

11. tétel

Berekméri Evelin

2019. június 10.

Kivonat

Számítógépes tanulás – Predikciós és klasszifikációs módszerek. Felügyelt és felügyelet nélküli tanítás. A tanítóhalmaz, a validáció és a túlfittelés. K-means, Support Vector Machine, Random Forest, k-NN-módszer.

Klasszifikációs módszerek

Klasszifikáció során kvalitatív válaszokat prediktálunk, amelyek egy véges halmazból vesznek fel értékeket. Klasszifikációs módszerek: logisztikus regresszió, lineáris diszkriminancia-analízis, KNN, tree, random forest, boosting, support vector machine.

Felügyelt és felügyelet nélküli tanítás

Felügyelt tanítás esetén egy statisztikai modellt építünk inputok alapján azért, hogy prediktáljuk vagy megbecsüljük az outputokat. Minden megfigyeléshez (x_i "predictor" érték(ek)hez) tartozik egy y_i "válasz" érték. Ilyen módszerek például a lineáris és logisztikus regresszió, boosting, support vector machine.

Felügyelet nélküli tanulás során nincsenek outputok, csak inputok, amik alapján az adathalmaz szerkezetéről és kapcsolatairól tudhatunk meg többet. Ilyen módszer például a klaszterezés.

A tanítóhalmaz, a validáció és a túlfittelés

A rendelkezésünkre álló adathalmazt felosztjuk tanítóhalmazra, validációs halmazra és teszhalmazra.

K-means, Support Vector Machine, Random Forest, k-NN-módszer

Hivatkozások