EVROPSKÁ UNIE
Evropské strukturální a investiční fondy
Operační program Výzkum, vývoj a vzdělávání

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

Název projektu	Rozvoj vzdělávání na Slezské univerzitě v Opavě
Registrační číslo projektu	CZ.02.2.69/0.0./0.0/16_015/0002400

Expertní systémy

Neurčitost – Bayesovské sítě

Jan Górecki

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Neurčitost (opakování)

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

- Neurčitost je charakteristickým rysem složitých systémů. Vlastní povaha reality způsobuje, že poznatky, které z ní získáváme, jsou neurčité či vágní.

Příčiny neurčitosti (opakování)

– problémy s daty; např.:
  - chybějící nebo nedostupná data
  - nespolehlivá data (např. z důvodu chyb měření)
  - nepřesná nebo nekonzistentní reprezentace dat

Příčiny neurčitosti (opakování)

– nejisté znalosti; např.:
  - znalost nemusí být platná ve všech případech
  - znalost může obsahovat vágní pojmy.

Bayesovské sítě

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

- ***Bayesovská síť*** (*Bayesian belief network*) je orientovaný acyklický graf, jehož uzlům odpovídají náhodné proměnné a vazby reprezentují kauzální závislosti mezi těmito proměnnými.

Bayesovské sítě

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

- Hrana $X \rightarrow Y$ znamená, že X kauzálně ovlivňuje Y (pozorování X poskytuje kauzální podporu Y, pozorování Y poskytuje diagnostickou podporu pro X). Bayesovská síť umožňuje provádět prediktivní i diagnostické inference.

Bayesovské sítě

- Každému uzlu je přiřazena tabulka rozdělení pravděpodobnosti. Jestliže uzel nemá žádné předchůdce (rodiče), jedná se o nepodmíněnou pravděpodobnost, v opačném případě jde o podmíněnou pravděpodobnost.

Pojmy potřebné pro definici bayesovské sítě

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Nechť $G = (V, E)$ je orientovaný acyklický graf a nechť $v \in V$.

Definujme následující množiny:

$$C(v) = \{u \in V \mid (u, v) \in E\},$$

$$D(v) = \{w \in V \mid \text{ existuje cesta z } v \text{ do } w\},$$

$$A(v) = \{x \in V \mid x \neq v \text{ a } x \notin C(v) \cup D(v)\}.$$

Množina $C(v)$ je množinou bezprostředních předchůdců (rodičů) uzlu v, $D(v)$ je množinou všech následníků uzlu v (nejen bezprostředních). V případě bayesovské sítě prvky $C(v)$ nazýváme také *příčinami*.

Definice bayesovské sítě

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Nechť (Ω, P) je pravděpodobnostní prostor, kde $\Omega = \Omega_1 \times \ldots \times \Omega_n$, a nechť X_i $(i = 1, \ldots, n)$ je projekce Ω na Ω_i (tj. $X_i : \Omega \to \Omega_i$ je náhodná proměnná). Nechť (V, E) je orientovaný acyklický graf, kde $V = \{X_1, \ldots, X_n\}$.

Definice bayesovské sítě

Řekneme, že (V, E, P) je ***bayesovská síť'***, jestliže pro všechna $X_i \in V$ a všechna $W \subseteq A(X_i)$ jsou X_i a W podmíněně nezávislé při daném $C(X_i)$. To znamená, že když $W = \{Y_1, \ldots, Y_k\}$, $C(X_i) = \{Z_1, \ldots, Z_m\}$,

$P(Y_1 \wedge \ldots \wedge Y_k \wedge Z_1 \wedge \ldots \wedge Z_m) \neq 0$, pak

$$P(X_i \mid Y_1 \wedge \ldots \wedge Y_k \wedge Z_1 \wedge \ldots \wedge Z_m) = P(X_i \mid Z_1 \wedge \ldots \wedge Z_m).$$

Definice bayesovské sítě

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Zjednodušeně řečeno, když známe příčiny X_i, nic jiného než X_i samotné nebo jeho následníci nám nemůže dát nějaké další informace o X_i. Místo pojmu bayesovská síť se někdy užívají pojmy *kauzální síť* či *influenční diagram*.

Poznámky k symbolice

Nechť X je náhodná proměnná s oborem hodnot $O(X)$ a P je pravděpodobnost. Symbolem $P(X)$ rozumíme funkci definovanou na $O(X)$ tak, že pro x $\in$ $O(X)$ je $P(x) = P(X = x)$.

Poznámky k symbolice

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Zápisy pravděpodobností můžeme dále zjednodušit následujícím způsobem. Nechť $V = \{X_1, \dots, X_n\}$, $C(X_i) = \{Z_1, \dots, Z_m\}$. Pak definujeme

$$P(V) = P(\{X_1\} \cup \dots \cup \{X_n\}) = P(X_1, \dots, X_n) = P(X_1 \wedge \dots \wedge X_n),$$

$$P(X_i \mid C(X_i)) = P(X_i \mid Z_1, \dots, Z_m) = P(X_i \mid Z_1 \wedge \dots \wedge Z_m).$$

Poznámky k symbolice

Pomocí uvedené symboliky můžeme definici bayesovské sítě přepsat takto:
Řekneme, že (V, E, P) je bayesovská síť, jestliže pro všechna

$X_i \in V$ a všechna $W \subseteq A(X_i)$, $P(W \cup C(X_i)) \neq 0$, platí, že

$P(X_i \mid W \cup C(X_i)) = P(X_i \mid C(X_i))$.

Vlastnosti bayesovské sítě

Nechť (V, E, P) je bayesovská síť. Pak platí:

$$P(V) = \prod_{\substack{X \in V \\ P(C(X)) \neq 0}} P\big(X \mid C(X)\big)$$

Vlastnosti bayesovské sítě

Nechť (V, E) je orientovaný acyklický graf, kde $V = \{X_1, \ldots, X_n\}$, přičemž X_i jsou proměnné s obory hodnot $O(X_i)$. Nechť $f(X|\, C(X))$ je nezáporná reálná funkce taková, že $\sum_{x \in O(X)} f(x \mid C(X)) = 1$ pro všechny kombinace hodnot proměnných z $C(X)$. Pak

$\Omega = O(X_1) \times \ldots \times O(X_n)$ a $P(V) = \prod_{X \in V} f(X \mid C(X))$

definují pravděpodobnostní prostor, pro nějž (V, E, P) je bayesovská síť. Přitom $P(X|\, C(X))$ je buď 0 nebo $f(X \mid C(X))$.

Konstrukce bayesovské sítě

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

1. Specifikujeme veličiny $X_1, \dots , X_n$ a jejich obory hodnot $O(X_i)$.

Konstrukce bayesovské sítě

1. Specifikujeme veličiny $X_1, \dots , X_n$ a jejich obory hodnot $O(X_i)$.

Konstrukce bayesovské sítě

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

2. Zkonstruujeme orientovaný acyklický graf (V, E), kde $V = \{X_1, \ldots, X_n\}$, vyjadřující kauzální závislosti mezi veličinami.

Konstrukce bayesovské sítě

3. Odhadneme pravděpodobnost P tak, že odhadneme $P(X|C(X))$ pro všechna X, všechny hodnoty X a všechny kombinace hodnot proměnných z $C(X)$. Podle předchozí věty je nezbytné splnění pouze těchto podmínek:

$$0 \leq P\left(X \mid C(X)\right) \leq 1$$

$$\sum_{x \in O(X)} P\left(x \mid C(X)\right) = 1$$

Problém pro bayesovskou síť

Problém řešený pomocí bayesovské sítě je možno zjednodušeně formulovat takto:

Problém pro bayesovskou síť

Nechť je dána bayesovská síť (V, E, P) a množiny $U \subseteq V$, $W \subseteq V$, $U \cap W = \emptyset$. Jsou-li zadány hodnoty proměnných z množiny U, je třeba zjistit $P(W \mid U)$.

Problém pro bayesovskou síť

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Po zadání hodnot některých proměnných se provádí *inference*, což znamená přepočet podmíněných pravděpodobností pro ostatní proměnné. Inference v bayesovské síti je založena na Bayesových vzorcích.

Problém pro bayesovskou síť

Obecně je výše zmíněný problém NP-těžký, což znamená, že pro něj neexistují algoritmy s polynomiální časovou složitostí. Pokud bayesovská síť nemá speciální strukturu, je nutné použít aproximační techniky, které jsou obvykle založeny na nějakých transformacích bayesovské sítě na jednodušší tvar.

Jednoduše souvislá bayesovská síť

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Řekneme, že bayesovská síť je ***jednoduše souvislá***, jestliže mezi každými dvěma uzly existuje právě jedna neorientovaná cesta.

Jednoduše souvislá bayesovská síť

Jednoduše souvislá síť se také nazývá *polystrom* nebo *les*. Zvláštním případem polystromu je *strom*, což je graf, kde každý uzel má nejvýše jednoho rodiče.

Jednoduše souvislá bayesovská síť

SLEZSKÁ UNIVERZITA
OBCHODNĚ PODNIKATELSKÁ FAKULTA V KARVINÉ

Pro polystromovou bayesovskou síť existují algoritmy pro inferenci, které mají polynomiální časovou složitost.

Typické použití Bayesovských sítí

- pro modelování a vysvětlení chování problémů v různých oblastech např. model chování vody v krajině

Typické použití Bayesovských sítí

- pro nalezení nejpravděpodobnějších konfigurací proměnných – např. při automatickém rozpoznávání řeči nebo při dekódování zakódovaných zpráv

Typické použití Bayesovských sítí

- pro podporu rozhodování při nejisté informaci - použití teorie maximalizace očekávaného užitku

Typické použití Bayesovských sítí

- pro nalezení dobrých strategií pro řešení problémů v oblastech s nejistotou - např. technická diagnostika laserových tiskáren, nebo návrh adaptivních testů.

Děkuji za pozornost

Některé snímky převzaty od:

RNDr. Jiří Dvořák, CSc. dvorak@fme.vutbr.cz