Homework 6

Due: 14 April 2016

Consider the following data set, representing 16 data points with attributes A, B, and C and "outcome" attribute R:

A	B	C	R
T	T	F	−
T	T	F	−
T	F	F	−
T	F	F	+
T	F	F	+
T	F	T	+
T	T	T	+
T	T	T	+
F	T	T	+
F	T	T	−
F	T	T	−
F	T	T	−
F	F	T	+
F	F	T	+
F	F	T	+
F	F	T	+

Problem 6.1

Give the entropy of the overall data set with respect to the random variable R; compute the information gain represented by attributes A, B, and C; and discuss what your computed gain tells us about the decision tree that would be built to predict R.

Problem 6.2

On this data set, how many decision nodes would the final decision tree have? Assuming that the data is distributed exactly like this training set, what is the best performance you could expect from the tree? How does this compare to the performance you'd expect with no tree, or the one-node only-the-root tree that you built in the previous problem? Discuss what else would be needed to further improve on that performance.