

18. Probabilistički grafički modeli II

Strojno učenje 1, UNIZG FER, ak. god. 2023./2024.

Jan Šnajder, vježbe, v2.3

1 Zadaci za učenje

1. [Svrha: Razumjeti i izvježbati egzaktno zaključivanje kod Bayesovih mreža. Postati svjestan složenosti egzaktnog zaključivanja.] Skicirajte Bayesovu mrežu iz zadatka 2 iz cjeline 17. Parametri modele neka su sljedeći. Za čvorove x_1 i x_2 parametri su $P(x_1 = \top) = 0.2$ i $P(x_2 = \top) = 0.6$. Tablice uvjetnih vjerojatnosti za preostale čvorove su:

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

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

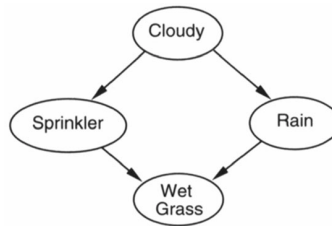
- (a) Postupkom egzaktnog zaključivanja izračunajte $P(y = \top | x_1 = \top, x_4 = 3)$.
- (b) Koja je razlika između posteriornog i MAP-upita? O kakvom tipu upita se radi u prošlom zadatku? Obrazložite.
- (c) Utječe li broj varijabli u mreži na učinkovitost zaključivanja? Zašto?
- (d) Objasnite ideju približnog zaključivanja uzorkovanjem. Koja je prednost tog postupka? U kratkim crtama objasnite kako biste uzorkovali $P(x_1, x_2, x_3, x_4, y)$ koristeći unaprijedno uzorkovanje (engl. *forward sampling*).
2. [Svrha: Razumjeti učenje Bayesovih mreža i njegovu povezanost s procjenom parametara. Znati kako pristupiti učenju modela ako su podatci nepotpuni.]
- (a) Što su parametri Bayesove mreže i na koji način ih učimo iz podataka?
- (b) Izvedite log-izglednost (proizvoljne) Bayesove mreže. Objasnite zašto je moguće procjenjivati parametre svakog čvora mreže zasebno.
- (c) Objasnite što to znači da neki model ima skrivene (latentne) varijable. Kako one utječu na postupak učenja modela?
3. [Svrha: Izvježbati procjenu parametara čvora Bayesove mreže na temelju zadanog skupa podataka. Izvježbati kako napisati izraz za egzaktno zaključivanje na temelju konkretne Bayesove mreže. Razumijeti prednosti i nedostatke egzaktnog zaključivanja naspram metoda uzorkovanja.] Skicirajte Bayesovu mrežu iz zadatka 4 iz cjeline 17. Parametre te mreže procjenjujemo na sljedećem skupu podataka:
- (a) Primjenom (Laplaceovog) MAP-procjenitelja procijenite $P(P|S, T)$.

S	P	T	R
<i>ženski</i>	⊤	1	<i>visok</i>
<i>ženski</i>	⊤	5	<i>umjeren</i>
<i>muški</i>	⊥	3	<i>nizak</i>
<i>ženski</i>	⊥	1	<i>umjeren</i>
<i>muški</i>	⊤	5	<i>nizak</i>
<i>ženski</i>	⊥	1	<i>nizak</i>

- (b) Korištenjem egzaktnog zaključivanja izvedite izraz za vjerojatnost visokog rizika oboljenja osobe koja je pušač i posjećuje teretanu pet puta tjedno. Za svaku od četiri varijable naznačite radi li se o varijabli upita, opaženoj varijabli ili varijabli smetnje.
- (c) Na ovoj mreži ilustrirajte prednosti i nedostatke metoda uzorkovanja nad metodom egzaktnog zaključivanja.
- (d) Na ovoj mreži ilustrirajte nedostatak unaprijednog uzorkovanja. Što su alternative unaprijednom uzorkovanju?

2 Zadatci s ispita

1. (N) Na slici ispod prikazana je Bayesova mreža za problem prskalice za travu, koji smo bili koristili na predavanjima. Varijable su: C (oblačno/*cloudy*), S (prskalice/*sprinkler*), R (kiša/*rain*) i W (mokra trava/*wet grass*). Dane su i tablice uvjetnih vjerojatnosti za svaki čvor.



C	$P(C)$	S	C	$P(S C)$	R	C	$P(R C)$	W	R	S	$P(W R, S)$
0	0.5	0	0	0.5	0	0	0.8	0	0	0	1.0
0	0.5	0	1	0.9	0	1	0.2	0	0	1	0.9
1	0.5	1	0	0.5	1	0	0.2	0	1	0	0.1
		1	1	0.1	1	1	0.8	0	1	1	0.01
								1	0	0	0.0
								1	0	1	0.1
								1	1	0	0.9
								1	1	1	0.99

Izračunajte aposteriornu vjerojatnost da pada kiša ako je trava mokra i nije oblačno.

☐ A 0.112 ☐ B 0.491 ☐ C 0.709 ☐ D 0.825

2. (N) Bayesovom mrežom s četiri varijable modeliramo konstrukte pozitivne psihologije. Koristimo binarne varijable *Ljubav* (L), *Sreća* (S), *Tjeskoba* (T), s vrijednostima 0 (nema) i 1 (ima), te ternarnu varijablu *Novac* (N), s vrijednostima 0 (nema), 1 (ima malo) i 2 (ima puno). Strukturu Bayesove mreže definirali smo tako da ona modelira sljedeće pretpostavljene kauzalne odnose: L uzrokuje S, a N uzrokuje S i T. Tako definiranu Bayesovu mrežu zatim treniramo na sljedećem skupu od $N = 7$ primjera:

L	N	S	T
1	0	1	0
1	0	1	0
0	2	0	1
1	2	1	1
1	1	1	0
0	0	0	0
0	2	1	0

Parametre modela procjenjujemo MAP-procjeniteljem sa $\alpha = \beta = 2$ (za binarne varijable) odnosno $\alpha_k = 2$ (za ternarnu varijablu), što je istovjetno Laplaceovom zaglađivanju MLE procjene. Na kraju nas, naravno, zanima koja je vjerojatnost života uz ljubav, sreću i malo novaca. Napravite potrebne MAP-procjene parametara. **Koliko iznosi zajednička vjerojatnost $P(L = 1, S = 1, N = 1)$?**

- ☐ A 0.023 ☐ B 0.074 ☐ C 0.143 ☐ D 0.833

3. (P) Razmotrite jednostavnu Bayesovu mrežu koja odgovara faktorizaciji $P(x, y, z) = P(x)P(y)P(z|x, y)$. Sve varijable su binarne. Vrijedi $P(x = 1) = 0.2$ i $P(y = 1) = 0.3$. Tablica uvjetne vjerojatnosti za čvor z je sljedeća:

z	x	y	$p(z x, y)$	z	x	y	$p(z x, y)$
0	0	0	0.1	1	0	0	0.9
0	0	1	0.2	1	0	1	0.8
0	1	0	0.5	1	1	0	0.5
0	1	1	0.9	1	1	1	0.1

Postupkom uzorkovanja s odbijanjem uzorkujemo iz aposteriorne distribucije $P(y|x = 1, z = 0)$. Uzorkovanje smo ponovili ukupno $N = 1000$ puta. **Koja je očekivana veličina uzorka, odnosno koliko slučajnih vektora nećemo morati odbaciti?**

- ☐ A 54 ☐ B 124 ☐ C 200 ☐ D 739

4. (N) Bayesovu mrežu koristimo za medicinsku dijagnostiku te modeliramo sljedeće kauzalne odnose. Upala grla ($U = 1$) može biti uzrokovana virusom ($V = 1$) ili bakterijom ($B = 1$). Povišena temperatura ($T = 1$) može biti uzrokovana upalom grla ili sunčanicom ($S = 1$). Sve varijable su binarne. Na temelju podataka o pacijentima procijenili smo parametre mreže: $P(V = 1) = 0.3$, $P(B = 1) = 0.1$ i $P(S = 1) = 0.05$. Uvjetne vjerojatnosti za čvorove U i T su:

V	B	$P(U = 1 V, B)$	U	S	$P(T = 1 U, S)$
0	0	0.2	0	0	0
0	1	0.5	0	1	0.2
1	0	0.4	1	0	0.4
1	1	0.7	1	1	0.4

Zanima nas koje je najvjerojatnije objašnjenje izravnog uzroka povišene temperature u pacijenata kod kojih nije dokazano prisustvo virusa. U tu svrhu računamo MAP-upit za par varijabli upita U i S uz opažene varijable $V = 0$ i $T = 1$, tj. računamo $\operatorname{argmax}_{U, S} P(U, S|V = 0, T = 1)$. Neka je p_1 vjerojatnost najvjerojatnijeg (MAP) objašnjenja za varijable U i S , a p_2 vjerojatnost drugog po redu najvjerojatnijeg objašnjenja za te varijable. **Koliko je puta najvjerojatnije objašnjenje vjerojatnije od drugog najvjerojatnijeg objašnjenja, tj. koliko iznosi p_1/p_2 ?**

- ☐ A 11.35 ☐ B 13.35 ☐ C 15.52 ☐ D 17.88

5. (N) Razmotrite Bayesovu mrežu koja odgovara faktorizaciji $P(w, x, y, z) = P(w)P(x)P(y|w, x)P(z|x)$. Sve varijable su binarne. Vrijedi $P(w = 1) = 0.1$, $P(x = 1) = 0.2$, $P(z = 1|x = 0) = 0.9$ i $P(z = 1|x = 1) = 0.7$. Tablica uvjetnih vjerojatnosti za čvor y je sljedeća:

w	x	$p(y = 1 w, x)$
0	0	0
0	1	0.4
1	0	0.2
1	1	0.7

Postupkom uzorkovanja s odbijanjem želimo procijeniti parametar μ uvjetne distribucije $P(x = 0|y = 1, z = 0)$. Uzorkovanje smo ponovili ukupno $N = 100$ puta, od čega smo neke vektore morali odbaciti, pa je naš uzorak manji od N . Na temelju dobivenog uzorka parametar μ procjenjujemo MAP procjeniteljem uz $\alpha = \beta = 2$. **Koliko iznosi očekivana MAP procjena parametra μ ?**

- ☐ A 0.0490 ☐ B 0.0786 ☐ C 0.1274 ☐ D 0.1877