CS 6375 Homework Assignment 1 - Due Monday, January 27 2020

1. This question is just to test the probability basics. Suppose you are in a game show and you have the choice of opening three doors. Behind one door is gold, the second is empty while the third has a tiger. Now you pick a door, say door 1 and the host, who knows what's behind the doors, opens a *different* door, say door 3 and shows that it is empty. He then says, "Do you want to pick door 2?". Should you switch? Use Bayes theorem to show what the optimal strategy is.

2. Consider the following data set

Example	x_1	x_2	x_3	x_4	y
1	0	0	1	0	0
2	0	1	0	0	0
3	0	0	1	1	1
4	1	0	0	1	1
5	0	1	1	0	0
6	1	1	0	0	0
7	0	1	0	1	0

   Find the smallest function that can **accurately** capture this data set. Start from just consider one feature to simple conjunctions and then expand to at least m-of-n rules. Explain why this problem is ill-posed.

3. In our class, perceptron was presented as a binary classifier. How can this be extended to a multi-class classification setting?

4. We used hinge loss in the perceptron algorithm. Suppose we have a new loss matrix with the cost of a false positive being $L(1, -1) = c_0$ and the cost of a false negative $L(-1, 1) = c_1$. Note that this setting can happen in several cases where one category is more important than the other (Spam filtering is a classic example of this scenario. It is fine if some junk email comes to our inbox but no important email should be moved to junk). And suppose we used the following loss function

$$J(\mathbf{w}) = \frac{1}{N}\sum_{i=1}^{N} z_i max(0, -y_i \mathbf{w}.x_i)$$

   for our objective function, where $z_i = c_0$ if $y = -1$ and $z_i = c_1$ if $y = 1$. Compute the gradient using this approximation and show how the batch perceptron algorithm has to be modified to incorporate this change.

5. For logistic regression, we defined

$$p_1(\mathbf{x}; \mathbf{w}) = \frac{1}{1 + exp(-\mathbf{w} \cdot \mathbf{x})}$$
$$p_0(\mathbf{x}; \mathbf{w}) = 1 - p_1(\mathbf{x}; \mathbf{w})$$

   Show that this is equivalent to

$$log\frac{p_1(\mathbf{x}; \mathbf{w})}{p_0(\mathbf{x}; \mathbf{w})} = \mathbf{w} \cdot \mathbf{x}$$

   Also, show that

$$p_1(\mathbf{x}; \mathbf{w}) = \frac{exp(\mathbf{w} \cdot \mathbf{x})}{1 + exp(\mathbf{w} \cdot \mathbf{x})} \quad (1)$$

   The equation we considered is called as a logistic function. The above expression is the other way most papers would present the logistic function.