Notes from CS 450

Daniel Craig
Department of Computer Science
BYU Idaho - Class of 2019

October 11, 2018

1 Notes from Chapter 1

I have not read Chapter 1 yet.

2 Notes from Chapter 2

"Accuracy is defined as the sum of the number of true positives and true negatives divided by the total number of examples."

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

"Sensitivity (also known as the true positive rate) is the ratio of the number of correct positive examples to the number classified as positive"

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

"specificity is the same ratio for negative examples"

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

"Precision is the ratio of correct positive examples to the number of actual positive examples"

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall is the ratio of the number of correct positive examples out of those that were classified as positive, which is the same as sensitivity.

$$\frac{TP}{TP + FN} \tag{5}$$

Bayes rule is the most important equation in machine learning, and it is as follows:

$$P(C_i|X_j) = \frac{P(X_j|C_i)P(C_i)}{P(X_j 1)} \tag{6}$$

"Bayes rule relates the posterior probability

$$P(C_i|X_j)$$

with the prior probability

$$P(C_i)$$

and class conditional probability

$$P(X_j|C_i)$$

. The denominator (the term on the bottom of the fraction) acts to normalise everything, so that all the probabilities sum to 1."

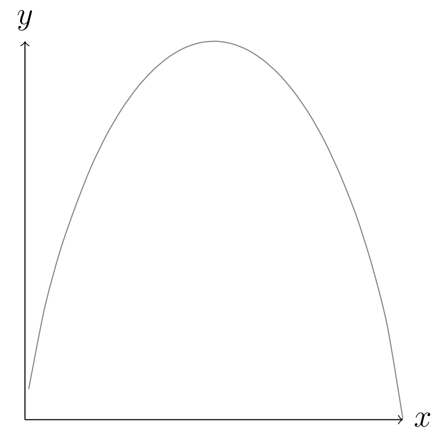

Figure 1: The Binary Entropy Function

3 Notes from Chapter 12

The highest information gain is achieved by selecting for the lowest entropy. See figure 1.

Entropy is defined as the following:

$$Entropy(p) = -\sum_{i} p_i \log_2 p_i \tag{7}$$

Information gain is defined as "the entropy of the whole set minus the entropy when a particular feature is chosen."

$$Gain(S, F) = Entropy(S) - \sum_{f \in values(F)} \frac{\mid S_f \mid}{\mid S \mid} Entropy(S_f) \tag{8}$$