

Machine Learning Module

Laboratory Course Work

Issued Monday 30th January

Submission Deadline Friday 17th March

Mark Girolami

`girolami@dcs.gla.ac.uk`

Department of Computing Science

University of Glasgow

January 29, 2006

This coursework is **based** on the laboratory exercise set in Week 3 of the course. You will be expected to submit a detailed report which describes the experiments you undertook and analyses the results obtained.

1 Using Cross Validation to Select the Optimal Prior and Radial-Basis Hyper-Parameters

In a real application there is no possibility to generate test-sets on which to assess how good your model is and we introduced LOOCV as a general model assessment and selection method. In our model we now have a pair of hyper-parameters which will index the possible model set, so we will have quite a large number of α & β combinations to consider.

Write a Matlab script which will do the following

1. Generate a *train* and *test* set for the sinc function for a pre-set and known noise level.
2. For a range of α & β values compute
 - (a) the *MSE* on the *training* data
 - (b) the *MSE* on the *test* data
 - (c) the LOOCV *MSE*
3. Identify the α , β pair that yields the minimum *train*, *test* and LOOCV errors
4. How well does LOOCV do in terms of locating optimal α , β pairings?
5. The results you obtain will be for one *training* set. Generate multiple *training* data sets and repeat 1 - 3 above for each one. Discuss how the values obtained in 2(a),(b),(c) and 3, vary across multiple training instances?