5512, Spring-2019
ASSIGNMENT 5:
Assigned: 04/22/19 Due: 05/05/19 at 11:55 PM (submit via Canvas, you may scan or take a picture of your paper answers) Submit only pdf or txt files (for non-code part), separate submission for code files
Show as much work as possible for all problems!

Problem 1. (20 points)
Write a program that would read in the examples (in csv format, example provided for problem 2) and outputs a decision tree using the heuristic we described in class based on entropy. Display the tree using an array (like how heaps are stored): if the parent is at index i, the left child is 2*i and right is 2*i+1. Assume the csv file is in the base directory of the code and is called "inTree.csv".

This assumes arrays start at index 1, so just make your array 1 size larger than you need and ignore index 0 (assuming your arrays actually start at 0). You may also assume the input file will be formatted like the example and only contain binary inputs and outputs.

The output should assume that "true" or "1" are the left branch and "0" is the right branch. If you are splitting, show the the attribute (you may assume less than 26, so it will be in the alphabet) or if it is a leaf, show the output ("1" or "0"). Use underscores, "_", to separate the nodes of the trees. If no node is present, put a space " ".

For example, the output:
`A_1_B_ _ _0_1`

... corresponds to the tree:
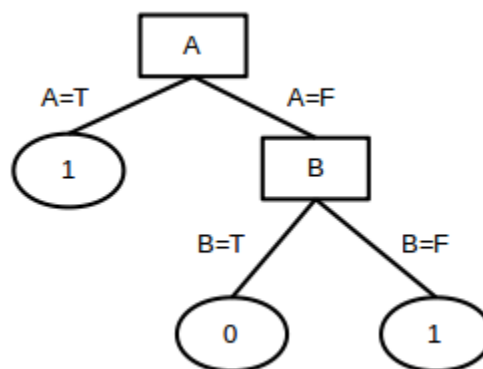
Problem 2. (15 points)

Consider the following table with attributes/inputs A, B, C, D, and E. Find each of the 5 depth-one trees that only use a single attribute/input to make a decision. If you use 3 of these depth-one trees together to vote for the correct output, which three give the best accuracy on the examples below. Discuss the results.

| example	A	B	C	D	E	output
1	0	1	0	1	1	1
2	1	0	1	1	1	1
3	0	0	0	0	1	0
4	1	1	0	0	1	1
5	0	1	0	0	0	0
6	0	0	1	1	0	0
7	1	1	1	1	1	1
8	1	1	1	0	0	0
9	1	0	0	0	1	1
10	1	1	0	0	1	1
11	0	0	0	1	0	1
12	0	1	0	1	0	1
13	0	1	1	1	1	0
14	0	0	1	1	0	0
15	0	1	0	0	0	0
16	0	0	0	0	0	0
17	0	1	0	1	1	1
18	1	1	1	1	0	1
19	1	0	1	0	0	0
20	1	1	0	1	1	1

Problem 3. (15 points)

Given the decision tree below, use the χ^2 (chi-squared) pruning we discussed in class to reduce the overfitting. Use the standard statistics threshold of 95% confident (or less than 5% random). The output of the classification is binary, and the true and false counts for the sub-tables are shown in the nodes. Make sure to not just show what was pruned, but also the resultant decision tree. Rectangles are splits and ovals are leaves.
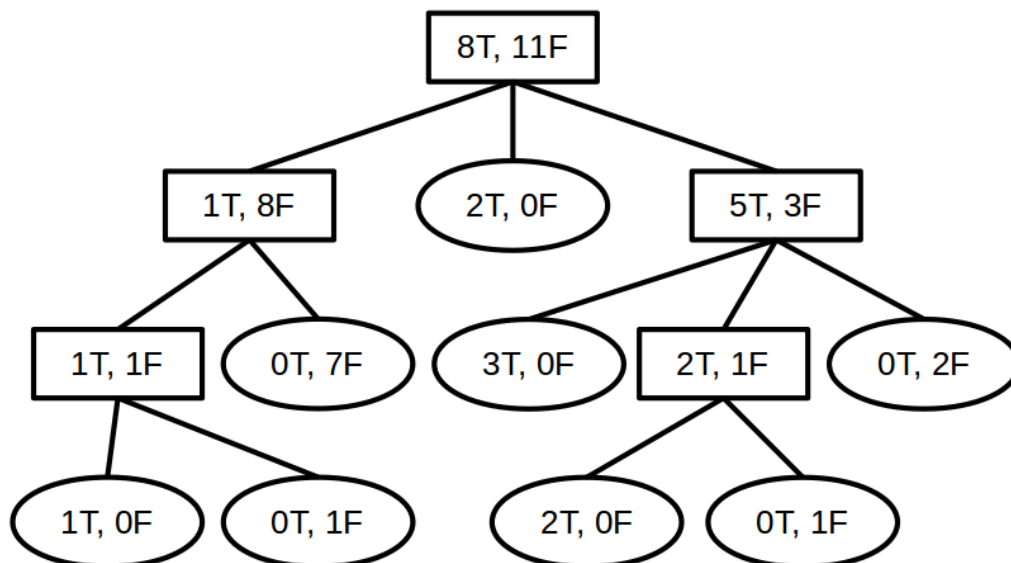
Problems 4 & 5 use the following neural network:
Note: some weights are negative.
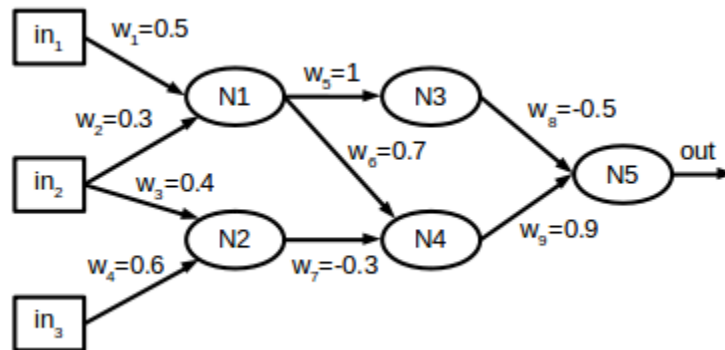
Problem 4. (15 points)
If input (in_1=0.6, in_2=0.5, in_3=0.3) is given and output (out=1) is expected, what should be the new value of weight w_9? (Assume α=0.5.)

Problem 5. (20 points)
If input (in_1=0.6, in_2=0.5, in_3=0.3) is given and output (out=1) is expected, what should be the new value of weight w_1? (Assume α=0.5.) Use the original value for weight w_9, not your updated weight from Problem 4.

Problem 6. (15 points)
Starting at (0,0,0,0), use gradient descent to find the minimum of the function below. (You decide an appropriate α).

$$f(a, b, c, d) = 2 \cdot a^2 \cdot c^4 - a \cdot b + e^{sin(d)+d} + 2 \cdot b^2 \cdot d^2 - 3 \cdot c^3 \cdot d$$