**Objectives for Midterm 2**

You should be able to answer the questions on the second test if you can do the things listed in the following. The test will only sample your knowledge, and so will be shorter than this list.

You may use your textbook, copies of the lecture slides, a one page HANDWRITTEN cheat sheet, and your own class notes, but no other materials. You can use a pocket calculator (not programmable!), but you cannot use a laptop.

* (P302) Use confusion and cost matrices to compute which of two classifiers is better for a data set.
* (P298) Explain the notion of an ROC curve and its meaning for classifier performance.
* (P298) Use ROC curves to compare performance of different classifiers.
* (P221) Construct sets of decision rules from decision trees, and decision trees from sets of decision rules.
* (P208) Explain and compare the notions of coverage and accuracy of classification rules.
* (P213) Explain how to construct sets of classification rules, and compare this process to methods for constructing decision trees.
* (P223) Explain and apply instance based classifiers.
* (P228) State and explain Bayes theorem and its use in updating probability distributions to incorporate new evidence, and use it to compute probabilities.
* (P229) Explain and use Bayes classifiers.
* (P231) Diagram and explain the naive Bayesian classification method.
* (P231) Compute probabilities in small data sets and use these values in a naive Bayesian classifier to classify data items.
* (P246) Explain and evaluate artificial neural network classification methods.
* (P256) Explain the notions of maximum margin hyperplanes and support vectors and their roles and uses in support vector machines.
* (P283) Explain and compare bagging and boosting ensemble methods for improving classifier performance.
* (P328) Explain and compare the notions of association and classification rules.
* (P330) Explain and compare the notions of support, and confidence of association rules.
* (P333) Describe and use the Apriori Principle and the Apriori Algorithm.
* (P344) Describe and apply hash trees.
* (P354) Define and use maximal frequent itemsets and closed itemsets.
* (P359) Explain methods for frequent itemset generation.
* (P333) Explain how to efficiently generate rules from frequent itemsets.
* (PPT) Explain and use multiple minimum support thresholds.
* (P373) Explain, compute, and critique pattern interestingness measures.
* () Describe the effect of support pruning on pattern interestingness measures.
* (P429) Define and use the notions of sequence and subsequence.
* (P431) Define the sequential pattern mining problem.
* (P433) Describe and apply the Generalized Sequential Pattern Mining Approach.
* (P436) Explain and use timing constraints for sequential pattern.
* (P437) Explain and perform mining for sequential pattern with timing constraints.