Computer Exercise 3

Sept. 26, 2019

Goal:

Download a free SVM package, and experiment with the given data to explore the effect of different choices of package on the performance. The choices include kernels, kernel parameters and the C.

Learning task:

Using 10 gene expression features to classify two cell types (fibroblast vs. endothelial cell)

Datasets	Sample size	Feature data file	Class label file
TrainingSet-1	40	train_10gene_sub.csv	train_10gene_label_sub.csv
TrainingSet-2	303	train_10gene.csv	train_label.csv
TestSet	100	test_10gene.csv	test_label.csv
TestSet-2	263	test2_10gene.csv	test2_label.csv

Note: TrainingSet-2 is the combination of samples in TrainingSet-1 and TestSet-2.

（Data are the same as those used in Ex.2.）

- **Experiment:**

  Find and download a free SVM package and learn its usage.

  Use TrainingSet-1 to train the SVM with linear kernel and with Gaussian or polynomial kernel with at least 3 different sets of parameters. Apply the trained SVM on TestSet and TestSet-2 separately. Calculate the training error, the cross-validation error on the training set, the test errors on the two test sets.

  Use TrainingSet-2 to train the SVM. Apply the trained SVM on TestSet. Calculate the training error, the cross-validation error on the training set, the test error.

  Analyze and discuss observations in the experiments.

- **Experiment Report**

  Write an experiment report to describe and analyze the experiment observations.

  Provide detailed supplementary materials that should include at least the following materials:

  - A readme file that contain information on all the supplementary files, programming environment and parameters used in the experiments (if any)
  - A brief introduction of the SVM package used, and cripts or command lines used in the experiments (should let TAs to be able to run the code and reproduce your experiments)
  - Experiment result files

  All files should be packed in one file for submission. Acceptable formats are .zip, .rar, and .7z.

Report due date: Oct. 9 (Wednesday), 23:59 (same for Ex2)