Inferențe în Rețele Bayesiene

Tudor Berariu
tudor.berariu@gmail.com

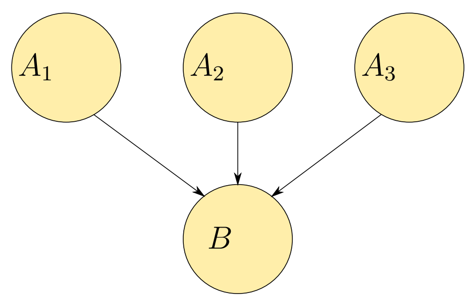
4 februarie 2015

1 Mai multe cauze ale aceluiași eveniment

1.1 Trei cauze

Problema

Fie rețeaua bayesiană din figura de mai jos. Aceasta este descrisă de următorii $[1 + 1 + 1 + 2^3 = 11]$ parametri:



$P(A_1)$
$P(A_2)$
$P(A_3)$
$P(B|A_1, A_2, A_3)$
$P(B|A_1, A_2, \neg A_3)$
$P(B|A_1, \neg A_2, A_3)$
$\dots$
$P(B|\neg A_1, \neg A_2, \neg A_3)$

Să presupunem că dorim să calculăm $P(A_1|B)$.

Primul pas

Conform Teoremei lui Bayes:

$$P(A_1|B) = \frac{\color{green}{P(B|A_1)} \cdot P(A_1)}{\color{purple}{P(B)}} \tag{1}$$

În Ecuația 1 avem două cantități necunoscute: $P(B|A_1)$ și $P(B)$.

Calculul lui $P(B|A_1)$

Putem introduce o variabilă nouă, aplicând formula probailităților totale.

$$P(X) = \sum_y P(X|Y=y) \cdot P(Y=y)$$

Pentru evenimente binare:

$$P(X) = P(X|Y) \cdot P(Y) + P(X|\neg Y) \cdot P(\neg Y)$$

O introducem pe A_2:

$$P(B|A_1) = P(B|A_1, A_2) \cdot P(A_2|A_1) + P(B|A_1, \neg A_2) \cdot P(\neg A_2|A_1) \tag{2}$$

Dar, cum [fără a fi B observat] A_1 și A_2 sunt condițional independente, $P(A_2|A_1) = P(A_2)$ și $P(\neg A_2|A_1) = P(\neg A_2)$, iar Ecuația 2 devine:

$$P(B|A_1) = P(B|A_1, A_2) \cdot P(A_2) + P(B|A_1, \neg A_2) \cdot P(\neg A_2) \tag{3}$$

Aplicând din nou formula probabilităților totale pentru a factoriza după A_3:

$$\begin{aligned}
P(B|A_1) &= P(B|A_1, A_2) \cdot P(A_2) + P(B|A_1, \neg A_2) \cdot P(\neg A_2) \\
P(B|A_1) &= \big(P(B|A_1, A_2, A_3) \cdot P(A_3|A_1, A_2) + \\
&\quad\; P(B|A_1, A_2, \neg A_3) \cdot P(\neg A_3|A_1, A_2)\big) \cdot P(A_2) + \\
&\quad \big(P(B|A_1, \neg A_2, A_3) \cdot P(A_3|A_1, \neg A_2) + \\
&\quad\; P(B|A_1, \neg A_2, \neg A_3) \cdot P(\neg A_3|A_1, \neg A_2)\big) \cdot P(\neg A_2)
\end{aligned}$$

Dar, aplicând același raționament ca înainte, A_3 este condițional independent de A_1 și A_2 atunci când B nu este observat și atunci:

- $P(A_3|A_1, A_2) = P(A_3)$
- $P(\neg A_3|A_1, A_2) = P(\neg A_3)$

- $P(A_3|A_1, \neg A_2) = P(A_3)$
- $P(\neg A_3|A_1, \neg A_2) = P(\neg A_3)$

și atunci rescriem ecuația anterioară:

$$\begin{aligned}
P(B|A_1) &= \big(P(B|A_1, A_2, A_3) \cdot P(A_3) + \\
&\quad P(B|A_1, A_2, \neg A_3) \cdot P(\neg A_3)\big) \cdot P(A_2) + \\
&\quad \big(P(B|A_1, \neg A_2, A_3) \cdot P(A_3) + \\
&\quad P(B|A_1, \neg A_2, \neg A_3) \cdot P(\neg A_3)\big) \cdot P(\neg A_2) \\
&= P(B|A_1, A_2, A_3) \cdot P(A_2) \cdot P(A_3) + \\
&\quad P(B|A_1, A_2, \neg A_3) \cdot P(A_2) \cdot P(\neg A_3) + \\
&\quad P(B|A_1, \neg A_2, A_3) \cdot P(\neg A_2) \cdot P(A_3) + \\
&\quad P(B|A_1, \neg A_2, \neg A_3) \cdot P(\neg A_2) \cdot P(\neg A_3)
\end{aligned}$$

Înlocuind ultimul rezultat în Ecuația 2:

$$\begin{aligned}
P(A_1|B) &= \Big(P(B|A_1, A_2, A_3) \cdot P(A_2) \cdot P(A_3) + \\
&\quad P(B|A_1, A_2, \neg A_3) \cdot P(A_2) \cdot P(\neg A_3) + \\
&\quad P(B|A_1, \neg A_2, A_3) \cdot P(\neg A_2) \cdot P(A_3) + \\
&\quad P(B|A_1, \neg A_2, \neg A_3) \cdot P(\neg A_2) \cdot P(\neg A_3)\Big) \cdot \frac{P(A_1)}{P(B)}
\end{aligned}$$

$P(B)$ este necunoscut

Mai rămâne o problemă: calculul lui $P(B)$. Folosind aceeași strategie de mai sus intuim că $P(B)$ s-ar exprima printr-o sumă de 8 termeni de tipul $P(B|A_1, A_2, A_3) \cdot P(A_1) \cdot P(A_2) \cdot P(A_3)$. Mai există o soluție prin care putem evita această desfășurare explicită a lui $P(B)$ în funcție de cauzele lui (evenimentele ce condiționează pe) B: calculăm $P(\neg A_1|B)$.

Urmând aceiași pași ca mai devreme, calculăm:

$$
\begin{aligned}
P(\neg A_1|B) \;=\; & \Big(P(B|\neg A_1, A_2, A_3) \cdot P(A_2) \cdot P(A_3) + \\
& P(B|\neg A_1, A_2, \neg A_3) \cdot P(A_2) \cdot P(\neg A_3) + \\
& P(B|\neg A_1, \neg A_2, A_3) \cdot P(\neg A_2) \cdot P(A_3) + \\
& P(B|\neg A_1, \neg A_2, \neg A_3) \cdot P(\neg A_2) \cdot P(\neg A_3) \Big) \cdot \frac{P(\neg A_1)}{P(B)}
\end{aligned}
$$

Știind că $P(A_1|B) + P(\neg A_1|B) = 1$, obținem un sistem de ecuații pe care îl putem rezolva fără a mai calcula $P(B)$.

1.2 N cauze

Dacă am avea nu 3, ci N evenimente A_i ce condiționează evenimentul B:

$$
\begin{aligned}
P(A_i|B) &= \frac{P(B|A_i) \cdot P(A_i)}{P(B)} \\
&= \frac{\left(\sum P(B|A_1, \ldots, A_N) \cdot P(A_1) \cdot \ldots \cdot P(A_{i-1}) \cdot P(A_{i+1}) \cdot \ldots \cdot P(A_N) \right) \cdot P(A_i)}{P(B)}
\end{aligned}
$$