CMPT-740 Fall 2003
Foundations of Data Mining
Martin Ester

Assignment 5

Total marks: 60
Due date: October 22, 2003

Assignment 5.1

(a) How does the decision surface (that separates the different classes in the data space) of the *k*-nearest neighbor classifier look like? Illustrate this by providing a two-dimensional example.

(b) Show a training data set of two-dimensional points that can be classified with 100 % accuracy with a *k*-nearest neighbor classifier but not with a linear SVM.

(c) In which real-world situations could clustering as preprocessing reduce the classification error of a *k*-nearest neighbor classifier (in particular, for $k = 1$)? Provide a two-dimensional example. Describe your whole method including the clustering step and the *k*-nearest neighbor classifier.

Assignment 5.2

(a) Design a perceptron with two inputs *A* and *B* that implements the Boolean function $A \wedge \neg B$.

(b) Design a two-layer network of perceptrons implementing *A* XOR *B*.