

Intelligens orvosi műszerek (VIMIA023)

Zárthelyi 2014. december 2.

Minden válaszhoz rövid, tömör indoklást is kérek, kivéve az igaz/hamis feleletválasztós feladatokat! Fontos javaslat az elmúlt évek tapasztalatai alapján: nem érdemes kapkodni, alaposan olvassa el a feladatot, gondolkodjon el rajta, értelmezze, mielőtt elkezd megoldani! (Pl. egy „nem” vagy egy „mindig” szó a feladat szövegében sokat változtathat a helyes válaszon...)

NÉV (nyomtatott betűvel): NEPTUN-KÓD:
.....

ALÁÍRÁS:

1. A következő állítások közül melyik hamis, melyik igaz?

- a. Ítéletkalkulust használó szabályalapú rendszerekben a rezolúciót nem alkalmazhatjuk minden szakterületen. a. Igaz Hamis
- b. Egy döntési eljárást jellemző ROC görbét a görbe alatti vízszintes tengelyszakasz hosszával jellemezzük. b. Igaz Hamis
- c. Egy betegség szűrésével elérhető nyereség függ attól, hogy mekkora az adott betegség előfordulási aránya a népességben. c. Igaz Hamis
- d. Az ítéletkalkulusnál a VAGY bevezetés korrekt következtetési lépés. d. Igaz Hamis
- e. A C_{11} költség mindig kisebb a C_{00} költségnél. e. Igaz Hamis
- f. A Modus Ponens következtetés szerint: ha A igaz és az $A \Rightarrow B$ szabály fennáll, akkor következtethetünk arra, hogy B is igaz. f. Igaz Hamis
- g. Egy arcképfelismerő rendszerben a szem megtalálása a képen lényegkiemelési lépés. g. Igaz Hamis
- h. A Pan-Tompkins algoritmusban a felüláteresztő szűrő a nullvonal lassú változásának kiszűrésére szolgál. h. Igaz Hamis
- i. A Pan-Tompkins algoritmus közvetlenül megadja a HRV értéket. i. Igaz Hamis
- j. Elvileg előfordulhat szélsőséges paraméterek esetén, hogy egy szűrési eljárás végén az eredményezi a legkisebb költséget, ha mindenkit betegnek diagnosztizálunk. j. Igaz Hamis
- k. A HRV a szívritmus hosszúidejű ingadozására jellemző. k. Igaz Hamis
- l. Az előre következtetésnél a szabályok feltétel részére próbáljuk illeszteni a már ismert vagy kikövetkeztetett tényeket. l. Igaz Hamis

(jó válasz ≤ 6 : 0 pont, $6 <$ jó válasz: (jó válaszok száma-6) pont, 12 jó válasz 6 pont)

A túloldalon is vannak még feladatok!

2. Vércukormérést végzünk, és ennek alapján szűrjük a cukorbetegséget. 18.172 embert vizsgáltunk meg, akikről tudtuk, hogy közülük kik – ebből a szempontból – egészségesek, és kik betegek. Összesen 7.384 cukorbeteg van a vizsgált mintában. Megmértük az éhomi vércukorszintjüket, és azt tapasztaltuk, hogy a táblázatban megadott limiteknél az alábbiak szerint alakult a limit fölötti cukorszinttel rendelkező betegek és egészségesek száma. (Persze ez a valóságos szűrés erősen leegyszerűsített, és ezáltal torzított képe ... ráadásul légbőlkapott adatokkal, szóval csak gyakorlásra jó.)

G [mmol/l]	0	2	6	10	14	18	22	26
Azon egészségesek száma, akiknek G fölé esik a mért vércukorszintje	10.788	10.788	1.570	115	0	0	0	0
Azon betegek száma, akiknek G fölé esik a mért vércukorszintje	7.384	7.071	6.811	3.373	1.557	775	261	0

Adja meg a $G=6$ küszöbszint (e fölött betegnek, ez alatt egészségesnek tekintjük) esetén az eljárás ezen vizsgálatból kiszámítható érzékenységét és specificitását! **(3 pont)**

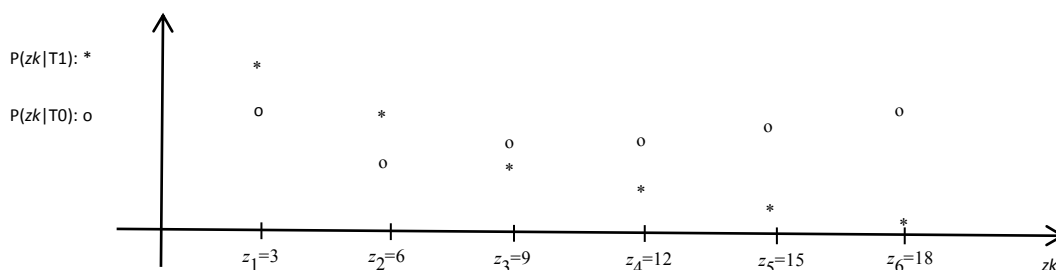
3. Egy „diagnosztikai eljárás” lényegében véletlen találgatást takar: a bal kéz középső ujjának hosszából következtetünk arra, hogy szívbeteg-e az illető. Ha egy adott küszöbnél kisebb az ujjhossza, akkor betegnek „diagnosztizálják”, különben egészségesnek. Egy vizsgálatot végeztek 103.871 emberen, akik közül 27.663 volt szívbeteg. Azt tapasztalták, hogy a legrövidebb középső ujj 2 cm-es volt, a leghosszabb 15 cm-es, és e két érték közt egyenletesen helyezkedtek el mind a betegek ujjhosszai, mind az egészségesek ujjhosszai. Rajzolja fel a vizsgálat alapján a „diagnosztikai eljárás” ROC görbéjét! Jelölje meg rajta a 2 cm-es ujjhossz küszöbhez és a 15 cm-es ujjhossz küszöbhez tartozó pontokat! **(3 pont)**
4. A következő két logikai mondat ugyanazt az igazságtáblát (a bemeneti változók igaz/hamis értékeihez tartozóan a logikai kifejezés igazsága/hamissága) hozza létre, vagy nem? (Magyarán: adja meg a két igazságtáblát!)

$$B \rightarrow A \wedge C$$

$$(\neg A \vee \neg B) \wedge C$$

(4 pont)

5. Egy z_k paramétert mérünk, és ennek alapján akarunk diagnosztizálni egy betegséget. A következő ábra azt mutatja, hogy milyen valószínűséggel mérhetjük az egyes z_k értékeket az esetben, ha valaki beteg (T_1), illetve az esetben, ha egészséges (T_0). Amennyiben a betegség a népesség 50%-ában fordul elő, és a felmerülő költségek (a szokásos jelölésekkel): $C_{00}=1.000$, $C_{10}=10.000$, $C_{11}=23.000$, $C_{01}=32.000$, akkor mely z_k értékeknél érdemes valakit betegnek tekinteni, és melyeknél egészségesnek? **(4 pont)**



Jó munkát!

A 3. és 5. feladatok rövid megoldása

3. feladat

Megbeszéltük órán, hogy a véletlen találgatás egy 45° -os egyenes ROC görbére vezet, mert ugyanolyan arányban nőnek a TP és FP minták az betegek és egészségesek csoportjain belül (hiszen a feladat szempontjából értelmetlen, irreleváns paraméter mentén osztályozunk). Ez különben szépen kijön abból az – amúgy részben felesleges – információból is, hogy a minták (ujjhosszak) mindkét csoportban egyenletesen oszlanak el. (Elég lenne azt tudnunk, hogy egyformán oszlanak el a két csoportban, de ez egyszerűsítés volt a feladat megoldásához.) Ha pl. centiméterenként felvesszük a küszöböt, akkor a 2, 3, ..., 15 centiméteres küszöbnél az egészségesek közül $103871/14 \Rightarrow$ kb. 7419 emberrel nő minden centiméterrel a betegnek tekintettek száma (FP), a betegekénél pedig $27663/14 \Rightarrow$ kb. 1976-al. Tehát:

Küsz.	2	3	4	5	...	14	15
FP (akiknek rövidebb az ujjuk a küszöbnél)	0	7419	14838	22257	...	96447	103871
$1-Sp=FP/103871$	0	0,07	0,14	0,21	...	0,93	1
TP (akiknek rövidebb az ujjuk a küszöbnél)	0	1976	3952	5928	...	25687	27663
$Sz=TP/27663$	0	0,07	0,14	0,21	...	0,93	1

3. feladat

Itt a kulcs az volt – amit órán szintén megbeszéltünk –, hogy egy adott z_k megfigyelés esetén vannak egészségesek is, betegek is, akik ilyen paramétert mutatnak. Persze nem tudjuk, hogy ki egészséges, ki beteg, csak eldönthetjük, hogy az összes ilyen értéket (pl. $37,5^\circ\text{C}$ testhőmérsékletet) mutató embert egészségesnek vagy betegnek tekintjük. A kérdés, hogy melyik a jobb (olcsóbb) döntés?

Ha a z_k értéket mutató embereket mind egészségesnek veszem, akkor ebből a csoportból az egészségesek a TN-be kerülnek, a betegek a FN-be.

Ha a z_k értéket mutató embereket mind betegnek veszem, akkor ebből a csoportból az egészségesek a FP-be kerülnek, a betegek a TP-be.

Akkor célszerű egészségesekhez sorolni az ilyen paraméterrel mért embereket, ha:

$$(C_{10} - C_{00}) * P(z_k|T_0) * P_0 > (C_{01} - C_{11}) * P(z_k|T_1) * P_1$$

Ez azt fejezi ki, hogy:

- $P(z_k|T_0) * P_0$ azon egészségesek aránya (száma), akinek ekkora a z paraméterértékük, tehát rájuk nézve $z = z_k$ (az előző idéttlen példa adataival pl. az ujjhosszuk 3 cm),
- ha őket egészségesnek vennénk, akkor C_{00} -ba kerülnének fejenként,
- ha betegnek vennénk, akkor C_{10} -ba,
- tehát, ha betegnek vennénk őket, akkor az egyenlőtlenség baloldal mutatja erre csoportra (egészségesek, akik a z_k értéket mutatják) a költségnövekedést (C_{00} helyett C_{10}).

Teljesen hasonló módon az egyenlőtlenség jobboldala:

- $(z_k|T_1) * P_1$ azon betegek száma, akinek ekkora a z paraméterértékük, tehát rájuk nézve $z = z_k$ (az előző idéttlen példa adataival pl. az ujjhosszuk 3 cm),
- ha őket betegnek vennénk, akkor C_{11} -be kerülnének fejenként,

- ha egészségesnek vennénk, akkor C_{01} -be,
- tehát, ha egészségesnek vennénk őket, akkor az egyenlőtlenség jobboldala mutatja erre csoportra (betegek, akik a z_k értéket mutatják) a költségnövekedést (C_{11} helyett C_{01}).

Ha a baloldal nagyobb, ez azt jelenti, hogy betegek tekinteni őket nagyobb vereség, mint a másik csoportot egészségesnek. (Nagyobb költségnövekedést okoz az FP, mint az FN csoport.)

Tehát az a költséghatékony, ha ezt a csoportot egészségesnek tekintjük.

Persze, ha az egyenlőtlenség fordított, akkor a következtetés is fordított, ezt a csoportot betegek olcsóbb tekinteni (kisebb költségnövekedéssel jár a FP, mint a FN csoport).

Példánk annyiban egyszerűbb, hogy

$$C_{10}-C_{00}=10000-1000=9000,$$

$$C_{01}-C_{11}=32000-23000=9000,$$

$$P_0=P_1=0,5$$

tehát az egyenlőtlenség mindkét oldala egyszerűsíthető, és csak azt kell vizsgálni, hogy

$$P(z_k|T_0) > P(z_k|T_1) \text{ ez esetben egészségesnek célszerű venni őket,}$$

vagy:

$$P(z_k|T_0) < P(z_k|T_1) \text{ ez esetben betegek célszerű venni őket.}$$