

Arian Jalalian - 99243026

1. Sigmoid works for binary classification but softmax is multiclass
2. If we use MSE, we get a non convex function in classification tasks resulting in gradient descent not working well and getting stuck in local minimums
3. It means that for the given email is 75 percent spam and 25 percent not spam but we classify it as spam
4. $x_1 + x_2 - 3 = 0$ is the decision boundary

