Supervised Learning 2004 — Assignment 4

Fernando Pérez-Cruz fernando@gatsby.ucl.ac.uk

Due Monday 6th December

Please hand in Q1 and Q2 and the results for Q3. Send by e-mail (subject Assignment 4) with the proposed code to solve Q3.

Q1 Mean and Median

Assume 8 people want to gather for a meeting; 5 of them live in Stuttgart and 3 in Munich. Where should they meet if:

a) they want the total distance travelled by all people to be minimal.

b) they want the average distance travelled per person to be minimal

c) they want the average squared distance to be minimal?

What happens to the meeting points if one of the 3 people moves from Munich to Sydney? (Question 3.3 from Schölkopf and Smola, 2002).

Q2 Cross-Validation

In this chapter, we have been concerned mainly with the case where the patterns are assigned to one of two classes, i.e., $y \in \{\pm 1\}$. Consider now the case where the assignment is not strict, i.e., $y \in [-1, 1]$. Modify the C-SVM and ν-SVM, such that

- whenever $y = 0$, the corresponding pattern has effectively no influence.
- if all labels are in $\{\pm 1\}$, the original algorithm is recovered.
- if $|y| < 1$, then the corresponding pattern has less influence than it would have for $|y| = 1$.

(Question 7.11 from Schölkopf and Smola, 2002).

Q3 Multiclass SVM

Propose a code that solves multi-class SVM with the techniques *one-vs-one* and *one-vs-all*. The code must allow to change the kernel and its parameter as well as the C value. As binary C-SVM, I would suggest the LibSVM (http://www.csie.ntu.edu.tw/∼cjlin/libsvm/), but any other can be used as well. Notice that LibSVM has many features, one of them in multi-class SVM. **DO NOT COPY IT, DO IT YOURSELF**.

Use your code to solve a ten dimensional digit recognition problem. The dataset for training and test will be available in the course webpage on Tuesday the 30th of December. Use a Gaussian Kernel:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$$

Report:

- The kernel width σ and the value of C, for the one-vs-all classifiers. (Use the same for all of them).
- The kernel width σ and the value of C, for the one-vs-one classifiers. (Use the same for all of them).
- Fill in the following table for both classifiers:

		True Label									
		0	1	2	3	4	5	6	7	8	9
P	0										
r	1										
e	2										
d	3										
	4										
L	5										
a	6										
b	7										
e	8										
l	9										