Name: VEN THON
ID: e20191250
Group: I5-GIC-C

Train two models, multinominal naive Bayes and binarized naive Bayes, both with add-1 smoothing, on the following document counts for key sentiment words, with positive or negative class assigned as noted.

doc	"good"	"poor"	"great"	(class)
d1.	3	0	3	pos
d2.	0	1	2	pos
d3.	1	3	0	neg
d4.	1	5	2	neg
d5.	0	2	0	neg

Use both naive Bayes models to assign a class (pos or neg) to this sentence:
A good, good plot and great characters, but poor acting. Do the two models agree or disagree?

Answer:
Count of positive documents (pos): 2
Count of negative documents (neg): 3
Total number of documents: 5
P(class = pos) = 2/5
P(class = neg) = 3/5

Calculating the probabilities of each word occurring in each class:
P(good|pos) = (3 + 1) / (3 + 1 + 1) = 4/5
P(good|neg) = (1 + 1) / (1 + 3 + 1) = 2/6 = 1/3

P(poor|pos) = (0 + 1) / (3 + 1 + 1) = 1/5
P(poor|neg) = (3 + 1) / (1 + 3 + 1) = 4/6 = 2/3

P(great|pos) = (3 + 1) / (3 + 1 + 1) = 4/5
P(great|neg) = (0 + 1) / (1 + 3 + 1) = 1/6

Now, to classify the sentence "A good, good plot and great characters, but poor acting" using the multinomial Naive Bayes model, we need to calculate the probabilities for each class.

P(class = pos|sentence) = P(pos) * P(good|pos)^2 * P(poor|pos) * P(great|pos)
= (2/5) * (4/5)^2 * (1/5) * (4/5)
= 128/625
P(class = neg|sentence) = P(neg) * P(good|neg)^2 * P(poor|neg) * P(great|neg)
= (3/5) * (1/3)^2 * (2/3) * (1/6)
= 2/45
So that P(class = pos|sentence) =128/625
P(class = neg|sentence) = 2/45