CMSC/PHIL 372 Artificial Intelligence
Problem Set 5
Due at the start of class on April 26*th*

1. (K-Nearest Neighbor – 25 pts) Consider the following labeled data set:

   ```
   5, 6, +
   3, 2, +
   4, 2, +
   3, 4, -
   6, 3, -
   ```

   (a) (15 pts) Plot the data (you can draw it by hand) and draw the decision surface that separates the two classes (+ and -) for k=1.

   (b) (10 pts) How would you classify a point at (4,1) for both k=1 and k=3?

2. (Linear Separability – 20 pts) [Rephrase of Alpaydin Ex.10.9] Consider data characterized by a single attribute $x \in \mathbb{R}$ and class label $y \in \{C_1, C_2\}$ with the following class assignments:

$$y = \left\{ \begin{array}{ll} C_1 & \text{if } x < 1 \text{ or } x > 6 \\ C_2 & \text{otherwise} \end{array} \right. \tag{1}$$

   Describe how you can use a linear discriminant (i.e., a linear SVM) to separate the two classes.

3. (Support Vectors – 25 pts) For an SVM, if we remove one of the support vectors from the training set, does the size of the maximum margin (the distance from the support vectors to the hyperplane, or the width of the fat separator) decrease, stay the same, or increase for that dataset? Why?

   Also justify your answer by providing a simple dataset (no more than 2-D) in which you identify the support vectors, draw the location of the maximum margin hyperplane, remove one of the support vectors, and draw the location of the resulting maximum margin hyperplane. (You can draw by hand.)

4. (Definitions – 25 pts) Provide a 1-2 sentence definition or description of the following (be brief!):

   (a) The Naive Bayes learning algorithm (e.g., how you learn all the parameters needed for Naive Bayes, in 1-2 sentences. There is no need to go over the entire equation or its derivation. Just tell me what parameters are needed and how you learn them from data.)

   (b) Deep Belief networks (describe them in one sentence)