Bayesovské sitě

Je to způsob grafického vyjádření závislosti mezi jery (proměnnými), měří závislost je podmíněna pravděpodobnost.

Slouží k rozhodování v situacích, které jsou zatíženy nějakou neurčitostí.

Jednoduché příklady

a) Mám auto a 4 proměnné: benzín (ano, ne)

   ukazatel paliva (prázdné, na život, plné)

   zapalování (cistě, zaneseno)

   start (ano, ne)

Vím, že např.: 

lidy nebude mít benzín, pravděpodobně nemostartuji

-/-

ukazatel pravděpodobně ubírá prázdně

lidy je zaneseno zapalování, pravděpodobně nemostartuji

lidy mám dost paliva a cistě zapalování, pravděpodobně

pomostartuji

Můj problém ale bude přesně opačný - jistila jsem, že auto nestartuje a odhadoval, čím to pravděpodobně bude

Podívej se na ukazatel paliva, jistím, že ukazuje 1/2 - 

můžu k tomu usoudit, že problém je v zásobování ?

nebo ukazatel ukazuje zprávně a je možné, že nemám benzín ?
graficky:

Benin ➔ Zapolevočí

Ukazatel ➔ Start

2) proměnné: pohlaví (muž, žena)
   výška (malá, velká)
   délka vlasů (krátké, dlouhé)

Vím, že: muž bude pravděpodobně výšší

— n —

mít krátké vlasy

Můj problém: vidím osobu s dlouhými vlasy
bude to spíš muž nebo žena?
bude spíš mensí nebo výšší?

graficky:

Pohlaví

∨

výška ➔ vlasy
3) proměnné: salmonela'za (ano, ne)
   chřipka (ano, ne)
   žaludeční nevolnost (ano, ne)
   bledost (ano, ne)

graf:

\[\begin{align*}
\text{Salmonela} & \quad \text{Chřipka} \\
\downarrow & \quad \downarrow \\
\text{Nevolnost} & \quad \text{Bledost}
\end{align*}\]

Bayesovský sítě je orientovaný acyklicky graf

vrcholy jsou náhodné proměnné s konečně mnoha náležejícími disjunktními stav

každé proměnné 'A', jejímiž rodiči jsou B₁, ..., Bₙ, je

přiřazena podmíněná pravděpodobnost \(P(A|B₁, ..., Bₙ)\)
Podmíněné pravděpodobnosti

Definice: \( P(A|B) = \frac{P(AB)}{P(B)} \) pro \( P(B) > 0 \)

Obecnější definice: \( P(A|B,C) = \frac{P(ABC)}{P(BC)} \)

Aplikuje se v případech, kdy všechny jevy uvažujeme jako podmíněné nezávislé v kontextu \( C \)

(Rényho axiomatika podmíněných pravděpodobností)

Z definice dale plyne: (Bayesův vzorec)

\[ P(B|A) = \frac{P(AB)}{P(A)} = \frac{P(AB)P(B)}{P(A)} \]

resp. \( P(B|A,C) = \frac{P(ABC)P(B|C)}{P(A|C)} \)

Jevy \( A,B \) jsou nezávislé, když \( P(AB) = P(A)P(B) \)

neboli \( P(A|B) = P(A) \)

Jevy \( A,B \) jsou podmíněné nezávislé při daném \( C \), když \( P(A|C) = P(A|BC) \)
Dále platí:

Věta o úplné pravděpodobnosti: Nechť $B_1, \ldots, B_n$ jsou jevy takové, že $\mathbb{P}(B_i) > 0$ pro $i = 1, \ldots, n$ a $\mathbb{P}(B_i B_j) = 0$ pro $i \neq j$, a $\sum_{i=1}^{n} \mathbb{P}(B_i) = 1$. Pak

$$
\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i) \mathbb{P}(B_i).
$$

Bayesova věta:

$$
\mathbb{P}(B_i \mid A) = \frac{\mathbb{P}(A \mid B_i) \mathbb{P}(B_i)}{\sum_{i=1}^{n} \mathbb{P}(A \mid B_i) \mathbb{P}(B_i)}.
$$

Terminologie: $\mathbb{P}(B_i)$ ... apriorní pravděpodobnost jevu $B_i$

$\mathbb{P}(B_i \mid A)$ ... aposeriorní

Věta o nasobení pravděpodobnosti: Nechť $\mathbb{P}(A_1, \ldots, A_m) > 0$. Pak

$$
\mathbb{P}(A_1 \ldots A_m) = \mathbb{P}(A_1) \mathbb{P}(A_2 \mid A_1) \mathbb{P}(A_3 \mid A_2 A_1) \ldots \mathbb{P}(A_m \mid A_1 \ldots A_m)
$$

Odp.: $\mathbb{P}(A_1) \frac{\mathbb{P}(A_2 A_1)}{\mathbb{P}(A_1)} \frac{\mathbb{P}(A_3 A_2 A_1)}{\mathbb{P}(A_2 A_1)} \ldots \frac{\mathbb{P}(A_m \ldots A_1)}{\mathbb{P}(A_{m-1} \ldots A_1)}$
Odpověď: V moravské aleci se jmenuje "čalba," a "čalba" v poměru 5:3.

Předpokládám, že platí statistické následování: výtvarní signálie "čalba" přijdou 4:5 jako "čalba," a významové signálie "čalba," přijdou 4:3 jako "čalba." Přijali jsme signály "čalba." Jaké je pravděpodobnost, že jsou významové výstup signálu "čalba"?

A: Jde o výstup signálu "čalba.

P(A₁) = \frac{5}{8}

P(A₂) = \frac{3}{8}

B: Jde o výstup signálu "čalba.

P(B|A₁) = \frac{3}{5}

P(B|A₂) = \frac{1}{3}

\begin{align*}
P(A₁|B) &= \frac{P(B|A₁)P(A₁)}{P(B|A₁)P(A₁) + P(B|A₂)P(A₂)} \\
&= \frac{\frac{3}{5} \cdot \frac{5}{8}}{\frac{3}{5} \cdot \frac{5}{8} + \frac{1}{3} \cdot \frac{3}{8}} \\
&= \frac{\frac{3}{5}}{\frac{5}{4}} = \frac{3}{5} = 0.95
\end{align*}
Dalšie evidenčné skúšanie je schopné pozvedať 10 očiavníků a 2 klepte. Radež a sekundárne,

**B.** Obraz je losuje.

**C.** Odtaniek štúna, ale obraz je očiavník.

\[
P(A) = \frac{10}{12}
\]

\[
P(B) = \frac{2}{12}
\]

\[
P(c|A) = \frac{5}{6}, \quad P(c|B) = \frac{1}{6}
\]

\[
P(A|c) = \frac{P(c|A)P(A)}{P(c|A)P(A) + P(c|B)P(B)} = \frac{\frac{5}{6} \cdot \frac{10}{12}}{\frac{5}{6} \cdot \frac{10}{12} + \frac{1}{6} \cdot \frac{2}{12}} = \frac{5 \cdot 10}{5 \cdot 10 + 2} = \frac{50}{52} = 0,96
\]
V Bayesovských sítích se kombinují 3 základní typy propojení:

1) sériové

\[ A \rightarrow B \rightarrow C \]

A ovlivňuje B, B ovlivňuje C, A ovlivňuje C prostřednictvím B. To platí, když máme možnost přímo pozorovat jevy A nebo C. Pokud je známo B, tak jev A už neovlivní C podmíněně (jevy A a C budou nezávislé při daném B).


2) divergentní

\[ A \leftarrow \downarrow \rightarrow B \quad C \quad \cdots \quad X \]


Příklad: pochvili nějakým způsobem můžu vliv na délku vlasů vyšší postavy.

Ale délka vlasů a výška postavy u mužů jsou nezávislé (nemusí pravda, že by těžba vyšší musí měli delší vlasy).
Príklad: A. množství benzinu v nádrži auto učer
B. — — dnes
C. — — zítra

S autem se jezdí, ale palivo se nepadává, tj. může jedině ubývat.

Vím-li, že dnes je nádrž plná, pak to, kolik bude benzinu zítra, už nezaisťuje tam, kolik ho bylo učera.
3) konvergentní

\[ \begin{array}{cc}
B & C \\
\downarrow & \downarrow \\
& A \\
\end{array} \]

Tady je to obražené - dobud nemí nic zmíno o A, jsou jený
B, C, ..., X neznáme.

Příklad: dobud nevím, jestli má člověk žaludeční potíže nebo ne, tak informace o tom, že nemá salmonelózu, nic neříka o možnosti, že má chřipky. Jakmile ale vím, že má žaludeční potíže a nemá salmonelózu, tak pravděpodobnost chřipky vzrůstá.

Totež ale bude platit, když místo nevolnosti budu evidovat bledost.

Seznažíci v tomto případě nemusí být přímo jen A, ale i jeho následník.

**Pojem evidence** - je to událost určitosti stavů náhodné proměnné

"hard evidence" - tu mám v případě, že proměnnou přímo pozorují a její stav zjistím (např. že pozorovaná osoba je muž)

"soft evidence" - ostatní případy
Další pojmy: certainty - jistota, určitost
belief - víra

(omo se to reálně formálně nedefinuje, spíš intuitivně a z kontextu je zřejmé, o co jde)

Jestliže např. proměnná A může mít stavy \( a_1, \ldots, a_n \) s pravděpodobnostmi \( x_1, \ldots, x_n \) a evidence říka, že momentálně může být jen ve stavech \( a_i \) nebo \( a_j \), pak

\[
P(A) = P(a_1, \ldots, a_n) = (x_1, \ldots, x_n)
\]

\[
P(A, e) = (0, \ldots, 0, x_i, 0, \ldots, 0, 0 \ldots 0) =
\]

\[= P(A) \times (0, \ldots, 1, 0, \ldots, 1, 0 \ldots 0) = P(A) \cdot e
\]

**finding (oteze ?)**

Příklady kombinací propojení:

a)
e u proměnných B a H značí, že jiné je přirozena evidence
(tj. jejich stavy jsou pozorovány, tedy pevné)
které stavy v síti může ovlivnit proměnná A a které ne
(tj. od kterých proměnných je A d-separováno ?)

Propojení \( A \rightarrow B \rightarrow E \) je sériové a B je pevné, tj.
toto spojení je blokováno a A nemůže ovlivnit E.

Ale: A je propojeno s \( D \rightarrow H \rightarrow K \)

propojení \( H \leftarrow K \rightarrow L \) je konvergentní, a protože
přírůst potomku K, tj. H je pevné, pak H je propojeno
s I a tudy i se zbytek síť kromě G
(c-E-I je sériové a není blokováno, tedy E-C-F-J-L,
ale E-G-J je konvergentní a J není pevné)

Přestože všechny proměnné v oboli E jsou pevné,
je E propojeno a F
(H je pevné a propojení E-H-F je konvergentní), tedy i s B
a taky s A (D je pevné,
propojení A-D-B je konvergentní)
Tady je F separované od A, E, G

Terminologie: luky proměnné A, B nejsou d-separované, nazývají se d- propojené
d-separované proměnné se nadý nazývají strukturování: nezávisle

 Platí: jsou-li A a B d-separované, pak změna "jistoty"
A nemá vliv na "jistotu" B.

Jsou-li A a B d-propojené, pak změna "viry" v A
nemusí nutně změnit "viry" v B.