⇒ EXAMPLES

PROBLEM 1

Example 1: The dataset $\tilde{\imath} = \{(x^{(i)}, y^{(i)})\}_{i=1}^{M}$, $x^{(i)} \in \mathbb{R}^3$, $y^{(i)} \in \mathbb{R}$ is a regression dataset because the target $y^{(i)}$ takes values from the real set $\mathbb{R}$. Because there are three features in the the following linear regression model would fit to it.

$$\hat{y}^{(i)} = w_0 + w_1 x_1^{(i)} + w_2 x_2^{(i)} + w_3 x_3^{(i)}.$$

That is one weight for each feature plus one additional weight (i.e., w_0) combined with the features as above.

Example 2 The data $\tilde{\imath} = \{(x^{(i)}, y^{(i)})\}_{i=1}^{N}$, $x^{(i)} \in \mathbb{R}$, $y^{(i)} \in \{0,1\}$ is classification data because the target $y^{(i)}$ assumes discrete values from the set $\{0,1\}$. It is binary classification data.

PROBLEM 2:

We will use the data in example 1 above. The loss function is,

$$E(x^{(i)}, y^{(i)}) = \frac{1}{2} \sum_{i=1}^{m} (y^{(i)} - \hat{y}^{(i)})^2.$$

we have to find $\nabla E = \frac{dE}{dw}$, the gradient of E wrt $W = [w_0\ w_1\ w_2\ w_3]^T$.