1. We consider the training examples shown in the following table for a binary classification problem.

Instance	a_1	a_2	a_3	Target Class
1	T	T	1	+
2	T	T	6	+
3	T	F	5	-
4	F	F	4	+
5	F	T	7	-
6	F	T	3	-
7	F	F	8	-
8	T	F	7	+
9	F	T	5	-

   a) What is the original entropy of this set of training instances?

   b) What are the information gains when a_1 and a_2 are used for partitioning the training set respectively?

2. We again consider the training examples shown in Q.1

   a) Calculate the respective changes in the Gini index value when a_1 and a_2 are used for partitioning the training set.

   b) Calculate the respective changes in the classification error when a_1 and a_2 are used for partitioning the training set.

c) For a_3, which is a continuous attribute, compute the information gain for every possible split. What is the best threshold for splitting the set of attribute values?