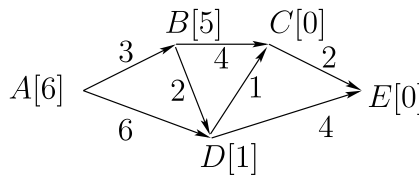
Mesterséges intelligencia minta ZH

2009.

1. Keresd meg a lenti gráfban a legrövidebb utat A-ból E-be az A* algoritmus segítségével!

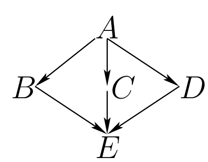


2. Az alábbi táblázatban 14 személy bizonyos adatai lettek összegyűjtve. Ezen minta alapján a Naiv Bayes módszert alkalmazva mit tippelsz: egy olyan illető, aki legfeljebb harminc éves, tanuló, közepes jövedelemmel rendelkezik és gyenge a hitelképessége, az inkább vesz számítógépet, vagy inkább nem?

kor	jövedelem	tanuló	hitelképesség	vesz számítógépet
≤ 30	magas	nem	jó	nem
≤ 30	magas	nem	kitűnő	nem
> 40	közepes	nem	kitűnő	nem
> 40	alacsony	igen	gyenge	nem
≤ 30	közepes	nem	jó	nem
31–40	alacsony	igen	kitűnő	igen
≤ 30	alacsony	igen	jó	igen
> 40	közepes	igen	jó	igen
≤ 30	közepes	igen	kitűnő	igen
≤ 30	közepes	nem	kitűnő	igen
31–40	magas	igen	jó	igen
31–40	magas	nem	kitűnő	igen
> 40	közepes	nem	jó	igen
> 40	alacsony	igen	gyenge	igen

3. Mi $P(B = i, C = h, D = i | A = h, E = i)$ értéke a lenti Bayes-hálóban?

x	y	z	$P(E = i \mid B = x, C = y, D = z)$
i	i	i	0
i	i	h	0
i	h	i	0,8
i	h	h	0
h	i	i	0
h	i	h	0,4
h	h	i	0,9
h	h	h	0



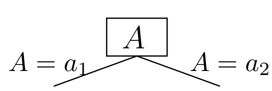
$P(A$=i$)$=0,8

	x=i	x=h
$P(B = i \mid A = x)$	0,4	0,7
$P(C = i \mid A = x)$	0,3	0,6
$P(D = i \mid A = x)$	0,2	0,7

4. Az ID3 futása közben a konstruálás alatt álló fa egyik csúcsánál a lenti ábrán látható döntési helyzet adódik — azaz a csúcs vagy A címkét kap, és ezáltal az a) eset áll elő, vagy B címkét kap, és a b) helyzet áll elő. Az entrópia helyett a $E(p) = 4p(1 - p)$ függvényt használva melyik attribútumot választja az ID3, és miért? (S a csúcshoz tartozó példák halmaza, ℓ első komponense pedig mindig a paraméterében lévő példahalmaz pozitív, a második pedig a negatív elemeinek száma.)

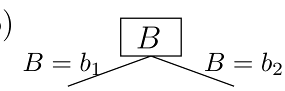
$$\ell(S) = (8, 12)$$

a)



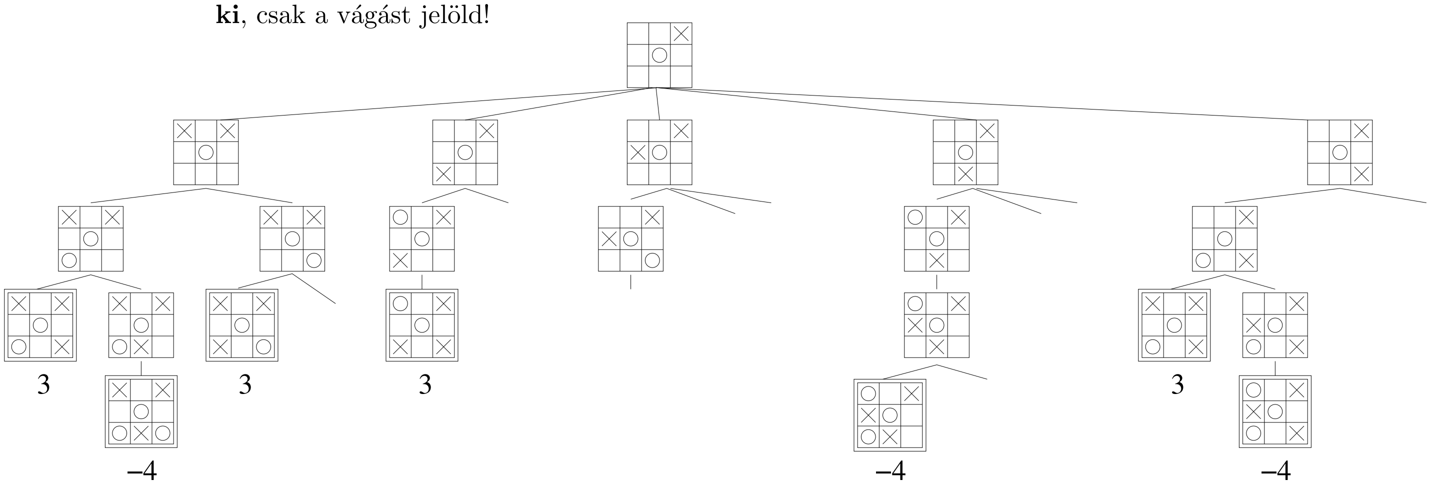
$\ell(S|_{A=a_1}) = (6, 7)$ $\quad \ell(S|_{A=a_2}) = (2, 5)$

b)



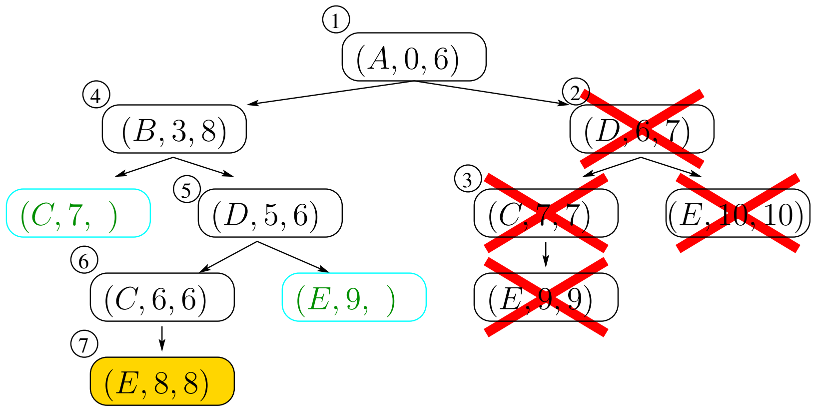
$\ell(S|_{B=b_1}) = (4, 2)$ $\quad \ell(S|_{B=b_2}) = (4, 10)$

5. Egészítsd ki a `min-max` algoritmussal a következő játék lenti ábrán látható játékfáját: $\times$ és $\circ$ felváltva raknak a 3×3-as tábla valamelyik mezőjére úgy, hogy $\times$ ne kerüljön $\times$ mellé, $\circ$ pedig $\circ$ mellé. A játékot $\times$ kezdi. A játék addig tart, míg egyikük már nem tud lépni; ekkor ő annyi \$-t ad ellenfelének, ahány lépés volt a játék során. Ahol alkalmazható α- vagy β-vágás, ott **a levágott részfákat ne dolgozd ki**, csak a vágást jelöld!



Megoldások

1. Jelölje a lenti keresőfa csúcsainak címkéiben az első komponens mindig a gráf megfelelő csúcsát, a második komponens a hozzá tartozó aktuális g', a harmadik pedig az f' értéket.



Ugyanez táblázatos formában megoldva:

Ny	Z
A(nil,0,6)	

Ny	Z
B(A,3,8)	A(nil,0,6)
D(A,6,7)	

Ny	Z
B(A,3,8)	A(nil,0,6)
E(D,10,10)	D(A,6,7)
C(D,7,7)	

Ny	Z
B(A,3,8)	A(nil,0,6)
E(C,9,9)	D(A,6,7)
	C(D,7,7)

Ny	Z
E(C,9,9)	A(nil,0,6)
D(B,5,6)	C(D,7,7)
	B(A,3,8)

Ny	Z
E(C,9,9)	A(nil,0,6)
C(D,6,6)	B(A,3,8)
	D(B,5,6)

Ny	Z
E(C,9,9)	A(nil,0,6)
	B(A,3,8)
	D(B,5,6)
	C(D,6,6)

Ny	Z
	A(nil,0,6)
	B(A,3,8)
	D(B,5,6)
	C(D,6,6)
	E(C,8,8)

A megadott gráfban tehát a 8 költségű A, B, D, C, E út az optimális.

2. Az adatok alapján a következő becsléseket írjuk fel:

$$\mathbb{P}[\text{vesz}] \cdot \mathbb{P}[\leq 30|\text{vesz}] \cdot \mathbb{P}[\text{tanuló}|\text{vesz}] \cdot \mathbb{P}[\text{közepes jöv.}|\text{vesz}] \cdot \mathbb{P}[\text{gyenge hit.}|\text{vesz}] \approx \frac{9}{14} \cdot \frac{3}{9} \cdot \frac{6}{9} \cdot \frac{4}{9} \cdot \frac{1}{9} = \frac{4}{567} \ ,$$

illetve

$$\mathbb{P}[\text{nem}] \cdot \mathbb{P}[\leq 30|\text{nem}] \cdot \mathbb{P}[\text{tanuló}|\text{nem}] \cdot \mathbb{P}[\text{közepes jöv.}|\text{nem}] \cdot \mathbb{P}[\text{gyenge hit.}|\text{nem}] \approx \frac{5}{14} \cdot \frac{3}{5} \cdot \frac{1}{5} \cdot \frac{2}{5} \cdot \frac{1}{5} = \frac{3}{875} \ ,$$

így a Naív-Bayes módszer szabályai szerint azt valószínűsítjük, hogy az illető vesz számítógépet.

3. Tekintve, hogy a feladatban megadott Bayes-hálóban bármely $a, b, c, d, e \in \{i, h\}$ esetén

$$\mathbb{P}[A = a, B = b, C = c, D = d, E = e] =$$

$$= \mathbb{P}[A = a] \cdot \mathbb{P}[B = b|A = a] \cdot \mathbb{P}[C = c|A = a] \cdot \mathbb{P}[D = d|A = a] \cdot \mathbb{P}[E = e|B = b, C = c, D = d] \ ,$$

ezért

$$\mathbb{P}[A = h, B = i, C = h, D = i, E = i] = 0,2 \cdot 0,7 \cdot 0,4 \cdot 0,7 \cdot 0,8 \ ,$$

továbbá

$$\begin{aligned}\mathbb{P}[A = h, E = i] =& \mathbb{P}[A = h, B = i, C = h, D = i, E = i] + \mathbb{P}[A = h, B = h, C = i, D = h, E = i] + \\ &+ \mathbb{P}[A = h, B = h, C = h, D = i, E = i] = \\ =& 0,2 \cdot 0,7 \cdot 0,4 \cdot 0,7 \cdot 0,8 + 0,2 \cdot 0,3 \cdot 0,6 \cdot 0,3 \cdot 0,4 + 0,2 \cdot 0,3 \cdot 0,4 \cdot 0,7 \cdot 0,9 \ ,\end{aligned}$$

ami alapján

$$\begin{aligned}\mathbb{P}[B=i,C=h,D=i|A=h,E=i]=&\frac{\mathbb{P}[A=h,B=i,C=h,D=i,E=i]}{\mathbb{P}[A=h,E=i]}=\\=&\frac{0,7\cdot 0,7\cdot 0,8}{0,7\cdot 0,7\cdot 0,8+0,3\cdot 0,6\cdot 0,3+0,3\cdot 0,7\cdot 0,9}=\\=&\frac{392}{392+54+189}\approx\\\approx&0,6173\ .\end{aligned}$$

4.

$$\begin{aligned}\mathrm{Gain}_E(S,A)=&E\left(\frac{8}{20}\right)-\frac{13}{20}\cdot E\left(\frac{6}{13}\right)-\frac{7}{20}\cdot E\left(\frac{2}{7}\right)=\\=&E\left(\frac{8}{20}\right)-\frac{13}{20}\cdot 4\cdot\frac{6}{13}\cdot\frac{7}{13}-\frac{7}{20}\cdot 4\cdot\frac{2}{7}\cdot\frac{5}{7}=\\=&E\left(\frac{8}{20}\right)-\frac{42}{65}-\frac{2}{7}\approx\\\approx&E\left(\frac{8}{20}\right)-0,9319\ ,\end{aligned}$$

$$\begin{aligned}\mathrm{Gain}_E(S,B)=&E\left(\frac{8}{20}\right)-\frac{6}{20}\cdot E\left(\frac{4}{6}\right)-\frac{14}{20}\cdot E\left(\frac{4}{14}\right)=\\=&E\left(\frac{8}{20}\right)-\frac{6}{20}\cdot 4\cdot\frac{4}{6}\cdot\frac{2}{6}-\frac{14}{20}\cdot 4\cdot\frac{4}{14}\cdot\frac{10}{14}=\\=&E\left(\frac{8}{20}\right)-\frac{4}{15}-\frac{4}{7}\approx\\\approx&E\left(\frac{8}{20}\right)-0,8381\ ,\end{aligned}$$

azaz $\mathrm{Gain}_E(S,A)<\mathrm{Gain}_E(S,B)$, így az ID3 a B attribútumot választja.

5.