

CS 181 - Homework 1

Lexi Ross & Ye Zhao

February 13, 2013

1 Decision Tree & ID3

2 ID3 with Pruning

3 Boosting

3.1 Information Gain Criterion using weights

In order to take into account weights when calculating the information gain of a given attribute, we use the following formula to calculate the entropy of the labels:

$$H(\text{labels}) = \sum_{c=1}^C \frac{W_c}{W} \log_2 \frac{W}{W_c}$$

Where

$$W_c = \sum_{n=1}^N I(y_n = c) w_n$$

and

$$W = \sum_{n=1}^N w_n$$