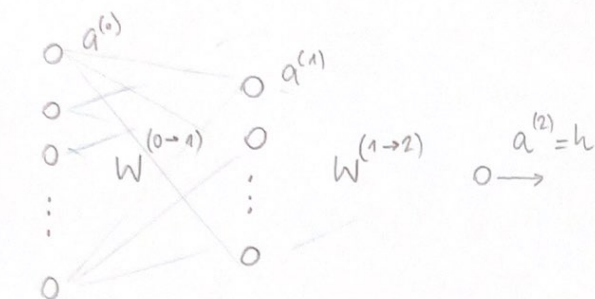


SENTIMENT ANALYSIS NETWORK | UDACITY

layer 0 INPUT layer 1 HIDDEN layer 2 OUTPUT (No bias)



74074 units
word
counts

10 units
No activation
function == no
nonlinearity

1 sigmoid

FORWARD

$$\underline{a}^{(0)} = \underline{x} \quad (74074, 1)$$

new units old units

$$\underline{z}^{(1)} = \underline{W}^{(0)} \cdot \underline{a}^{(0)} \quad \underbrace{(10, 74074)}_{\underline{W}^{(0)}} \times \underbrace{(74074, 1)}_{\underline{a}^{(0)}} = (10, 1)$$

$$\underline{a}^{(1)} = f(\underline{z}^{(1)}) = \underline{z}^{(1)} ; f: x \rightarrow x, f(x) = x : \text{no activation function}$$

$$\underline{z}^{(2)} = \underline{W}^{(1)} \cdot \underline{a}^{(1)} : \underbrace{(1, 10)}_{\underline{W}^{(1)}} \times \underbrace{(10, 1)}_{\underline{a}^{(1)}} = (1, 1)$$

$$\underline{a}^{(2)} = \sigma(\underline{z}^{(2)}) : \text{sigmoid activation}$$

$$h = \underline{a}^{(2)}$$

BACKWARD

$$\underline{e}^{(2)} = \underline{y} - \hat{\underline{y}} \quad (1, 1)$$

$$\underline{\delta}^{(2)} = \underline{e}^{(2)} \cdot \sigma'(\underline{z}^{(2)}) ; (1, 1) \quad \sigma' = \sigma(1-\sigma) \text{ sigmoid derivative}$$

$$\underline{e}^{(1)} = (\underline{W}^{(1)})^T \cdot \underline{\delta}^{(2)} : \underbrace{(10, 1)}_{(\underline{W}^{(1)})^T} \times \underbrace{(1, 1)}_{\underline{\delta}^{(2)}} = (10, 1)$$

$$\underline{\delta}^{(1)} = \underline{e}^{(1)} \cdot f'(\underline{z}^{(1)}) \quad (10, 1)$$

1 (no activation function)

one epoch loop =
for each
example m:

$$\underline{\Delta W}^{(l)} = \underline{\Delta W}^{(l)} + \underline{\delta}^{(l+1)} \cdot \underline{a}^{(l)T}$$

$\left\{ \begin{array}{l} \underline{W}^{(0)} : (10, 1) \times (1, 74074) \\ \underline{W}^{(1)} : (1, 1) \times (10, 1) \end{array} \right.$

m loop or after
epoch

$$\underline{W}^{(l)} = \underline{W}^{(l)} + \frac{\alpha}{m} \cdot \underline{\Delta W}^{(l)}$$