

*Nem pontosan ezek lesznek a vizsgafeladatok, csak jellegükre hasonlók!  
Természetesen a feladatok stílusára lehet következtetni a zh és pótzh feladatokból,  
illetve a félév közben megoldott gyakorló feladatokból is.*

1. Testtömegindex (BMI) mérést végzünk: célunk, hogy ennek alapján szűrjük a kóros soványságot. A szűrési eljárás jellemzése céljából összesen 23.417 embert vizsgáltunk meg, akikről tudjuk, hogy kik – ebből a szempontból – egészségesek, és kik (kórosan sovány) betegek. Összesen 16.917 egészséges van a vizsgált mintában. Megmértük a vizsgálatban résztvevők testtömegindexét, és azt tapasztaltuk, hogy a táblázatban megadott B értékek alatti értéknél az alábbiak szerint alakult a betegek, ezen B értékek feletti értéknél az egészségesek száma. (Persze ez a valóságos szűrés erősen leegyszerűsített, és ezáltal torzított képe... ráadásul légbőlkapott adatokkal, szóval csak ROC görbe gyakorlásra jó.)

B	8	16	18	20	25
Egészségesek, akiknek B <u>fölé</u> esik a mért BMI-je	16.917	16.917	15.863	12.132	0
Kórosan sovány betegek, akiknek B <u>alá</u> esik a mért BMI-je	0	3.879	5.803	6.194	6.500

Rajzolja fel ezen vizsgálat alapján a – B-nél kisebb érték esetén beteg, a fölötti értéknél egészséges – eljárást jellemző ROC görbét!

2. Az X betegségben a népesség 1,17%-a szenved. Ha a betegséget korai fázisban kezeljük, akkor a kezelés átlagos költsége 23.000 Ft/fő, ha csak a késői fázisában, akkor 843.000 Ft/fő.

Két lehetőség közt választhatunk:

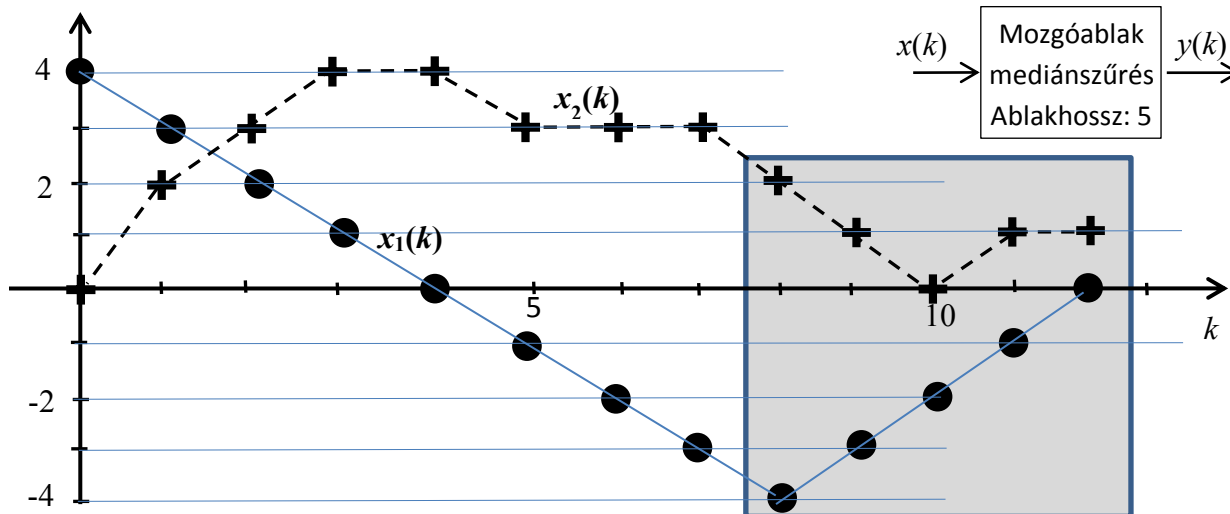
- a) Szűrést végzünk, amelynek költsége 4.000 Ft fejenként, specificitása 0,95; szenzitivitása 0,92. Az itt betegnek vélelmezett embereknél (nyilván akár valóban betegek, akár nem – ezt nem tudhatjuk előre) egy további – átlagosan 8.000 Ft/fő költségű vizsgálatot eldöntjük, hogy tényleg beteg-e, és csak a tényleg betegnek bizonyult embereket kezeljük.
  - b) 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 a legkevesebb pénzbe?

3. Egy ötmillió célcsoport szűrésének lehetőségét vizsgáljuk. Az adott betegségben a népesség 0,4%-a szenved. A szűrés költsége fejenként 1.700 forint, tehát ez mindenkinél jelentkezik. Az előzetes vizsgálatok alapján a szűrési eljárás segítségével 83,7% biztonsággal felismerhetők a betegek. Ugyanakkor 12,3%-ban az egészségeseket is betegnek mutatja a vizsgálat, ezt a tévedést csak egy újabb, 8.200 Ft-os kontrollvizsgálattal lehet korrigálni. Ezért ezt a kontrollvizsgálatot minden betegnek diagnosztizálnál elvégezzük. (Akar valóban betegek, akár csak mi annak tartjuk őket, hiszen nem tudjuk, hogy kik valóban betegek, csak azt tudjuk, hogy a szűrés kiket mutatott annak.) Ha idejében felismerjük a betegséget a szűrés eredményeképpen, akkor a kezelés átlagosan 58.400 forintba kerül. Ha nem ismerjük fel – szűréssel – idejében a betegséget, és csak később kezdjük el kezelni, akkor az átlagosan 773.000 forintba kerül a kezelés. Mennyi lesz az ezen betegség által okozott egy főre jutó költség, illetve a teljes népességre jutó költség?

4. Az ábrán látható két jelet mediánszűréssel szűrjük egy 5 pontos mozgóablakkal. A k-dik időpontban a szűrés eredménye, ha a bemenetére kapcsolt jel  $x(k)$ :

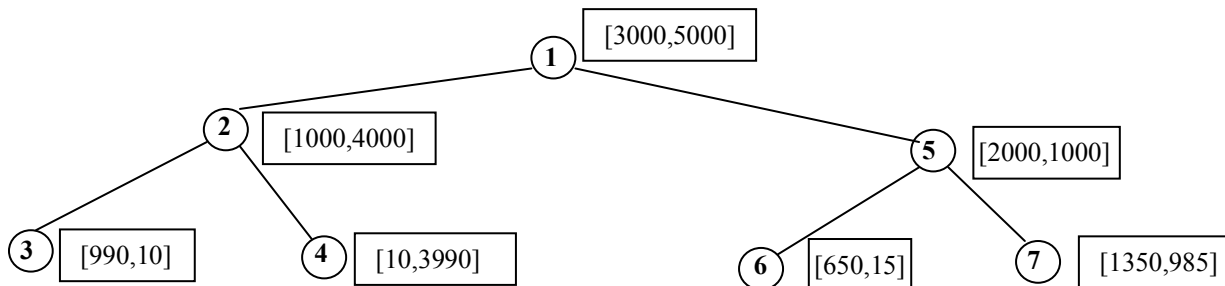
$$y(k) = \text{medián}(x(k - 2), x(k - 1), x(k), x(k + 1), x(k + 2))$$

Mi lesz a mediánszűrés eredménye az  $x_1(k)$  jelre, az  $x_2(k)$  jelre és az  $x_1(k)+x_2(k)$  jelre, amikor az ábrán megjelölt szakasz esik az ablakba ( $k=10$ )?



5. Az alábbi ábrán egy 8000 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. A döntési fával elérhető osztályozás hibaarányának százszorosa az ezzel kapcsolatos költségünk. (Ha a hibaarány 0, akkor a költségünk is 0, ha a hibaarány 1, akkor 100 a költség, ha pl. a hibaarány 0,5 a megfelelő költségünk 50 lenne.) A szokásos módon meghatározott komplexitás 6 költséget okoz nekünk levelenként. Érdemes-e felhasználnunk a döntési fát, vagy jobban járunk, ha egyszerűen bezárjuk a gyökér csomópontot, és a gyökér csomópontot levélnek tekintve adjuk meg az osztályba sorolásra a választ?



6. Egy  $y(k)=x(k)+z(k)$  jelet mérünk, Az  $x(k)$  szinuszos jelsorozat periódusideje  $K_{\text{per}}=7$ . Az  $x(k)$  értékeket sorban becsüljük a teljes periódusra  $k=1,2,\dots, K_{\text{per}}$ . A mérést torzító  $z(k)$  nulla várható értékű (átlaga véges számú zajminta esetén közelít a nullához, ha nagyon sok mintát veszünk, akkor gyakorlatilag nulla) sztochasztikus zaj, amelynek egyes időpontokban az értékei függetlenek egymástól és a jeltől, ezért megfelelő átlagolással csökkenthető a torzító hatása. A következő  $y(k)$  sorozatot mértük:

$k$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$y(k)$	1,51	2,63	-2,24	0,09	-0,66	-1,76	-0,02	1,31	1,37	2,79	-2,12	-0,76	0,27	0,35	1,69
$k$	16	17	18	19	20	21	22	23	24	25	26	27	28	29	31
$y(k)$	0,59	-0,10	0,72	0,43	0,97	1,09	-0,24	1,51	1,65	-0,28	0,06	0,27	0,11	1,26	0,01

Adja meg ezen mérések alapján  $x(2)$  átlagolással javított becslését!

7. Egy tudottan konstans jelet sokszor megmértünk, és a méréseket átlagoltunk, mert a mérésünk zajjal terhelt. Az első 22 mérés a következő adatsorozatot eredményezte:

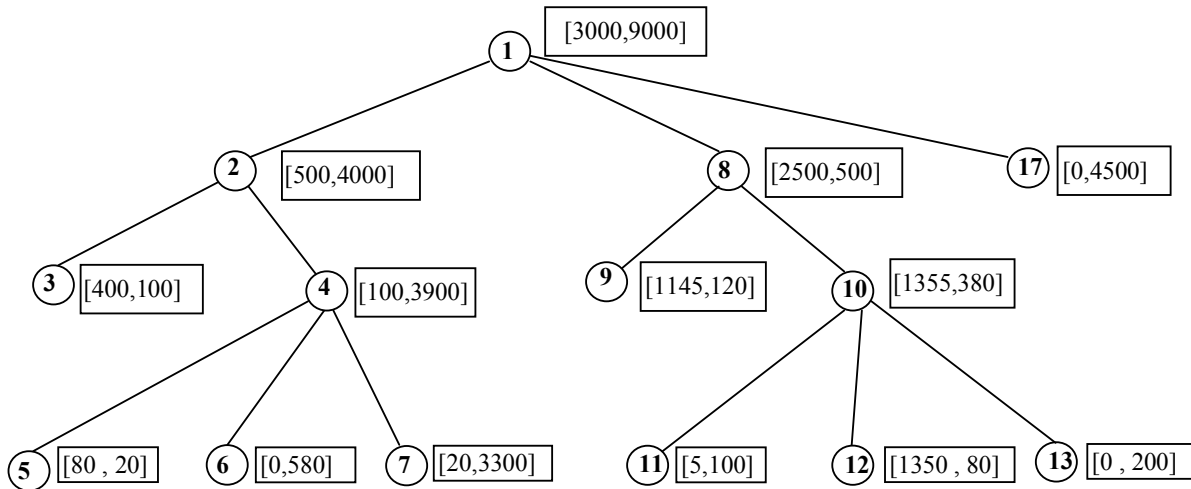
13.3147 13.4058 12.6270 13.4134 13.1324 12.5975 12.7785 13.0469 13.4575 13.4649  
 12.6576 13.4706 13.4572 12.9854 13.3003 12.6419 12.9218 13.4157 13.2922 13.4595  
 13.1557 12.5357

Ennek alapján a becslésünk 13,1151 volt. Végeztünk egy újabb, huszonharmadik mérést is, ennek eredménye 12,7210 volt. Mennyi az új becslésünk az ismeretlen konstansra?

8. Előzetes vizsgálatot végeztünk egy diagnosztikus eljárás érzékenysége és specificitására, a megvizsgált 75.210 emberből 40.105 volt egészséges, a többi beteg. A szűrési eljárással összesen 36.321 embert észleltünk betegnek, de közülük csak 31.980 volt ténylegesen az. Ezen előzetes vizsgálat eredménye alapján számolva várhatóan mi lesz a szűrési eljárásunk érzékenysége és specificitása?
9. Az X betegségben a népesség 0,15%-a szenved, a szűrés specificitása 0,88; szenzitivitása 0,92? A szűrés költsége 4.000 Ft, a nem diagnosztizált X betegség 850.000 Ft-ba kerül személyenként, a szűrés során felfedezett X betegség kezelése 56.000 Ft/fő. A tévesen betegnek diagnosztizált X betegség 7.600 Ft/fő többletköltséget okoz a szűrés költségén túl, mire kiderül az illetőről, hogy egészséges. (Ezt a 7600 forintos megerősítő vizsgálatot minden betegnek diagnosztizálton elvégezzük.)

Mekkora átlagos költségmegtakarítást jelent a társadalombiztosításnak fejenként, ha a szenzitivitást sikerül 0,95-re javítanunk, miközben a többi adat változatlan?

10. 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?



11. Mutassa meg, hogy a 3x3-as ablakkal végzett 2D mediánszűrésre ezen két kép esetén fennáll-e a szuperpozíció! ( A szélső sorokat, oszlopokat nem változtatjuk a szűrés során.)

0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	1	1	0	0	0	0	0
0	0	1	1	0	0	1	0	1	0	0	0	0
0	0	0	0	1	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0

0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0
0	1	1	1	0	0	0	0	0	0	0	1	0
0	1	0	0	1	0	0	0	0	0	0	1	0
0	1	1	1	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0

12. Adjon meg 2 olyan 2-dimenziós jelet (képet), amelyre fennáll a 3x3-as ablakkal végzett 2D mediánszűrés elvégzése esetén a szuperpozíció! Adjon meg 2 olyant is, amelyeken van nulla értékű és nem nulla értékű pixel is, de nem áll fenn a szuperpozíció!