1) Read chapter 4 from Haykin's book (2nd edition), starting from section 4.1 to 4.6(including 4.6). Summarize or sketch your insights in mind-map or an outline or a summary.

2) For this task you have to program 2-2-1 multi layered perceptron (MLP) with sigmoid activation function on XOR data.
a. Experiments with initial weights
i. Train the network with zero initial weights i.e. $w_{ij} = 0$.
ii. Train with random initial weights
Compare and comment on the convergence.
b. Experiment with different learning rates e.g. 0.1, 0.3, 0.5, 0.9..

Compare the convergence and plot some resulting surfaces.
You are not allowed to use any neural network toolbox for this solution.

3) [Haykin(2nd edition) 4.16]

Investigate the use of back-propagation learning using a sigmoidal nonlinearity to achieve one-to-one mappings, as described here:
1. F(x) = 1/x 1<=x<=100
2. F(x) = log10(x) 1<=x<=10
3. F(x) = exp(-x) 1<=x<=10
4. F(x) = sin(x) 0<=x<=pi/2

For each mapping, do the following:
(a) Set up two sets of data, one for network training, and the other for testing.
(b) Use the training data set to compute the synaptic weights of the network, assumed to have a single hidden layer.
(c) Evaluate the computation accuracy of the network by using the test data. Use a single hidden layer but with a variable number of hidden neurons. Investigate how the network performance is affected by varying the size of the hidden layer.

You can use any neural network toolbox (MATLAB or python or …) for solving this problem or you can use your own implementation of MLP from previous question.