CHAPTER 3

Output: Knowledge Representation

Outline

- Tables
- Linear Models
- Trees
- Rules
- Instance-based Representation
- Clusters

Tables

Decision or regression table

Table 1.2 Weather Data

Outlook	Temperature	Humidity	Windy	Play
Sunny	hot	high	false	no
Sunny	hot	high	true	no
Overcast	hot	high	false	yes
Rainy	mild	high	false	yes
Rainy	cool	normal	false	yes
Rainy	cool	normal	true	no
Overcast	cool	normal	true	yes
Sunny	mild	high	false	no
Sunny	cool	normal	false	yes
Rainy	mild	normal	false	yes
Sunny	mild	normal	true	yes
Overcast	mild	high	true	yes
Overcast	hot	normal	false	yes
Rainy	mild	high	true	no

Classification learning

將test data 進行training
前四個attribute作為input
"Play"為output

Exemplar view

Linear Models (1/2)

PRP = 37.06 + 2.47*CACH

Numeric prediction

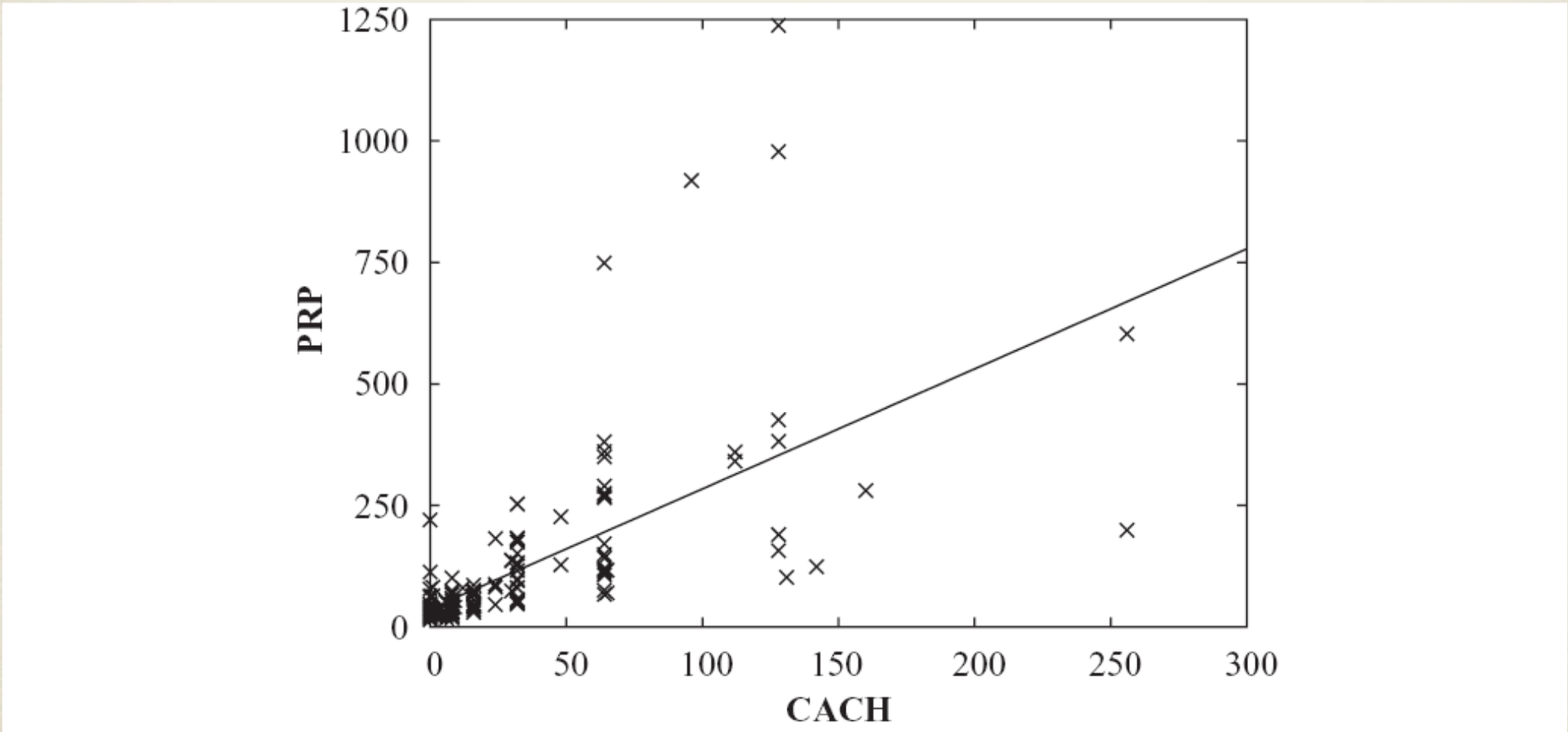


FIGURE 3.1

A linear regression function for the CPU performance data.

Linear Models (2/2)

Classification learning

2.0 - 0.5*PETAL-LENGTH - 0.8*PETAL-WIDTH

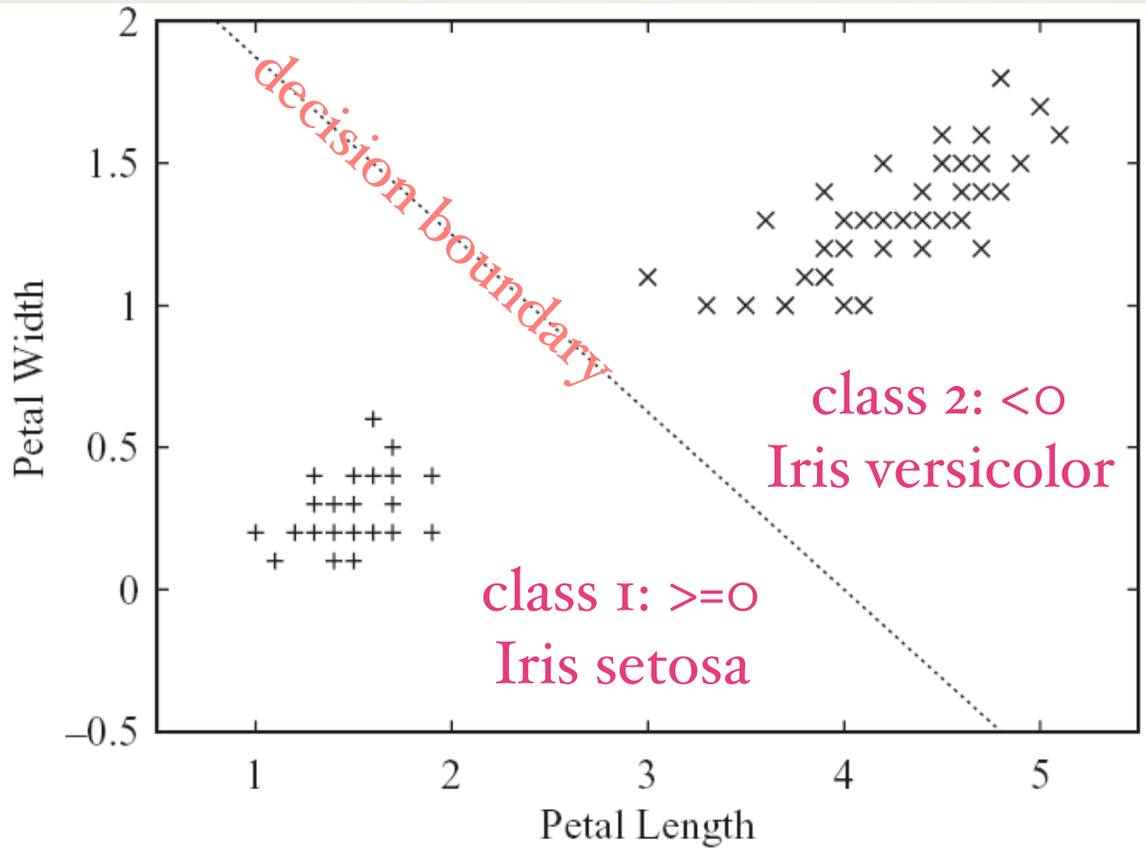


FIGURE 3.2

A linear decision boundary separating *Iris setosas* from *Iris versicolors*.

Tree (1/3)

Decision Tree

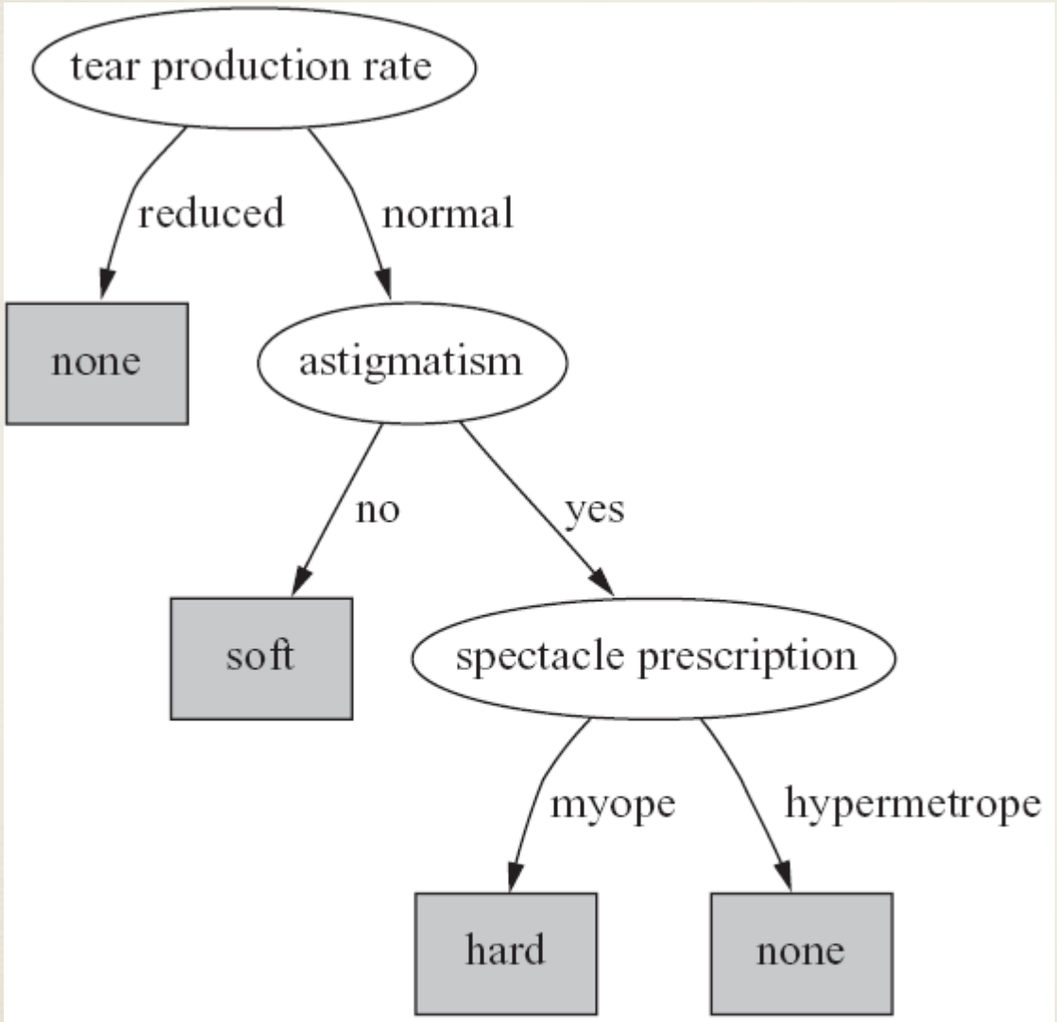


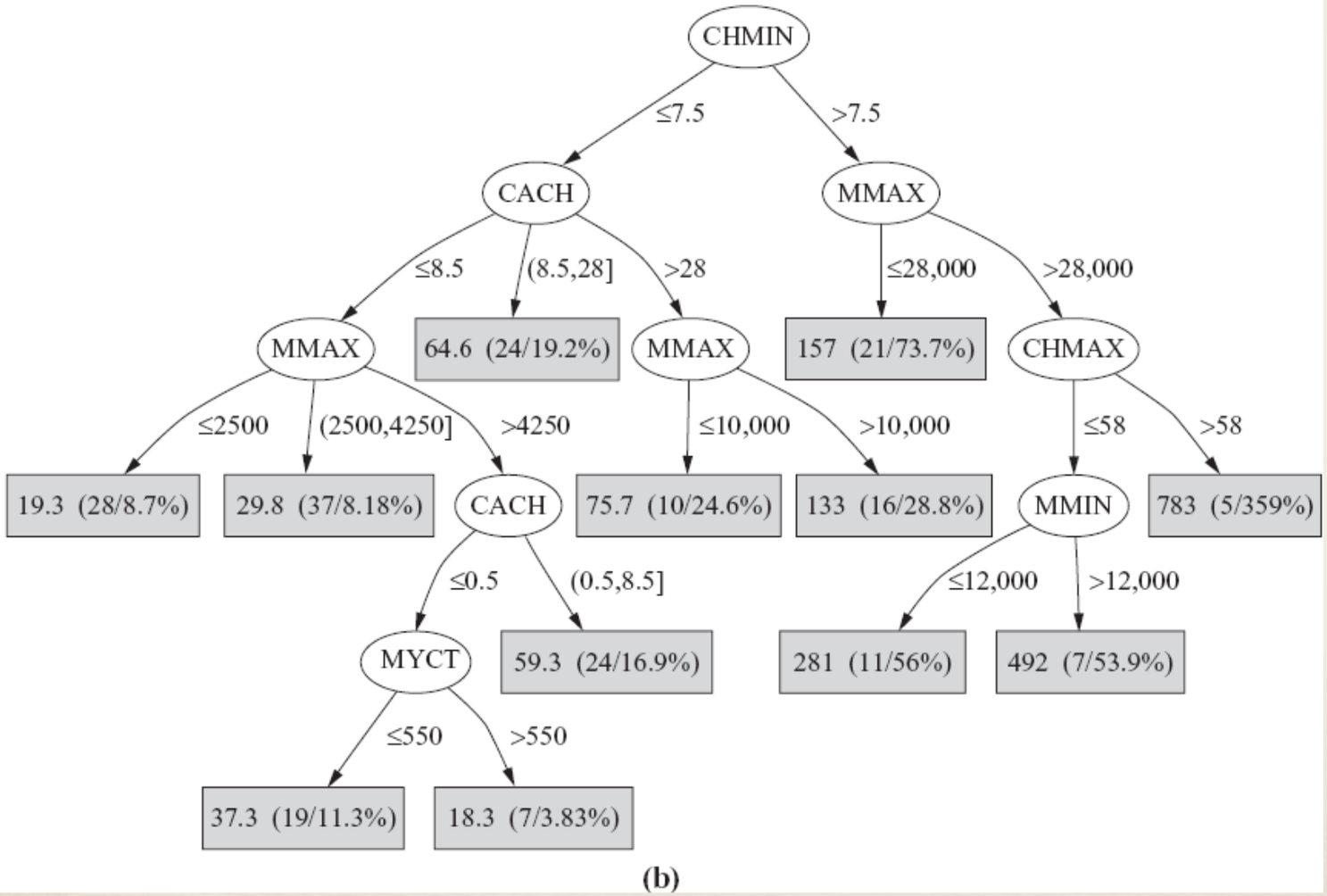
FIGURE 1.2

Decision tree for the contact lens data.

Classification learning

Tree (2/3)

Regression Tree



(b)

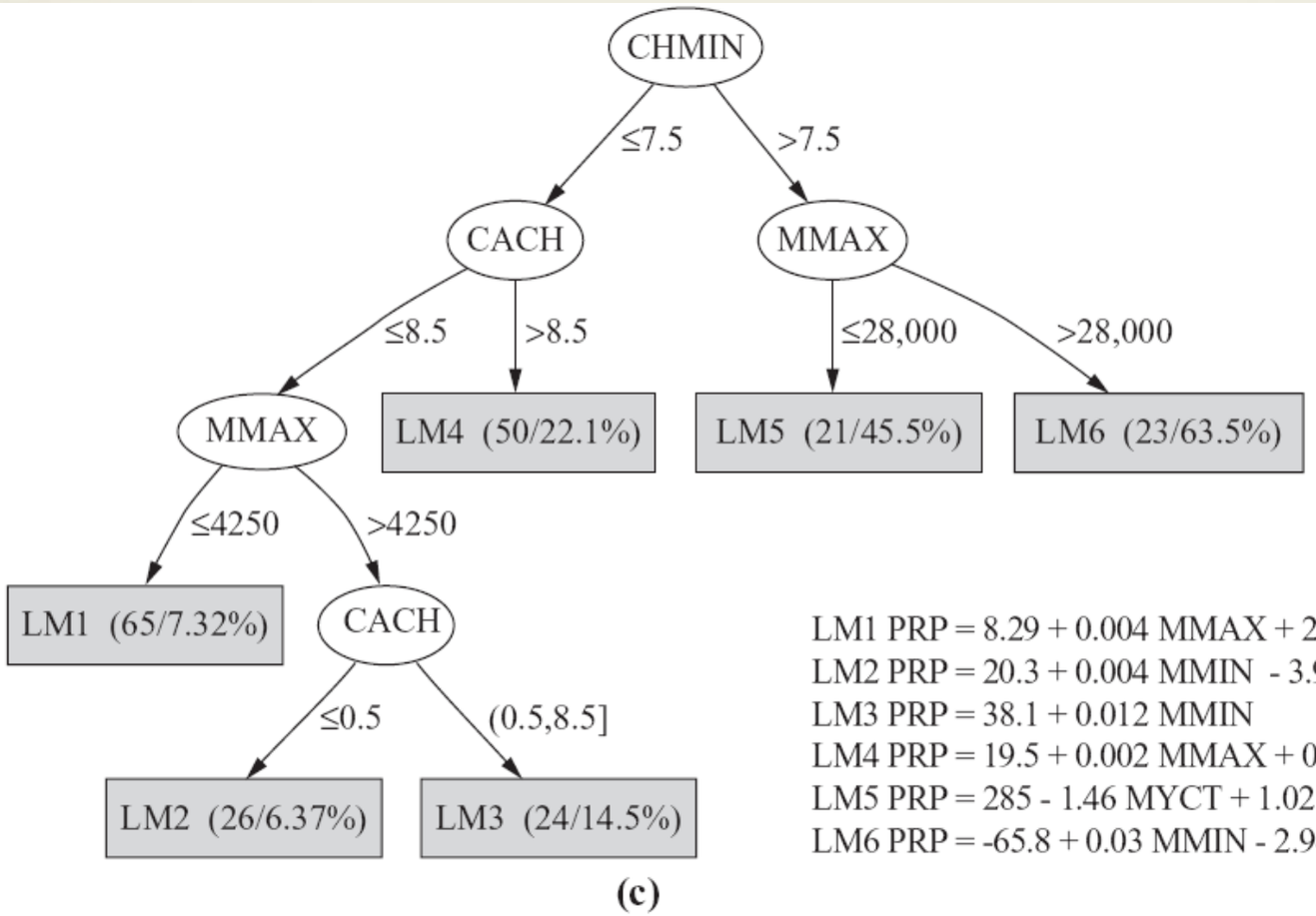
Predict numeric quantities

averaged numeric values

Tree (3/3)

Model Tree

Predict numeric quantities



LM1 PRP = 8.29 + 0.004 MMAX + 2.77 CHMIN
LM2 PRP = 20.3 + 0.004 MMIN - 3.99 CHMIN + 0.946 CHMAX
LM3 PRP = 38.1 + 0.012 MMIN
LM4 PRP = 19.5 + 0.002 MMAX + 0.698 CACH + 0.969 CHMAX
LM5 PRP = 285 - 1.46 MYCT + 1.02 CACH - 9.39 CHMIN
LM6 PRP = -65.8 + 0.03 MMIN - 2.94 CHMIN + 4.98 CHMAX

最後的leaf是linear model #different from regression tree

FIGURE 3.4

Models for the CPU performance data: (a) linear regression, (b) regression tree, and (c) model tree.

Rules

Classification Rule

If outlook=sunny and humidity=high	**then** play=no
If outlook=rainy and windy=true	**then** play=no
If outlook=overcast	**then** play=yes
If humidity=normal	**then** play=yes
If none of the above	**then** play=yes

decision list

有 if...else...then的感覺
又稱decision list

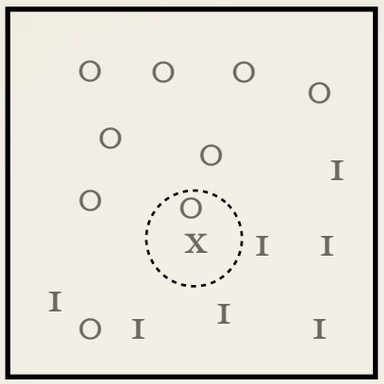
Association Rule

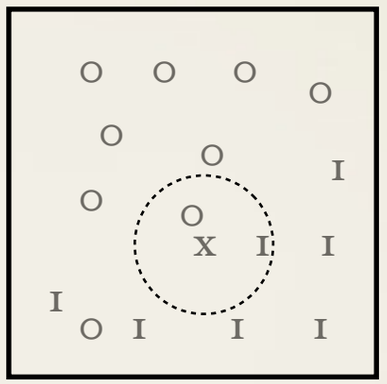
If temperature=cool	**then** humidity=normal
If humidity=normal and windy=false	**then** play=yes
If outlook=sunny and play=no	**then** humidity=high
If windy=false and play=no	**then** outlook=sunny and humidity=high

Instance-based Representation

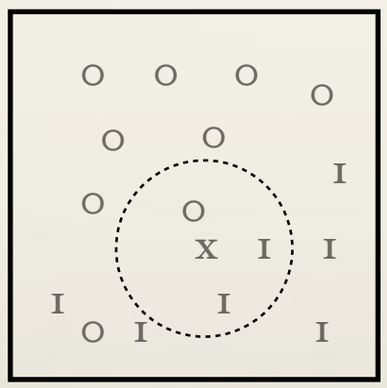
Classification or Numeric prediction

Exemplar view

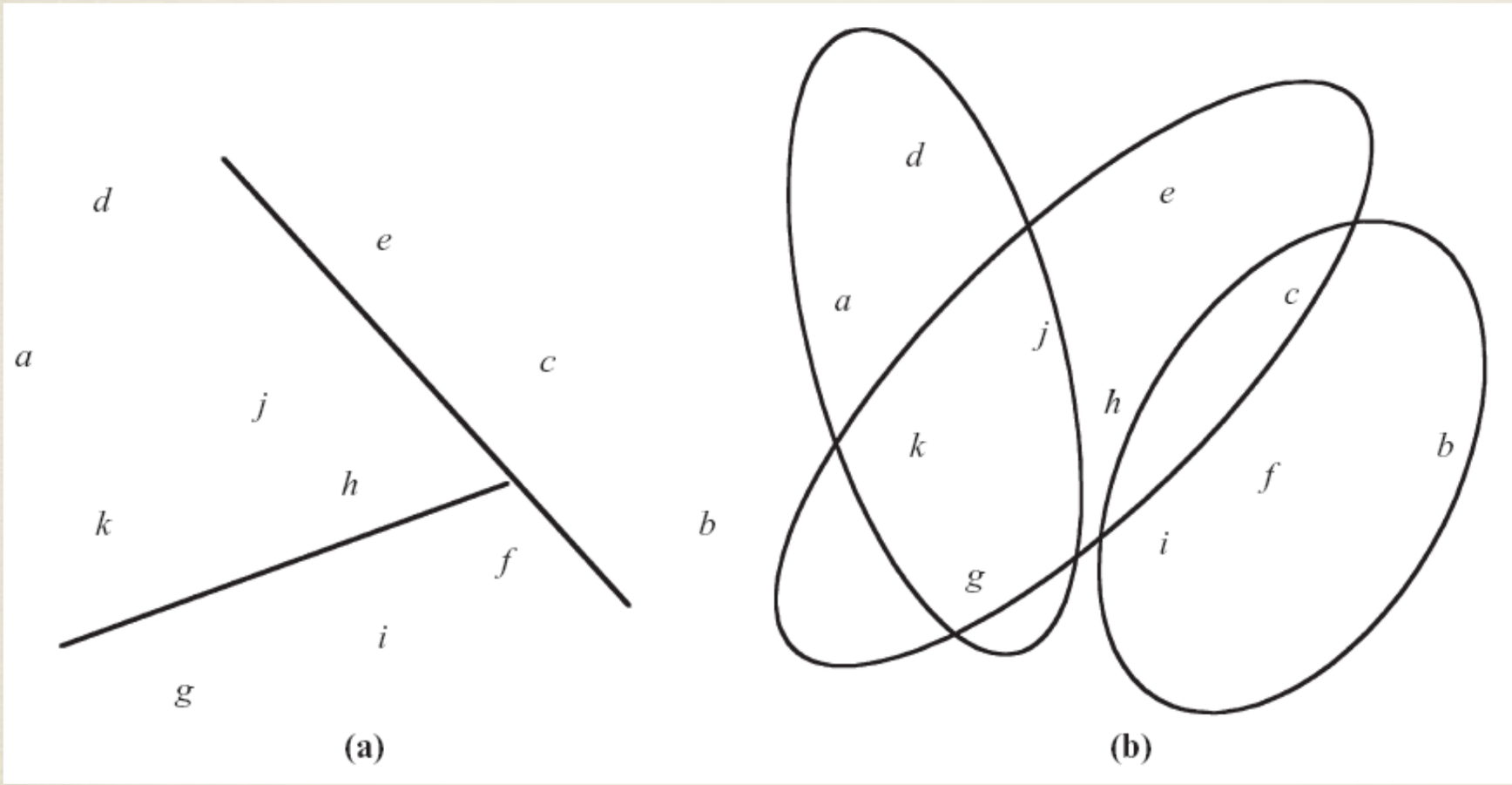




Randomly pick one or select the closest one



Clusters (1/2)



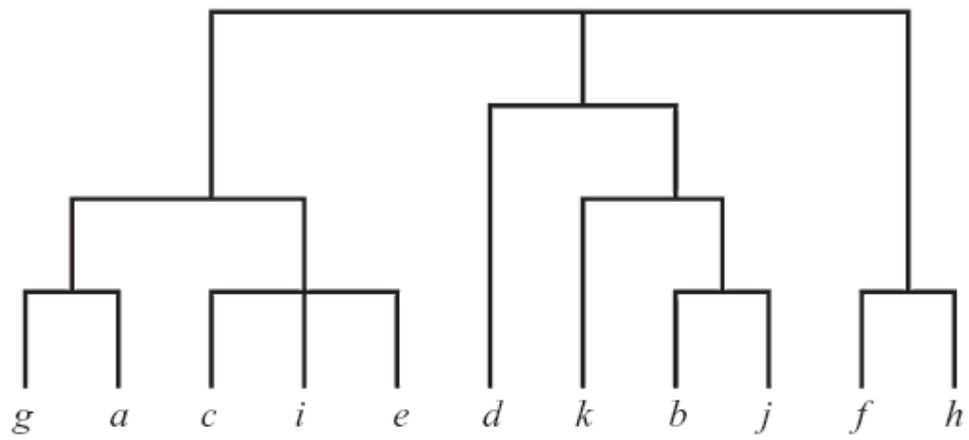
probability

Clusters (2/2)

	1	2	3
a	0.4	0.1	0.5
b	0.1	0.8	0.1
c	0.3	0.3	0.4
d	0.1	0.1	0.8
e	0.4	0.2	0.4
f	0.1	0.4	0.5
g	0.7	0.2	0.1
h	0.5	0.4	0.1
...			

probability

(c)



(d)

FIGURE 3.11

Different ways of representing clusters.