CSE 151: Machine learning **Winter 2019**

Homework 7

Submission instructions:

- Please type up your solutions.
- If a problem asks for a numerical answer, you need only provide this answer. There is no need to show your work, unless you would like to.
- Upload the PDF file for your homework to `gradescope` by 6pm on Tuesday February 26.

Part A: Linear classification

1. Consider the following small data set in $\mathbb{R}^2$:

   - Points $(1,2), (2,1), (2,3), (3,2)$ have label -1.
   - Points $(4,5), (5,4), (5,6), (6,5)$ have label $+1$.

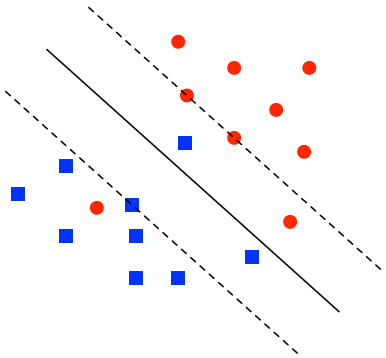
   Now, suppose (hard margin) SVM is run on this data.

   (a) Sketch the resulting decision boundary.
   (b) What is the (numerical value of the) margin, exactly?
   (c) What are w and b, exactly?

2. An SVM classifier is learned for a data set in $\mathbb{R}^2$. It is given by $w = (3,4)$ and $b = -12$.

   (a) Draw the decision boundary, making sure to clearly indicate where it intersects the axes.
   (b) Draw the left- and right-hand boundaries, also clearly making where they intersect the axes.
   (c) What is the margin of this classifier?
   (d) How would the point $(2,2)$ be classified?
   (e) It turns out that the data set has two distinct support vectors of the form $(1, ?)$. What are they?

3. The picture below shows the decision boundary obtained upon running soft-margin SVM on a small data set of blue squares and red circles.



(a) Copy this figure and mark the support vectors. For each, indicate the approximate value of the corresponding slack variable.

(b) Suppose the factor C in the soft-margin SVM optimization problem were increased. Would you expect the margin to increase or decrease?

4. A linear predictor is used to solve a classification problem with three classes. The data is two-dimensional and the linear functions for each class are:

   - Class 1: $w_1 = (1, 1)$, $b_1 = 0$
   - Class 2: $w_2 = (1, 0)$, $b_2 = 1$
   - Class 3: $w_3 = (0, 1)$, $b_3 = -1$

   Draw the resulting decision boundary and clearly mark the region corresponding to each class.

Part B: Programming problems

1. *Support vector machine.* As you did last week, use the Iris data set, but this time use features 0 and 2, and labels 1,2.

   (a) Is this data linearly separable?

   (b) Use `sklearn.svm.SVC` to fit a support vector machine classifier to the data. You will need to invoke the option `kernel='linear'`. Try at least 10 different values of the slack parameter C. In your writeup, include a table that shows these values of C and for each of them gives the training error and the number of support vectors.

   (c) Which value of C do you think is best? For this value, include a plot of the data points and the linear decision boundary.

2. *Multiclass Perceptron.* Implement the multiclass Perceptron algorithm from class.

   (a) Load in the data set `data0.txt`. This has 2-d data in four classes (coded as 0,1,2,3). Each row consists of three numbers: the two coordinates of the data points and the label.

   (b) Run the multiclass Perceptron algorithm to learn a classifier. Create a plot that shows all the data points (with different colors and shapes for different labels) as well as the decision regions.