

Algorithms for Data Analysis

Statistics – Machine Learning



Matthieu Gilson
Chaire de Professeur Junior

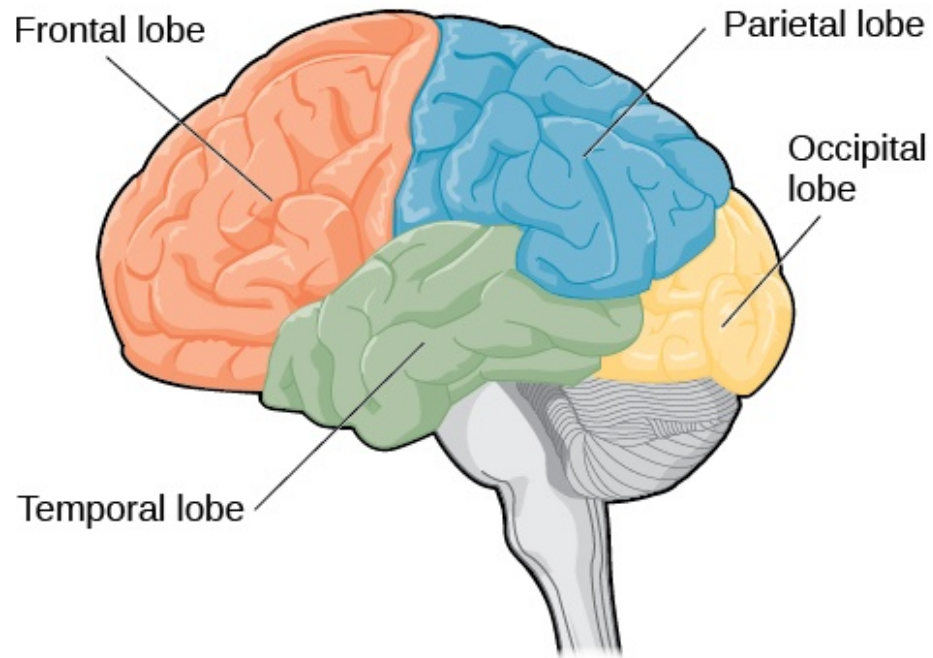


Formatting Data in Scikit-Learn

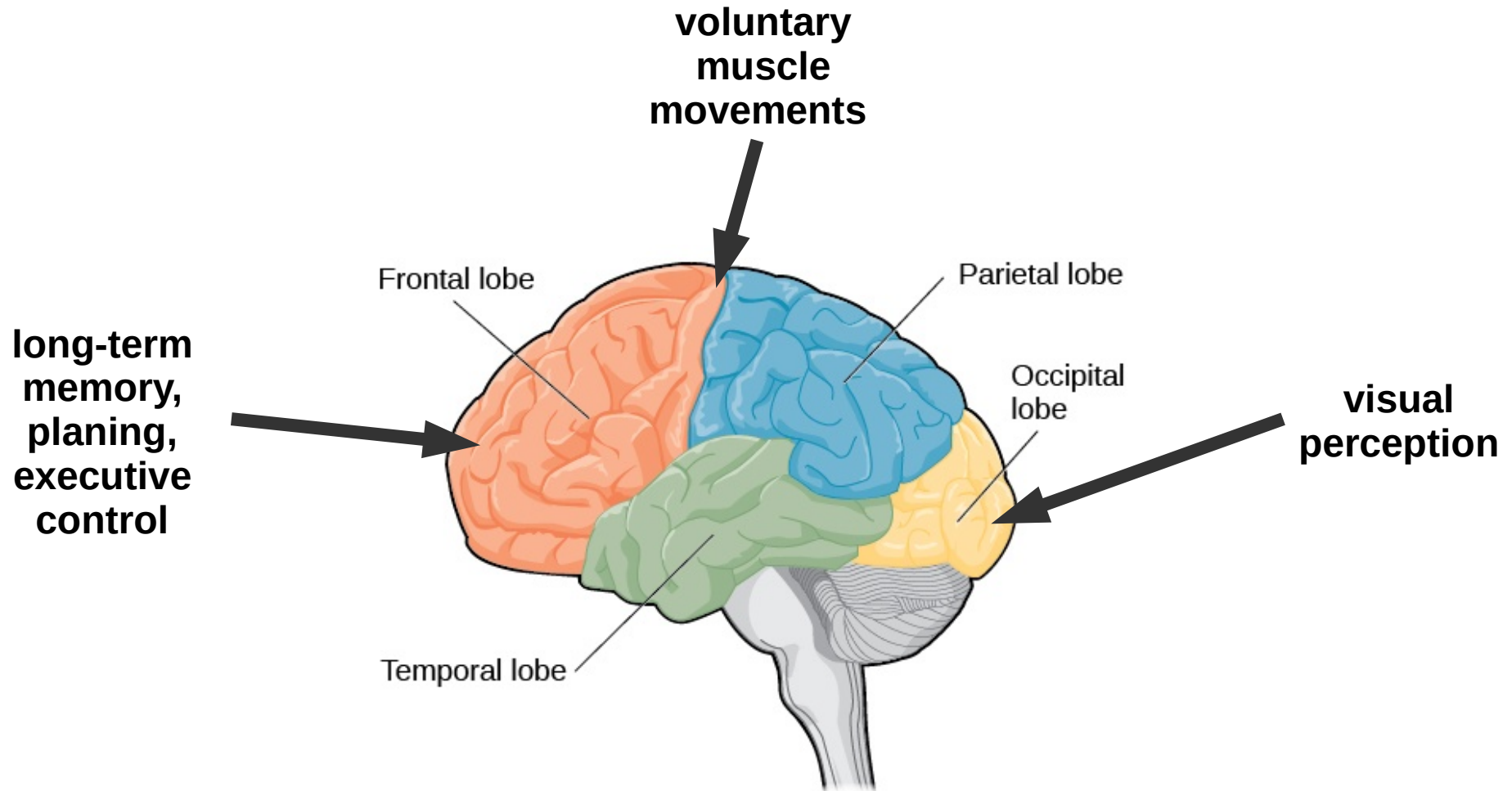
Introduction to Models in Neuroscience

- **The brain as a distributed and complex network system**
- Neuroimaging: quantifying the brain
- Statistical analysis versus classification
- Example 1: diagnosis / prognosis in stroke
- Example 2: whole-brain modeling, The Virtual Brain
- Example 3: characterize structure in multivariate data
- Scikit-learn: formatting data

Localizing Behavioral and Cognitive Functions



Localizing Behavioral and Cognitive Functions



Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response

**Context 1:
same color?**



Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response

**Context 1:
same color?**



oui

Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response

**Context 1:
same color?**



oui

non

Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response

**Context 1:
same color?**



oui



non

**Context 2:
same shape?**



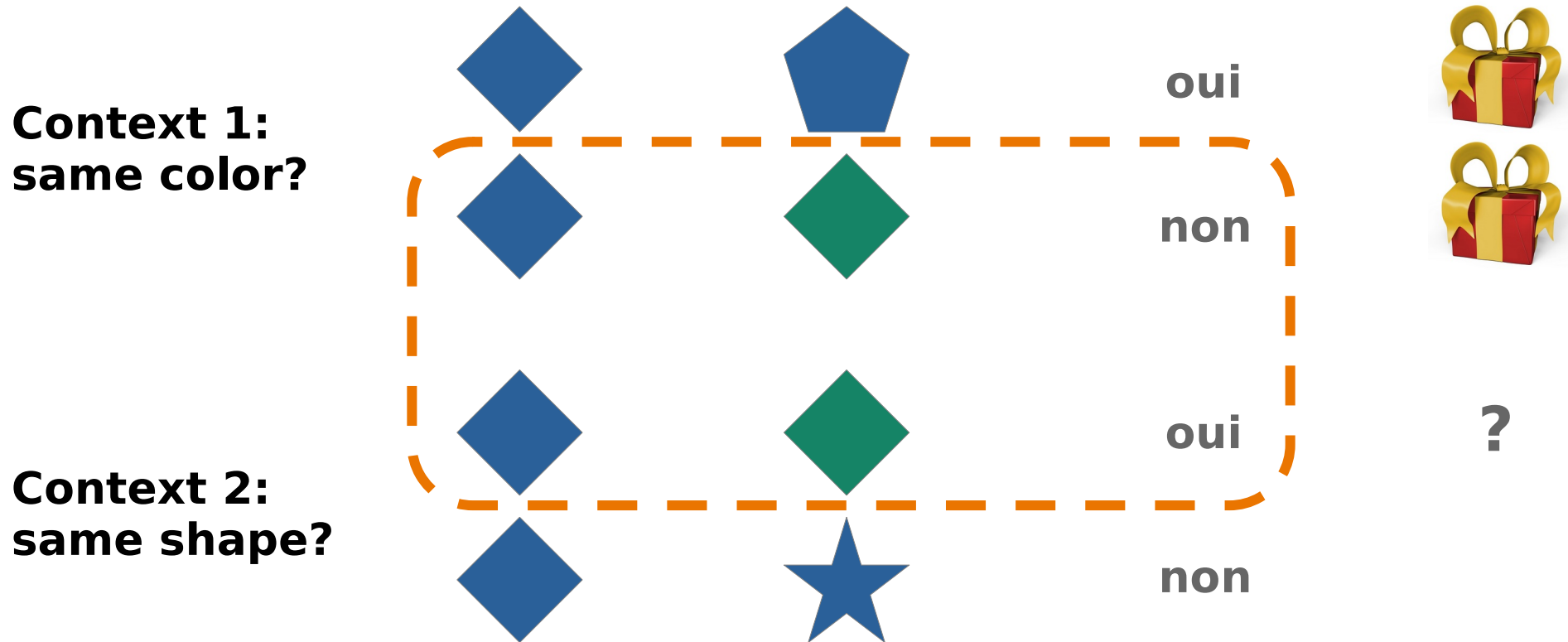
oui



non

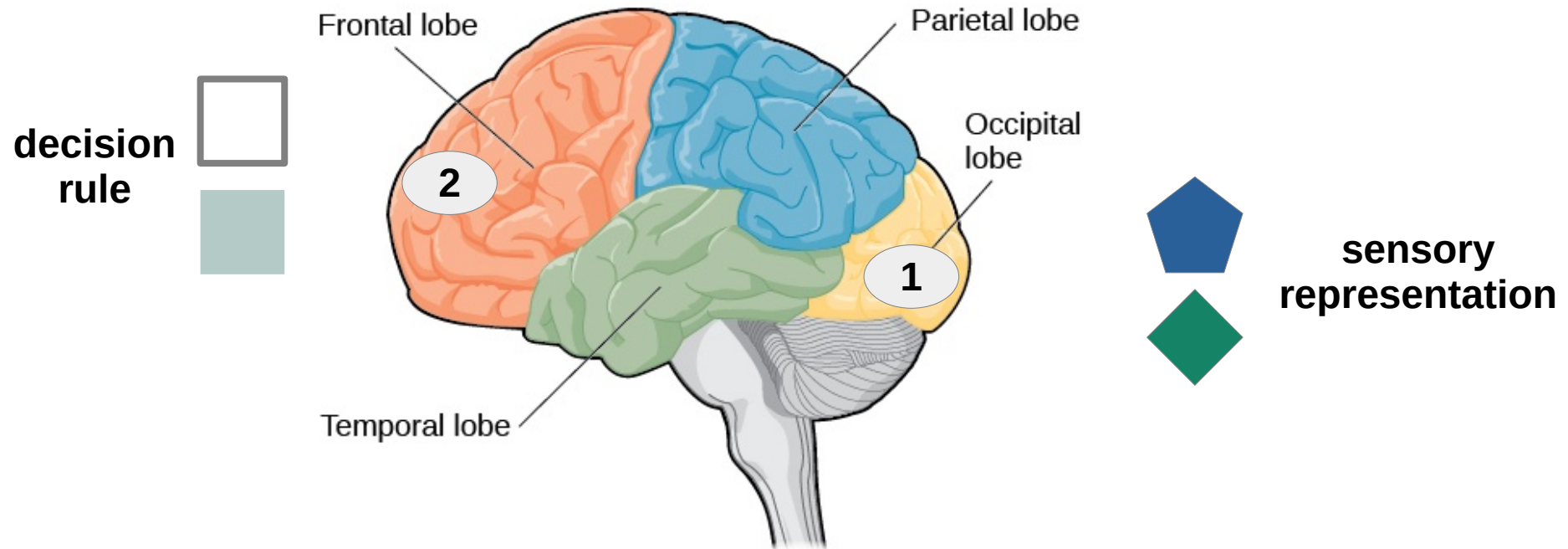
Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response



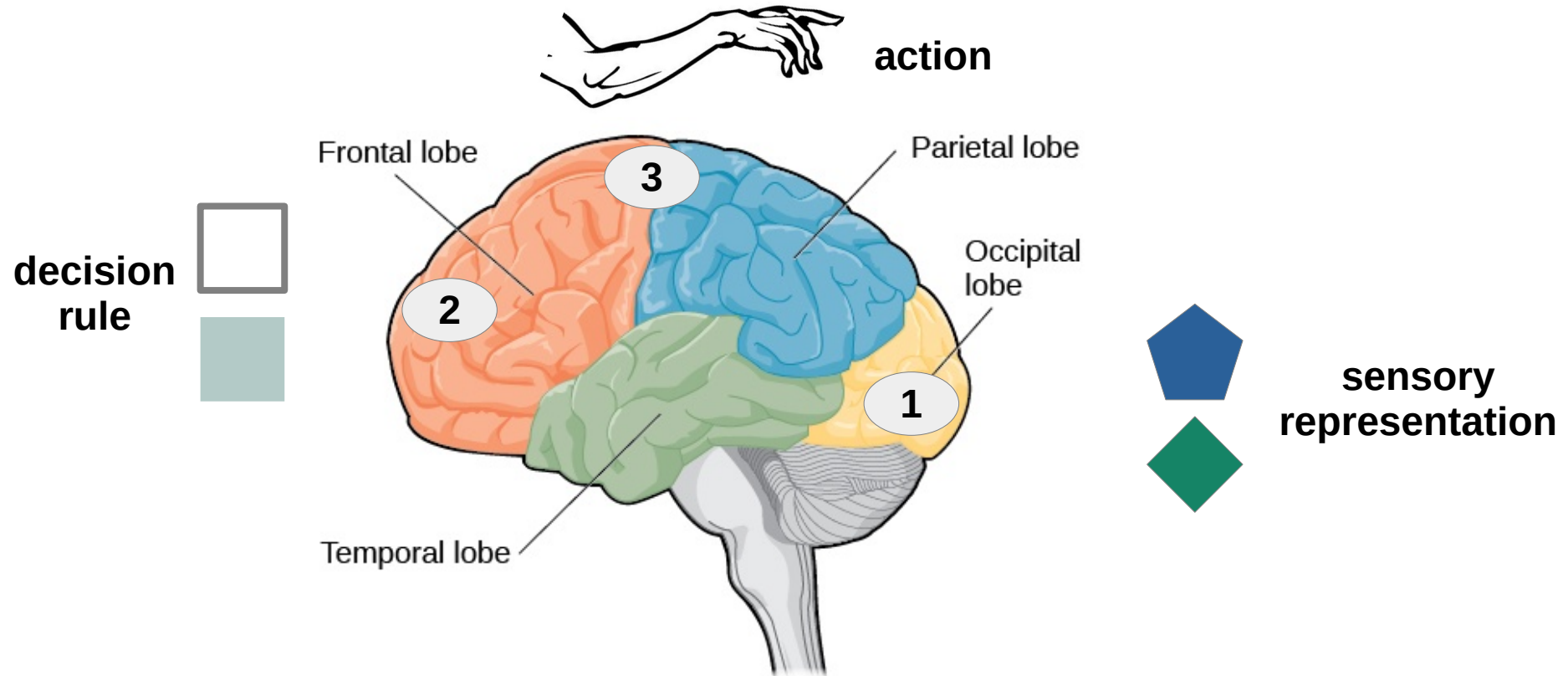
Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response
- Activation of brain regions and interactions (functional hierarchy)



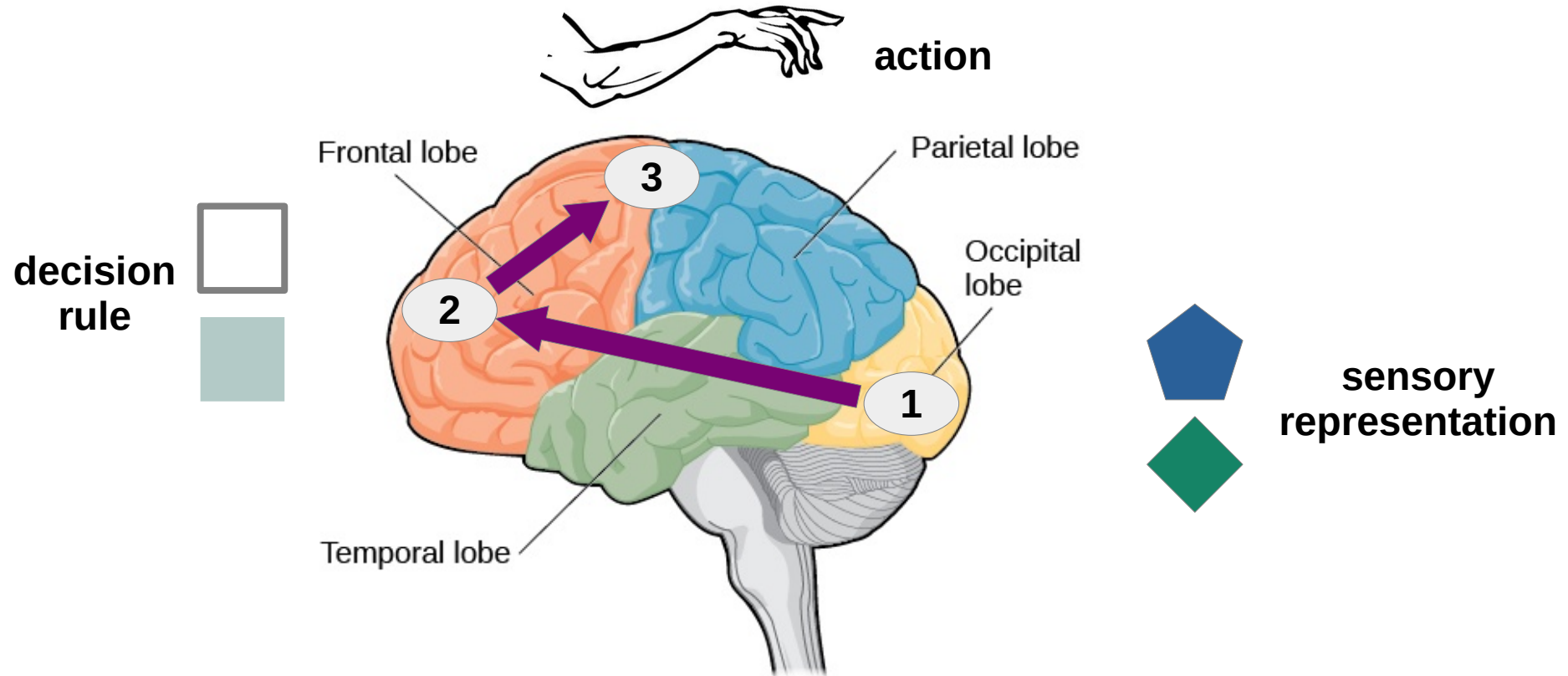
Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response
- Activation of brain regions and interactions (functional hierarchy)



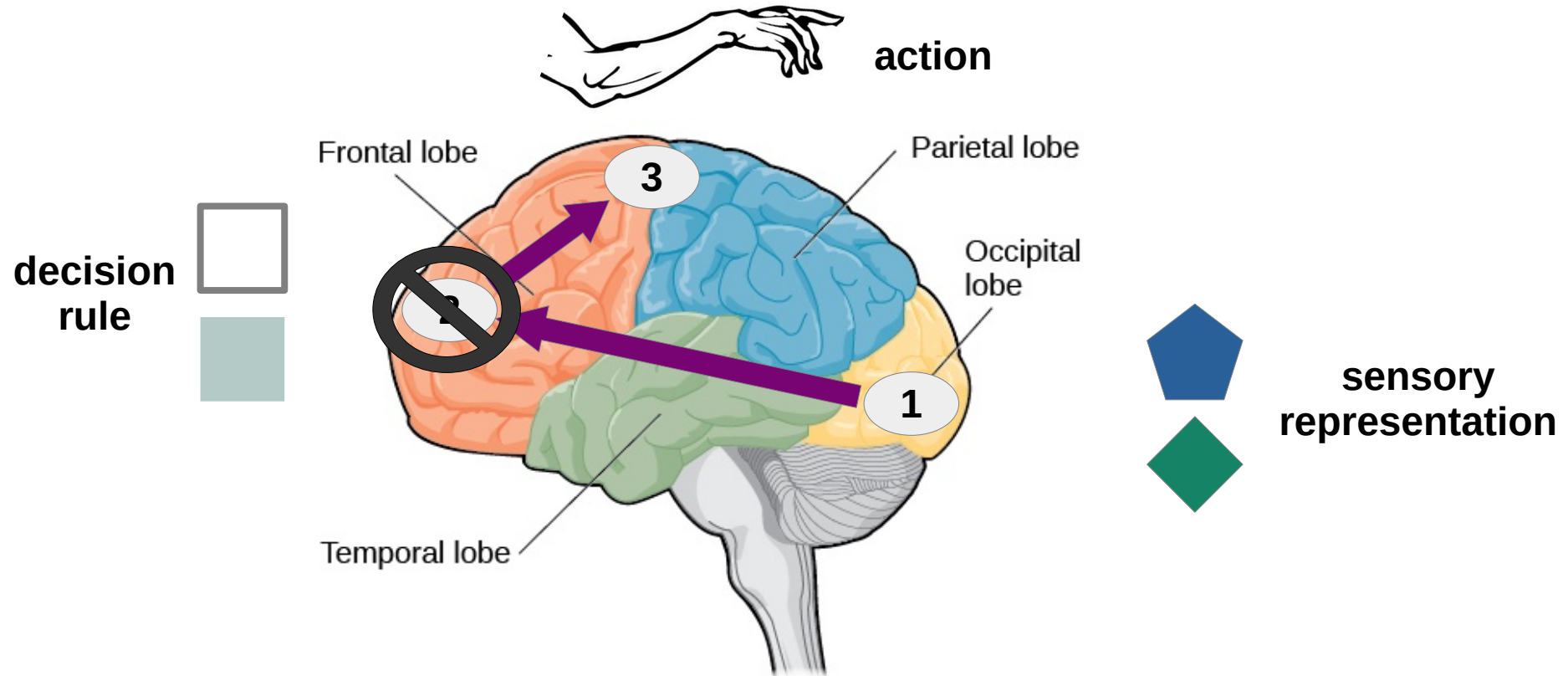
Cognitive Neuroscience: Neuronal Implementation of Functions

- Example: matching of visual objects depending on color or shape, motor response
- Activation of brain regions and interactions (functional hierarchy)



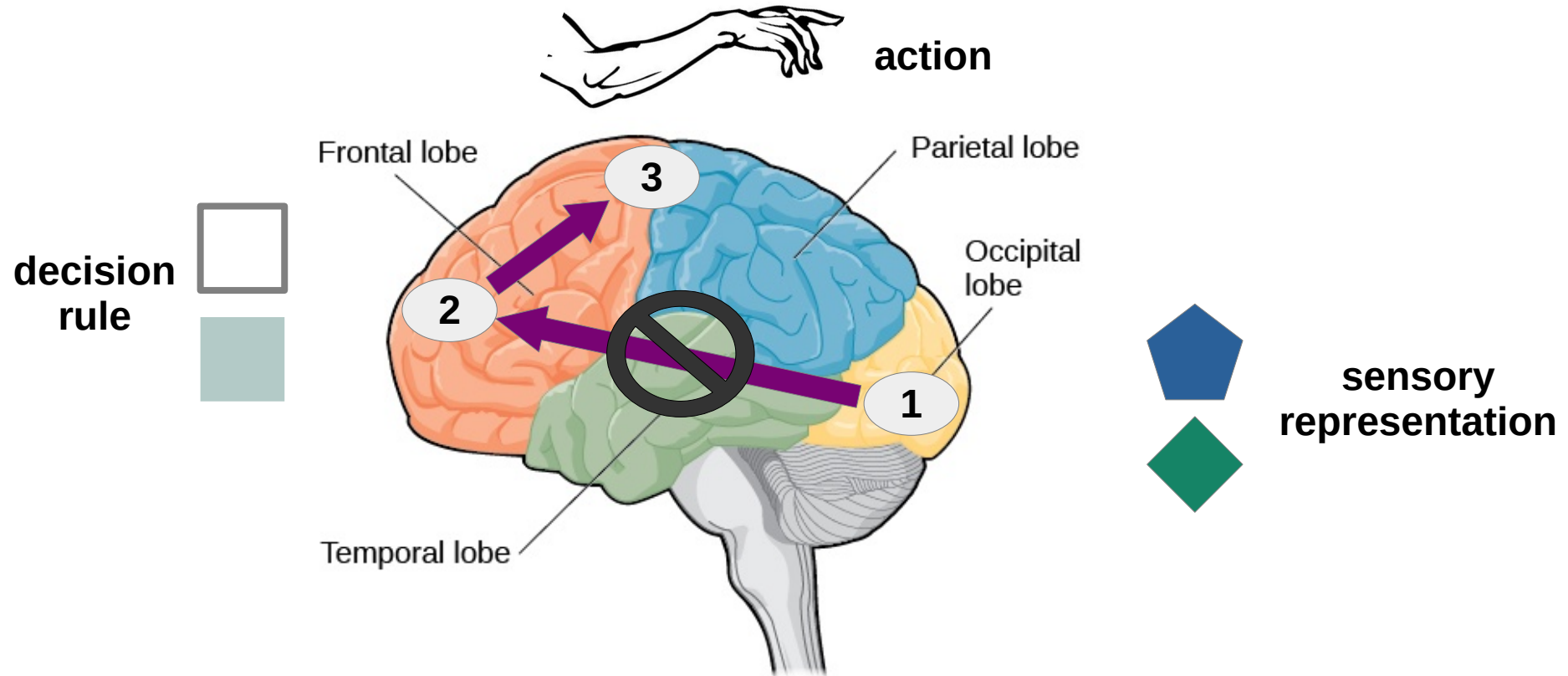
Characterization of Pathological Alterations

- Example: matching of visual objects depending on color or shape, motor response
- Activation of brain regions and interactions (functional hierarchy)



Characterization of Pathological Alterations

- Example: matching of visual objects depending on color or shape, motor response
- Activation of brain regions and interactions (functional hierarchy)

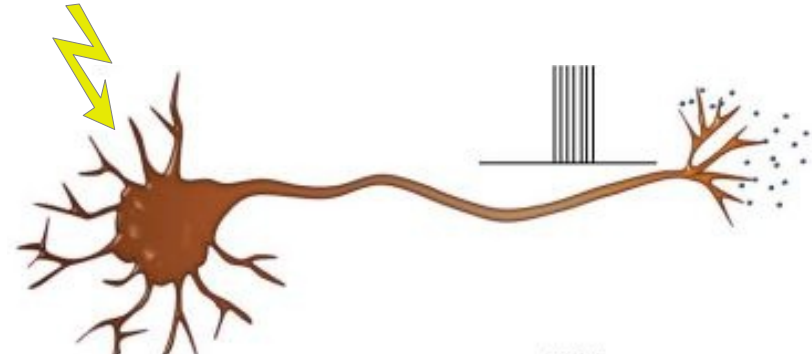


Multiscale Organization of the Brain

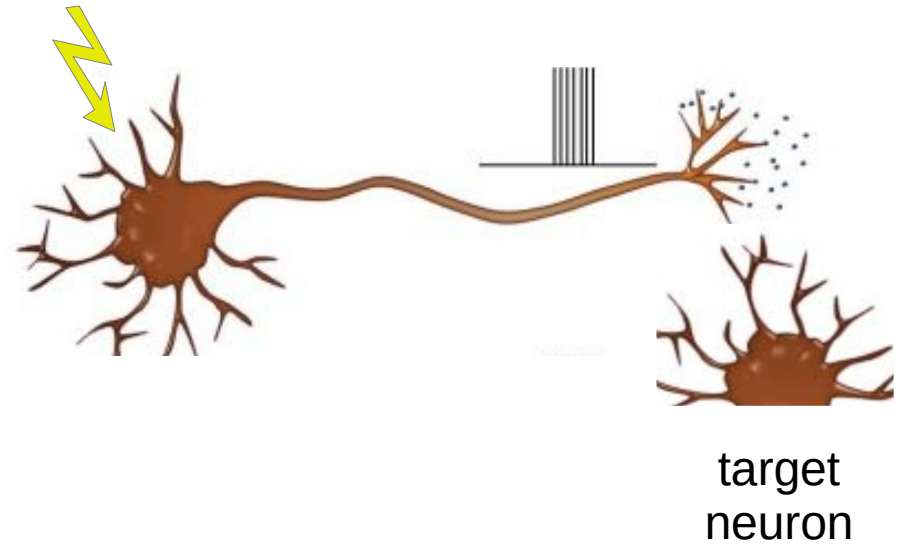


- 10^{11} neurons
 - 20% in cortex, 80% in cerebellum
- about 10000 synapses by neurons

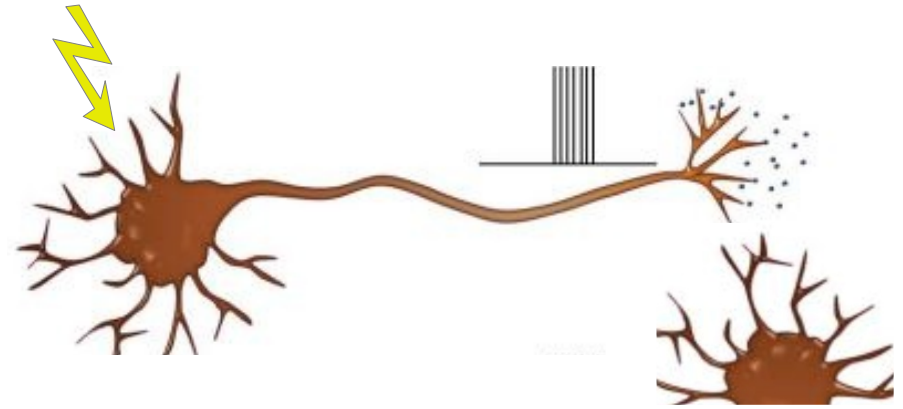
Multiscale Organization of the Brain



Multiscale Organization of the Brain

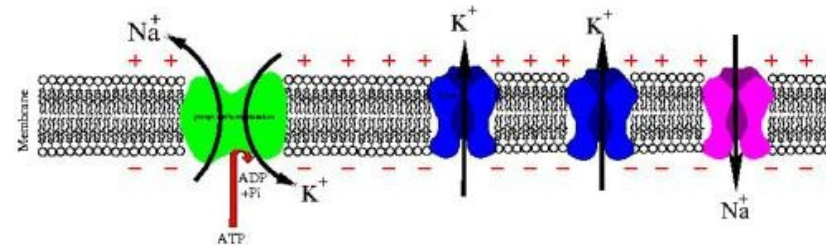


Multiscale Organization of the Brain

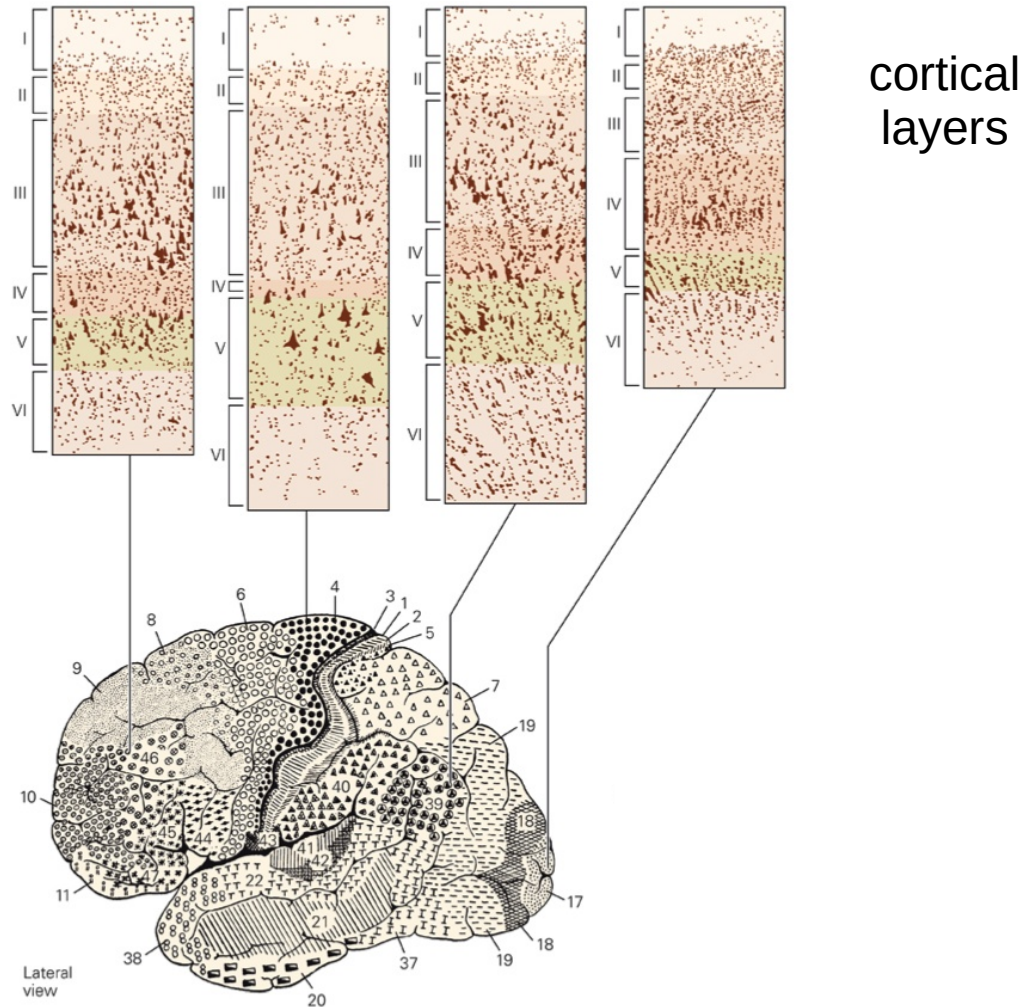


target
neuron

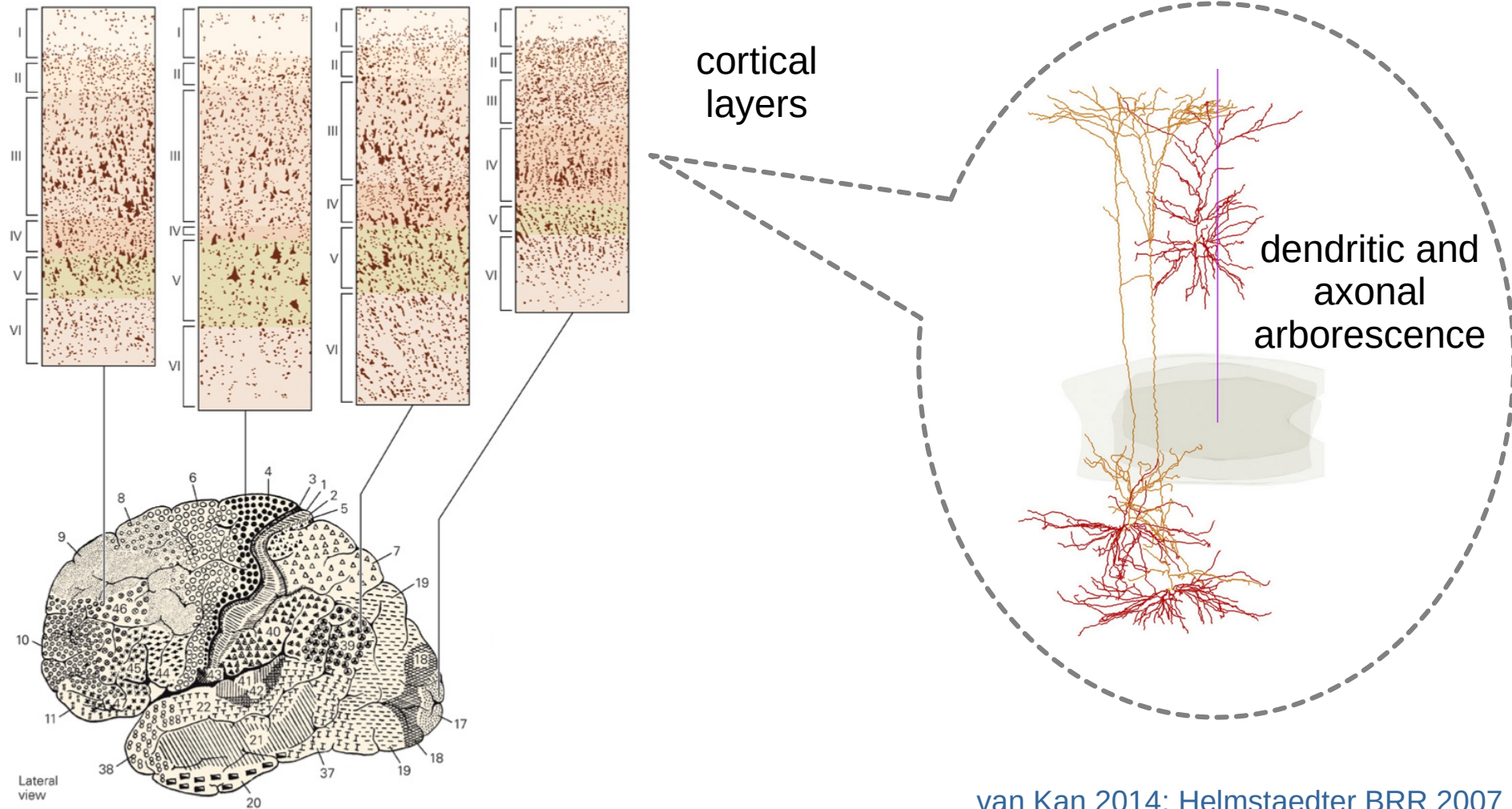
active ionic channels on neuronal cell membrane



Multiscale Organization of the Brain



Multiscale Organization of the Brain

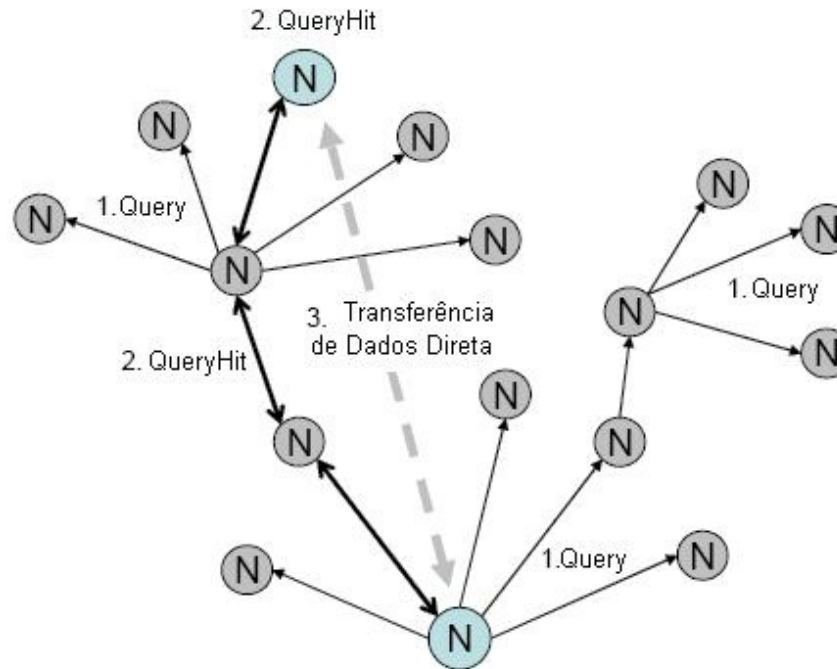


Analogy with Communication Network

- Complexity = vulnerability or resilience?
 - example in engineering: telecommunication, transport, logistic, financial, ... networks
 - potentially fragile (internet cut, traffic jams, COVID, 2007 crisis)

Gnutella network

- peer-to-peer file sharing scheme
- efficiency: short paths

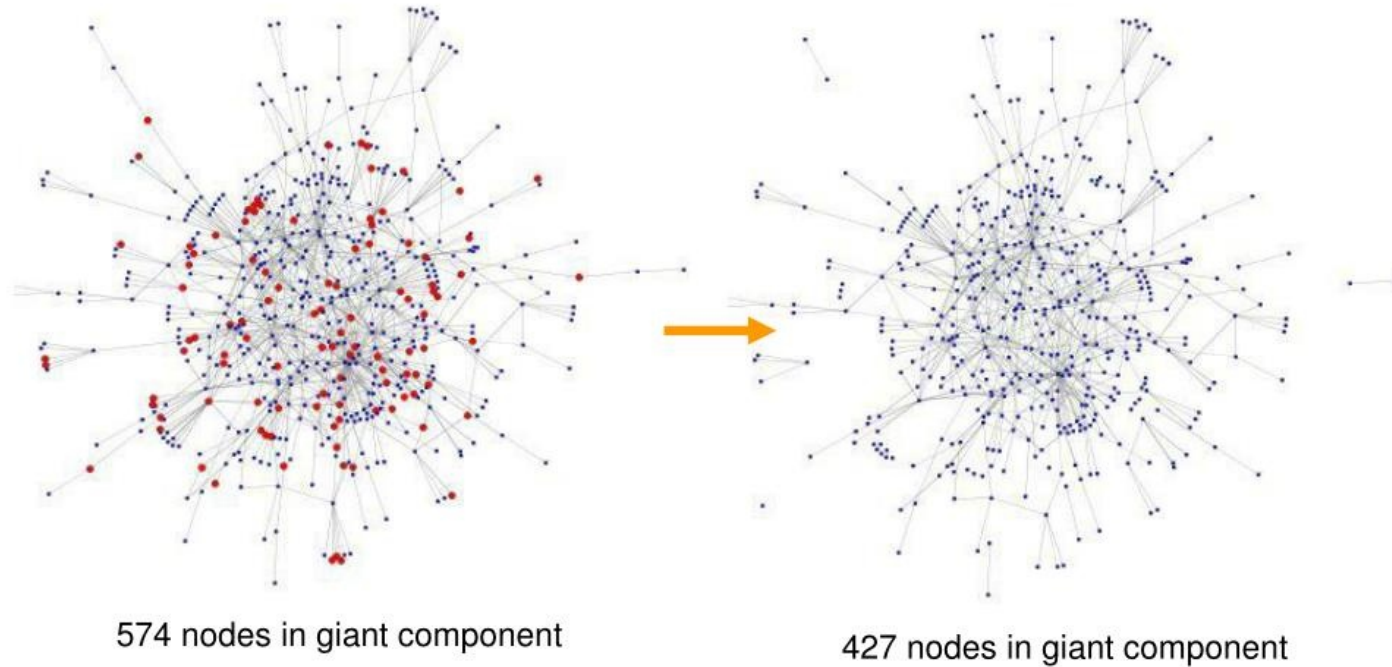


Analogy with Communication Network

- Complexity = vulnerability or resilience?
 - example in engineering: telecommunication, transport, logistic, financial, ... networks
 - potentially fragile (internet cut, traffic jams, COVID, 2007 crisis)

Gnutella network

- peer-to-peer file sharing scheme
- efficiency: short paths
- resilience to node breakdown / attack
- scale-free topology



- Historically functions thought to be localized (cf. Phrenology)
- But modern viewpoint focuses on understanding how distributed neuronal implementations (high-level cognition)
- Limits of analogy with telecommunication network
 - heterogeneous neuron types
 - multiple scales of organization
 - many molecular mechanisms in parallel: synaptic transmission + neuromodulation + metabolism + ...
- Still, some concepts remain important
 - redundant network structure (supporting robustness)
 - dynamic aspect of processes
- How are these properties that affected in neuropathologies?
 - systemic approach at network level

Introduction to Models in Neuroscience

- The brain as a distributed and complex network system
- **Neuroimaging: quantifying the brain**
- Statistical analysis versus classification
- Example 1: diagnosis / prognosis in stroke
- Example 2: whole-brain modeling, The Virtual Brain
- Example 3: characterize structure in multivariate data
- Scikit-learn: formatting data

Magnetic Resonance Imaging (MRI) to Reveal Brain Structure



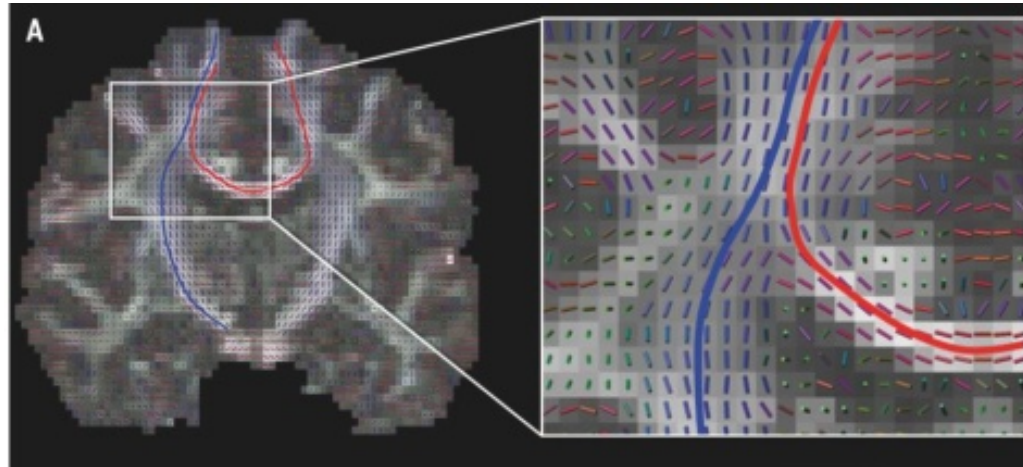
Magnetic Resonance Imaging (MRI) to Reveal Brain Structure



- Identification tissus neuronaux (matière grise versus blanche ; riches en eau, lipides, ...)
- Séquence du scanner : excitation certains atomes (hydrogène)

Tractography: Axonal Fibers = Connectome

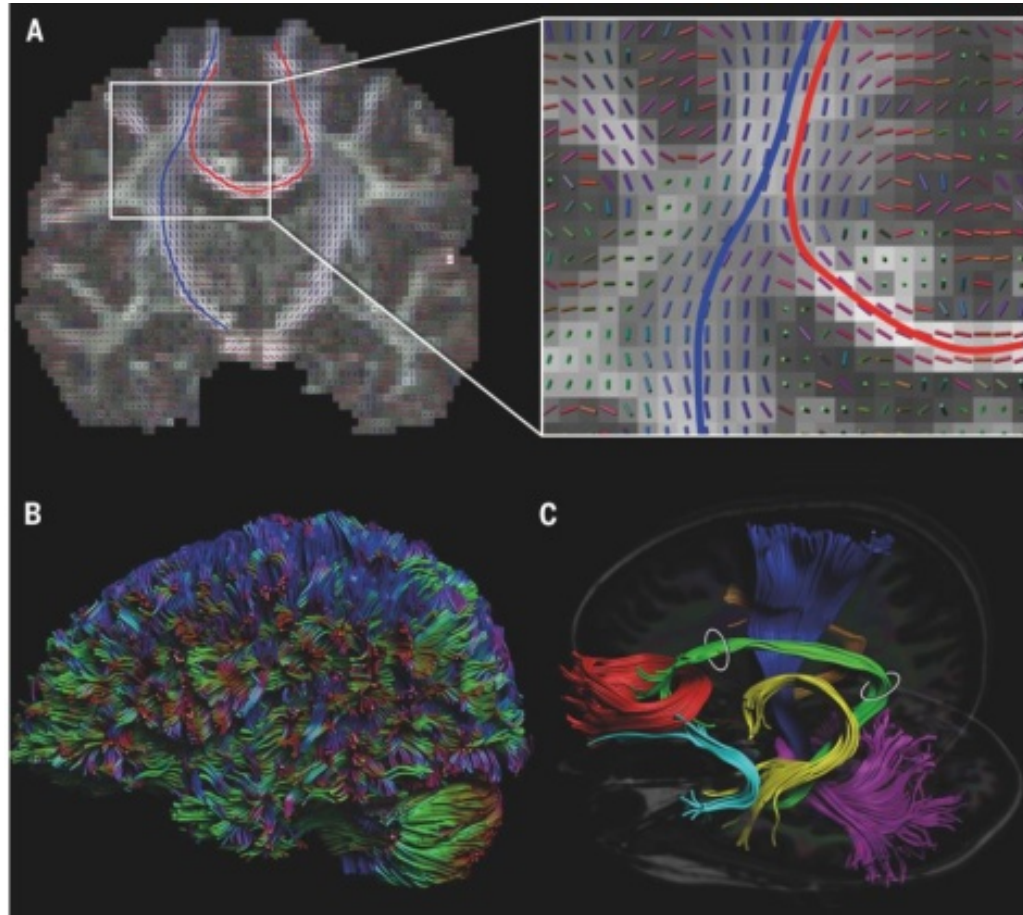
diffusion IRM
(movement of
water molecules)



- corticospinal tract
- corpus callosum

Tractography: Axonal Fibers = Connectome

diffusion IRM
(movement of
water molecules)



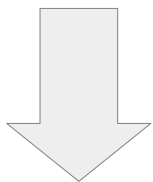
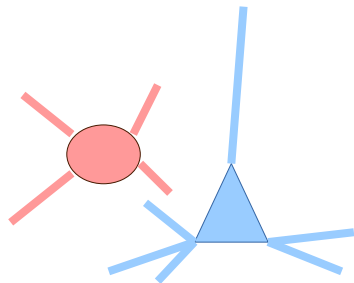
- corticospinal tract
- corpus callosum

whole-brain
connectome

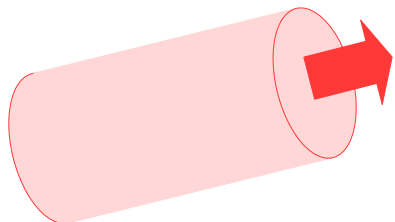
Example fiber
bundles (white
matter)

Functional MRI: Proxy for Neuronal Activity

neuronal activity

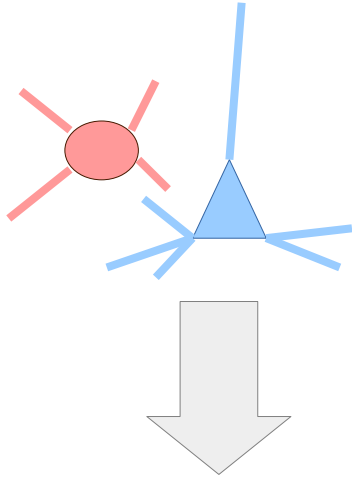


blood influx
(oxygenation)

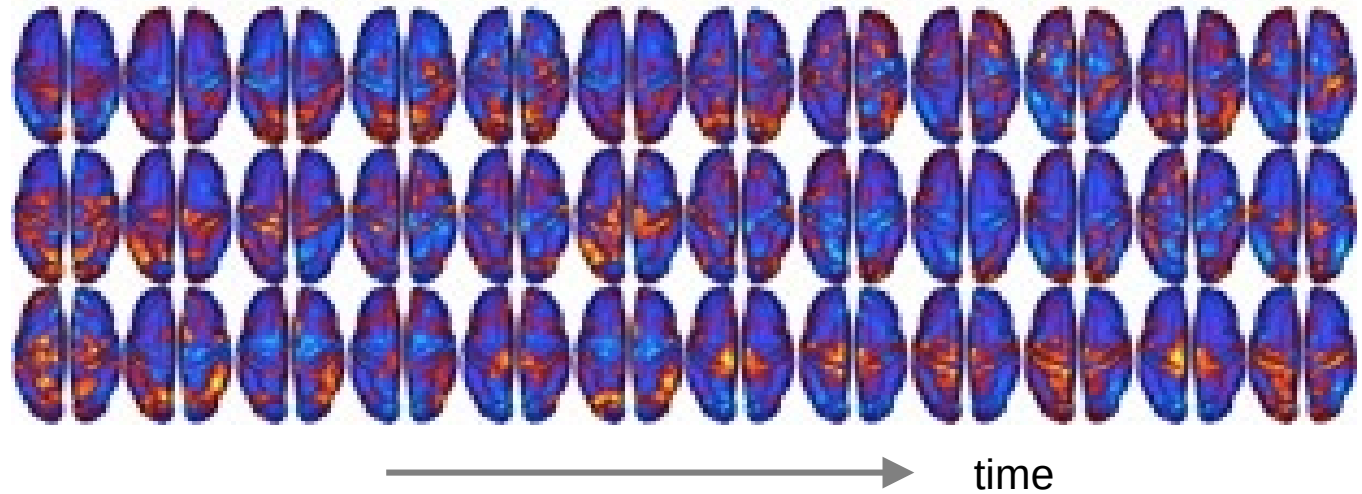
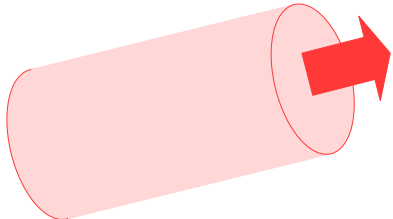


Functional MRI: Proxy for Neuronal Activity

neuronal activity



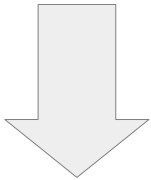
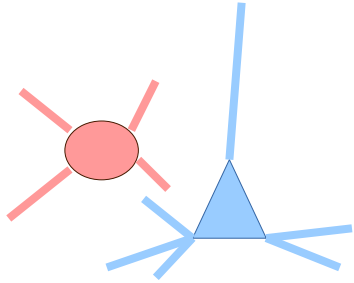
blood influx
(oxygenation)



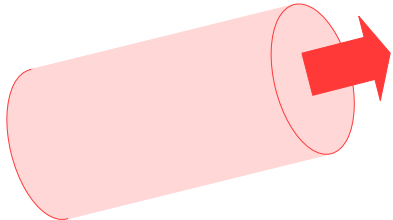
- blood-oxygen-level dependent (BOLD) signal
- indirect measure of neuronal activation
- spatio-temporal structure

Functional MRI: Proxy for Neuronal Activity

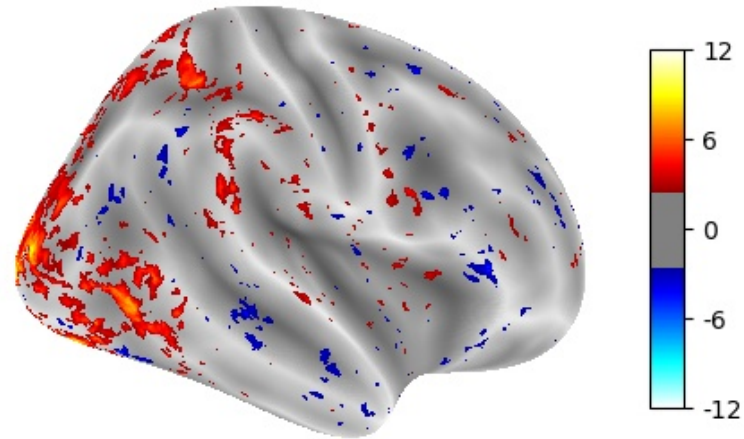
neuronal activity



blood influx
(oxygenation)



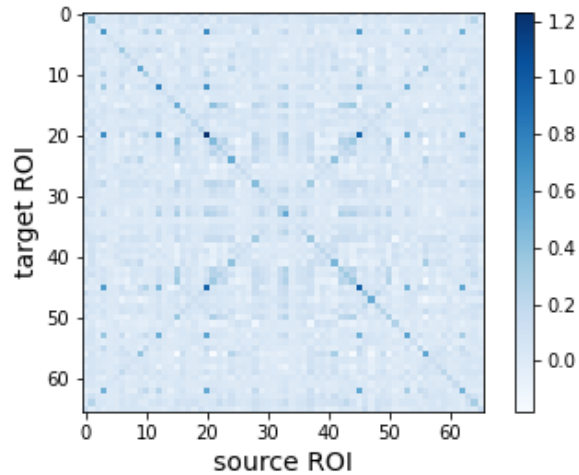
change in neuronal activity during emotional task



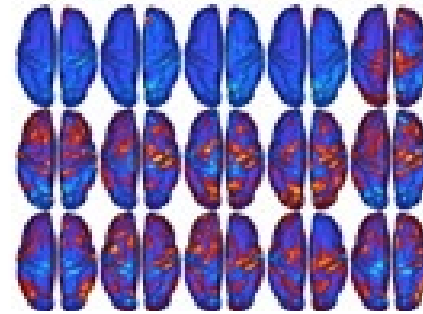
Introduction to Models in Neuroscience

- The brain as a distributed and complex network system
- Neuroimaging: quantifying the brain
- **Statistical analysis versus classification**
- Example 1: diagnosis / prognosis in stroke
- Example 2: whole-brain modeling, The Virtual Brain
- Example 3: characterize structure in multivariate data
- Scikit-learn: formatting data

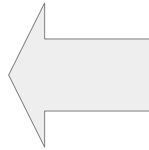
Functional Connectivity: Signature of Brain Dynamic Activity



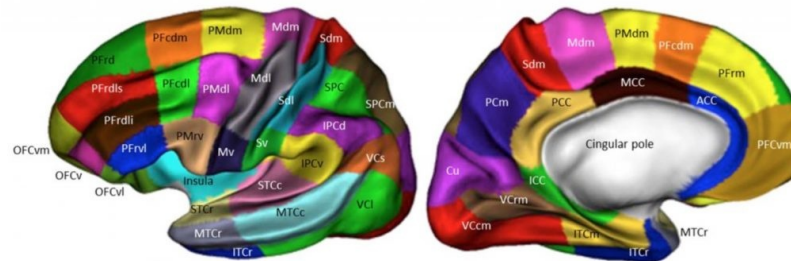
- 100 regions
- 5000 matrix elements



BOLD images
(voxel space)

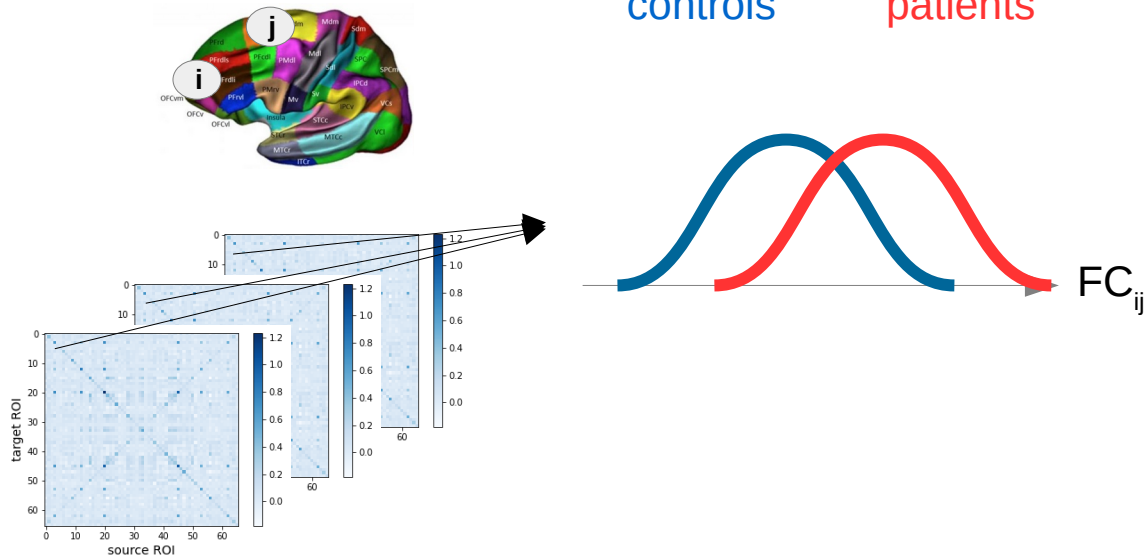


Brain Parcellation



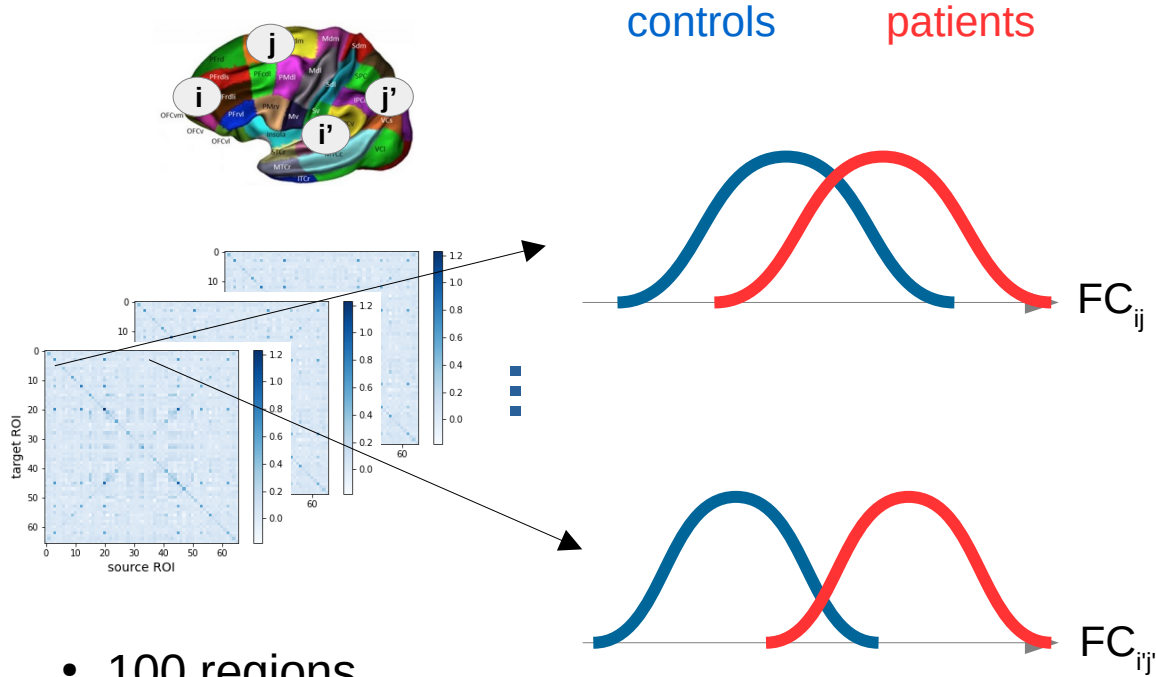
MarsAtlas, Auzias, Brovelli, et al.

Descriptive Statistics



- 100 regions
- 5000 matrix elements
- patients et healthy controls

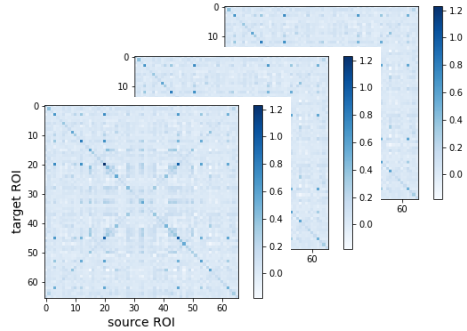
Descriptive Statistics



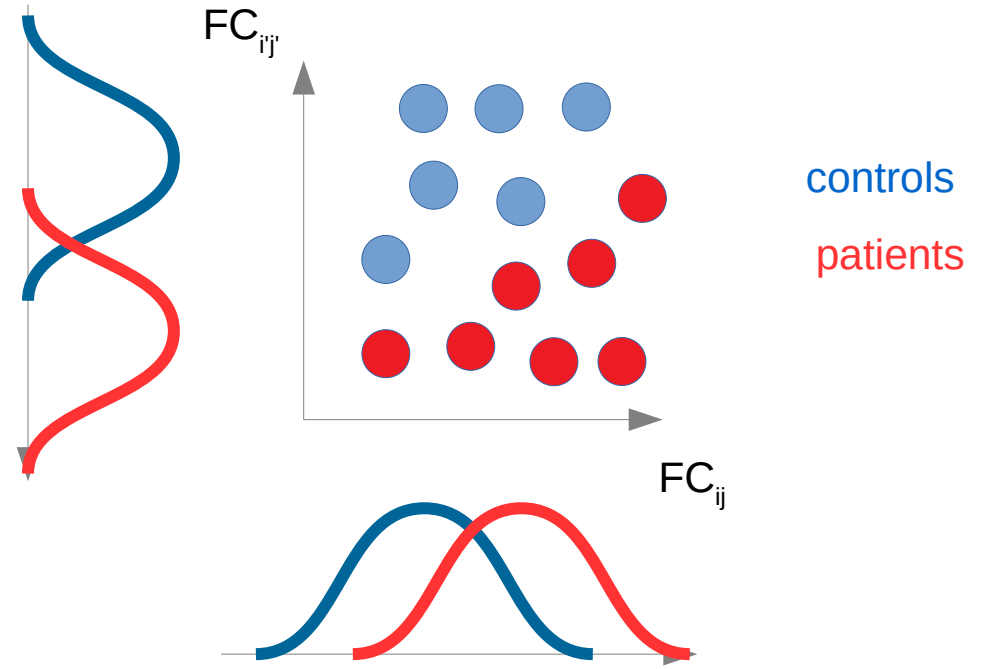
- 100 regions
- 5000 matrix elements
- patients et healthy controls

Which interactions change most across patients and controls?

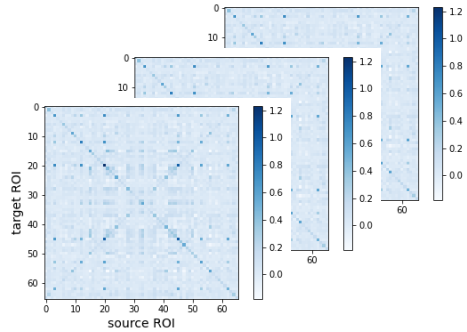
Descriptive Statistics



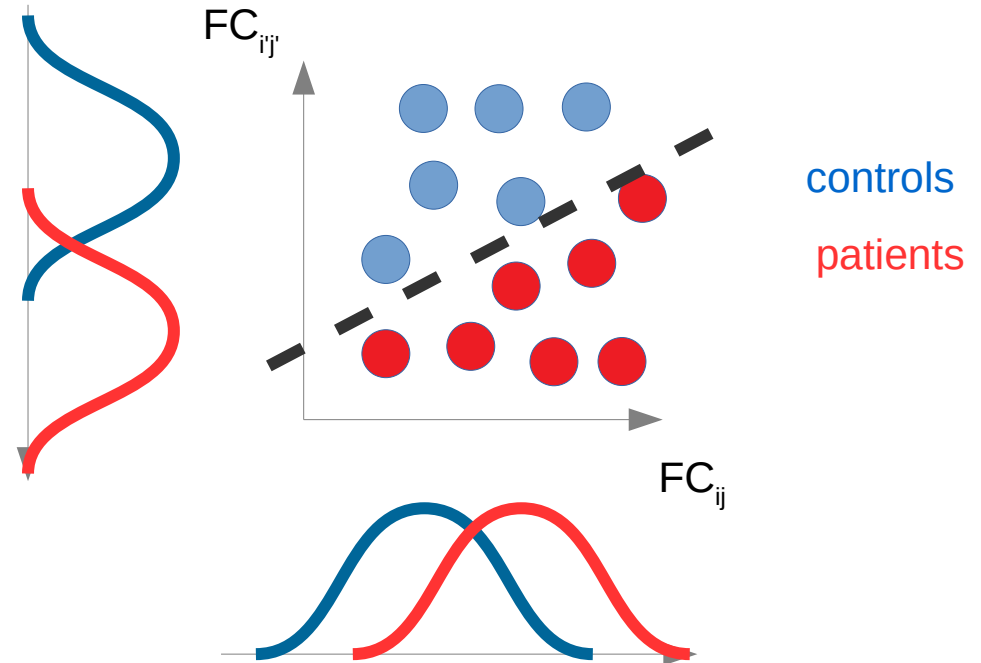
- 100 regions
- 5000 matrix elements
- patients et healthy controls



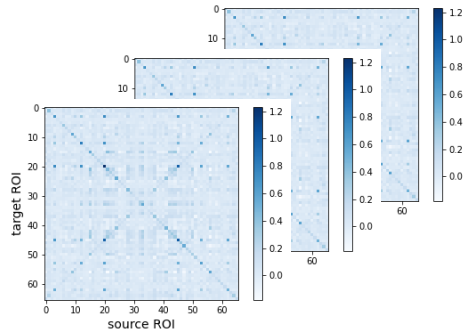
Descriptive Statistics



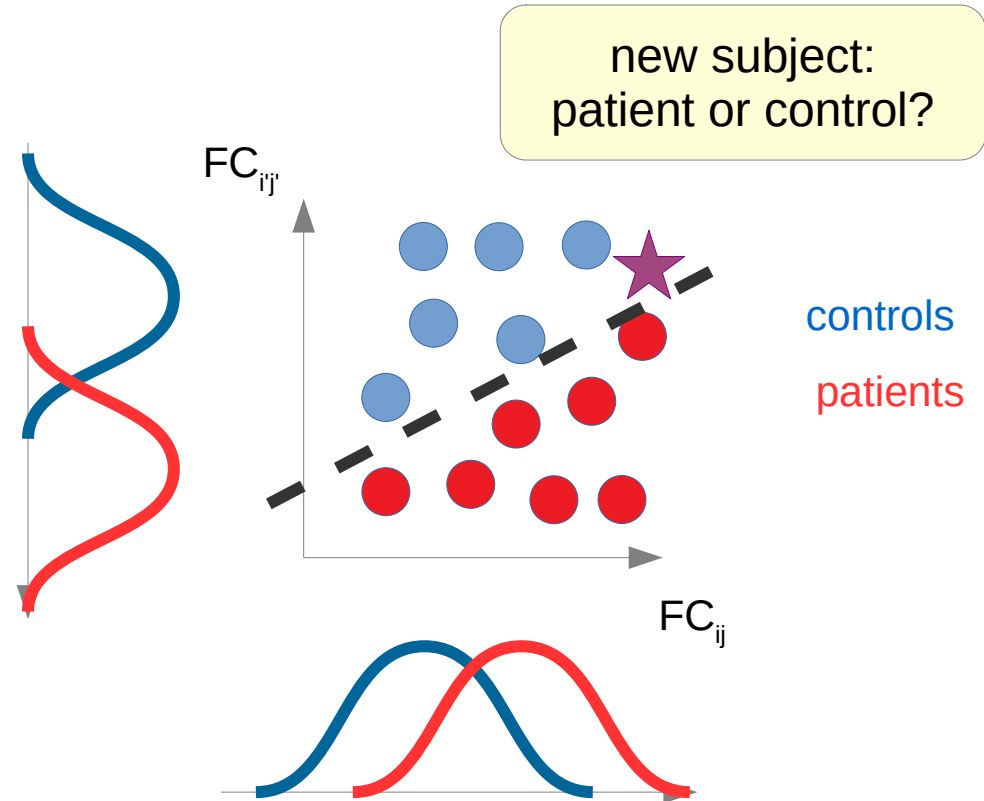
- 100 regions
- 5000 matrix elements
- patients et healthy controls



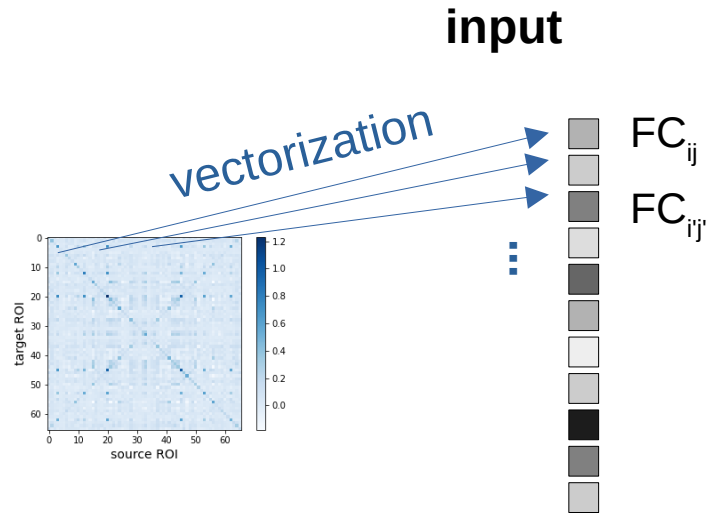
Predictive Statistics



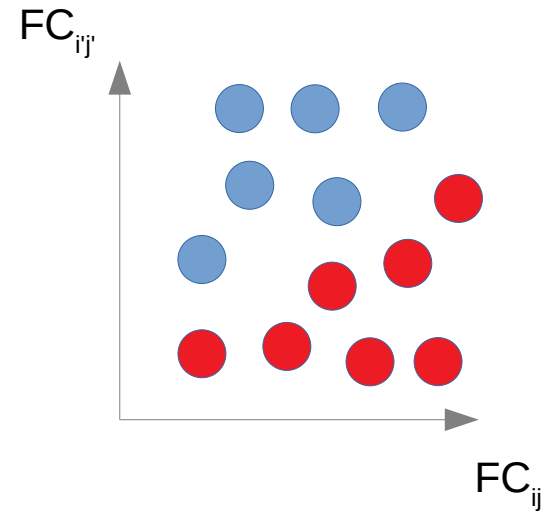
- 100 regions
- 5000 matrix elements
- patients et healthy controls



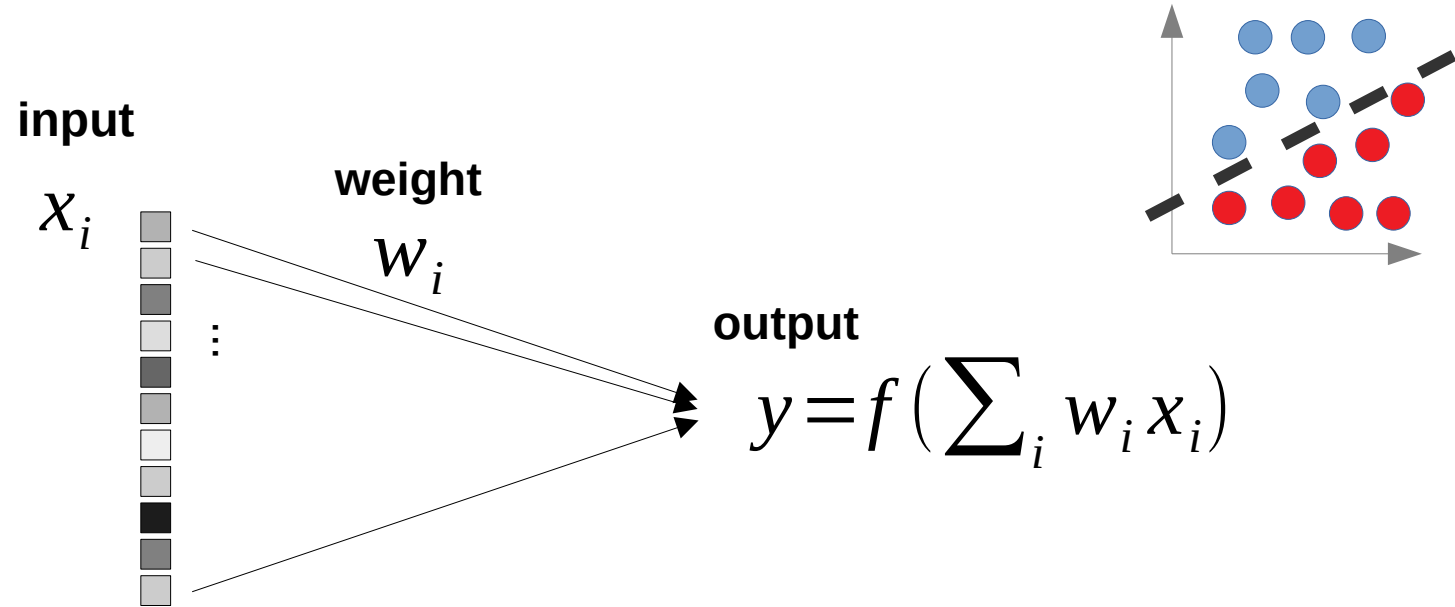
Predictive Statistics



- 100 regions
- 5000 matrix elements

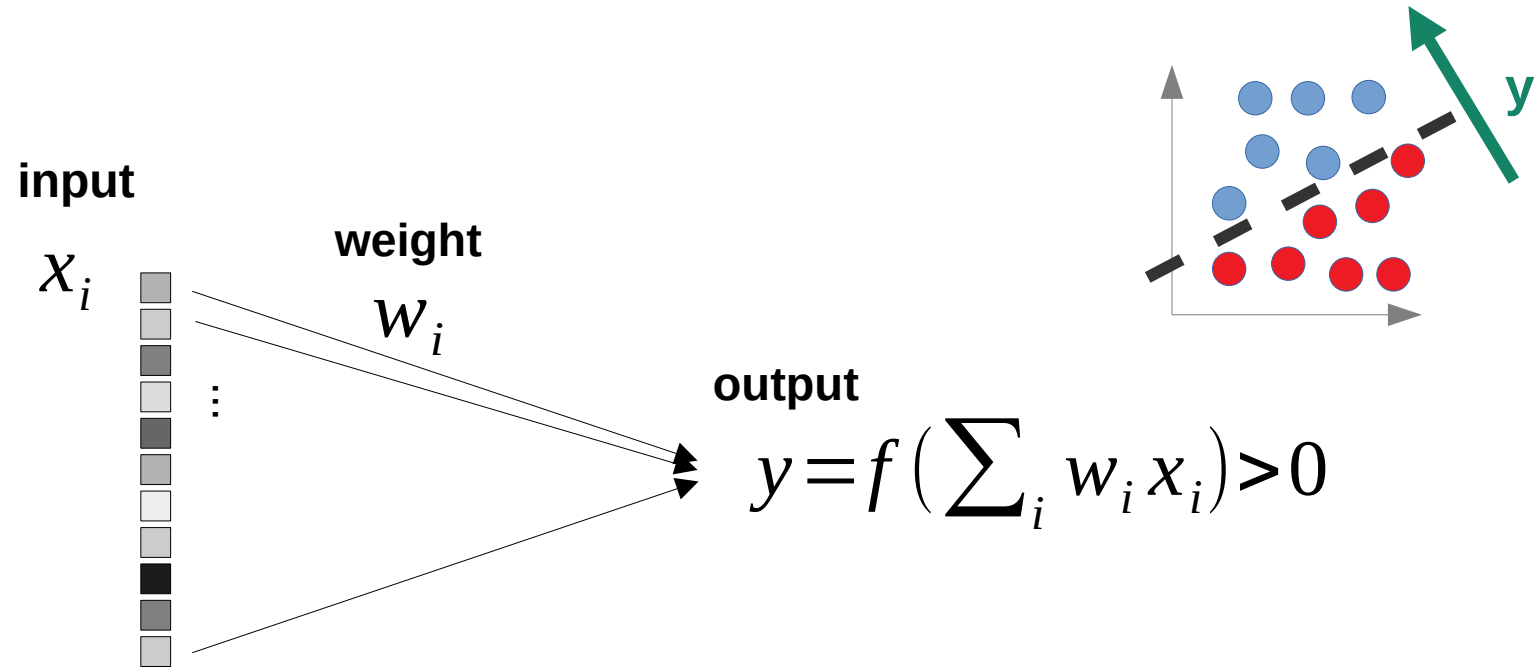


Predictive Statistics



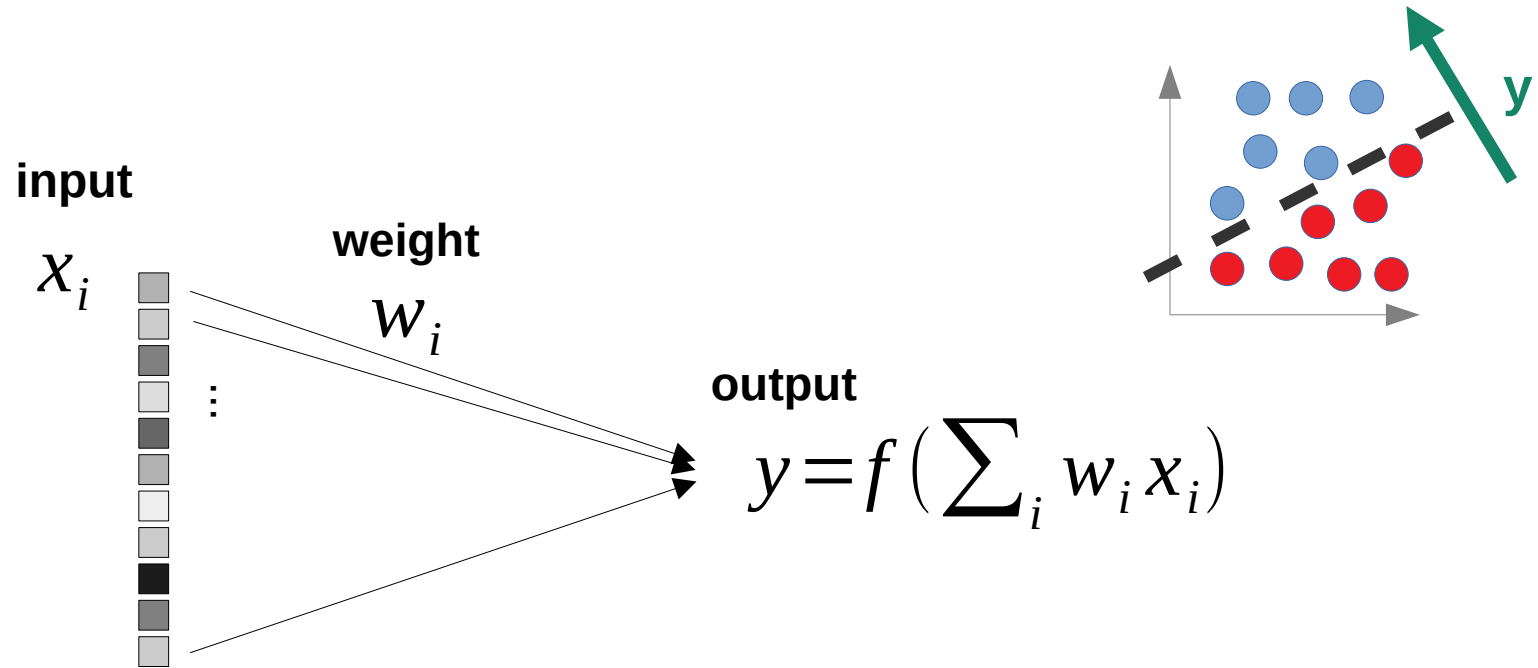
- Output indicates class/category for input (subject signature)

Predictive Statistics



- Output indicates class/category for input (subject signature)

Predictive Statistics

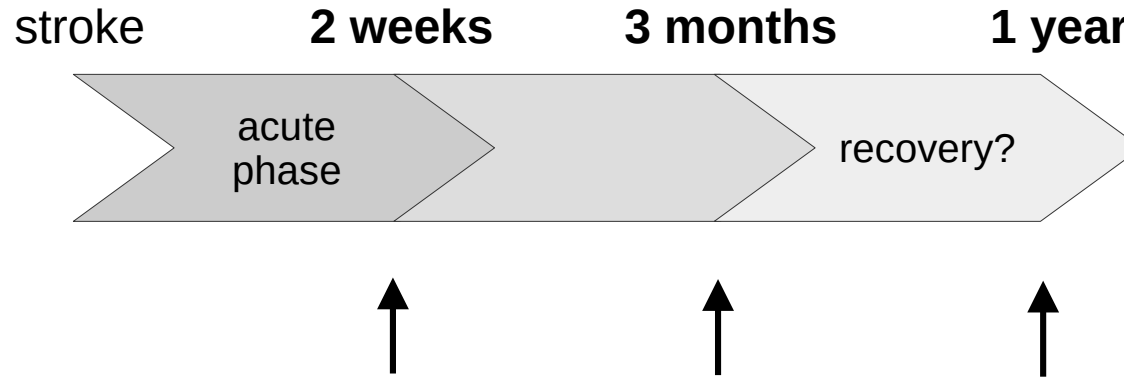


- Output indicates class/category for input (subject signature)
- Training weights for batch of inputs to obtain best possible prediction

Introduction to Models in Neuroscience

- The brain as a distributed and complex network system
- Neuroimaging: quantifying the brain
- Statistical analysis versus classification
- **Example 1: diagnosis / prognosis in stroke**
- Example 2: whole-brain modeling, The Virtual Brain
- Example 3: characterize structure in multivariate data
- Scikit-learn: formatting data

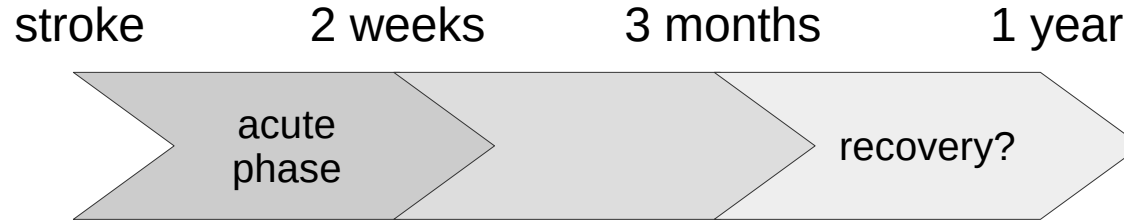
Monitoring and Predicting Recovery for Stroke Patients



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



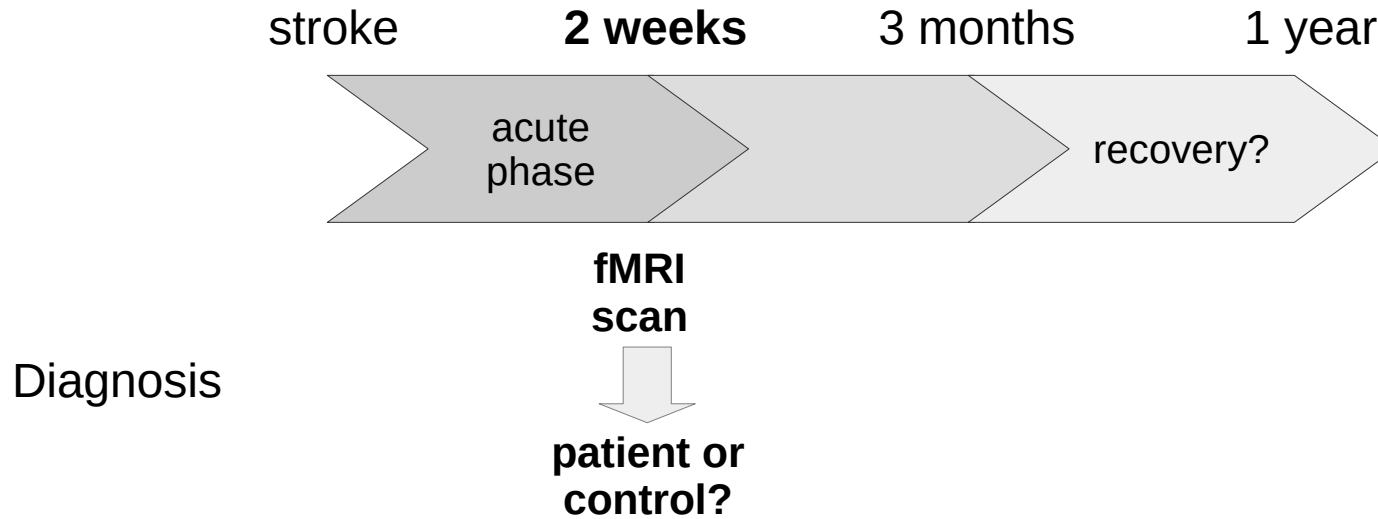
Monitoring and Predicting Recovery for Stroke Patients



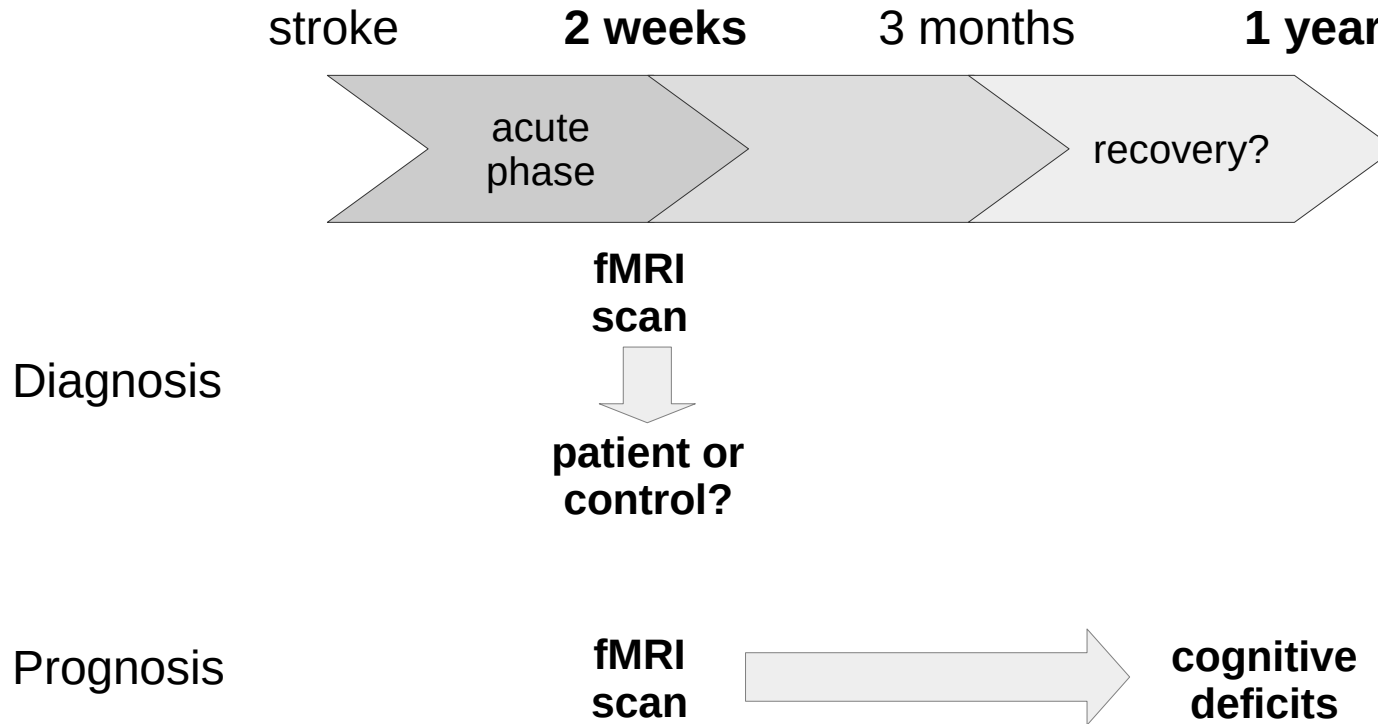
- 132 patients, 25 controls
- 300 brain regions (ROIs)
- 80%-20% train-test (stratified split)
- classifier: logistic regression



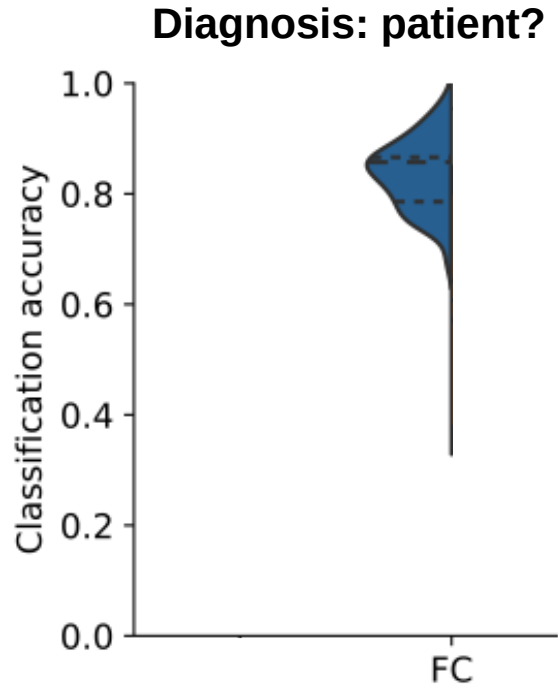
Monitoring and Predicting Recovery for Stroke Patients



Monitoring and Predicting Recovery for Stroke Patients

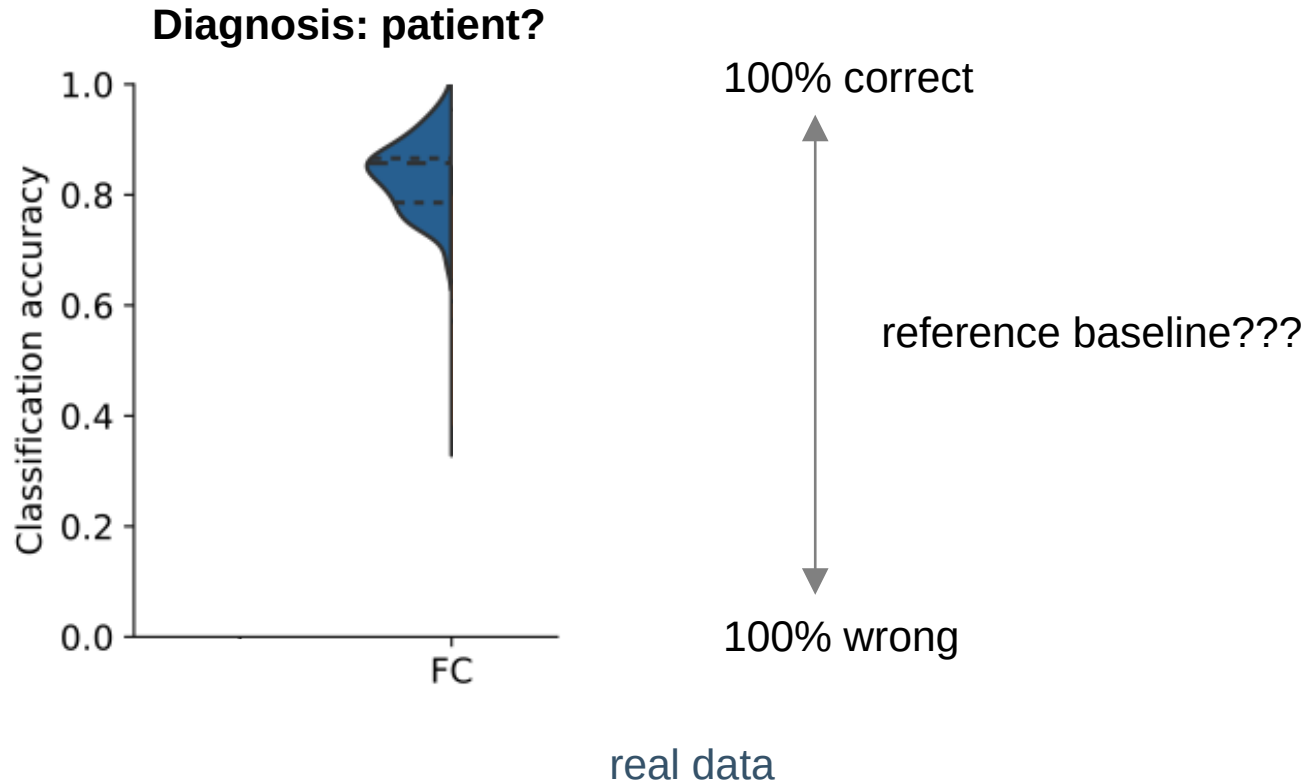


Prediction from FC

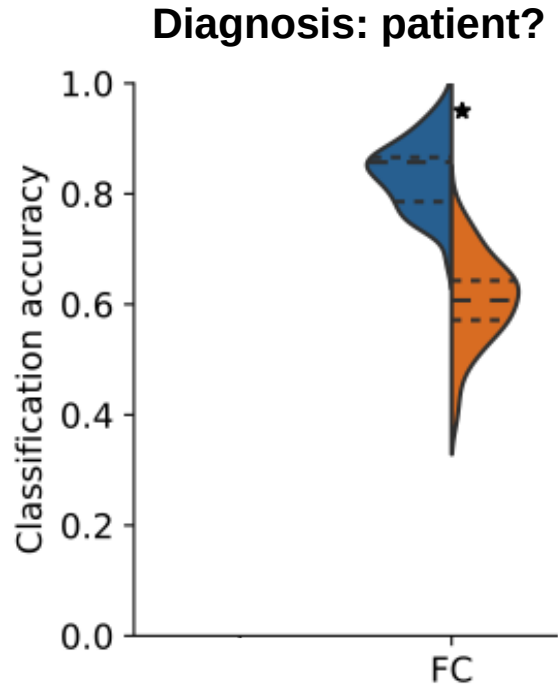


real data

Prediction from FC

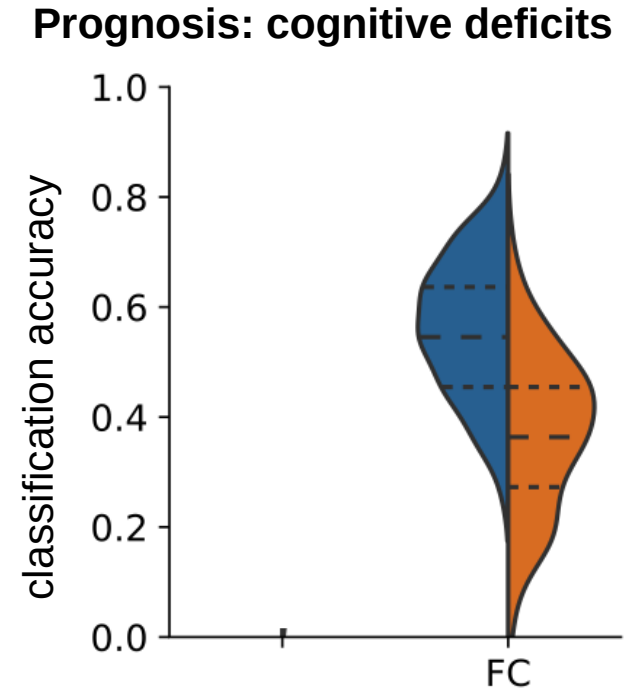
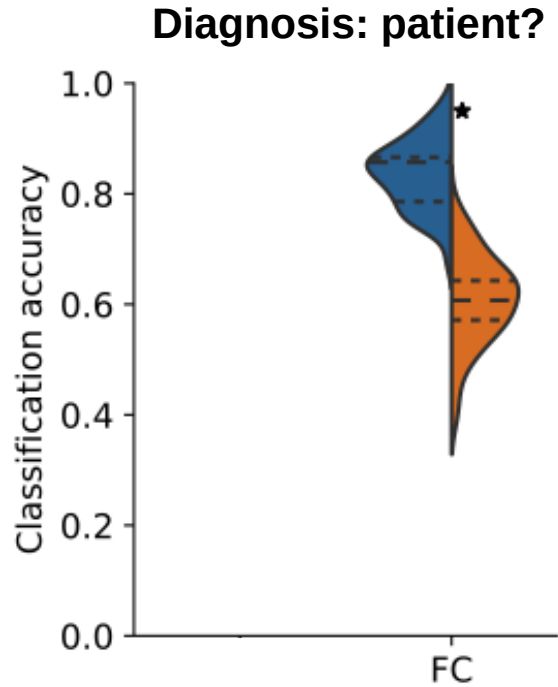


Prediction from FC



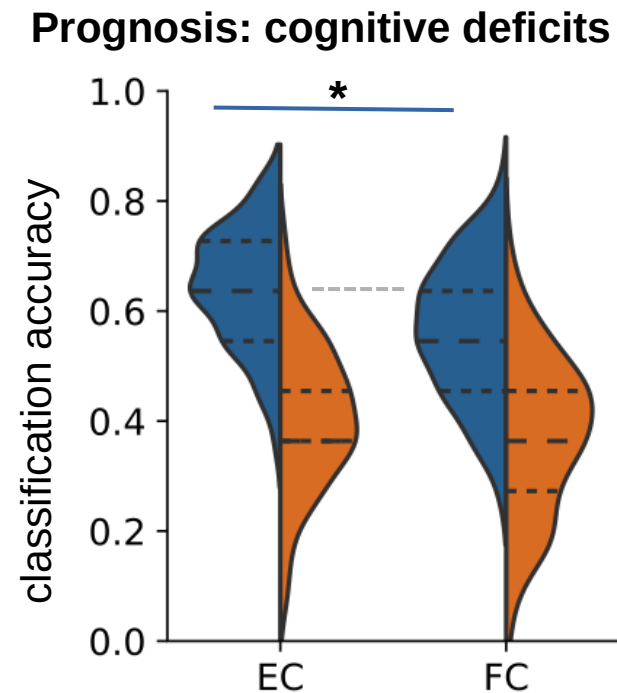
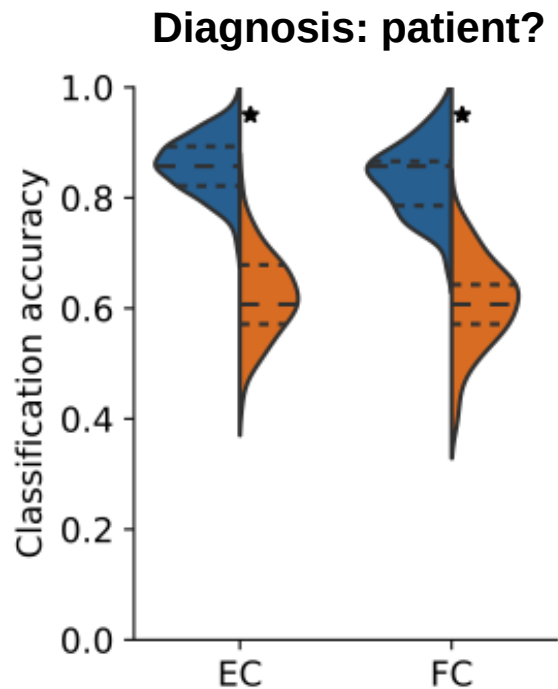
real data
chance level
(shuffling surrogates)

Prediction from FC



real data
chance level
(shuffling surrogates)

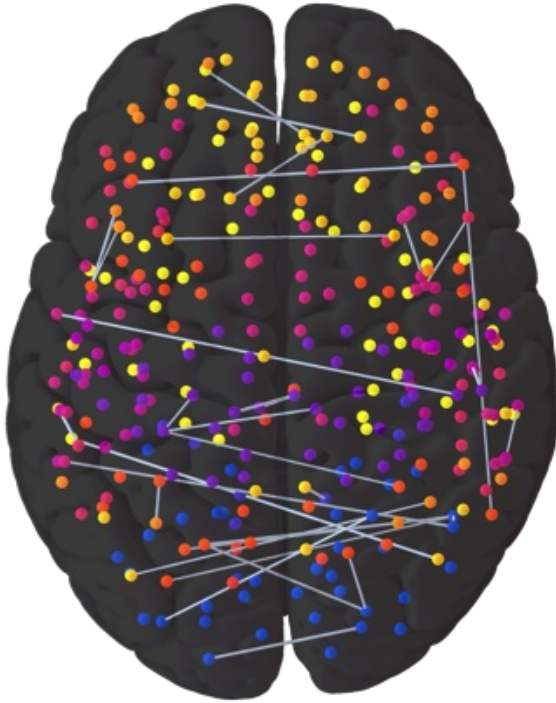
Model-based EC can Outperform FC



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

Introduction to Models in Neuroscience

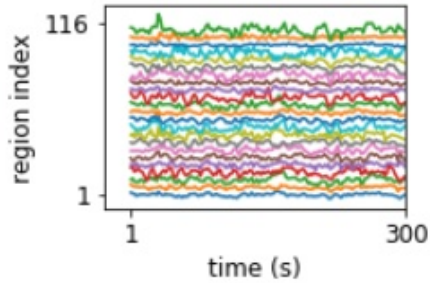
- The brain as a distributed and complex network system
- Neuroimaging: quantifying the brain
- Statistical analysis versus classification
- Example 1: diagnosis / prognosis in stroke
- **Example 2: whole-brain modeling, The Virtual Brain**
- Example 3: characterize structure in multivariate data
- Scikit-learn: formatting data

Model-Based Analysis of Neuroimaging Data

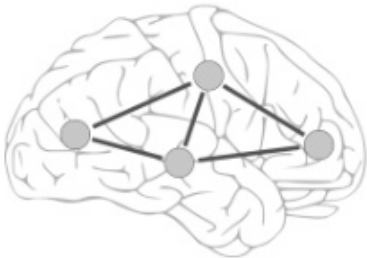
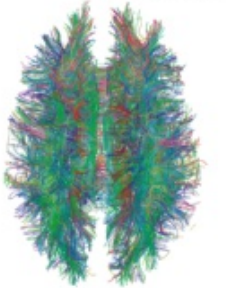
parcellation



fMRI/BOLD signals



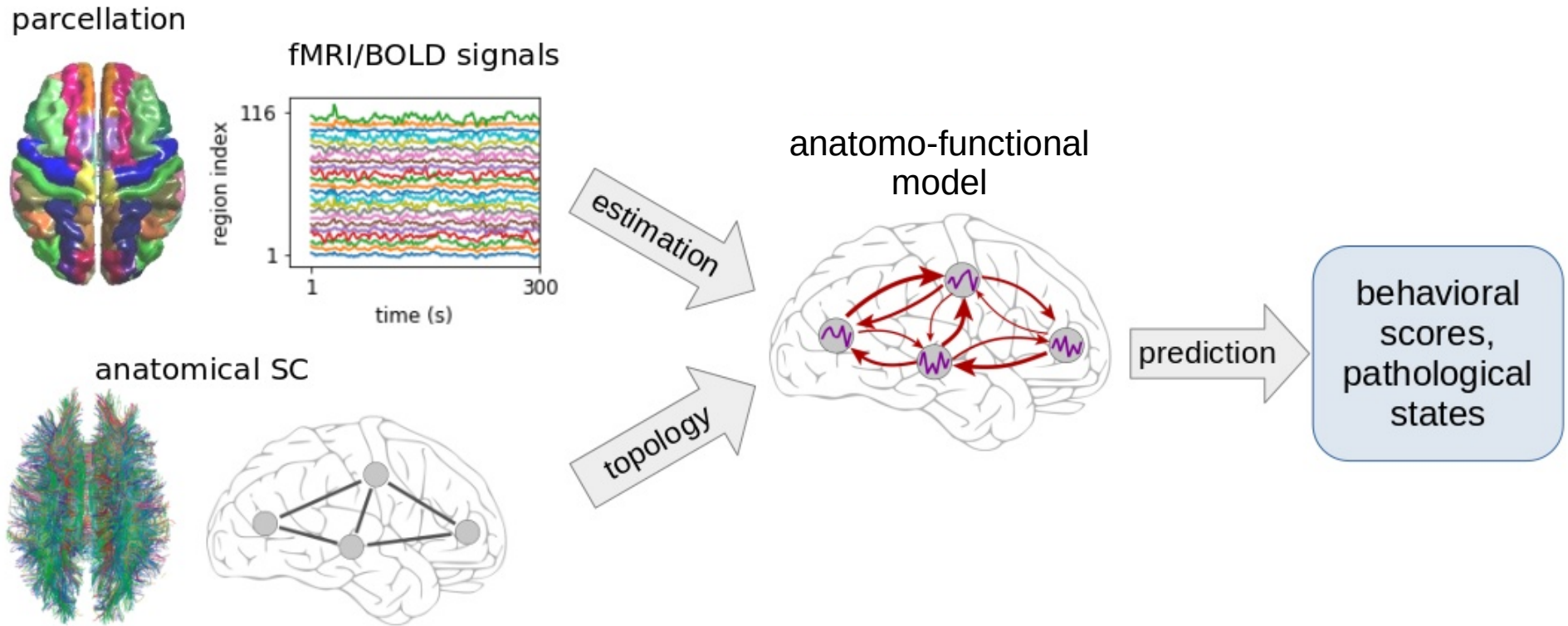
anatomical SC



prediction

behavioral
scores,
pathological
states

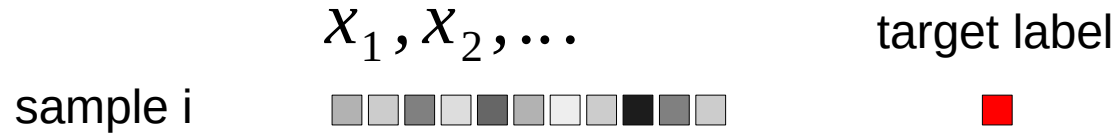
Model-Based Analysis of Neuroimaging Data



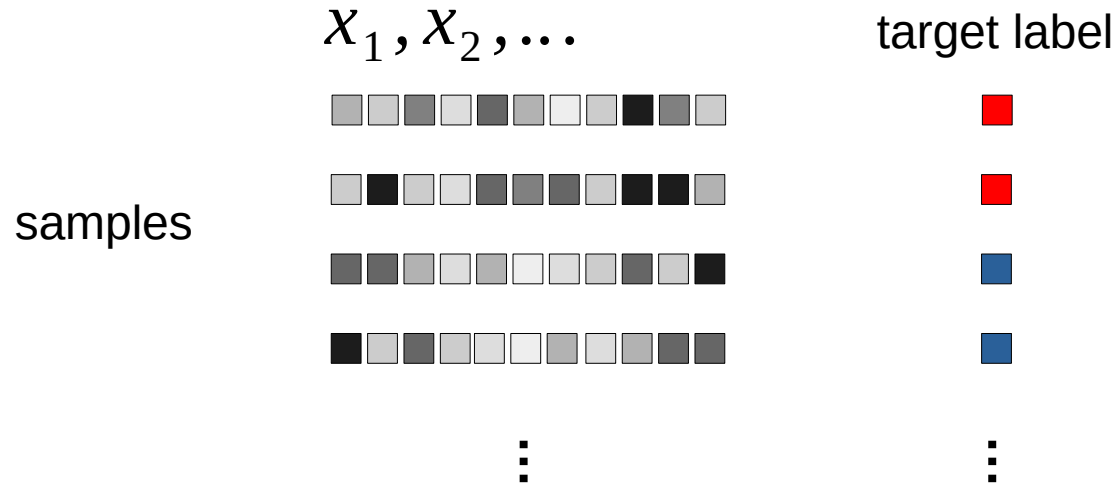
Introduction to Models in Neuroscience

- The brain as a distributed and complex network system
- Neuroimaging: quantifying the brain
- Statistical analysis versus classification
- Example 1: diagnosis / prognosis in stroke
- Example 2: whole-brain modeling, The Virtual Brain
- Example 3: characterize structure in multivariate data
- **Scikit-learn: formatting data**

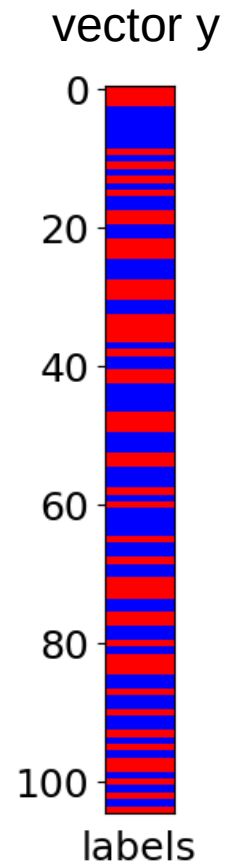
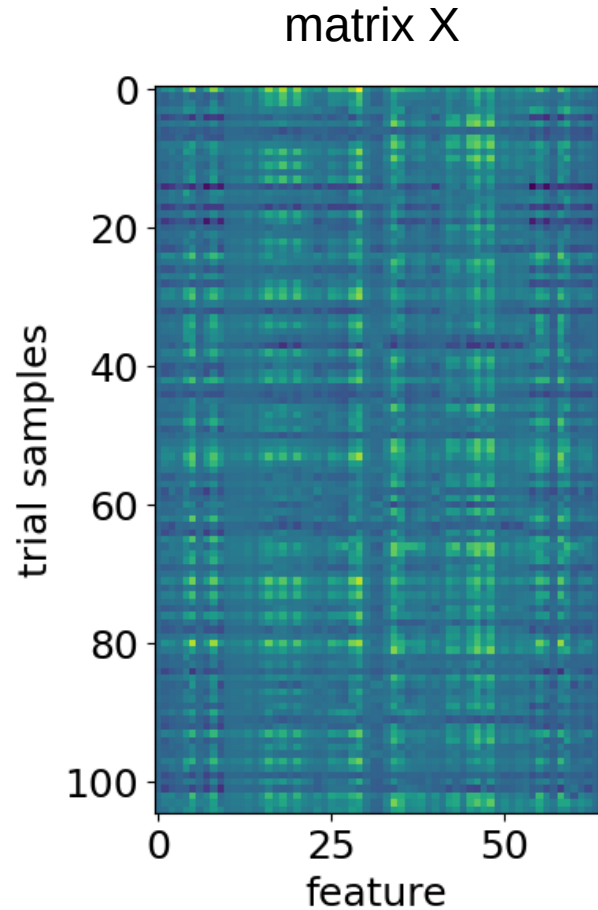
Organization of Data for Classification



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.

fit	train the using data (X,y), sets the transformation (PCA with a number of components), etc.
transform	maps the X to y (prediction for classifier, new coordinates for PCA, etc.)
fit_transform	
score	evaluation (classification accuracy, clustering quality, etc.)

<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

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

Practice

- Anaconda installation
- Environment yml file: required libraries
- Use jupyter-notebook (or jupyter-lab)