Termen voor Data mining

* Feature construction
* Feature space
* Hypothesis space
* Forward/background feature selection
* Experimental methodology
  + Cross-validation
  + Parameter tuning
  + Accuracy, ROC, precision-recall
* Decision trees
* Rule induction
* Naïve Bayes
* Neural networks
* Instance-based learning
* Computationally sound
  + Scalability time and space complexity
  + Parallelizable, e.g., MAP-Reduce and Hadoop?
* Statistically sound
  + Are patterns meaningful?
  + Do our results generalize to new data?
* Ergonomically sound
  + Presents results in a comprehensible manner
  + Does it need 6 PhDs to run it?