

hw3

Hongyi Guo 516030910306

March 22, 2019

1

`make` in the `libsvm` folder and `make` in the `python` folder.

2

Grid search parameter c and γ in $\{2^{-10}, 2^{-9}, \dots, 2^{10}\}$. The i^{th} line means c equals to 2^{i-10} . The j^{th} column means γ equals to 2^{j-10} . Note that the results are tested on test set for convenience. I do not use a validation set although that would be better from my perspective.

[illegible]

Figure 1: One-versus-rest with linear kernel

[illegible]

Figure 2: One-versus-one with linear kernel

[illegible]

Figure 3: One-versus-rest with RBF kernel

[illegible]

Figure 4: One-versus-one with RBF kernel

3

Results are as in Fig.1, Fig.2, Fig.3, Fig.4

4

4.1 OvR v.s. OvO

One-versus-rest method runs faster than one-versus-one when there are more classes with no doubt. But in our case where only three classes are considered, the total number of models used in both methods are equal (and are 3). But because each model of the one-versus-rest method needs all data from training set. It may trains slightly slower than one-versus-one method.

The biggest problem in OvO is that the voting may cause a tie, which means -1 defeats 0, 0 defeats 1 and 1 defeats -1. This is rare but not impossible. OvR use the probability(confidence score) to vote thus doesn't have to consider ties but suffers from the problem of unbalanced distributions of 'one' and 'rest'.

In my experiment, OvO always outperforms OvR given any kernel.

4.2 Linear v.s. RBF

Typically, RBF kernel should outperform linear kernel because it's more complex and can handle more difficult situation. In my experiment, both linear kernel and RBF kernel can

reach a score of about 83.1%.

From the visualization in Fig.5, -1 and 0 are very close in distribution. After deeper inspection, I find that both linear and RBF kernel fails to classify -1 and 0 in the test set. I think the training set may not be large enough to build a very robust classifier.

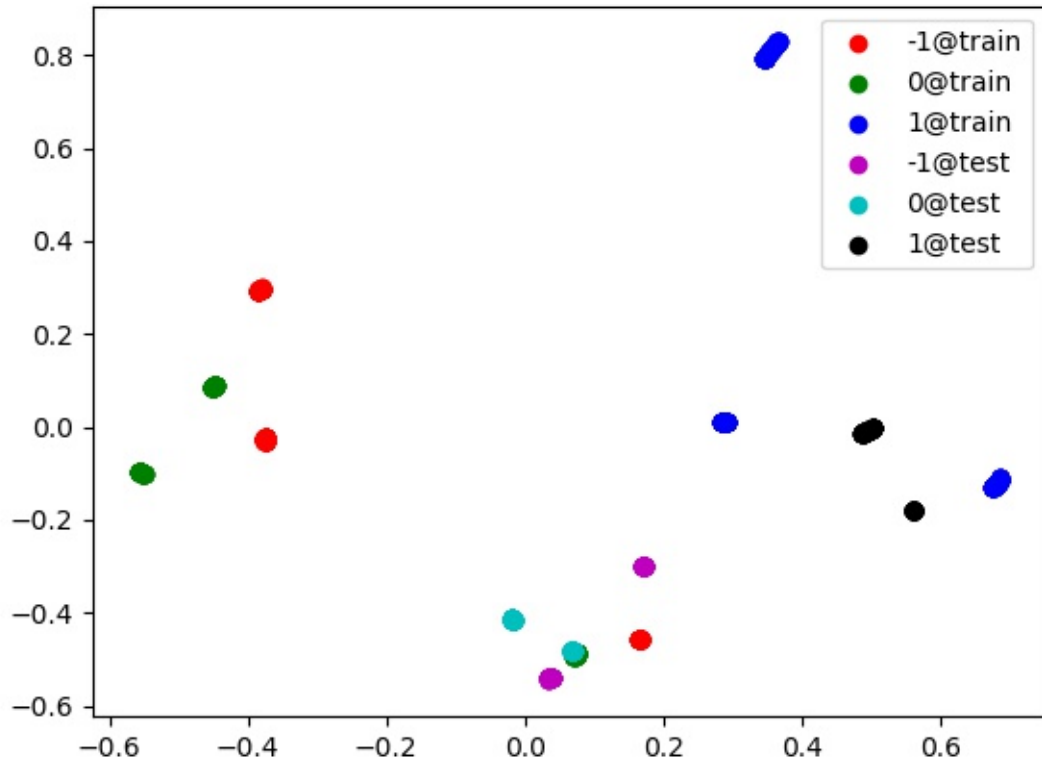


Figure 5: Data distribution