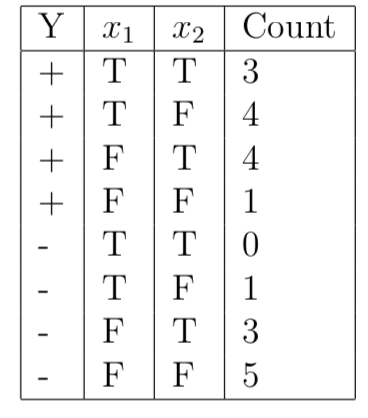
Chris Kasper

CS 383-002

Homework #4

**1. Theory**

1.



1. Entropy of system:
2. Information Gains for x1 and x2:

For x1:

T – (7/8, +), (1/8, -)

F – (5/13, +), (8/13, -)

For x2:

T – (7/10, +), (3/10, -)

F – (5/11, +), (6/11, -)

1. Decision tree using ID3 algorithm:

X1

F

T

X2

X2

T

F

-

+

F

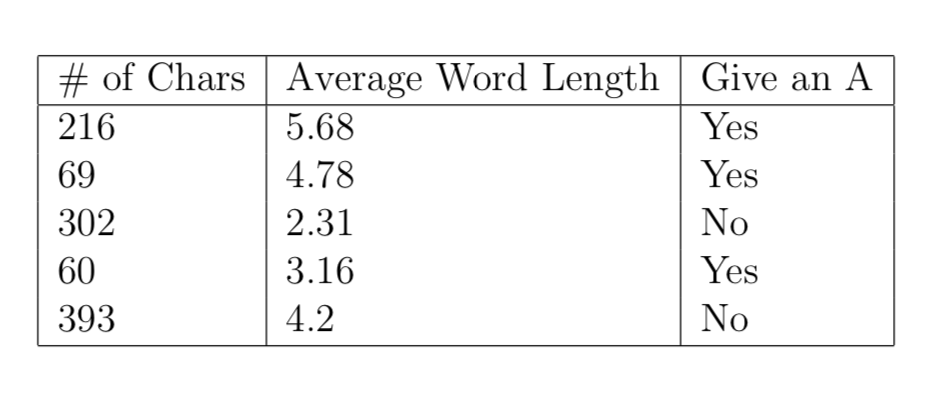
T

+

+

Note: x2 nodes, when x1 is T, could be collapsed to just “+”

2.



1. Class priors:
2. Finding Gaussian parameters to do Gaussian Naïve Bayes classification on the decision to give an A or not:

,

Gaussian Parameters for A=Y models:

Gaussian Parameters for A=N models:

1. Given an essay with 242 characters and an average word length of 4.56, determine whether or not it would get an A.

-------

3.

1. Augment log likelihood objective function to deal with the range of -1 <= tanh(z) <= 1

For : we need to account for the negative values that could be returned from because the equation is bound from -1 to 1. When , we want the probability to 1. When , we want the probability to 0. When , we want the probability to 0.5, hence why I add 1 to and then divide by 2. The exponent accounts for the two binary outcomes, 1 and -1. When y=1, the exponent = 1, and when y= -1, the exponent = 0.

For : We just need to subtract 1 by. The exponent accounts for the two binary outcomes, 1 and -1. When y=1, the exponent = 0, and when y= -1, the exponent = 1.

Therefore, the new log likelihood objective function is …

1. Show that:

Using quotient rule…

1. Derivative of log likelihood function

**2. Naïve Bayes Classifier**

|  |  |
| --- | --- |
| Precision: | 68.16% |
| Recall: | 95.75% |
| F-Measure: | 79.63% |
| Accuracy: | 81.23% |