

Intelligens orvosi műszerek

VIMIA023

A 2. zh előtti gyakorlás

2018 ősz

<http://www.mit.bme.hu/oktatas/targyak/vimia023>

dr. Pataki Béla

pataki@mit.bme.hu

(463-)2679

2. zh. előtti gyakorlás 1A

A következő állítások közül melyik hamis, melyik igaz?

- a. Az FN és a TN esetek összege az egészségesek összes számát adja meg. **a. Igaz Hamis**
- b. A ROC görbe analízisnél az egyes küszöbszintekhez tartozó valós negatív arány (TNR, specificitás) nem befolyásolja a görbe alakját. **b. Igaz Hamis**
- c. Ha egy csoportnál (pl. a 17-nél kisebb testtömegindexűek csoportja) nagyobb költségnövekedést okoznak a hamis pozitív diagnózisok, mint a hamis negatív diagnózisok, akkor ezt a csoportot egészségesnek célszerű tekinteni. **c. Igaz Hamis**
- d. A döntési eljárás költségkomponensei közül C_{00} mindig nagyobb, mint C_{11} . **d. Igaz Hamis**
- e. Ha egy csoportot (pl. a $90 < x < 100$ Hgmm diasztolés vérnyomásúak csoportja) betegnek tekintünk a diagnosztikai eljárás során, akkor a hamis pozitív diagnózisok okoznak ebben a csoportban költségnövekedést. **e. Igaz Hamis**

2. zh. előtti gyakorlás 1B

A következő állítások közül melyik hamis, melyik igaz?

f. Ha mindenkit betegnek tekintünk, akkor a specificitás 1 lesz.

f. **Igaz Hamis**

g. A ROC görbe analízisnél a véletlen találgatás kb. 0,5 görbe alatti területet eredményez.

g. **Igaz Hamis**

h. A döntésünk összköltségének helyes becslését alapvetően befolyásolja, hogy jól tudjuk-e becsülni az egyes esetek (TP, TN, FP, FN) költségét.

h. **Igaz Hamis**

i. Ha a döntési küszöböt változtatjuk, akkor a specificitás és a szenzitivitás általában ugyanabban az irányban változik (mindkettő nő vagy mindkettő csökken).

i. **Igaz Hamis**

j. Az ideális ROC görbe alatti terület 1.

j. **Igaz Hamis**

2. zh. előtti gyakorlás 2.

Egy z paramétert mérünk, amelynek 8 lehetséges értéke van, és ez alapján akarunk diagnosztizálni egy betegséget. A népesség 5%-a szenved ebben a betegségben. A szűrés költsége 100 bitcoin, a szűrés során tévesen betegnek tekintett emberek 1000 bitcoin költséget jelentenek (ebben már benne van a szűrés is, tehát $900+100$), a tévesen egészségesnek tekintett emberek 10^5 bitcoinba kerülnek (ebben is benne van a szűrés költsége), a helyesen diagnosztizáltak gyógyítása (szintén a szűréssel együtt) 10^4 bitcoin.

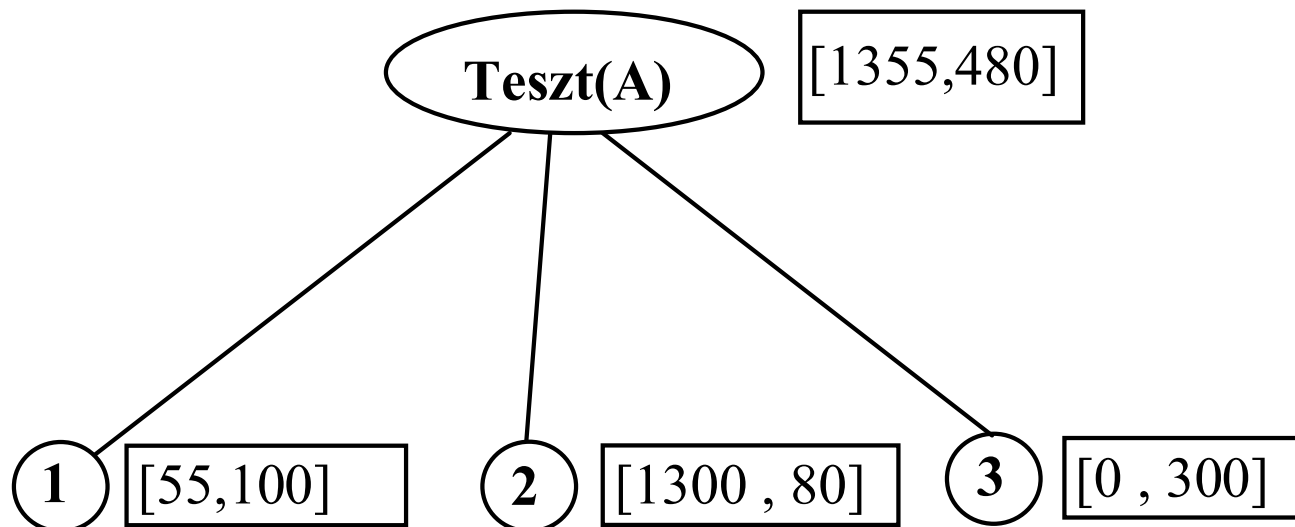
2A. A táblázat 4-es csoportjába tartozó embereket (akiknek a z értéke 39) hova érdemesebb sorolni, az egészségesek vagy a betegek közé?

2B. Ha a limit alatti z értékkel rendelkezőket tekintjük egészségesnek, akkor a 38,5 vagy a 39,5 limit esetén lesz kisebb a költség?

i	1	2	3	4	5	6	7	8
z_i	36	37	38	39	40	41	42	43
$P(z_i T0)$	0,8	0,1	0,05	0,01	0,01	0,01	0,01	0,01
$P(z_i T1)$	0,05	0,1	0,5	0,12	0,1	0,04	0,03	0,01

2. zh. előtti gyakorlás 3.

Az alábbi ábrán egy ismert besorolású minták alapján tanítással előállított döntési fa építésének első lépése látható. A csomópontok mellett, mindig tőlük jobbra, bekeretezve szögletes zárójelben látható két szám, ez a C1, illetve C2 osztályba jutó tanítóminták száma az adott csomópontban. Ebben az első lépésben az „A” attribútumot teszteljük, a tesztnek 3 kimenetele lehetséges. Mekkora lesz a teszt információnyeresége, ha az ábrán látható mintaarányok jól tükrözik az egyes esetek arányát a teszt előtt és után?



2. zh. előtti gyakorlás 4.

Az X betegségben a népesség 0,23%-a szenved. Ha a betegséget korai fázisban kezeljük, akkor a kezelés átlagos költsége 17.000 Ft/fő, ha csak a késői fázisában, akkor 910.000 Ft/fő.

Két lehetőség közt választhatunk:

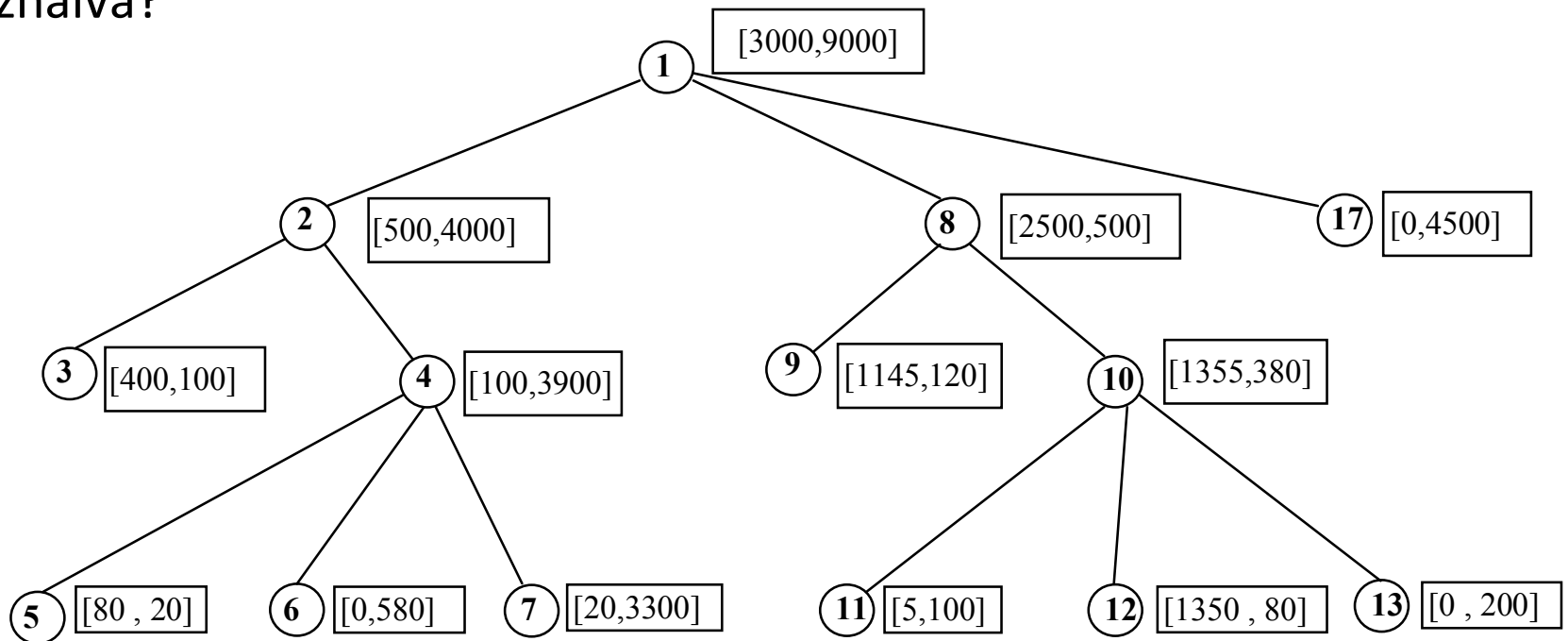
1. Szűrést végzünk, amelynek költsége 2.000 Ft fejenként, specificitása 0,88; szenzitivitása 0,96. Az itt betegnek vélelmezett embereknél (nyilván akár valóban betegek, akár nem, ezt nem tudhatjuk előre) egy további – átlagosan 5.000 Ft/fő költségű – vizsgálattal eldöntjük, hogy tényleg beteg-e, és csak a tényleg betegnek bizonyult embereket kezeljük.

2. Nem szűrünk, hanem mindenkit egészségesnek tekintünk, és vállaljuk, hogy minden beteget majd csak a későbbi fázisban tudunk gyógyítani.

A két lehetőség közül melyik kerül a társadalombiztosításnak kevesebb pénzbe?

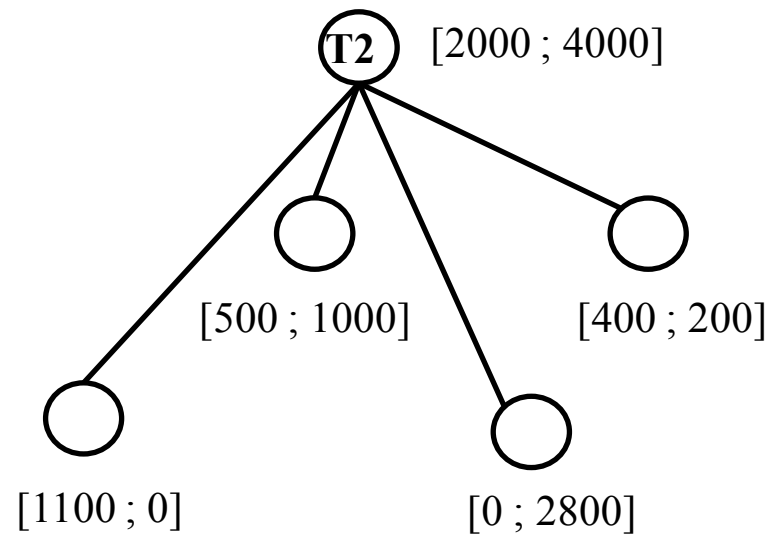
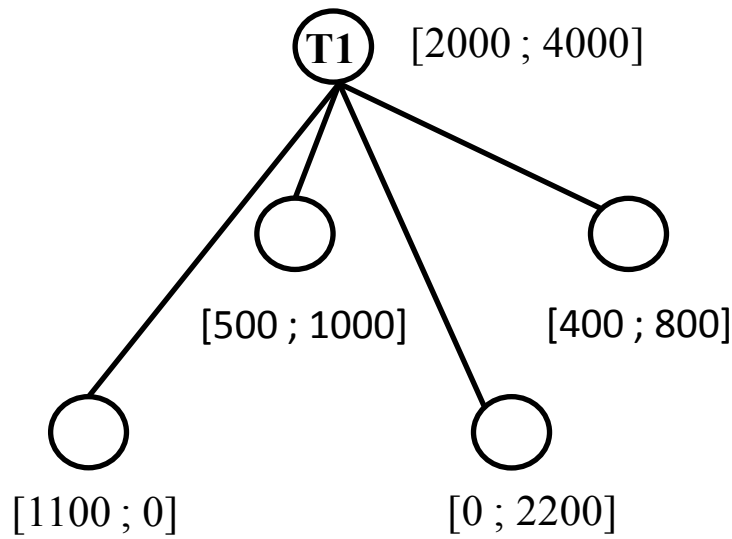
2. zh. előtti gyakorlás 5.

Az alábbi ábrán egy 12.000 minta alapján tanítással előállított döntési fa látható. A csomópontok mellett, mindig tőlük jobbra bekeretezve szögletes zárójelben látható két szám, ez a C1, illetve C2 osztályba jutó tanítóminták száma az adott csomópontban. Mind az osztályozás hibaaránya, mind az eszköz komplexitása költséget jelent nekünk, az egységnyi komplexitás növekedés 10-szer akkor költség, mint az egységnyi hibaarány növekedés. Két csomópont bezárása merült fel: a 8-as vagy a 2-es csomópontot érdemes bezárnunk a hibaarány-komplexitás kompromisszum elvét használva?



2. zh. előtti gyakorlás 6.

Az alábbi két teszt (T1, illetve T2) közül melyikkel kapunk nagyobb információnyereséget? A minták az A, illetve a B osztályba tartoznak, eredetileg 2000, illetve 4000 tanítóminta volt a két osztályból. A gyökércsomópontok mellett (jobbra), illetve a tesztek után létrejövő gyermekcsomópontok alatt található két-két szám azt mutatja, hogy abba a csomópontba hány A, illetve B osztálybeli tanítóminta jutott el [#A,#B] sorrendben. Természetesen a választását indokolja!



2. zh. előtti gyakorlás 7.

Egy adott betegség a népesség 0,7%-át érinti. A kifejlesztett diagnosztikai eljárásunk egy z paraméter mérésén alapul. Azt tapasztaltuk az eljárásnak egy 34.558 fős mintán történő vizsgálata során, hogy a mintában található 13.347 betegből 12.679-et sikerült helyesen betegnek diagnosztizálni, a többieket egészségesnek minősítette az eljárás. A vizsgálatba bevont 21.211 egészséges ember közül viszont 873-at betegnek diagnosztizáltunk, csak a többieket minősítettük egészségesnek. Tapasztalataink szerint az így talált arányok jól leírják a teljes népesség esetén is a szenzitivitás, specificitás jellemzőket.

Ha ezt az eljárást alkalmazzuk, akkor mennyibe fog kerülni a társadalombiztosításnak a 9.000.000 fős népességből az ebben a betegségben nem szenvedők (ebből a szempontból egészségesek) szűrése, és a közülük tévesen betegnek diagnosztizáltak kezelése? (Nem a teljes költség!)

A szokásos jelölésekkel a különböző csoportokban az egy főre jutó költségek: $C_{00}=2.700$ Ft; $C_{10}=15.300$ Ft; $C_{01}=182.500$ Ft; $C_{11}=71.700$ Ft. (Itt C_{10} , C_{01} , C_{11} már tartalmazza a vizsgálat C_{00} költségét is.)