CS181 Assignment 1

Joy Ming and Alisa Nguyen (9 February 2013)

1. Decision Trees and ID3

   (a) ID3 will chose to split on $\boxed{A}$ because it has a higher information gain.

      - Splitting on B will have an information gain of $Gain(X_k, B) = H(B) - Remainder(X_k, B) = ?? - \frac{2}{7}(\frac{1}{2}\log_2 2 + \frac{1}{2}\log_2 2 + \frac{3}{7}(\frac{2}{3}\log_2\frac{3}{2} + \frac{1}{3}\log_2 3)$

2. ID3 and Pruning

3. Boosting

4. Tree Analysis