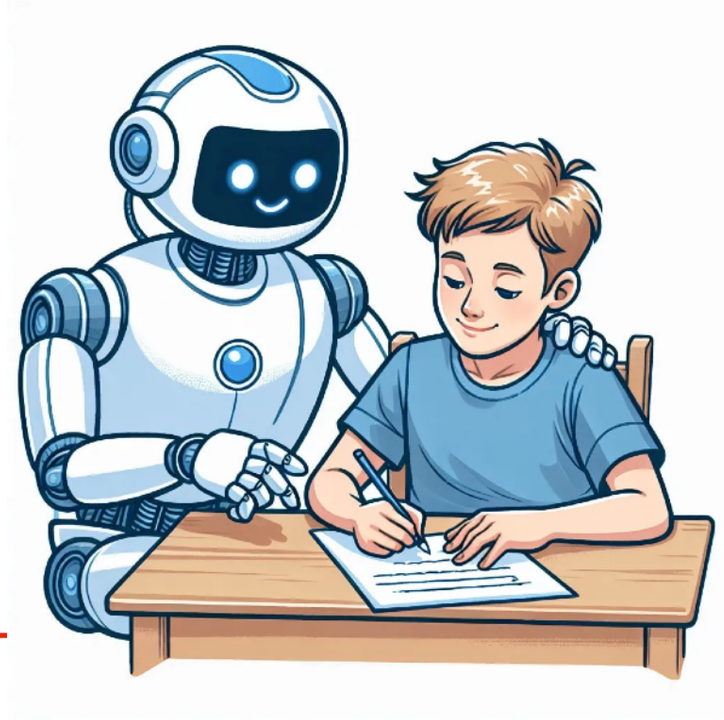
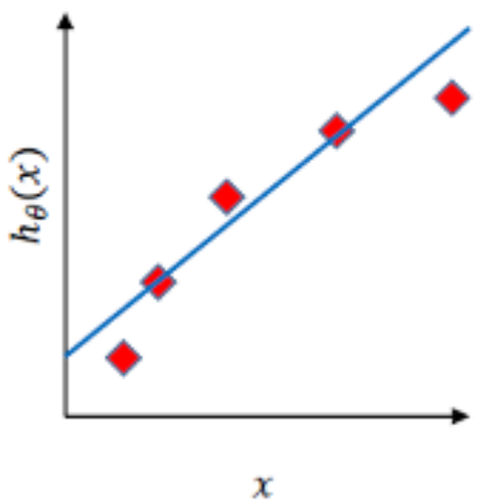
Content

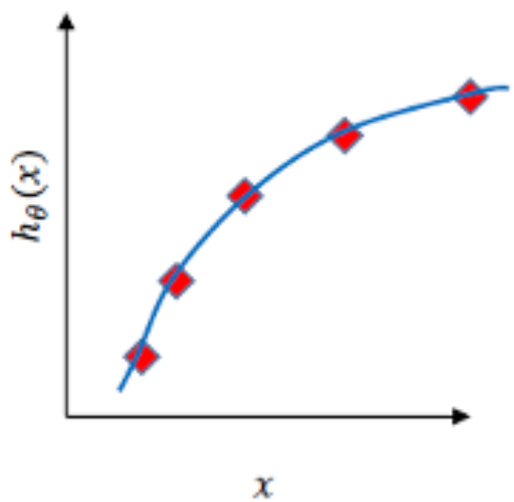


The quality of fitting

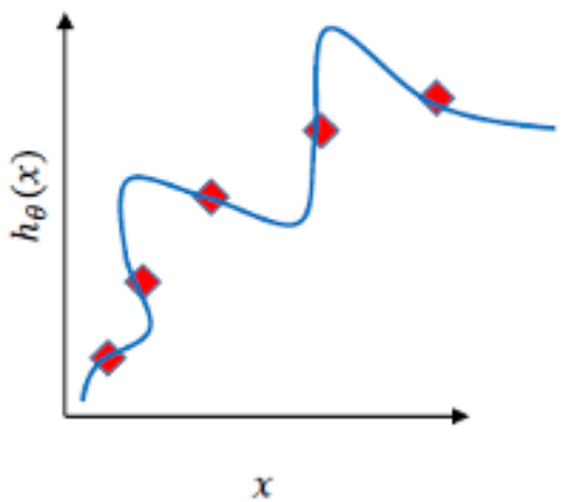
Underfitting (high bias)



High bias



Overfitting (high variance)

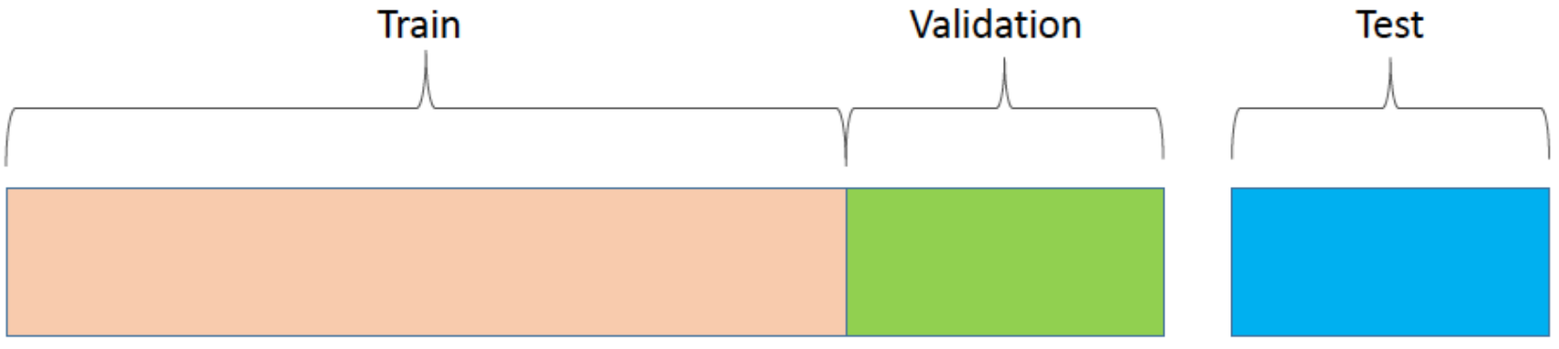


High variance

Polynomial regression

Forecastability

$$h_\theta^{[d]}(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \cdots + \theta_d x^d = \theta^T x$$

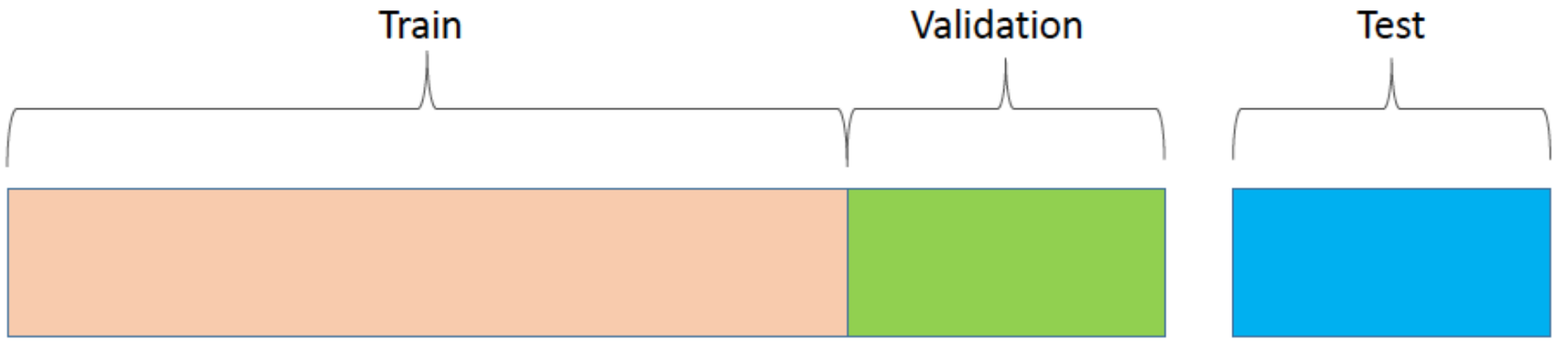


$$J_{\text{train}}(\theta) = \frac{1}{2M_{\text{train}}} \sum_{i_{\text{train}}=1}^{M_{\text{train}}} \left(y^{(i_{\text{train}})} - h_\theta(x)\right)^2$$

$$J_{\text{cv}}(\theta) = \frac{1}{2M_{\text{cv}}} \sum_{i_{\text{cv}}=1}^{M_{\text{cv}}} \left(y^{(i_{\text{cv}})} - h_\theta(x)\right)^2$$

$$J_{\text{test}}(\theta) = \frac{1}{2M_{\text{test}}} \sum_{i_{\text{test}}=1}^{M_{\text{test}}} \left(y^{(i_{\text{test}})} - h_\theta(x)\right)^2$$

Learning Curves and Regularization

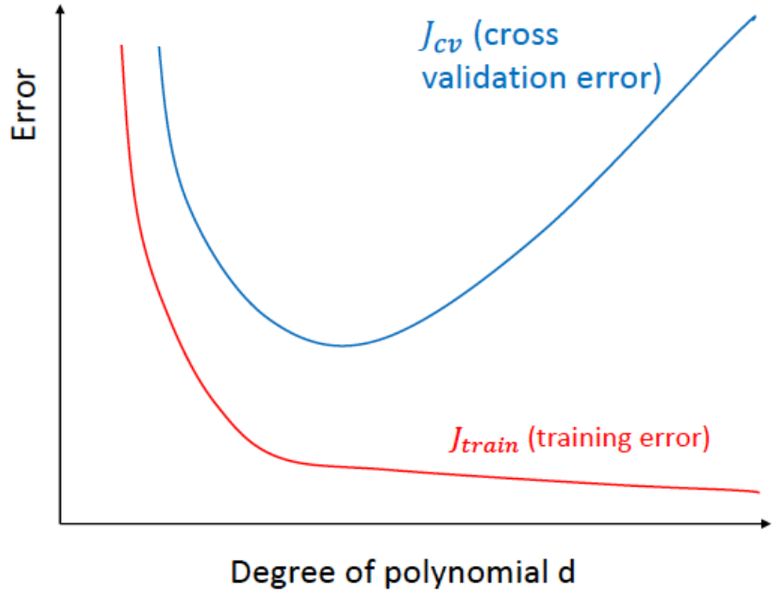


Forecastability

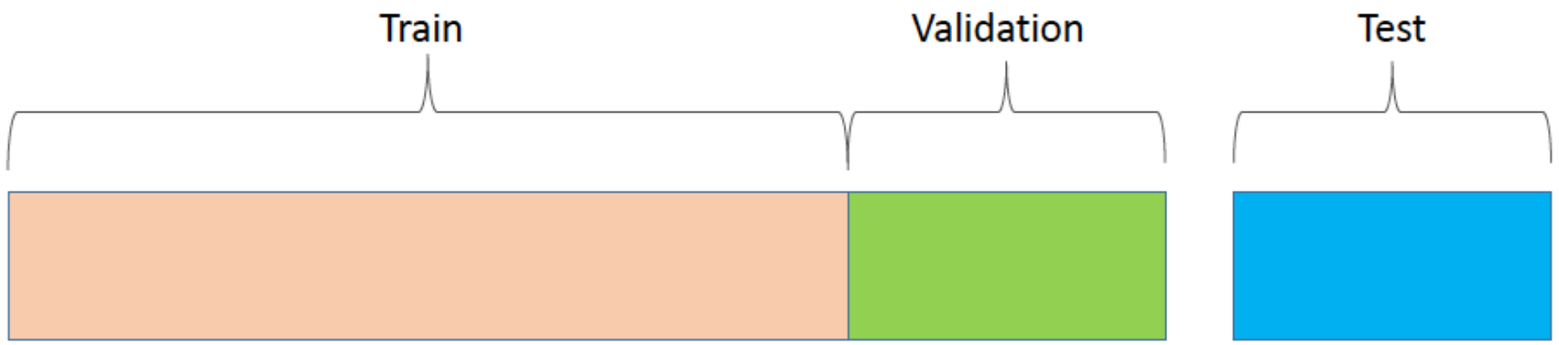
$$h_\theta^{[d]}(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \cdots + \theta_d x^d = \theta^T x$$

Polynomial regression

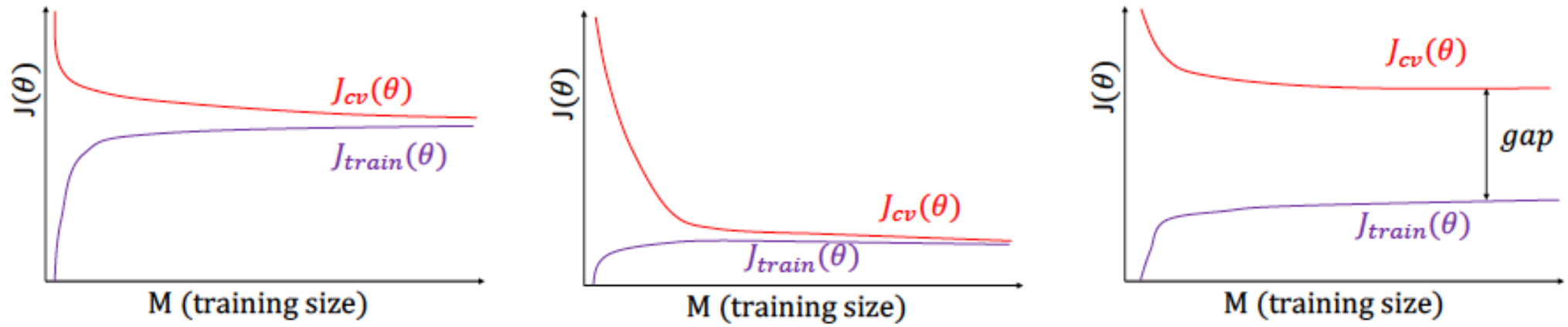
Model Selection



Learning Curves and Regularization



Learning curves



Regularization

$$J(\theta) = \frac{1}{2M}\left[\sum_{i=1}^{M}(h_\theta(x^{(i)}) - y^{(i)})^2 + \lambda\sum_{j=1}^{N}\theta_j^2\right] \qquad \lambda \geq 0$$