CSE 473: Introduction to Artificial Intelligence

HW 9: Exercise 10 pts

Due November 21 in class. We'll go over it. Early turn-in OK

1. (5 pts) You are given the training set from slide 23 of the Decision Trees packet.

   ```
   X Y Z C
   1 1 1 I
   1 1 0 I
   0 0 1 II
   1 0 0 II
   ```

   In class, I went over the computation of information gain for variables Y (which had the most information) and Z (which had the least). The slide I created in class is posted on Nov 2 as Gain Example. Show how to computer the information gain for variable X. You should get the same value (.31) that was obtained using the alternate formulation for information content.

2. (5 pts) Change the classes to be 1 and 0 instead of I and II. Now we'll try the first epoch of training for a neural net.

   ```
   X Y Z C
   1 1 1 1
   1 1 0 1
   0 0 1 0
   1 0 0 0
   ```

   The neural net has 3 inputs (X, Y, and Z) and one output (C). The initial weights W_X, W_Y, W_Z should be set to 1/3, 1/3, 1/3. The activation function should be the sigmoid function

   $$g(x) = \frac{1}{1+e^{-x}} = \frac{1}{1+\frac{1}{e^x}}$$

   The derivative of the sigmoid function is

   $$g'(x) = x(1-x)$$

   Use the algorithm given in Figure 20.21 to train the network (by hand, no programming needed) with $\alpha = 1$. The stopping criterion is after one epoch. It will show you the direction in which the weights are going, not the final values. Your final answer is just the value of the 3 weights after one epoch. Sorry it's so numeric, but I find this the best way for me to understand what is going on.