

report

in part 1, when i was training the dataset, i use nested loops to store the conditional probability of each features given each label.  
i named them feature result ,there are 57 features and each feature has 7 types.

so the parameter list i create for the first part is  $57 \times 7 \times 3$ , because there are 57 tables each has 7 types of value, which is 0 1 2 3 4 5 6, and three labels 0 1 2

for testing, i use naive bayes formula to multiply each conditional probability of each types of feature, and multiply the probability of each label.  
then compare the probability of each labels of that sample, to check with the valid label. Then calculate the accuracy.

for the second part, i store the mean of yes , mean of no, standard deviation of yes, standard deviation of no, so i need to store 4 values for each of 57 features.  
and i create a  $2 \times 2 \times 57$  table after the training.

for testing, i use the Gaussian models to calculate each conditional probability for each feature, and multiply them together, then times the probability of each label 0 and 1. Compare the probability of each label and choose the highest one. compare to the valid label to calculate the accuracy.

for the last part, i pick 460 values in spambase.data for testing and use the others for training. i do it 10 times and record the accuracy for each 461 test samples. It is almost the same as the previous one.

the test result for part1 shows the accuracy is 0.865652  
the test result for part2 shows the accuracy is 0.782826 for testing avc\_c2  
the test result for part2 shows the mean of accuracy is 0.815835 and the standard deviation of accuracy is 0.149262 for testing spambase.data

in part1, the zero condition will decrease the value of accuracy, because in naive bayes formula, when i multiply the conditional probability, one zero value will cause the result to be zero, so in some case it will make the result to be wrong.

in part2, the zero condition has the same effect, the accuracy get higher after i fixed.