Bayesian Classification

Bayer's Theorem

- $P(H|X) = \dfrac{P(X|H)\ P(H)}{P(X)}$

Hypothesis

$H: C = C_1$

- We can determine the prior probability in several ways.

Eg: If we know that there are 30% applications are safe and 70% are risky.

$P(C_i|X) = \dfrac{P(X|C_i)\ P(C_i)}{P(X)}$

$P(X)$ is constant for all classes. Hence, only $P(X|C_i)\ P(C_i)$

$P(X|C_i) = P(X_1, X_2, X_3, \dots, X_n | C_i)$

$= P(X_1|C_i)\ P(X_2|C_i) \cdots P(X_n|C_i)$

- $P(X_k|C_i) = N(x_k - 0.5 \leq x_k \leq x_k + 0.5, \mu_{C_i}, \sigma_{C_i})$

Example

X = (age = youth, income = medium, student = yes, credit_rating = fair)

Let C_1 = buys a computer C_2 = do not buy a computer

$P(C_1|X) = ?$ $P(C_2|X) = ?$

$P(C_1|X) = \dfrac{P(X|C_1)\ P(C_1)}{P(X)}$

$= \dfrac{P(x_1 = \text{youth}, x_2 = \text{medium}, x_3 = \text{yes}, x_4 = \text{fair} \mid C_1)\ P(C_1)}{P(X)}$

$P(X) = P(X_1 \cap X_2 \cap X_3 \cap X_4)$

$P(C_1) = \dfrac{9}{14}$ $P(C_2) = \dfrac{5}{14}$

$P(x_1|C_1) = \dfrac{2}{9}$ $P(x_2|C_1) = \dfrac{4}{9}$ $P(x_3|C_1) = \dfrac{6}{9}$ $P(x_4|C_1) = \dfrac{6}{9}$

$P(X|C_1)P(C_1) = \dfrac{2}{9} \times \dfrac{4}{9} \times \dfrac{6}{9} \times \dfrac{6}{9} \times \dfrac{9}{14} = 0.0282$

$$P(C_2|X) = \frac{P(X|C_2)\; P(C_2)}{P(X)}$$

$$= \frac{P(x_1|C_2)\, P(x_2|C_2)\, P(x_3|C_2)\, P(x_4|C_2)\; P(C_2)}{P(X)}$$

$P(x_1|C_2) = \frac{3}{5}$ $\quad P(x_2|C_2) = \frac{2}{5}$ $\quad P(x_3|C_2) = \frac{1}{5}$ $\quad P(x_4|C_2) = \frac{2}{5}$

$$P(X|C_2)P(C_2) = \frac{3}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{5}{14} = \frac{12 \times 5}{(5)^3 \times 14} = 0.0343$$

- $P(X|C_1)P(C_1) < P(X|C_2)P(C_2)$.

Therefore, the given observation belongs to the class C_2//.

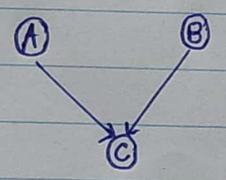
Bayesian Belief Networks (BBN)

- $P(X|Y,Z) = P(X|Z)$ means x, y, z are independent

if A is independent and B depends on A, then

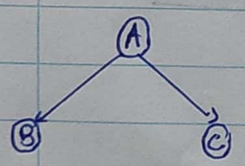
$P(X_A, X_B) = P(B|A)\,P(A)$

↑ joint probability



A - independent, B - independent, C - has 2 parents

$P(A,B,C) = P(C|A,B)\,P(A)\,P(B)$



$P(A,B,C) = P(B|A)\,P(C|A)\,P(A)$

$A \rightarrow B \rightarrow C$ $\quad P(A,B,C) = P(C|B)\,P(B|A)\,P(A)$

The Alarm Example

B = a burglary occurs $\quad$ J = John calls

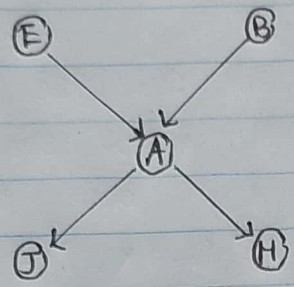
E = an earthquakes occurs $\quad$ M = Mary calls

A = alarm ringing

No: Date: / /

Find the probability of John and Harry report the alarm and neither earthquakes nor burglary occur.

A = Y, B = N, E = N, J = Y, H = Y.



$P(A, B', E', J, H) = ?$

$= P(J \mid A) \times P(H \mid A) \times P(A \mid B', E') \times P(E') P(B')$

$= 0.9 \times 0.7 \times 0.001 \times 0.998 \times 0.999$

$= 0.00063$ //

• $P(J,A,B,E) = P(J \mid A) \times P(A \mid B, E) \times P(B) P(E)$

$P(J,A,B,E') = P(J \mid A) \times P(A \mid B, E') \times P(B) P(E')$

$P(J,A,B',E) = P(J \mid A) P(A \mid B', E) P(B') P(E)$

$P(J,A,B',E') = P(J \mid A) P(A \mid B', E') P(B') P(E')$

$P(J,A,B,E) = 0.9 \times 0.95 \times 0.001 \times 0.002 = 1.71 \times 10^{-6}$

$P(J,A,B,E') = 0.9 \times 0.94 \times 0.001 \times 0.998 = 8.443 \times 10^{-4}$

$P(J,A,B',E) = 0.9 \times 0.29 \times 0.999 \times 0.002 = 5.2148 \times 10^{-4}$

$P(J,A,B',E') = 0.9 \times 0.001 \times 0.999 \times 0.998 = 8.973 \times 10^{-4}$

$P(J,A',B,E) = P(J \mid A') P(A' \mid B, E) P(B) P(E)$

$= 0.05 \times 0.05 \times 0.001 \times 0.002 = 5 \times 10^{-9}$

$P(J,A',B,E') = P(J \mid A') P(A' \mid B, E') P(B) P(E')$

$= 0.05 \times 0.06 \times 0.001 \times 0.998 = 2.994 \times 10^{-6}$

$P(J,A',B',E) = P(J \mid A') P(A' \mid B', E) P(B') P(E)$

$= 0.05 \times 0.71 \times 0.999 \times 0.002 = 7.0929 \times 10^{-5}$

$P(J,A',B',E') = P(J \mid A') P(A' \mid B', E') P(B') P(E')$

$= 0.05 \times 0.999 \times 0.999 \times 0.998 = 0.0498$

$= 521.38763 \times 10^{-4} = 0.05214$ //