# Regression on LETOR



## Introduction

The aim of this entry is to provide a competitive baseline in the *pointwise* approach category. Compared to plain regression, this entry includes two rather standard enhancements:

- A regularizer; in other words, we perform ridge regression.
- A reweighting in such way that the relevant and non-relevant documents have the same total weight. In the context of SVM classifier, this trick has for instance been used by [Morik et al. '99] (see equation (4)).

## Learning Parameters

The details of the implementation can be found by looking at the Matlab source code.

In particular, the ridge is chosen to be a constant times the mean value of the diagonal elements of the covariance matrix; and this constant is selected on the validation set in the set

$$\{0.0001, 0.001, 0.01, 0.1, 1, 10\}.$$

| Dataset | Regularization constant (from Fold1 to Fold5) |
|---|---|
| OHSUMED | 1, 1, 0.1, 0.1, 0.1 |
| TD2003 | 0.01, 0.001, 0.1, 0.1, 0.001 |
| TD2004 | 0.0001, 0.01, 0.0001, 0.01, 0.01 |
| HP2003 | 0.1, 0.1, 0.1, 0.1, 1 |
| HP2004 | 1, 10, 0.01, 0.1, 1 |
| NP2003 | 1, 0.01, 0.1, 0.1, 0.01 |
| NP2004 | 1, 10, 10, 1, 10 |

## Notes

This document was written by Olivier Chapelle, and the experiments were conducted by Olivier Chapelle. If any problem, please contact letor@microsoft.com or chap@yahoo-inc.com