Name: Hong Pengfei ID: 1002949

I used sklearn.svm.SVC as my SVM, and I used the default kernel “RBF” kernel in it.

I trained on training set and test on validation set, And got the following result:

|  |  |  |
| --- | --- | --- |
| C value | Class wise Accuracy | accuracy |
| 0.01 | 0.6 | 0.77 |
| 0.1 | 0.61 | 0.77 |
| 0.32 | 0.61 | 0.77 |
| 1 | 0.62 | 0.78 |
| 3.16 | 0.61 | 0.79 |
| 10 | 0.6 | 0.79 |
| … | … | … |

The C that gives the best performance is C=1.

I trained on validation set + training set and tested on test set, and got the following result:

classwise accuracy on test set: 0.5482886732886733

vanilla accuracy on test set: 0.7529411764705882

Why split classwise

Split classwise means split the class proportionally between training and test set.

Split randomly is just randomly split the dataset with no constraints so it cannot make sure that the proportions of the class in training, validation and test set is the same.

You asked us to split classwise so that training set is a good representation/reflection of the test set.

For more information, please see the code in jupyter notebook.