1 Question 1

given = the historical data of results of soccer matches playing against ajax
goal = predict whether a team will win/ draw/ lose against ajax at any given moment
learning task = supervised learning, classification

2 Question 2

a)

Given data:

x	y
3	6
5	7
6	10

Initial values:

$$m = 3$$
$$\alpha = 0.1$$
$$\theta_0 = 0$$
$$\theta_1 = 1$$
$$h_\theta(x) = \theta_0 + \theta_1 x^{(i)}$$

gradient descent algorithm:
repeat until convergence

$$\theta_0 := \theta_0 - \alpha * 1/m \sum_{i=1}^{m} (h_\theta(x)^{(i)} - y^{(i)})$$
$$\theta_1 := \theta_1 - \alpha * 1/m \sum_{i=1}^{m} (h_\theta(x)^{(i)} - y^{(i)}) * x^{(i)}$$

first iteration $theta_0$

$$\theta_0 := 0 - 0.11/3 \sum_{i=1}^{3} (0 + (1 * 3) - 6) + (0 + (1 * 5) - 7) + (0 + (1 * 6) - 10)$$
$$\theta_0 := -1/30 * -9$$
$$\theta_0 := 0.3$$

first iteration $theta_1$

$$\theta_1 := 1 - 0.11/3 \sum_{i=1}^{3} (0 + (1 * 3) - 6)(3) + (0 + (1 * 5) - 7)(5) + (0 + (1 * 6) - 10)(6)$$
$$\theta_1 := 2.43$$

new values are:

$$\theta_0 := 0.3$$
$$\theta_1 := 2.43$$

Updates:

$$\theta_0 := 0.3 - 0.1 * 1/3((0.3 + (2.43 * 3) - 6) + (0.3 + (2.43 * 5) - 7) + (0.3 + (2.43 * 6) - 10))$$
$$\theta_0 := 0.3 - 0.1 * 1.3(1.59 + 5.45 + 4.88)$$
$$\theta_0 := -0.10$$
$$\theta_1 := 2.43 - 0.1 * 1/3 \sum_{i=1}^{3} (0.3 + (2.43 * 3) - 6)(3) + (0.3 + (2.43 * 5) - 7)(5) + (0.3 + (2.43 * 6) -$$
$$\theta_1 := 2.43 - 0.1 * 1/3 \sum_{i=1}^{3} (4.77 + 27.25 + 29.28)$$
$$\theta_1 := 0.39$$

new values after second iteration

$$\theta_0 := -0.10$$
$$\theta_1 := 0.39$$

b)

I am assuming that the question means that we should calculate z-scores to fit a sigma of 1 and a mean of 0.

$$\mu(x) = 4.67$$
$$\sigma(x) = 1.25$$
$$z - score = (X - \mu)/\sigma$$

x	y
-1.576	-0.566
0.024	-0.227
0.824	0.79

$$m = 3$$
$$\alpha = 0.1$$
$$\theta_0 = 0$$
$$\theta_1 = 1$$
$$h_\theta(x) = \theta_0 + \theta_1 x^{(i)}$$

gradient descent algorithm:
repeat until convergence

$$\theta_0 := \theta_0 - \alpha * 1/m \sum_{i=1}^{m} (h_\theta(x)^{(i)} - y^{(i)})$$
$$\theta_1 := \theta_1 - \alpha * 1/m \sum_{i=1}^{m} (h_\theta(x)^{(i)} - y^{(i)}) * x^{(i)}$$

first iteration z-values $theta_0$

$$\theta_0 := 0 - 0.11/3 \sum_{i=1}^{3} (0 + (1 * 1.576) + 0.566) + (0 + (1 * 0.024) + 0.227) + (0 + (1 * 0.824) - 0.79)$$
$$\theta_0 := -1/30 * 2.417$$
$$\theta_0 := -0.081$$

first iteration z-values $theta_1$

$$\theta_1 := 1 - 0.11/3 \sum_{i=1}^{3} (0 + (1 * 1.576) + 0.566)(1.576) + (0 + (1 * 0.024) + 0.227)(0.024) + (0 + (1 *$$
$$\theta_1 := -1/30 * 0.085$$
$$\theta_1 := -0.003$$

new values for $theta_0$ are -0.081 and for $theta_1$ is -0.003

second iteration z-values $theta_0$

$$\theta_0 := -0.081 - 0.11/3 \sum_{i=1}^{3} (-0.081 + (-0.003 * 1.576) + 0.566) + (-0.081 + (-0.003 * 0.024) + 0.$$

$\theta_0 := -1/30 * -0.247$

$\theta_0 := 0.0082$

second iteration z-values $theta_1$

$$\theta_1 := 0.003 - 0.11/3 \sum_{i=1}^{3} (-0.081 + (-0.003 * 1.576) + 0.566) * 1.576 + (-0.081 + (-0.003 * 0.024$$

$\theta_1 := -1/30 * (-0.089 - 0.08 - 0.734)$

$\theta_1 := 0.0301$

the value for $theta_0$ is now 0.0082 , and the value for $theta_1$ is now 0.0301.

Otherwise, if we assume that the question gives us our mu and $sigma$, then by using the z-score ((X - mu)/$sigma$)we get the same values for x, and $theta_0$ and $theta_1$.

3 question 4

we set the derivative of the cost function to 0. We already know the derivative of this because it is part of gradient descent, and, as we are using $theta_1$ we take that derivative.

$$1/m \sum_{(i=1)}^{m} (h_\theta(x)^{(i)} - y)(x)^{(i)} = 0 \tag{1}$$

$$\sum_{(i=1)}^{m} (\theta_0 + \theta_1(x)^{(i)} - y^{(i)})(x)^{(i)} = 0 \tag{2}$$

$$\sum_{(i=1)}^{m} (\theta_0 * (x)^{(i)} + \theta_1((x)^{(i)})^2 - y^{(i)} * (x)^{(i)}) = 0 \tag{3}$$

$$\tag{4}$$

then we use summation laws to break the sum into smaller parts, and take out $theta_1$

$$\sum_{(i=1)}^{m} \theta_1((x)^{(i)})^2 = -\sum_{(i=1)}^{m} (\theta_0 * (x)^{(i)} - y^{(i)} * (x)^{(i)}) \tag{5}$$

$$\theta_1 = -\sum_{(i=1)}^{m} (\theta_0 * (x)^{(i)} - y^{(i)} * (x)^{(i)}) / \sum_{(i=1)}^{m} ((x^{(I)})^2 \tag{6}$$