

## HOMEWORK 8

### NO DUE DATE

Do the following exercises without turning them in. There is no credit for this assignment, but doing these problems will help you prepare for the end sem exam.

**Problem 1.** For  $x \in \mathbb{R}^D$  and  $W \in \mathbb{R}^{K \times D}$ , define the function  $f(x, W) = \text{Softmax}(Wx)$ .

- (a) Write down the  $c$ -th component of  $f$  explicitly.
- (b) Show that if add a vector  $a^T \in \mathbb{R}^D$  to each row of  $W$  then  $f$  is unchanged.
- (c) Calculate  $\frac{\partial f}{\partial w_{kj}}$  and  $\frac{\partial f}{\partial x_j}$ .

**Problem 2.** Consider the following data set, where the first two components are the x- and y- component in the plane and the letter is the class label.

(0,0,A) (0,1,A) (0,2,A) (0,3,A) (0,4,A)  
(1,0,A) (1,1,B) (1,2,B) (1,3,B) (1,4,A)  
(2,0,A) (2,1,B) (2,2,A) (2,3,B) (2,4,A)  
(3,0,A) (3,1,B) (3,2,B) (3,3,B) (3,4,A)  
(4,0,A) (4,1,A) (4,2,A) (4,3,A) (4,4,A)

Draw these points into the plane.

- (a) Draw the decision boundaries of the KNN classifier when  $K = 1$ .
- (b) Draw the decision boundaries of the KNN classifier when  $K = 3$ . How does this algorithm classify the point (2,2)?

**Problem 3.** Count in how many different ways five data points can be assigned to three clusters.

**Problem 4.** Count in how many different ways  $N$  data points can be assigned to two clusters.

**Problem 5.** Suppose the dissimilarity matrix is

$$\begin{bmatrix} & 0.3 & 0.4 & 0.7 \\ 0.3 & & 0.5 & 0.8 \\ 0.4 & 0.5 & & 0.45 \\ 0.7 & 0.8 & 0.45 & \end{bmatrix}$$

- (a) On the basis of this dissimilarity matrix, sketch the dendrogram that results from hierarchically clustering these four observations using complete linkage. Be sure to indicate on the plot the height at which each fusion occurs, as well as the observations corresponding to each leaf in the dendrogram.
- (b) Repeat (a), this time using single linkage clustering.
- (c) Suppose that we cut the dendrogram obtained in (a) such that two clusters result. Which observations are in each cluster?
- (d) Suppose that we cut the dendrogram obtained in (b) such that two clusters result. Which observations are in each cluster?

**Problem 4.** Suppose that for a particular data set, we perform hierarchical clustering using single linkage and using complete linkage. We obtain two dendrograms.

- (a) At a certain point on the single linkage dendrogram, the clusters  $\{1, 2, 3\}$  and  $\{4, 5\}$  fuse. On the complete linkage dendrogram, the clusters  $\{1, 2, 3\}$  and  $\{4, 5\}$  also fuse at a certain point. Which fusion will occur higher on the tree, or will they fuse at the same height, or is there not enough information to tell?
- (b) At a certain point on the single linkage dendrogram, the clusters  $\{5\}$  and  $\{6\}$  fuse. On the complete linkage dendrogram, the clusters  $\{5\}$  and  $\{6\}$  also fuse at a certain point. Which fusion will occur higher on the tree, or will they fuse at the same height, or