

Adatbányászati technikák (VISZM185) 2018 tavaszi félév

Csima Judit

BME, VIK
Számítástudományi és Információelméleti Tanszék

2018. február 5.

Előadás

- Csima Judit, csima@cs.bme.hu, IB 136/b
- minden hétfőn 8.15-9.55 QBF08 és
minden páratlan héten csütörtökön 10.15-12.00 QBF08
- weboldal (vetített fóliák, segédanyagok): www.cs.bme.hu/dm

Labor

- Csima Judit, csima@cs.bme.hu
Kabódi László, kabodil@gmail.com
- két csoport, helyszín minden csoportnak **R4K**
 - 11-es kurzus: páros heteken csütörtökön 8.15-9.45 (nincs szünet)
 - 12-es kurzus: páratlan heteken csütörtökön 8.15-9:45 (nincs szünet)
- laborhasználathoz BME címtáras azonosító kell
ha nincs még: <https://login.bme.hu/admin>
(beiratkozott, Aktív státusz kell hozzá, **lehet, hogy 1-2 nap, amíg elkészül!!**)
- lehet saját gépet is használni, de akkor (az első feléhez) kell rá R és RStudio, letölthető innen: <https://cran.r-project.org> és www.rstudio.com, a második részét a labornak majd hirdetjük később

Az előadás tematikája

- témák címszavakban: bevezetés, adattípusok, előfeldolgozás, megjelenítés, osztályozók, klaszterezés, asszociációs szabályok
- az előadás egy részében vetített (magyar és angol nyelvű) slide-ok is, bővebb szóbeli magyarázattal
- nagyrészt táblás előadás, jegyzetelni célszerű, elmélet, sok példával
- a tárgy anyaga nagyrészt: Tan, Steinbach, Kumar: Data Mining (Addison-Wesley)
(<http://www-users.cs.umn.edu/~kumar/dmbook/index.php>)

Labor-tematika

- két nagy rész, 3-3 alkalom jut mindkettőre (egy labor elmarad a 11-es kurzuson, ezt majd megbeszéljük)
- utolsó alkalommal beadott házik ellenőrzése
- első rész: R programozási nyelv és környezet: ábrázoláshoz, statisztikai számításokhoz
- második rész: Weka adatbányászati szoftver alkalmazása

Zárthelyi követelmények

- két ZH a félév során: március 22. és május 15., este 6-kor
- mindkét ZH 35 pontos, mindkettőből el kell érni 40%-ot (14 pont)
- egy ZH pótolható pótZHval (javító céllal is megírható): május 22.
- utolsó pótlási lehetőség (pótpótzh) valamikor később
- 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

- nincs beugró
- labor végén nem kell leadni jegyzőkönyvet
- két beadandó feladat(sor) a félév során (R és Weka)
- mindkét feladat(sor) 15 pontos (nincs minimumkövetelmény)
- utolsó laboron ellenőrzés (belekérdezünk)
- ha kiderül, hogy nem saját kódot adott be valaki: elégtelen az egész tárgyra
- 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