HW1. Due 3/1/19
Instructor: Ranga Raju Vatsavai
Student Name:

This is not a programming a h/w, you don't have to write any code.

1. Decision Trees

- Total 3 attributes (Department, Age, Salary)
    - Department {Sales, Systems, Marketing}
    - Age {21-30, 31-40, 41-50}
    - Salary {Low, Medium, High}
- **Class label variable**: Status {Junior, Senior}
- Total records or objects or instances = 12
- Use Multiway split when constructing the decision tree

Table 1

Department	Age	Salary	Status
Sales	31-40	Medium	Senior
Sales	21-30	Low	Junior
Sales	31-40	Low	Junior
Systems	21-30	Medium	Junior
Systems	31-40	High	Senior
Systems	21-30	Medium	Junior
Systems	41-50	High	Senior
Marketing	31-40	Medium	Senior
Marketing	31-40	Medium	Junior
Marketing	41-50	High	Senior
Marketing	21-30	Low	Junior
Marketing	21-30	Medium	Junior

Using data given in Table 1 as training data, answer the following question:

- **a.** Construct decision tree using Gini index. Show all work and draw the resulting tree.
- **b.** Using the training data given in Table 1, compute accuracy and f-measure (that is, use the tree predict labels, then compare those labels with labels given in the table.
- **c.** For the following test data, predict the class label for each instance using the tree constructed in (a)

Sales	21-30	High	Predicted label
Systems	21-30	Medium	
Marketing	41-50	High	
Marketing	31-40	Low	

2. Naive Bayes Classification

a. State the key assumption made by Naive Bayes Classifier, and justify the assumption.

b. Consider the following dataset given below in **Table 2.**

A	B	C	Class
0	0	0	+
0	0	1	-
0	1	1	-
0	1	1	-
0	0	1	+
1	0	1	+
1	0	1	-
1	0	1	-
1	1	1	+
1	0	1	+

Table 2: Dataset for Naive Bayes Classifier

Estimate the conditional probabilities for

P(A = 0 | +) =	P(A = 0 | -) =
P(A = 1 | +) =	P(A = 1 | -) =
P(B = 0 | +) =	P(B = 0 | -) =
P(B = 1 | +) =	P(B = 1 | -) =
P(C = 0 | +) =	P(C = 0 | -) =
P(C = 1 | +) =	P(C = 1 | -) =

c. (1 point) Predict class label when A=0, B=1, and C=0 using the probabilities computed from (b).

d. (3 points) Estimate the following conditional probabilities using m-estimate approach, with p = 0.5, m = 4.

P(A = 0 | +) =	P(A = 0 | -) =
P(A = 1 | +) =	P(A = 1 | -) =
P(B = 0 | +) =	P(B = 0 | -) =
P(B = 1 | +) =	P(B = 1 | -) =
P(C = 0 | +) =	P(C = 0 | -) =
P(C = 1 | +) =	P(C = 1 | -) =

e. (1 point) Predict class label when A=0, B=1, C=0) using the probabilities computed from (d).

f. (1 point) Compare the two methods for estimating probabilities. Which method is better and why?

Part B (will not be graded, but I encourage everyone to work on the problem and submit solution (will be given extra credit)):

1. Design a neural network for the XOR problem, for the data shown below:

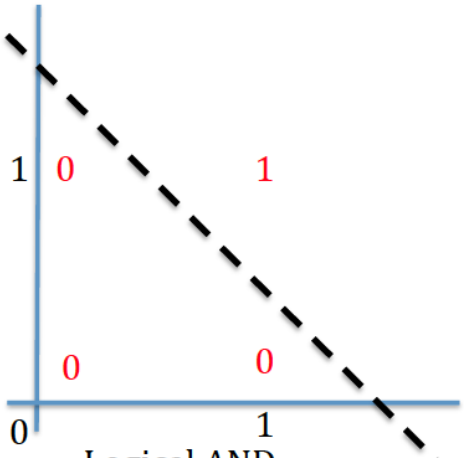
Truth table for XOR problem:

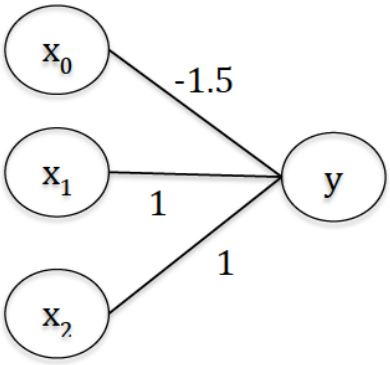
X1	X2	XOR	Class
0	1	1	A
1	0	1	A
0	0	0	B
1	1	0	B

Hint: We know a simple perceptron can't solve this problem. However, a 2 - layer perceptron (with one hidden layer) can. Draw a 2-d scatter plot and draw separating planes.

Your objective is to design the network, show the weights and biases, and show resulting line equations of planes. Using your network, predict class for each of the four inputs i the table.

Here is an example solution for Logical AND, data is represented in the following table (logical AND values for each input pair is shown in red color)





Perceptron Solution

$x_1+x_2-1.5 = 0$

2. Consider the following table:

For the given table, compute the (i) construct contingency table; (ii) compute overall accuracy, (ii) precision, (iv) recall, and (v) F-measure.

Classification Prediction 1	**Ground-truth (Actual)**
+	+
-	-
+	+
+	+
+	+
-	+
-	+
-	-
-	+
-	-
-	+
-	-
-	+
+	+
+	-

	Predicted +	Predicted -
Actual +		
Actual -		