Advanced Python for Neuroscientists

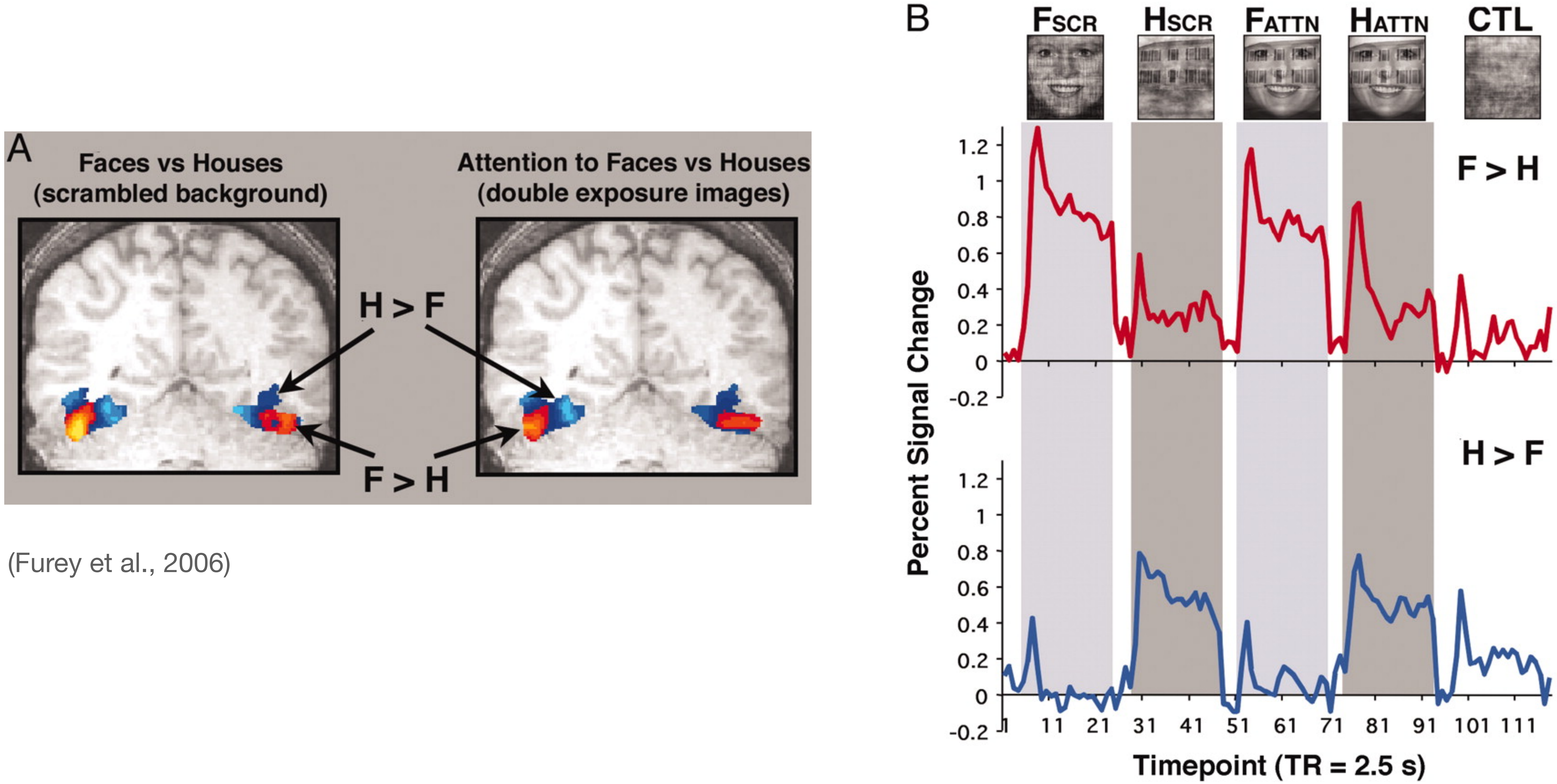
Lecture 4: Classification (Decoding) / GLM (Encoding)

2022/07/07

Classification (Decoding)

An example in fMRI

- “Traditional” view of neural response - selective response



(Furey et al., 2006)

Classification (Decoding)

An example in fMRI

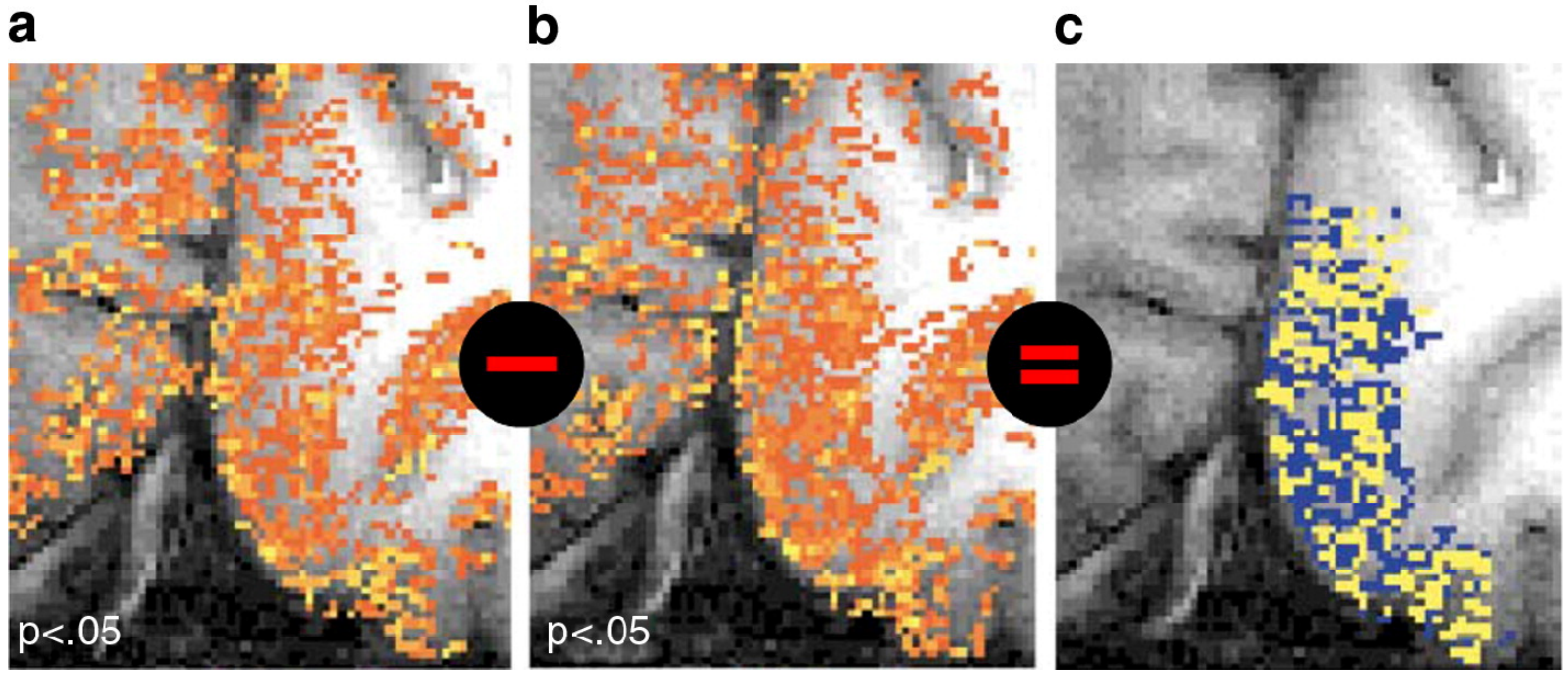
Analyzing for information, not activation, to exploit high-resolution fMRI

Nikolaus Kriegeskorte* and Peter Bandettini

Section on Functional Imaging Methods, Lab of Brain and Cognition, National Institute of Mental He

Received 15 September 2006; revised 22 December 2006; accepted 6 February 2007
Available online 27 February 2007

- Searchlight method in fMRI
    - Real information is in $a - b$, $a + .5b$,…



Classification (Decoding)

1. fMRI voxel activation
2. Gene expressions
3. Single neuron spike count
4. Pixel luminance patten

…

Data → **Classes**

1. Visual stimulus
2. Cell type
3. Attend left / right
4. Dog / cat

…

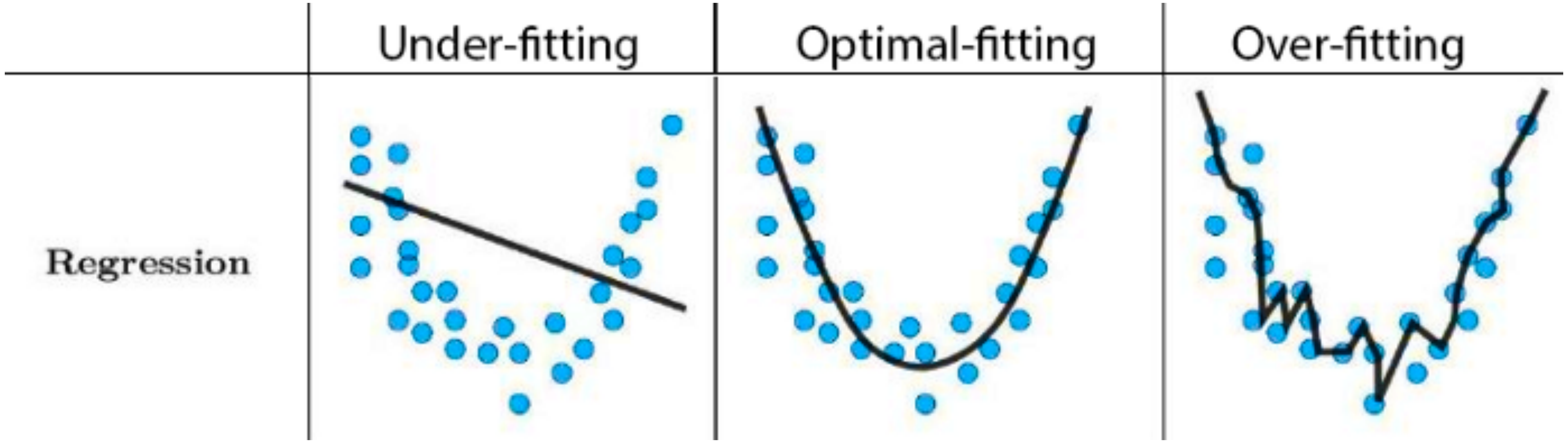
Basic machine learning concepts

Basic machine learning concepts

- Cross-validation

Basic machine learning concepts

- Cross-validation
- Bias-variance trade-off



Basic machine learning concepts

- Cross-validation
- Bias-variance trade-off
- Supervised vs. unsupervised learning

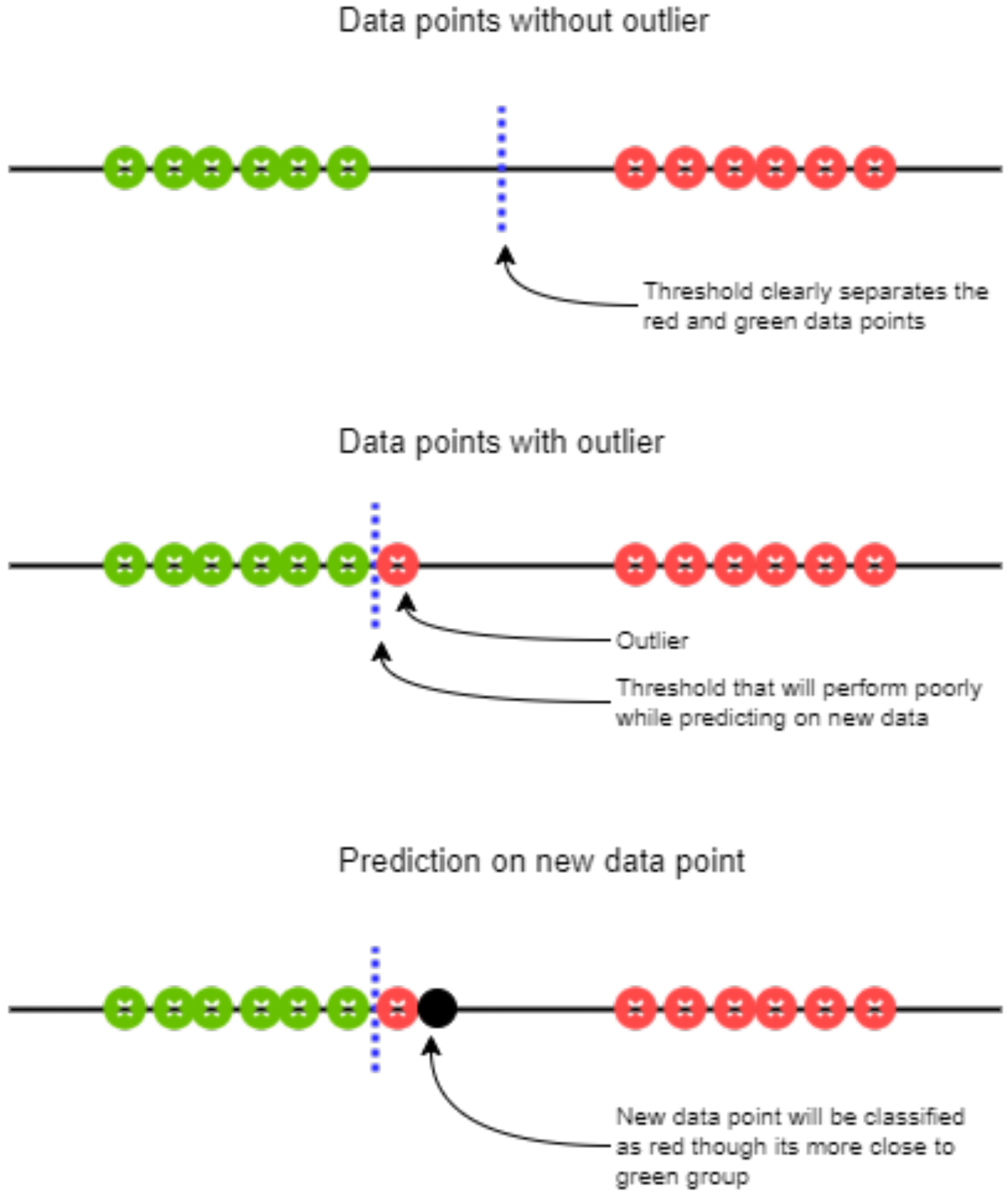
Basic machine learning concepts

- Cross-validation
- Bias-variance trade-off
- Supervised vs. unsupervised learning
- Cost function - the thing you're trying to minimize
  - E.g. $\sum (y - \hat{y})^2, \ \sum (y - \hat{y})^2 + \lambda |\mathbf{w}|$
  - Misclassification
    - 0-1 loss; hinge loss (SVM); log loss (logistic regression)

Support Vector Machine

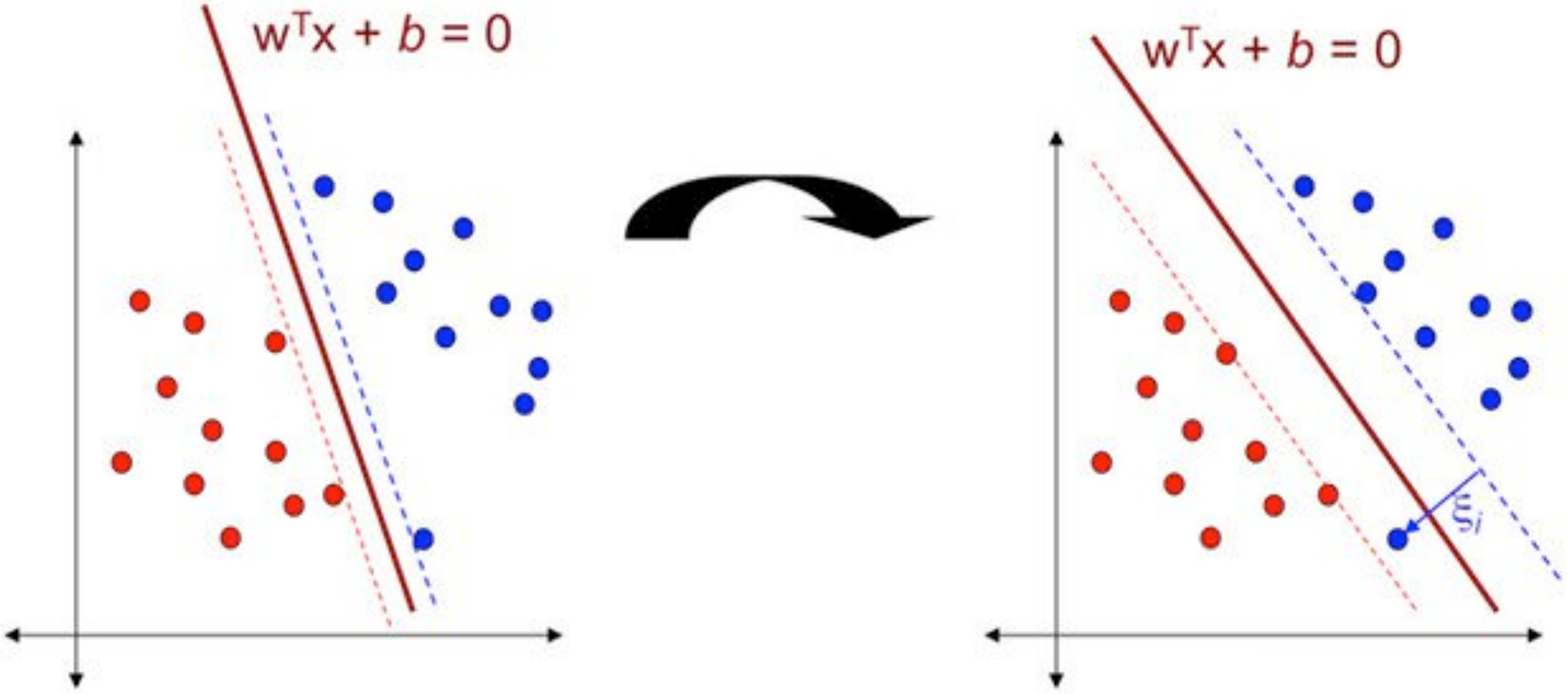
Support Vector Machine

- Soft margin
  - Maximum margin classifier
  - Support vector classifier



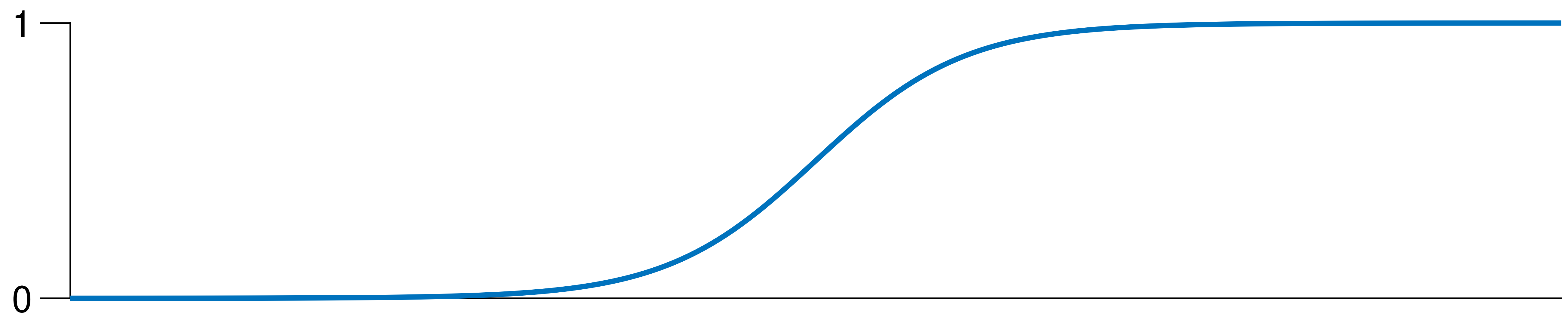
Support Vector Machine

- Soft margin
    - Hyper-parameter: how much do you care about misclassification



Support Vector Machine

- SVM vs logistic regression
  - Different cost function
  - SVM performs better with data that is less structured



Naive Bayes

Bayesian Inference

- Probability based
  - Prior
  - likelihood
  - posterior

$$\overset{\text{Posterior}}{p(H=h \mid Y=y)} = \frac{\overset{\text{prior}}{p(H=h)}\,\overset{\text{Likelihood}}{p(Y=y \mid H=h)}}{p(Y=y)}$$

$$\propto p(H=h)\,p(Y=y \mid H=h)$$

- Weighted average of prior
- Update prior belief with likelihood, get posterior

Bayesian Inference

Example: COVID testing

- sensitivity (true positive rate) $p(Y = 1 \mid H = 1)$
- specificity (true negative rate) $p(Y = 0 \mid H = 0)$

Likelihood

- Prevalence of the disease $p(H = 1)$ Prior

- Estimate $p(H \mid Y)$ Posterior

Bayesian Inference

Example: COVID testing

- How likely is that you actually have COVID when receiving a positive test?
  - If you know ~ 10% of people are infected in your area

$$p(H=1 \mid Y=1) = \frac{p(Y=1 \mid H=1)p(H=1)}{p(Y=1 \mid H=1)p(H=1) + p(Y=1 \mid H=0)p(H=0)}$$

$$= \frac{\text{TPR} \times \text{prior}}{\text{TPR} \times \text{prior} + \text{FPR} \times (1 - \text{prior})}$$

$$= \frac{0.875 \times 0.1}{0.875 \times 0.1 + 0.025 \times 0.9} = 0.795$$

Bayesian Inference
Example: COVID testing

- How likely is that you actually have COVID when receiving a positive test?
  - If you know ~ 1% of people are infected in your area

$$p(H=1 \mid Y=1) = \frac{p(Y=1 \mid H=1)p(H=1)}{p(Y=1 \mid H=1)p(H=1) + p(Y=1 \mid H=0)p(H=0)}$$

$$= \frac{\text{TPR} \times \text{prior}}{\text{TPR} \times \text{prior} + \text{FPR} \times (1 - \text{prior})}$$

$$= \frac{0.875 \times 0.01}{0.875 \times 0.01 + 0.025 \times 0.99} = 0.2612$$

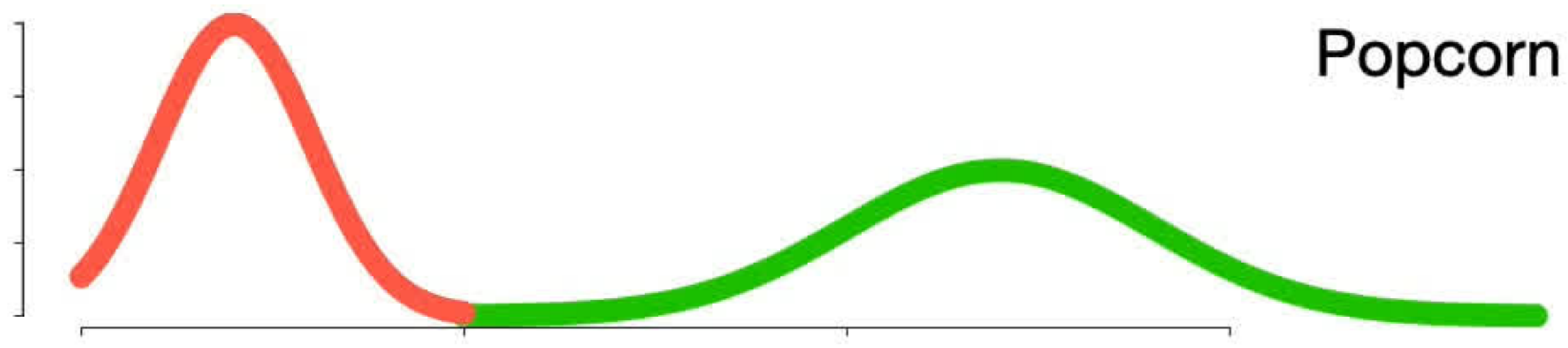
Naive Bayes Classifier

- $p(\text{Class} \mid \text{Data}) = \dfrac{p(\text{Class})p(\text{Data} \mid \text{Class})}{p(\text{Data})}$
- $p(C_k \mid \mathbf{x}) = \dfrac{p(C_k)p(\mathbf{x} \mid C_k)}{p(\mathbf{x})} \propto p(C_k)p(\mathbf{x} \mid C_k)$

Naive Bayes Classifier

- $p(C_k \mid \mathbf{x}) \propto p(C_k)\boxed{p(\mathbf{x} \mid C_k)}$
- Assume independence
- $p(\mathbf{x} \mid C_k) = \prod_i p(x_i \mid C_k)$

Popcorn (grams)	Soda Pop (ml)	Candy (grams)
24.3	750.7	0.2
28.2	533.2	50.5
etc.	etc.	etc.



Popcorn (grams)	Soda Pop (ml)	Candy (grams)
2.1	120.5	90.7
4.8	110.9	102.3
etc.	etc.	etc.

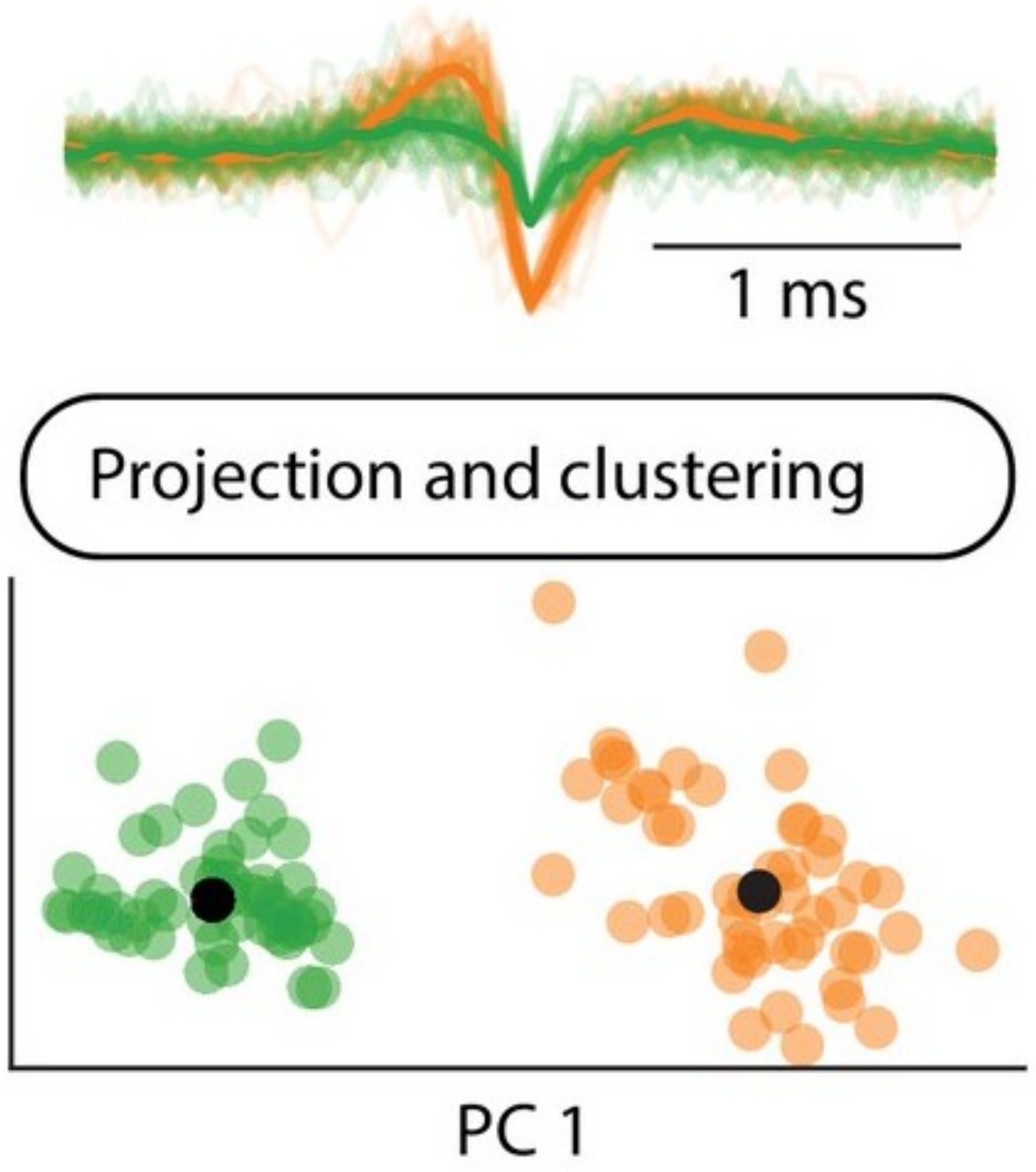
Gaussian Naive Bayes is named after the **Gaussian** distributions that represent the data in the **Training Dataset**.

K-means clustering

K-means clustering

Example: spike sorting

- Unsupervised learning
  - The data should separate themselves



GLM (Encoding)

1. fMRI voxel activation
2. Gene expressions
3. Single neuron spike count
4. Pixel luminance patten

…

Data ← **External Variables**

1. Visual stimulus
2. Cell type
3. Attend left / right
4. Dog / cat

…