Computer Exercise 1

Sept. 12, 2019

Goal:

Write a computer program of the Perceptron Algorithm in Matlab, Octave or Python, and conduct experiments on the given data to explore the behavior of the algorithm in different situations. The program should include both the "Fixed increment rule" and the "Variable increment rule", and should include a strategy to force convergence when the samples are not linearly separable.

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

- **Experiment 1:**

  Use TrainingSet-1 to train the Perceptron, compare the numbers of iterations needed for the learning process to converge using the "Fixed increment rule" and the "Variable increment rule", respectively. Apply the trained perceptrons on TrainingSet-2 and on TestSet, and compare the error rates on the two sets.

- **Experiment 2:**

  Use TrainingSet-2 to train the Perceptron with only one option of the increment rules, without forcing the training to stop for a large number of iterations. Stop arbitrarily after a long period of training, calculate the error rate on the training data, and apply the obtained perceptron on TestSet to get the test error rate.

  Redo the above experiment but taking certain strategy to force the training process to converge after a reasonable period of training. Check the error rate on the training data, and apply the obtained perceptron on TestSet to get the test error rate.

- **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)
  - Source code (should allow TAs to execute the code)
  - Experiment result files

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

Report due date: Sept. 29 (Saturday), 23:59, 2019