CISC 684
Introduction to Machine Learning
Homework 2
Due Date 11:59pm Wednesday March 22
(Please note this homework is due on Wednesday 22nd)
No Late Submissions

Q1 (individual work)
Let X_1, …, X_4 be Boolean variables. Consider the Boolean expression ($\sim X_1 \vee X_2 \vee \sim X_3 \vee \sim X_4$). Provide a decision tree consistent with this expression.

Q2, Consider the training set given below, where X_1, X_2, X_3 and X_4 are attributes and Y is the target value.

Y	X_1	X_2	X_3	X_4
+1	0	1	0	1
+1	1	0	1	0
+1	1	1	1	0
+1	0	0	0	1
+1	1	1	1	0
-1	0	0	1	1
-1	0	0	0	0
-1	0	0	1	0
-1	1	0	0	0
-1	0	0	1	1

a. (Group work) Implement the ID3 algorithm and show what tree results. Please make sure that you are implementing ID3 and are not using some decision tree algorithm/package.
   Make sure your program, at each split point, prints out the entropy at that time and the information gain of each attribute at that time.
b. (Individual work) Design by hand a tree with 4 leaf nodes, 3 internal nodes and depth bounded by two which is consistent with this training data.

Q3, (Group Work) Implement the ID3 algorithm and apply it on the PlayTennis training data. Next, add one or two additional instances (no more than two) to the training data such that "Outlook" is no longer attribute chosen at the root. Provide the extra instances you added and what attribute is chosen at the root. The main point of this exercise is to illustrate that small changes in the training data can result in significant changes to to the model.

Q4. (Individual work) Give a MLP for XOR over 3 boolean-valued variables, x_1, x_2, and x_3. The output of XOR over three variables is 1 if and only if odd number of the variables are 1.

Q5.(Individual work) Show there is a MLP that has outputs < 1,0,0,0>, < 0,1,0,0>, < 0,0,1,0> and < 0,0,0,1> for inputs < 1,0,0,0>, < 0,1,0,0>, < 0,0,1,0> and < 0,0,0,1> respectively. There is no constraints on the output for other input instances. The MLP should have four inputs and four output units.
There should be one hidden layer with two units. The hidden and output units should be thresholded units (not sigmoid units).
Hint. This is not a difficult question. Think about the hidden units values for the 4 inputs and corresponding output node values. Design a perceptron for each of them.