Lab exersizes

1

A

Most frequent:

Validation: 14%

Testing: 14%

Naïve bayes:

Validation: 69%

Testing: 55%

Smoothing parameter: 2.00000000

Most frequents correct rate depends entirely on the amount of options and the frequency of said option, 14% could very well be achieved by having 8 options with equal chance, or 2000 options where one pops up 14% of the time. Naïve bayes will do a lot better because it also looks at the properties of said options, it doesn’t relate those to eachother so the degree of error is still quite high. For instance measuring holes in numbers will achieve the same amount of holes, height and with for both 6 and 9 making it prone to errrors for those kind of cases.

B

Laplace smoothing makes sure that if a single property is particularly weird that doesn’t affect the total chance too badly. This makes sure that if something is very out of the ordinary it will still look at the other properties instead of choosing something somewhat at random.

C

This option tries a large array of smoothing parameters to see which one works best and then uses the optimal smoothing parameter. this will generally lead to better results, which it also does here in comparison to a. (74 compared to 69)

D

There be 784. Log 784 is less then 784.

E

Validation is probs on the data that you tuned it on, test on different data, so you see whether it’s overfitted and stuff