Neural decoding of spike trains and local field potentials with machine learning in python

Omar Costilla-Reyes, PhD
Miller Lab
Picower institute for learning and memory
Brain and Cognitive Sciences
MIT

Let's get started!

- Get Anaconda (30 min install approx.):
- https://www.anaconda.com/distribution/

- Get the LFP and spike data:
- http://bit.ly/spikes-lfp-decoding

- Notebooks on github as well

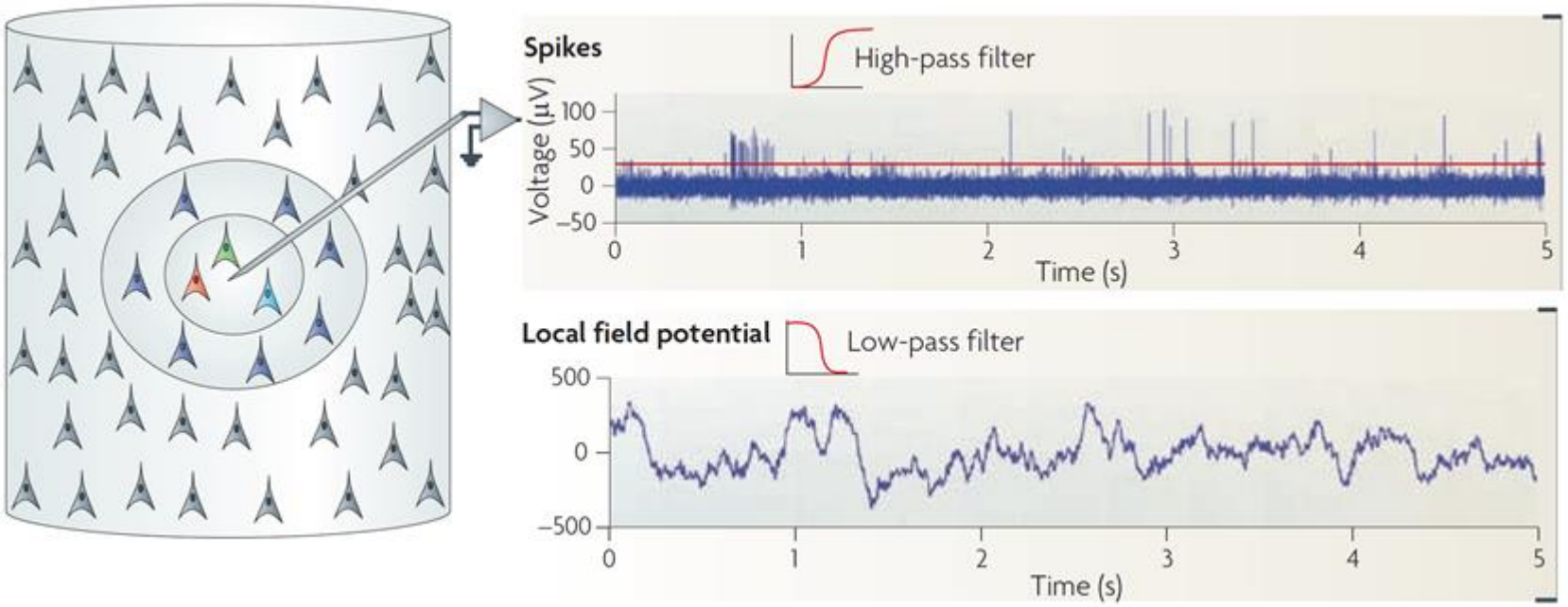
- If using your own Python distribution or older versions of Anaconda
  - Update scikit-learn to version 0.20.3
  - Update matplotlib to version 3.0.3

```
pip install scikit-learn –upgrade
pip install matplotlib --upgrade
```

Outline

- Intro and background
- Hands-on python tutorial
  - LFPs
  - Spikes
- If we have time:
  - Play with the notebooks: test suggested changes from you: changes in model, data, etc. it's ok to break the code ☺
- Feel free to ask questions

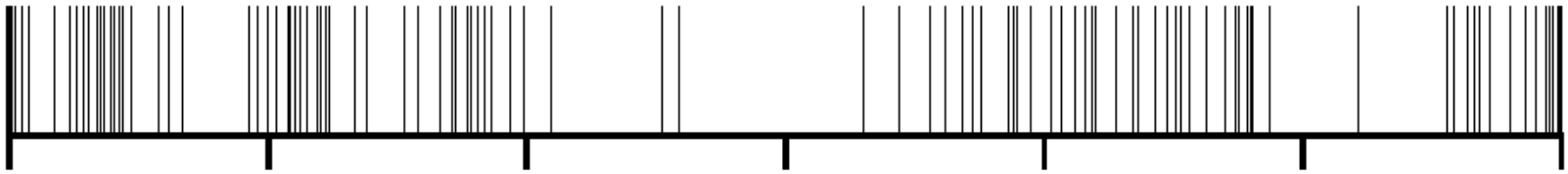
Spikes and local field potentials



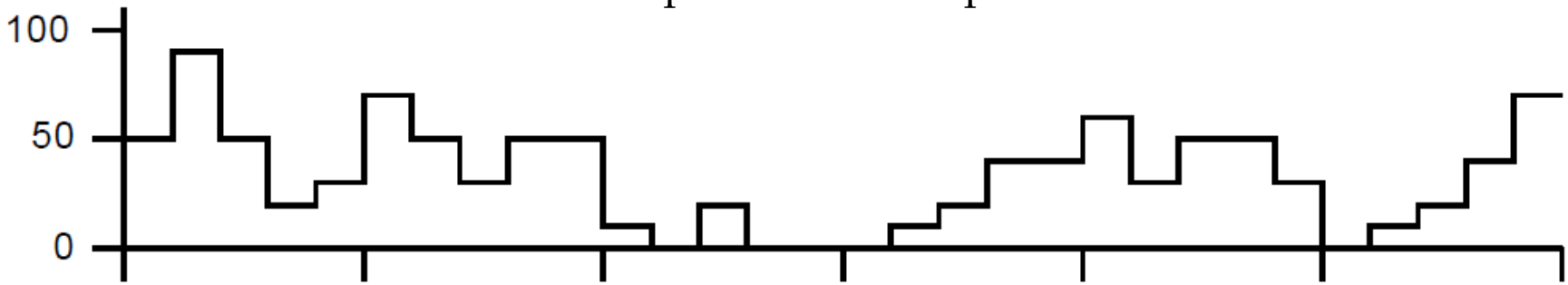
- **Spike:** single neuron recording
- **LFP:** summed electric current flowing from multiple nearby neurons

Quiroga, Rodrigo Quian, and Stefano Panzeri. "Extracting information from neuronal populations: information theory and decoding approaches." Nature Reviews Neuroscience 10.3 (2009): 173.

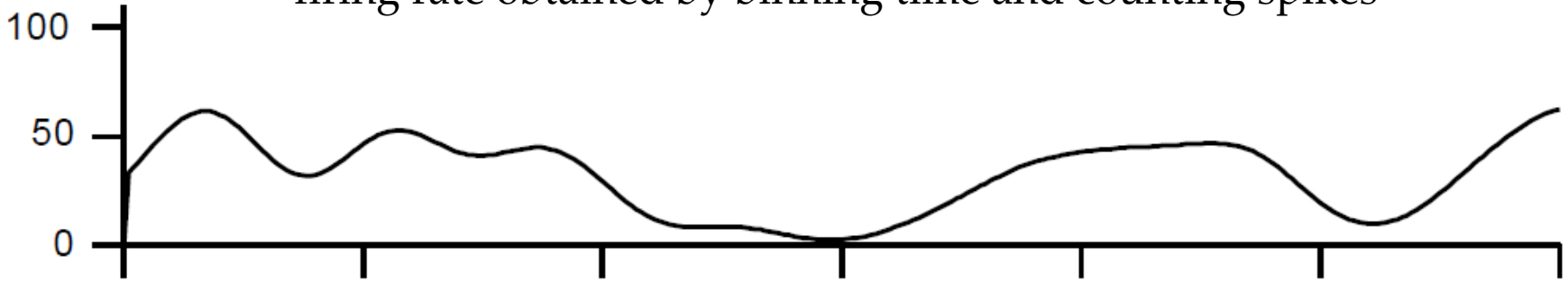
Spike rate



Spike train example



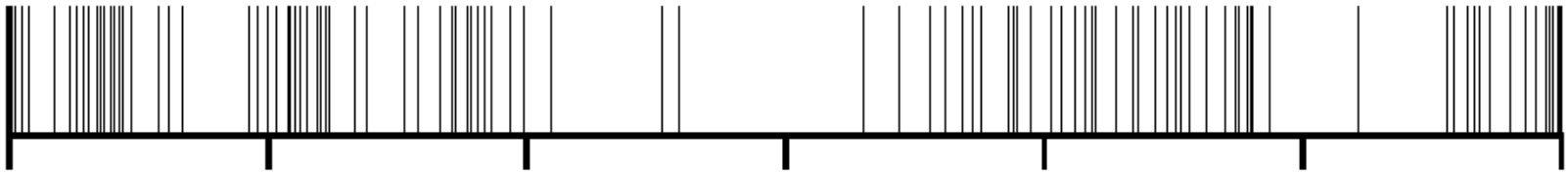
firing rate obtained by binning time and counting spikes



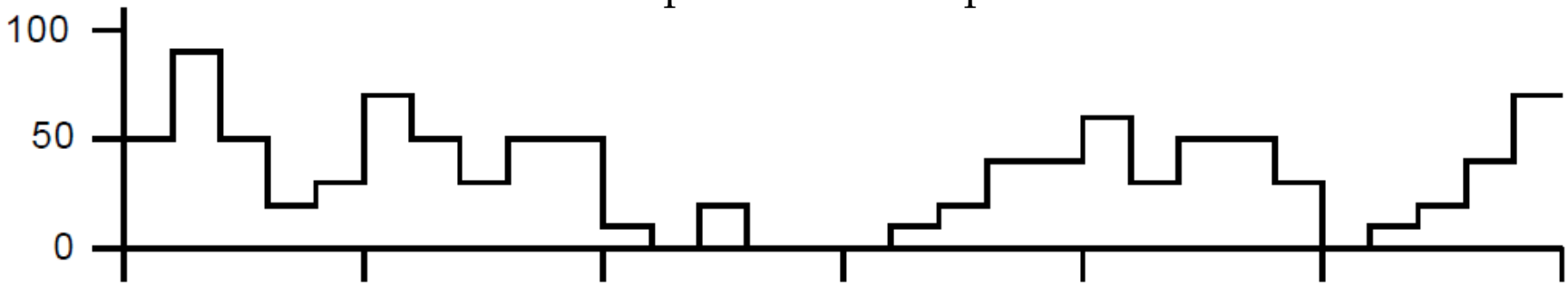
Approximate firing rate using a Gaussian window of 50 ms

Dayan, Peter, and Laurence F. Abbott. "Theoretical neuroscience: computational and mathematical modeling of neural systems." (2001).

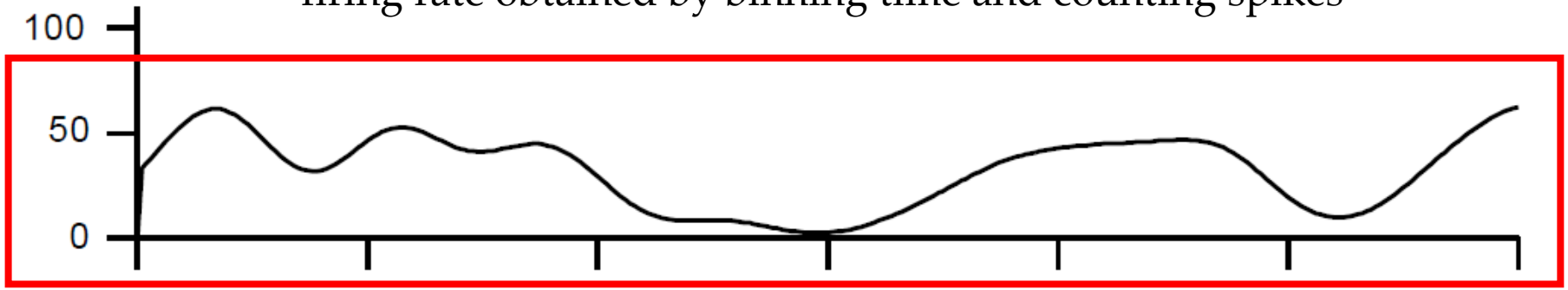
Spike rate



Spike train example



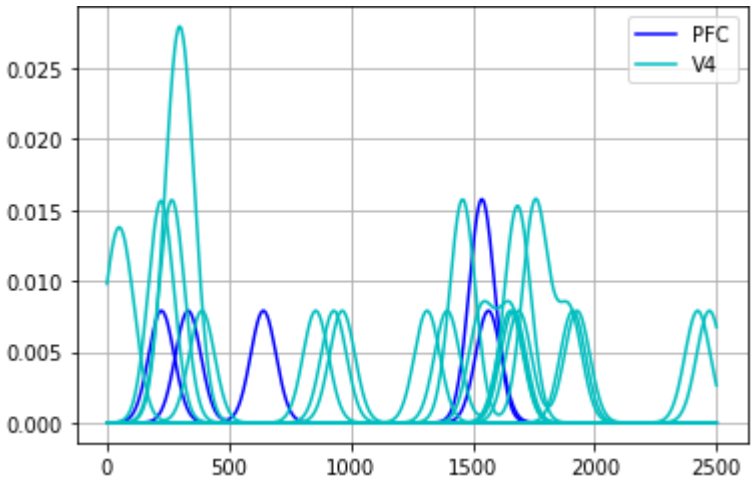
firing rate obtained by binning time and counting spikes



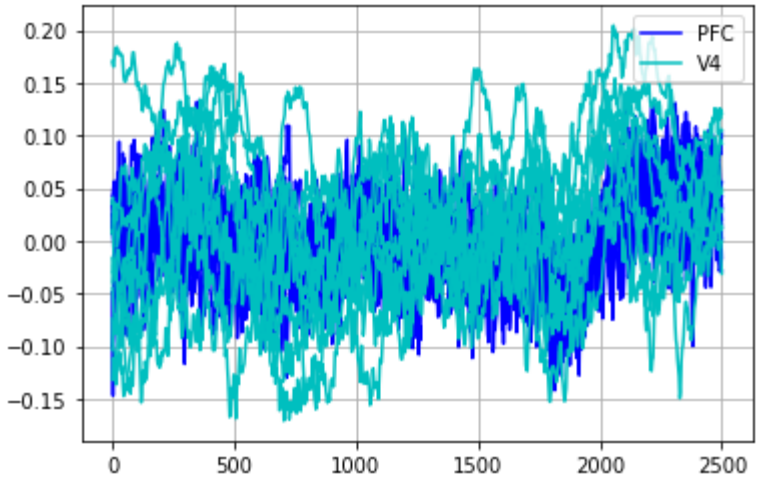
Approximate firing rate using a Gaussian window of 50 ms

Dayan, Peter, and Laurence F. Abbott. "Theoretical neuroscience: computational and mathematical modeling of neural systems." (2001).

Dataset's spike rate example



Dataset's LFP rate example

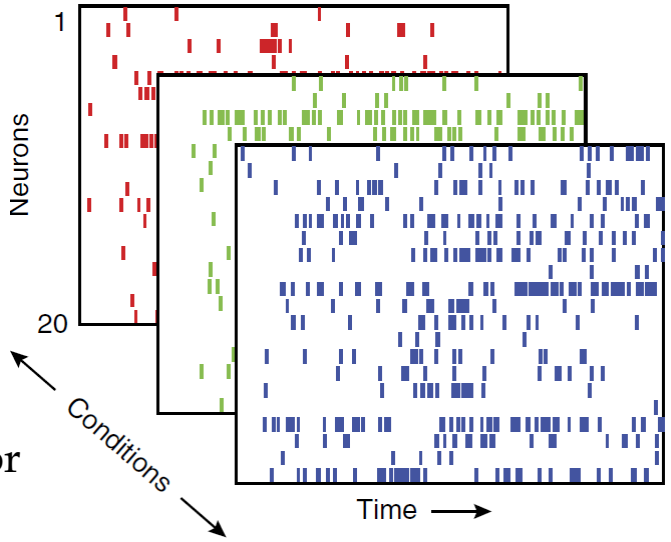


Neural encoding and decoding

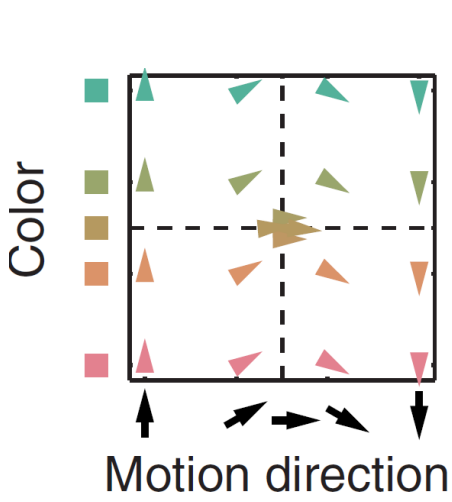
Encoding (brain)

Decoding

Predict stimulus or behavior
from neural activity



Neural encoding and decoding

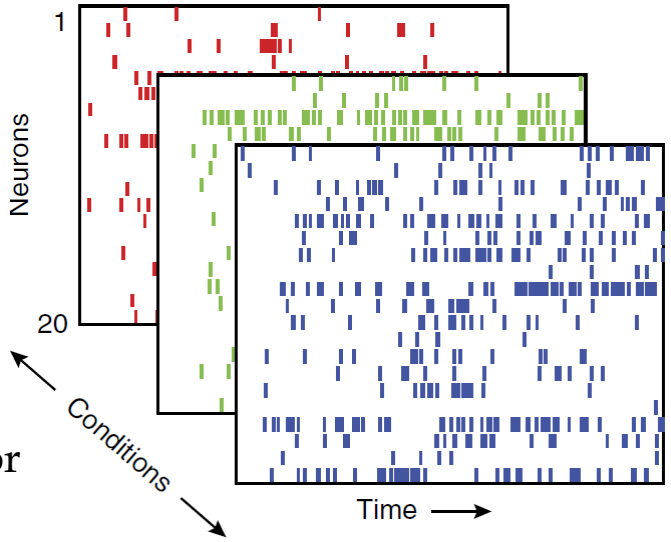


(Dataset stimulus
Shown to a monkey)

Encoding (brain)

Decoding

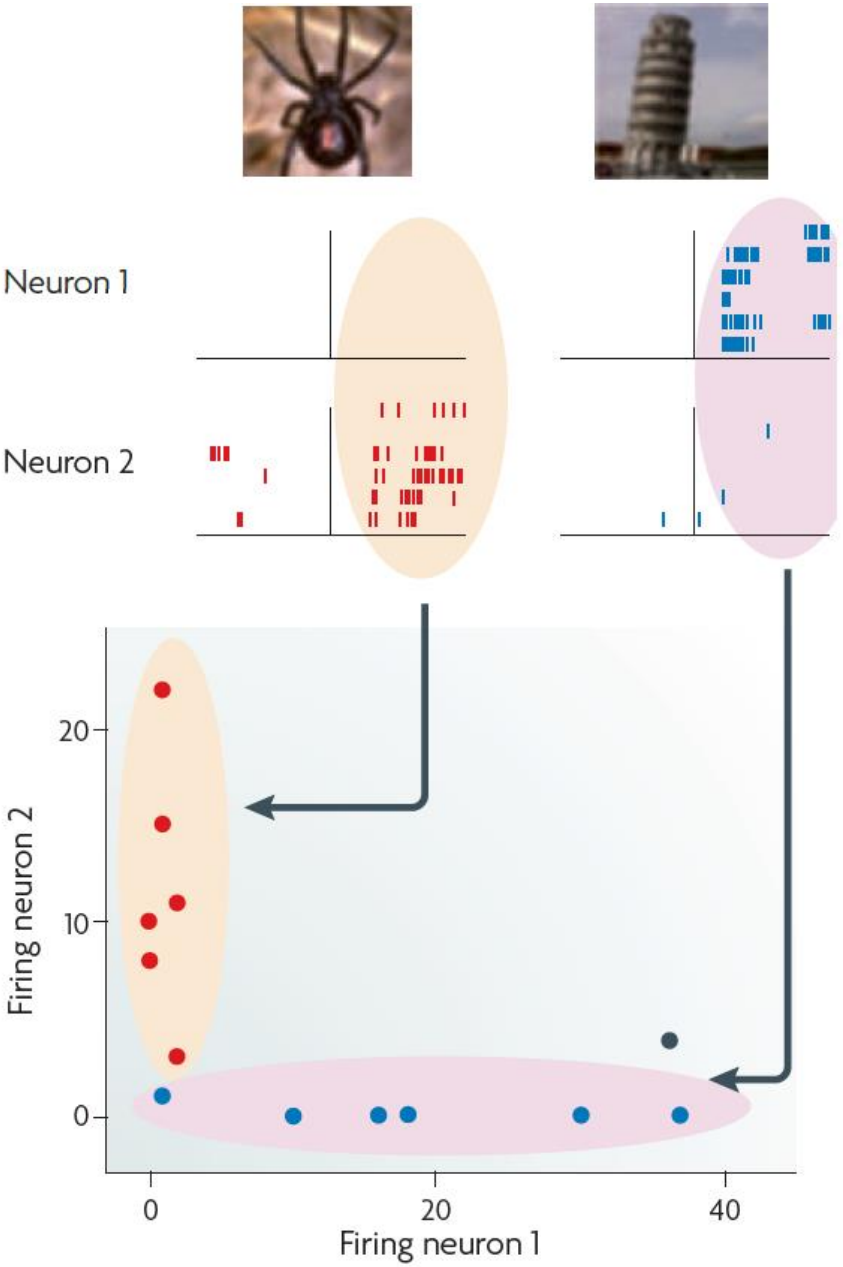
Predict stimulus or behavior
from neural activity



Decoding analysis

Responses of two neurons to two stimuli.

Each **trial** is represented as a point in a 2-dimensional space.



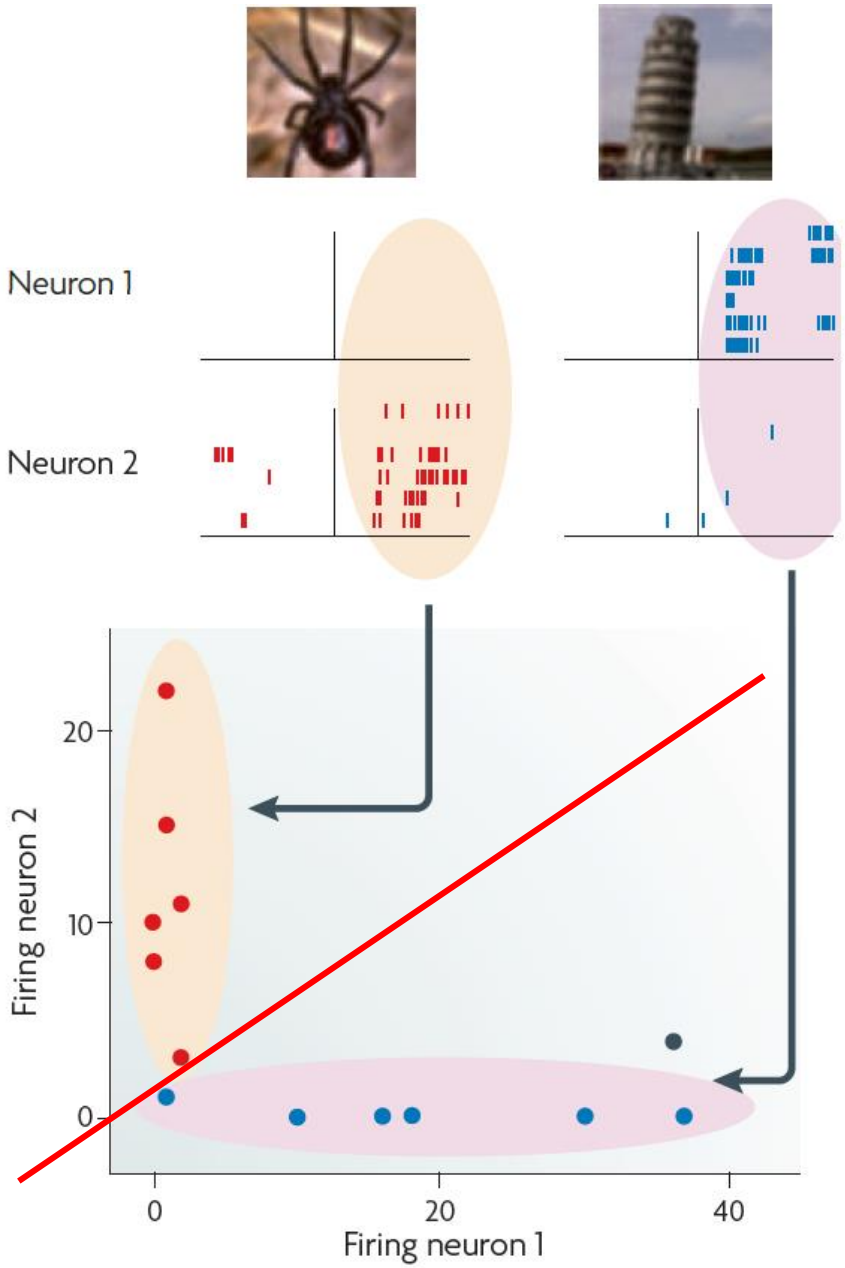
Quiroga, Rodrigo Quian, and Stefano Panzeri. "Extracting information from neuronal populations: information theory and decoding approaches." Nature Reviews Neuroscience 10.3 (2009): 173.

Decoding analysis

Responses of two neurons to two stimuli.

Each trial is represented as a point in a 2-dimensional space.

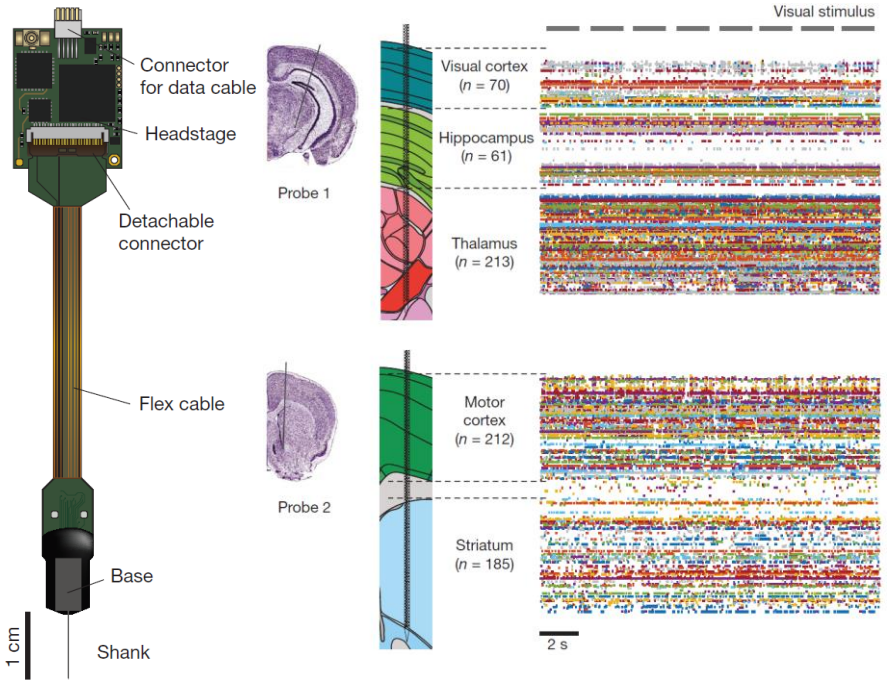
Linear machine learning model

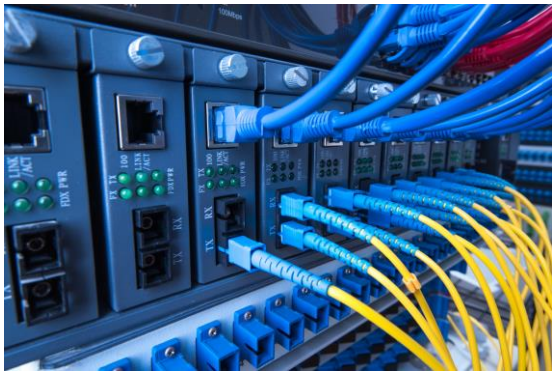


Quiroga, Rodrigo Quian, and Stefano Panzeri. "Extracting information from neuronal populations: information theory and decoding approaches." Nature Reviews Neuroscience 10.3 (2009): 173.

Why neural decoding *now*?

Single neuron vs multiple neuron analysis (neuropixels)



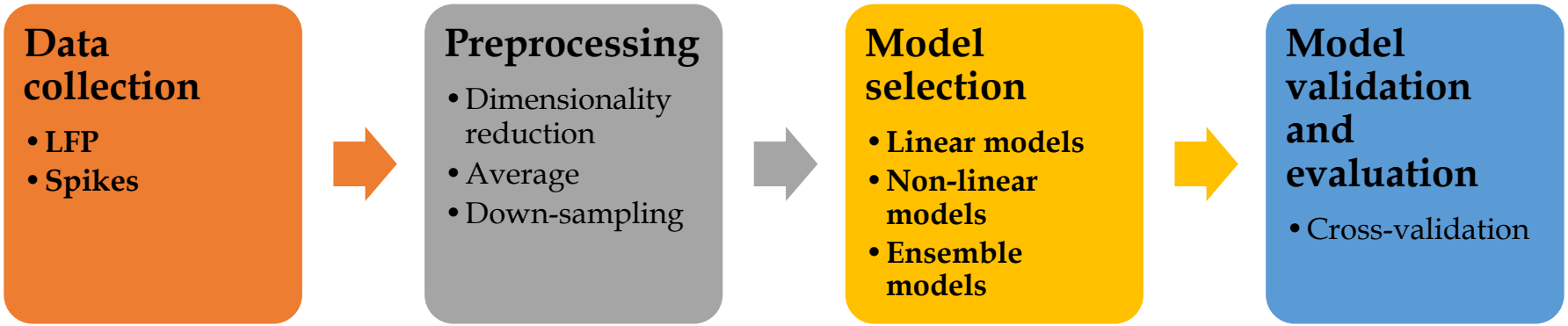


python

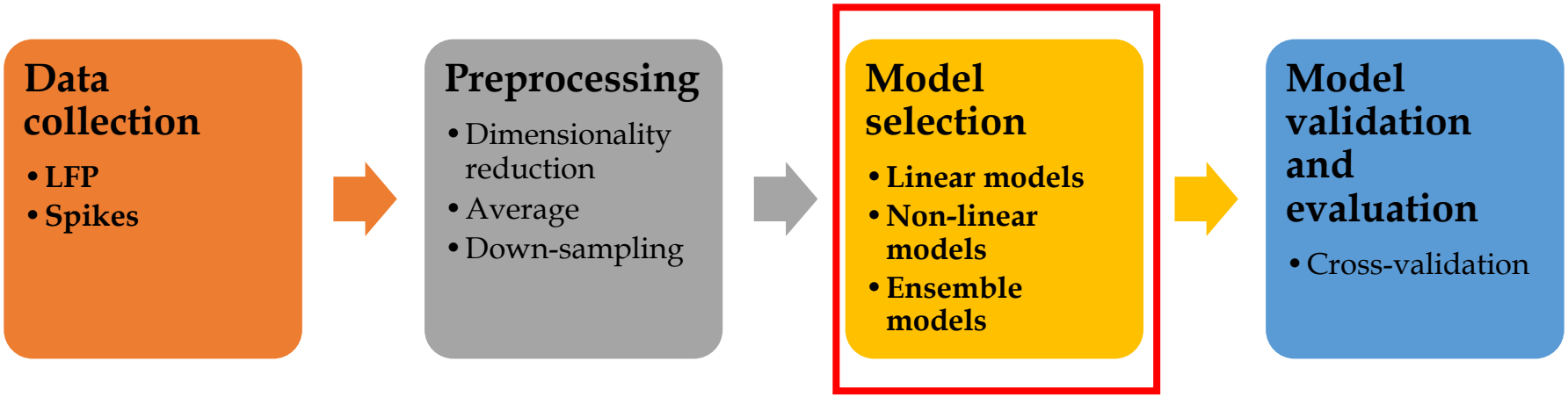
741 neurons were recorded simultaneously

Jun, James J., et al. "Fully integrated silicon probes for high-density recording of neural activity." *Nature* 551.7679 (2017): 232.

Decoding analysis pipeline

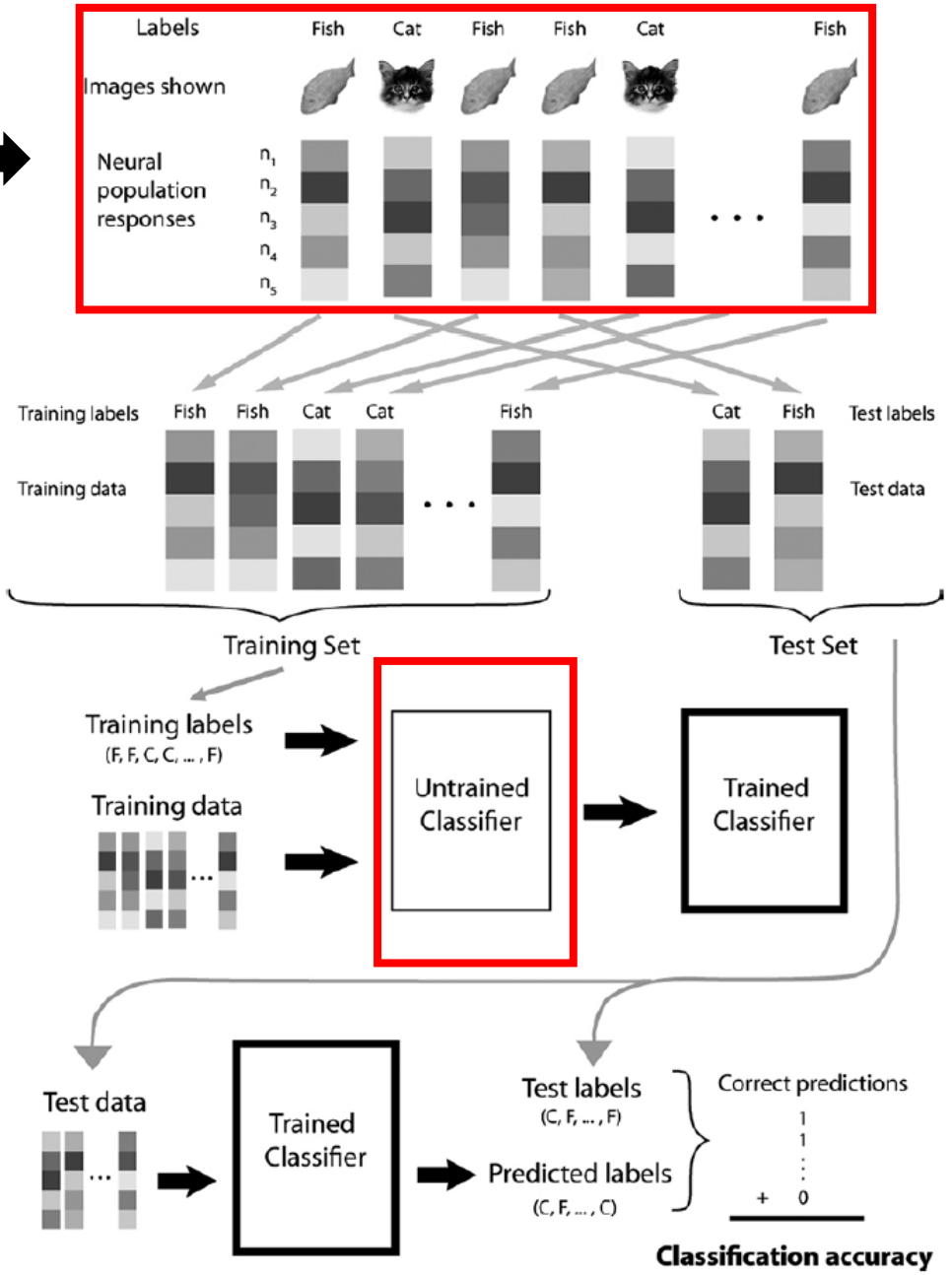


Decoding analysis pipeline



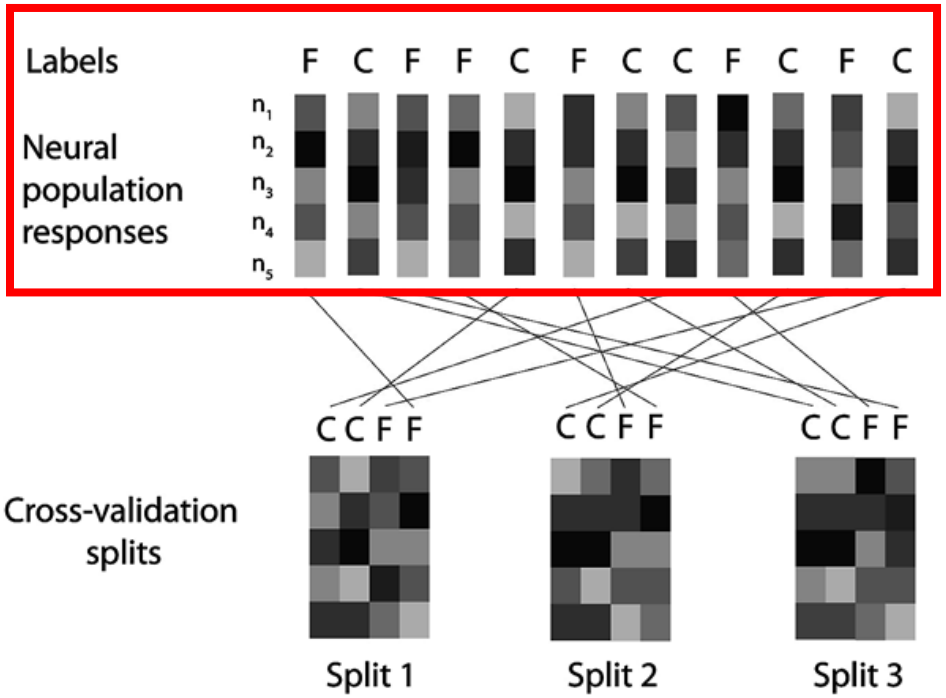
Fish/cat dataset

Model selection and training

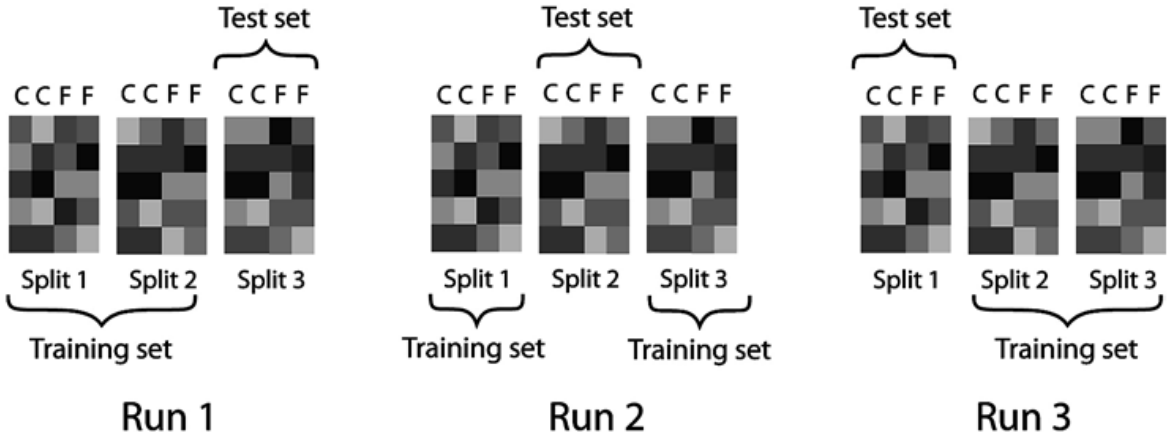


Meyers, E. M., and Gabriel Kreiman. "Tutorial on pattern classification in cell recording." *Visual population codes*(2012): 517-538.

Fish/cat dataset



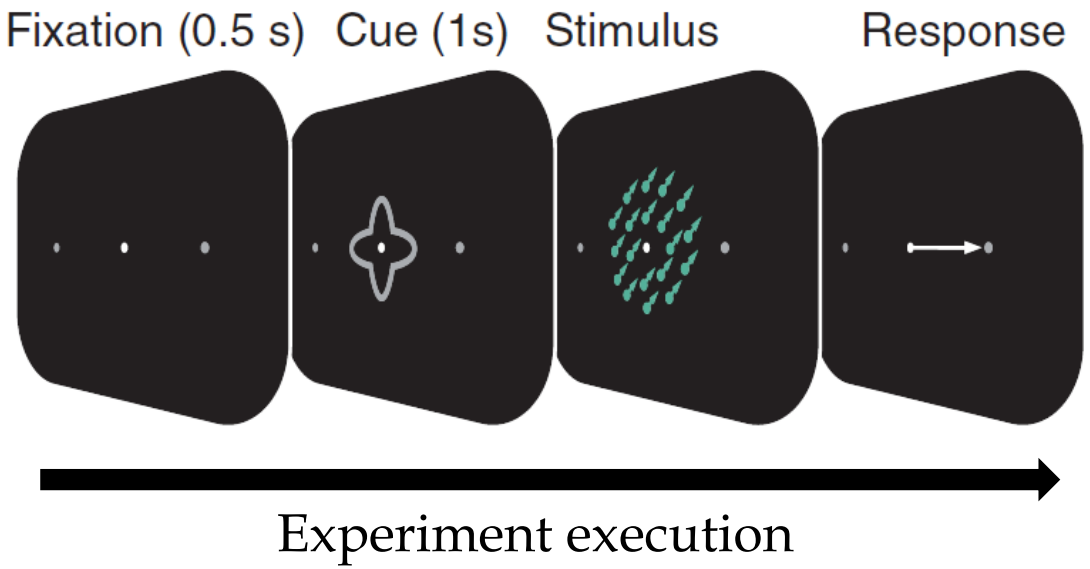
3-fold Cross-validation



Meyers, E. M., and Gabriel Kreiman. "Tutorial on pattern classification in cell recording." *Visual population codes*(2012): 517-538.

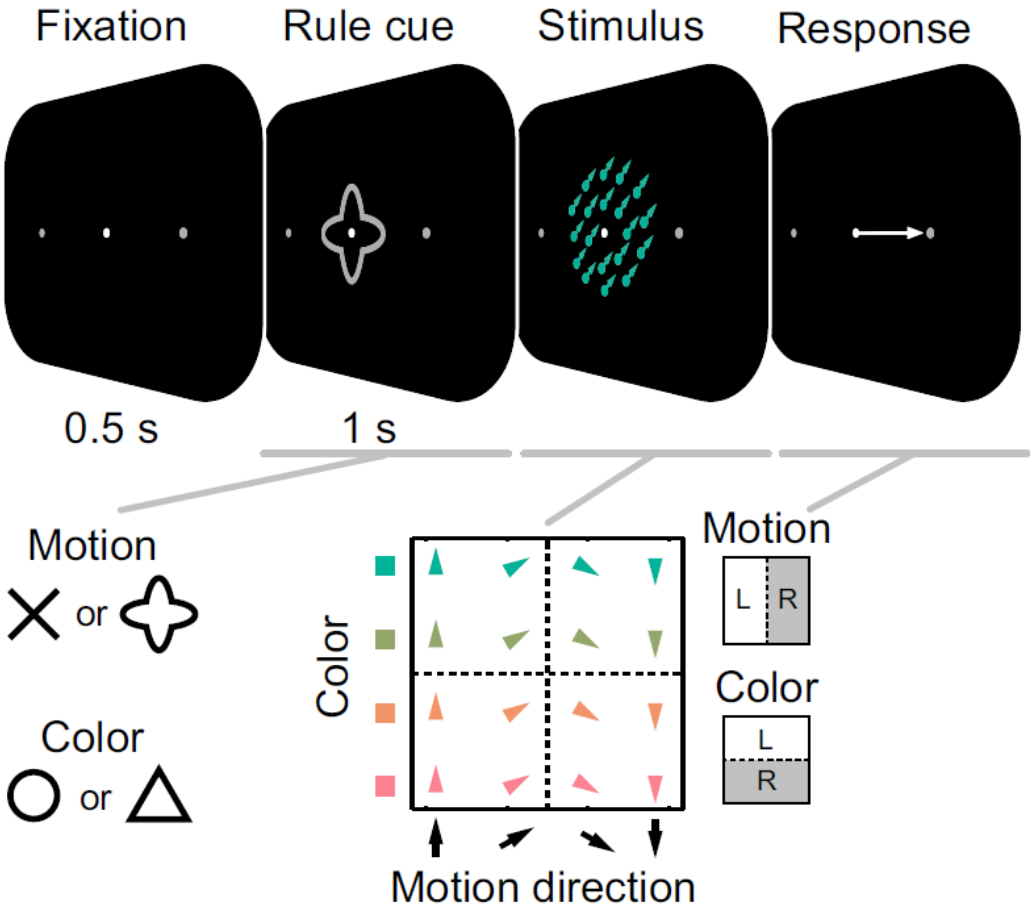
Motion/color categorization dataset

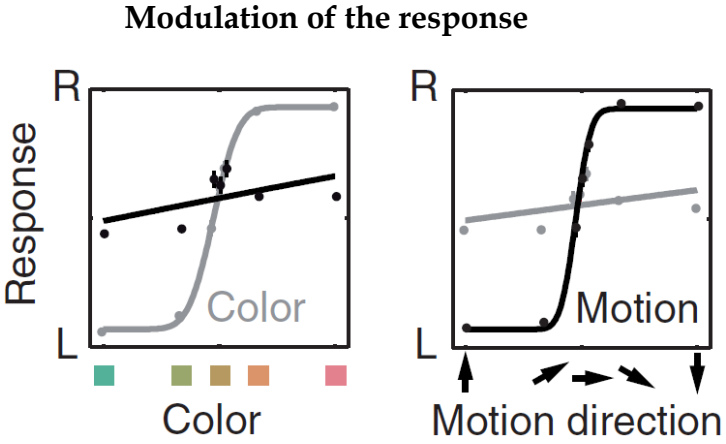
- monkeys categorized motion or color of centrally presented stimuli.



Siegel, Markus, Timothy J. Buschman, and Earl K. Miller. "Cortical information flow during flexible sensorimotor decisions." *Science* 348.6241 (2015): 1352-1355.

Experiment rules



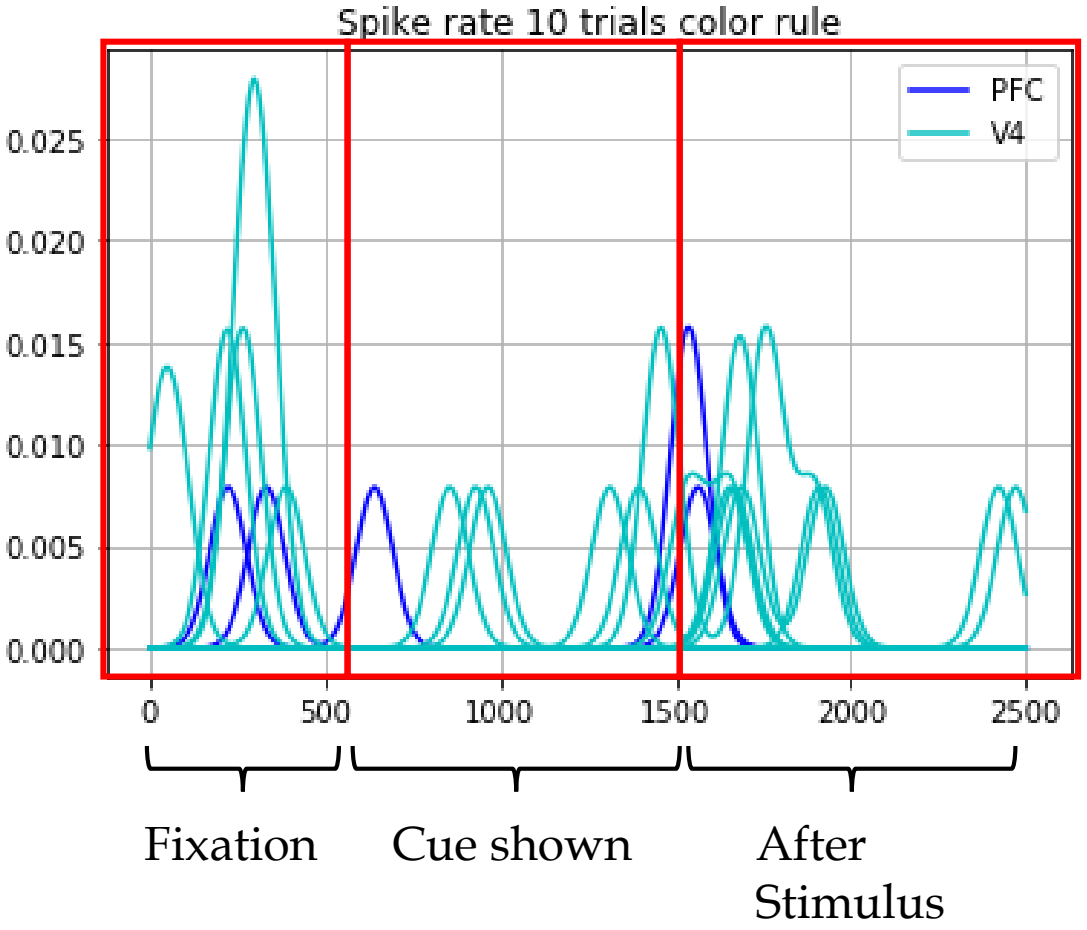


Two different cue shapes cued each task.

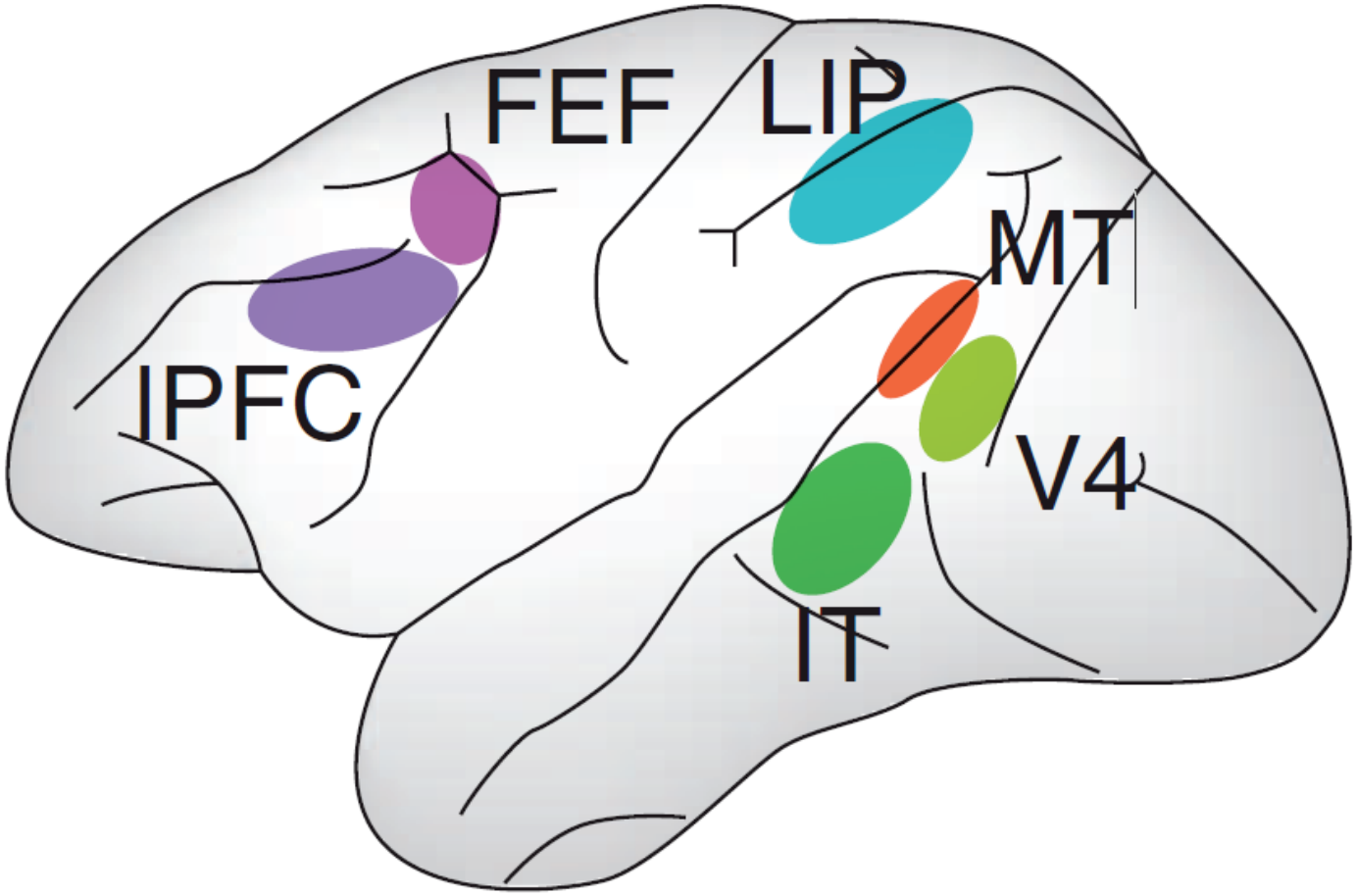
Stimuli systematically sampled motion direction (upward to downward) and color (green to red).

Brincat, Scott L., et al. "Gradual progression from sensory to task-related processing in cerebral cortex." *Proceedings of the National Academy of Sciences* 115.30 (2018): E7202-E7211.

Experimental progression in spike trials (10 samples)

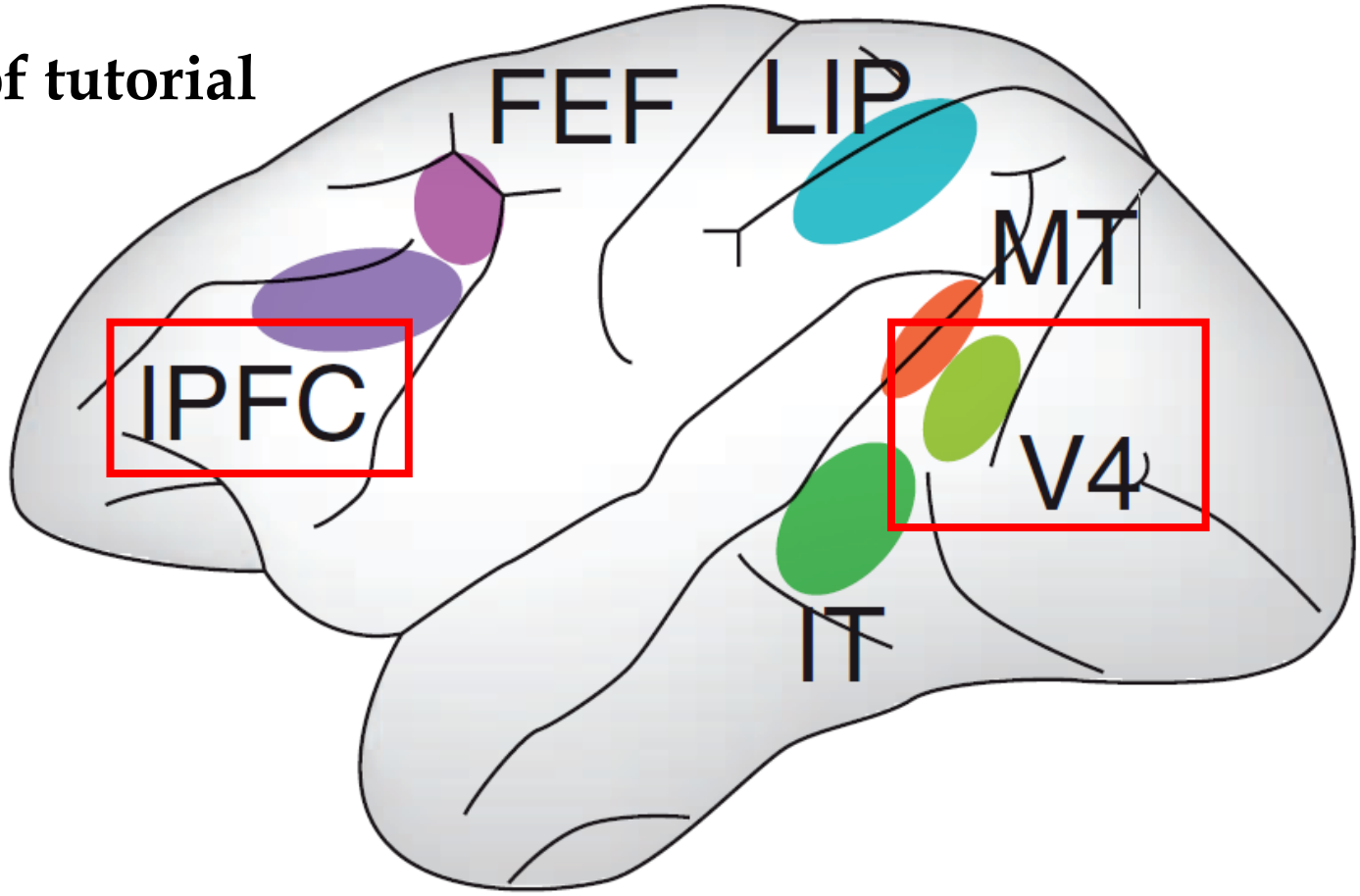


Recorded brain regions



Recorded brain regions

Focus of tutorial



Tutorial objectives

- Spikes and LFPs decoding with Jupyter notebooks

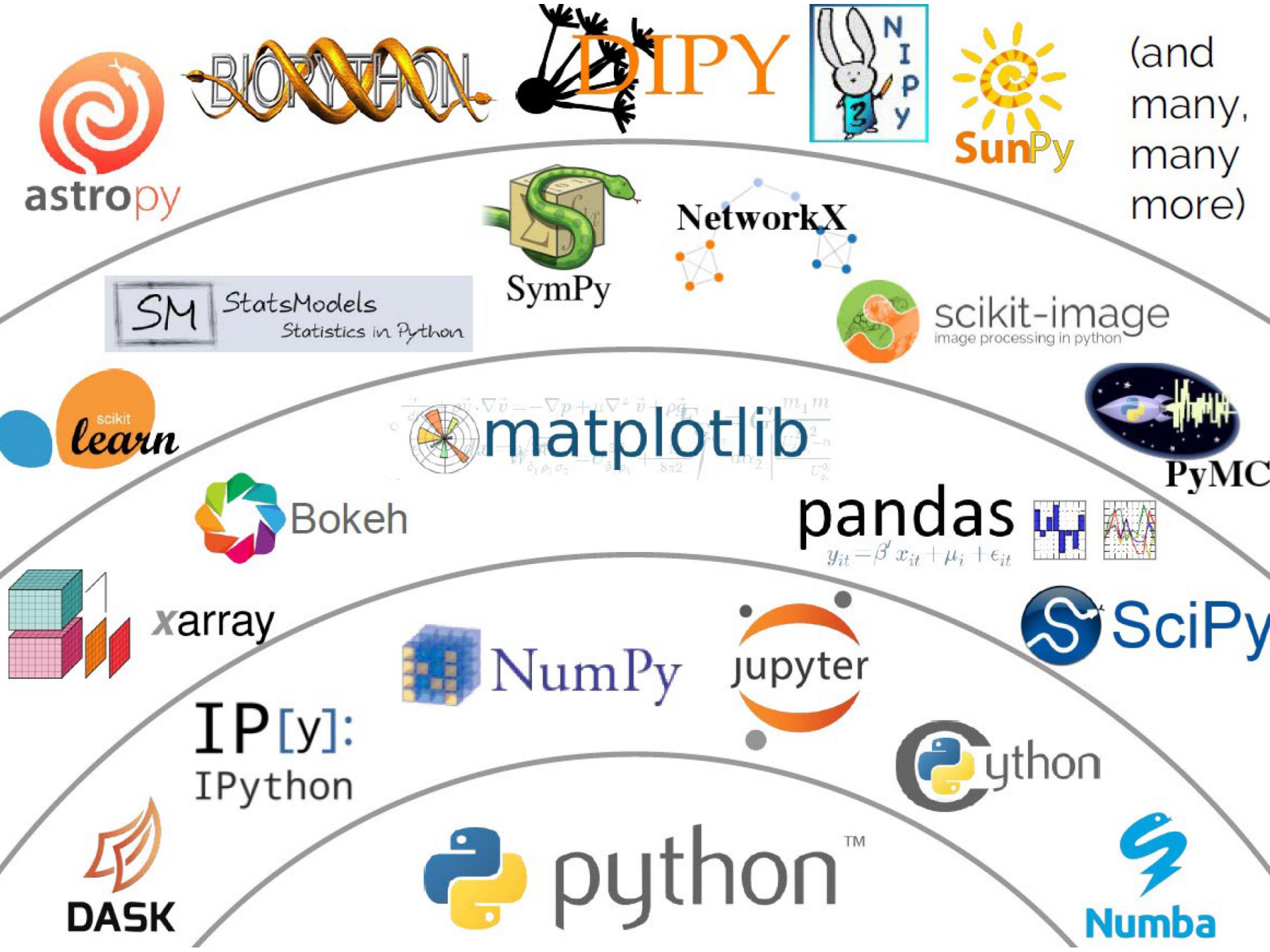
- **Decode color and motion** in:
  - V4 and PFC
  - Single experimental session, small dataset (out of 44)

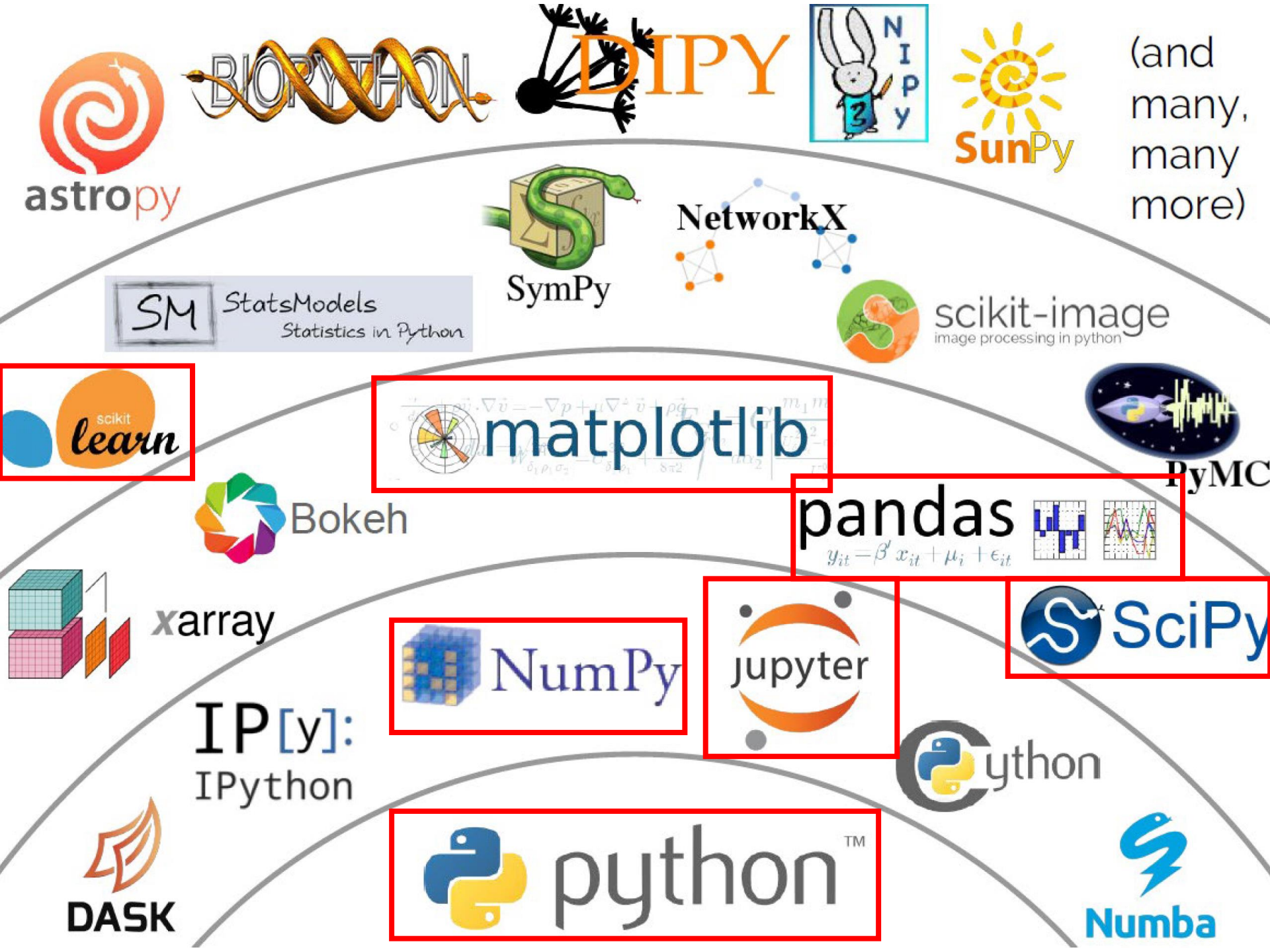
- Perform cross-validation

- Evaluate the accuracy of the model with f-score:
  - Considers both the precision and recall
  - https://scikit-learn.org/stable/modules/generated/sklearn.metrics.f1_score.html

Why python?

Slide credit: https://twitter.com/jakevdp


astropy
BIOPYTHON
DIPY
NIPY
SunPy
(and many, many more)
SymPy
NetworkX
SM StatsModels Statistics in Python
scikit-image image processing in python
scikit learn
matplotlib
PyMC
Bokeh
pandas
xarray
NumPy
jupyter
SciPy
IP[y]: IPython
Cython
DASK
python™
Numba


astropy
BIOPYTHON
DIPY
NIPY
SunPy
(and many, many more)
SymPy
NetworkX
SM StatsModels Statistics in Python
scikit-image image processing in python
scikit learn
matplotlib
PyMC
Bokeh
pandas
xarray
NumPy
jupyter
SciPy
IP[y]: IPython
Cython
DASK
python™
Numba

Free software machine learning library for Python

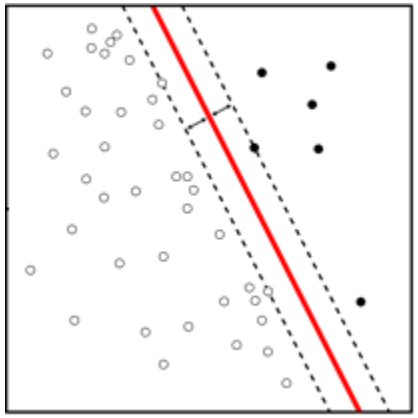
https://scikit-learn.org/stable/

Machine learning model: linear

- *sklearn.svm.LinearSVC*
- Example: 2D data. Constructs a hyper-plane, which can be used for classification.

- The hyper-plane has the largest distance to the nearest training data points of any class (so-called functional margin)

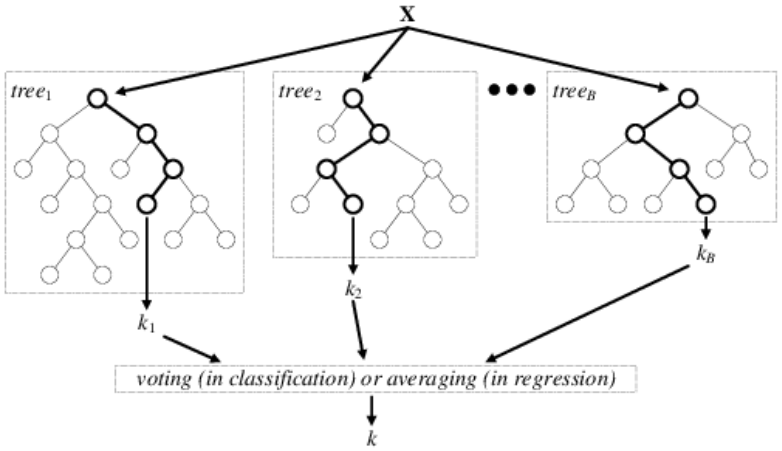
- The larger the margin the lower the generalization error of the classifier.



https://scikit-learn.org/stable/modules/svm.html#svm-classification

Machine learning model: Ensemble

- *sklearn.ensemble.ExtraTreesClassifier*
  - *F*ast computational version of the random forest
- combines the predictions of several base estimators
- Prediction of the ensemble is given as the averaged prediction of the individual classifiers (decision trees)
- Improve generalizability / robustness over a single estimator.



https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.ExtraTreesClassifier.html

cross_validate function

- *sklearn.model_selection.cross_validate(estimator, X_window, y_motion_color_area, scoring=scoring, cv=cv, return_train_score=False, n_jobs=-1)*

- Evaluate metric(s) by cross-validation and also record fit/score times.

- **X:** data
- **y:** labels
- **scoring:** *make_scorer a* callable object for f-score scoring in motion (label 0) and color (label 1)
- ***cv:*** cross-validation splitting strategy
- **n_jobs:** The number of CPUs to use to do the computation **-1** means using all processors.

https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.cross_validate.html

Demo

- **Unzip LFP files and open:**
  - BCS_decoding_tutorial-v5-lfp.ipynb

- **Unzip Spikes files and open:**
  - BCS_decoding_tutorial-v5-spikes.ipynb

Further resources

- MATLAB decoding toolbox
  - www.readout.info
- CBMM Tutorial: Using Decoding to Understand Neural Algorithms (MATLAB)
  - https://cbmm.mit.edu/learning-hub/tutorials/computational-tutorial/decoding-analyses-understand-neural-content-and-coding
- Data Analysis Baseline Library
- (Python - higher level than scikit-learn)
  - https://amueller.github.io/dabl/index.html

Thank you!

Q&A

costilla@mit.edu

@konet