

4a 20% training & 80% testing [n = 120]

trial 1

	prediction		
	S	ver	vir
actual	S	39	0
	ver	12	25
	vir	0	37

overall 86% accurate

trial 2

	prediction		
	S	ver	vir
actual	S	38	0
	ver	2	41
	vir	4	30

overall 90% accurate

→ 50% split [n = 75]

trial 1

	prediction		
	S	ver	vir
actual	S	23	0
	ver	1	26
	vir	0	23

overall 96% accurate

trial 2

	prediction		
	S	ver	vir
actual	S	27	0
	ver	2	24
	vir	0	22

overall 97% accurate

note: trial 2 is in white as the first trial ran was more important to note than the second; we may have only have one try in a real world scenario. also I noticed that sometimes the smaller training groups ($p=0.2$ for ex) was so biased that it was able to perfectly predict on some trials.

→ 70% train, 30% test.

trial 1

	prediction		
	S	ver	vir
actual	S	16	0
	ver	0	14
	vir	0	14

overall 97% accurate

trial 2

	prediction		
	S	ver	vir
actual	S	17	0
	ver	1	9
	vir	0	18

overall 97% accurate

→ 90% train, 10% test.

trial 1

	prediction		
	S	ver	vir
actual	S	6	0
	ver	0	2
	vir	0	7

overall 100% accurate

trial 2

	prediction		
	S	ver	vir
actual	S	6	0
	ver	0	7
	vir	1	0

overall 93% accurate

note: as noted in the ipynb, not even a 90/10 split can guarantee 100% accuracy, especially with random splits. This is because some bias may be introduced, and/or some samples could be huge outliers.