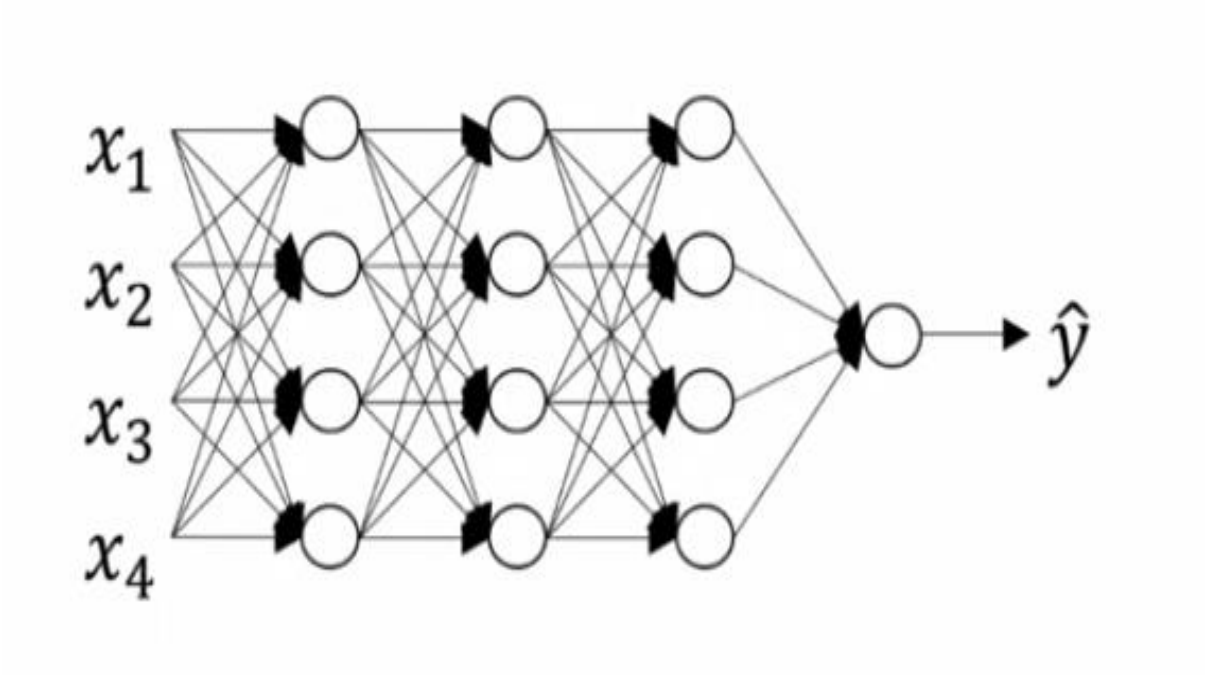
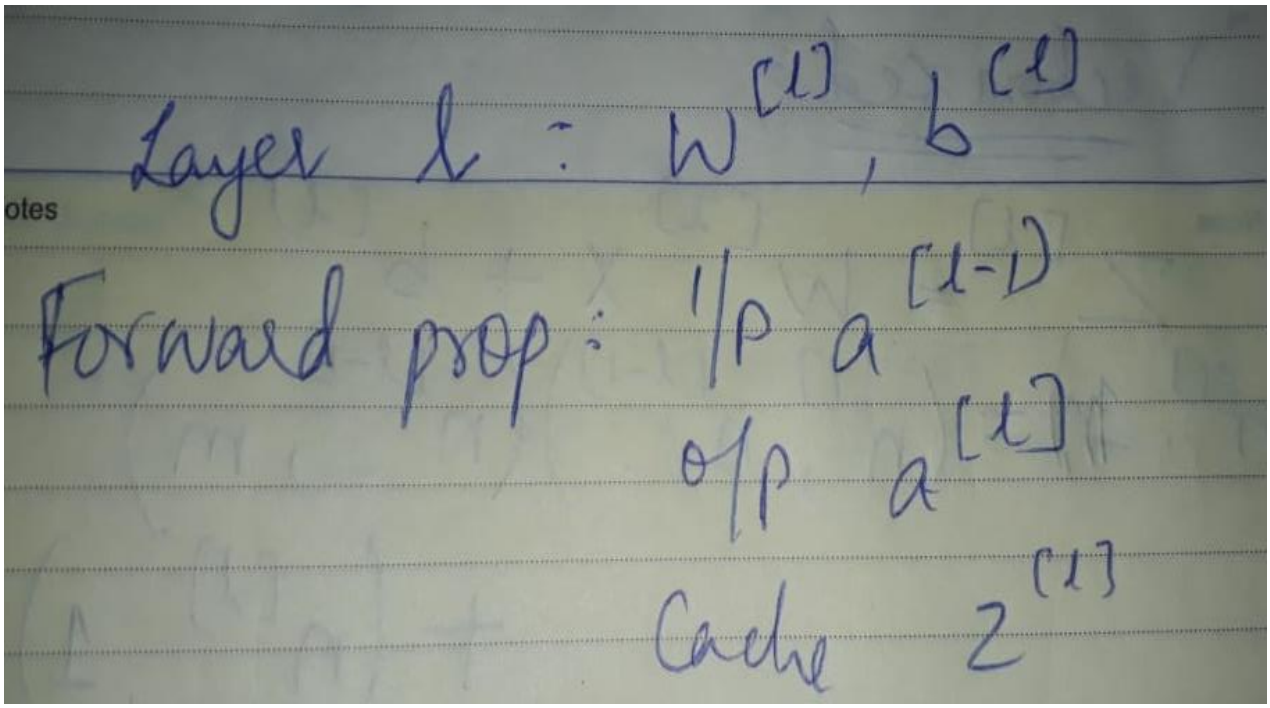
AUTHOR: SYEDA DARAQSHAN

WEEK 4

NEURAL NETWORKS AND DEEP LEARNING

BUILDING BLOCKS OF DEEP NEURAL NETWORK





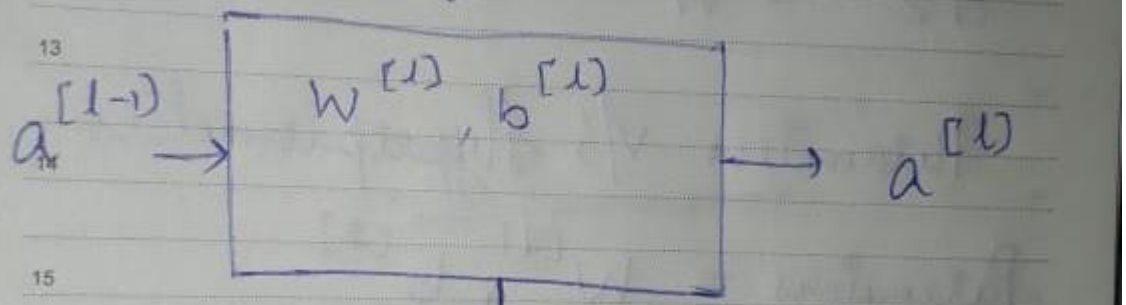
Backward prop

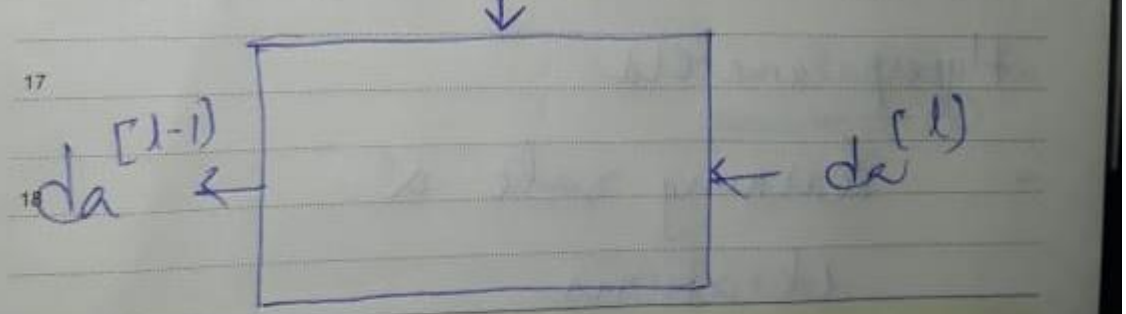
I/P $da^{[l]}$

O/P $da^{[l-1]}$ cache $z^{[l]}$

$dw^{[l]}$, $db^{[l]}$

layer l





Notes

$$dz^{[l]} = da^{[l]} * g^{[l]'}(z^{[l]})$$

$$dZ^{[L]} = A^{[L]} - Y$$

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

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

$$dZ^{[L-1]} = W^{[L]^T} dZ^{[L]} * g'^{[L-1]}(Z^{[L-1]})$$

Note that * denotes element-wise multiplication)

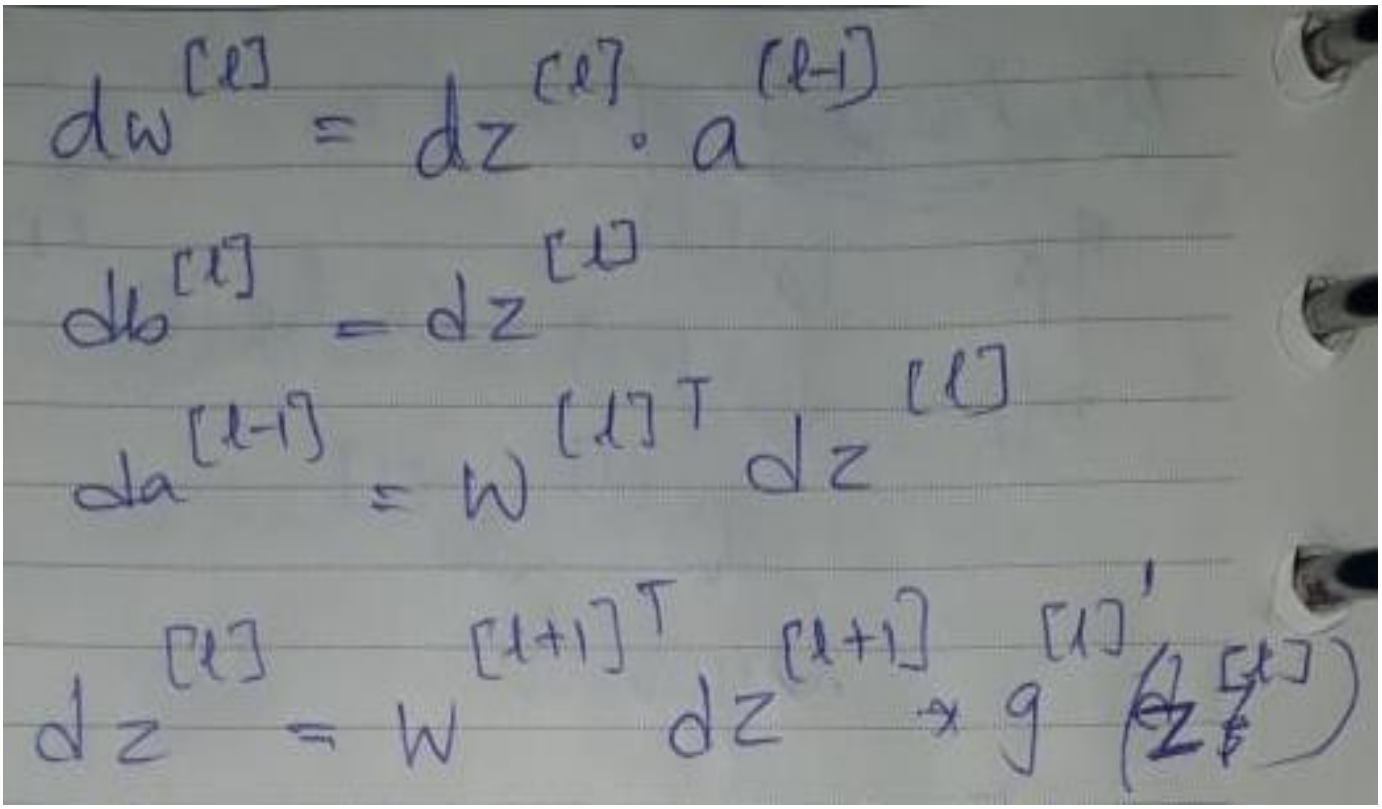
$$\vdots$$

$$dZ^{[1]} = W^{[2]} dZ^{[2]} * g'^{[1]}(Z^{[1]})$$

$$dW^{[1]} = \frac{1}{m} dZ^{[1]} A^{[0]^T}$$

Note that $A^{[0]^T}$ is another way to denote the input features, which is also written as X^T

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



PARAMETERS V/S HYPERPARAMETERS

Hyperparameters control the parameters,hence the name given!