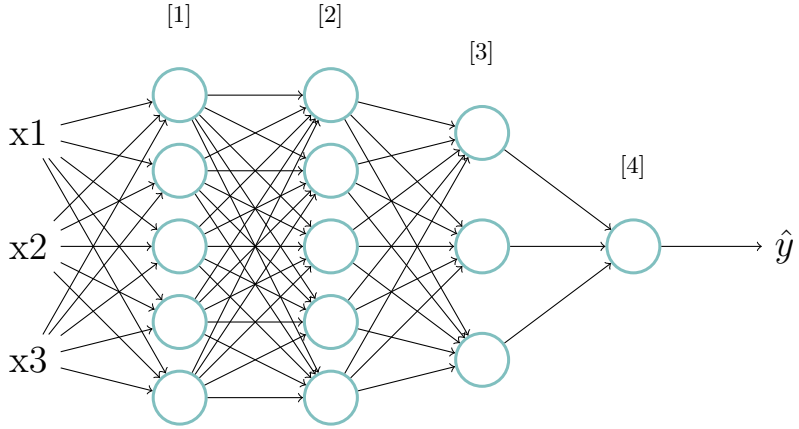
1 Week1

Improving Deep Neural Network

1.1 Setting up your ML application

Train/dev/test sets



$$Z^{[1]} = W^{[1]}A^{[0]} + b^{[1]} \quad _{where\ A^{[0]}=X}$$

$$A^{[1]} = g^{[1]}(Z^{[1]})$$

$$Z^{[2]} = W^{[2]}A^{[1]} + b^{[2]}$$

$$A^{[2]} = g^{[2]}(Z^{[2]})$$

$$\dots$$

$$Z^{[4]} = W^{[4]}A^{[3]} + b^{[4]}$$

$$A^{[4]} = g^{[4]}(Z^{[4]}) = \hat{Y}$$

Using for loops on $[l]$ layer for $l = 1, 2, \dots, L$:

$$Z^{[l]} = W^{[l]}A^{[l-1]} + b^{[l]}$$

$$A^{[l]} = g^{[l]}(Z^{[l]})$$

1.2 Forward and Backward Propagation

Forward propagation for layer l: $a^{[l-1]} \rightarrow a^{[l]}, z^{[l]}, w^{[l]}, b^{[l]}$

$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}$

$A^{[l]} = g^{[l]}(Z^{[l]})$

(for $i = 1, \ldots, L$ with initial value $A^{[0]} = X$)

Backward propagation for layer l: $da^{[l]} \rightarrow da^{[l-1]}, dW^{[l]}, db^{[l]}$

$dZ^{[l]} = dA^{[l]} * g^{[l]'}(Z^{[l]})$

$dW^{[l]} = \frac{1}{m} dZ^{[l]} A^{[l-1]^T}$

$db^{[l]} = \frac{1}{m} np.sum(dZ^{[l]}, axis = 1, keepdims = True)$

$dA^{[l-1]} = W^{[l]^T} dZ^{[l]} = \frac{dJ}{dA^{[l-1]}} = \frac{dZ^{[l]}}{dA^{[l-1]}} \frac{dJ}{dZ^{[l]}} = \frac{dZ^{[l]}}{dA^{[l-1]}} dZ^{[l]}$

(with initial value $dZ^{[L]} = A^{[L]} - Y$)