Resources for practicing PR4NI pragmatic cursory overview

Yaroslav O. Halchenko

Center for Open Neuroscience, Center for Cognitive Neuroscience, Dartmouth College, USA

Geneva, Switzerland 2016

http://www.pymvpa.org

http://NeuroDebian.net

http://datalad.org

http://duecredit.org

Visit our DataLad/NeuroDebian exhibit table and posters #1855, #1870

Acknowledgements

What is the most valuable "Resource"?
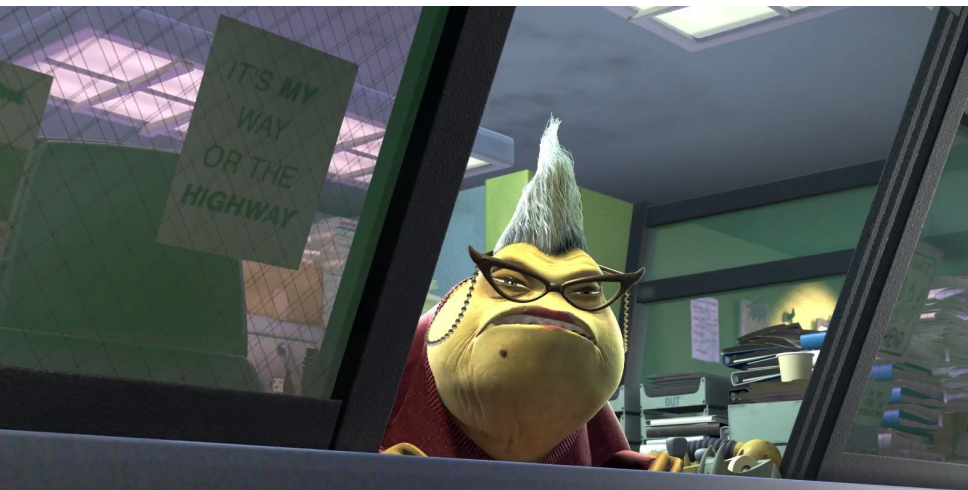
What is the most valuable "Resource"? Humans
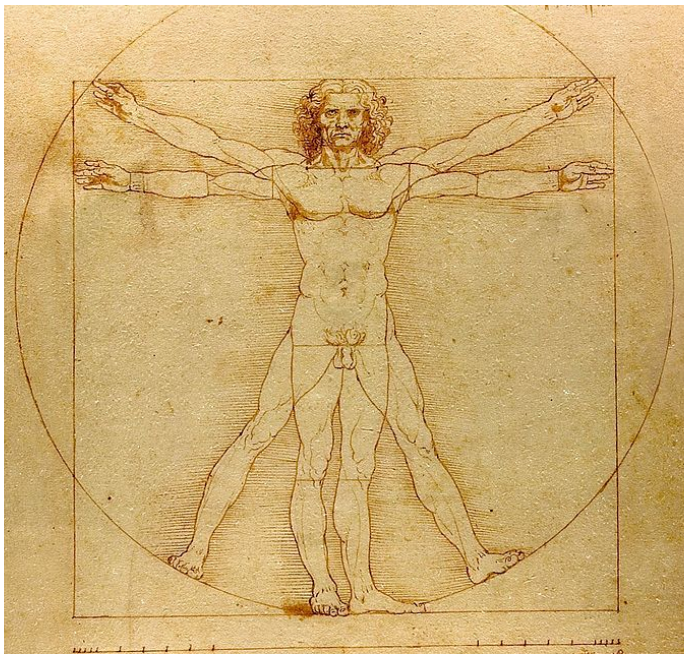
Humans virtue #1

- no "collective mind" (yet, read Nexus trilogy?)
- condensed presentation of articles is rarely sufficient and not reproducible as such

- no "collective mind" (yet, read Nexus trilogy?)
- condensed presentation of articles is rarely sufficient and not reproducible as such
- we need to work together
- share and contribute back

Humans virtue #2

- no time warping machine (yet)
- hard to stretch time: 24x7 is kinda fixed

Humans virtue #2

- no time warping machine (yet)
- hard to stretch time: 24x7 is kinda fixed
- we need to avoid repeating ourselves or others twice
- automate and stay efficient

Humans virtue #3

- impossible to clone "your own Vapnik" (yet or ever)
- condensed presentation of articles is rarely sufficient and not reproducible as such
- we think we (or students) can do the same thing

Halchenko, Y. O. (2015). Overview of statistical evaluation techniques adopted by publicly available MVPA toolboxes. Organization of Human Brain Mapping Annual Meeting, Honolulu HI, USA. Talk
http://www.pymvpa.org/files/OHBM2015_Halchenko.pdf

Humans virtue #3

- impossible to clone "your own Vapnik" (yet or ever)
- condensed presentation of articles is rarely sufficient and not reproducible as such
- we think we (or students) can do the same thing
- we need to do our best to do "correct science"
- test and validate, be skeptical

Halchenko, Y. O. (2015). Overview of statistical evaluation techniques adopted by publicly available MVPA toolboxes. Organization of Human Brain Mapping Annual Meeting, Honolulu HI, USA. Talk
http://www.pymvpa.org/files/OHBM2015_Halchenko.pdf

- automate and stay efficient
- test and validate
- share and contribute back (*i.e.*, collaborate)

Research parasite

Become a Symbiotic Research parasite

https://en.wikipedia.org/wiki/Symbiosis

... In 1877 Albert Bernhard Frank used the word symbiosis (which previously had been used to depict people living together in community) ...

- Knowledge and thought transfer
  - Social media
  - Publications

- Reusable artifacts
  - Software
  - Data

Social media: News and discussions

Rapid, often unverified, but interactive communication channels

Twitter/G+/Facebook/... :

- follow your favorite researchers and projects
- no specific recommendations, but I am yarikoptic@Twitter
- I have heard that Nikolaus Kriegeskorte is active on Facebook (please choose media which is open)

Mailing lists/Forums :

- Project mailing lists
- NeuroStars.org
    - Q&A site for neuro-scientists
    - main forum for some projects (nipype)
    - over 500 registered users
    - easy to follow/subscribe to specific tags and/or posts
- neuroimaging@python.org mailing list

Social media: Blogs

Semi-regular informal presentations and explorations of curious minds.

Jo Etzel mvpa.blogspot.com :
  Posts on various aspects of PR4NI/MVPA in neuroimaging (R, Python)

Andrew Jahn andysbrainblog.blogspot.com :
  Introductions (blog + YouTube videos) to software tools in neuroimaging – installation and basic use

Jeanette Mumford mumfordbrainstats.tumblr.com :
  Extensive collection (blog + YouTube videos) on various aspects of neuroimaging from univariate to MVPA

Russel Poldrack, Tal Yarkoni, ...

Workshops, tutorials, summer schools, ...

You can *attend* advanced workshops (even post-fact)

Prior OHBMs :
> www.pathlms.com/ohbm/courses
> Videos or slides for prior OHBM tutorials (including PR4NI) and
> the main conference

CCN@Dartmouth Summer Workshops :
> www.dartmouth.edu/%7Eccn/workshops/index.html
> Videos from yearly workshops on "Brain Decoding" topics.
> This year (coming in August) – "Predictive Decoding"

NeuroImaging Training Program @UCLA :
> www.brainmapping.org/NITP/NITPSummerProgram.php
> Video/slides archives on various aspects of neuroimaging,
> including pattern recognition

Knowledge transfer: Publications
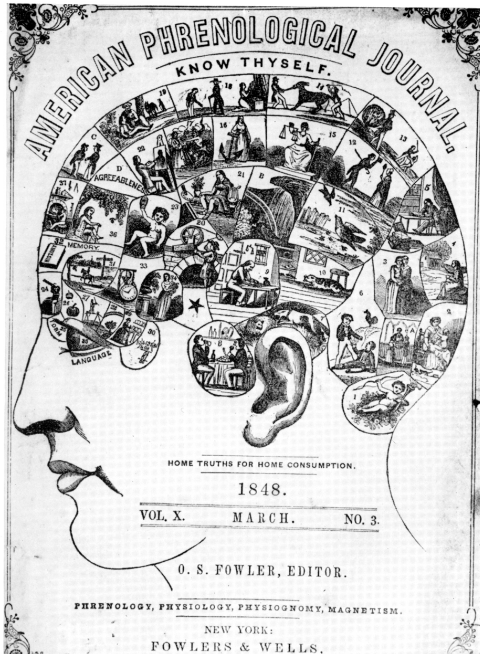
CONTRIBUTE

Free books on Machine Learning :
   github.com/datalad/mlbooks (git-annex repository)
   inspired by 12 Best Free Ebooks for Machine Learning
   (devzum.com/2015/05/best-free-machine-learning-ebooks)

Knowledge transfer: Pre-print servers

Whenever possible, please share your works for early feedback.
Provide feedback in return on freshly "published" pre-prints

- arxiv.org - the Pioneer of all pre-print servers.
  Originally Physics only, now also carries "Machine Learning" and
  "Neurons and Cognition"

- biorxiv.org - pre-print server for Biological sciences provided by
  Cold Spring Harbor Laboratory. Has "Neuroscience" section

Recent example arxiv.org/abs/1606.02840

Grootswagers, T., Wardle, S. G., and Carlson, T. A. (2016). Decoding
dynamic brain patterns from evoked responses: A tutorial on
multivariate pattern analysis applied to time-series neuroimaging data.
ArXiv e-prints

Reusable artifacts

Reusable artifacts: Selected toolkits

Selection criteria: toolkit must ...

- be an active Free and Open Source Software (FOSS) project
- be geared directly toward PR4NI:
    - I/O for common formats
    - bi-directional masking, basic pre-processing
    - basic parallelism
- have extended publicly available QA (unit-/regression/etc tests)
- have user-oriented documentation (tutorials, examples, etc.)
- streamlined ways for **you** to contribute back (e.g., public VCS)

Google spreadsheet with some others, and additional information:
https://goo.gl/qxX3Y2

CoSMoMVPA (cosmomvpa.org)

CoSMoMVPA

A multi-modal, multi-variate pattern analysis (MVPA) toolbox in Matlab / GNU Octave for cognitive neuroscientists.

CoSMoMVPA

- State-of-the art, yet simple to use MVPA implementations.
- Runs on the Matlab and GNU Octave platform.
- Handles fMRI volumetric, fMRI surface-based, and MEEG data through a uniform data structure.
- Support for a wide range of data formats.
- Searchlights in the volume, on the surface, over sensors, time bins, and frequency bands.
- Multiple-comparison correction using Threshold-Free Cluster Enhancement Monte Carlo simulations.
- Extensive documentation, including a variety of runnable scripts and implementation exercises (with solutions).
- Is Free/Open Source Software (MIT License).

This Page

Show Source

Quick search

Examples

Philosophy

Get started

Download

Modules

CoSMoMVPA (cosmomvpa.org)

C⊙SM⊙MVPA

CONTRIBUTE

Vital signs

Language/Interface : Matlab/Octave

Neuroimaging modalities : *MRI, M/EEG

Public VCS : github.com/CoSMoMVPA/CoSMoMVPA

Authors (90%) + Contributors : 1+6

Life span, version : Sometime in 2013/June 2016, 20160620

Tests coverage : 90%

Documentation : cosmomvpa.org/documentation.html

Workshops/Tutorials : cosmomvpa.org/prni2016.html

Canonical DOI : 10.1101/047118

Show Source

Quick search

CoSMoMVPA (cosmomvpa.org)

CoSMoMVPA

A multi-modal, multi-variate pattern analysis (MVPA) toolbox in Matlab / GNU Octave for cognitive neuroscientists.

5 "killer features"

- flexible searchlights (any combination of M/EEG time and frequency, volume, surface, channel and source space)
- bootstrapped cluster level + TFCE thresholding methods
- surface based analyses
- good i/o support (NIFTI, GIFTI, AFNI, SPM, BrainVoyager, EEGLAB, FieldTrip, PyMVPA)
- cross-time generalization (e.g., M/EEG trials)

This Page

Show Source

Quick search

Nilearn (nilearn.github.io)

Nilearn (nilearn.github.io)

Nilearn:

Machine learning for Neuro-Imaging in

SVM   Ward clustering
Searchlight   ICA
Nifti IO   Datasets

Vital signs

Language/Interface : Python

Neuroimaging modalities : *MRI

Public VCS : github.com/nilearn/nilearn

Authors (90%) + Contributors : 12+37

Life span, version : February 2015/June 2016, 0.2.5

Tests coverage : 93%

Documentation : nilearn.github.io

Workshops/Tutorials : prni2016.wix.com/prni2016#!blank-6/x350s

Canonical DOI : 10.3389/fninf.2014.00014

Nilearn is part of the NiPy ecosystem.

Development
Nilearn on GitHub

Nilearn (nilearn.github.io)

Nilearn:

Machine learning for Neuro-Imaging in Python

Google Custom Search | Search | ×

5 "killer features"

- scikit-learn API (yeah, fit-transform!)
- automatic fetching of open datasets
- functional connectivity: inverse-covariance estimation, tangent space embedding
- searchlights, SpaceNet
- nice plotting for stat maps, masks, connectomes, connectivity matrices

User Guide
Browse the full documentation
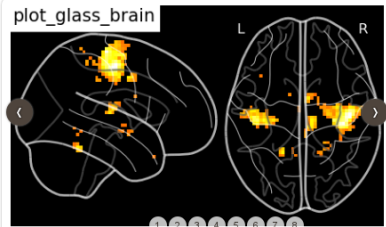
1 2 3 4 5 6 7 8

Installation

Development

Nilearn on GitHub

Nilearn is part of the NiPy ecosystem.

PyMVPA (pymvpa.org)

PyMVPA

CCN Center for Cognitive Neuroscience at Dartmouth

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

Multivariate Pattern Analysis in Python

PyMVPA is a Python package intended to ease statistical learning analyses of large datasets. It offers an extensible framework with a high-level interface to a broad range of algorithms for classification, regression, feature selection, data import and export. It is designed to integrate well with related software packages, such as scikit-learn, shogun, MDP, etc. While it is not limited to the neuroimaging domain, it is eminently suited for such datasets. PyMVPA is free software and requires nothing but free-software to run.

PyMVPA stands for **M**ulti**V**ariate **P**attern **A**nalysis (MVPA) in **Py**thon.

Installation   Tutorial   Documentation   Support

News

Table Of Contents

Next topic

Quick links

Search mailing list

PyMVPA (pymvpa.org)

Vital signs

Language/Interface : Python, Command line

Neuroimaging modalities : *MRI, E/MEG (some)

Public VCS : github.com/PyMVPA/PyMVPA

Authors (90%) + Contributors : 3+36

Life span, version : February 2008/May 2016, 2.5.0

Tests coverage : 80%

Documentation : www.pymvpa.org

Workshops/Tutorials : http://www.pymvpa.org/courses.html

Canonical DOI : 10.1007/s12021-008-9041-y

it!)

♡ ⟶   13 May

PyMVPA@MLOSS.org
PyMVPA@INCF

Search mailing list

PyMVPA (pymvpa.org)

Multivariate Pattern Analysis in Python

5 "killer features"

- adapters to R libs (e.g. glmnet), MDP, sklearn, MDP, ..., CoSMoMVPA, AFNI
- dataset construct to bind data and attributes together
- flexible Searchlighting (volume, surface, space/time)
- cluster level thresholding for searchlights
- functional alignment: Hyperalignment

PyMVPA Team @pymvpa

Fresh release (2.5.0) of #PyMVPA with SearchlightHyperalignment (dx.doi.org/10.1093/cercor...) is out! (#NeuroDebian got it!)

13 May

Bug tracker
Mailing list archive
Who is using PyMVPA?
Dataset Archive
PyMVPA@MLOSS.org
PyMVPA@INCF

Search mailing list

Reusable artifacts: other software

See also

mloss.org :

>   The catalog of Open Source Machine Learning toolboxes and
>   platforms

nitrc.org :

>   The catalog (and often primary hosting) of Neuroimaging
>   resources

thunder-project.org :

>   Modular Apache Spark-based platform for the analysis of image
>   and time series data in Python.

doi: 10.3389/fnins.2016.00248 :

>   Mahmud, M. and Vassanelli, S. (2016). Processing and analysis
>   of multichannel extracellular neuronal signals: State-of-the-art
>   and challenges. *Frontiers in Neuroscience*, 10(248)

search for... ✕   Go

Tweet   G+1   918   37 Comments | next

The Ultimate Neuroscience Software Platform

NeuroDebian provides a large collection of popular neuroscience research software for the Debian operating system as well as Ubuntu and other derivatives. Popular packages include *AFNI*, *FSL*, *PyMVPA* and *many others*. While we do strive to maintain a high level of quality, we make no guarantee that a given package works as expected, so use them at your own risk. If you do encounter problems or you just like to say thanks, simply *send us an email*.

Learn more about NeuroDebian, the goals of this project, and help us *spread the word about NeuroDebian*!

Halchenko, Y. O. & Hanke, M. (2012). Open is not enough. Let's take the next step: An integrated, community-driven computing platform for neuroscience. *Frontiers in Neuroinformatics*, 6:22.

[more publications]

[[...] The] only way to conduct reliable and reproducible science is to use open source software [...]. NeuroDebian is by far the most advanced undertaking for such a scientific approach in the neuroscience community.
—- **Valentin Haenel** [2010-09-17]
Psignifit and pyoptical developer, Modelling Kognitiver Prozesse, Technische Universität, Berlin, Germany

[more testimonials]

Get NeuroDebian

First select what kind of operating system you are using, and then choose a download server close to you:

Select your operating system ▾   Select a download server ▾

News

Tweets by @NeuroDebian

What are the *inside outs* of NeuroDebian?

CONTRIBUTE

Mind map:

- **QA/Testing**
  - Build-time tests
  - Maintainer tests
  - Interoperability testing
- **Virtual machine**
  - Compatibility
    - Windows
    - Mac OS X
    - other
  - Reproducable research
  - Portability
  - Teaching
- **Packaging**
  - Mentoring
  - Maintenance
    - Proxy bugreports
    - Portability patches
    - Library/compiler transitions
  - Legal checks
  - Modularisation
- **Communication**
  - Online
    - Publications
    - Mailing lists
    - IRC (#neurodebian on OFTC)
    - identi.ca/twitter (neurodebian)
    - NeuroDebian Insider Blog
    - team@neuro.debian.net
  - Personal
    - Conference booth
    - Talks
    - Stalking
- **Repository**
  - Software
    - Debian
      - i386
      - amd64
    - Ubuntu
      - i386
      - amd64
  - Data
  - Usage statistics
  - Snapshoting
  - Mirrors
    - DE
    - GR
    - US-CA
    - US-NH
    - US-TN
- **Debian**
  - Blends
    - Debian Science
    - Debian Med
  - BTS
  - QA
  - Archive
    - i386
    - amd64
    - armel
    - mips(el)
    - kfreebsd-i386/amd64
    - ...
  - Derivatives

Aforementioned software
is nearly useless
without data

Human Connectome Project db.humanconnectome.org :
Multi-institution effort collecting a rich curated MRI, fMRI, DTI, MEG data (current release - 900 subjects, pre-processed data available) (planned for DataLad #579)

Reusable artifacts: Data providers #2

OpenfMRI openfmri.org :
> a wide (49 datasets ATM) collection of (primarily fMRI/MRI) datasets (not pre-processed) across different tasks with a unified file system hierarchy (available through DataLad)

Study Forrest studyforrest.org :
> multifaceted MRI, fMRI (including 7T), DTI, EEG data with stimuli annotations while subjects experienced rich natural stimulation (Forrest Gump movie) (available through DataLad)

Academic Torrents academictorrents.com :
> a growing collection of datasets (including MRI) distributed as Torrents (planned for DataLad #30)

NITRC Image Repository www.nitrc.org/ir :
> hosts 14 data projects (including 1000 FCP)

INDI fcon_1000.projects.nitrc.org :
> The International Neuroimaging Datasharing Initiative (INDI) collecting and curating submissions for a wide variety of datasets (planned for DataLad #580)

Reusable artifacts: Data access/(re-)sharing

CONTRIBUTE

nilearn.datasets :
> Unified API for downloading some popular atlases and datasets (e.g. Haxby 2001)

Nidata nidata.github.io :
> Python module with unified access to a good range of popular datasets (from OpenfMRI, HCP, etc)

DataLad datalad.org :
> uses your favorite git (+git-annex) to

- provide unified access to various data sources
- automate scraping of such datasets from the web resources
- publish your new or derived data (again, under git control)

Discover more about DataLad at our DataLad/NeuroDebian exhibit table,
poster 1855 (June 28: 12:45 PM-02:45 PM), talk (11:08 AM, Room K, Level 2)

Git-RDM github.com/ctjacobs/git-rdm :
> a git "plugin" to publish data to Zenodo or Figshare

Let's become a "collective PR4NI mind"

- resources (humans, software, data) for PR4NI are vast
- the field evolves with tremendous pace and classical models (dead tree papers, "research non-parasites", etc.) are inefficient and do not scale
- do not just (ab)use resources – become a (part of the) resource!
- when or where possible – (re)use and/or enhance existing resources instead of investing into a "new" duplicate

- automate and stay efficient
- test and validate
- share and contribute back (*i.e.*, collaborate)

Brain Download: iz compltes.

Thank you!

References

Grootswagers, T., Wardle, S. G., and Carlson, T. A. (2016). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time-series neuroimaging data. *ArXiv e-prints*.

Halchenko, Y. O. (2015). Overview of statistical evaluation techniques adopted by publicly available MVPA toolboxes. Organization of Human Brain Mapping Annual Meeting, Honolulu HI, USA. Talk.

Mahmud, M. and Vassanelli, S. (2016). Processing and analysis of multichannel extracellular neuronal signals: State-of-the-art and challenges. *Frontiers in Neuroscience*, 10(248).

Who is NeuroDebian for?

You want to . . .

- use a **rock-solid** operating system
- have **readily usable and the latest** software at your fingertips
- **try something new**, without investing much time
- offer **students** a fully functional "take-away" research environment
- **efficiently collaborate** with other researchers
- **waste less time** maintaining computers
- have **your own software** easily available for others to use
- **develop software** without worrying about dependencies