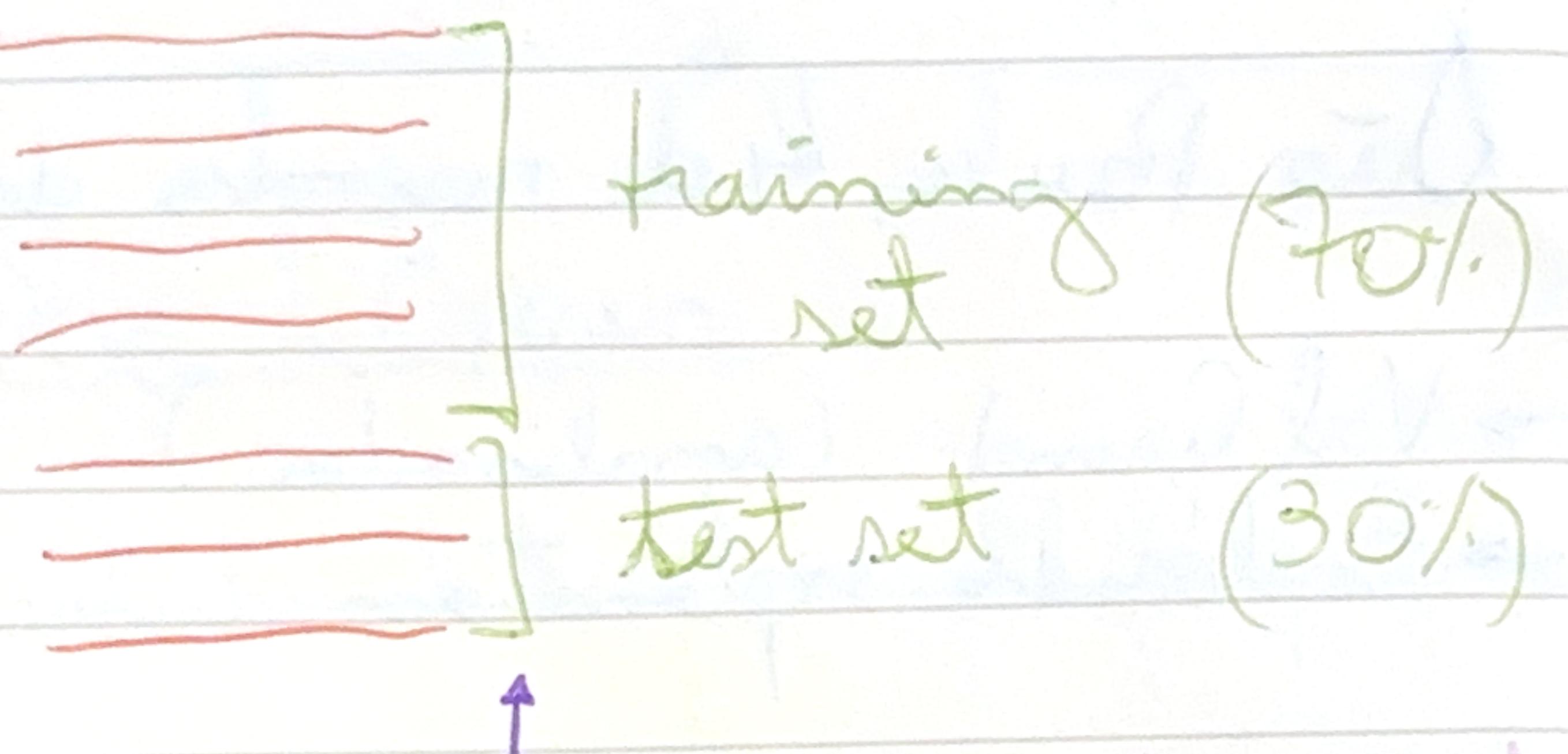


"Diagnostics can take time to implement, but doing so can be a very good use of your time" Andrew Ng

Dataset (100%)



81 of ab randomly selected

Uma outra forma para calcular o erro cometido na regressão logística é a seguinte:

$$\text{err}(h_0(x), y) = \begin{cases} 1, & \text{se } h_0(x) < 0.5 \text{ e } y = 1 \text{ ou} \\ & \text{se } h_0(x) \geq 0.5 \text{ e } y = 0 \\ 0, & \text{caso contrário} \end{cases}$$

$$\text{Test error} = \frac{1}{m_{\text{test}}} \sum_{i=1}^{m_{\text{test}}} \text{err}(h_0(x_{\text{test}}^{(i)}), y_{\text{test}}^{(i)})$$

Isto resultará na fração de exemplos-testes que foram classificados incorretamente.

No caso de regressões lineares, usar a média quadrática dos erros no training set continua uma boa

Model selection

$$d=1 \quad h_0(x) = \theta_0 + \theta_1 x \quad \theta^{(1)}$$

$$d=2 \quad h_0(x) = \theta_0 + \theta_1 x + \theta_2 x^2 \quad \theta^{(2)}$$

$$d=3 \quad h_0(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 \quad \theta^{(3)}$$

$$d=10 \quad h_0(x) = \theta_0 + \dots + \theta_{10} x^{10} \quad \theta^{(10)}$$

(1) → Calcular $J_{\text{train}}(\theta^{(i)})$ para $i=1:100$

DATASET 100% dividido em 60% train

