

# Machine Learning Course



Part 1: Models in Computational Neuroscience

Part 2: Vision (L Perrinet)

Part 3: Supervised Learning

Part 4: Unsupervised Learning

Part 5: Time Series Classification

[https://etulab.univ-amu.fr/gilson.m/courseml\\_phd2023/](https://etulab.univ-amu.fr/gilson.m/courseml_phd2023/)

# **Part 1: Models in Computational Neuroscience**

Part 2: Vision (L Perrinet)

Part 3: Supervised Learning

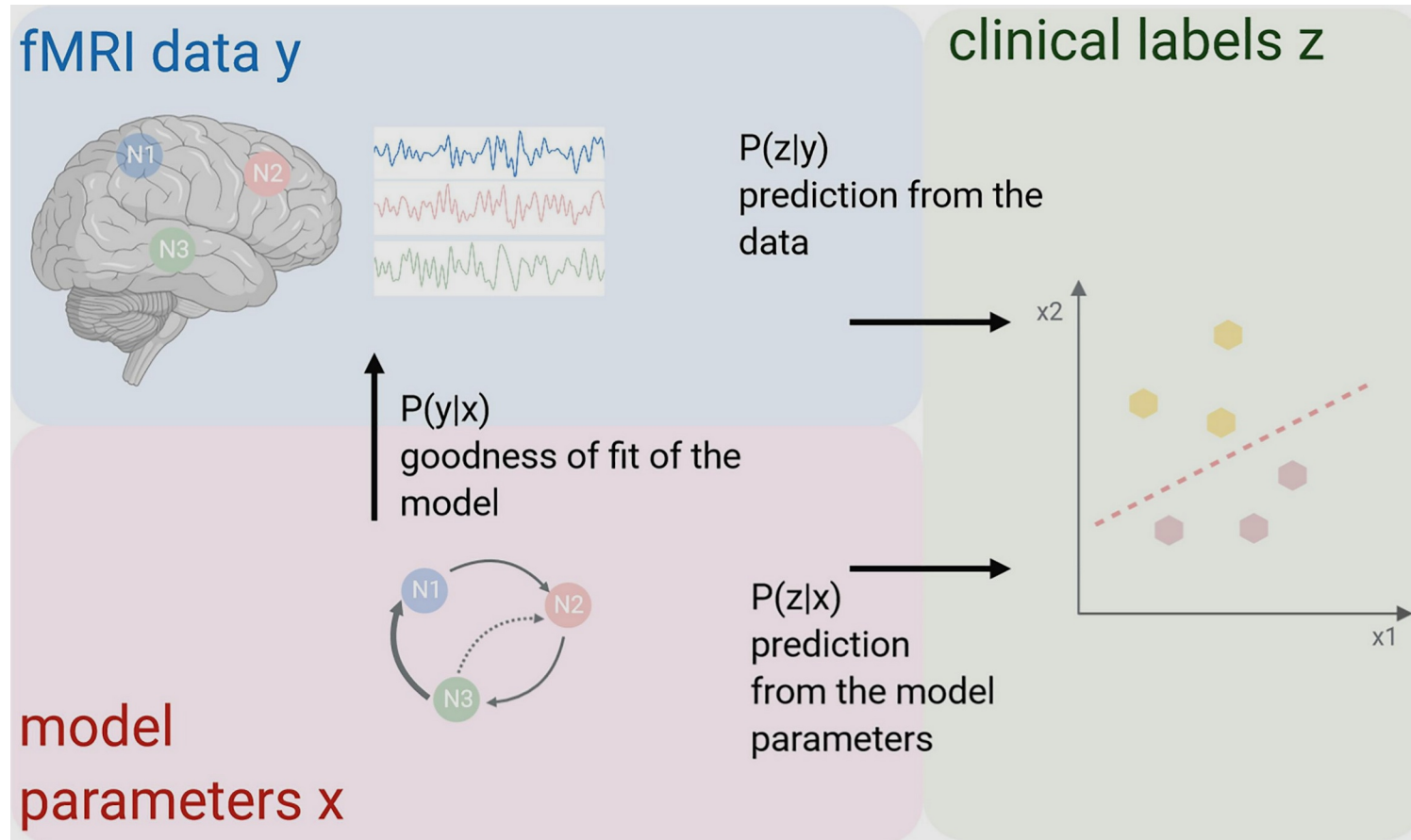
Part 4: Unsupervised Learning

Part 5: Time Series Classification

# Part 1: Models in Computational Neuroscience

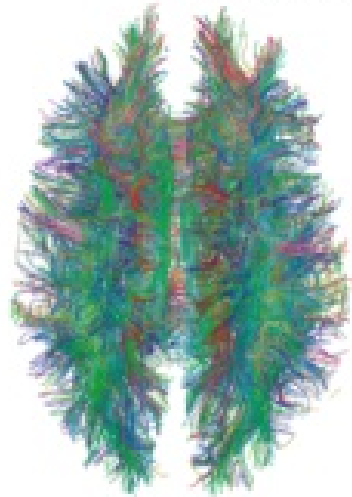
- Brief review of models in computational neuroscience
- Example 1: classification for diagnosis / prognosis
- Example 2: characterize structure in multivariate data
- Scikit-learn: formating data

# Models: Statistical, Biophysical, etc.

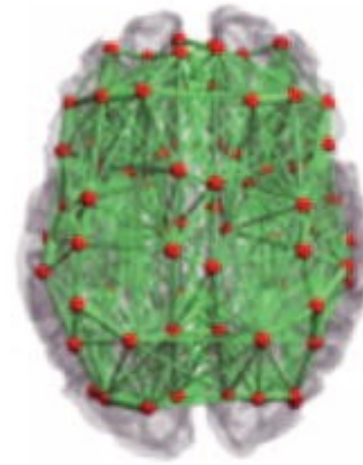


# Example of Biophysical Model: The Virtual Brain

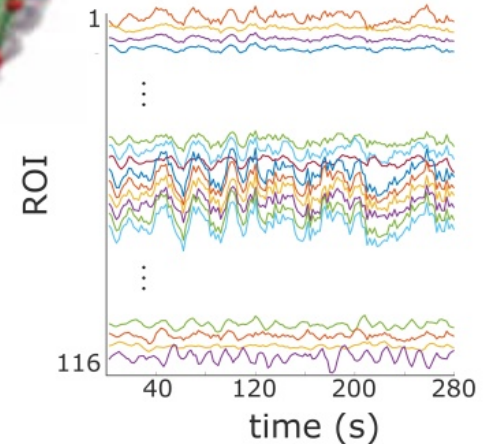
Anatomical connectivity



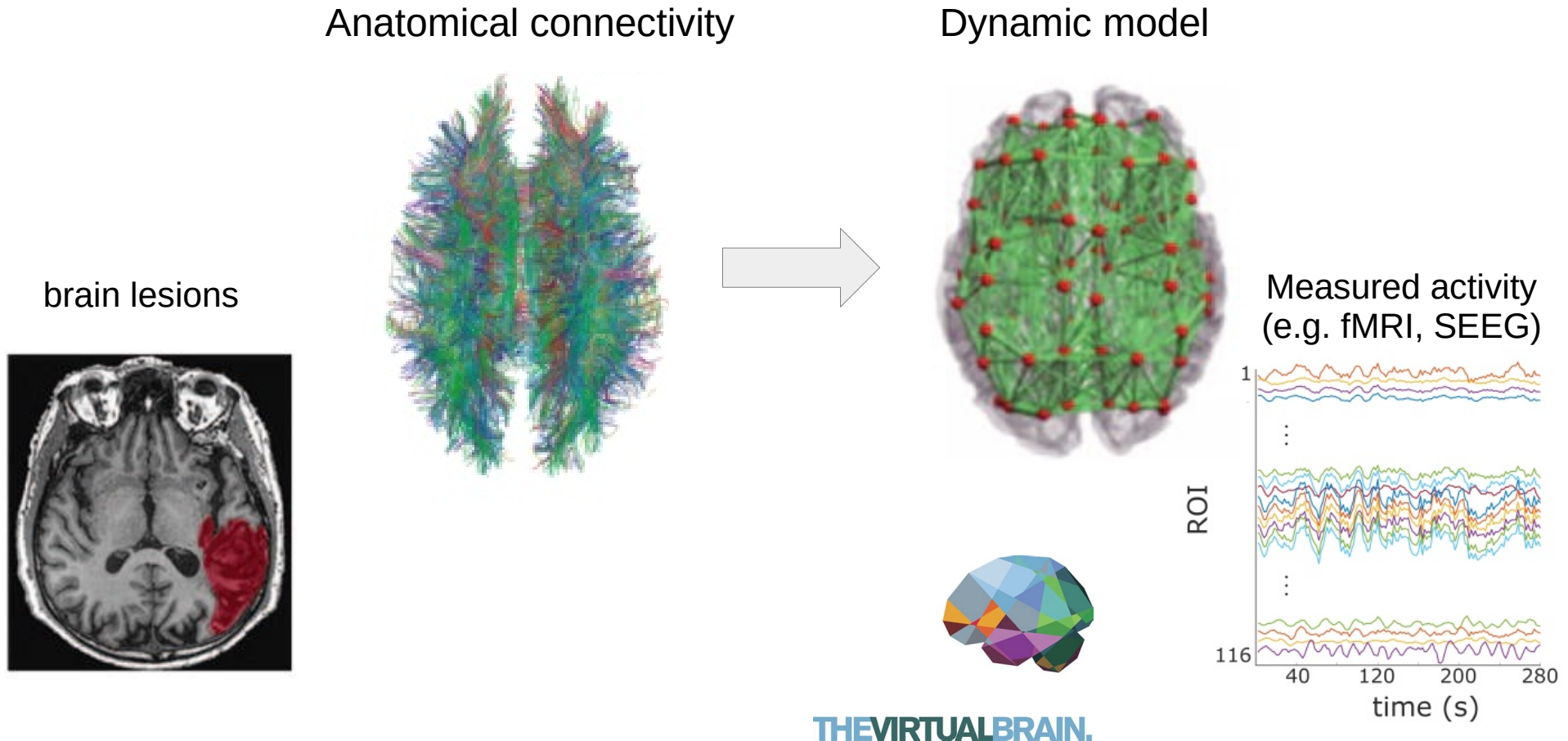
Dynamic model



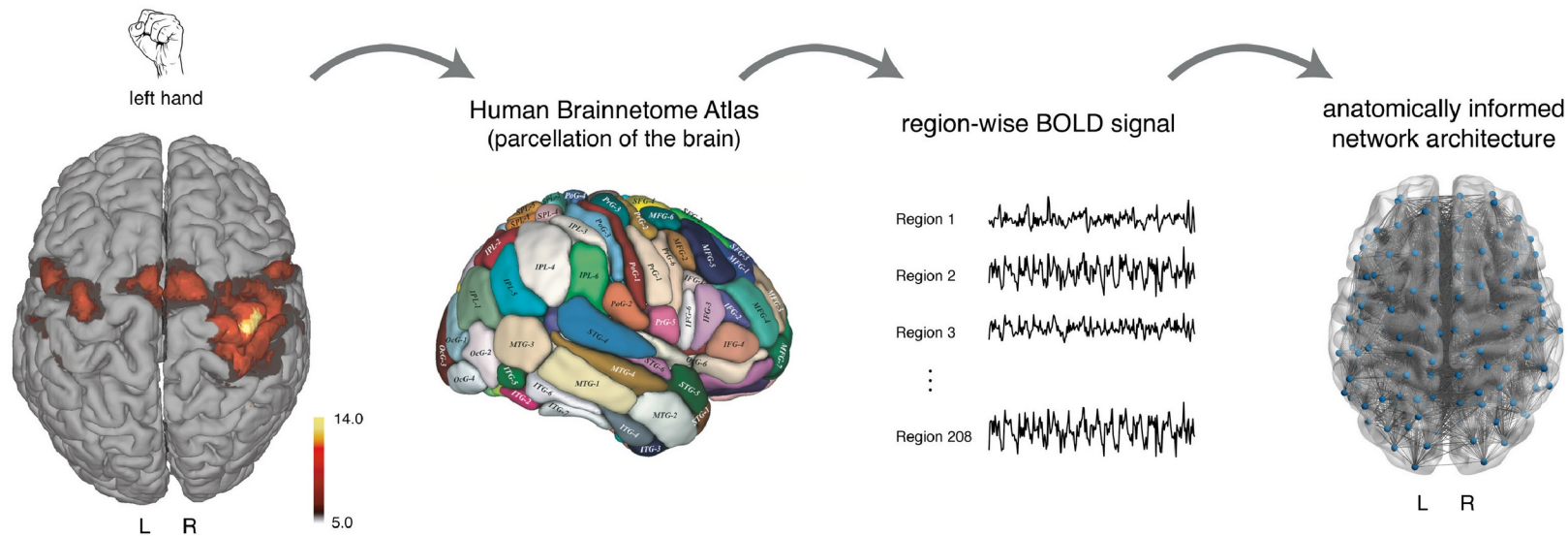
Measured activity  
(e.g. fMRI, SEEG)



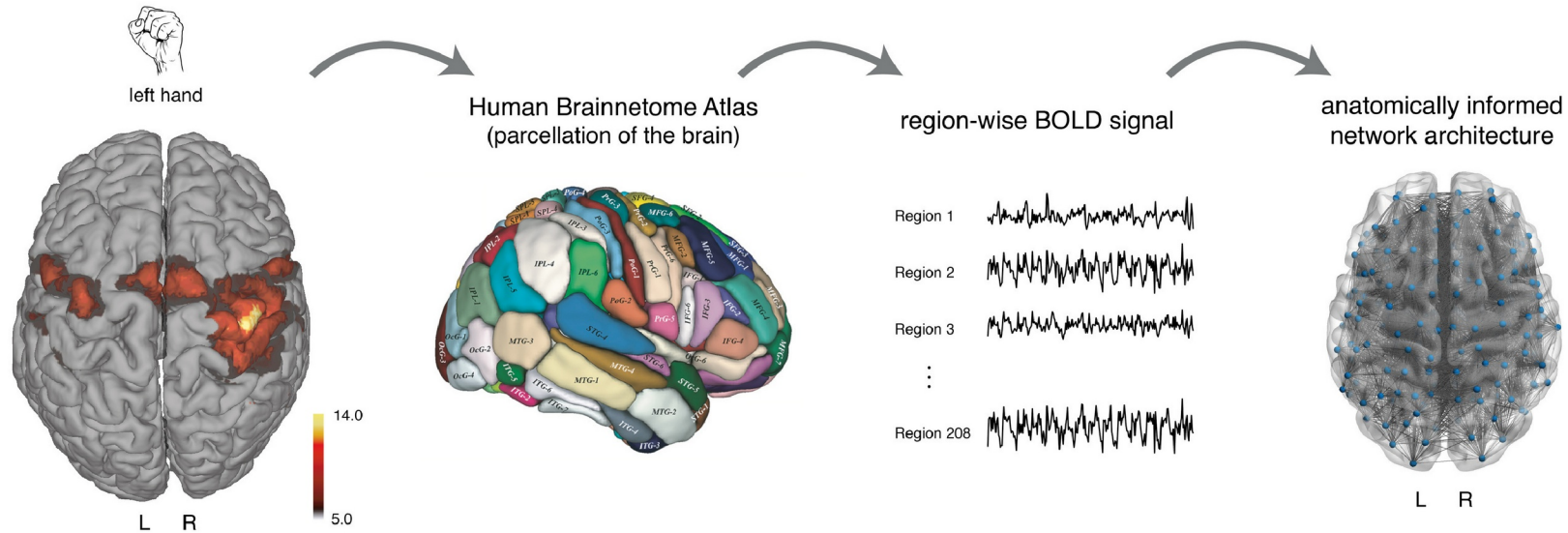
# Example of Biophysical Model: The Virtual Brain



# Example of Biophysical Model: Dynamic Causal Model

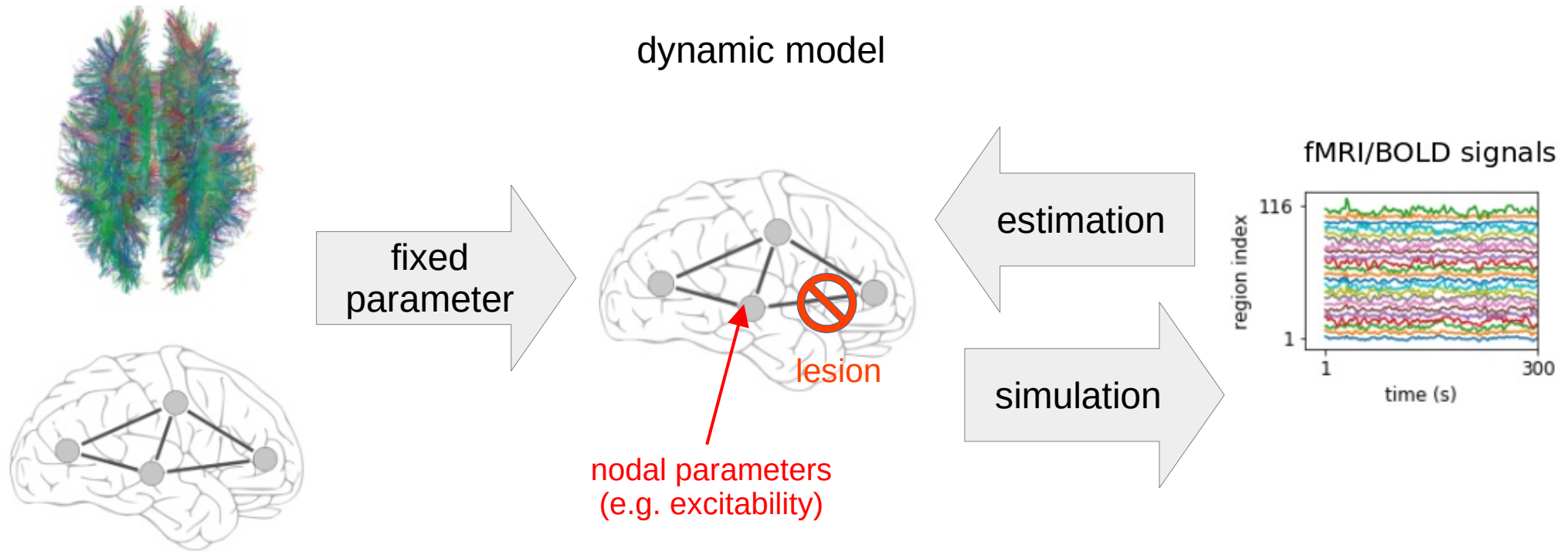


# Example of Biophysical Model: Dynamic Causal Model

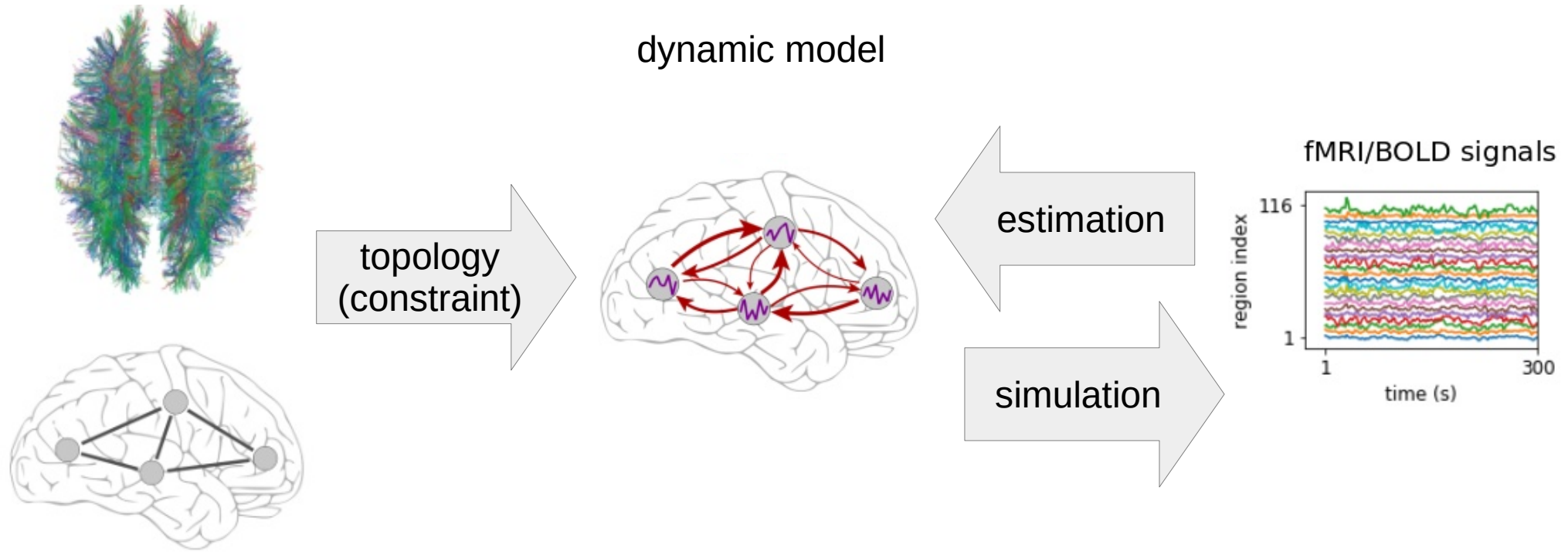


- Which brain interactions explain the data?
- Processing hierarchy of brain regions

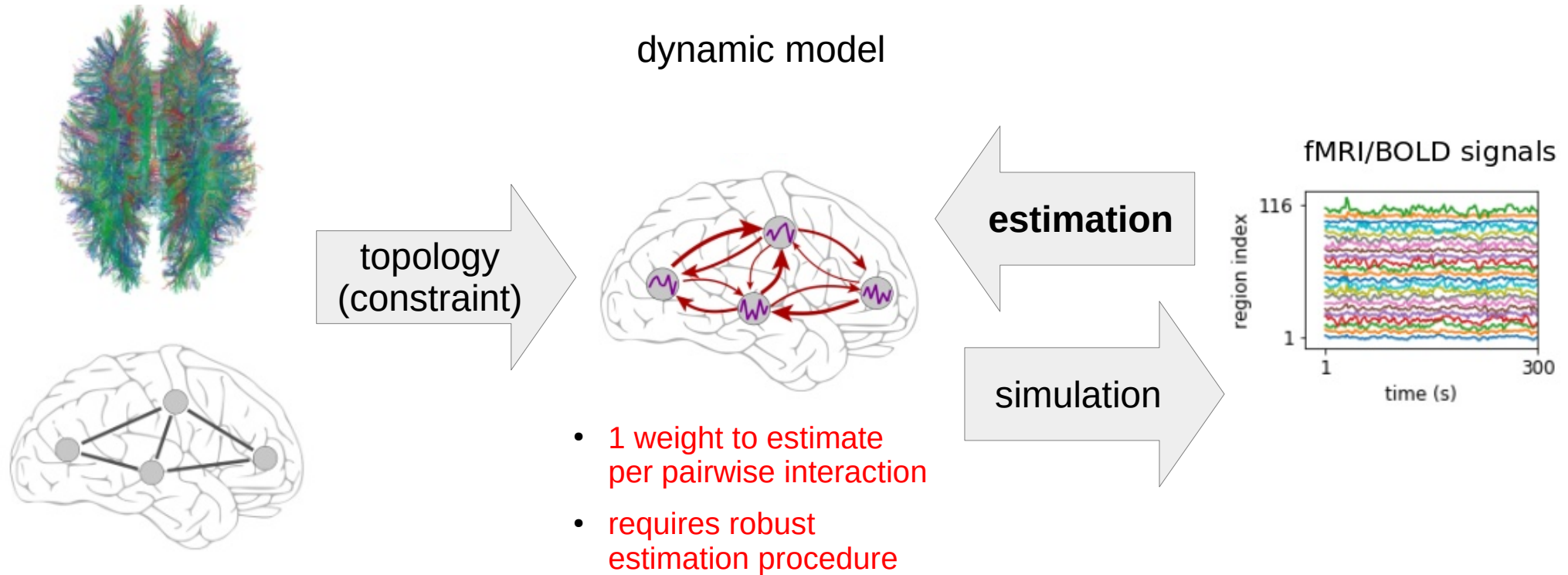
# Different Aims, Different Methods...



# Different Aims, Different Methods...



# Different Aims, Different Methods...

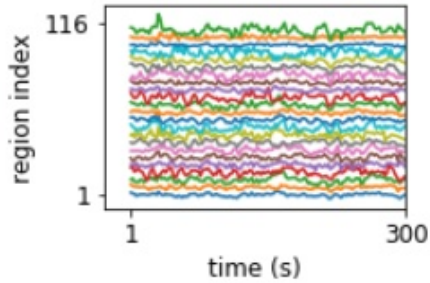


# Statistical Models to Relate Data

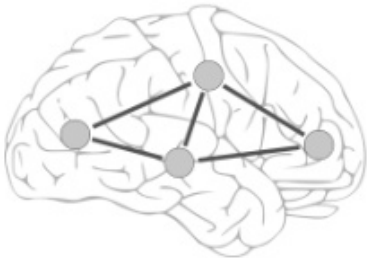
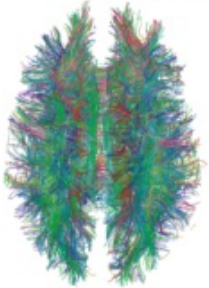
parcellation



fMRI/BOLD signals



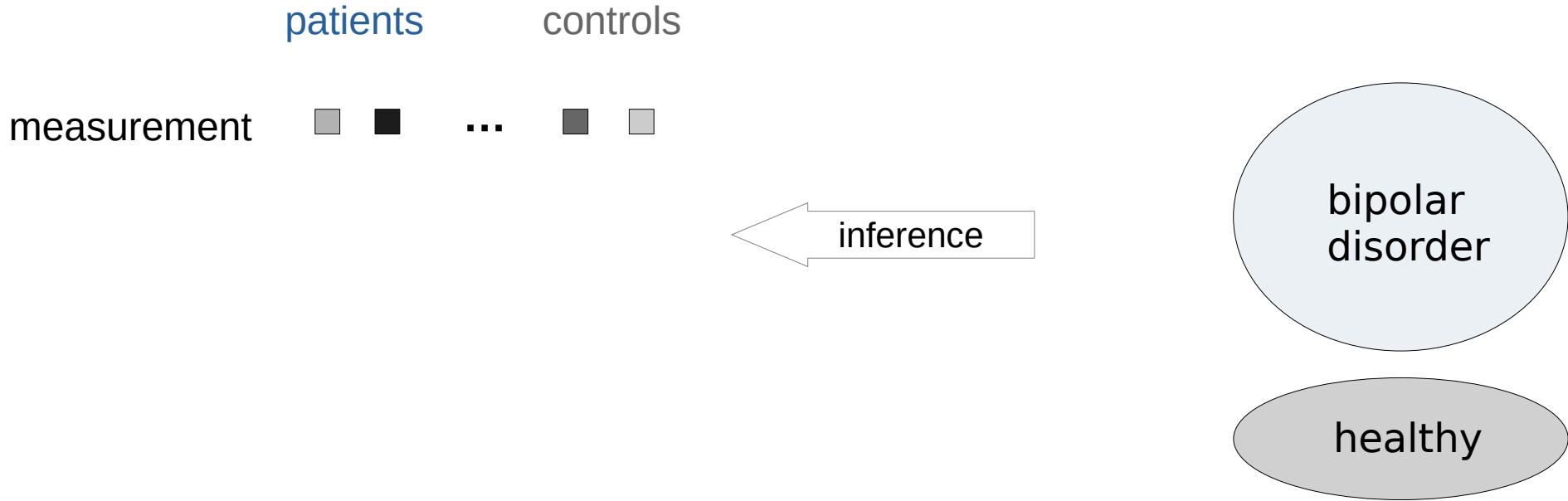
anatomical SC



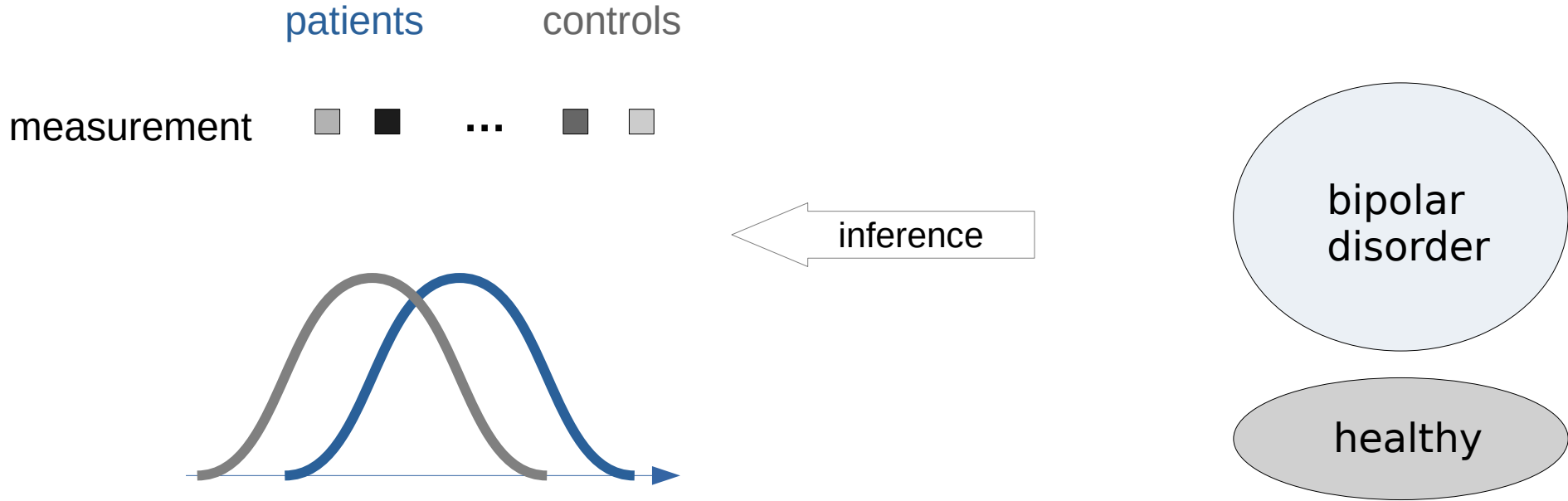
prediction

behavioral  
scores,  
pathological  
states

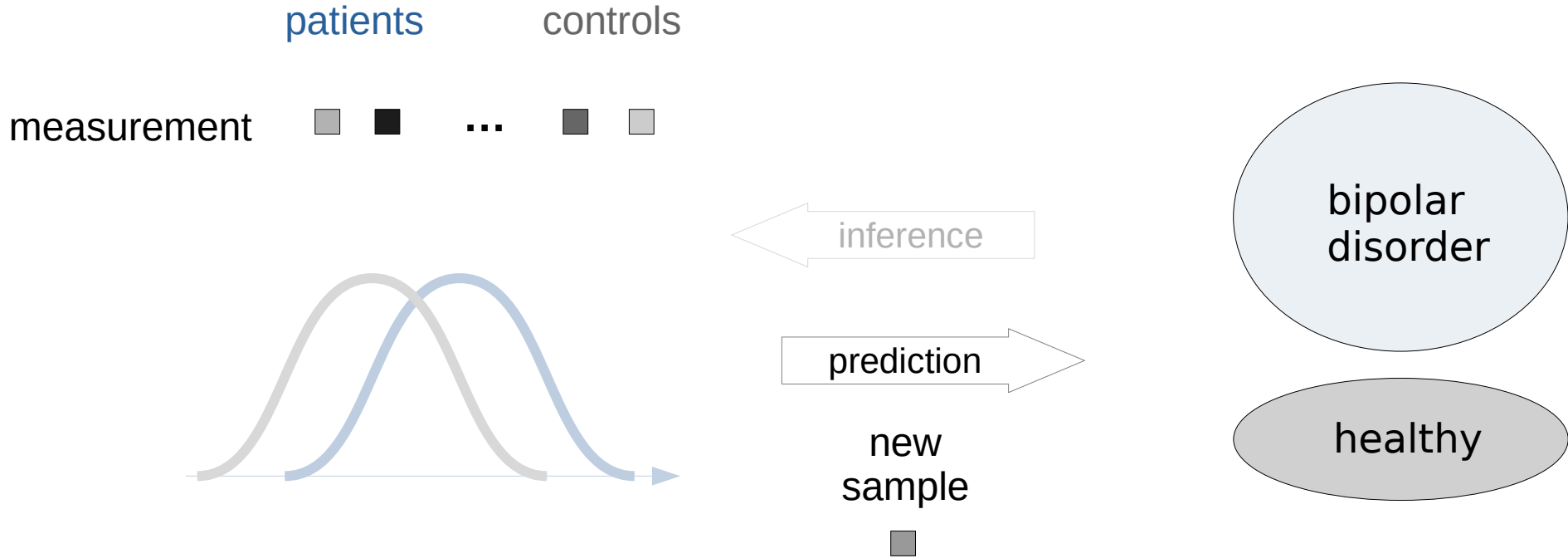
# Statistical Analysis versus Machine Learning



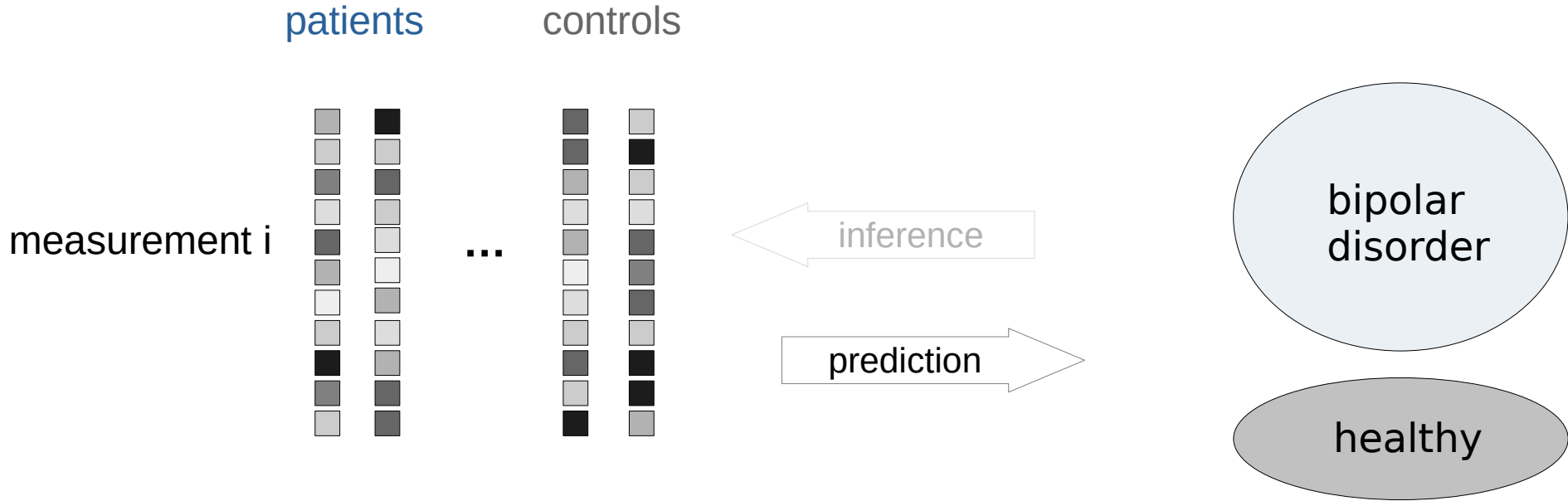
# Statistical Analysis versus Machine Learning



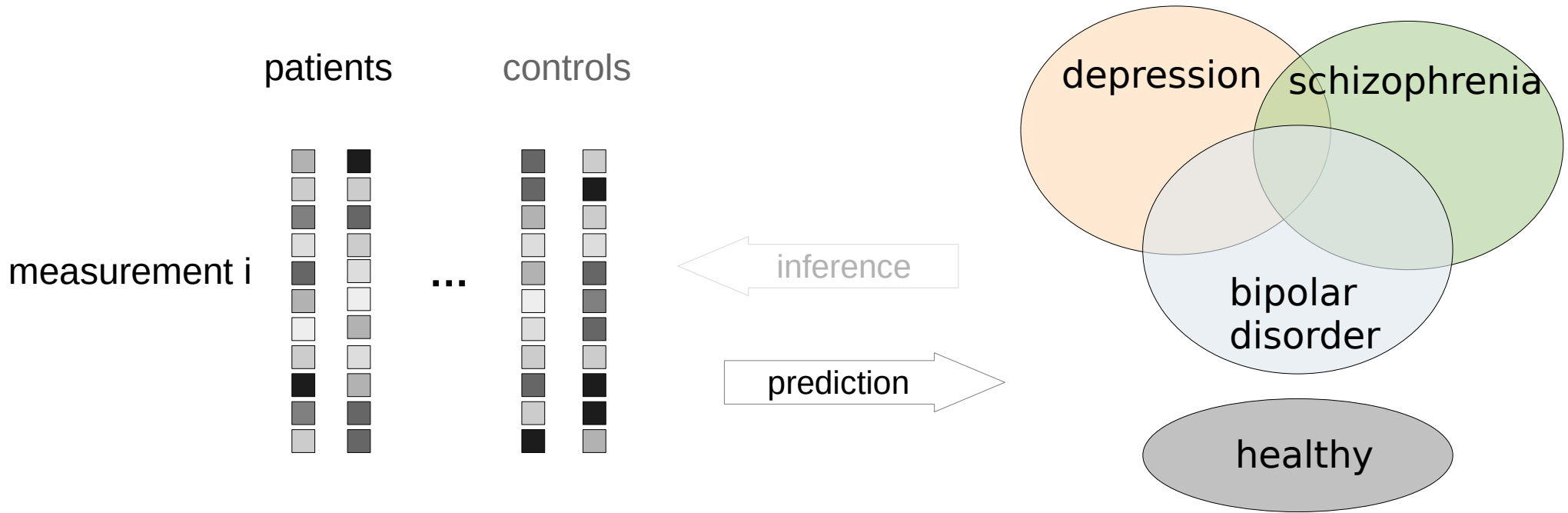
# Statistical Analysis versus Machine Learning



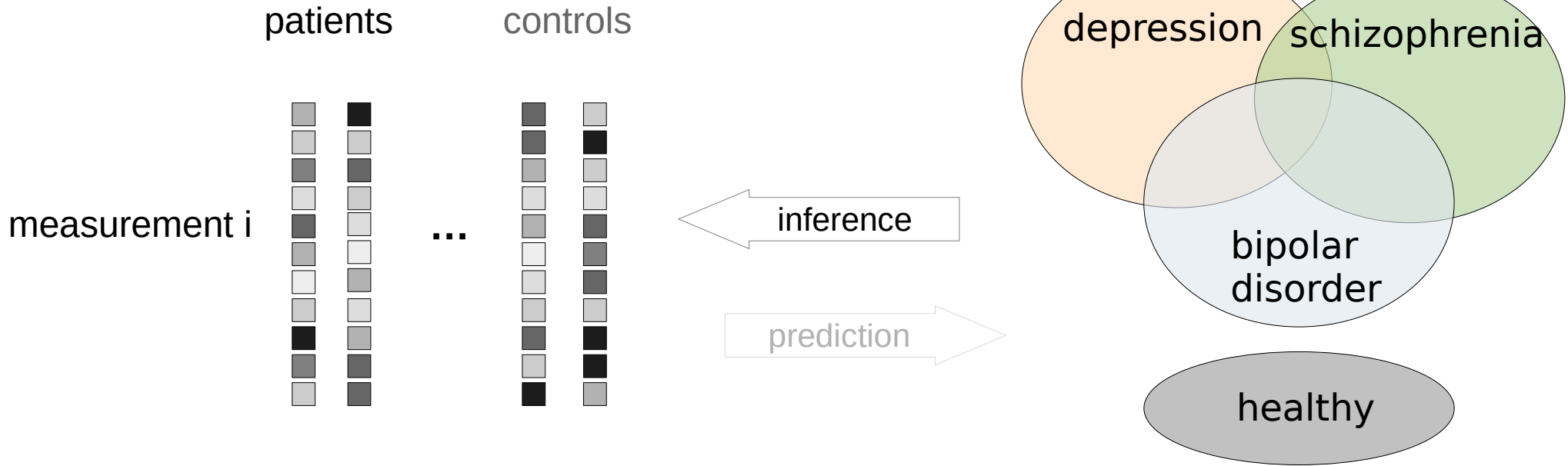
# Statistical Analysis versus Machine Learning



# Statistical Analysis versus Machine Learning



# Statistical Analysis versus Machine Learning



- multiple comparisons
- multiple classes

# When to choose Machine Learning?

- When data are in high dimension (e.g. neuroimaging)
- When there are many classes (or loosely defined, cf. semi-supervised learning)
- It controls for false positive (cross validation)

- But...





Journal of Multivariate Analysis  
Volume 190, July 2022, 104946



---

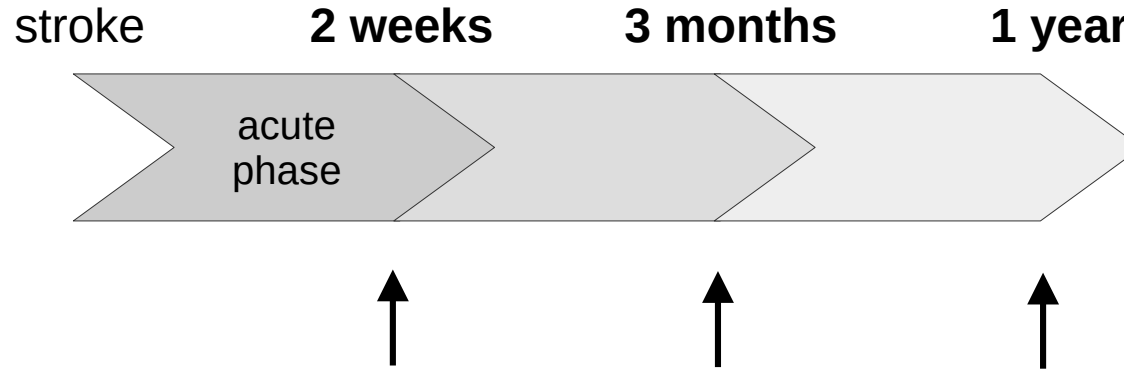
A generalized Wilcoxon–Mann–Whitney type  
test for multivariate data through pairwise  
distance

Jiamin Liu<sup>a, b</sup>, Shuangge Ma<sup>c</sup>, Wangli Xu<sup>a</sup>  , Liping Zhu<sup>d</sup>

# Part 1: Models in Computational Neuroscience

- Brief review of models in computational neuroscience
- **Example 1: classification for diagnosis / prognosis**
- Example 2: characterize structure in multivariate data
- Scikit-learn: formating data

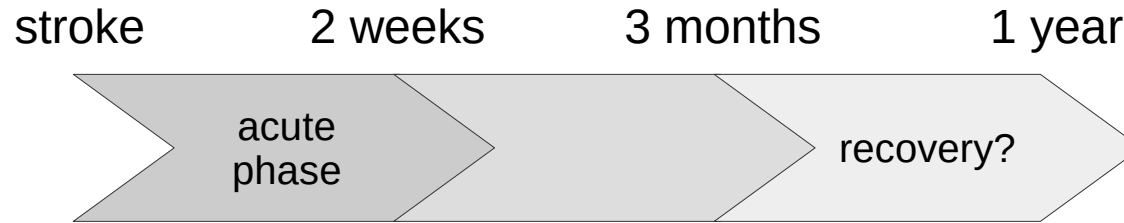
# Monitoring and Predicting Recovery for Stroke Patients



- Structural and functional MRI scans
- cognitive tests: memory, motor task, ...

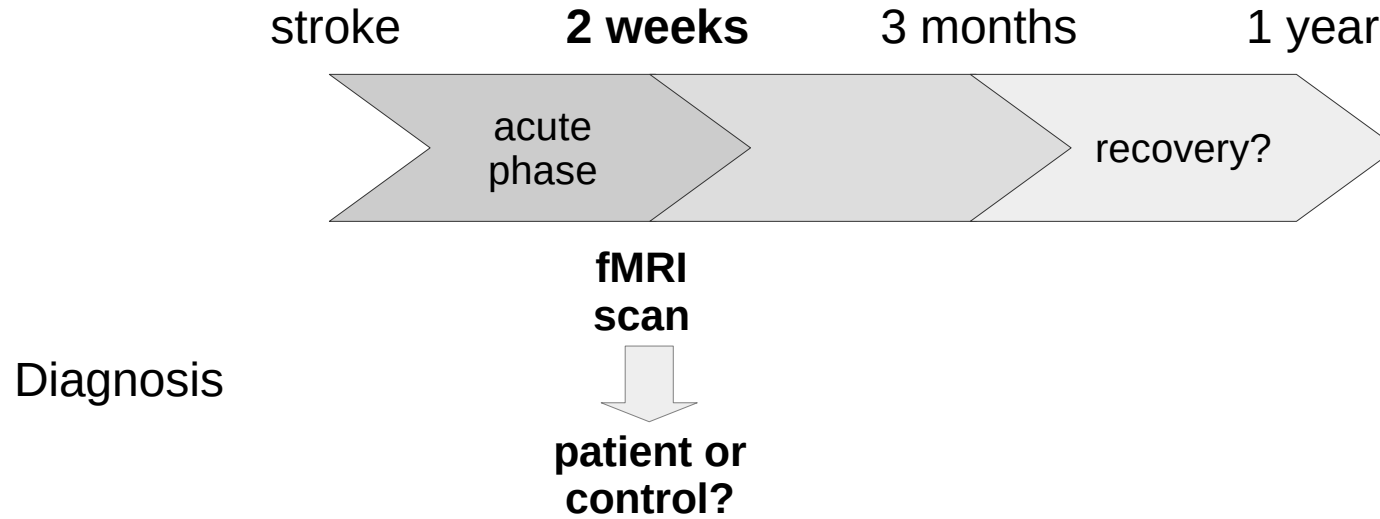


# Prediction for Stroke Patients

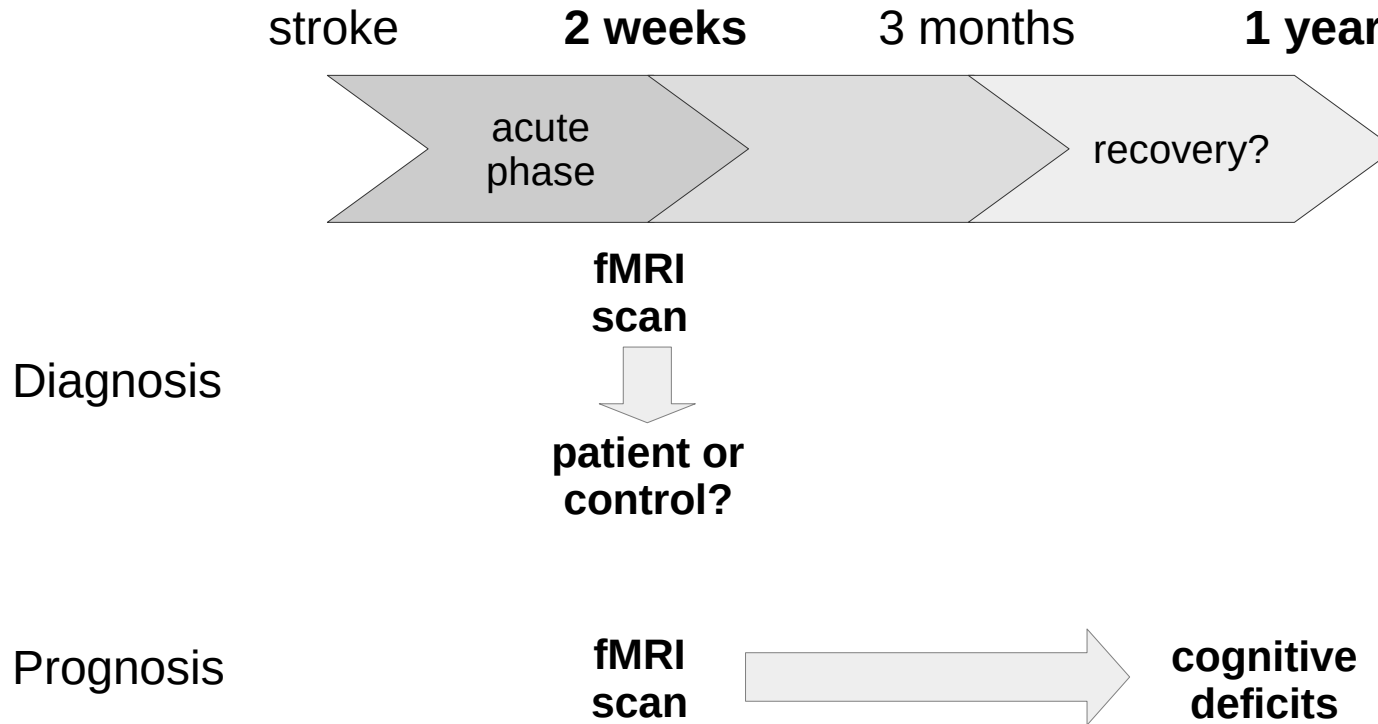


- 132 patients, 25 controls
- 300 brain regions (ROIs)
- 80%-20% train-test (stratified split)
- classifier: logistic regression

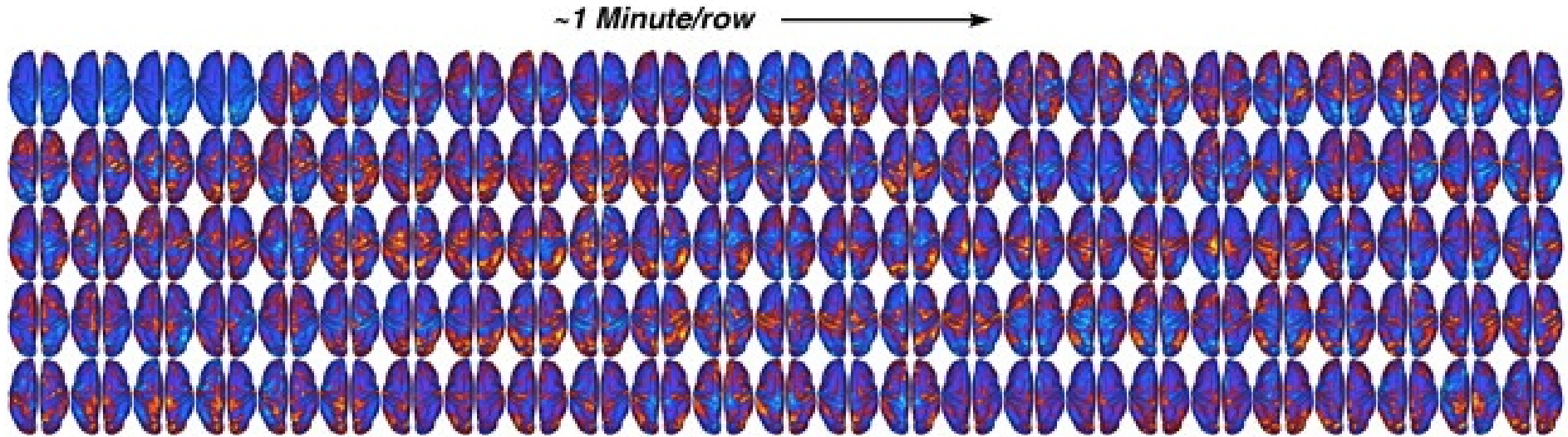
# Prediction for Stroke Patients



# Prediction for Stroke Patients



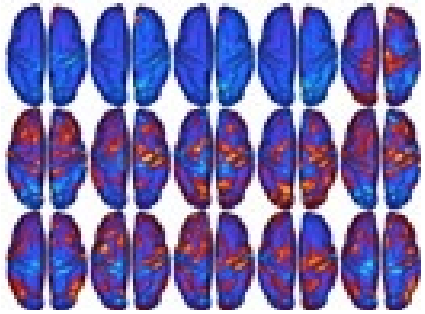
# Functional MRI



- blood-oxygen-level dependent (BOLD) signal
- indirect measure of neuronal activation
- spatio-temporal structure

# Classification Pipeline for Personalized Prognosis

BOLD images  
(voxel space)



prediction

Patient or  
healthy control?

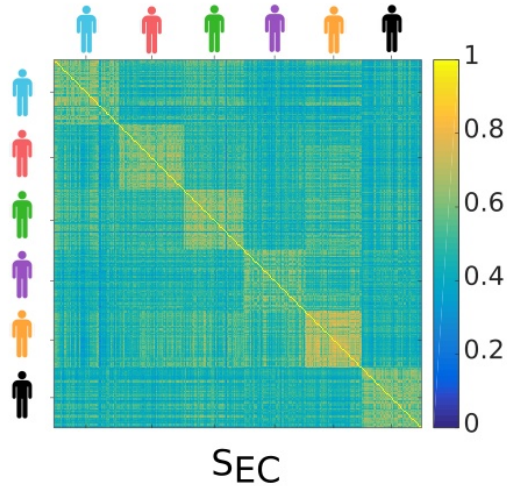




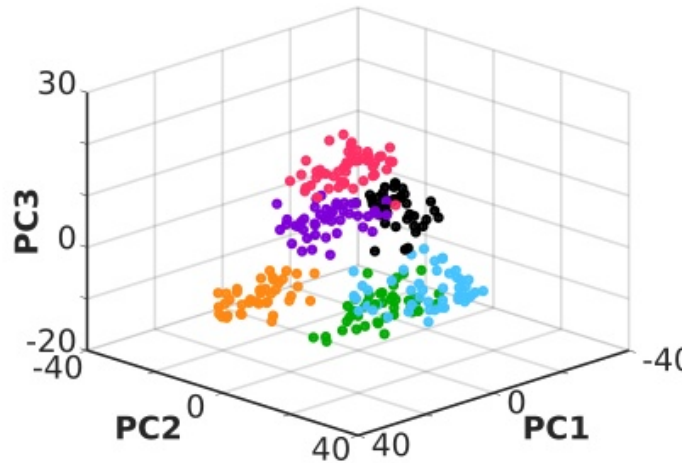


# Types of Variabilities in FC Data

Similarity across  
resting-state sessions



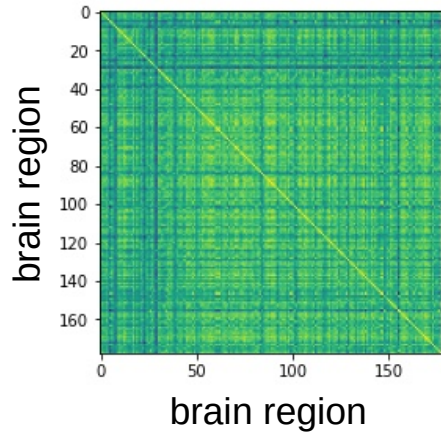
PCA representation  
(1 dot = 1 session)



- Strong bias related to subject identity
- Signature for conditions needs to generalize across subjects
- Ignore session-to-session variability

# Predictive Statistics Model: Classifier

FC matrix



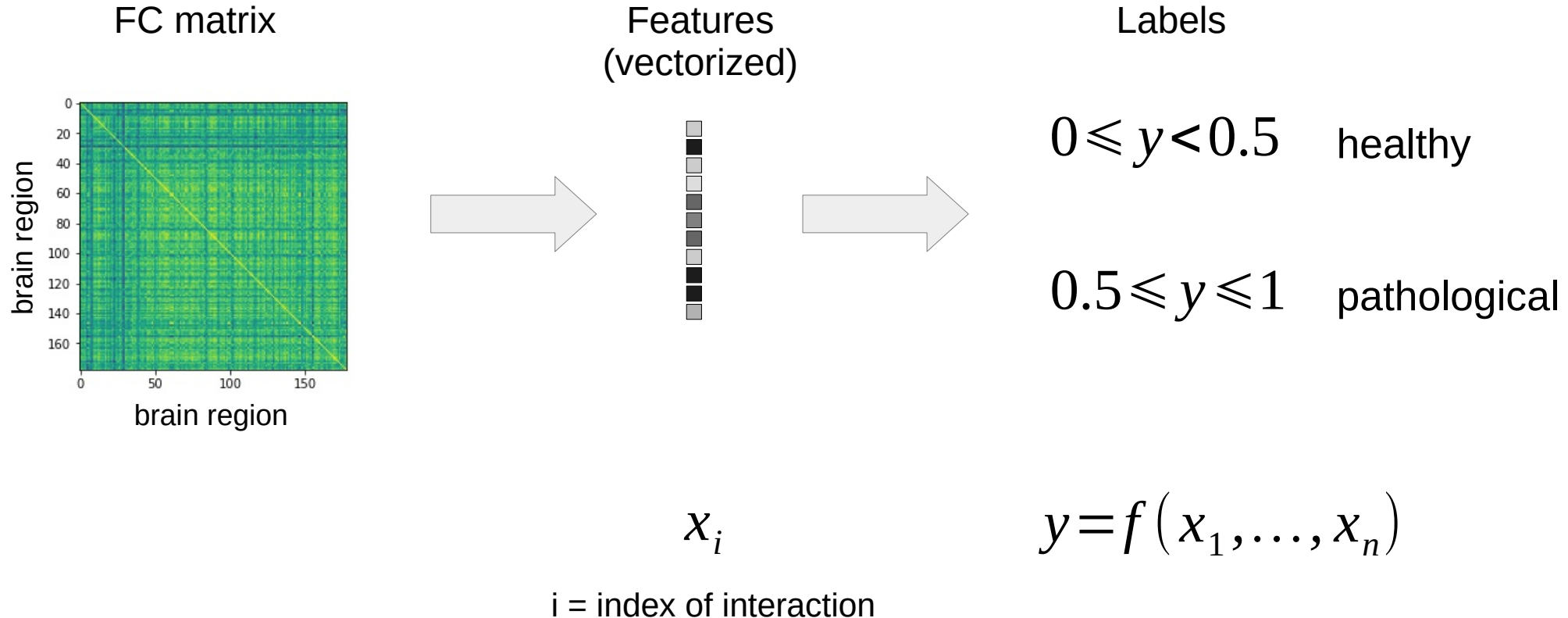
Features  
(vectorized)



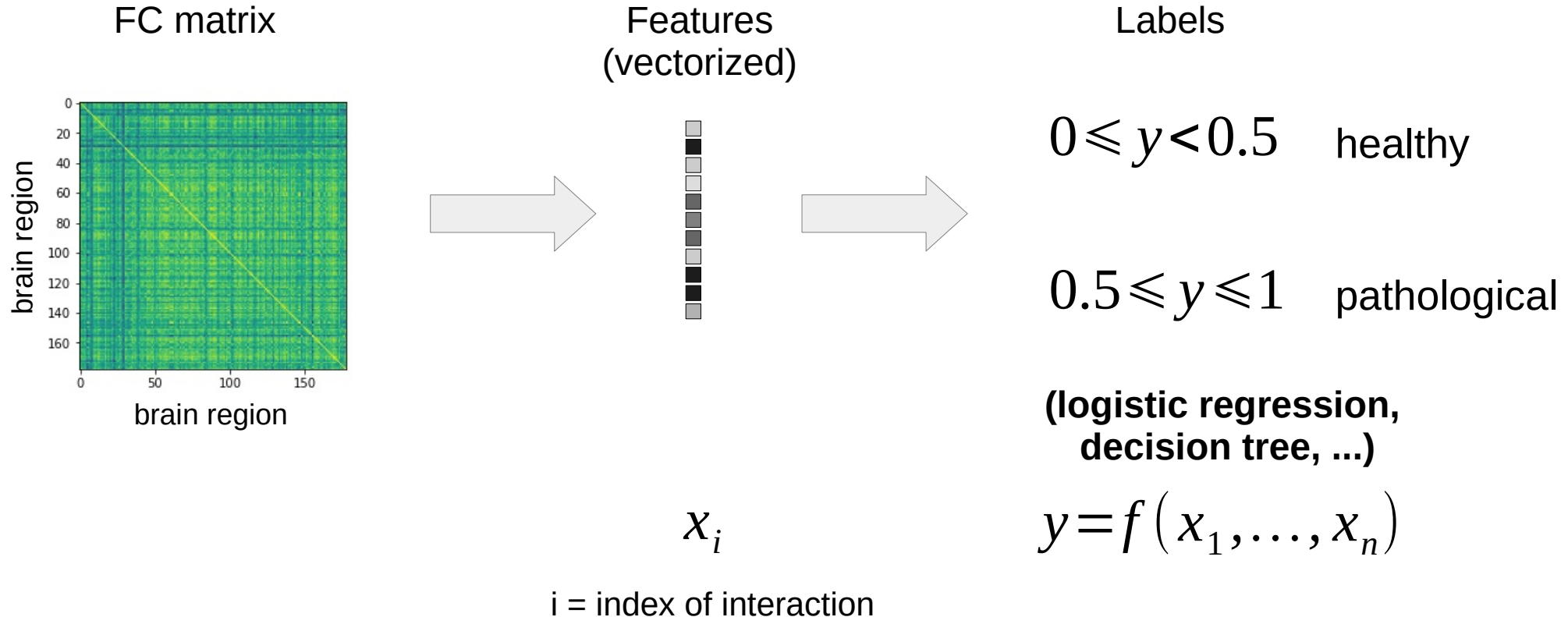
$$X_i$$

i = index of interaction

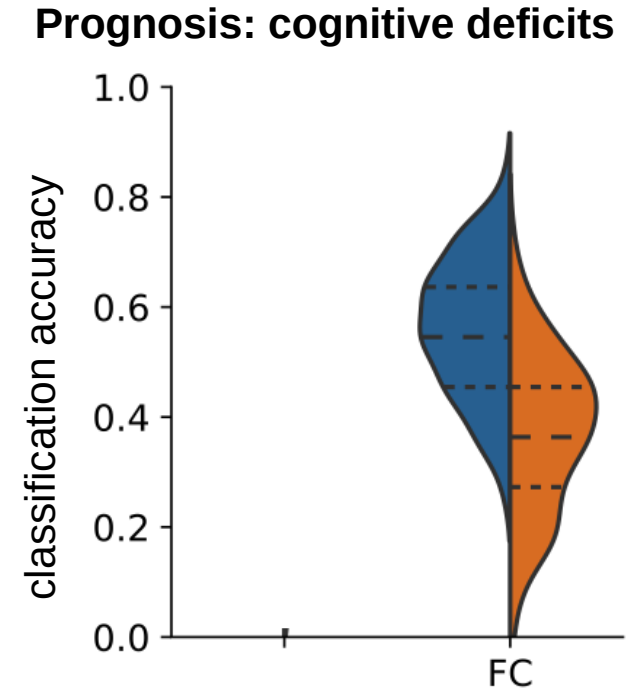
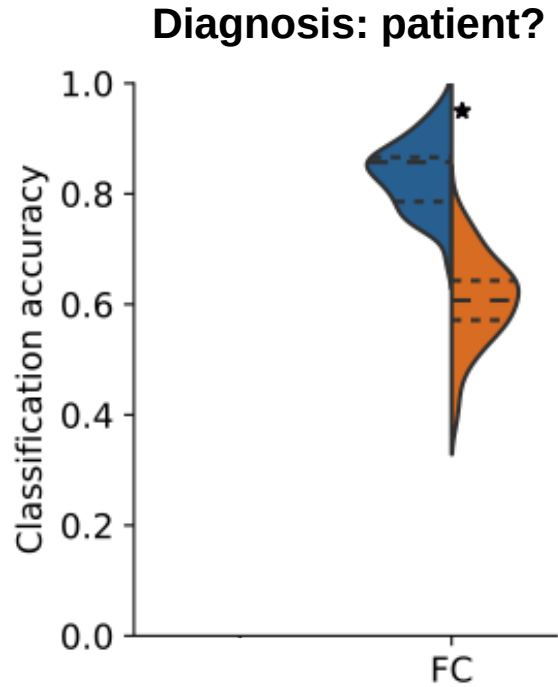
# Predictive Statistics Model: Classifier



# Predictive Statistics Model: Classifier

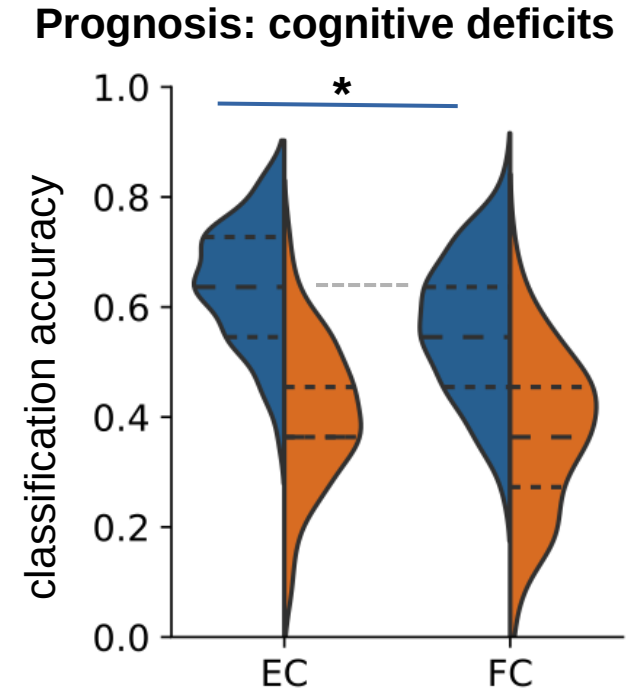
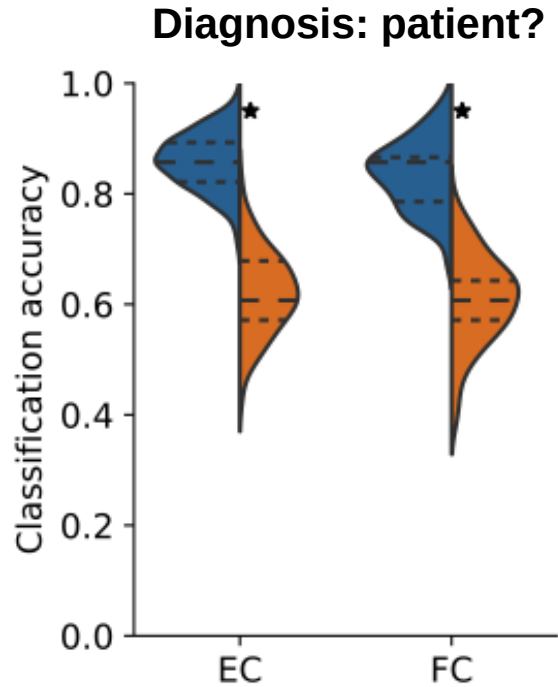


# Prediction from FC



real data  
chance level  
(shuffling surrogates)

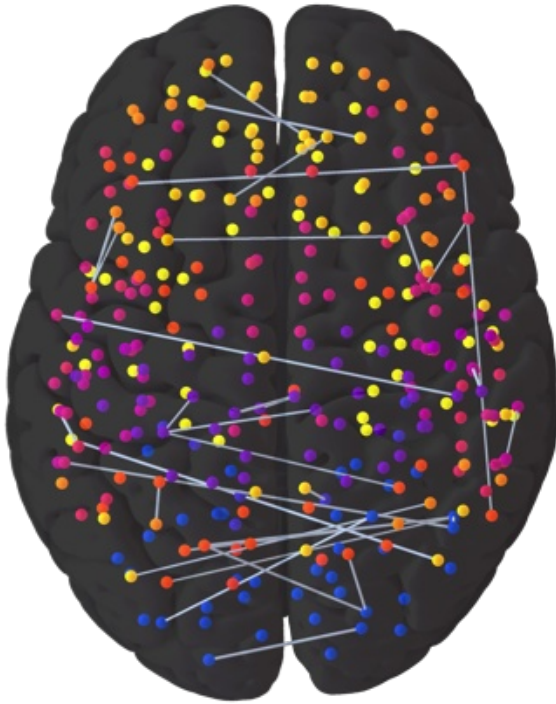
# Model-based EC can Outperform FC



real data  
chance level  
(shuffling surrogates)

# Interpretation as Cortico-Cortical Reconfiguration

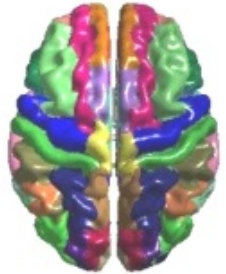
Informative EC links



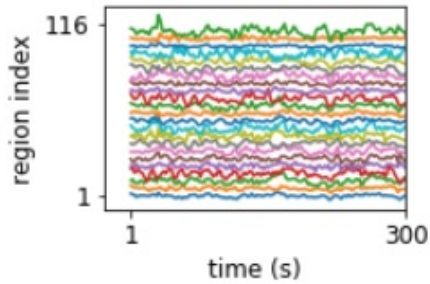
- What supports the recovery of cognitive deficits?
- Not only close to lesions
- Inter-hemispheric connections

# Model-Based Analysis of Neuroimaging Data

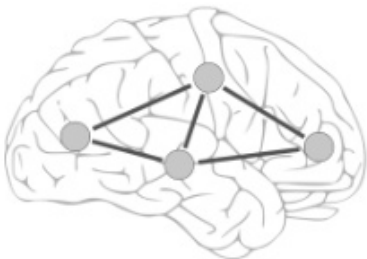
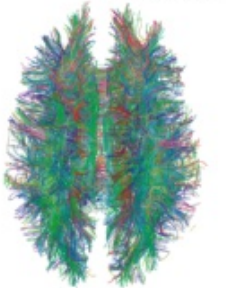
parcellation



fMRI/BOLD signals



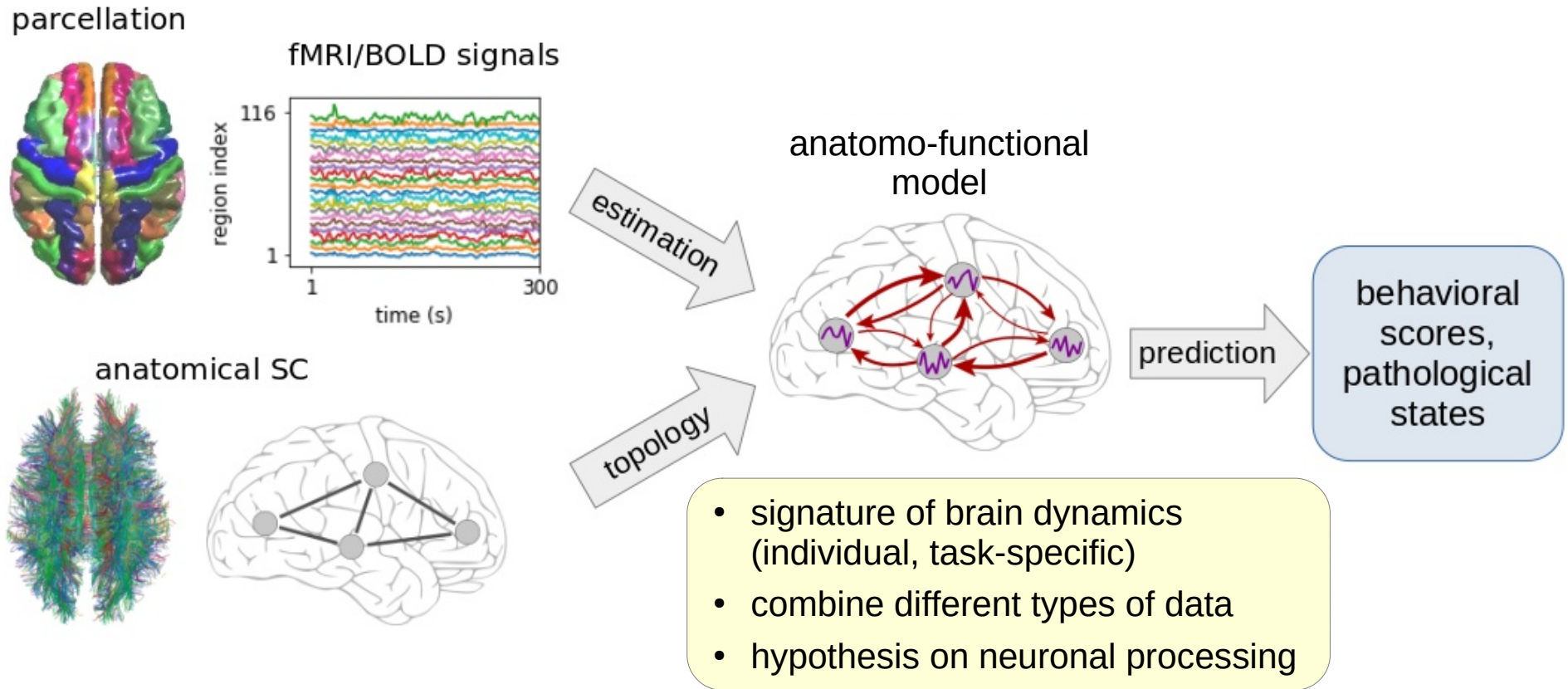
anatomical SC



prediction

behavioral  
scores,  
pathological  
states

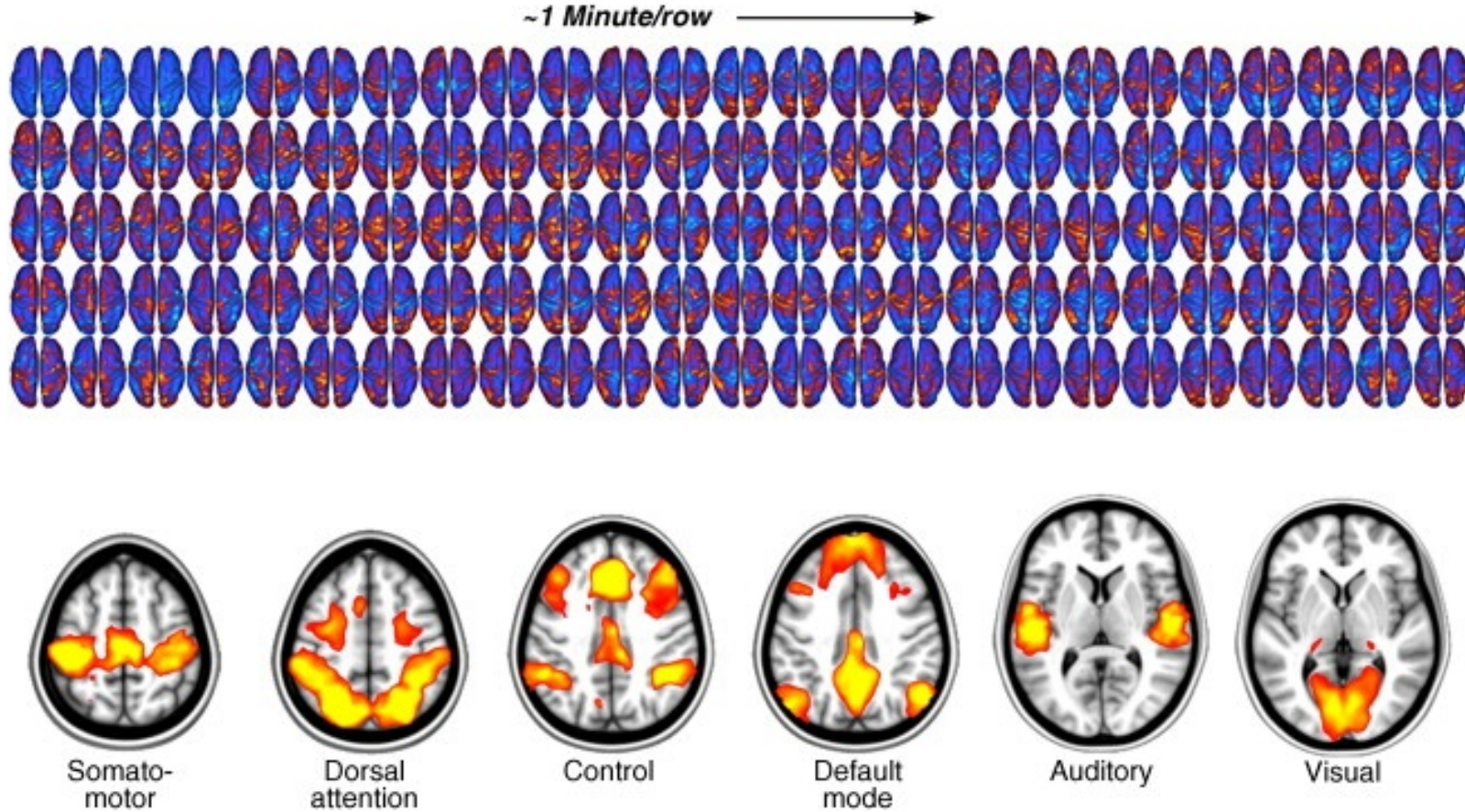
# Model-Based Analysis of Neuroimaging Data



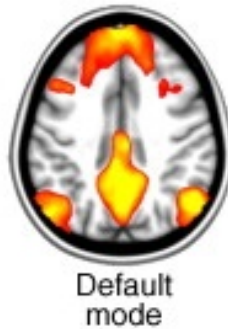
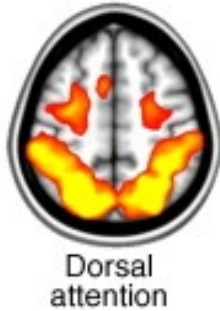
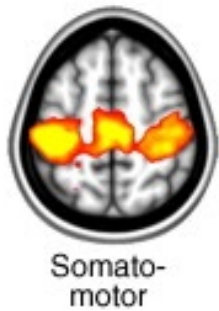
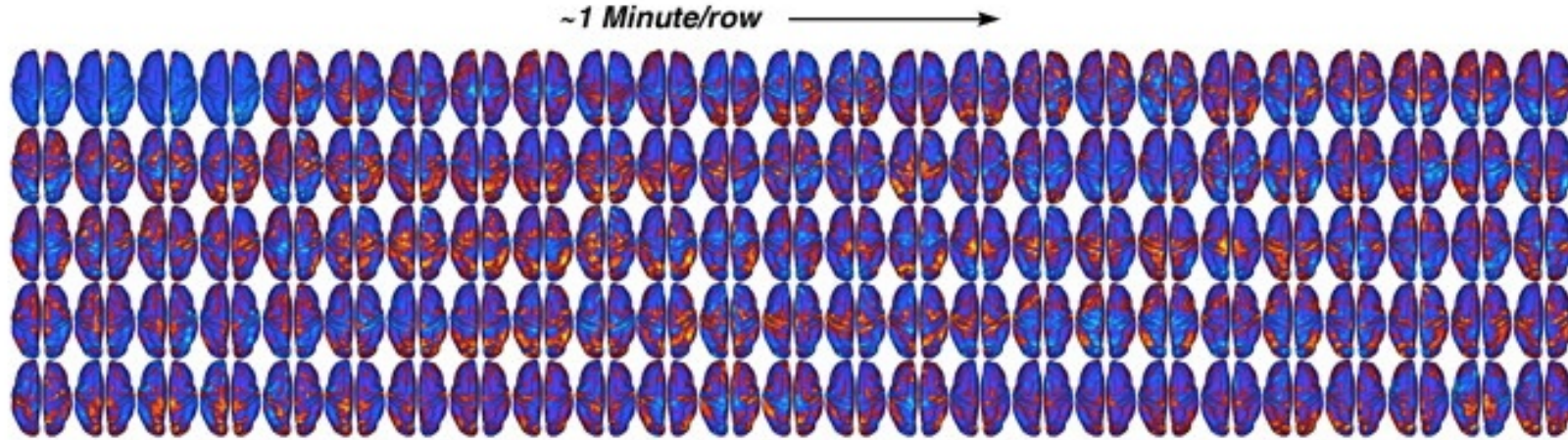
# Part 1: Models in Computational Neuroscience

- Brief review of models in computational neuroscience
- Example 1: classification for diagnosis / prognosis
- **Example 2: characterize structure in multivariate data**
- Scikit-learn: formating data

# Resting-State Networks



# Resting-State Networks

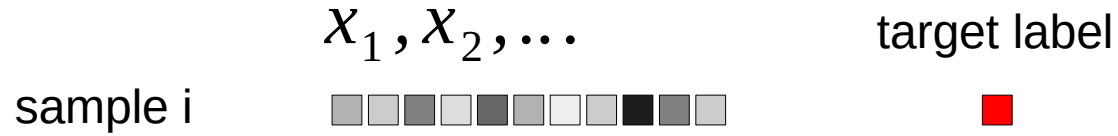


**independent  
component  
analysis**

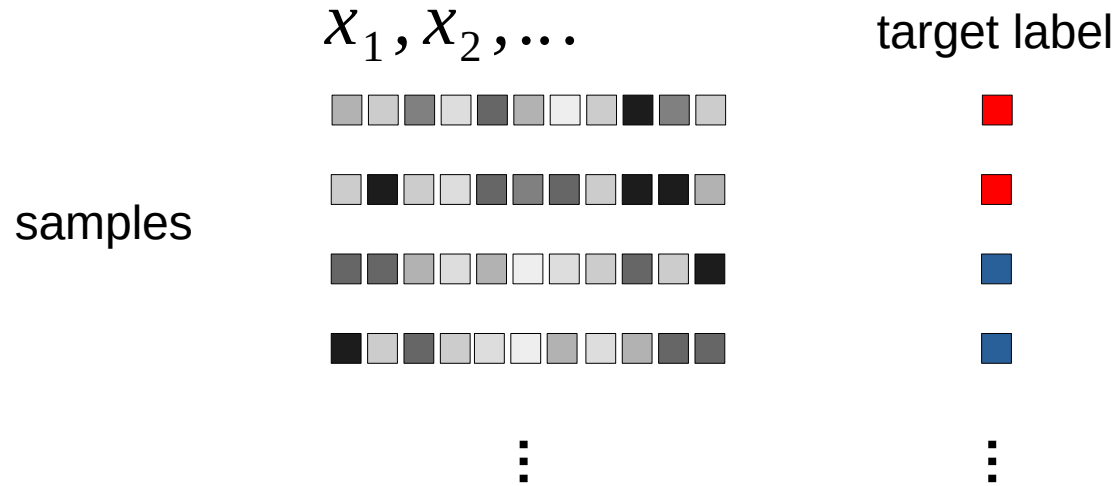
# Part 1: Models in Computational Neuroscience

- Brief review of models in computational neuroscience
- Example 1: characterize structure in multivariate data
- Example 2: classification for diagnosis / prognosis
- **Scikit-learn: formating data**

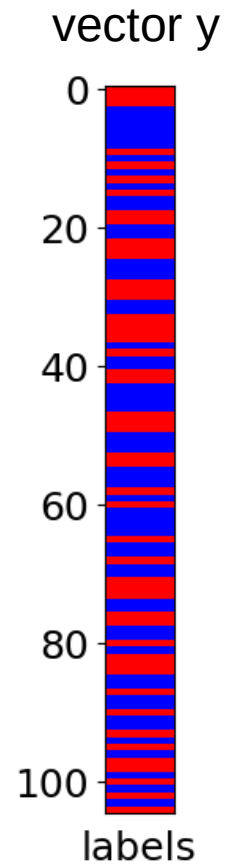
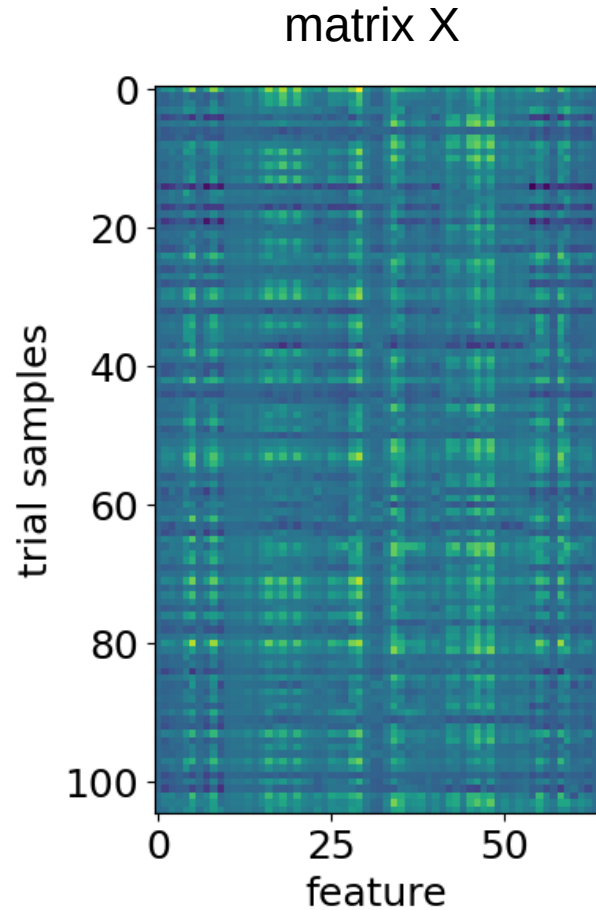
# Organization of Data for Classification



# Organization of Data for Classification



# Organization of Data for Classification



Classifier model:

$$y = f(X)$$

# Supervised learning, unsupervised learning, etc.

Classification:  $f(X) = y$  labels

Regression:  $f(X) = y$  explained variables

Clustering:  $f(X) = y$  labels

Dimensionality reduction:  $f(X) = y$  new coordinates

# Supervised learning, unsupervised learning, etc.

Classification:	$f(X) = y$	labels
Regression:	$f(X) = y$	explained variables
Clustering:	$f(X) = y$	labels
Dimensionality reduction:	$f(X) = y$	new coordinates

Optimize  $f$   
for  $(X, y)$

# Supervised learning, unsupervised learning, etc.

<b>fit</b>	train the using data (X,y), sets the transformation (PCA with a number of components), etc.
<b>transform</b>	maps the X to y (prediction for classifier, new coordinates for PCA, etc.)
<b>fit_transform</b>	

<https://scikit-learn.org>



- simple and efficient tools for predictive data analysis
- accessible to everybody, and reusable in various contexts
- built on NumPy, SciPy, and matplotlib
- open source, commercially usable - BSD license

<b>classification</b> MLR, SVM	<b>regression</b>	<b>clustering</b> k-means
<b>dim reduction</b> PCA, ICA	<b>model selection</b> CV, grid search	<b>preprocessing</b>

# Practice

- Anaconda installation
- Environment yml file: required libraries
- Use jupyter-notebook (or jupyter-lab)
- TO DO: load\_fmri\_data (in data\_fmri)