

Adatbányászati technikák (VISZM185) 2014 tavaszi félév

Csima Judit

BME, VIK,
Számítástudományi és Információelméleti Tanszék

2014. február 12.

Előadás

- Csima Judit, csima@cs.bme.hu, IB 136/b
- minden szerdán 8.30-10.00 (nincs szünet), QBF10 és minden páratlan héten csütörtökön 10.30-12.00 (nincs szünet), QBF12
- weboldal (vetített fóliák, segédanyagok): www.cs.bme.hu/dm

Labor

- Csima Judit, csima@cs.bme.hu
Bagyinszki Bence, bagyibence@gmail.com
- két csoport, páros (01-es kurzus) és páratlan (02-es kurzus) heteken, csütörtökön 8.30-10.00 (nincs szünet), mindkét csoportnak R4J (már ezen a héten is)
- laborhasználathoz kell HSZK account
ha nincs még: <https://accadmin.hszk.bme.hu/index.php>
(beiratkozott, Aktív státusz kell hozzá)
- lehet saját gépet is használni, de akkor (az első feléhez) kell RStudio, letölthető innen: www.rstudio.com

Előadás-tematika

- témák címszavakban : bevezetés, adattípusok, megjelenítés, asszociációs szabályok, osztályozók, klaszterezés, anomália detektálás
- anyaga nagyrészt: Tan, Steinbach, Kumar: Data Mining (Addison-Wesley)
(<http://www-users.cs.umn.edu/~kumar/dmbook/index.php>)
- néha (az elején) vetített (magyar és angol nyelvű) slide-ok
- legtöbbször táblás előadás, jegyzetelni kell, elmélet, sok példával

Labor-tematika

- két nagy rész, 3-3 alkalom jut mindkettőre
- utolsó alkalommal beadott házik ellenőrzése
- első rész: R programozási nyelv és környezet: ábrázoláshoz, statisztikai számításokhoz
- második rész: Weka adatbányászati szoftver alkalmazása és Java programba integrálása

Zárthelyi követelmények

- két ZH a félév során: március 26., május 14., az előadás idejében
- mindkét ZH 35 pontos, mindkettőből el kell érni 40%-ot (14 pont)
- egy ZH pótolható pótZHval (javító céllal is megírható): május 21.
- utolsó pótlási lehetőség (pótpótzh) a pótlási héten
- az egyik ZHnak meg kell lennie elsőre, a zh eredeti időpontjában
- anyag: elméleti kérdések az előadás anyagából és feladatok
- a ZH előtti előadáson nézünk típuspéldákat, mintakérdéseket

Labor követelmények

- két beadandó feladat(sor) a félév során (R és Weka/Java)
- mindkét feladat(sor) 15 pontos
- utolsó laboron ellenőrzés (belekérdezünk)
- ha kiderül, hogy nem saját kódot adott be valaki: elégtelen
- laboron katalógus, az első 6 alkalomból legfeljebb egyszer lehet hiányozni

Félévi jegy

- ha a két ZH sikeres (akár pótlással)
- összesen max. 100 pont (35-35 pont a két Zh, 15-15 pont a két labor)
- 40 ponttól kettes, 55 ponttól hármas, 70 ponttól négyes, 85 ponttól ötös