clc; clear all; close all tic

x = [0 0;0 1;1 0;1 1]; % feature vector

t = [0;1;1;0]; % desired output value

[a, b] = size(x); % acquire number of training patterns and feature vectors

numOut = size(t,2); % output layer's unit size lr = 0.3; % learning rate

max\_iteration = 100000; % Maximum iteration if the program does not break mse = 1; % MSE initialization

w1 = 2 \* rand(2,3)- 1; % input-hidden layer weight range -1 to 1 w2 = 2 \* rand(1,3) - 1; % Hidden-output neuron weight range -1 to 1 x = [ones(a,1), x]; % add bias 1

for epoch = 1 : max\_iteration

for i = 1:a

% Forworad phase

% calculate hidden layer

z(i,:) = 1 ./ (1 + exp(-w1 \* x(i,:)')); % Binary Sigmoid

% cauculate output layer

o(i,:) = 1 ./ (1 + exp(-w2 \* [1; z(i,:)'])); % Binary Sigmoid

%Backward phase

% calculate gragient of the output layer

del\_2(i,:) = (t(i,:) - o(i,:)) .\* o(i,:) .\* (1 - o(i,:));

% calculate gragient of the hidden layer

del\_1(i,:) = z(i,:) .\* (1 - z(i,:)) .\* (del\_2(i,:) \* w2(:,2:end)); w2\_new(:,:,i) = lr \* del\_2(i,:)' \* [1, z(i,:)];

if i>=2

w2\_new(:,:,i)=w2\_new(:,:,i)+0.1\*w2\_new(:,:,i-1); % Momentum(0.1)term end

w1\_new(:,:,i) = lr \* del\_1(i,:)' \* x(i,:); if i>=2

w1\_new(:,:,i)=w1\_new(:,:,i)+0.1\*w1\_new(:,:,i-1); % Momentum(0.1)Term end

end

% sum of training pattern w2\_new = sum(w2\_new,3); w1\_new = sum(w1\_new,3);

%update weights

w2 = w2 + w2\_new; % update w2 w1 = w1 + w1\_new; % update w2

% mean square error calculation mse(1,epoch) = sum((o-t).^2)/(a\*numOut);

%stopping criteria

if mse(1,epoch) <=0.001 break

end

end

%Plot the learning curve figure

plot(1:epoch, mse, 'b'); grid on

title(['Learning curve at learning rate =',num2str(lr)]) xlabel('epoch');

ylabel('mse')

accuracy=(100-(sum(abs(o-t))/4)\*100) toc