

APRENANTATGE ESTADÍSTIC I MINERIA DE DADES

Objectius: Com es processen les grans masses de dades? Les bases de dades de molts gigabytes, per exemple les recollides en una Web de e-venda, contenen molt de "soroll" inútil, però també una certa porció d'informació important. Estudiarem els procediments i mètodes que permeten detectar-la i aïllar-la, des d'una perspectiva basada en exemples treballats en l'ordinador.

Prerequisits recomanats: Estadística bàsica. Elements de programació.

TEMARI

1. Aprenentatge supervisat. Classificació i regressió lineals. El perceptró.
2. No linealitats: splines i nuclis.
3. Models additius. Xarxes neuronals, amb perceptrons i amb funcions radials.
4. Aprenentatge amb nuclis. Màquines de vectors suport.
5. Models basats en arbres.
6. Aprenentatge no supervisat. Regles d'associació. Clusters. Mapes auto-organitzats.

BIBLIOGRAFIA

Trevor Hastie, Robert Tibshirani and Jerome Friedman: "The elements of Statistical Learning. Data Mining, Inference, and Prediction", Springer, New York (2001).

Ian H. Witten and Eibe Frank: "Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations", Morgan Kaufmann Publishers, San Francisco (1999).

David Hand, Heikki Mannila and Padhraic Smyth: "Principles of Data Mining", The MIT Press, Cambridge, Massachusetts (2001).

Nello Cristianini and John Shawe-Taylor: "An Introduction to Support Vector Machines and other kernel-based learning methods", Cambridge University Press (2000).