Support vector machines

Martynas Lukosevicius, Alejo Perez Gomez, Shwetha Vandagadde Chandramouly

07/12/2020

2. SUPPORT VECTOR MACHINES

Test results:

	error	trained on	tested on
filter0	0.07	train	validation
filter1	0.0848938826466916	train	test
filter2	0.083645443196005	train + validation	test
filter3	0.0337078651685393	all data (spam)	test

1. Which model do we return to the user ? filter0, filter1, filter2 or filter3 ? Why ?

Model should be returned with smallest error, however its hard to comapre filter3 with others because it was tested on data which was used for training. Also filter0 was not tested on test dataset, as a result error might differ.

Taking into account all arguments, filter2 should be selected as it was trained on bigger dataset (test + validation) and error is smallest between those which can be compared. (filter1 and filter2)

2. What is the estimate of the generalization error of the model selected ? err0, err1, err2 or err3 ? Why ?

Generalization error is the error of the model on data which was not used for training, so err3 is not a generalization error. For chosen model (filter2) generalization error is err2 - 0.0836454