Example 7.4 **Predicting a class label using naive Bayesian classification:** We wish to predict the class label of an unknown sample using naive Bayesian classification, given the same training data as in Example 7.2 for decision tree induction. The training data are in Table 7.1. The data samples are described by the attributes *age*, *income*, *student*, and *credit_rating*. The class label attribute, *buys_computer*, has two distinct values (namely, {*yes*, *no*}). Let C_1 correspond to the class *buys_computer* = "*yes*" and C_2 correspond to *buys_computer* = "*no*". The unknown sample we wish to classify is

$$X = (age = \text{"<=30"}, income = \text{"medium"}, student = \text{"yes"}, credit\_rating = \text{"fair"}).$$

We need to maximize $P(X|C_i)P(C_i)$, for $i = 1, 2$. $P(C_i)$, the prior probability of each class, can be computed based on the training samples:

$$P(buys\_computer = \text{"yes"}) = 9/14 = 0.643$$

$$P(buys\_computer = \text{"no"}) = 5/14 = 0.357$$

To compute $P(X|C_i)$, for $i = 1, 2$, we compute the following conditional probabilities:

$$\begin{aligned}
&P(age = \text{"<30"} \mid buys_computer = \text{"yes"}) &&= 2/9 = 0.222\\
&P(age = \text{"<30"} \mid buys_computer = \text{"no"}) &&= 3/5 = 0.600\\
&P(income = \text{"medium"} \mid buys_computer = \text{"yes"}) &&= 4/9 = 0.444\\
&P(income = \text{"medium"} \mid buys_computer = \text{"no"}) &&= 2/5 = 0.400\\
&P(student = \text{"yes"} \mid buys_computer = \text{"yes"}) &&= 6/9 = 0.667\\
&P(student = \text{"yes"} \mid buys_computer = \text{"no"}) &&= 1/5 = 0.200\\
&P(credit_rating = \text{"fair"} \mid buys_computer = \text{"yes"}) &&= 6/9 = 0.667\\
&P(credit_rating = \text{"fair"} \mid buys_computer = \text{"no"}) &&= 2/5 = 0.400
\end{aligned}$$

Using the above probabilities, we obtain

$$\begin{aligned}
&P(X|buys_computer = \text{"yes"}) = 0.222 \times 0.444 \times 0.667 \times 0.667 = 0.044\\
&P(X|buys_computer = \text{"no"}) = 0.600 \times 0.400 \times 0.200 \times 0.400 = 0.019\\
&P(X|buys_computer = \text{"yes"})P(buys_computer = \text{"yes"}) = 0.044 \times 0.643 = 0.028\\
&P(X|buys_computer = \text{"no"})P(buys_computer = \text{"no"}) = 0.019 \times 0.357 = 0.007
\end{aligned}$$

NOTE: DON'T NEED P(X)

Therefore, the naive Bayesian classifier predicts *buys_computer* = "*yes*" for sample X. ■

7.4.3 Bayesian Belief Networks

The naive Bayesian classifier makes the assumption of class conditional independence, that is, given the class label of a sample, the values of the attributes are