William Frank
4/18/19
HW10: Classification

Maximum Likelihood Parameters:

Word	P(Word | Bob)	P(Word | Clarence)
all	1/28	1/18
assignment	2/28	2/18
by	0/28	1/18
come	1/28	0/18
did	2/28	1/18
fail	0/28	1/18
graded	0/28	1/18
grading	1/28	0/18
great	2/28	0/18
help	1/28	0/18
is	0/28	1/18
my	1/28	0/18
need	1/28	0/18
no	0/28	1/18
office	1/28	0/18
on	1/28	1/18
one	0/28	1/18
should	2/28	1/18
students	2/28	1/18
ta	2/28	1/19
talk	1/28	0/18
the	2/28	1/18
this	2/28	1/18
time	0/28	1/18
to	2/28	0/18
you	1/28	0/18

1. “you did great”:

P(Bob | “you did great”) ≈ (4/7) (1/28) (2/28) (2/28) = 1/9604
P(Clarence | “you did great”) ≈ (3/7) (0/18) (1/18) (0/18) = 0
Classified as from Bob

2. “no students should fail”:

P(Bob | “no students should fail”) ≈ (4/7) (0/28) (2/28) (2/28) (0/28) = 0
P(Clarence | “no students should fail”) ≈ (3/7) (1/18) (1/18) (1/18) (1/18) = 1/244944
Classified as from Clarence

3. Did the classifier do what you think it should? If not, why not?

The classifier did what I expected it to do, but maybe not what it should do. The positive sentiment in the second message seems much more like Bob, but because he has never used the words “fail” or “no” before he was given a probability of 0. The classifier uses word frequency rather than some deeper understanding of sentiment.

Laplace Smoothing Parameters:

Word	P(Word | Bob)	P(Word | Clarence)
all	2/54	2/44
assignment	3/54	3/44
by	1/54	2/44
come	2/54	1/44
did	3/54	2/44
fail	1/54	2/44
graded	1/54	2/44
grading	2/54	1/44
great	3/54	1/44
help	2/54	1/44
is	1/54	2/44
my	2/54	1/44
need	2/54	1/44
no	1/54	2/44
office	2/54	1/44
on	2/54	2/44
one	1/54	2/44
should	3/54	2/44
students	3/54	2/44
ta	3/54	2/44
talk	2/54	1/44
the	3/54	2/44

this	3/54	2/44
time	1/54	2/44
to	3/54	1/44
you	2/54	1/44

1. “you did great”:

P(Bob | “you did great”) ≈ (4/7) (2/54) (3/54) (3/54) = 1/15309 ≈
P(Clarence | “you did great”) ≈ (3/7) (1/44) (2/44) (1/44) = 3/298144 ≈
Classified as from Bob

2. “no student should fail”:

P(Bob | “no students should fail”) ≈ (4/7) (1/54) (3/54) (3/54) (1/54) = 1/1653372 ≈ 6.05e-7
P(Clarence | “no students should fail”) ≈ (3/7) (2/44) (2/44) (2/44) (2/44) = 3/1639792 ≈ 1.83e-6
Classified as from Clarence

3. Did the classifier do what you think it should? If not, why not? Did the classifications change?

The classifier still made the same classifications as before, but it did not give anyone a 0 probability. The second sentence gave a much higher probability to Bob than it did with no smoothing, but it was still classified as coming from Clarence. Whether or not this is a correct classification is unknown, but the bag of words approach does not capture sentiment very well.