CS 6375 - Quiz 2

Name _____

1. You are playing a betting game, if you win you get $1 and if you lose you give $3. According to the odds you have 80% chance of winning, what is the expected winning amount for you.

A. $1.4
B. $1
C. $0.2
D. $0

2. True- False: Overfitting is more likely when you have huge amount of data to train?
   A. TRUE
   B. FALSE

3. In which of the following scenario a gain ratio is preferred over Information Gain?
   A. When a categorical variable has very large number of category
   B. When a categorical variable has very small number of category
   C. Number of categories is the not the reason
   D. None of these

4. There are 3 boxes, A, B and C. The probability of choosing box A, B and C are 0.5, 0.3 and 0.2 respectively. Box A, B and C contains 20%, 30% and 50% rotten apples. What is the probability that an apple is drawn from box A given that it is rotten? To solve this kind of problem, what theorem should we use?

   Let A = event that an apple is drawn from box A, B = event that an apple is drawn from box B, C = event that an apple is drawn from box C, R = a rotten apple is drawn.

   Given P(A) = 0.5, P(B) = 0.3, P(C) = 0.2.
   Also, P(R|A) = 0.2, P(R|B) = 0.3, P(R|C) = 0.5.
   Our aim is to find P(A|R) = P(A intersect R)/P(R)
   = P(A) x P(R|A) / [P(A) x P(R|A) + P(B) x P(R|B) + P(C) x P(R|C)]
   = 0.5 x 0.2 / (0.5 x 0.2 + 0.3 x 0.3 + 0.2 x 0.5)
   = 0.3448.

   Bayesian inference is a branch of statistic inference that arises from this theorem.

5. Give decision tree to represent the boolean function: A ∨ [B ∧ ¬C]