Homework # 3

Foundations of Computer and Data Science CS-596

Problem 1: We would like to test the ability of the first kernel method (discussed in the class) to approximate pdfs. Therefore we generate 1000 realizations of a random variable uniformly distributed in $[0,1]$. Approximate the corresponding pdf using the Gaussian kernel

$$\kappa(x,h) = \frac{1}{\sqrt{2\pi h}} e^{-\frac{1}{2h}x^2}.$$

Plot the resulting approximations for different values of h. Do not limit your graph in $[0,1]$ but extend it beyond these two limits (right and left). What do you observe as far as capturing the support of the random variable (the interval $[0,1]$) and the constant value of the pdf (equal to 1 in $[0,1]$) is concerned? b) Complete the same steps for the case of the Laplacian kernel

$$\kappa(x,h) = \frac{1}{2h} e^{-\frac{1}{h}|x|}.$$

Problem 2: The Matlab data file data3-2.mat contains two matrices: `stars` and `circles` each being a list of vectors of length 2. Each length-two vector identifies a point in 2-D which is labeled either star or circle. We are interested in developing a classifier that distinguishes between the two sets. If you place the two sets on the plane verify that they cannot be separated with a straight line. We would therefore like to use the kernel method to find a nonlinear separating boundary. To achieve this we assign the label "1" to `stars` and the label "-1" to `circles` and if $\phi(X), X = [x_1, x_2]^T$ is the transformation we would like to apply to the data then we want to solve the following minimization problem to find the optimum $\phi(X)$

$$\min_{\phi} \left\{ \sum_{X_i \in \text{stars}} \big(1 - \phi(X_i)\big)^2 + \sum_{X_j \in \text{circles}} \big(1 + \phi(X_j)\big)^2 + \lambda \|\phi(X)\|^2 \right\}.$$

Explain what this minimization tries to achieve. The previous minimization if performed over functions $\phi(X)$ that belong to the space of functions generated by the Gaussian kernel

$$\kappa(X,Y) = e^{-\frac{1}{h}\|X-Y\|^2} = e^{-\frac{1}{h}\{(x_1-y_1)^2 + (x_2-y_2)^2\}}.$$

a) Use the main (last) theorem we mentioned in the class to find the optimum $\phi(X)$ by reducing it into a problem for identifying a set of parameters. b) Once you identify $\phi(X)$ explain how you are going to use it to classify a new point X_{new} as "star" or "circle" given, of course, that $\phi(X_{\text{new}})$ will not be exactly equal to 1 or -1. c) Once you have specified your classification rule in b) find (numerically) the separating boundary for the two classes in the 2-D space (also place your the training points to verify the quality of your boundary). d) Repeat the same process for different values of h and λ. e) Repeat all previous questions using the following (far) simpler kernel function

$$\kappa(X,Y) = (1 + x_1 y_1 + x_2 y_2)^2.$$

In addition to the questions a),b),c),d), also specify the type of functions contained in the Hilbert space generated by this kernel. *Hint: Remember that the Hilbert space contains functions of the form* $\phi(X) = \sum_{i=1}^{m} \kappa(X, X_i)$ *where* $X_1, \ldots, X_m$ *are vectors from the original space while* X *is a free vector. Try to specify more explicitly the form of the functions* $\phi(X)$.

We will meet on Friday, November 30, at 6PM, in Core-101 to discuss the problems.

Your answers, in *hard copy*, must be submitted on Monday, December 03, in CBIM room #5, between 5 and 6PM to Mr Neelesh Kumar.