DSE 220: Machine learning

Worksheet 9 — Linear classification

The Perceptron algorithm

1. Draw the decision boundary in $\mathbb{R}^2$ that corresponds to the prediction rule $\text{sign}(2x_1 - x_2 - 6)$. Make sure to clearly indicate where this boundary intersects the axes. Show which side of the boundary is classified as positive and which side as negative.

2. A particular labeled data of n points is randomly permuted and then the Perceptron algorithm is run on it, repeatedly cycling through the points until convergence. It converges after making k updates. For each of the following statements, say whether it is **definitely true** or **possibly false**, and give a brief reason.

   (a) The data set is linearly separable.

   (b) If the process were repeated with a different random permutation, it would again converge.

   (c) If the process were repeated with a different random permutation, it would again converge after making k updates.

   (d) k is at most n.

3. The Perceptron algorithm is run on a data set, and converges after performing $p+q$ updates. Of these updates, p are on data points whose label is -1 and q are on data points whose label is $+1$. What is the final value of the parameter b?

Support vector machines

4. 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?

5. 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.

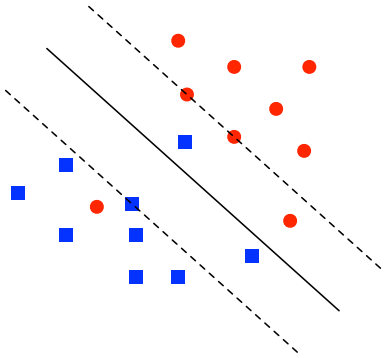
   (b) Draw the left- and right-hand boundaries, also clearly making where they intersect the axes.

   (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?

6. 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?