MACHINE LEARNING COMS 4771, HOMEWORK 2
Assigned September 30, 2014. Due October 14, 2013 before 1pm.

Please submit separate files for a) write-up, b) Matlab source files and c) figures (if you choose to include them separately from the writeup). Do not include any other files. Your write-up should be in ASCII plain text format (.txt) or Postscript (.ps) or the Adobe Portable Document Format (.pdf). Please do not submit Microsoft Office documents, LaTeX source code, or something more exotic since we will not be able to read it. LaTeX is preferred and highly recommended, but it is not mandatory. You can use any document editing software you wish, as long as the final product is in .ps or .pdf. Even if you do not use LaTeX to prepare your document, you can use LaTeX notation to mark up complicated mathematical expressions, for example, in comments in your Matlab code. See the Tutorials page for more information on LaTeX. All your code should be written in Matlab. Please submit all your source files, each function in a separate file. Clearly denote what each function does, which problem it belongs to, and what the inputs and the outputs are. Do not resubmit code or data provided to you. Do not submit code written by others. Identical submissions will be detected and both parties will get zero credit. In general, shorter code is better. Sample code is available on the Tutorials web page. Datasets are available from the Handouts web page. You may include figures directly in your write-up or include them separately as .jpg, .gif, .ps or .eps files, and refer to them by filename.

Submit your work via courseworks.columbia.edu.

1 Problem 1 (15 points)

VC Dimension

(a) Prove that $VC(H) \leq \log_2 |H|$ where H is a hypothesis space. (A hypothesis on a set of n points, defines which of two classes can each point belong to. A hypothesis space is a family of all possible hypotheses).

(b) Consider a perceptron in $\mathbb{R}^d$. How many points can it shatter or more specifically what is the VC dimension of this perceptron? Explain your answer.

(c) Find the VC-dimension of a class of circles in $\mathbb{R}^2$.

2 Problem 2 (10 points)

Kernels

(A) A stationary kernel is defined as translation invariant: $K(x, z) = K_S(x - z)$ that is, it depends only on the lag vector separating the two examples x and z, but not on the examples themselves. Consider the following stationary kernels

1) $\exp(-\frac{||x-z||}{\theta})$ (exponential kernel)

2) $\exp(-\frac{||x-z||^2}{\theta})$ (Gaussian kernel)

which are positive definite in R^d.

(i) Are they Mercer kernels according to the Bishop book's definition? (Explain yes or no).

(ii) Prove that sum and product of the two kernels is also a stationary kernel.

(B) For the kernel $K(x, y) = \exp(-\frac{1}{2} \parallel x - y \parallel^2) = \varphi(x) \cdot \varphi(y)$, write an explicit formula for φ.

3 Problem 3 (10 points)

$f(x;\lambda,k) = \begin{cases} \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}e^{-(x/\lambda)^k} & x \geq 0 \\ 0 & x < 0 \end{cases}$

Define $F(x;\lambda,k) = \int_0^x f(x;\lambda,k)dx$.

For the points $\{(0.25, 0.3, 0.4), (0.2, 0.4, 0.25), (0.3, 0.03, 0.2), (0.1, 0.2, 0.15), (0.03, 0.05, 0.7)\}$, using the Kernel function $K(x_i, x_j) = F(x_i \cdot x_j, 0.75, 3)$, compute the Gram Matrix.

4 Problem 4 (15 points)

You will build an SVM to classify data and use cross-validation to find the best SVM kernel and regularization value. Try different polynomials and RBF kernels (varying polynomial order from 1 to 5) and varying sigma in the RBF kernel. Also, try different values of C in the SVM. First, extract the support vector machine Matlab code from Steve Gunn's software package here:

```
http://www.isis.ecs.soton.ac.uk/resources/svminfo/
In svc.m replace [alpha lambda how] = qp(...);
with             [alpha lambda how] = quadprog(H,c,[],[],A,b,vlb,vub,x0);
```

Clearly denote the various components and the function calls or scripts that execute your Matlab functions. Note, to save the current figure in Matlab as a postscript file you can type:

```
print -depsc filename.eps
```

To test your SVM, you will build a simple object recognition system. Download the data file svm-dataset.mat from the Handouts web page. This loads a matrix X of size 100 x 4 and a vector Y of size 100 x 1. The X matrix consists of 100 data points of dimensionality 4 that belong to one of two classes. Their actual labels are given in the Y vector. Train your SVM on half of the examples and test on the other half (or other random subsets of the examples if you see fit). Show performance of the SVM for linear, polynomial and RBF kernels with different settings of the polynomial order, sigma and C value. Try to hunt for a good setting of these parameters to obtain high recognition accuracy.