

Adatbányászat

Tételsor

1. Az adatbányászat alapfogalmai, módszertana és alkalmazási területei. Adatbányászati módszerek és összehasonlításuk.
2. Adatbányászati előkészületek: mintavétel, feltáró adatelemzés, adatmódosítás.
3. Lineáris regressziós modellek. A lineáris modell és a legkisebb négyzetek módszere. A paraméterek értelmezése és az input változók fontosságának mérése.
4. Logisztikus regresszió. A modell: link és aktivációs függvény. Paraméterbecslés a maximum likelihood módszerrel: az iterált újrásúlyozott legkisebb négyzetek módszere.
5. Döntési fák. A CART, CHAID és C4.5 algoritmusok
6. Neurális hálók. Előrecsatolt neurális hálózatok. Aktivációs és kombi-nációs függvény. Tanítási módszerek, nemlineáris optimalizálás.
7. Társítási szabályok. Alapfogalmak: gyakori tételcsoport és társítási szabály, támogatottság és megbízhatóság. Érdekességi mérőszámok: lift-érték. Az Apriori algoritmus.
8. Klaszterezés: K -közép és hierarchikus módszerek.
9. Látens változós modellek. Folytonos látens változó: főkomponensana-lízis. Diszkrét látens változó: Gauss keverékek felbontása az EM-algoritmussal.

Debrecen, 2008. május 17.

Dr. Ispány Márton