

## Adatbányászati technikák 2. zárthelyire gyakorló feladatok 2018. május 14.

1. Készítsen perceptront az A OR NOT B függvényhez vagy mutassa meg, hogy ez nem lehetséges. Ez utóbbi esetben készítsen neurális hálózatot.

2. Az alábbi táblázattal adott függvényhez készítsen perceptront vagy mutassa meg, hogy ez nem lehetséges. Ez utóbbi esetben készítsen neurális hálózatot a függvény megvalósítására.

A	B	C	címke
0	0	0	1
0	0	1	0
0	1	0	0
1	0	0	1
0	1	1	1
1	0	1	0
1	1	0	0
1	1	1	0

3. A 2, 3, 4, 7, 8, 10, 12, 14, 20 számokat akarjuk Kmeans algoritmussal három klaszterbe sorolni.

(a) Írja le, hogy hogyan fut le az algoritmus, ha a 2, 7, 10 kezdő-centroidokat használjuk!

(b) Írja le, hogy hogyan fut le az algoritmus, ha a 2, 3, 4 kezdő-centroidokat használjuk!

4. Az alábbi mátrix alapján rajzolja fel dendrogramon, hogy a MIN (single link) hierarchikus klaszterező algoritmus hogyan klaszterez. A dendrogram elkészítésének és a hasonlóságmátrix változásának minden fázisát mutassa be.

	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$
$p_1$	1	0.8	0.3	0.2	0.9	0.1
$p_2$	0.8	1	0.2	0.25	0.7	0.15
$p_3$	0.3	0.2	1	0.8	0.1	0.05
$p_4$	0.2	0.25	0.8	1	0.05	0.6
$p_5$	0.9	0.7	0.1	0.05	1	0.7
$p_6$	0.1	0.15	0.05	0.6	0.7	1

5. (a) Az előző mátrix alapján futtassa le a MAX (complete link) hierarchikus klaszterező algoritmust.

(b) Hogyan néz ki a proximity mátrix az első kör után a fenti mátrixból kiindulva, ha a GROUP AVERAGE módszert használjuk?

6. Tekintsük az alábbi 15 tranzakciót az 123456 elemhalmazon: 123, 13456, 1235, 146, 136, 1234, 2346, 2456, 246, 346, 356, 456, 46, 5, 56. Az Apriori algoritmussal határozza meg a gyakori elemhalmazokat,  $min\_sup = 0.2$  küszöbvel! Mutassa be, hogy hogyan alakulnak a  $C_k$  és  $F_k$  halmazok az algoritmus futása során.

7. Az előző feladat tranzakcióit használva határozza meg az órán tanult algoritmussal, hogy melyek a 246 gyakori elemhalmazból kapható nagy megbízhatóságú szabályok, ha  $min\_conf = 0.7$ .
8. A 6. feladatban talált gyakoriakból melyek a maximális gyakoriak? Melyek a zárt gyakoriak?
9. Határozza meg az összes gyakori elemhalmazt az  $\{1, 2, 3, 4, 5, 6\}$  elemhalmazban, ha tudjuk hogy a zárt gyakoriak és az abszolút gyakoriságuk (a zárójelben) a következők:

146(2), 346(2), 13(3), 23(2), 24(3), 26(2), 45(2), 14(4), 34(3), 36(39), 46(5), 2(4), 3(5), 4(8), 6(7)

10. Tekintsük az alábbi kontingencia-táblázatot, mely a nem és a tollaslabdázás, mint hobbi kapcsolatát mutatja egy 5000 fős mintán.

	nő	férfi
tollasozik	2000	1750
tollasozik	1000	250

- (a) Mennyi a (nő  $\rightarrow$  tollasozik) asszociációs szabály Lift-mutatója?
- (b) A kapott érték ismeretében érdemes-e tollasozással kapcsolatos reklámmal megkeresni a nőket?