

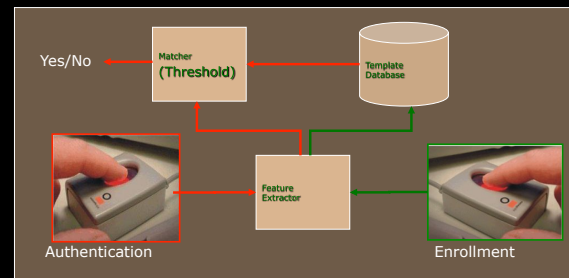


Pattern classification

Biometrics CSE 190 Lecture 3

CSE190a, Winter 2011

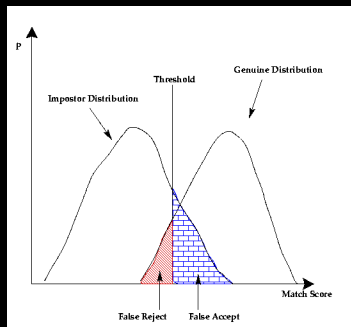
Biometrics: A Pattern Recognition System



- False accept rate (**FAR**): Proportion of imposters accepted
- False reject rate (**FRR**): Proportion of genuine users rejected
- Failure to enroll rate (**FTE**): portion of population that cannot be enrolled
- Failure to acquire rate (**FTA**): portion of population that cannot be verified

CSE190a, Winter 2011

Error Rates

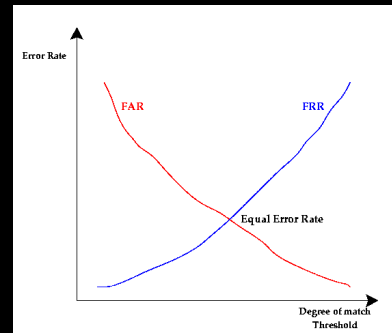


False Match (False Accept): Mistaking biometric measurements from two different persons to be from the same person; **False Non-match (False reject):** Mistaking two biometric measurements from the same person to be from two different persons

CSE190a, Winter 2011

(c) Jain 2004

Error vs Threshold

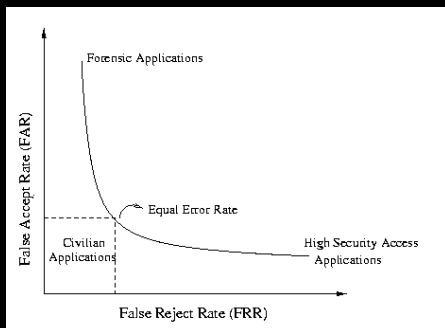


FAR: False accept rate
FRR: False reject rate

CSE190a, Winter 2011

(c) Jain 2004

ROC Curve



Accuracy requirements of a biometric system are application dependent

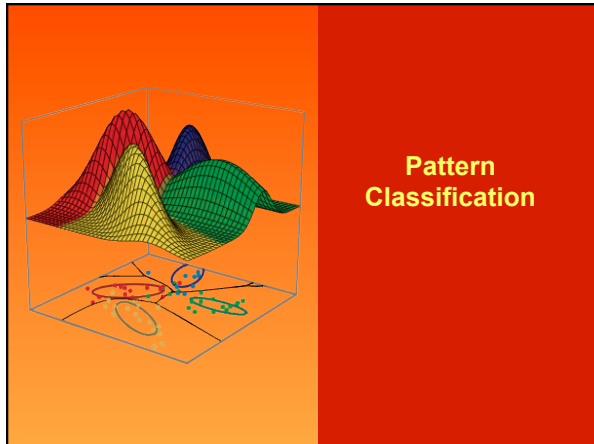
CSE190a, Winter 2011

(c) Jain 2004

Announcements

- Readings
- Project assignment on web page

CSE190a, Winter 2011

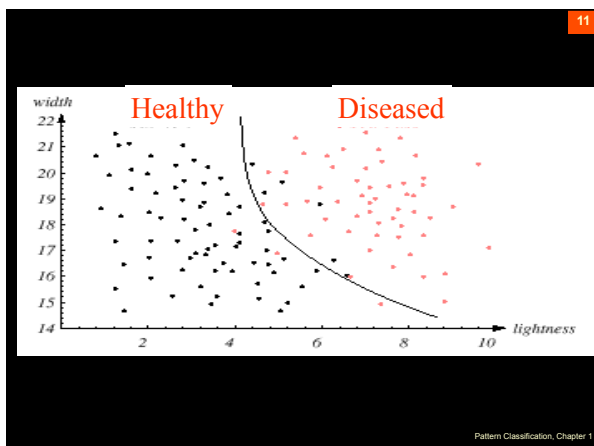
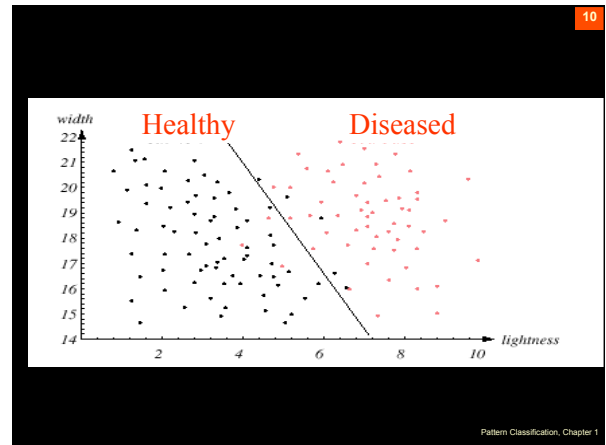
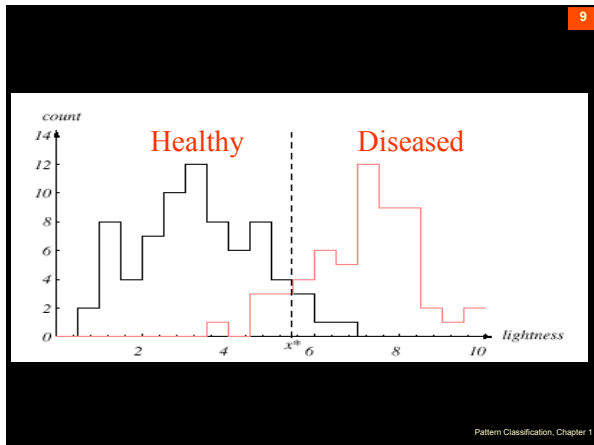


An Example

- “Sorting incoming Duck on a conveyor according to Healthy vs. diseased

Duck $\begin{cases} \text{Healthy} \\ \text{Diseased} \end{cases}$

Pattern Classification, Chapter 1



Bayesian Decision Theory

Continuous Features

(Sections 2.1-2.2)

Introduction

- The sea bass/salmon example
 - State of nature, prior
 - State of nature is a random variable
 - The catch of salmon and sea bass is equiprobable (The healthy of ducks probably isn't equiprobable).

• $P(\omega_1), P(\omega_2)$ Prior probabilities

• $P(\omega_1) = P(\omega_2)$ (uniform priors)

• $P(\omega_1) + P(\omega_2) = 1$ (exclusivity and exhaustivity)

Pattern Classification, Chapter 1

- Decision rule with only the prior information
 - Decide ω_1 if $P(\omega_1) > P(\omega_2)$ otherwise decide ω_2
- Use of the class-conditional information
 - $P(x | \omega_1)$ and $P(x | \omega_2)$ describe the difference in lightness between populations of healthy and diseased ducks

Pattern Classification, Chapter 1

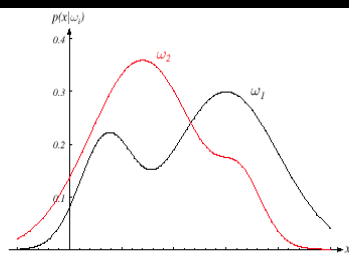


FIGURE 2.1. Hypothetical class-conditional probability density functions show the probability density of measuring a particular feature value x given the pattern is in category ω_i . If x represents the lightness of a fish, the two curves might describe the difference in lightness of populations of two types of fish. Density functions are normalized, and thus the area under each curve is 1.0. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Pattern Classification, Chapter 1

- Posterior, likelihood, evidence
 - $P(\omega_j | x) = (P(x | \omega_j) * P(\omega_j)) / P(x)$ (**BAYES RULE**)
 - In words, this can be said as:
Posterior = (Likelihood * Prior) / Evidence
 - Where in case of two categories

$$P(x) = \sum_{j=1}^{j=2} P(x | \omega_j) P(\omega_j)$$

Pattern Classification, Chapter 1

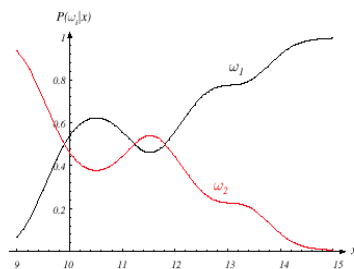
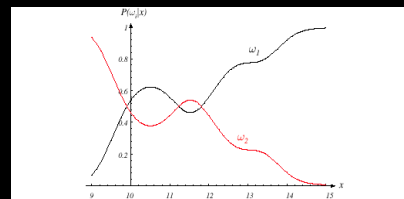


FIGURE 2.2. Posterior probabilities for the particular priors $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$ for the class-conditional probability densities shown in Fig. 2.1. Thus in this case, given that a pattern is measured to have feature value $x = 14$, the probability it is in category ω_2 is roughly 0.08, and that it is in ω_1 is 0.92. At every x , the posteriors sum to 1.0. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Pattern Classification, Chapter 1



- Intuitive decision rule given the posterior probabilities:
Given x :
if $P(\omega_1 | x) > P(\omega_2 | x)$ → True state of nature = ω_1
if $P(\omega_1 | x) < P(\omega_2 | x)$ → True state of nature = ω_2

Why do this?: Whenever we observe a particular x , the probability of error is :

$$P(\text{error} | x) = P(\omega_1 | x) \text{ if we decide } \omega_2$$

$$P(\text{error} | x) = P(\omega_2 | x) \text{ if we decide } \omega_1$$

Pattern Classification, Chapter 1

- Since decision rule is optimal for each feature value X , there is not better rule for all x .

Bayesian Decision Theory – Continuous Features

Generalization of the preceding ideas

- Use of more than one feature
- Use more than two states of nature
- Allowing actions and not only decide on the state of nature
- Introduce a loss of function (more general than the probability of error)
- Allowing actions other than classification primarily allows the possibility of rejection
- Refusing to make a decision in close or bad cases!
- Letting loss function state how costly each action taken is

Bayesian Decision Theory – Continuous Features

- Let $\mathbf{X}$ be a vector of features.
- Let $\{\omega_1, \omega_2, \dots, \omega_c\}$ be the set of c states of nature (or “classes”)
- Let $\{\alpha_1, \alpha_2, \dots, \alpha_a\}$ be the set of possible actions
- Let $\lambda(\alpha_i | \omega_j)$ be the loss for action α_i when the state of nature is ω_j

What is the Expected Loss for action α_i ?

For any given x the expected loss is

$$R(\alpha_i | x) = \sum_{j=1}^{j=c} \lambda(\alpha_i | \omega_j) P(\omega_j | x)$$

$R(\alpha_i | x)$ is called the Conditional Risk (or Expected Loss)

Overall risk

$$R = \text{Sum of all } \underbrace{R(\alpha_i | x)}_{\text{Conditional risk}} \text{ for } i = 1, \dots, a$$

Minimizing $R \iff$ Minimizing $R(\alpha_i | x)$ for $i = 1, \dots, a$

$$R(\alpha_i | x) = \sum_{j=1}^{j=c} \lambda(\alpha_i | \omega_j) P(\omega_j | x)$$

for $i = 1, \dots, a$

Given a measured feature vector x , which action should we take?

Select the action α_i for which $R(\alpha_i | x)$ is minimum

$\longrightarrow$ R is minimum and R in this case is called the Bayes risk = best performance that can be achieved!

Two-Category Classification

25

α_1 : deciding ω_1

α_2 : deciding ω_2

$\lambda_{ij} = \lambda(\alpha_i | \omega_j)$

loss incurred for deciding ω_i when the true state of nature is ω_j

Conditional risk:

$$R(\alpha_1 | x) = \lambda_{11}P(\omega_1 | x) + \lambda_{12}P(\omega_2 | x)$$

$$R(\alpha_2 | x) = \lambda_{21}P(\omega_1 | x) + \lambda_{22}P(\omega_2 | x)$$

Pattern Classification, Chapter 1

Our rule is the following:

if $R(\alpha_1 | x) < R(\alpha_2 | x)$

$$\lambda_{11}P(\omega_1 | x) + \lambda_{12}P(\omega_2 | x) < \lambda_{21}P(\omega_1 | x) + \lambda_{22}P(\omega_2 | x)$$

action α_i ; "decide ω_i " is taken

This results in the equivalent rule :

decide ω_1 if:

$$(\lambda_{21} - \lambda_{11}) P(x | \omega_1) P(\omega_1) > (\lambda_{12} - \lambda_{22}) P(x | \omega_2) P(\omega_2)$$

and decide ω_2 otherwise

Pattern Classification, Chapter 1

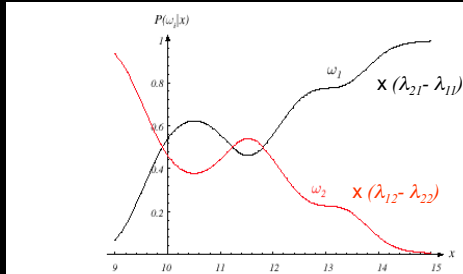


FIGURE 2.2. Posterior probabilities for the particular priors $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$ for the class-conditional probability densities shown in Fig. 2.1. Thus in this case, given that a pattern is measured to have feature value $x = 14$, the probability it is in category ω_2 is roughly 0.08, and that it is in ω_1 is 0.92. At every x , the posteriors sum to 1.0. From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.

Pattern Classification, Chapter 1