شناسایی آماری الگو
بخش یازدهم
(۴۱-۱۰-۷۱۱-۰۱)

مقدمه‌ای بر PGM

Bayesian networks
belief networks
Probabilistic networks

دانشگاه شهید بهشتی
پژوهشکده‌ی فضای مجازی
بهار ۱۳۹۶
احمد محمودی ازناوه

http://faculties.sbu.ac.ir/~a_mahmoudi/

فهرست مطالب

- مدل‌های گرافیکی
  - پیش‌گفتار
  - انواع اتصالات
  - استنتاج علی و تشخیصی
- یادگیری
- دیاگرام تأثیر

پیش‌گفتار

- استفاده از مدل‌های گرافیکی احتمالی به نوعی ترکیب تئوری گراف و احتمالات برای **تصمیم‌گیری** و **استنتاج** تحت شرایط احتمالاتی است.
- یک «نمودار دیداری» برای متغیرهای تصادفی ارائه می‌دهد.
- در زمینه‌های مختلفی نظیر یادگیری ماشین، بینایی ماشین، پرازش زبان طبیعی و بیوانفورماتیک کاربرد دارد.
- از کاربردهای آن می‌توان تشخیص‌های پزشکی، قطعه‌بندی تصویر و تعیین میزان اطمینان اشاره کرد.

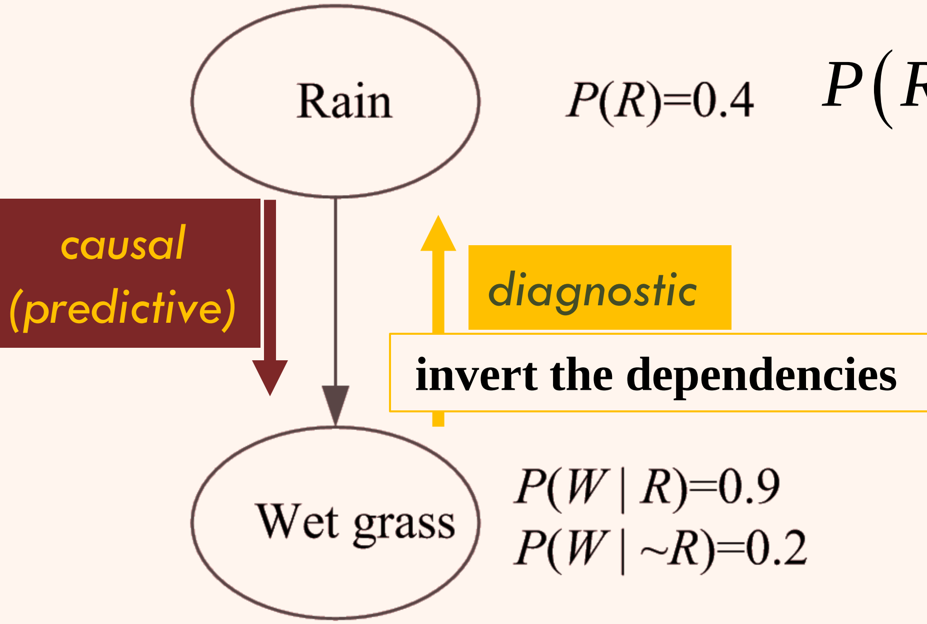
مدل‌های گرافیکی

- این شبکه‌های از یک سری «گره» و «یال» تشکیل شده است.
- هر گره معادل یک «متغیر تصادفی» است و «احتمال آن متغیر تصادفی» به آن گره نسبت داده می‌شود.
- در صورتی که بین گره‌ی X و Y یک یال وجود داشته باشد($X \rightarrow Y$)، به معنی **اثرگذاری مستقیم** X بر روی Y است.

direct influence

مثال

- در صورتی که بدانیم چمن‌ها خیس هستند، با چه احتمالی باران باعث این پدیده شده است؟



$P(R)=0.4$

$P(W \mid R)=0.9$

$P(W \mid \sim R)=0.2$

$$P(R|W) = \frac{P(W|R)P(R)}{P(W)} = \frac{P(W|R)P(R)}{P(W|R)P(R)+P(W|\sim R)P(\sim R)} = \frac{0.9\times 0.4}{0.9\times 0.4+0.2\times 0.6} = 0.75$$

- احتمال آن که هم چمن‌ها خیس باشند و هم باران بیارد:

$$P(R,W)=P(R)P(W|X)$$

یادآوری

- X و Y مستقلند اگر:

$$P(X,Y)=P(X)P(Y)$$

- X و Y مشروط به رخداد Z مستقل هستند اگر:

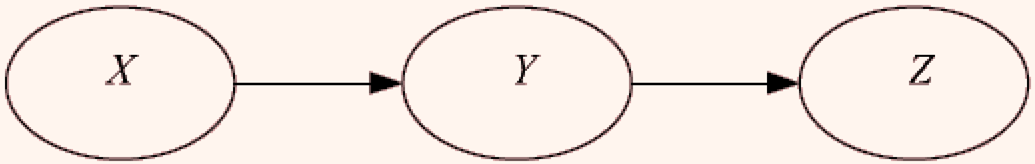
$$P(X,Y|Z)=P(X|Z)P(Y|Z)$$

یا

conditionally independent given Z

$$P(X|Y,Z)=P(X|Z)$$

Case 1: Head-to-Tail



- در این حالت سه رویداد به صورت سری به هم متصل می‌شوند.

$$P(X,Y,Z)=P(X)P(Y|X)P(Z|Y)$$

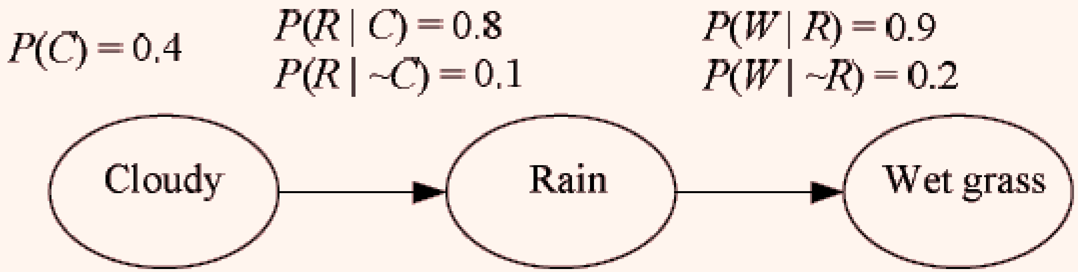
- در صورت دانستن Y، X و Z مستقل از یکدیگر خواهند بود:

$$P(Z|X,Y)=\frac{P(X,Y,Z)}{P(X,Y)}=\frac{P(X)P(Y|X)P(Z|Y)}{P(X)P(Y|X)}=P(Z|Y)$$

- در واقع با دانستن Y، Z همه چیز در مورد X را می‌داند.

Case 1: Head-to-Tail

مثال



$$P(R) = P(R|C)P(C) + P(R|\sim C)P(\sim C) = 0.38$$

$$P(W) = P(W|R)P(R) + P(W|\sim R)P(\sim R) = 0.48$$

- احتمال مرطوب بودن به شرط ابری بودن:

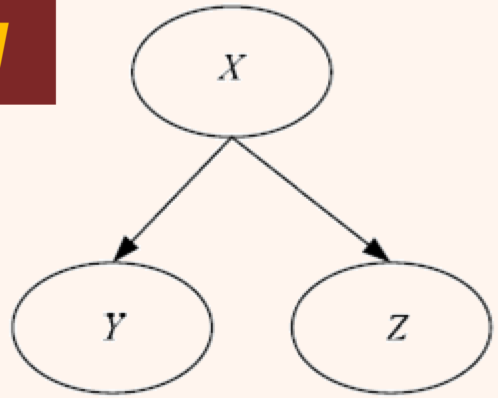
$$P(W|C)=$$

$$=P(W|R)P(R|C)+P(W|\sim R)P(\sim R|C)$$

- با استفاده از قانون بیز می‌توانیم به صورت معکوس استنتاج کنیم:

$$P(C|W) = \frac{P(W|C)P(C)}{P(W)} = 0.65$$

Case 2: Tail-to-Tail



- در این حالت یک گره، بر روی دو گره به صورت مستقیم اثرگذار است:

$$P(X,Y,Z)=P(X)P(Y/X)P(Z|X)$$

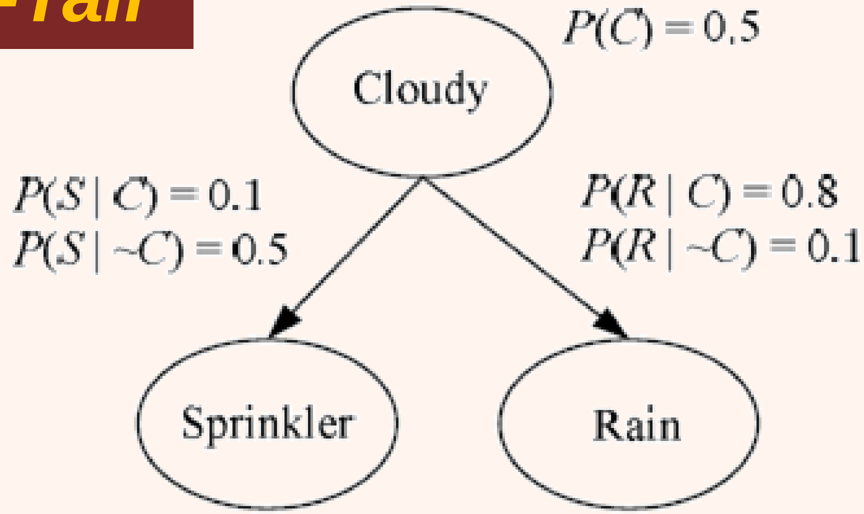
- در این حالت Y و Z از طریق X به هم وابسته هستند. در صورت وجود X نسبت به هم مستقل خواهند بود:

$$P(Y,Z|X)=\frac{P(X,Y,Z)}{P(X)}=\frac{P(X)P(Y|X)P(Z|X)}{P(X)}=P(Y|X)P(Z|X)$$

conditionally independent given *X*

مثال

Case 2: Tail-to-Tail



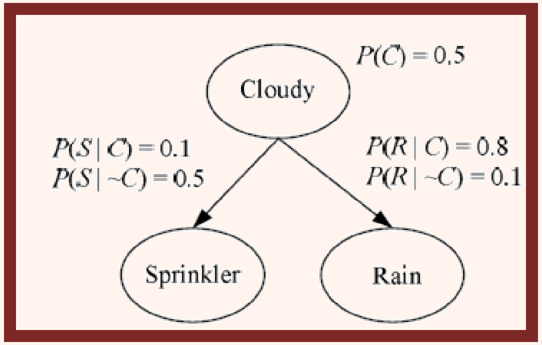
- در صورتی که بدانیم باران آمده است، احتمال ابری بودن هوا($P(C|R)$):

$$
\begin{aligned}
P(C|R) &= \frac{P(R|C)P(C)}{P(R)} = \frac{P(R|C)P(C)}{\sum_C P(R,C)} \\
&= \frac{P(R|C)P(C)}{P(R|C)P(C) + P(R|\sim C)P(\sim C)} = 0.89
\end{aligned}
$$

مثال

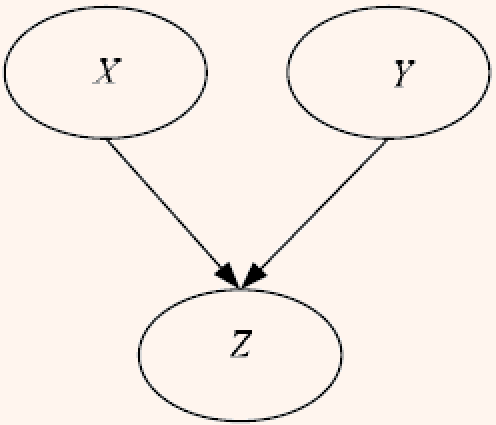
- در صورتی که تنها در مورد یکی از دو گره‌ی فرزند اطلاعاتی داشته باشیم، به عنوان مثال بارش باران می‌توان از طریق گره‌ی والد به اطلاعاتی در خصوص دیگر گره دست یافت($P(R|S)$):

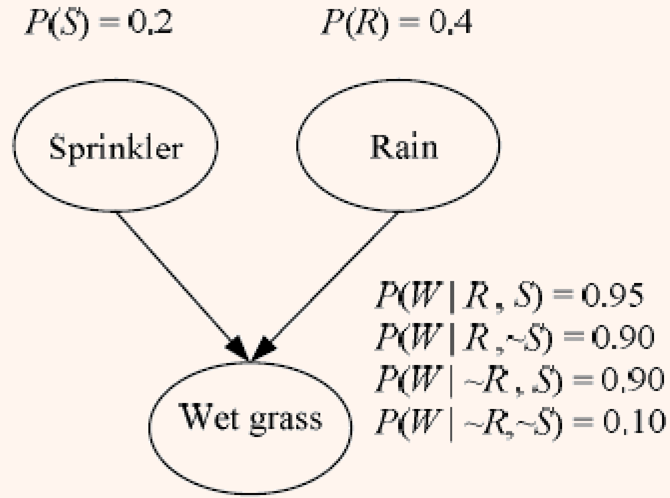
$$
\begin{aligned}
P(R|S) &= \sum_{C} P(R,C|S) = P(R|C)P(C|S) + P(R|{\sim}C)P({\sim}C|S) \\
&= P(R|C)\frac{P(S|C)P(C)}{P(S)} + P(R|{\sim}C)\frac{P(S|{\sim}C)P({\sim}C)}{P(S)} \\
&= 0.22
\end{aligned}
$$



Case 3: Head-to-Head

- در این حالت یک گره دارای دو والد است:



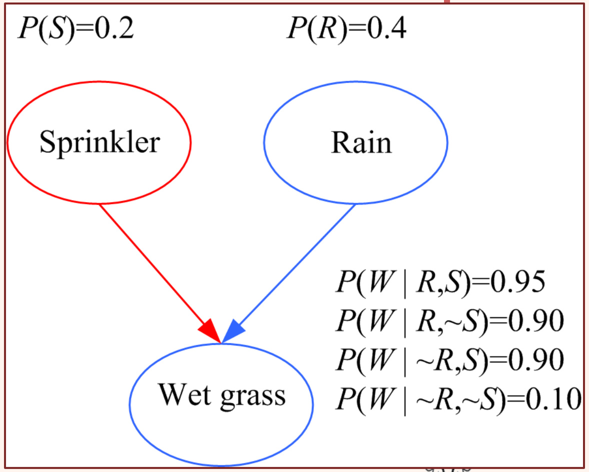


$$P(X,Y,Z)=P(X)P(Y)P(Z|X,Y)$$

مثال

Case 3: Head-to-Head

- در صورتی که در خصوص دو متغیر دیگر چیزی ندانیم، احتمال خیس بودن چمن(($P(W)$):

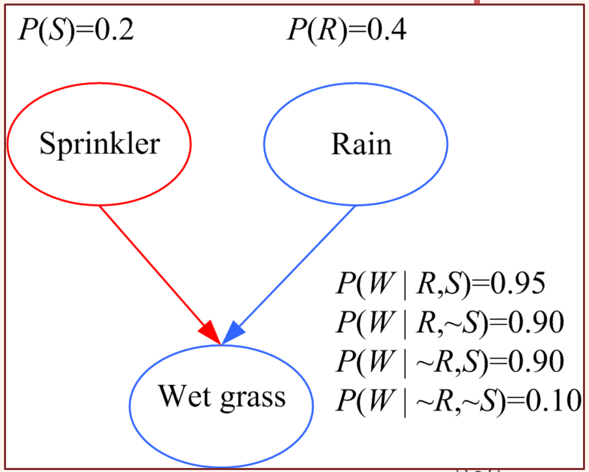


$$
\begin{aligned}
P(W) &= \sum_{R,S} P(W,R,S) \\
&= P(W|R,S)P(R,S) + P(W|\sim R,S)P(\sim R,S) \\
&\quad + P(W|R,\sim S)P(R,\sim S) + P(W|\sim R,\sim S)P(\sim R,\sim S) \\
&= P(W|R,S)P(R)P(S) + P(W|\sim R,S)P(\sim R)P(S) \\
&\quad + P(W|R,\sim S)P(R)P(\sim S) + P(W|\sim R,\sim S)P(\sim R)P(\sim S) \\
&= 0.52
\end{aligned}
$$

مثال

- احتمال خیس بودن چمن به شرط روشن بودن آب‌پاش:

$$P(W|S) = \sum_R P(W,R|S)$$



$$P(W|S) = P(W|R,S)\, P(R|S) + P(W|\sim R,S)\, P(\sim R|S)$$

$$= P(W|R,S)\, P(R) + P(W|\sim R,S)\, P(\sim R)$$

$$= 0.95\ 0.4 + 0.9\ 0.6 = 0.92$$

- و استنتاج تشخیصی روشن بودن آب‌پاش در صورت خیس بودن چمن

$$P(S|W) = \frac{P(W|S)P(S)}{P(W)} = 0.35$$

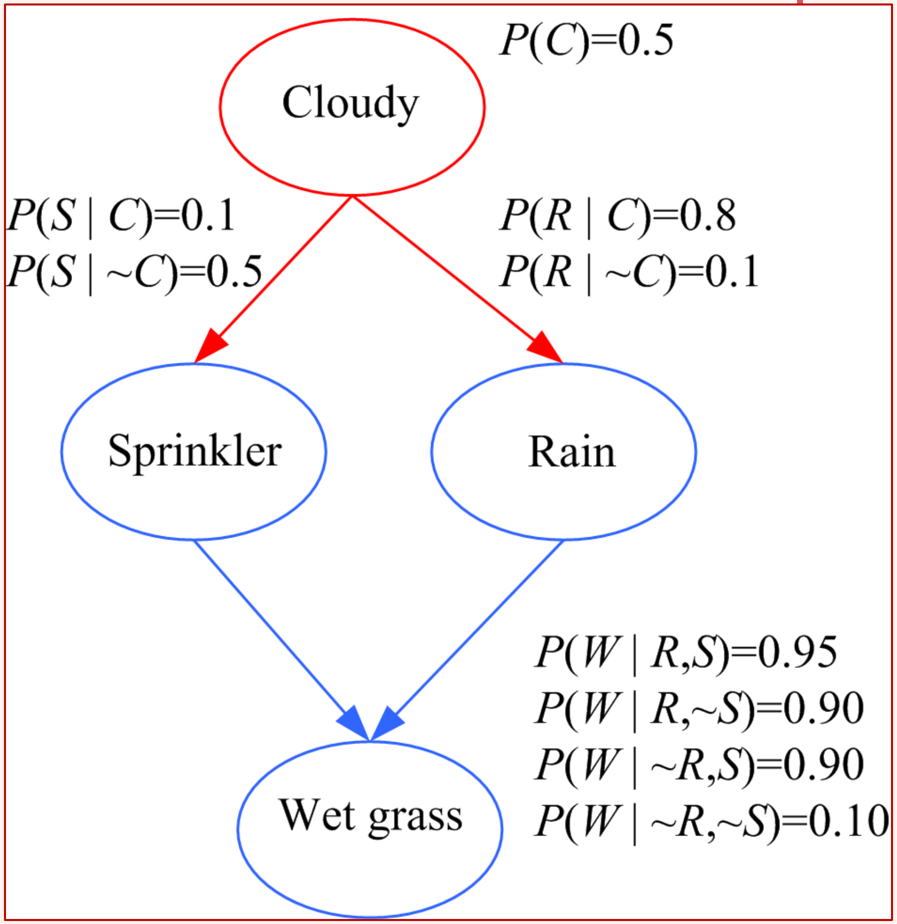
گراف کامل

Causal inference:

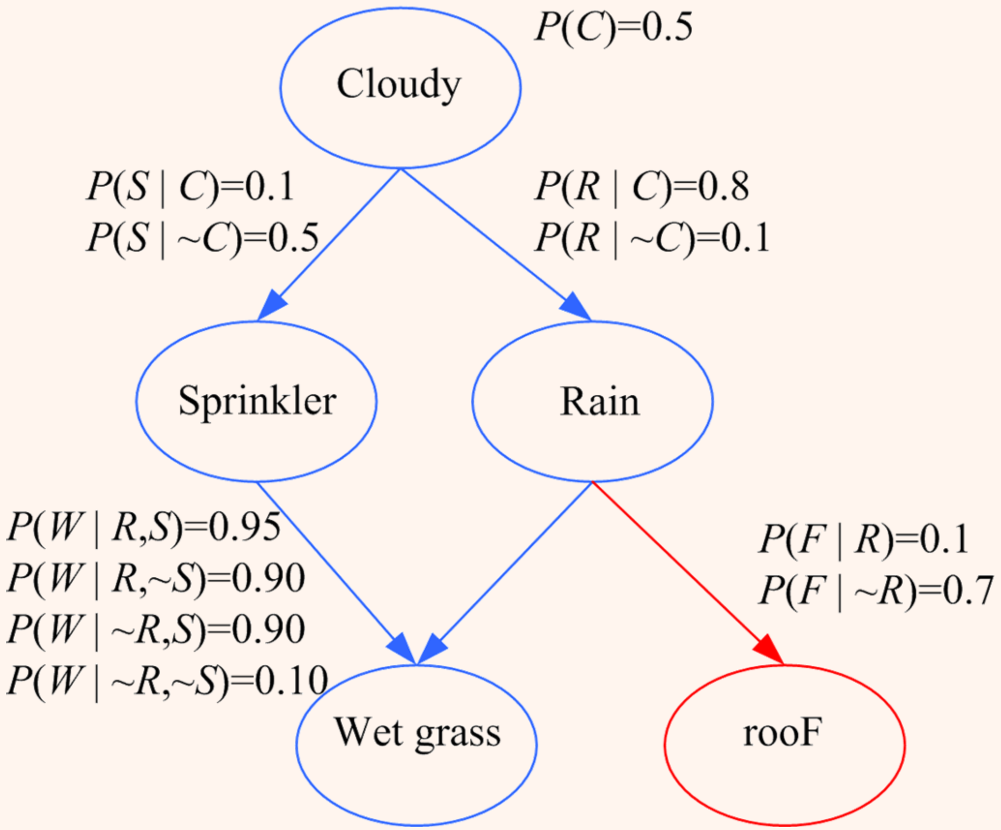
$$P(W|C) = P(W|R,S)\,P(R,S|C) + P(W|\sim R,S)\,P(\sim R,S|C) + P(W|R,\sim S)\,P(R,\sim S|C) + P(W|\sim R,\sim S)\,P(\sim R,\sim S|C)$$

$$P(R,S|C) = P(R|C)\,P(S|C)$$

$$P(C|W) = \frac{P(W|C)P(C)}{P(W)}$$



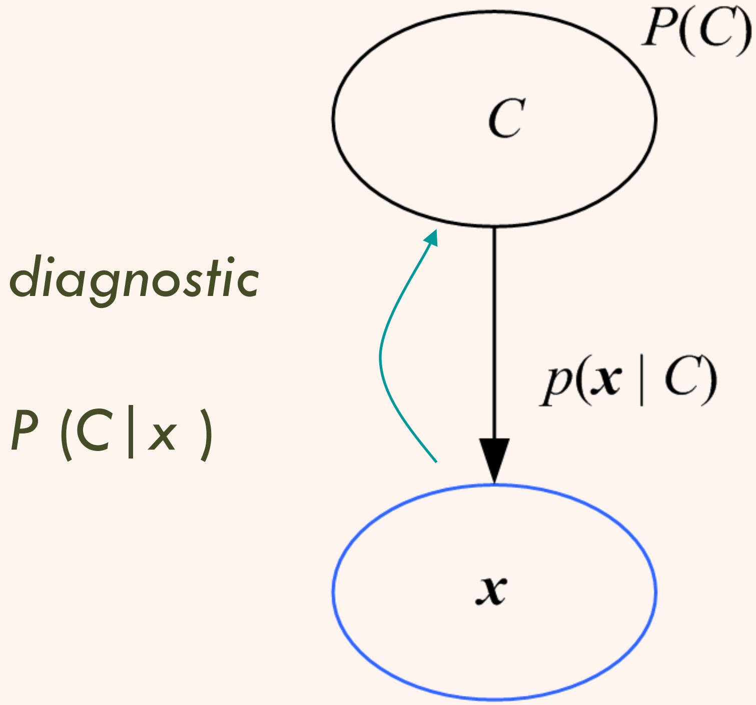
Exploiting the Local Structure



$$P(C,S,R,W,F) = P(C)P(S|C)P(R|C)P(W|S,R)P(F|R)$$

$$P(X_1,\ldots X_d) = \prod_{i=1}^{d} P(X_i \,|\, \text{parents}(X_i))$$

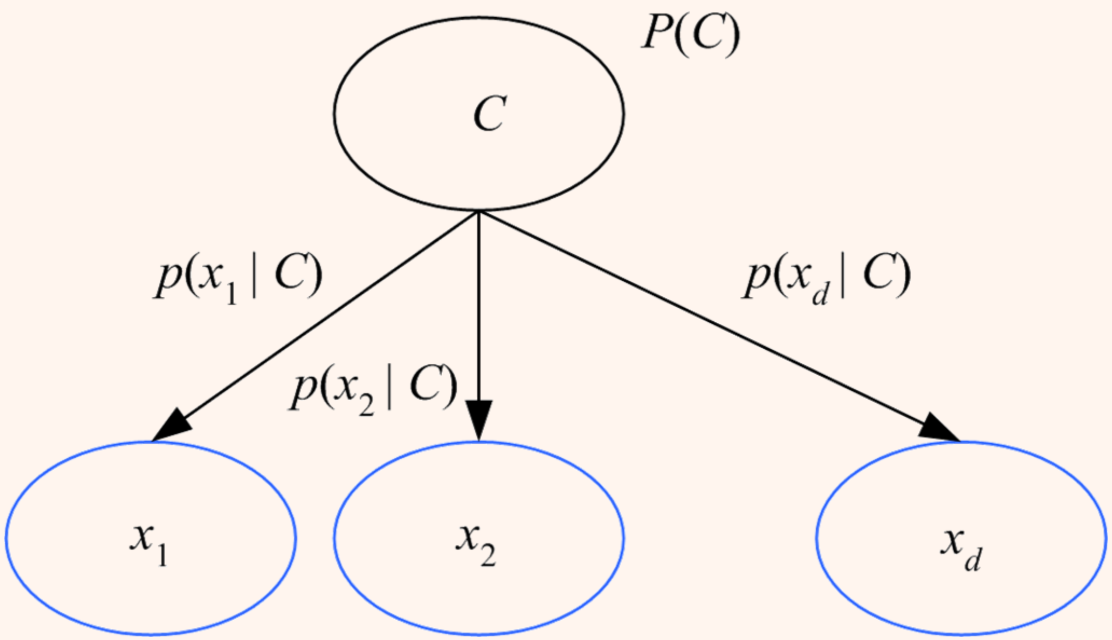
مثال(دسته‌بندی)



Bayes' rule inverts the arc:

$$P(C \mid \mathbf{x}) = \frac{p(\mathbf{x} \mid C) P(C)}{p(\mathbf{x})}$$

Naive Bayes' Classifier



Given C, x_j are independent:

$$p(\mathbf{x}|C) = p(x_1|C)\, p(x_2|C) \ldots p(x_d|C)$$

یادگیری مدل گرافیکی

- یادگیری احتمالات شرطی:
  - براساس ML یا رویکرد بیزی پارامترها آموزش می‌بینند.
- یادگیری ساختار:
  - مدل‌های مختلف بر اساس یک تابع معیار مورد بررسی قرار می‌گیرند.

با استفاده از این مدل‌ها لزومی ندارد، یک دسته ورودی و خروجی تعریف کنیم؛ رابطه‌ی یک‌سری متغیر تصادفی به دست می‌آید.