# EE 657 : Quiz 2

**Duration: 1 hour** **Date: Apr 8' 2017** **Marks:12.5**

Please outline all the steps systematically in the supplementary sheet.Careful
transfer the required answers on the answer sheet

1. Consider a training set of five data points $\{(\mathbf{x}_i, t_i)\}_{i=1}^{5}$, where each input $\mathbf{x}_i$ is 2-dimensional and its corresponding target $t_i$ is a real valued scalar. We wish to learn a suitable regression function, that can perfectly pass through all these points. To this goal, consider adopting a Multi layer perceptron (MLP) neural network with single hidden layer.

   (a) Draw a schematic representation of the neural network for this regression problem. Clearly specify:

      i. the number of nodes in the input layer

      ii. the number of basis functions in the hidden layer

      iii. the number of nodes in the output layer.

   (b) What is the training error for this network. Justify your answer.

   (c) How many weights are to be learnt for this regression task.

   **[1.5 + 1 + 1.5 = 4 marks]**

2. You are given 2 training examples (comprising 2 dimensions) for two equally likely classes $\omega_1$ and $\omega_2$ .

$$\omega_1 : \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \qquad \omega_2 : \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

A two dimensional Gaussian probability density function (with spherical covariance matrix) is centred on each of the examples and the estimation is performed using the Parzen Window. Classify the vector $\mathbf{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ by stating the maximum posterior probability value. Assume the zero-one loss function case.

The Gaussian distribution in $d$-dimensions is given by

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{0.5}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}$$

$\phi(x - x_i)$

[3 marks]

3. You are given a data set $\{\mathbf{x}_i, y_i\}_{i=1}^N$ of size $N$. Each input $\mathbf{x}_i$ is $d$-dimensional and its corresponding target $y_i$ takes any real value. The formulation of the SVM for the regression problem can be written as :

$$\min_{\mathbf{w},b,\xi,\hat{\xi}} \quad \frac{1}{2}||\mathbf{w}||^2 + \frac{C}{2}\sum_{i=1}^{N}(\xi_i^2 + \hat{\xi}_i^2)$$

subject to the constraints for $i = 1, 2, 3.....N$

$$\mathbf{w}^T\phi(\mathbf{x}_i) + b - y_i \leq \epsilon + \xi_i$$
$$y_i - \mathbf{w}^T\phi(\mathbf{x}_i) - b \leq \epsilon + \hat{\xi}_i$$
$$\xi_i \geq 0, \hat{\xi}_i \geq 0$$

(a) Write down the Lagrangian Function $L(\mathbf{w}, b, \boldsymbol{\xi}, \hat{\boldsymbol{\xi}})$ corresponding to this optimization problem. Here $\boldsymbol{\xi} = \{\xi_1, \xi_2, ....\xi_N\}$ , and $\hat{\boldsymbol{\xi}} = \{\hat{\xi}_1, \hat{\xi}_2, ....\hat{\xi}_N\}$.

(b) State the KKT conditions, that are to be satisfied at optimality, together with expressions for weights, biases, violations , complementary slackness conditions and other constraints (if necessary).

[1+4.5 = 5.5 marks]