

17. Probabilistički grafički modeli

Strojno učenje 1, UNIZG FER, ak. god. 2021./2022.

Jan Šnajder, vježbe, v2.0

1 Zadatci za učenje

1. [Svrha: Razumjeti što je to probabilistički grafički model. Shvatiti specifičnosti modela Bayesove mreže te kako taj model predstavlja zajedničku distribuciju. Shvatiti koje induktivne pretpostavke ovakva reprezentacija koristi.]

- Navedite tri osnovna aspekta svakog probabilističkog grafičkog modela (PGM).
- Je li PGM parametarski ili neparametarski model? Je li generativni ili diskriminativni? Obrazložite odgovore.
- Pretpostavite zajedničku distribuciju četiriju varijabli $p(x, y, w, z)$. Faktorizirajte ovu distribuciju primjenom osnovnih pravila vjerojatnosti te skicirajte Bayesovu mrežu koja odgovara toj faktorizaciji. Topološki uređaj uzmite da je x, y, w, z .
- Ponovite isto, ali ovaj put pretpostavljajući $y \perp w | x$ i $x \perp z | y, w$. Kojoj vrsti induktivne pristranosti odgovaraju ove pretpostavke o nezavisnosti? Obrazložite motivaciju za uvođenjem dodatnih pretpostavki u model.
- Formalno definirajte uređajno Markovljevo svojstvo i topološki uređaj čvorova mreže. Primjenom uređajnog Markovljevog svojstva izvedite uvjetne nezavisnosti kodirane Bayesovom mrežom koja odgovara faktorizaciji

$$P(x, y, w, z) = P(x)P(y|x, z)P(z)P(w|y).$$

- Nacrtajte Bayesovu mrežu Skrivenog Markovljevog modela (HMM) i napišite pripadnu faktorizaciju zajedničke vjerojatnosti $p(\mathbf{x}, \mathbf{z})$. Koje je svrha latentnih varijabli \mathbf{z} i koje su uvjetne nezavisnosti kodirane ovom mrežom?

2. [Svrha: Izvježbati iščitavanje Bayesove mreže i uvjetnih nezavisnosti iz zadane faktorizacije zajedničke vjerojatnosti. Razumjeti kako uvjetne nezavisnosti, broj varijabli i njihovih vrijednosti određuju ukupan broj parametara Bayesove mreže.] Gradimo Bayesovu mrežu koja predviđa hoće li student/ica uspješno položiti SU. Mreža sadrži pet varijabli: pohađa li osoba konzultacije (x_1), je li osoba dobra u Pythonu (x_2), rješava li osoba samostalno domaće zadaće i laboratorijske vježbe (x_3), ocjenu iz predmeta UI (x_4) te varijablu koja govori je li osoba položila SU (y). Pritom vrijedi $x_1, x_2, x_3, x_5 \in \{\top, \perp\}$ i $x_4 \in \{2, 3, 4, 5\}$. Vrijedi $P(x_1 = \top) = 0.2$ i $P(x_2 = \top) = 0.6$. Dane su i tablice uvjetnih vjerojatnosti.

x_1	x_2	$P(x_3 = \top)$	x_3	$P(y = \top)$
\perp	\perp	0.3	\perp	0.2
\perp	\top	0.5	\top	0.9
\top	\perp	0.8		
\top	\top	0.9		

x_2	$P(x_4 = 2)$	$P(x_4 = 3)$	$P(x_4 = 4)$	$P(x_4 = 5)$
\perp	0.4	0.2	0.3	0.1
\top	0.2	0.1	0.1	0.6

- (a) Skicirajte Bayesovu mrežu ako je faktorizacija zajedničke distribucije sljedeća:

$$P(x_1, x_2, x_3, x_4, y) = P(x_1)P(x_2)P(x_3|x_1, x_2)P(x_4|x_2)P(y|x_3).$$

- (b) Koji je ukupan broj parametara ove mreže?
 (c) Koje su uvjetne nezavisnosti kodirane u strukturu ove mreže?
3. [Svrha: Razumjeti ideju d-odvajanja i kako se ona može provesti grafički. Shvatiti motivaciju iza ispitivanja uvjetne nezavisnosti parova varijabli.]
- (a) Zašto bismo htjeli znati koji parovi varijabli su uvjetno nezavisni? Nije li ta informacija već kodirana unutar strukture mreže? Objasnite.
 (b) Formalno definirajte d-odvajanja i objasnite koji uvjeti (i kada) moraju vrijediti da bi neke dvije varijable bile uvjetno nezavisne.
 (c) Na temelju Bayesove mreže iz zadatka 2, odredite pod kojim uvjetima su varijable prolaza SU (y) i ocjene iz predmeta UI (x_4) uvjetno nezavisne.
 (d) Svojim riječima objasnite efekt objašnjavanja (engl. *explaining away*) koristeći za primjer varijable x_1 , x_2 i x_3 .

4. [Svrha: Izvežbati iščitavanje uvjetnih nezavisnosti iz Bayesove mreže te određivanje (ne)zavisnosti proizvoljnog para varijabli primjenom pravila d-odvajanja.] Bayesovom mrežom modeliramo vjerojatnost oboljenja od kardiovaskularnih bolesti. Mreža sadrži četiri varijable: spol osobe (S), koliko često osoba tjedno odlazi u teretanu (T), je li osoba pušač (P) te varijablu koja govori o kakvom se riziku radi (R). Pritom vrijedi $s \in \{\text{muški, ženski}\}$, $p \in \{\perp, \top\}$, $t \in \{1, 3, 5\}$ i $r \in \{\text{nizak, umjeren, visok}\}$. Zajednička razdioba faktorizirana je kao:

$$P(S, T, P, R) = P(S)P(T)P(P|S, T)P(R|P)$$

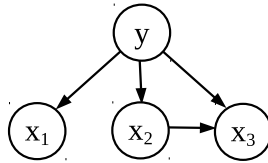
- (a) Skicirajte Bayesovu mrežu koja predstavlja izvedenu faktorizaciju. Primjenom uređajnoga Markovljevog svojstva izvedite pretpostavke uvjetne nezavisnosti varijabli koje su ugrađene u strukturu Bayesove mreže.
 (b) Koristeći pravila d-odvajanja, odredite pod kojim uvjetima su varijable x_1 i x_2 uvjetno nezavisne. Kako nam ta informacija može biti od koristi?

2 Zadatci s ispita

1. (T) Za Bayesovu mrežu kažemo da je generativni i parametarski model. **Zašto?**
- A Generativni jer definira zajedničku vjerojatnost svih varijabli, i opaženih i skrivenih, a parametarski jer se parametri modela mogu dobiti MLE-procjenom za svaki čvor Bayesove mreže zasebno, budući da se log-izglednost dekomponira po strukturi mreže
- B Generativni jer se može koristiti za generiranje skupa primjera na temelju zajedničke distribucije, a parametarski jer su broj čvorova mreže i njihovo povezivanje (dakle graf) definirani parametrima koji se mogu ugađati na skupu za učenje, čime se mogu dobiti različite strukture mreže
- C Generativni jer svaki čvor odgovara uvjetnoj vjerojatnosti koja je, na temelju Markovljevog uređajnog svojstva, generirana distribucijama čvorova roditelja, a parametarski jer Bayesova mreža zapravo definira zajedničku distribuciju koja je opisana skupom parametara
- D Generativni jer opisuje postupak kojim se mogu generirati podaci koji se pokoravaju određenoj zajedničkoj vjerojatnosnoj distribuciji, a parametarski jer svaki čvor Bayesove mreže definira uvjetnu vjerojatnost preko teorijske distribucije koja je opisana svojim parametrima
2. (T) Bayesove mreže na sažet način prikazuju zajedničku distribuciju te kodiraju uvjetne stohastičke nezavisnosti između varijabli. No, kao i svaki model strojnog učenja, tako se i Bayesove mreže

možu prenaučiti. **Koja je veza između uvjetnih nezavisnosti varijabli u Bayesovoj mreži i opasnosti od prenaučivosti?**

- A Uvođenje pretpostavki o uvjetnoj nezavisnosti pojednostavljuje strukturu Bayesove mreže i smanjuje broj parametara, čime se smanjuje i mogućnost prenaučivosti
 - B Uvođenjem pretpostavki o uvjetnoj nezavisnosti povećava se broj čvorova mreže, a time i broj parametara, što model čini složenijim i time sklonijim prenaučivosti
 - C Uvjetne nezavisnosti određuju strukturu mreže na način da definiraju koji su čvorovi mreže međusobno povezani, međutim to nema utjecaja na složenost modela niti na sklonost prenaučivosti
 - D Pretpostavke o uvjetnoj nezavisnosti čine induktivnu pristranost modela, pa što je više uvjetnih nezavisnosti, to je veća pristranost i model je lako prenaučiti
3. (P) Na slici ispod prikazana je Bayesova mreža koja odgovara polunajvnom Bayesovom klasifikatoru. Pretpostavite da su značajke x_1 , x_2 i x_3 binarne varijable te da je oznaka klase y također binarna varijabla. Označimo ovaj model sa \mathcal{H}_2 . Model \mathcal{H}_2 može se pojednostaviti ako se ukloni brid između varijabli x_2 i x_3 . Označimo takav model sa \mathcal{H}_1 . S druge strane, od modela \mathcal{H}_2 može se napraviti još složeniji model koji odgovara potpuno povezanom acikličkom grafu. Označimo takav model sa \mathcal{H}_3 .



Razmotrite koliko parametara imaju modeli \mathcal{H}_1 , \mathcal{H}_2 i \mathcal{H}_3 . **Koliko model \mathcal{H}_2 ima više parametara od modela \mathcal{H}_1 , a koliko manje parametara od modela \mathcal{H}_3 ?**

- A 2 više, 3 manje
 - B 2 više, 6 manje
 - C 4 više, 4 manje
 - D 4 više, 8 manje
4. (P) Razmotrite Bayesovu mrežu koja zajedničku vjerojatnost faktorizira na sljedeći način:

$$P(w, x, y, z) = P(w)P(y)P(x|w, y)P(z|w)$$

Odredite topološki uređaj varijabli. Ako postoji više mogućih topoloških uređaja, izaberite onaj koji po leksičkom poretku dolazi prvi (npr. x, y, z dolazi prije x, z, y). Zatim primijenite uređajno Markovljevo svojstvo te izvedite sve uvjetne nezavisnosti koje su kodirane u ovoj Bayesovoj mreži. **Koje sve uvjetne nezavisnosti vrijede u ovoj Bayesovoj mreži?**

- A $w \perp y, z \perp \{x, y\} | w$
 - B $x \perp y | z, z \perp w | y$
 - C $w \perp y, z \perp x, x \perp w | \{z, y\}$
 - D $y \perp w, y \perp x | \{w, z\}$
5. (P) Bayesova mreža ima pet varijabli, od kojih su v , w i z binarne, a x i y ternarne varijable. Topološki uređaj varijabli neka je v, w, x, y, z . Uz takav uređaj, u mreži vrijede sljedeće marginalne i uvjetne nezavisnosti:

$$v \perp w \quad w \perp x | v \quad v \perp y | \{w, x\} \quad \{v, w\} \perp z | \{x, y\}$$

Izvedite faktorizaciju zajedničke distribucije koja odgovara ovoj Bayesovoj mreži. **Koliko parametara ima dotična Bayesova mreža?**

- A 10
 - B 22
 - C 25
 - D 27
6. (P) Bayesovom mrežom s pet binarnih varijabli modeliramo prometne prilike u gradu Zagrebu. U našoj mreži, jutarnje doba dana (J) i loše vrijeme (V) utječu na nastanak prometne gužve (G), u smislu da oba događaja povećavaju vjerojatnost nastanka prometne gužve. Loše vrijeme također utječe na nastupanje prometne nesreće (N), u smislu da povećava vjerojatnost prometne nesreće. Nadalje, nastupanje prometne nesreće utječe na nastanak prometne gužve, u smislu da povećava vjerojatnost nastanka prometne gužve. Loše vrijeme također utječe na zastoj tramvaja (T), u smislu da povećava vjerojatnost zastoja tramvaja. Međutim, nestanak struje (S) također

uzorkuje zastoj tramvaja. Konačno, zastoj tramvaja uzrokuje masovno pješačenje putnika (P), što opet povećava vjerojatnost prometne nesreće. U ovom kauzalnom modelu može nastupiti efekt objašnjavanja. **Kako bi se efekt objašnjavanja konkretno manifestirao?**

A $P(V = 1|P = 1, T = 1) < P(V = 1|T = 1)$

B $P(G = 1|J = 1, V = 1) > P(G = 1|J = 1)$

C $P(V = 1|P = 1, N = 1) < P(V = 1|N = 1)$

D $P(T = 1|V = 1, P = 1) > P(T = 1|V = 1)$