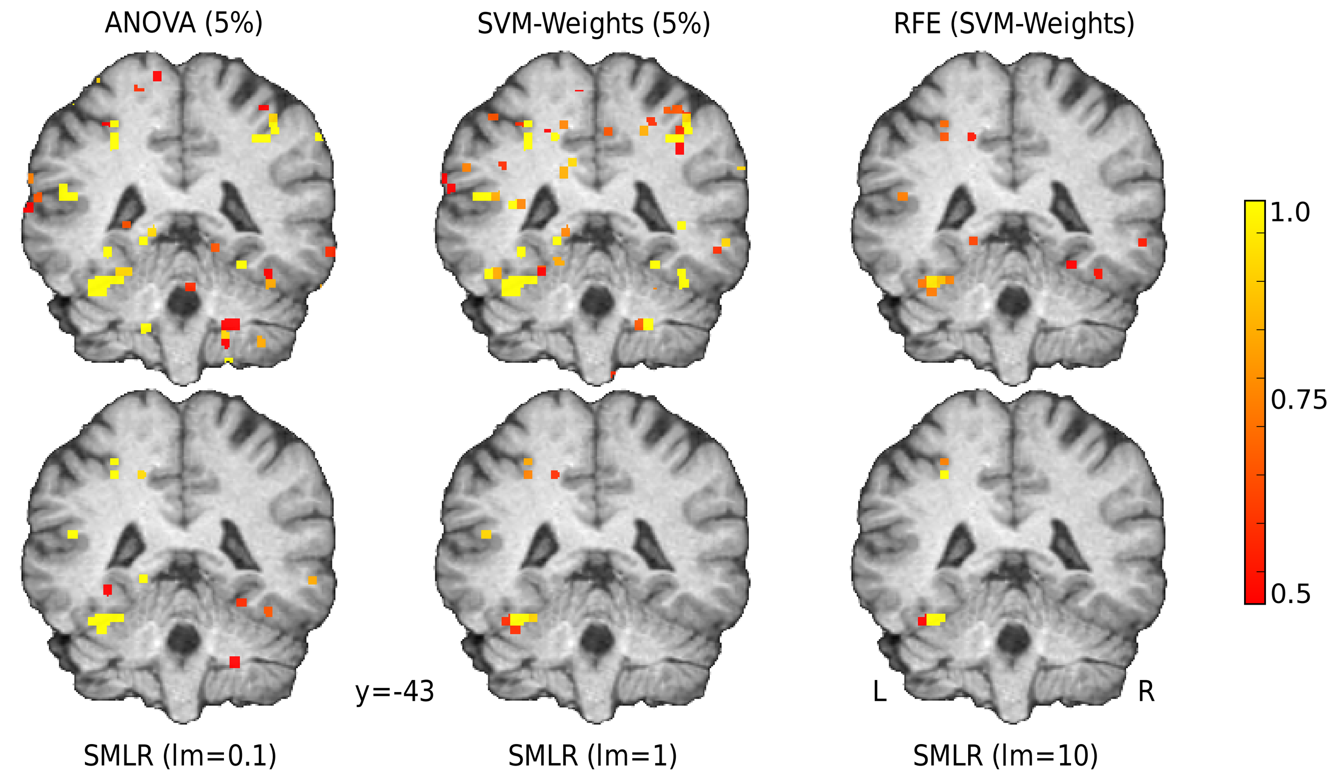
Example: Multiple ROI analysis with a Searchlight

```
attr = SampleAttributes('sample_attr_filename.txt')
dataset = NiftiDataset(
              samples='subj1_bold.nii.gz',
              labels=attr.labels, chunks=attr.chunks,
              mask='subj1_roi_mask.nii.gz')
cv = CrossValidatedTransferError(
        TransferError(LinearCSVMC()), OddEvenSplitter())
sl = Searchlight(cv, radius=5)
sl_map = sl(dataset)
dataset.map2Nifti(array(sl_map)).save('slight_5mm.nii.gz')
```



Example: Flexible feature selection

```
attr = SampleAttributes('sample_attr_filename.txt')
dataset = NiftiDataset(samples='subj1_bold.nii.gz',
                       labels=attr.labels, chunks=attr.chunks)
clf = LinearCSVMC()
clf = FeatureSelectionClassifier( clf,
      SensitivityBasedFeatureSelection( OneWayAnova(),
         FractionTailSelector(0.05,  mode='select')))
cv = CrossValidatedTransferError( TransferError(clf),
        NFoldSplitter(), enable_states=['confusion'])
error = cv(dataset)
```



CATS vs. SCISSORS classification. Data from one participant from Haxby et al., 2001.

The PyMVPA toolbox is available at

http://pkg-exppsy.alioth.debian.org/pymvpa/

PyMVPA mailing list

pkg-exppsy-pymvpa@lists.alioth.debian.org

References

Detre, G., Polyn, S. M., Moore, C., Natu, V., Singer, B., Cohen, J., Haxby, J. V., Norman, K. A. (2006). The multi-voxel pattern analysis (MVPA) toolbox. Poster presented at the Annual Meeting of the Organization for Human Brain Mapping (Florence, Italy).
Hanson, S., Matsuka, T., Haxby, J. (2004). Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: Is there a "face" area? Neuroimage, 23, 156–166.
Hanson, S. J., Halchenko, Y. O. (2008). Brain reading using full brain support vector machines for object recognition: there is no "face" identification area. Neural Computation, 20, 486–503.
Haxby, J., Gobbini, M., Furey, M., Ishai, A., Schouten, J., Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science, 293, 2425–2430.
Haynes, J.-D., Sakai, K., Rees, G., Gilbert, S., Frith, C., Passingham, R. E. (2007). Reading hidden intentions in the human brain. Current Biology, 17, 323–328.
Kriegeskorte, N., Bandettini, P. (2007). Analyzing for information, not activation, to exploit high-resolution fMRI. Neuroimage, 38, 649–662.
LaConte, S., Strother, S., Cherkassky, V., Anderson, J., Hu, X. (2005). Support vector machines for temporal classification of block design fMRI data. Neuroimage 26, 317–329.
O'Toole, A. J., Jiang, F., Abdi, H., Penard, N., Dunlop, J. P., Parent, M. A. (2007). Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. Journal of Cognitive Neuroscience, 19, 1735–1752.