

Terminologies:

m : number of samples

n : number of feature

$x^{(i)}$: the i th feature vector

$y^{(i)}$: the i th label value

$\hat{y}^{(i)}$: the i th label value (predicted by model)

θ : A model parameter

