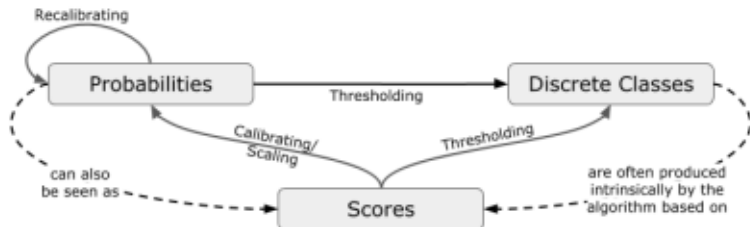


## PROBABILISTIC CLASSIFIERS / 2

- Both scoring and probabilistic classifiers can output classes by thresholding (binary case) / selecting the class with the maximum score (multiclass)
- Thresholding:  $h(\mathbf{x}) := [\pi(\mathbf{x}) \geq c]$  or  $h(\mathbf{x}) = [f(\mathbf{x}) \geq c]$  for some threshold  $c$ .
- Usually  $c = 0.5$  for probabilistic,  $c = 0$  for scoring classifiers.
- There are also versions of thresholding for the multiclass case



# DISCRIMINANT APPROACH

The **discriminant approach** tries to optimize the discriminant functions directly, usually via empirical risk minimization.

$$\hat{f} = \arg \min_{f \in \mathcal{H}} \mathcal{R}_{\text{emp}}(f) = \arg \min_{f \in \mathcal{H}} \sum_{i=1}^n L(y^{(i)}, f(\mathbf{x}^{(i)})).$$

**Examples:**

- Logistic regression (discriminant, linear)
- Neural networks
- Support vector machines

