Decision Tree Practice Problems:

1. You have a friend who only does one of four things on every Saturday afternoon: go shopping, watch a movie, play tennis, or just stay in. You have observed your friend's behavior over 11 different weekends. On each of these weekends you have noted the weather (sunny, windy, or rainy), whether her parents visit (visit or no-visit), whether she has drawn cash from an ATM machine (rich or poor), and whether she had an exam during the coming week (exam or no-exam). You have built the following data table:

# ex.	Weather	Parents	Cash	Exam	Decision
1	sunny	visit	rich	yes	cinema
2	sunny	no-visit	rich	no	tennis
3	windy	visit	rich	no	cinema
4	rainy	visit	poor	yes	cinema
5	rainy	no-visit	rich	no	stay-in
6	rainy	visit	poor	no	cinema
7	windy	no-visit	poor	yes	cinema
8	windy	no-visit	rich	yes	shopping
9	windy	visit	rich	no	cinema
10	sunny	no-visit	rich	no	tennis
11	sunny	no-visit	poor	yes	tennis

You now want to build a decision tree to predict the activity of your friend on any future Saturday afternoon from the observed values of Weather, Parents, Cash, and Exam.

i. Suppose that you want to build a very simple decision tree that allows you to predict the value of Decision from a single observable attribute (Weather, Parents, Cash, or Exam). What would be this attribute? If there is a tie among several attributes give each of them. For each attribute that you have selected for each value of this attribute, give the answer that the decision tree would give using the majority rule, as well as the number of misclassified examples.

ii. Draw the full tree that correctly classifies all the examples.

2. At the beginning of an exam, you try to predict whether each problem is easy or difficult (D = + if it is difficult and – if it is easy). Let us assume that you use two observable problem attributes (predicates):
   - The text length L of the problem (L = 1 if it is long, 0 otherwise)
   - The amount M of math in the text (M = 1 if there is a lot of math, 0 otherwise)

   For training data, assume that you have examined 12 previous problems from the homeworks, and have collected the following data:

L	M	D	#
0	0	–	4
0	0	+	1
0	1	–	0
0	1	+	3
1	0	–	1
1	0	+	2
1	1	–	1
1	1	+	0

The first line of this table reads as follows: 4 problems for which L = 0 and M = 0 were

not difficult (D = –). The second line says: 1 problem for which L = 0 and M = 0 was difficult (D = +). Etc… Note that you observed no problem for which L = 0 and M = 1, or L = 1 and M = 1.

Based on this training data, you want to compute a representation of a difficult problem (D) in the form of a decision tree using the two binary attributes L and M. Construct the best decision tree you can for the training data.

3. Consider the following table of observations:

No.	Outlook	Temperature	Humidity	Windy	Play Golf?
1	sunny	hot	high	false	N
2	sunny	hot	high	true	N
3	overcast	hot	high	false	Y
4	rain	mild	high	false	Y
5	rain	cool	normal	false	Y
6	rain	cool	normal	true	N
7	overcast	cool	normal	true	Y
8	sunny	mild	high	false	N
9	sunny	cool	normal	false	Y
10	rain	mild	normal	false	Y
11	sunny	mild	normal	true	Y
12	overcast	mild	high	true	Y
13	overcast	hot	normal	false	Y
14	rain	mild	high	true	N

From the classified examples in the above table, construct two decision trees for the classification "Play Golf." For the first tree, use Temperature as the root node. (This is a really bad choice.) Continue, using your best judgment for selecting other attributes. Remember that different attributes can be used in different branches on a given level of the tree. For the second tree, follow the Decision-Tree-Learning algorithm. As your Choose-Attribute function, choose the attribute with the highest information gain. Draw the decision trees.

Proof that if the same attribute appears twice in a decision tree, than the information gain equals zero when you apply it for the second time. You should assume that it is possible that other attributes are tested between the two applications of this attribute.