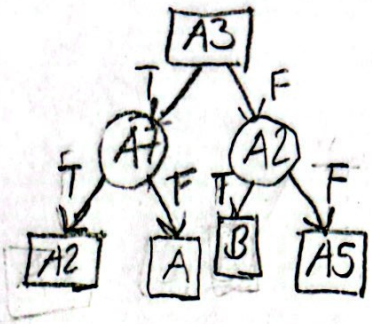
1. Here[1] is some randomly generated data for two classes, A and B, with eight boolean attributes. Generate the first two levels of a decision tree for these classes. You should write some code to help you; how much you want to automate is up to you but you can reuse it code later in Lab3. Specifically, you should identify which of the eight attributes to test first, based on information gain, and for either value of that attribute, which one to test next - then, for each of these four cases, which prediction to give (A or B). Draw your tree in the space provided below.



2. If you were to run Adaboost on the preceding problem, using decision stumps, assuming the use of information gain, consider the first iteration of the Adaboost algorithm. The first stump should be the same as the root of your previous tree.

(a) 0.00667 What would be the initial weights of each example?

$W_i = \frac{1}{N} = \frac{1}{150}$

(b) 54 How many items were misclassified by the first stump?

38+16 = 54

(c) 0.360 What would the error rate of the first stump be?

(d) 0.2877 What would be the hypothesis weight of the first stump?

(e) 0.005 What would be the new weights of each example that was classified *correctly*?

(f) 0.008 What would be the new weights of each example that was classified *incorrectly*?

[1] https://cs.rit.edu/~jro/courses/intelSys/hw/dtada/dtree-data.dat