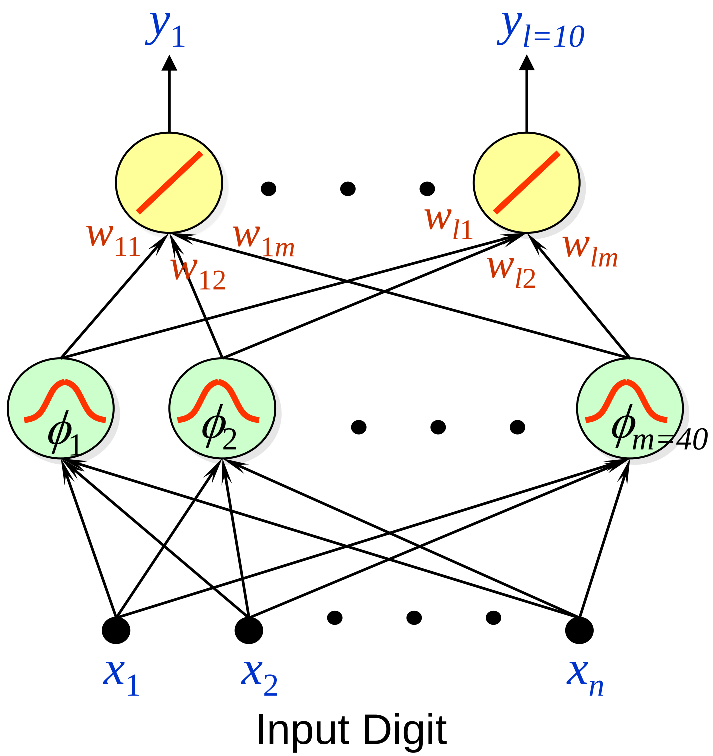
HW# 12
Application of Neural Network (CpE 520)
Due date: Dec. 7th , 2021

Read pp. 239-250 on RBF, Simon Haykin "Neural Networks and Learning Machines" 3rd addition, Prentice Hall

Problem 1:

1. Implement an RBF network (see the diagram) with m=40 centroids with ten binary outputs (number of classes). To get the centroids of the Gaussian function use the results from HW# 4 (unsupervised K-means algorithm on MNIST dataset with 40 centroids).
2. Find the weights in the second layer using the Pseudo-inverse method.
3. Initialize the centroids of the Gaussian functions using the K-means centroids and then fine tune the centroids + the weights of the second layer using the gradient decent method.
4. Plot the learning curve for the gradient decent method.
5. For each approach report your classification performance in a confusion matrix form of 10x10.



http://yann.lecun.com/exdb/mnist/

Problem 2:

Consider a Radial Basis Function network (RBF) where we want to find its network parameters (centers, spread and last layer weights).

1. Apply the stochastic gradient descent method for finding centers, spread and weights, by minimizing the (instantaneous) squared error cost function:

$$E(\mathbf{x}^{(k)}) = \frac{1}{2}\{y^{(k)} - \sum_{i=1}^{m} w_i \exp[-\left\|\mathbf{x}^{(k)} - \boldsymbol{\mu}_i\right\|^2 / 2\sigma_i^{\,2}]\}^2$$

The network parameters are updated as follows:

$$\boldsymbol{\mu}_i^{t+1} = \boldsymbol{\mu}_i^{t} - \eta_{\mu_i} \frac{\partial E}{\partial \boldsymbol{\mu}_i^{t}}$$

$$\sigma_i^{t+1} = \sigma_i^{t} - \eta_{\sigma_i} \frac{\partial E}{\partial \sigma_i^{t}}$$

$$w_i^{t+1} = w_i^{t} - \eta_{w_i} \frac{\partial E}{\partial w_i^{t}}$$

Find the following gradients of the cost function $E(\mathbf{x}^{(k)})$ with respect to each network parameter, show your derivation in detail do not copy the answers from my notes:

$$\frac{\partial E}{\partial \boldsymbol{\mu}_i^{t}} = ?, \quad \frac{\partial E}{\partial \sigma_i^{t}} = ?, \quad \frac{\partial E}{\partial w_i^{t}} = ?$$