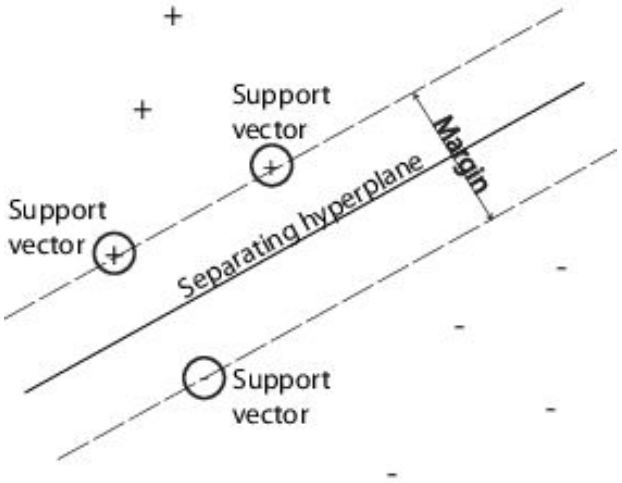
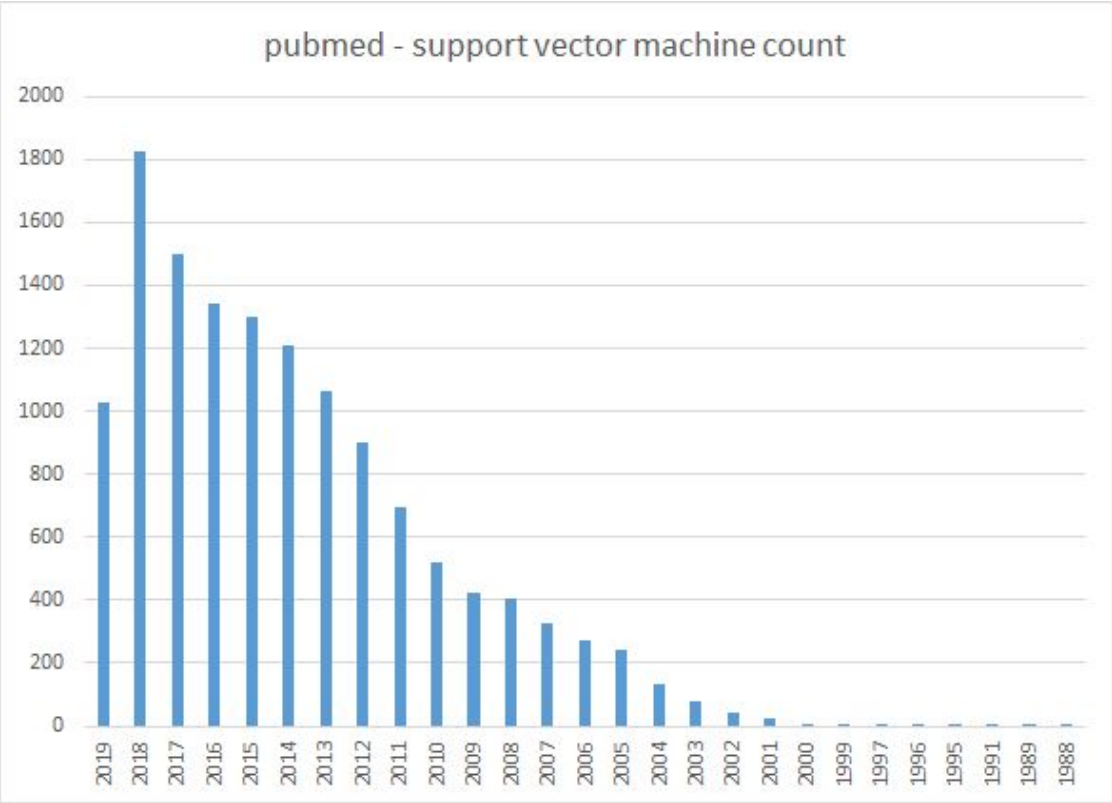
Intro to Linear Support Vector Machines and Binary Classification

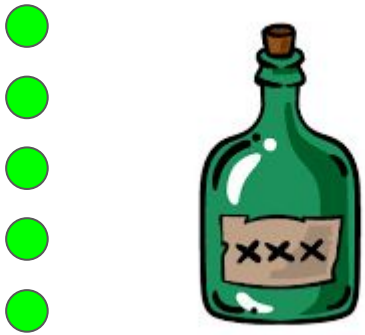


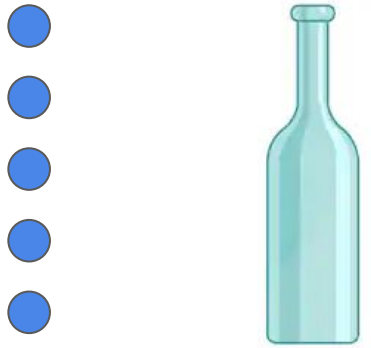
By Matt Flounders
June 19, 2019

Support vector what?

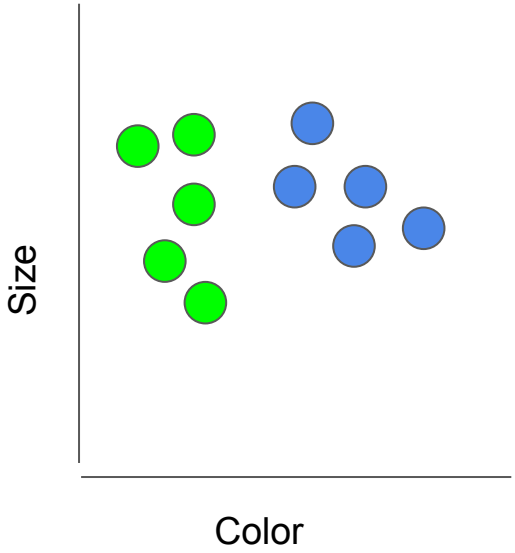


Classification Example

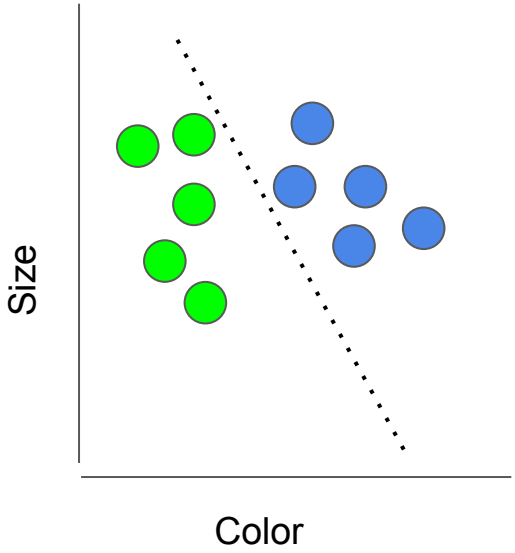




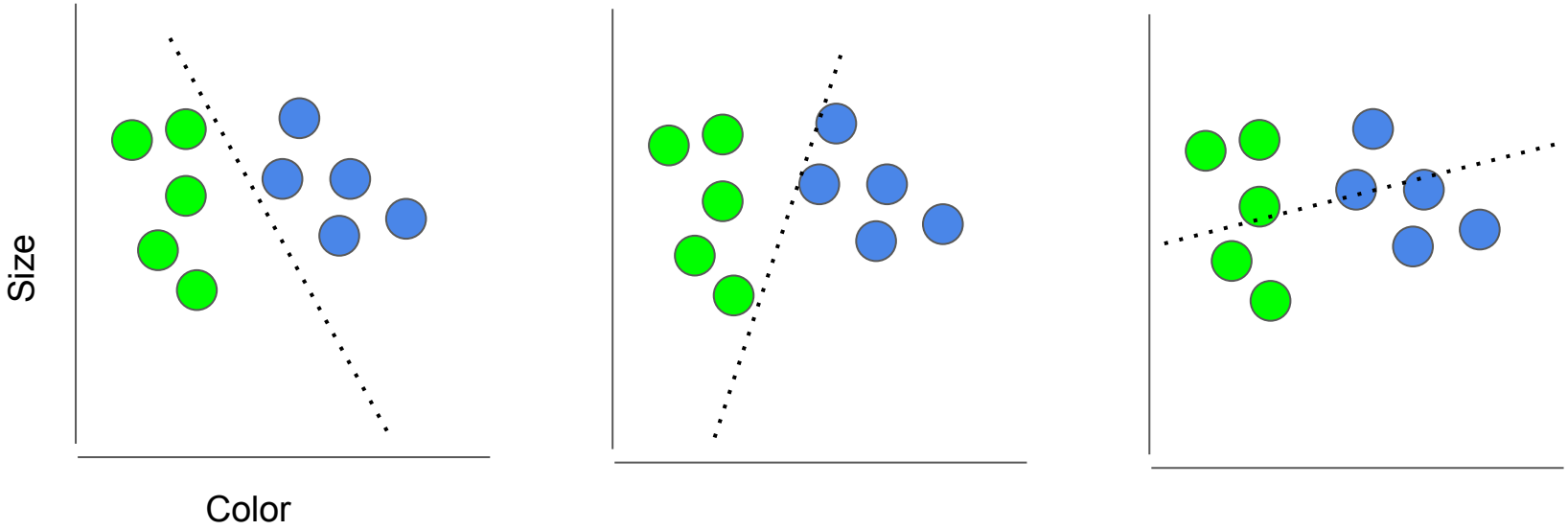
Classification Example



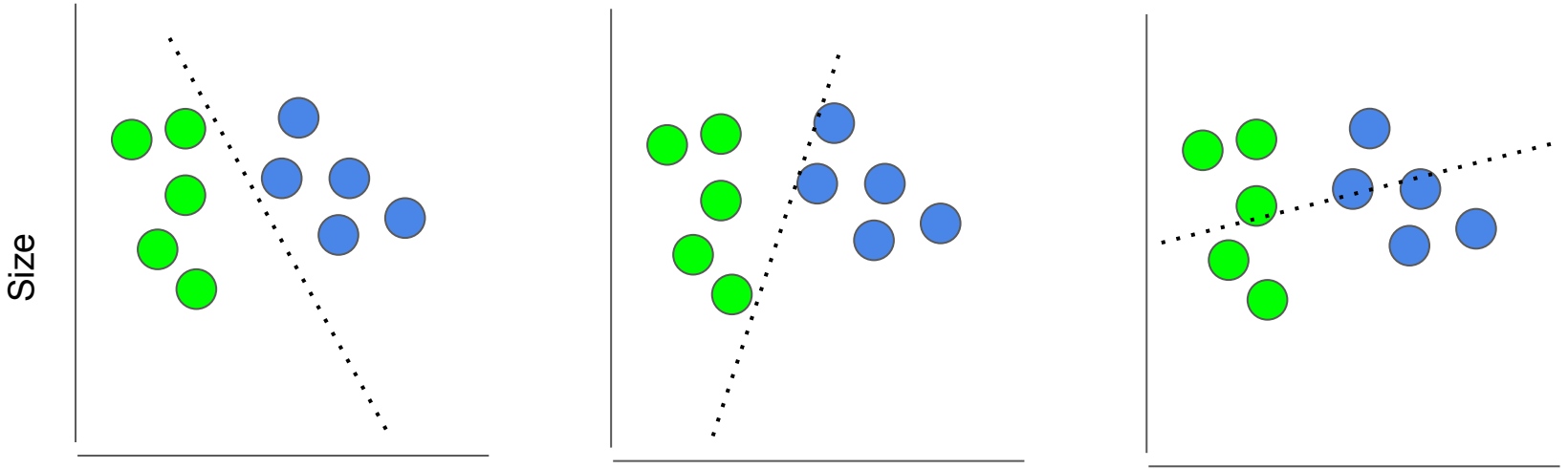
Classification Example



Hyperplane: best, worst, and ugly ….?

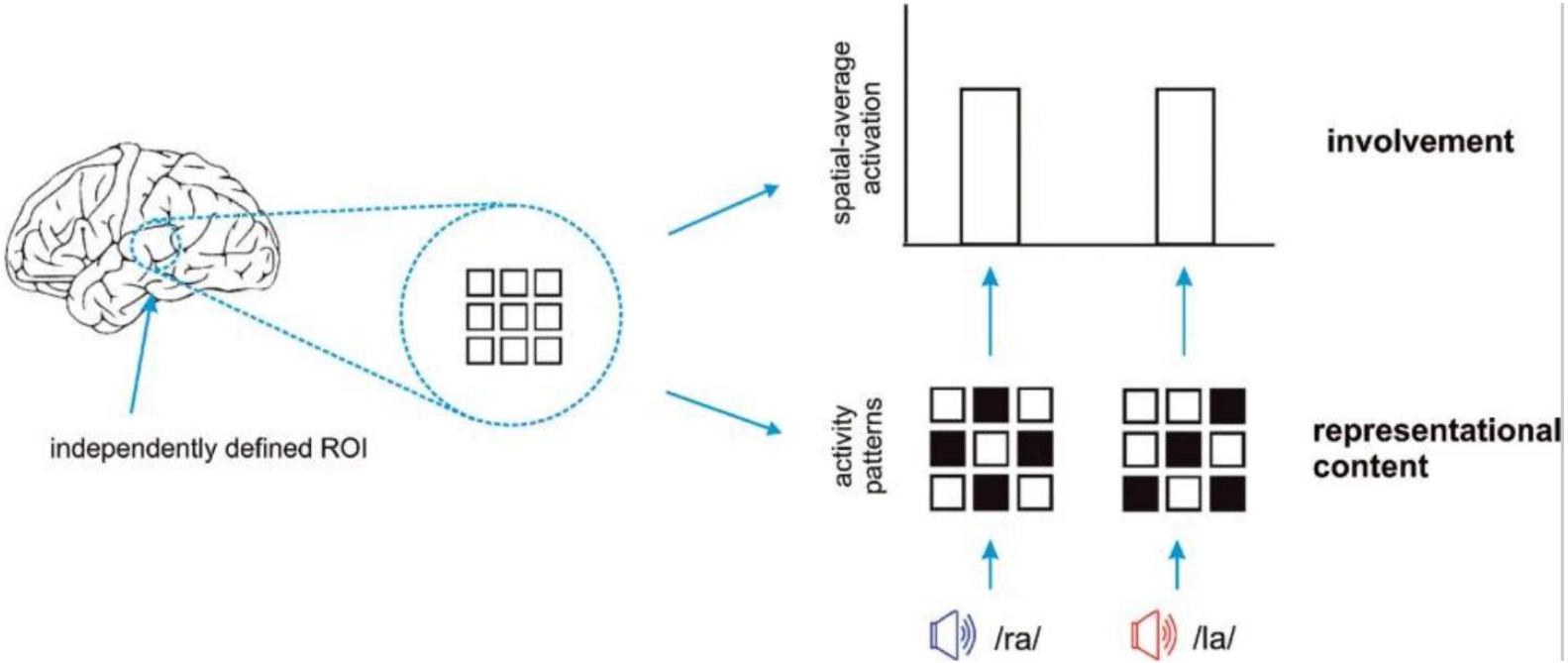


Hyperplane: best, worst, and ugly ….?



vs.

But, why does this matter for neuroscience?



Razida et al., 2008

But, why does this matter for neuroscience?

Article | Published: 24 April 2005

Predicting the orientation of invisible stimuli from activity in human primary visual cortex

John-Dylan Haynes ✉ & Geraint Rees

Nature Neuroscience **8**, 686–691 (2005) | Download Citation ⬇

But, why does this matter for neuroscience?

This Issue | Views 9,291 | Citations 55 | Altmetric 34

PDF | More | Cite | Permissions

Original Investigation FREE

June 2016

Prediction of Individual Response to Electroconvulsive Therapy via Machine Learning on Structural Magnetic Resonance Imaging Data

Ronny Redlich, PhD[1]; Nils Opel, MD[1]; Dominik Grotegerd, MSc[1]; et al

» Author Affiliations | Article Information

JAMA Psychiatry. 2016;73(6):5

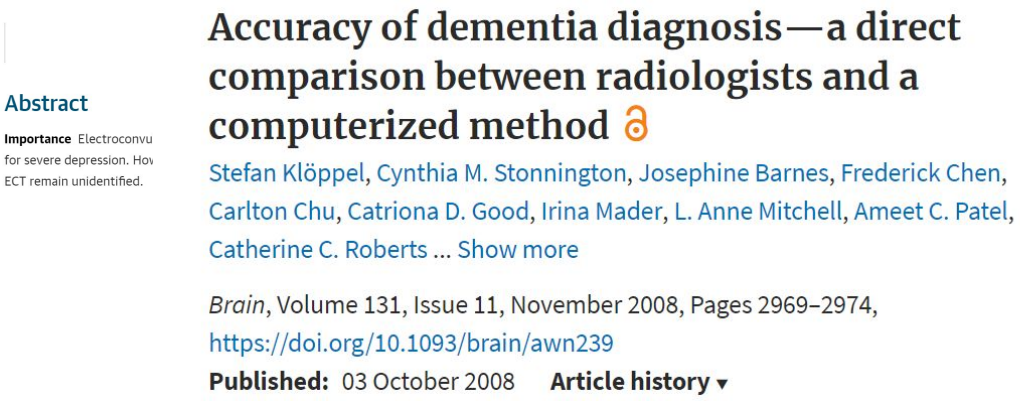
nature materials

Perspective | Published: 18 April 2019

Exploiting machine learning for end-to-end drug discovery and development

Sean Ekins, Ana C. Puhl, Kimberley M. Zorn, Thomas R. Lane, Daniel P. Russo, Jennifer J. Klein, Anthony J. Hickey & Alex M. Clark

Nature Materials **18**, 435–441 (2019) | Download Citation



Abstract

Importance Electroconvu
for severe depression. Hov
ECT remain unidentified.

Accuracy of dementia diagnosis—a direct comparison between radiologists and a computerized method

Stefan Klöppel, Cynthia M. Stonnington, Josephine Barnes, Frederick Chen, Carlton Chu, Catriona D. Good, Irina Mader, L. Anne Mitchell, Ameet C. Patel, Catherine C. Roberts ... Show more

Brain, Volume 131, Issue 11, November 2008, Pages 2969–2974, https://doi.org/10.1093/brain/awn239

Published: 03 October 2008 **Article history**

But, why does this matter for neuroscience?

This Issue Views 9,291 | Citations 55 | Altmetric 34

Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder FREE

Benson Mwangi, Klaus P. Ebmeier, Keith Matthews,

Brain, Volume 135, Issue 5, May 2012, Pages 1508–15

https://doi.org/10.1093/brain/aws084

Published: 27 April 2012 **Article history**

Abstract

Importance Electroconvulsive therapy (ECT) is one of the most effective treatments for severe depression. However, biomarkers that accurately predict a response to ECT remain unidentified.

computerized meth

Stefan Kloppel, Cynthia M. Stonnir

Carlton Chu, Catriona D. Good, Irir

Catherine C. Roberts ... Show more

Brain, Volume 131, Issue 11, Nover

https://doi.org/10.1093/brain/awr

Published: 03 October 2008 Ar

nature

hine learning for end-to-very and development

ey M. Zorn, Thomas R. Lane, Daniel P. Russo, Jennifer J. Klein,

MENU Translational Psychiatry

Original Article | OPEN | Published: 25 October 2016

A support vector machine model provides an accurate transcript-level-based diagnostic for major depressive disorder

J S Yu, A Y Xue, E E Redei & N Bagheri

Translational Psychiatry **6**, e931 (2016) | Download Citation

Table of contents

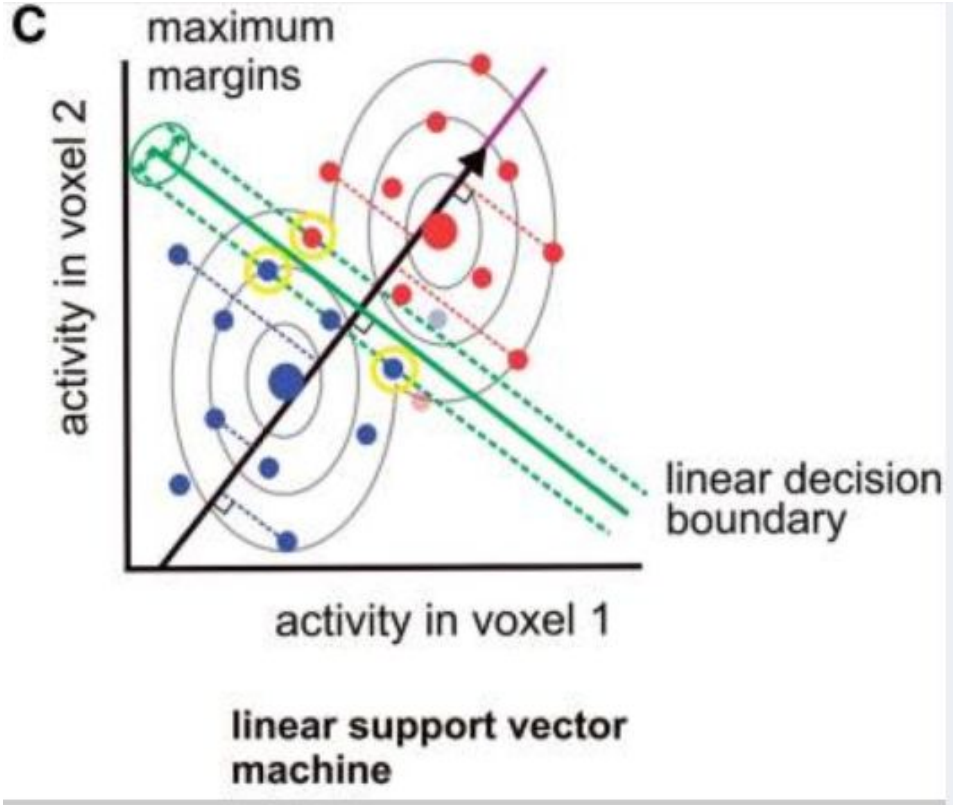
Terminology

Hyperplane: line discerning between data points, a decision boundary

Support vectors: data points closest to the separating hyperplane, identify the boundary of the slab

Margin: maximal width of the slab parallel to the hyperplane that has no interior data points

- **Soft margin:** slab parallel to the hyperplane, where many data points are separated but may contain some data points according to an a priori criteria



Mur et al., 2009

Soft margin

- A soft margin utilizes a priori criteria
  - Reducing “c” makes misclassification less important
  - Increasing “c” aims to stricly separate between classes

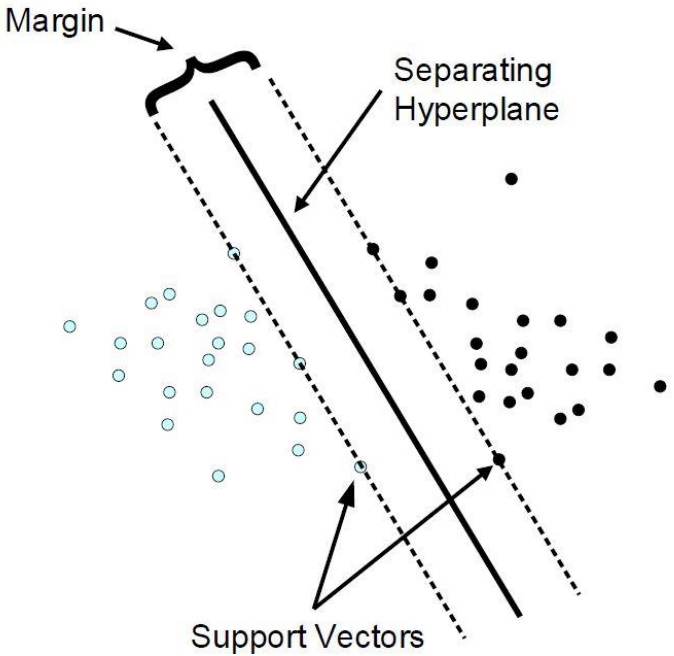
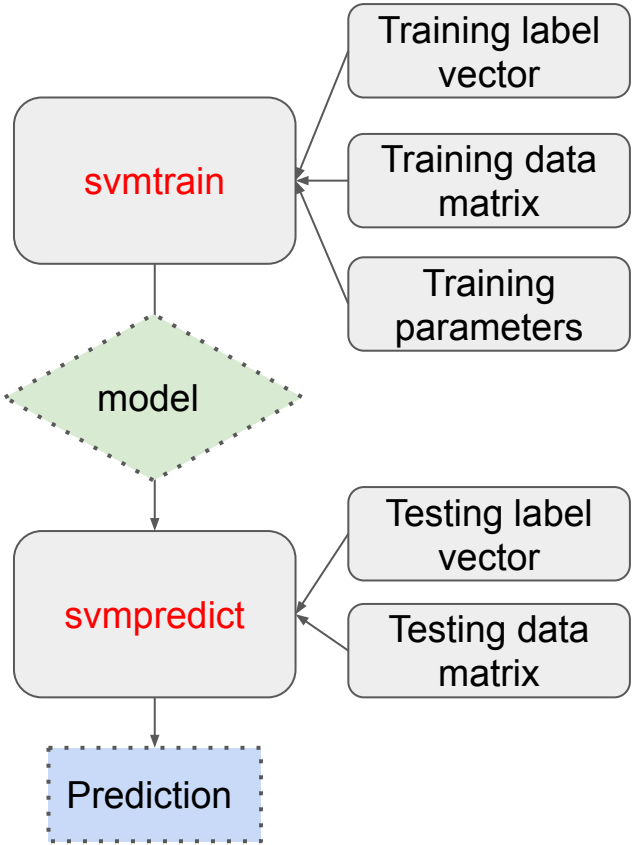


Figure 1: Classification (linear separable case)

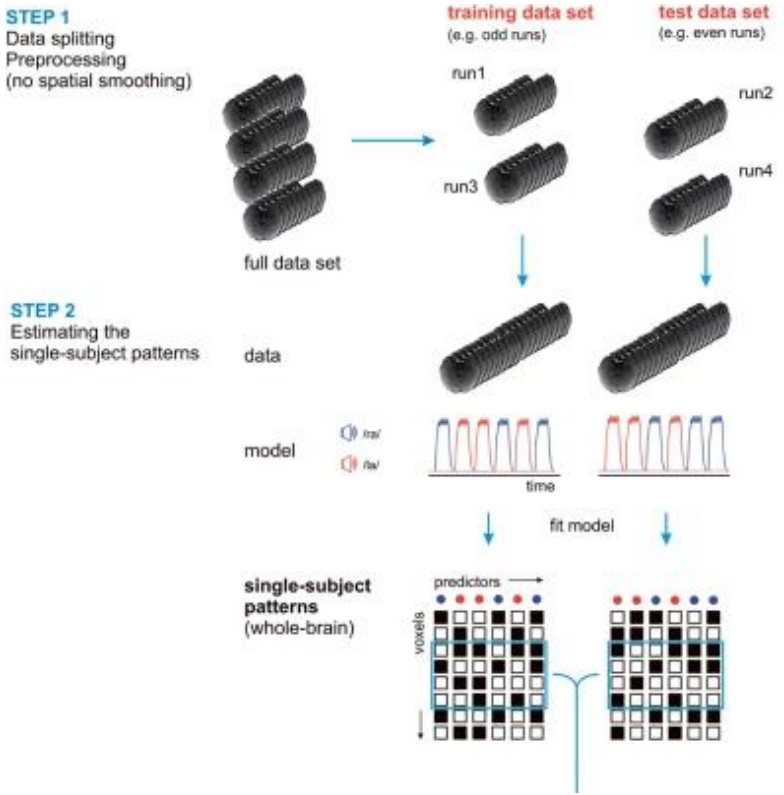
LibSVM

- A package implemented in many programming languages
- Tool to train, test, and optimize various classifications, regressions, and distributed estimation
  - Kernel options:
    - Linear
    - Polynomial
    - Radial Bias
    - Sigmoid
- Allows for multi-class classification, probability estimates, cross-validation, and “weighted” SVM



Workflow: Preprocess and Segment Data

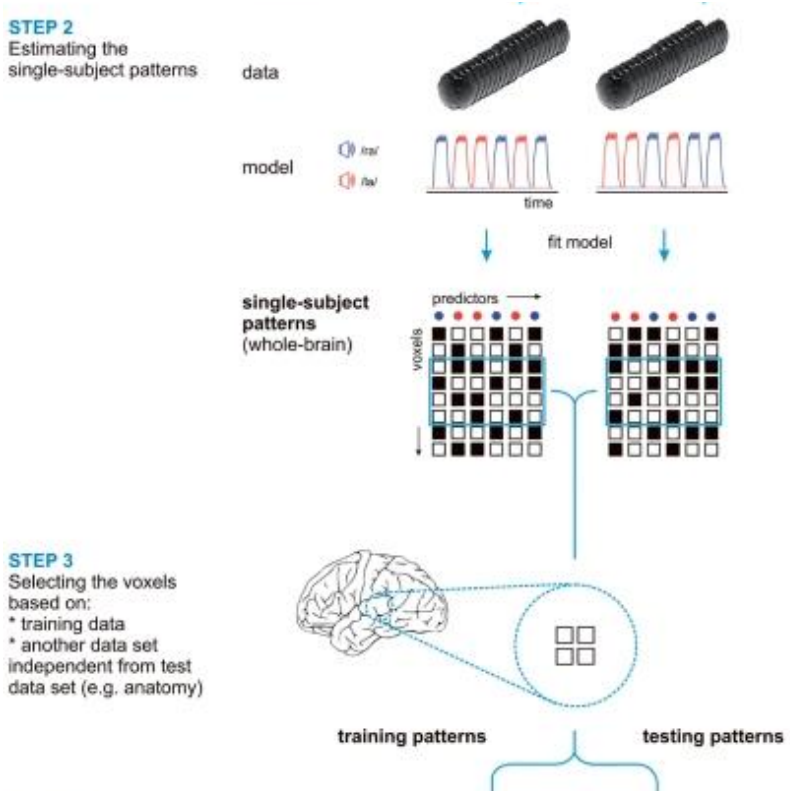
1. Split data into training and test sets
2. Within subject, utilize GLM to generate whole-brain activity patterns consisting of beta-estimates



Mur et al., 2009

Workflow: Preprocess and Segment Data

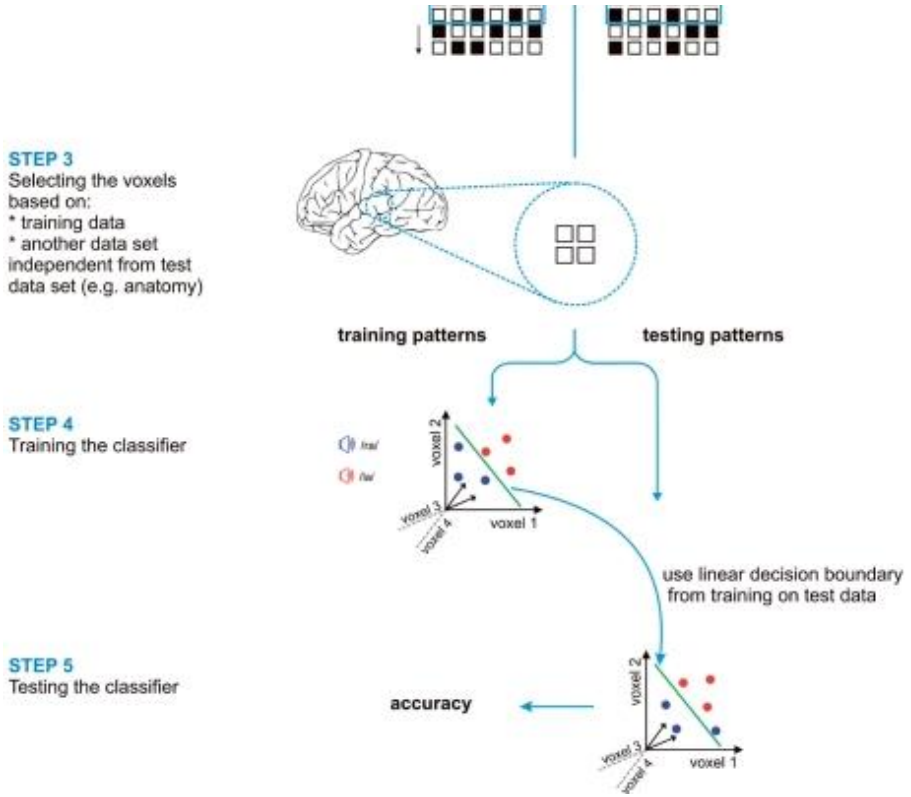
1. Split data into training and test sets
2. Within subject, utilize GLM to generate whole-brain activity patterns consisting of beta-estimates
3. Select voxels based on anatomy, function, both



Mur et al., 2009

Workflow: Train

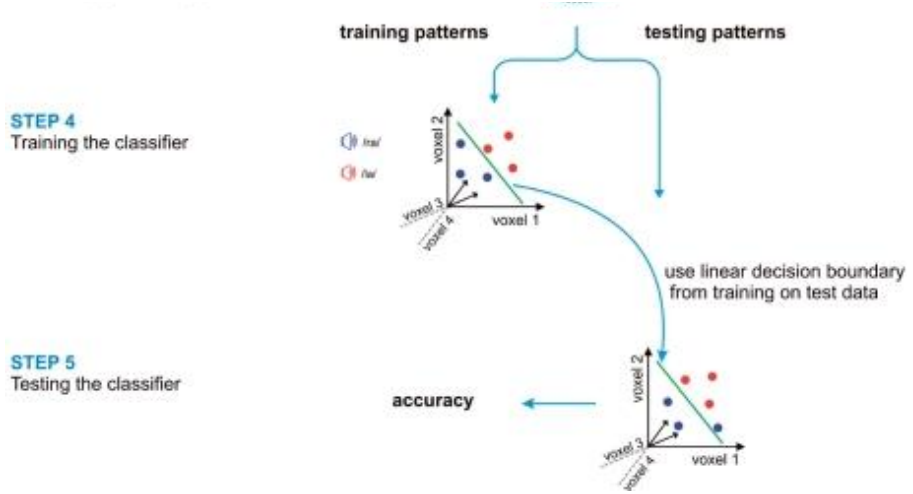
1. Split data into training and test sets
2. Within subject, utilize GLM to generate whole-brain activity patterns consisting of beta-estimates
3. Select voxels based on anatomy, function, both
4. Train a classifier



Mur et al., 2009

Workflow: Classify/Test/Predict

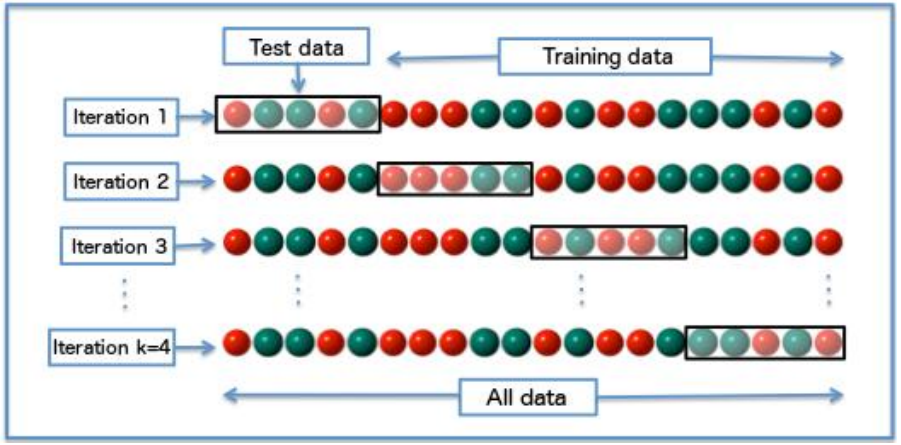
1. Split data into training and test sets
2. Within subject, utilize GLM to generate whole-brain activity patterns consisting of beta-estimates
3. Select voxels based on anatomy, function, both
4. Train a classifier
5. Test classifier



Mur et al., 2009

Workflow: Cross-validation

1. Split data into training and test sets
2. Within subject, utilize GLM to generate whole-brain activity patterns consisting of beta-estimates
3. Select voxels based on anatomy, function, both
4. Train a classifier
5. Test classifier
    a. k-fold cross-validation

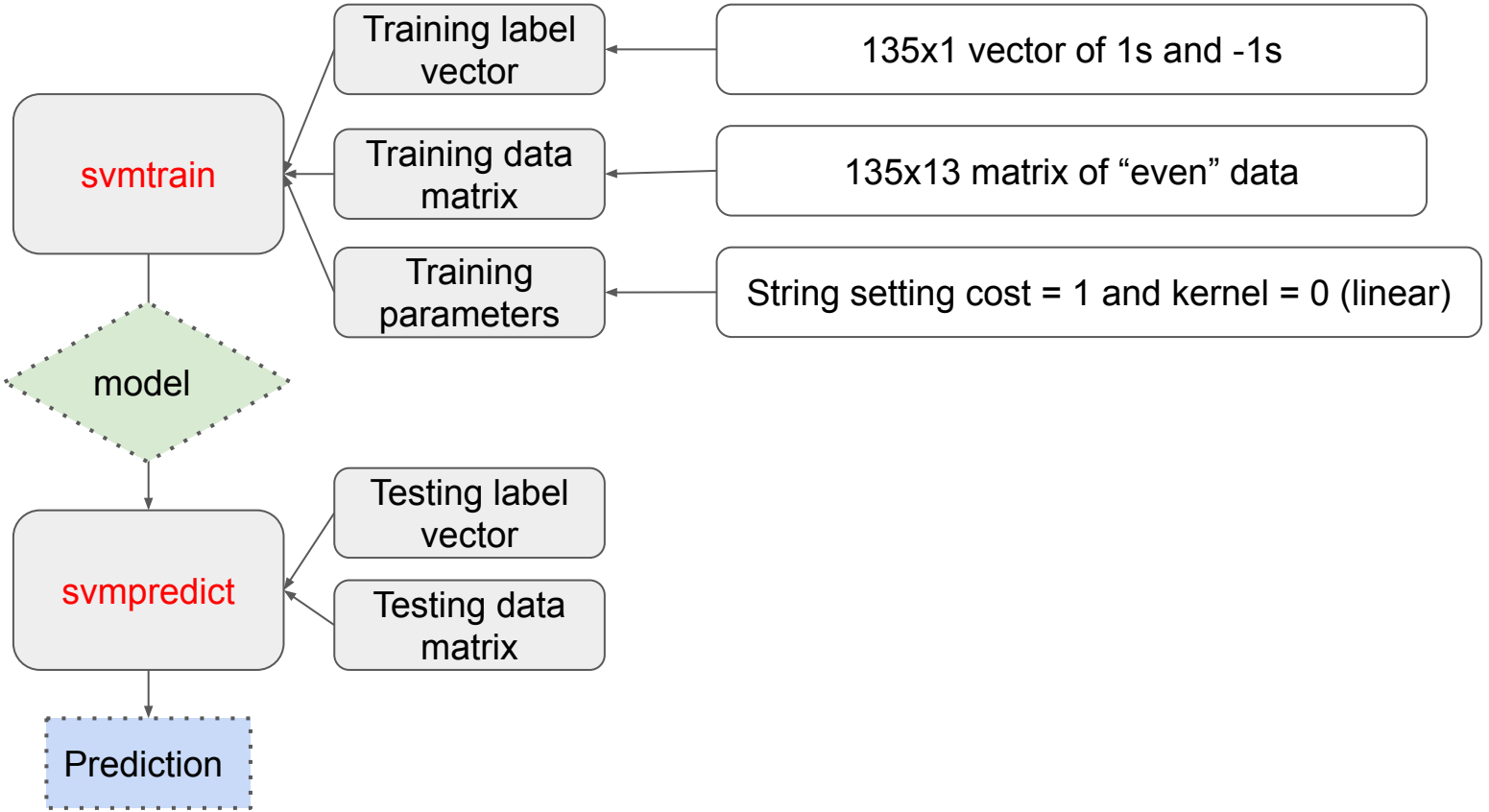


Example: heart_scale

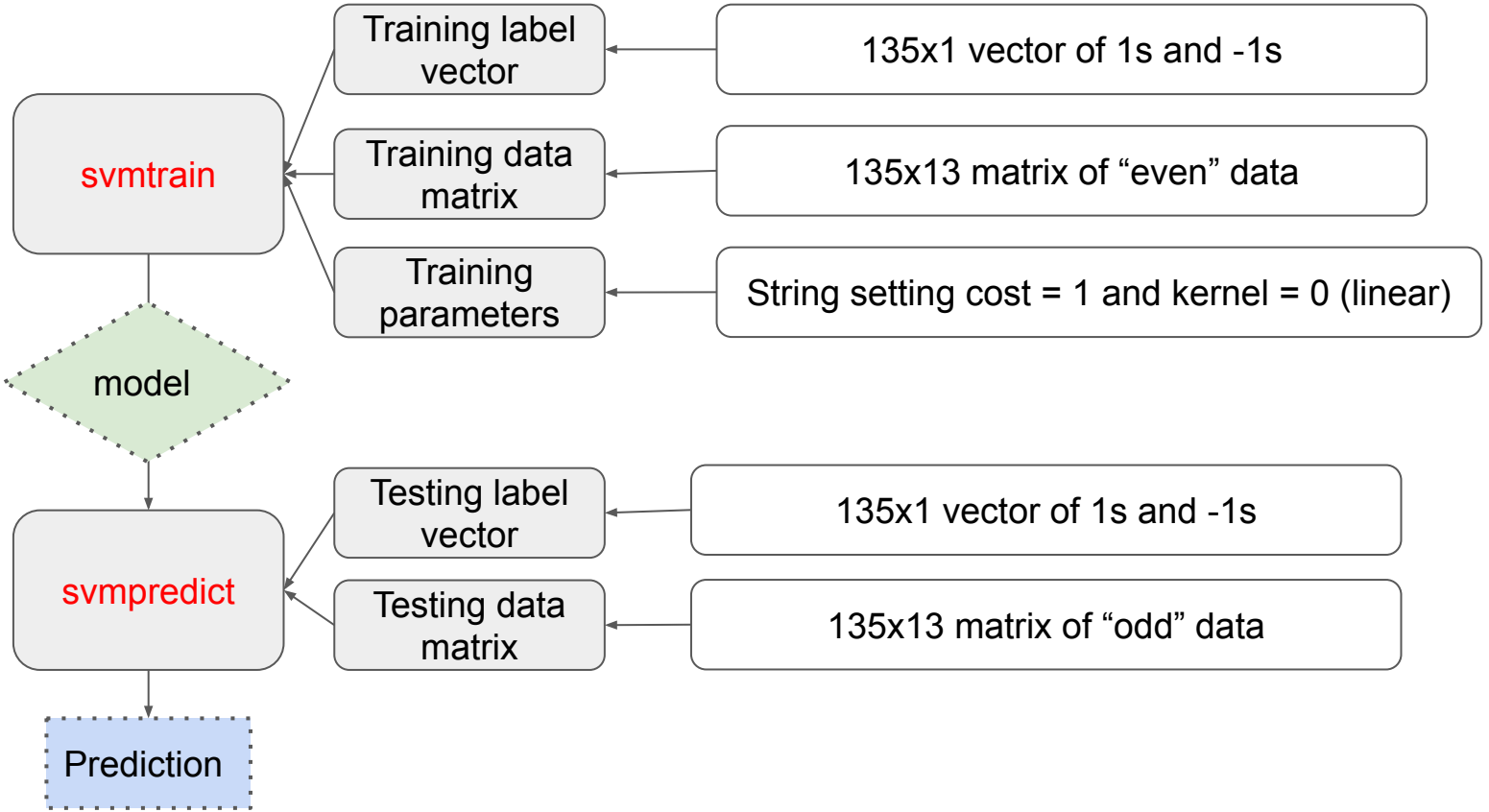
- Dataset: 270 examples with 14 features
  - 135 $heart_a$ and 135 $hearts_b$
  - 14 pieces of information
- Classifying $heart_a$ from $heart_b$

	BPM	...	Tricuspid Diameter
Heart 1			
...			
Heart 270			

Example: heart_scale



Example: heart_scale



Example: neuroscience

Example: neuroscience

Example: neuroscience

Example: neuroscience

Example: neuroscience

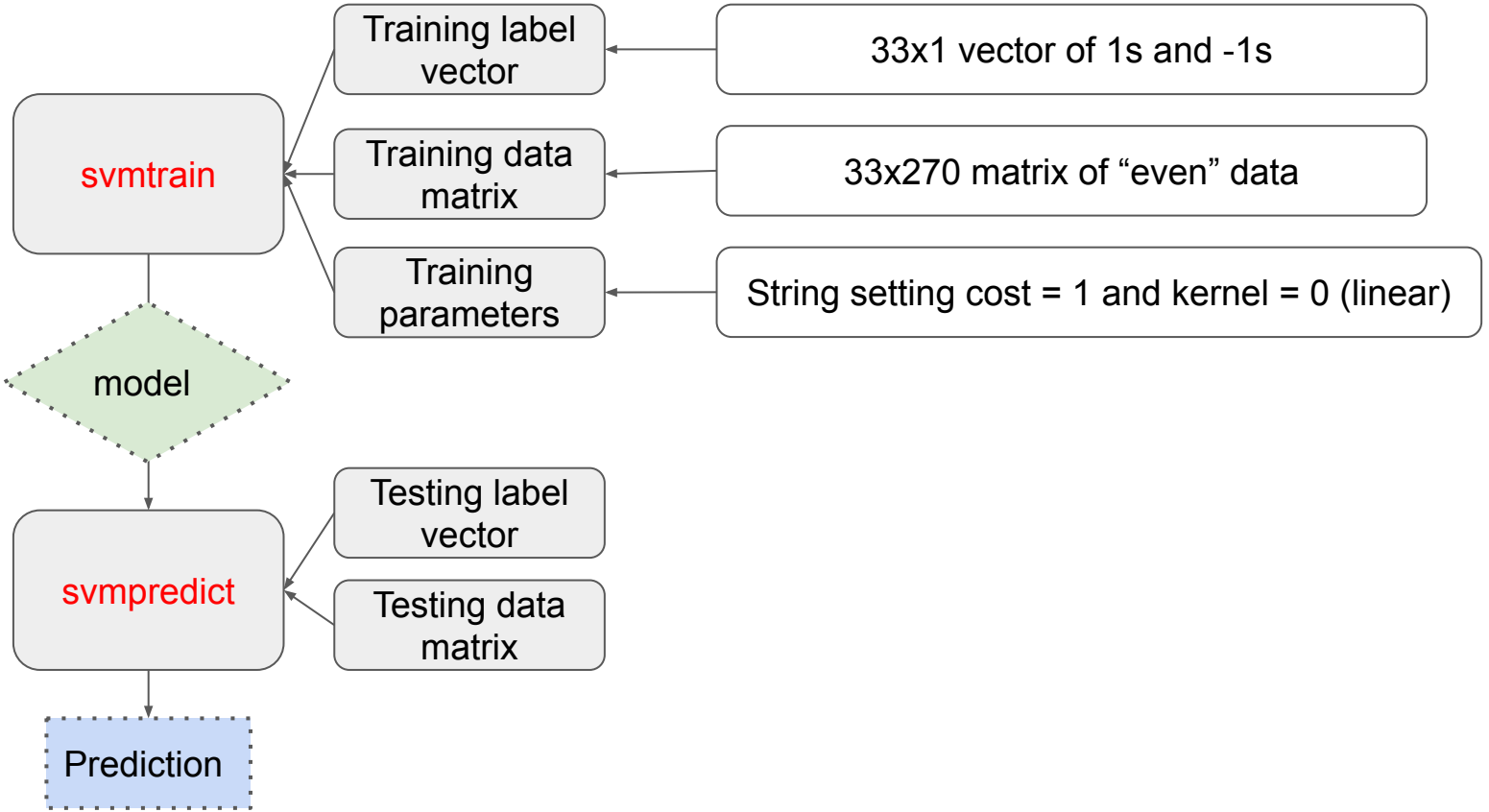
Example: neuroscience

- Dataset: 66 examples with 270 features
  - 33 images “Pre”, and 33 images “Post”
  - 270 MEG sensors
- Classifying Pre vs. Post, given identical stimulus input with differing prior

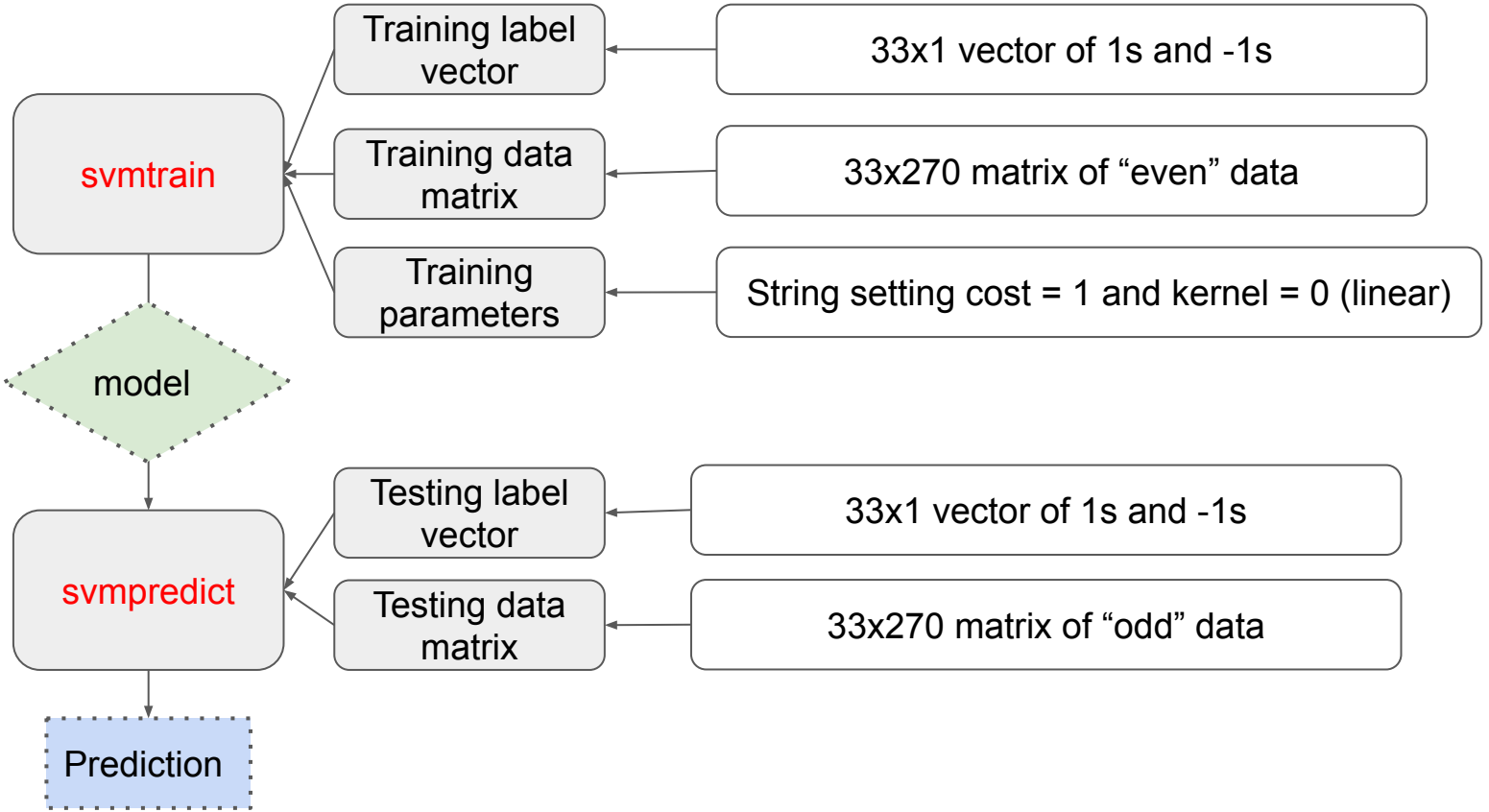
Pre Post

	MEG Sensor 1	...	MEG Sensor 270
Image1			
...			
Image2			

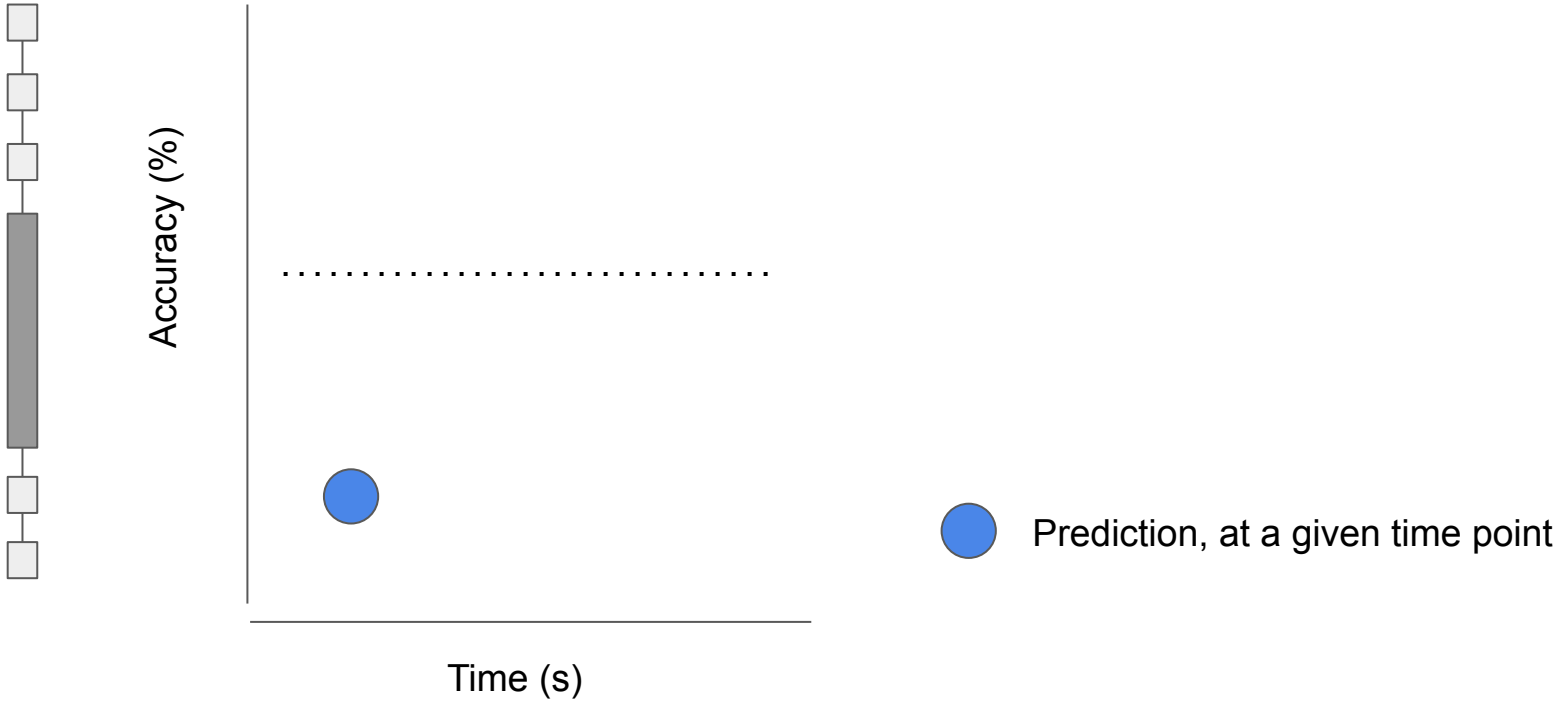
Example: neuroscience



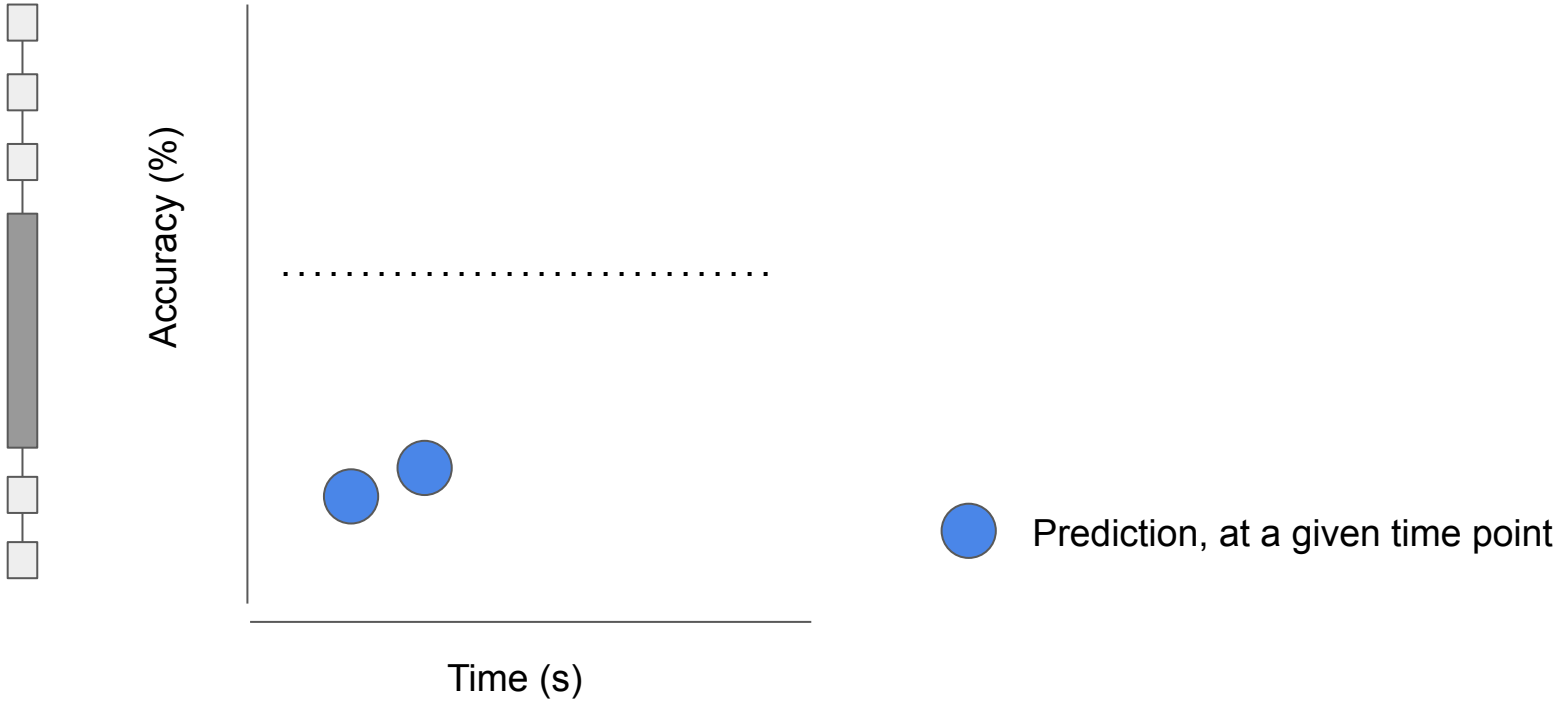
Example: neuroscience



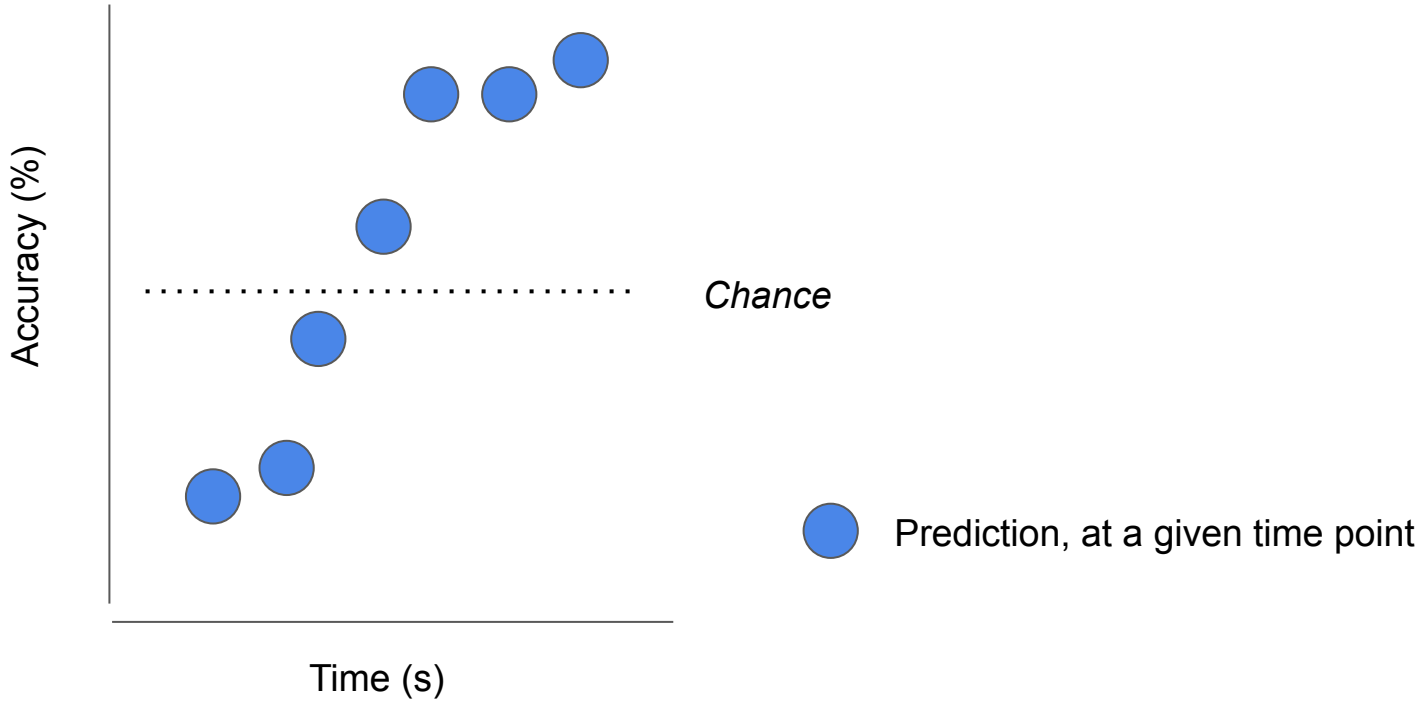
Example: neuroscience



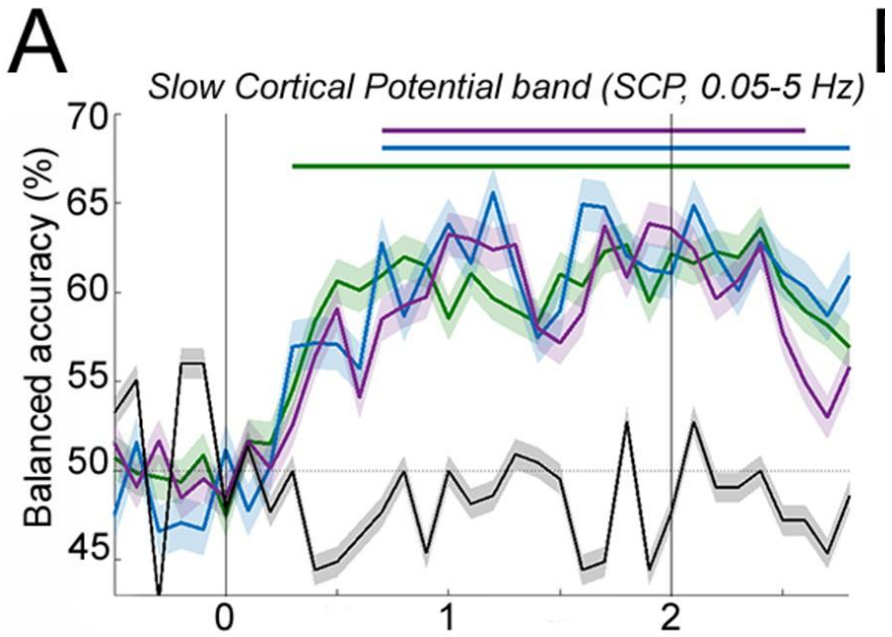
Example: neuroscience



Example: neuroscience



Example: neuroscience

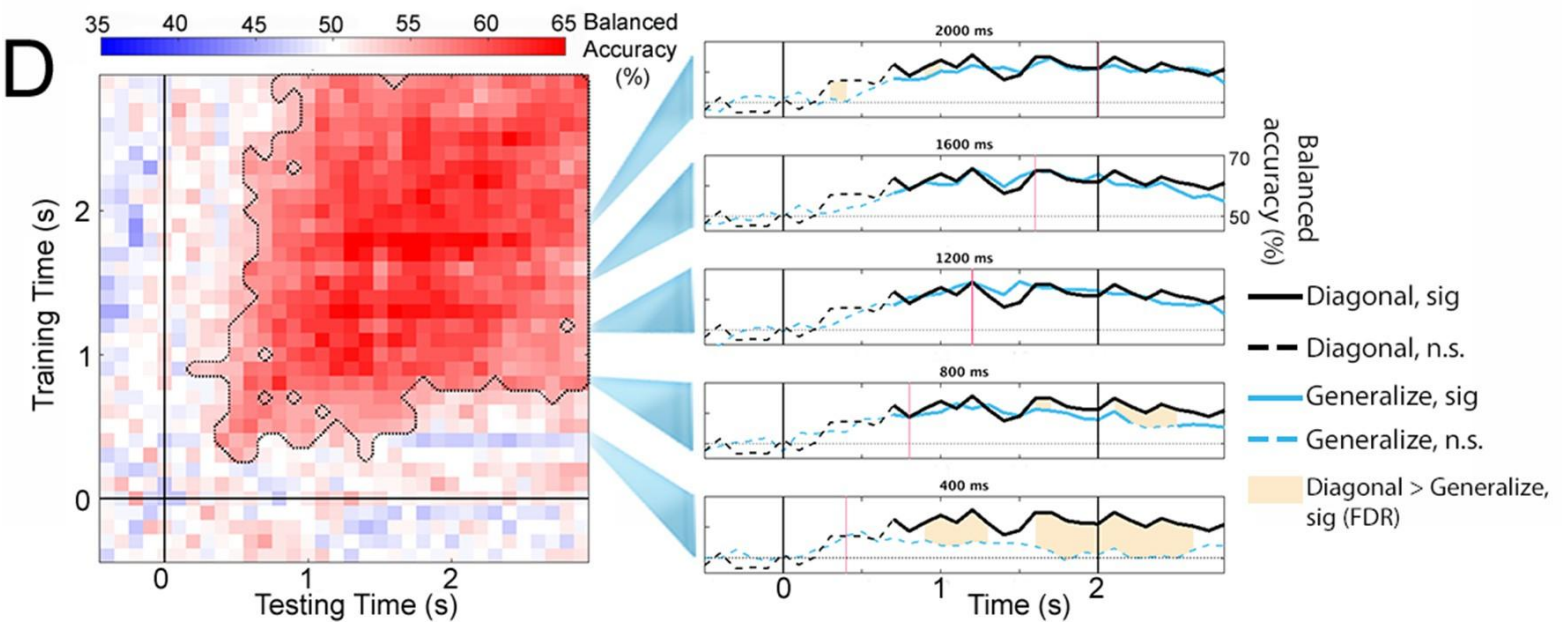


Pre vs. Post, prior presented

Pre vs. Post, no prior presented

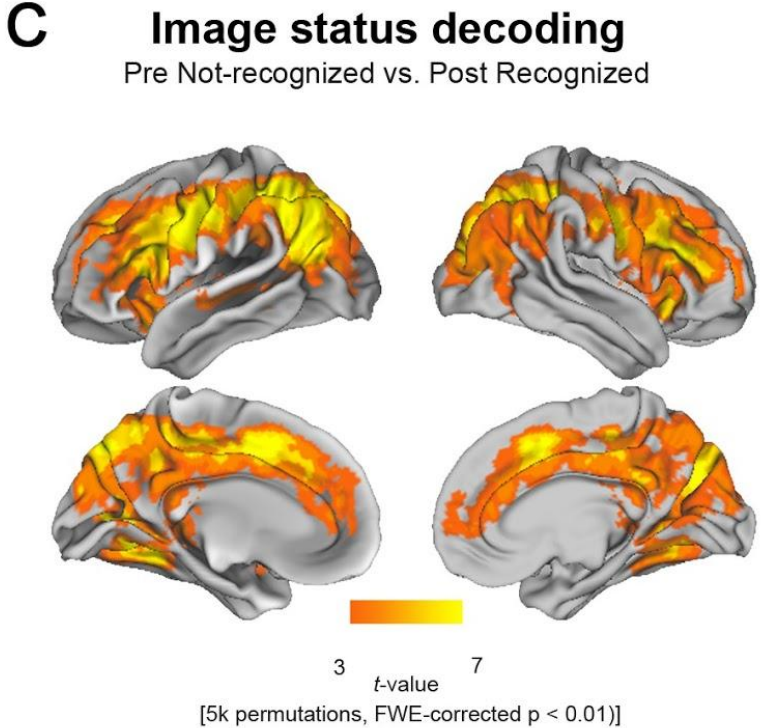
Flounders et al., 2019

Example: neuroscience



Flounders et al., 2019

Example: neuroscience



Gonzalez-Garcia et al., 2018

Conclusions

- SVM will always do the job, but it’s up to you to constrain it
- SVM interpretation is limited to input
- SVM is NOT the best method of machine learning, it is one method
- There are many ways to use SVM, LibSVM is one of them
  - SciKit Learn
  - PyMVPA

Future directions

- July 17, 2019: “Random forest: theory and application to neuroimaging” by Meichen Yu
  - Discussion of random forest technique
  - Comparison to Support Vector Machine
  - Best Practices and Pitfalls

References

- Pereira et al., 2009, Machine learning classifiers and fMRI: A tutorial overview, *Neuroimage* (link)
- Norman et al., 2006, Beyond mind-reading: multi-voxel pattern analysis of fMRI data, *TiCS* (link)
- Mur et al., 2009, Revealing representational content with pattern-information fMRI—an introductory guide, *Social Cognitive and Affective Neuroscience* (link)
- Gonzalez-Garcia et al., 2018, Content-specific activity in frontoparietal and default-mode networks during prior-guided visual perception, *eLife* (link)
- Flounders et al., 2019, Neural dynamics of visual ambiguity resolution by perceptual prior, *eLife* (link)

LibSVM package resources

- LibSVM Docs
  - Paper: https://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf
  - Beginners Guide: https://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf
  - Piaip's LIBSVM tutorial: https://www.csie.ntu.edu.tw/~piaip/svm/svm_tutorial.html#
  - MATLAB
    - Download directly from site
    - Example/datasets: https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/
      - https://www.youtube.com/watch?v=gePWtNAQcK8
  - R
    - Download: https://cran.r-project.org/web/packages/e1071/e1071.pdf
    - Example/datasets: https://www.csie.ntu.edu.tw/~cjlin/libsvm/R_example.html

MATLAB Resources

- General guide
  - https://www.mathworks.com/help/stats/support-vector-machines-for-binary-classification.html
- Functions of interest (with examples)
  - https://www.mathworks.com/help/stats/fitcsvm.html
- Example data and validation guidelines:
  - https://www.mathworks.com/help/stats/select-data-and-validation-for-classification-problem.html
- Choosing classifers:
  - https://www.mathworks.com/help/stats/choose-a-classifier.html#bunt0n0-1