**EX-3**

1. **IrcostFunction**

hyp = sigmoid(X\*theta);

J = (1/m)\*sum(-y.\*log(hyp) - (1-y).\* log(1-hyp)) + (lambda/(2\*m))\*sum(theta(2:end).^2);

grad(1) = (1/m)\*(X(:,1))'\*(hyp-y);

grad(2:end) = (1/m)\*(X(:,2:end))'\*(hyp-y) + (lambda/m)\*theta(2:end);

1. **oneVSAll**

initial\_theta = zeros(n+1,1);

options = optimset('GradObj', 'on', 'MaxIter', 50);

for i=1:num\_labels

all\_theta(i,:) = fmincg(@(t)(lrCostFunction(t,X,(y==i),lambda)),initial\_theta,options);

end

1. **predict**

a1=[ones(size(X,1),1) X];

z2=a1\*Theta1';

a2=sigmoid(z2);

a2=[ones(size(a2,1),1) a2];

z3=a2\*Theta2';

a3=sigmoid(z3);

[probability, p] = max(z3,[],2);

1. **predictoneVsAll**

probabilites = X\*all\_theta';

[probability,p]=max(probabilites,[],2);