

Adattudományi sáv záróvizsga-témakörök

Bevezetés az adattudományba 1.

- attribútumok típusai
 - o pl. I./1.
- hasonlóságok, távolságok (Minkowski-távolságok, Mahalanobis-távolság, egyszerű egyezőségi együttható (SMC), Jaccard-együttható, koszinusz távolság, dinamikus idővetemítés (DTW))
 - o pl. I./2.,3.,4.,5.,6.,7.
- döntési fa felépítése, működése (félreosztályozási hiba, Gini-index, entrópia alapján)
 - o pl. II./2., 3., 5., III./4.
- osztályozás kiértékelése, teljesítménymutatók (precision, recall, F-mérték, ROC, AUC)
 - o pl. II./3.,4.,5.
- naiv Bayes-módszer (Laplace-bebecslés, m-bebecslés)
 - o pl. II./6.
- osztályozó algoritmusok (kNN, döntési fa, naiv Bayes) működésének bemutatása és teljesítményének értékelése vizuálisan megadott kis mintaadatokon
 - o pl. II./7., 8.
- számolás Bayes-hálóval, feltételes függetlenség
 - o pl. III./1.
- lineáris regresszió
 - o pl. III./2.
- gradiens módszer, sztochasztikus gradiens módszer
 - o pl. III./2.
- perceptron működése, neurális hálók felépítése
 - o pl. III./3., 6., 8.
- lineáris szeparálhatóság, maximális margó elve
 - o pl. III./5., 7.
- asszociációs szabályok (támogatottság, megbízhatóság, lift mutató, apriori algoritmus, maximális, zárt elemhalmazok)
 - o pl. IV./1.
- k-közép klaszterező működése
 - o pl. IV./2.,4.,6.
- hierarchikus klaszterező működése (MIN, MAX távolság)
 - o IV./4.,5.
- klaszterező algoritmusok (hierarchikus klaszterező (MIN, ill. MAX távolság esetén), k-közép, Gauss-keverékmodell, DBSCAN) működésének bemutatása és teljesítményének értékelése vizuálisan megadott kis mintaadatokon
 - o pl. IV./7.
- ajánlórendszer rejtett faktormodellel
 - o pl. IV./3.
- PageRank algoritmus működése
 - o pl. IV./8.,9.,10.