



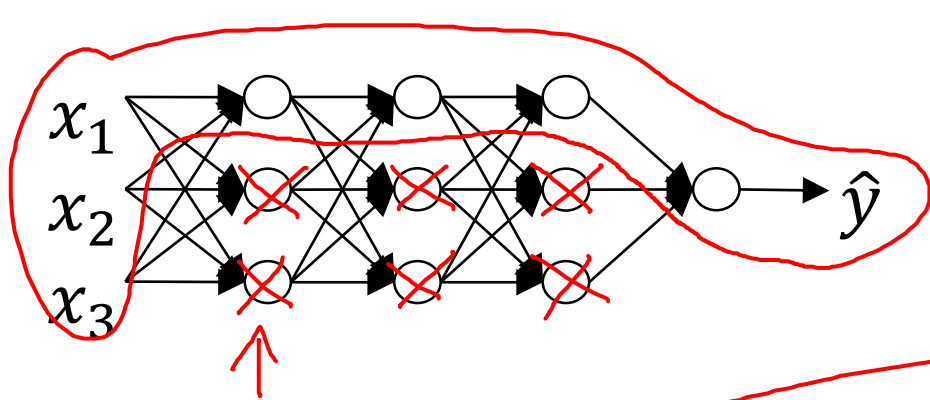
deeplearning.ai

# Regularizing your neural network

---

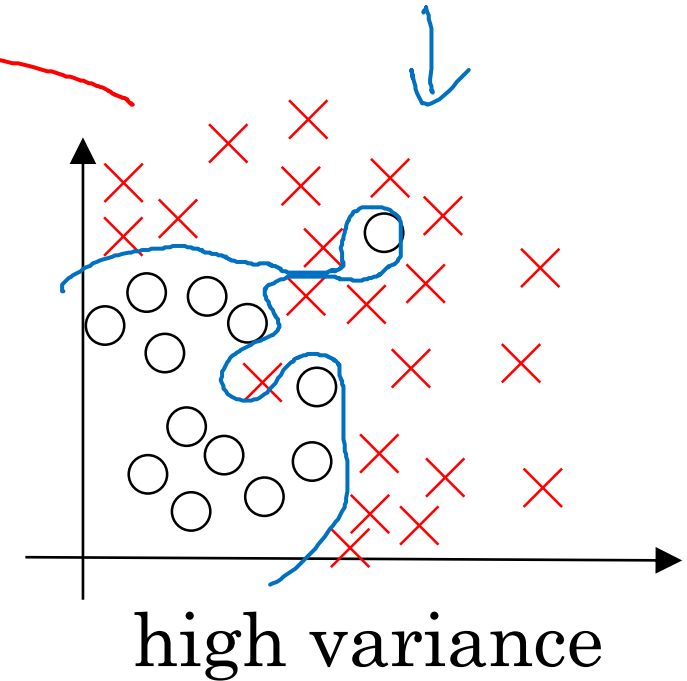
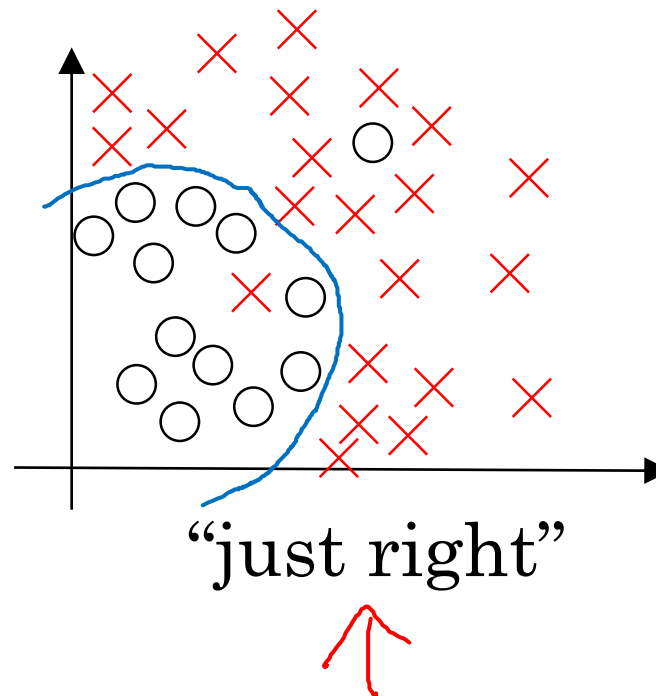
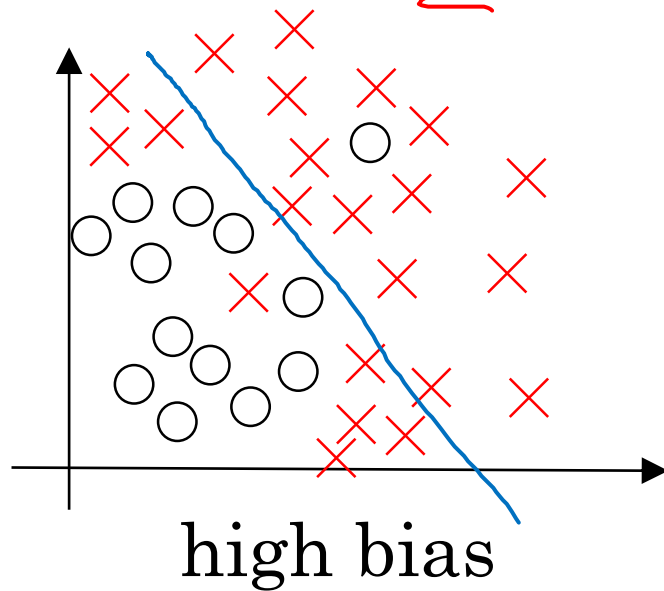
## Why regularization reduces overfitting

# How does regularization prevent overfitting?

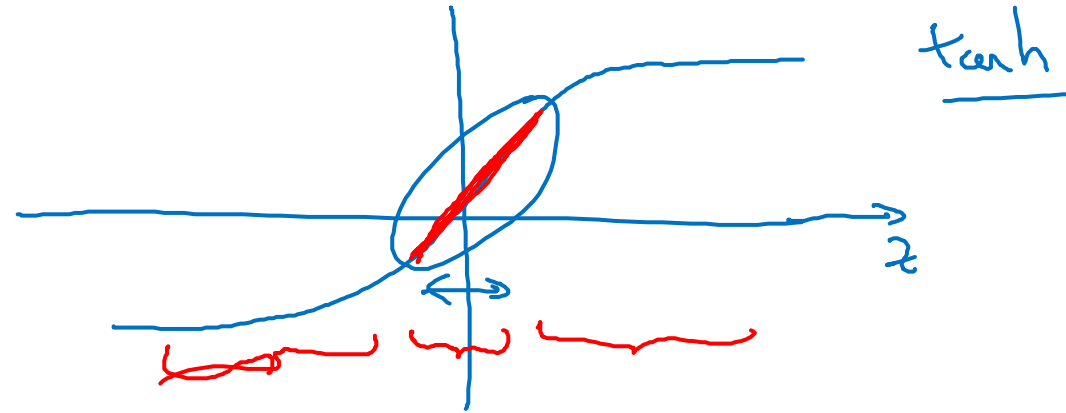


$$J(w^{(L)}, b^{(L)}) = \frac{1}{n} \sum_{i=1}^n \ell(y^{(i)}, \hat{y}^{(i)}) + \frac{\lambda}{2n} \sum_{l=1}^L \|w^{(l)}\|_F^2$$

$$w^{(L)} \approx 0$$



# How does regularization prevent overfitting?



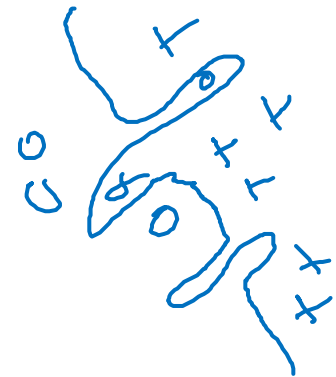
$$g(z) = \tanh(z)$$

$$\lambda \uparrow$$

$$W^{[L]} \downarrow$$

$$z^{[L]} = W^{[L]} a^{[L-1]} + b^{[L]}$$

Every layer  $\approx$  linear.



$$J(\dots) = \underbrace{\sum_i \mathcal{L}(\hat{y}^{(i)}, y^{(i)})}_{\text{training loss}} + \underbrace{\frac{\lambda}{2m} \sum_l \|W^{[l]}\|_F^2}_{\text{regularization term}}$$

