COMP9444 and COMP9844: Assignment 3

Due in week 10 (7 October 2013 23:59)

Achim Hoffmann, Anthony Knittel

Submission details

This assignment is to be submitted electronically as a document, and submitted using the "classrun" system. To submit, log in to your account on a CSE machine, and run

`9444 classrun -give svm myfile.pdf`

(replace `myfile.pdf` with the name of your document). The same command should be used for people enrolled in the extended course. Further information on using `classrun` is available at:

`http://www.cse.unsw.edu.au/help/doc/primer/node26.html`

The deadline for submission is October 7 23:59. Maximum file size is 3MB however smaller files are preferred. Late penalty is one mark off the maximum mark for each day late, further details are on the class website.

If you are familiar with writing equations using LaTeXor a word processor (or any other math typesetting program), you can use such an approach to write your solutions. An overview of how to use LaTeXis available at: `http://tobi.oetiker.ch/lshort/lshort.pdf`

An alternative approach is to write solutions by hand and submit them as an electronic document, for example by scanning. Please be careful of the file size if you use this approach, and ensure the submission is a single document in PDF format. Scanning at UNSW is available using many of the photocopying machines, such as in the main library.

Please ensure all relevant working is shown and that all writing is clear and easy to read.

Support Vector Machines

Determine the Optimal hyperplane given the following data by determining the Lagrange multipliers (α-values) for each support vector.

Negative training data ($y_i = -1$):

$$\vec{x_1}^T = \langle 1, 3 \rangle$$

Positive training data ($y_i = +1$):

$$\begin{aligned} \vec{x_2}^T &= \langle 2, 1 \rangle \\ \vec{x_3}^T &= \langle 3, 3 \rangle \end{aligned}$$

Note that the following condition has to be met:

$$\sum_{i=1}^{3} y_i \alpha_i = 0$$

This means that in your optimisation problem for n support vectors there are only $n-1$ independent variables to be determined using a system of $n-1$ equations.

Your submission should contain answers to the following points:

1. Provide the Lagrangian function which you need to optimize and show your calculations for finding each of the α-values.

2. Provide the weight vector as well as the bias for the Optimal hyperplane and how they are derived.

3. Draw the Optimal hyperplane along with the training data in the plane and mark all support vectors.

4. Independent of the above, given a Kernel function of $K(\vec{x_i}, \vec{x_j}) = (\vec{x_i}^T \vec{x_j} + 3)^2$, what is the implied mapping function $\Phi(\vec{x})$ into the feature space for the two-dimensional vector $\vec{x} = \langle 1, 3 \rangle^T$? Provide the result in the generic form using $x_i(1)$ and $x_i(2)$ as the first and second dimension components of $\vec{x_i}$, as well as using the two given values 1 and 3 respectively.

5. Using the following Kernel $K(\vec{x_i}, \vec{x_j}) = (\vec{x_i}^T \vec{x_j} + 1)^3$, provide the dual of the Lagrangian function which you need to optimize $(max_\alpha \mathcal{L}(\vec{\alpha}))$, and determine the Lagrange multipliers (α-values) for the following learning problem and show your calulations:

   Positive training data ($y_i = +1$):

   $$\begin{aligned} \vec{x_1}^T &= \langle -1, -1 \rangle \\ \vec{x_2}^T &= \langle 1, 1 \rangle \end{aligned}$$

   Negative training data ($y_i = -1$):

   $$\vec{x_3}^T = \langle 0, 0 \rangle$$

   Note that the training data is not linearly separable in the input space.

   Determine the class of the following data points and show your calculations: $\vec{x_4} = \langle 1, 0 \rangle$ and $\vec{x_5} = \langle 2, 2 \rangle$.

   Recommendation: Check your results by ensuring that your obtained classifier gets the correct results for the training data as well.