Digital Image Processing

Object recognition

(Nhận dạng đối tượng)

Thanh-Hai Tran

Electronics and Computer Engineering
School of Electronics and Telecommunications

Hanoi University of Science and Technology
1 Dai Co Viet - Hanoi - Vietnam

Outline

- **Pattern and pattern classes**
- **Recognition based on Decision-Theoretic Methods**
  - **Matching**
    - Minimum distance classifier
    - Matching by correlation
  - **Optimal statistical classifiers**
    - Bayes classifier for Gaussian pattern class
  - **Neural network**
    - Multilayer Perceptron
    - Deep Neural Network

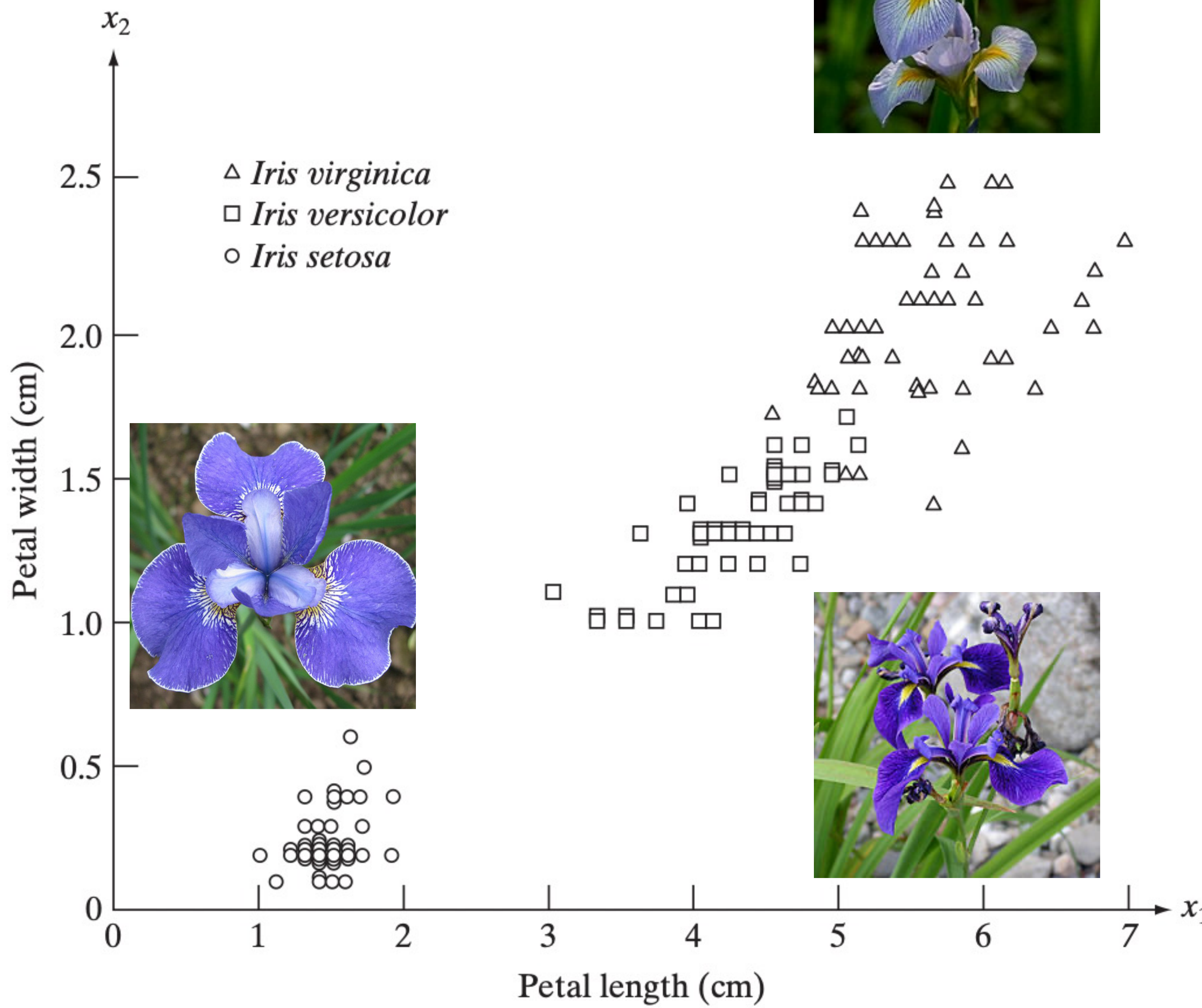
Pattern and Pattern classes

- A pattern is an arrangement of descriptors
- The name feature is used often in the pattern recognition literature to denote a descriptor
- A pattern class is a family of patterns that share some common properties
- Pattern classes are denoted $w_{1,}$ w_2 , $w_{n-1,}$ w_n
- Three common pattern arrangements:
  - Vector: (quantitative descriptor)
  - String and tree: (structural descriptor)

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Pattern and pattern classes

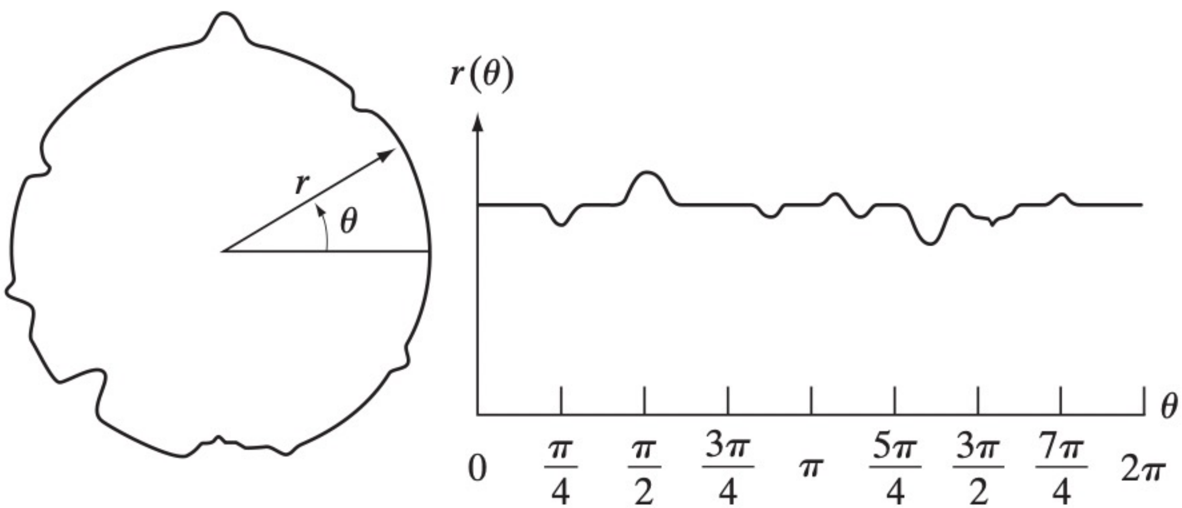
FIGURE 12.1
Three types of iris flowers described by two measurements.



Pattern and pattern classes

- **Object signature: describe each signature simply by its sampled amplitude values**

$$x_1 = r(\theta_1), x_2 = r(\theta_2), \ldots, x_n = r(\theta_n)$$

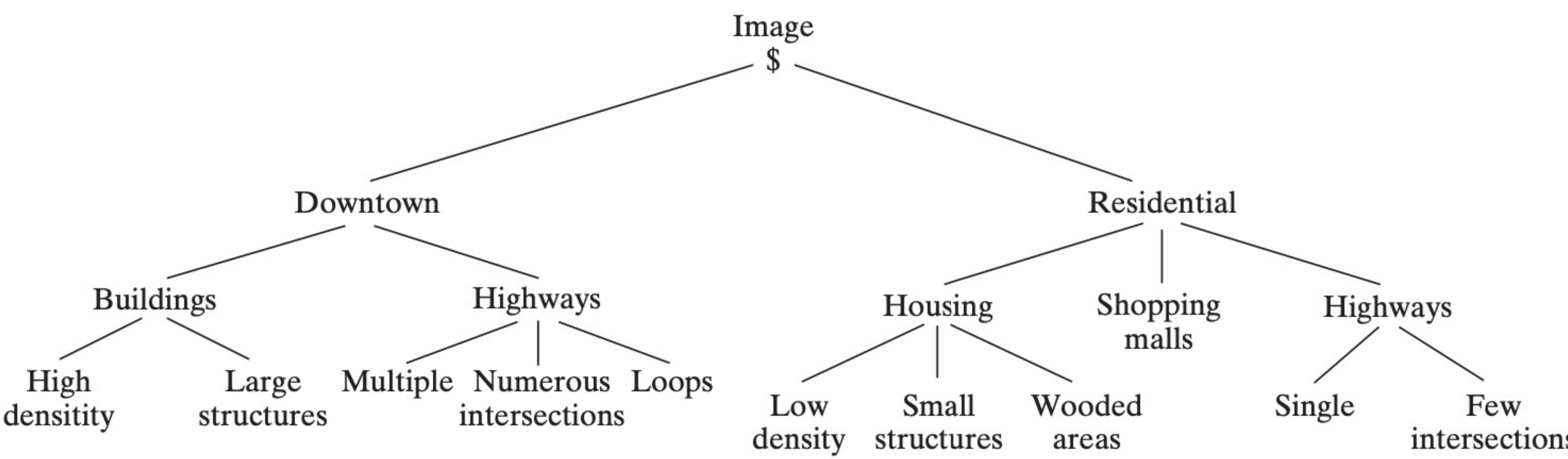


a b

FIGURE 12.2 A noisy object and its corresponding signature.

Pattern classes

- **Structural representation**
  - A tree description of the image



Recognition based on Decision theoretic methods

- Decision-theoretic approaches to recognition are based on the use of decision (or discriminant) functions.
- Let $x = (x_1, x_2, \ldots, x_n)^T$ represent an *n*-dimensional pattern vector
- the basic problem in decision-theoretic pattern recognition is to find W decision functions $d_1(\mathbf{x}), d_2(\mathbf{x}), \ldots, d_W(\mathbf{x})$
- If x belongs to w_i, then:

$$d_i(\mathbf{x}) > d_j(\mathbf{x}) \qquad j = 1, 2, \ldots, W; j \neq i$$

- The decision boundary separating class $w_{i \text{ from}}$ w_j

$$d_i(\mathbf{x}) - d_j(\mathbf{x}) = 0$$

Method 1: Matching

- Recognition techniques based on matching represent each class by a prototype pattern vector
- An unknown pattern is assigned to the class to which it is closest in terms of a predefined metric
- The simplest approach is the minimum distance classifier
- **Solution 1: Minimum distance classifier**
  - the prototype of each pattern class to be the mean vector of the patterns of that class

$$\mathbf{m}_j = \frac{1}{N_j} \sum_{\mathbf{x} \in \omega_j} \mathbf{x}_j \qquad j = 1, 2, \ldots, W$$

  - Where N_j is the number of pattern in the class w_j

Method 1: Matching

- Using the Euclidean distance to determine closeness reduces the problem to computing the distance measures

$$D_j(\mathbf{x}) = \|\mathbf{x} - \mathbf{m}_j\| \qquad j = 1, 2, \ldots, W$$

- Euclidian distance:

$$\|\mathbf{a}\| = (\mathbf{a}^T\mathbf{a})^{1/2}$$

- Select the smallest distance <=> evaluating the function

$$d_j(\mathbf{x}) = \mathbf{x}^T\mathbf{m}_j - \frac{1}{2}\mathbf{m}_j^T\mathbf{m}_j \qquad j = 1, 2, \ldots, W$$

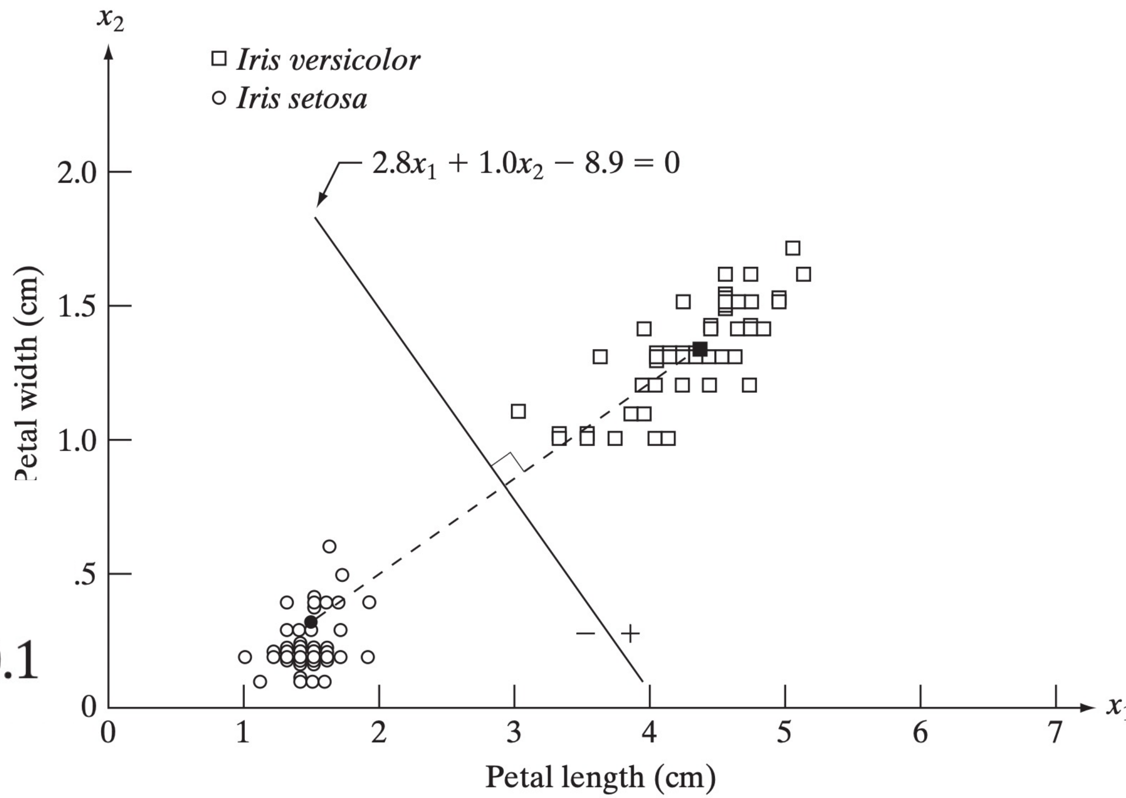
- Assign x to the class w_i w if $d_j(x)$ yields the maximal value.

Method 1: Matching

- **Two iris classes**
- **Means**
  - m1 ={4.3, 1.3}
  - m2 ={1.5, 0.3}
- **Decision functions**

$$d_1(\mathbf{x}) = \mathbf{x}^T\mathbf{m}_1 - \frac{1}{2}\mathbf{m}_1^T\mathbf{m}_1$$
$$= 4.3x_1 + 1.3x_2 - 10.1$$

$$d_2(\mathbf{x}) = \mathbf{x}^T\mathbf{m}_2 - \frac{1}{2}\mathbf{m}_2^T\mathbf{m}_2$$
$$= 1.5x_1 + 0.3x_2 - 1.17$$



- **Boundary equation:**

$$d_{12}(\mathbf{x}) = d_1(\mathbf{x}) - d_2(\mathbf{x})$$
$$= 2.8x_1 + 1.0x_2 - 8.9 = 0$$

Method 1: Matching

- **Solution 2: Matching by correlation**
- Convolution <=> Product in Frequency domain

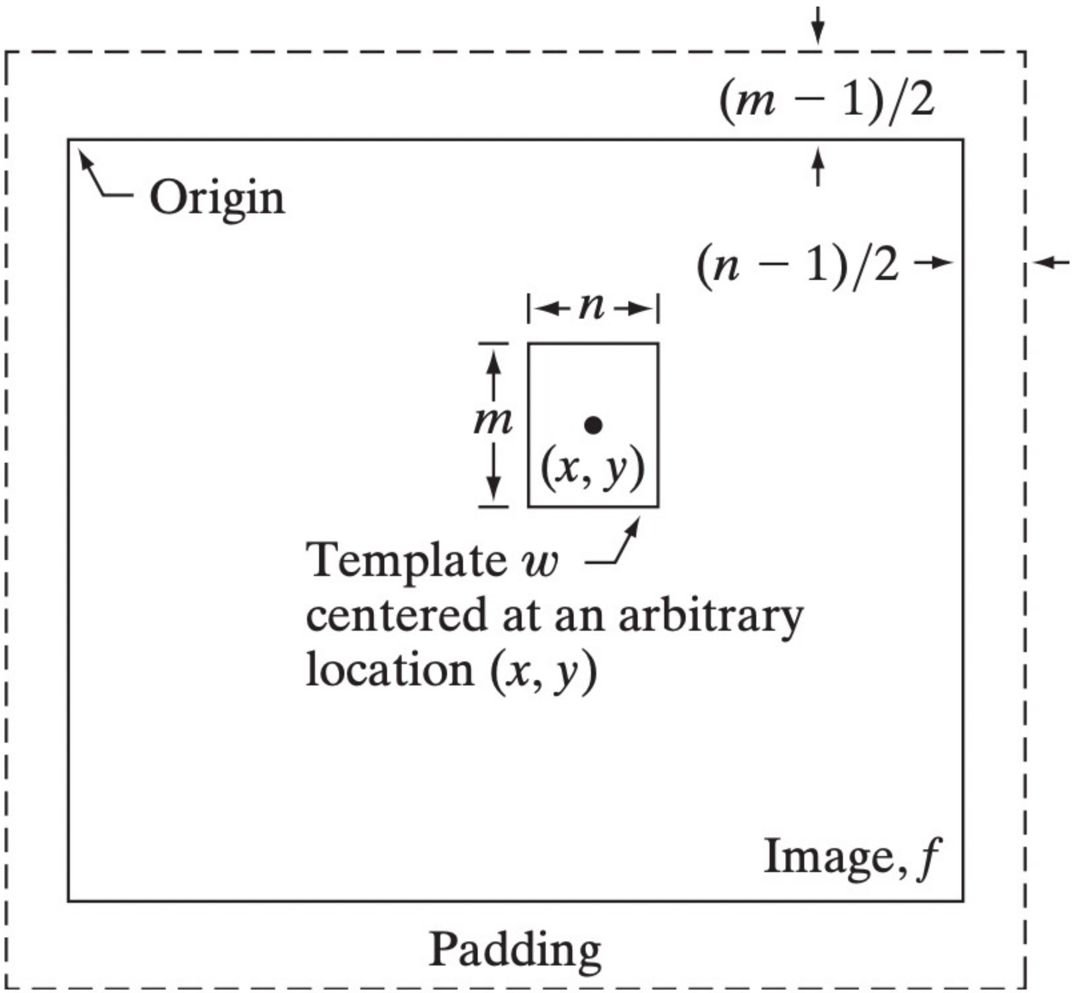
$$f(x, y) \star w(x, y) \Leftrightarrow F^*(u, v)W(u, v)$$

- Normalized correlation coefficient

$$\gamma(x, y) = \frac{\sum_s \sum_t \left[w(s, t) - \overline{w}\right] \sum_s \sum_t \left[f(x+s, y+t) - \overline{f}(x+s, y+t)\right]}{\left\{\sum_s \sum_t \left[w(s, t) - \overline{w}\right]^2 \sum_s \sum_t \left[f(x+s, y+t) - \overline{f}(x+s, y+t)\right]^2\right\}^{\frac{1}{2}}}$$

- When w is a template, then the method becomes template matching

- **Template matching**



Template matching

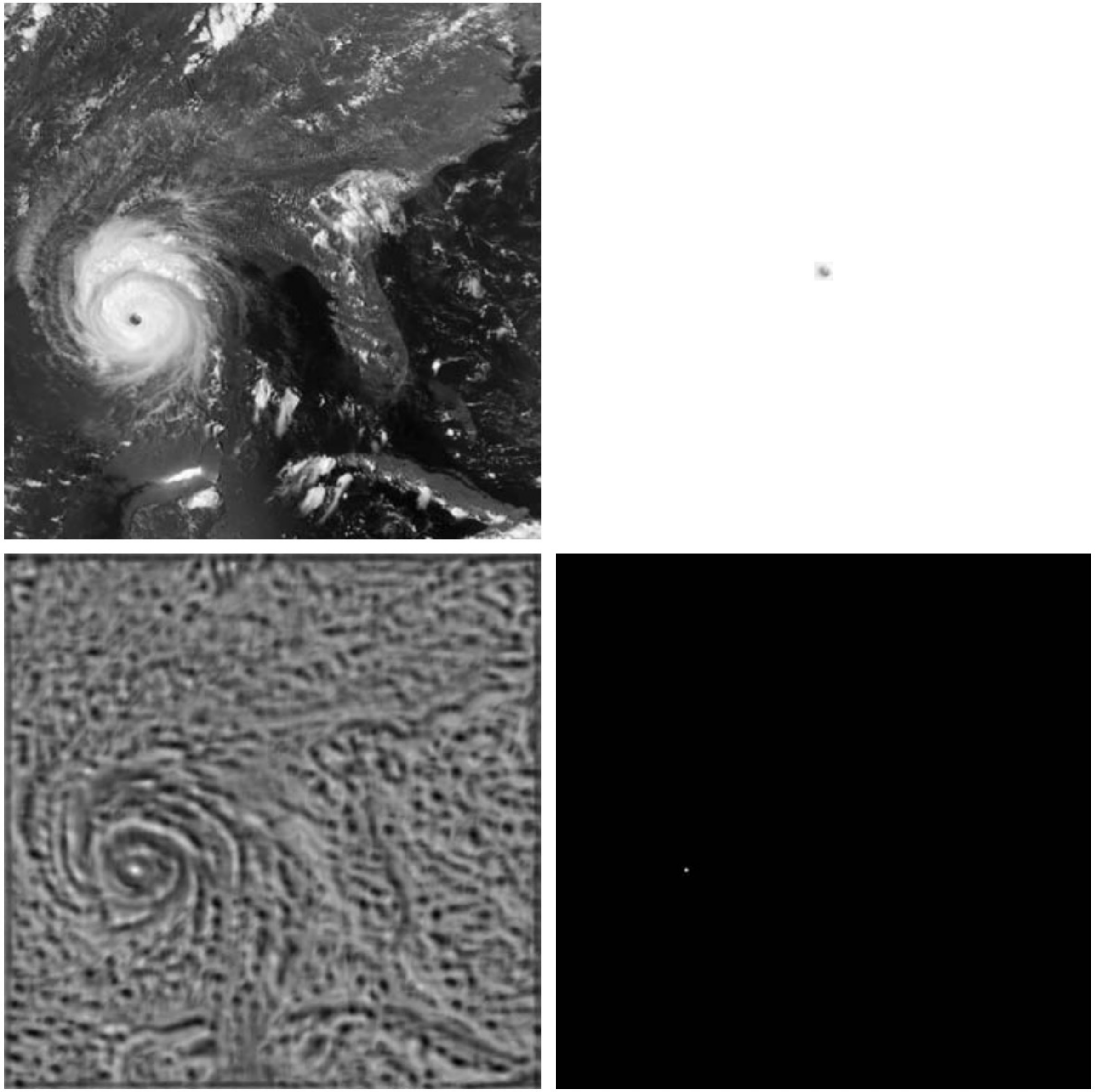


FIGURE 12.9 (a) Satellite image of Hurricane Andrew, taken on August 24, 1992. (b) Template of the eye of the storm. (c) Correlation coefficient shown as an image (note the brightest point). (d) Location of the best match. This point is a single pixel, but its size was enlarged to make it easier to see. (Original image courtesy of NOAA.)

Method 2:Optimum Statistical Classifiers

- The probability that a particular pattern x comes from class w_i is denoted i : $p(w_i/x)$
- If the pattern classifier decides that x came from w_j when it actually came from w_i, it incurs a loss, denoted L_{ij}
- the average loss incurred in assigning x to class w_j is

$$r_j(\mathbf{x}) = \sum_{k=1}^{W} L_{kj} p(\omega_k/\mathbf{x})$$

- From the basic probabilitistic , we have: p(A/B) = [p(A)p(B/A)]/p(B) then

$$r_j(\mathbf{x}) = \frac{1}{p(\mathbf{x})} \sum_{k=1}^{W} L_{kj} p(\mathbf{x}/\omega_k) P(\omega_k)$$

- We can drop p(x) then:

$$r_j(\mathbf{x}) = \sum_{k=1}^{W} L_{kj} p(\mathbf{x}/\omega_k) P(\omega_k)$$

Method 2: Optimum Statistical Classifiers

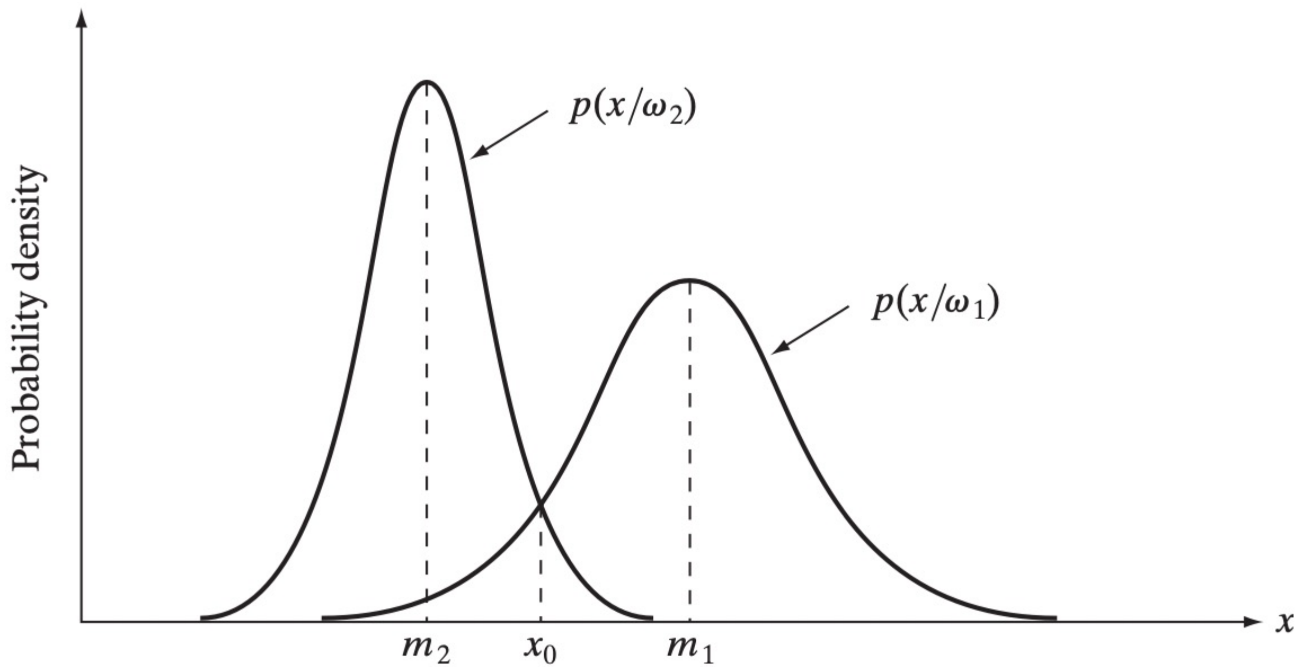
- The classifier has W possible classes to choose from for any given unknown pattern
- If it computes $r_1(x), r_2(x), \dots, r_W(x)$ for each pattern x and assigns the pattern to the class with the smallest loss, the total average loss with respect to all decisions will be minimum.
- The classifier that minimizes the total average loss is called the Bayes classifier
- X is assigned to w_i when

$$\sum_{k=1}^{W} L_{ki}\, p(\mathbf{x}/\omega_k) P(\omega_k) < \sum_{q=1}^{W} L_{qj}\, p(\mathbf{x}/\omega_q) P(\omega_q)$$

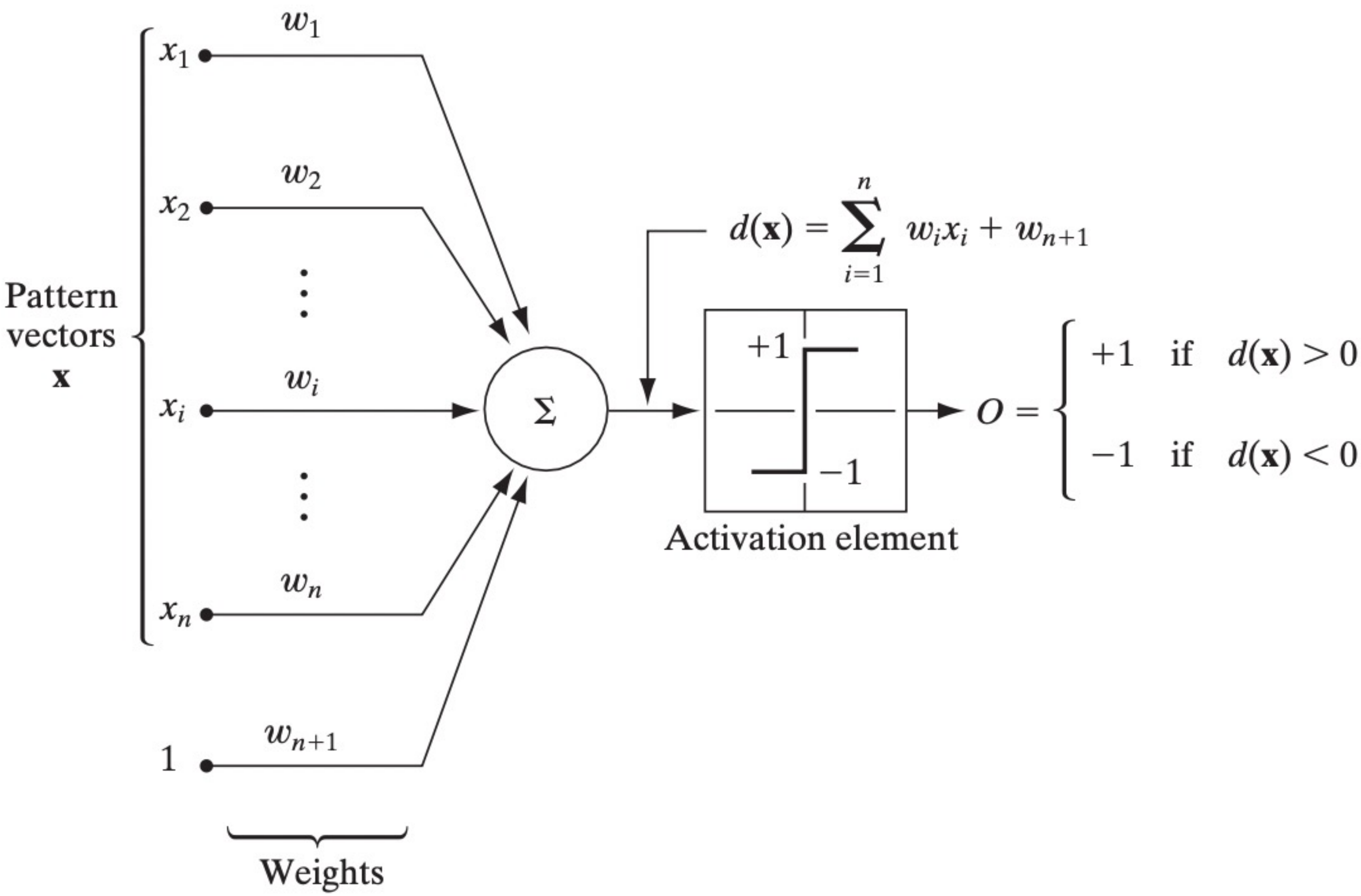
- For all j so that j <> i

Bayes classifier for Gaussian pattern classes

- 1-D problem
- Two pattern classes: W = 2 gonverned by 2 Gaussian Density
  - m1 and m2 are means
  - sigma1, sigma2 are variances



Neural network



Neural network

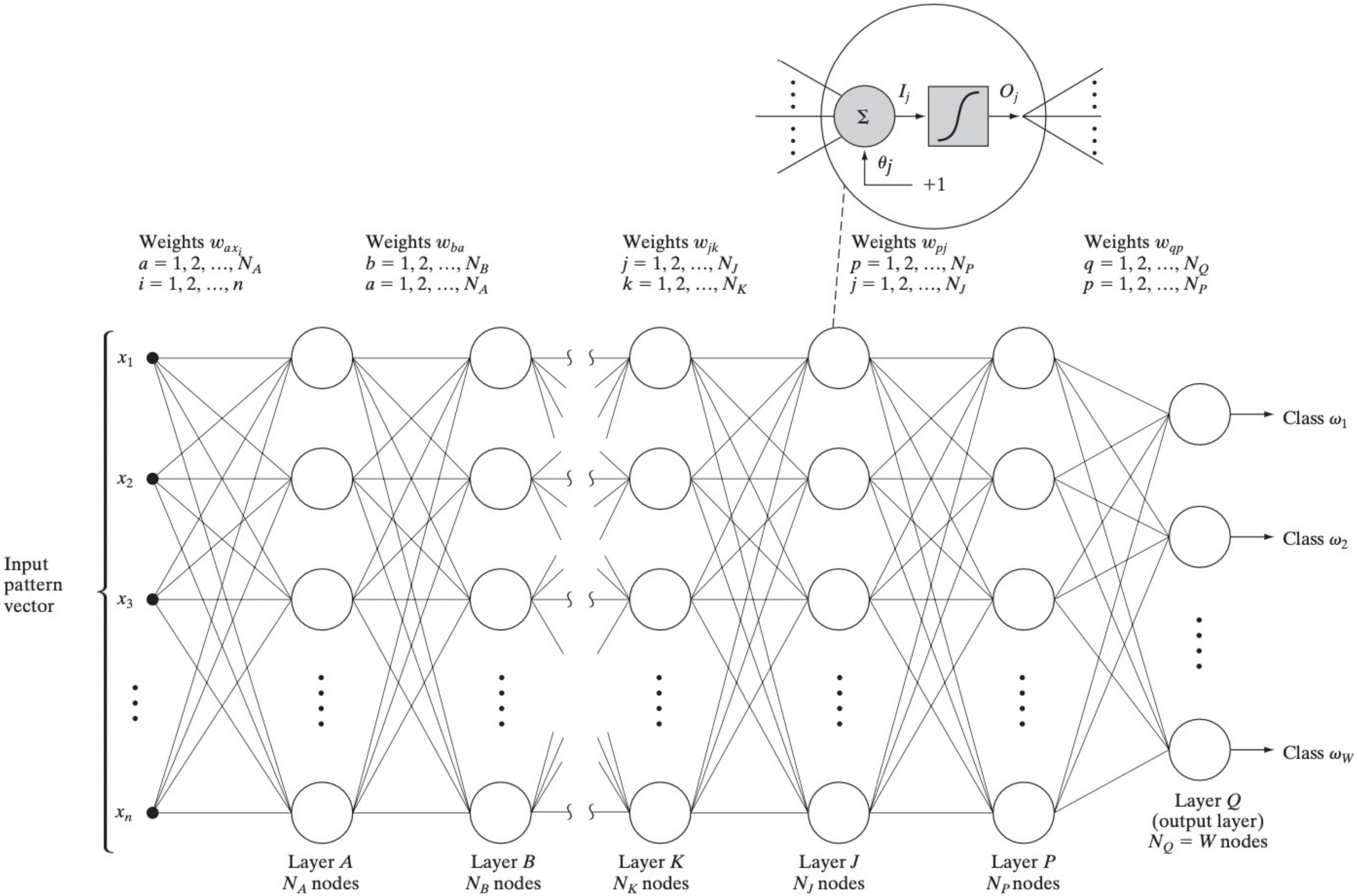
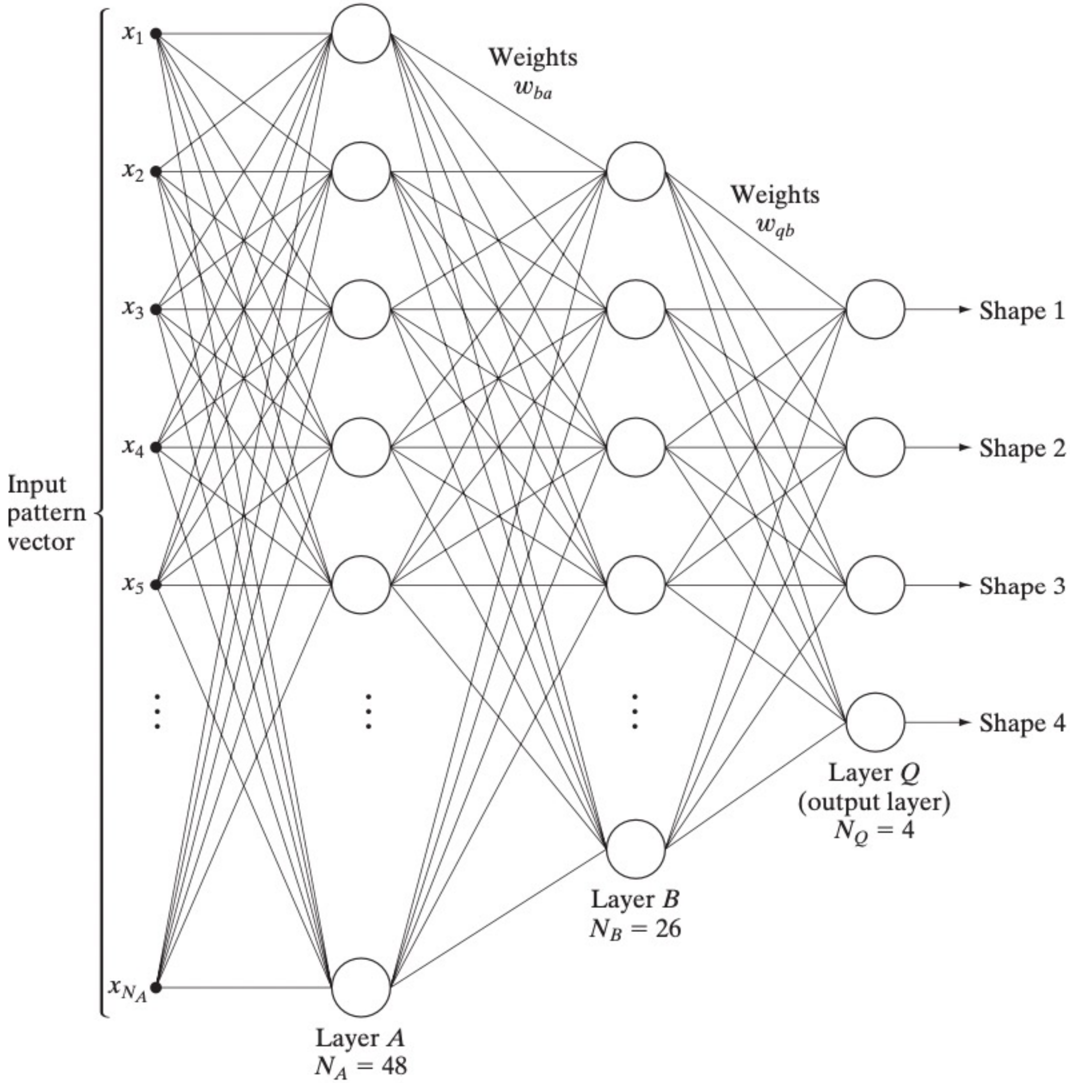
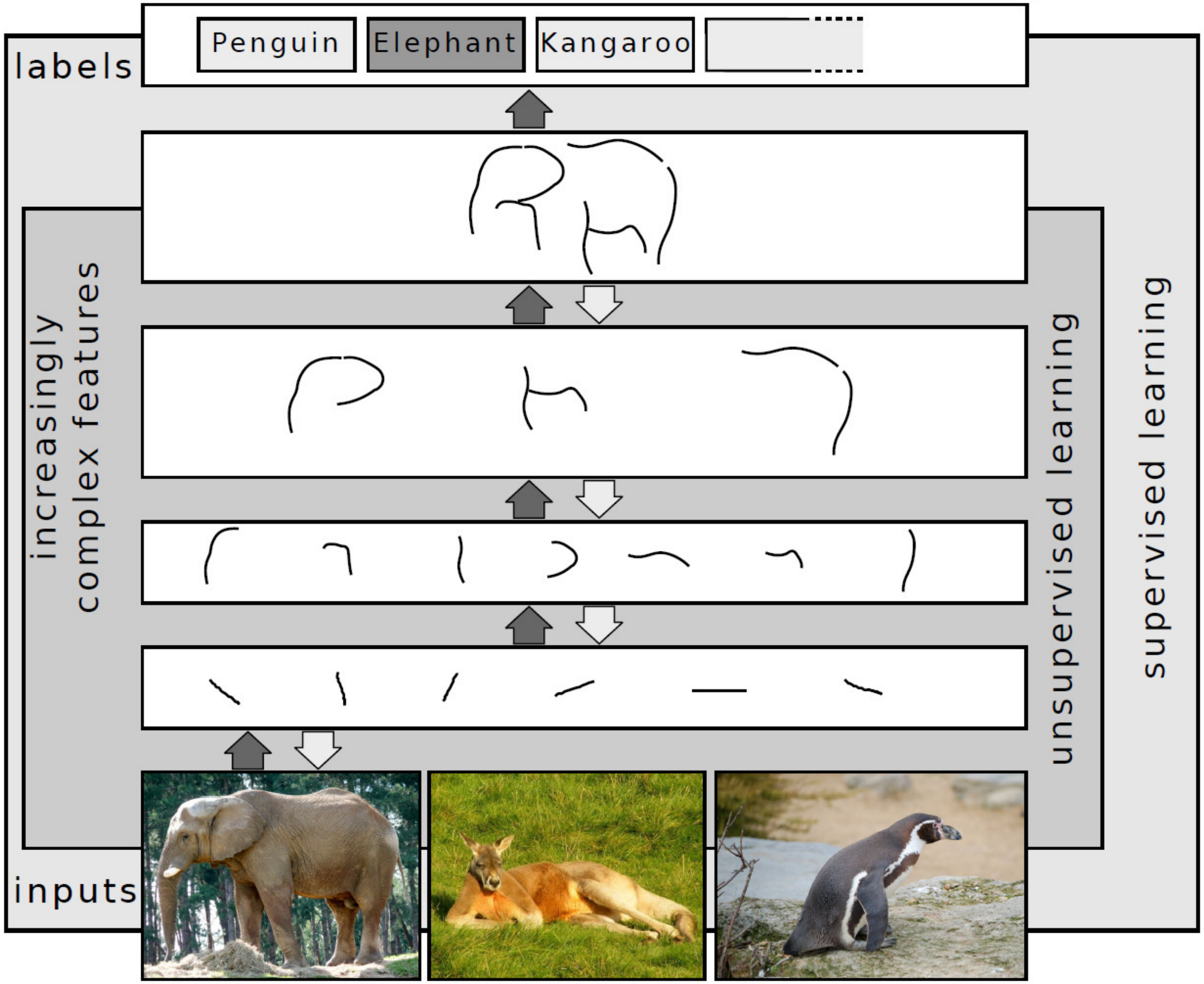


FIGURE 12.16 Multilayer feedforward neural network model. The blowup shows the basic structure of each neuron element throughout the network. The offset, θ_j, is treated as just another weight.

Neural network – an example



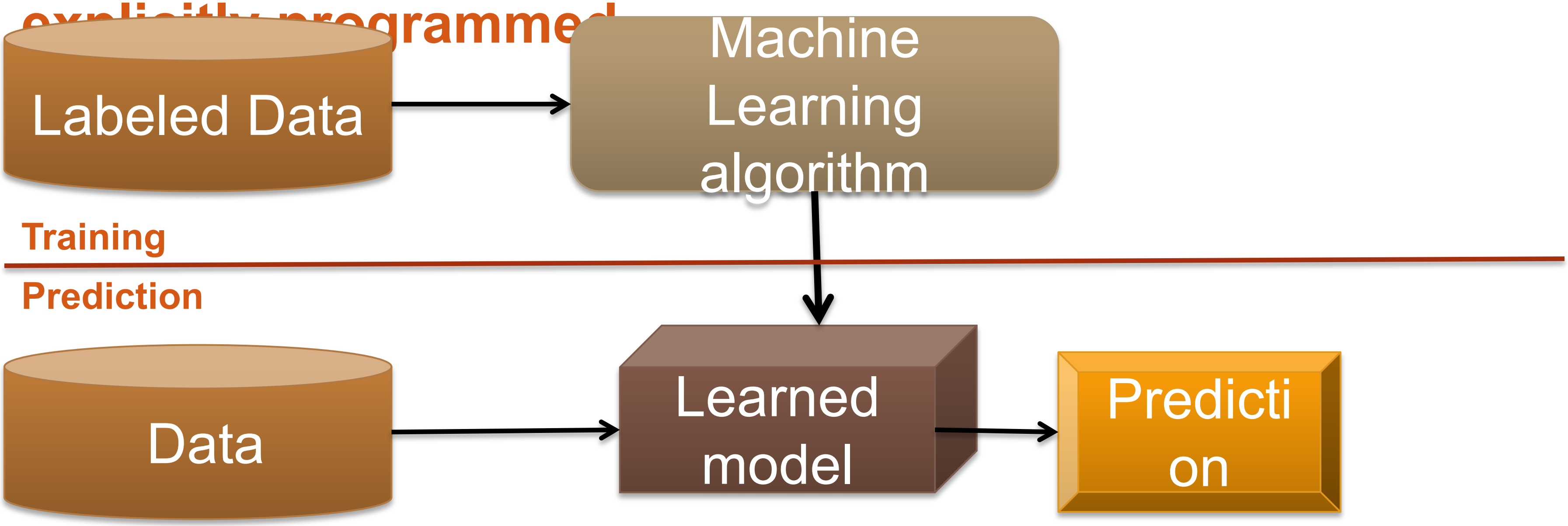
Object recognition



https://en.wikipedia.org/wiki/Deep_learning#/media/File:Deep_Learning.jpg

Machine Learning Basics

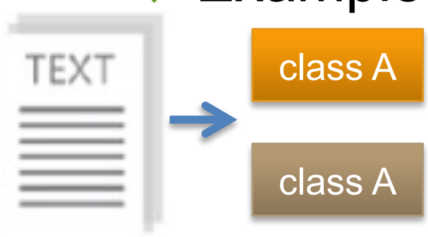
Machine learning is a field of computer science that gives computers the ability to **learn without being explicitly programmed**



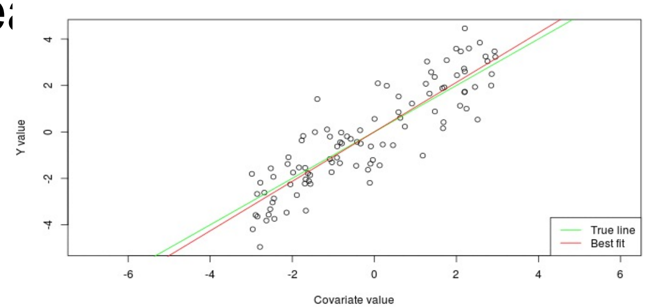
Methods that can learn from and make predictions on data

Type of learning

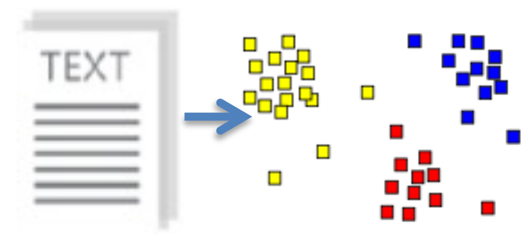
- **Supervised: Learning with a labeled training set**
  - Example: email *classification* with already labeled emails
- **Unsupervised: Discover patterns in unlabeled data**
  - Example: *cluster* similar documents based on text
- **Reinforcement learning: learn to act based on feedback/reward**
  - Example: le *in or lose*



Classification



Regression



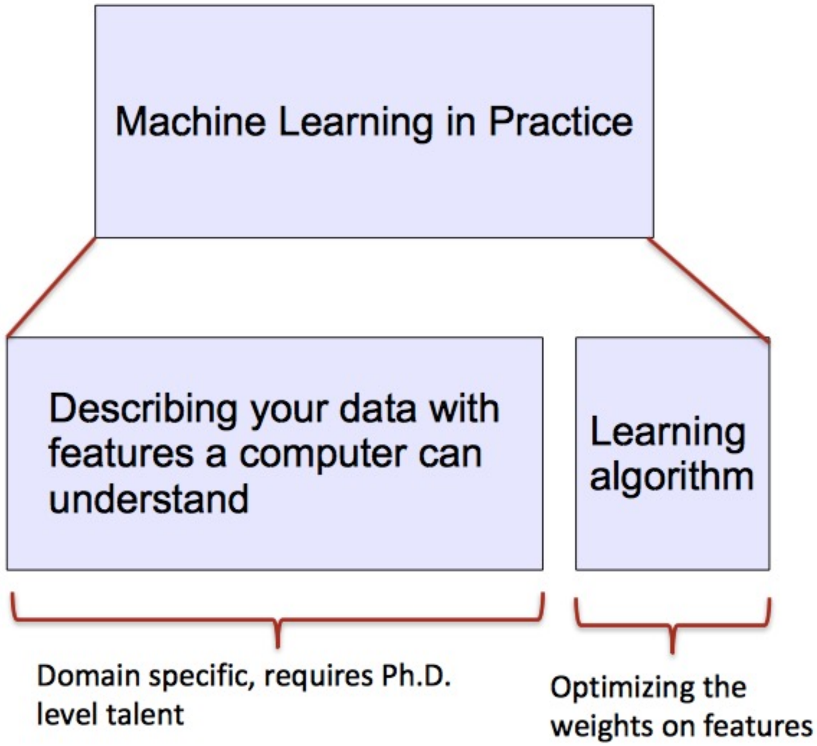
Clustering

Anomaly Detection
Sequence labeling
…

http://mbjoseph.github.io/2013/11/27/measure.html

ML vs. Deep Learning

- Most machine learning methods work well because of human-designed representations and input features
- ML becomes just optimizing weights to best make a final prediction

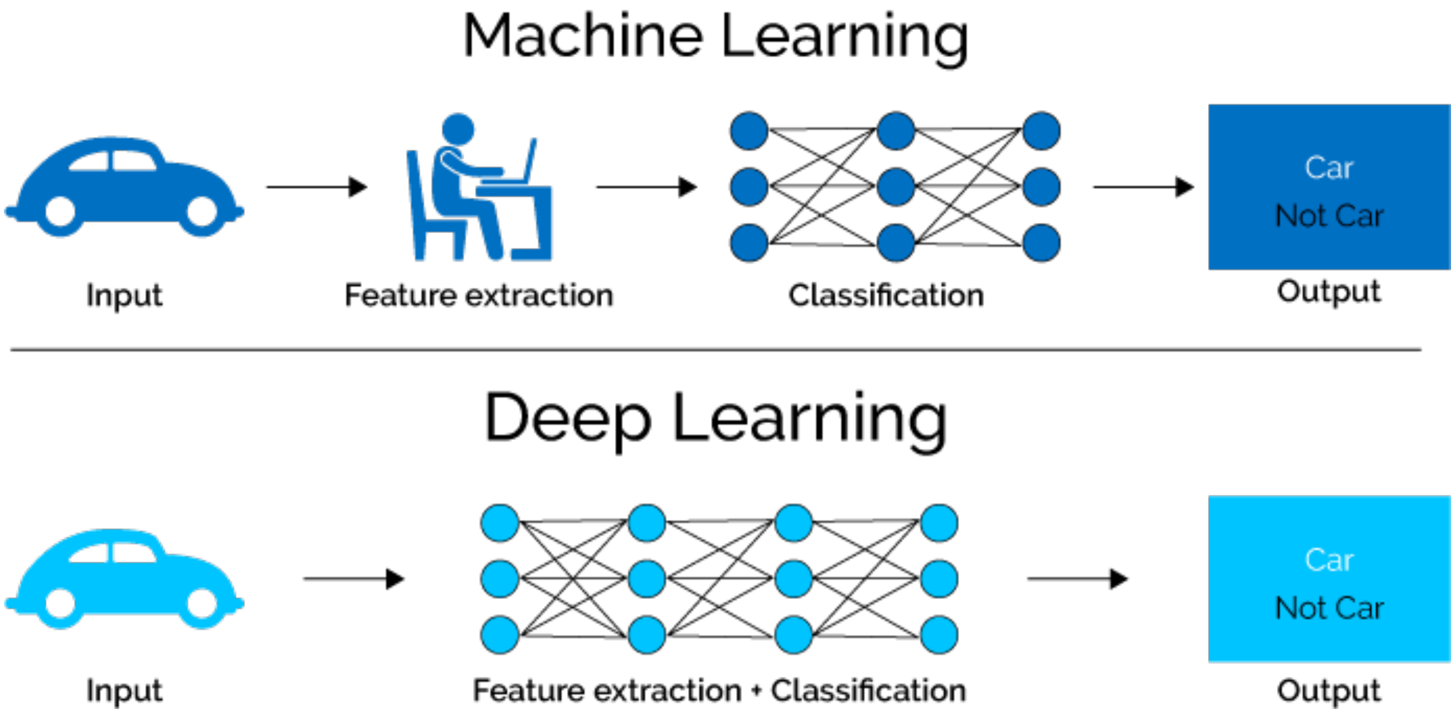


Feature	NER
Current Word	✓
Previous Word	✓
Next Word	✓
Current Word Character n-gram	all
Current POS Tag	✓
Surrounding POS Tag Sequence	✓
Current Word Shape	✓
Surrounding Word Shape Sequence	✓
Presence of Word in Left Window	size 4
Presence of Word in Right Window	size 4

Deep learning

- A machine learning subfield of learning **representations** of data. Exceptional effective at **learning patterns**.
- Deep learning algorithms attempt to learn (multiple levels of) representation by using a **hierarchy of multiple layers**
- If you provide the system **tons of information**, it begins to understand it and respond in useful ways.

What is Deep Learning?



https://www.xenonstack.com/blog/static/public/uploads/media/machine-learning-vs-deep-learning.png

To better understand Deep Learning, please register to the course "AI, DL" Computer Vision

References

- **Chapter 12,** Digital Image Processing, R.C. Gonzales and R. E. Woods
- Slide – Image Processing – Standford University
- **Courtesy of Hung-yi Lee**