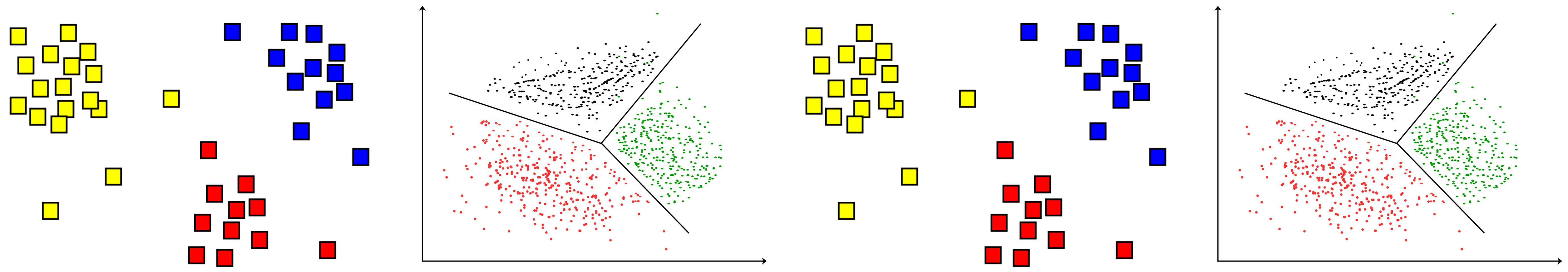
CPSC 425: Computer Vision



Lecture 27: Classification

Classification

Problem:

Assign new observations into one of a fixed set of categories (classes)

Key Idea(s):

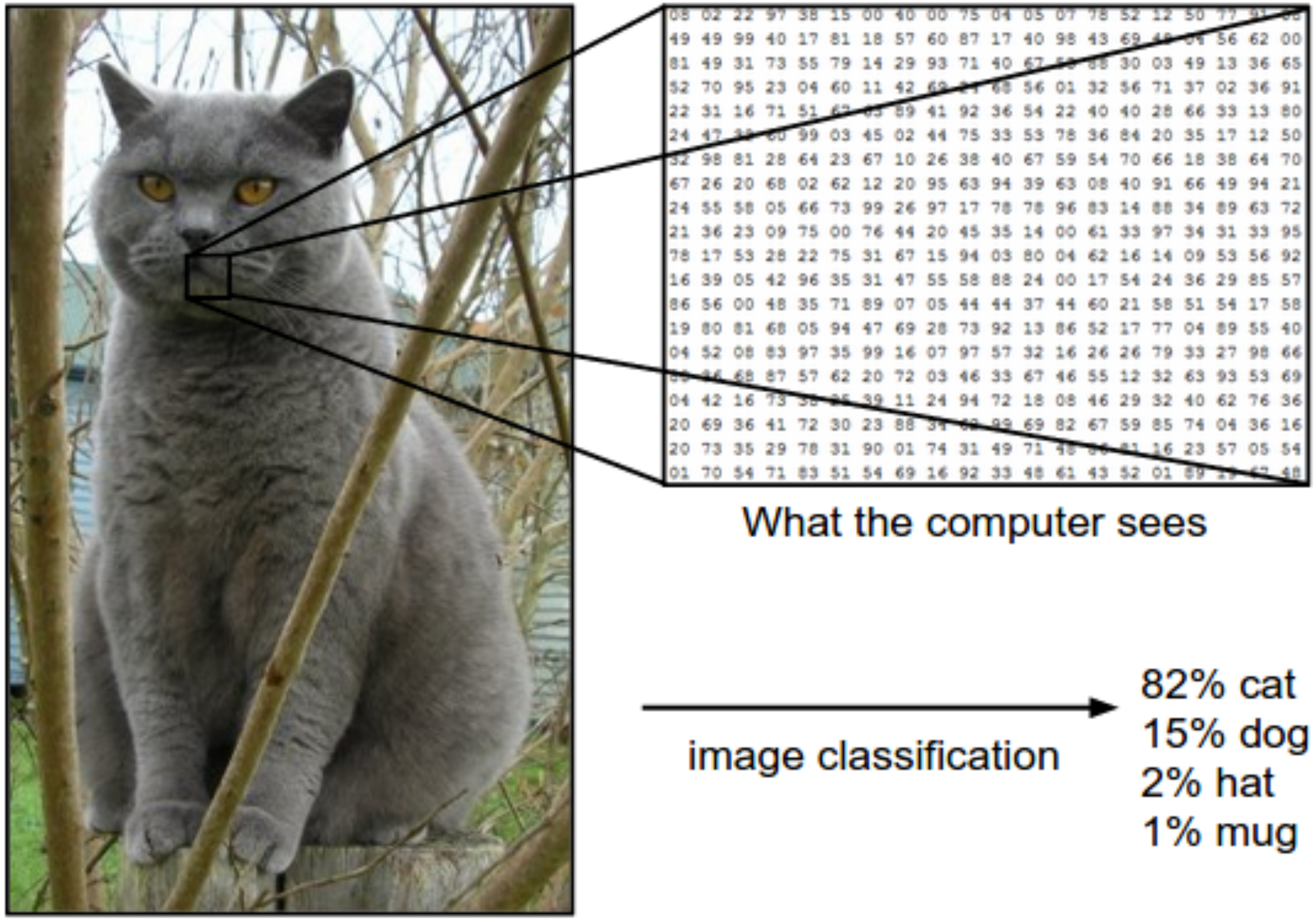
Build a model of data in a given category based on observations of instances in that category

Classification

(assume given set of discrete labels)
{dog, cat, truck, plane, ...}

→ cat

Classification



Classification

A **classifier** is a procedure that accepts as input a set of features and outputs a class **label**

Classifiers can be binary (face vs. not-face) or multi-class (cat, dog, horse, ...).

We build a classifier using a **training set** of labelled examples $\{(\mathbf{x}_i, y_i)\}$, where each $\mathbf{x}_i$ is a feature vector and each y_i is a class label.

Given a previously unseen observation, we use the classifier to predict its class label.

Classification

- Collect a database of images with labels
- Use ML to train an image classifier
- Evaluate the classifier on test images

Example training set

Label →

Feature vector computed from the image →

Example 1: A Classification Problem

Categorize images of fish
— “Atlantic salmon” vs “Pacific salmon”

Use **features** such as length, width, lightness, fin shape & number, mouth position, etc.

Given a previously unobserved image of a salmon, use the learned classifier to guess whether it is an Atlantic or Pacific salmon

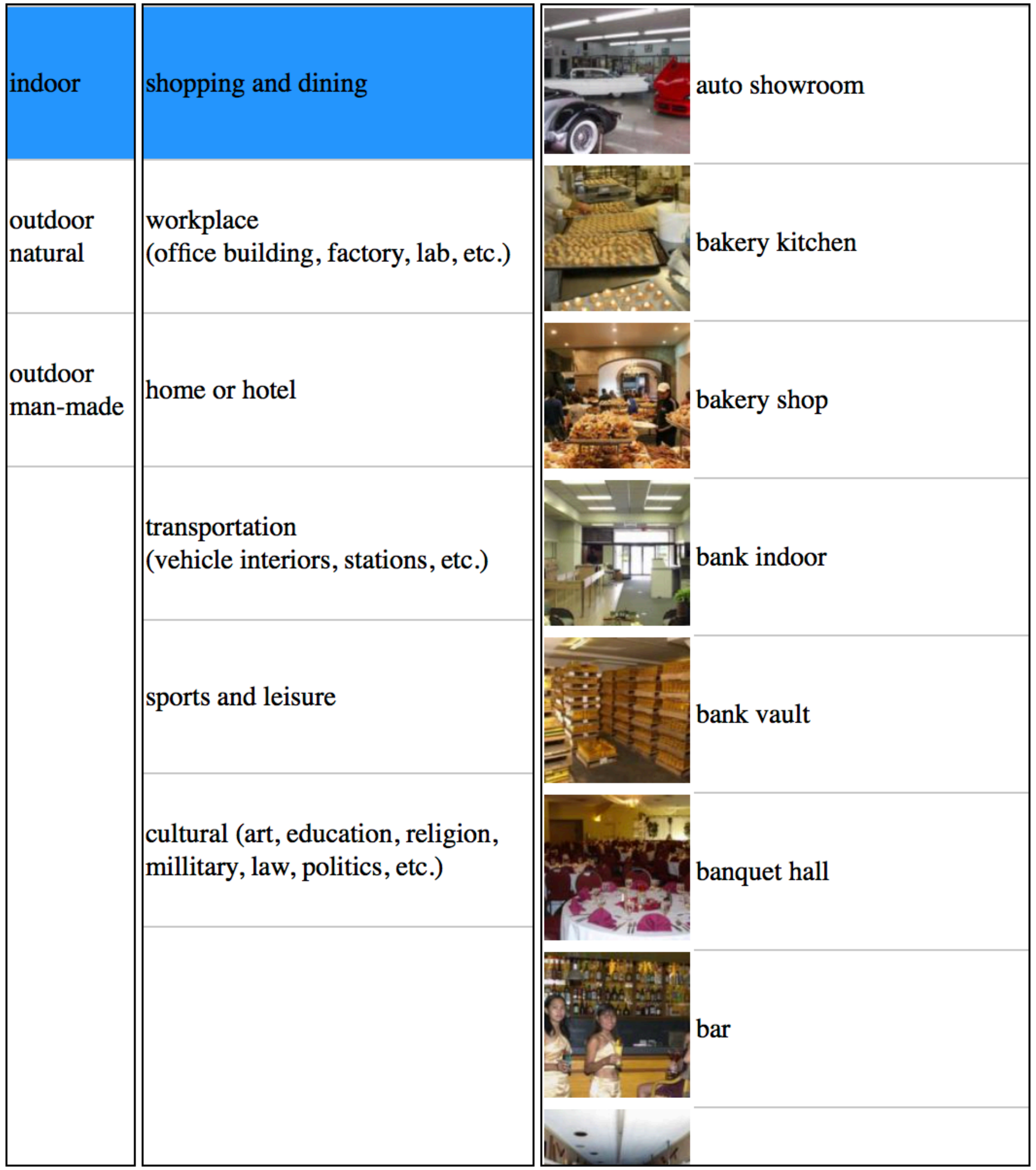
Figure credit: Duda & Hart

Example 2: Real Classification Problem

SUN Dataset

- 131K images

- 908 **scene** categories

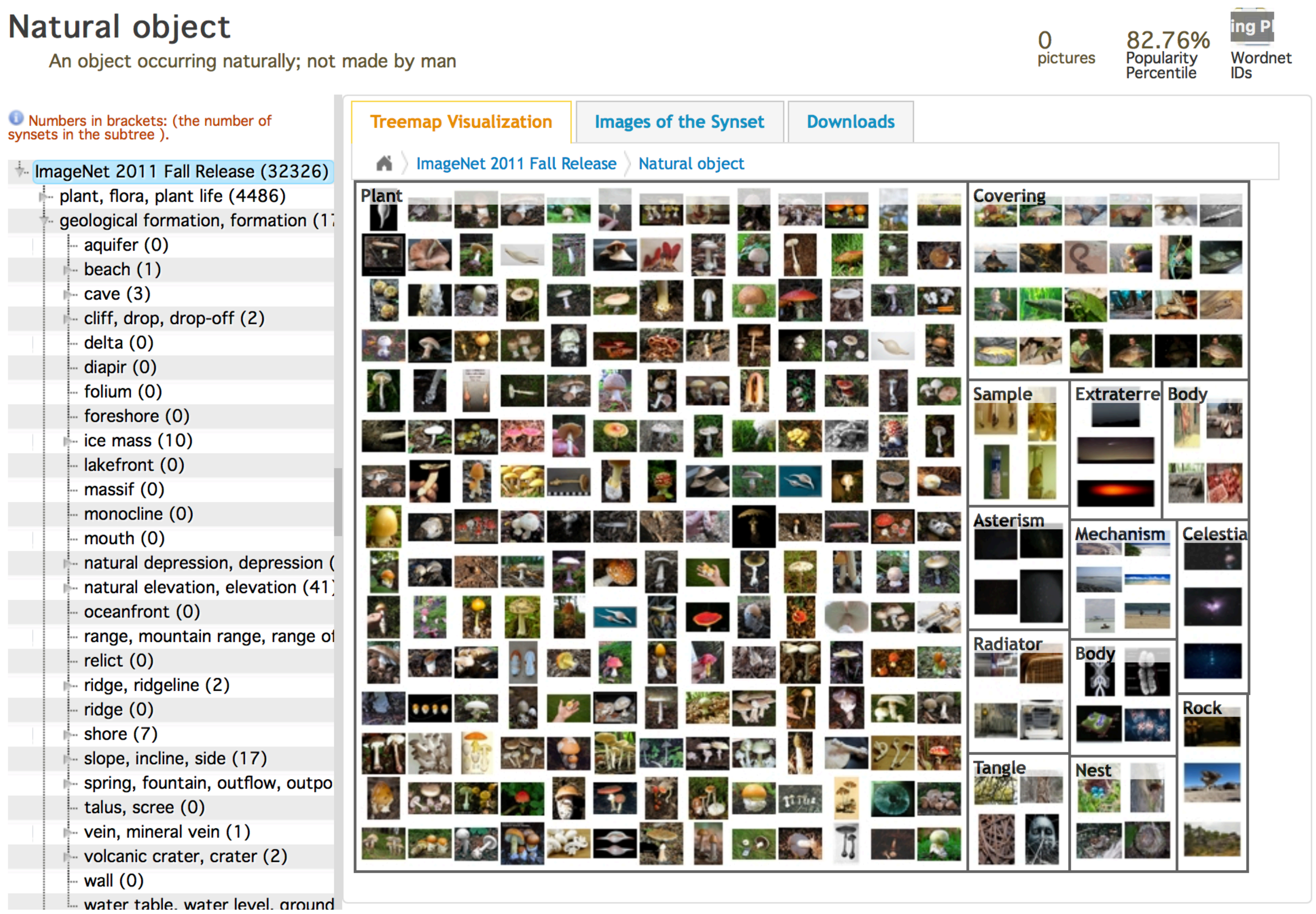


Example 3: Real Classification Problem

ImageNet Dataset

- 14 Million images

- 21K **object** categories



Bayes Rule (Review and Definitions)

Let c be the class label and let x be the measurement (i.e., evidence)

class–conditional probability
(a.k.a. likelihood)

prior probability

$$P(c|x) = \frac{P(x|c)p(c)}{P(x)}$$

posterior probability

unconditional probability
(a.k.a. marginal likelihood)

Bayes Rule (Review and Definitions)

Let c be the **class label** and let x be the **measurement** (i.e., evidence)

Simple case:

— binary classification; i.e., $c \in \{1, 2\}$

— features are 1D; i.e., $x \in \mathbb{R}$

$$P(c|x) = \frac{P(x|c)p(c)}{P(x)}$$

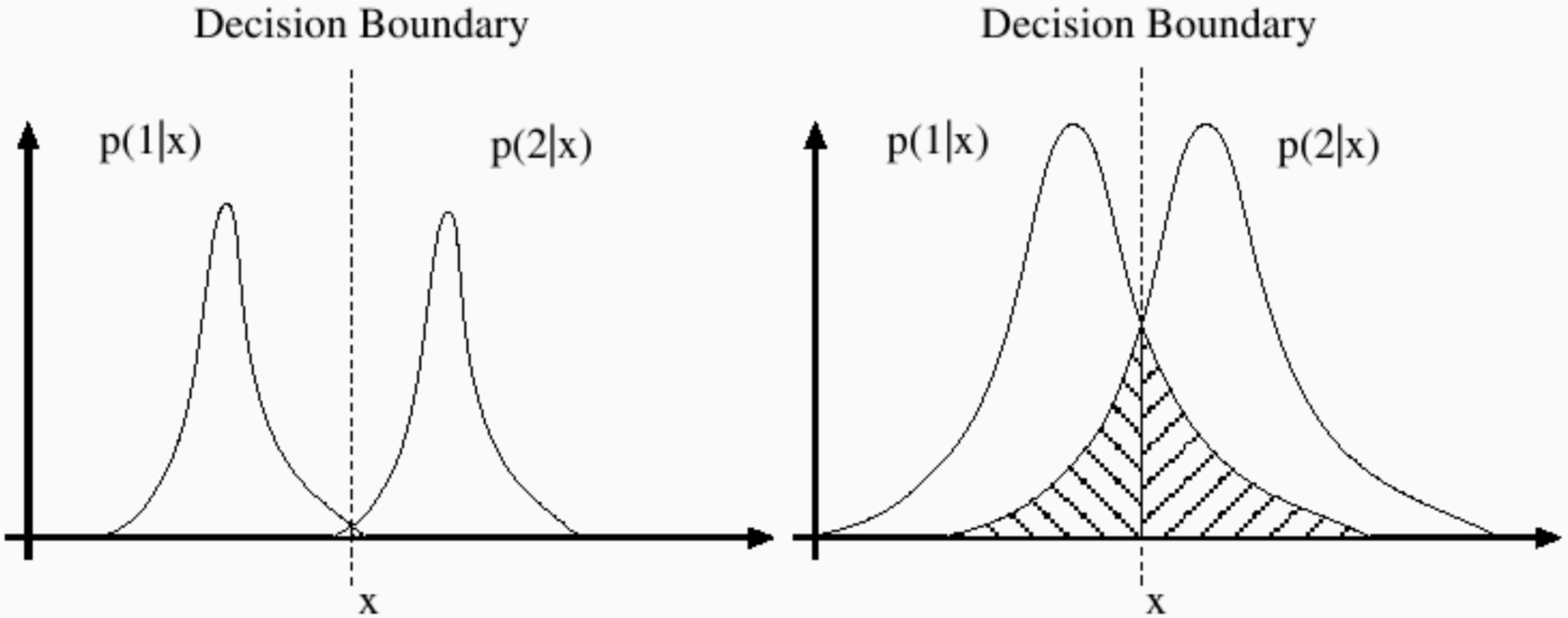
General case:

— multi-class; i.e., $c \in \{1, ..., 1000\}$

— features are high-dimensional; i.e., $x \in \mathbb{R}^{2,000+}$

Bayes’ Risk

Some errors may be inevitable: the minimum risk (shaded area) is called the **Bayes’ risk**



Forsyth & Ponce (2nd ed.) Figure 15.1