Xiao Zhou zhouxiao@bu.edu

CS 542 Problem Set 4

Discussed with classmates

4/12/2017

1. Written Problems
   1. Bishop 6.2

If we set w = 0, then the last w can be represented as

Where specify the number of times pattern n is used in training

The prediction is given by

Where is the kernel function, and the learning algorithm would be

For misclassified patterns, we can have

Then

Which can be considered as

Thus the learning algorithm only depend on the elements of the gram matrix.

* 1. Bishop 7.3

Suppose a data set has only 2 points, , each belong to one class, then maximum margin hyperplane can be find by

Where

Using Lagrange multipliers,

So we can have

and

* 1. Bishop 7.4

Since ,

For max margin solution,

Thus from 7.10 we can have

Consequently,

1. Programming

Codes and reports are in ‘.m’ files and ‘report.docx/pdf’ file.

The code takes a while to complete, time varies based on the hardware.