神经网络第四章

2022年7月14日 星期四　01:29

logistic → "shallow" model

5 hiden layers → "deep"

$n^{[l]}$, $a^{[l]}$, $W^{[l]}$ 的 meaning

Forward Progation in a Deep Network

$X: z^{[1]} = W^{[1]}x + b^{[1]}$

$a^{[1]} = g^{[1]}(z^{[1]})$

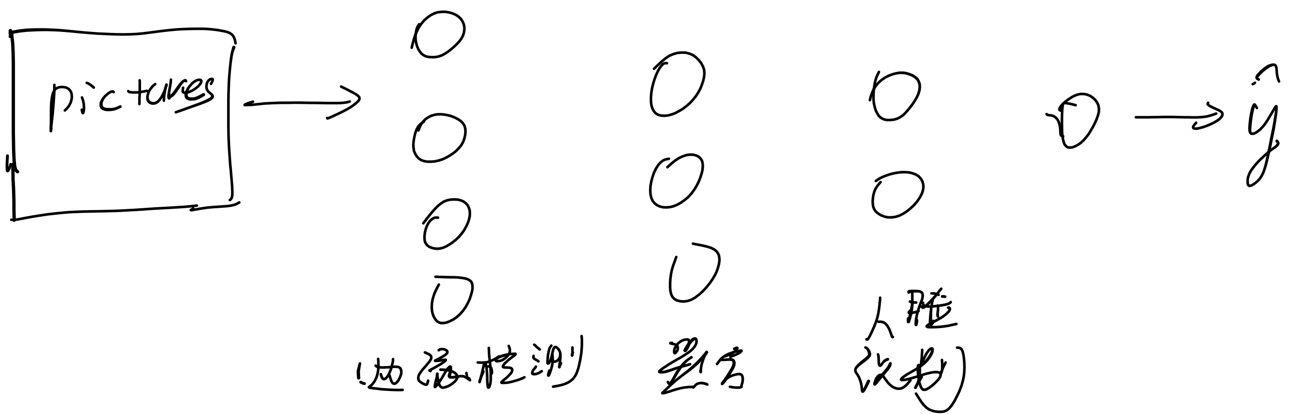
$z^{[l]} = W^{[l]} a^{[l-1]} + b^{[l]}$

$a^{[l]} = g^{[l]}(z^{[l]})$

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

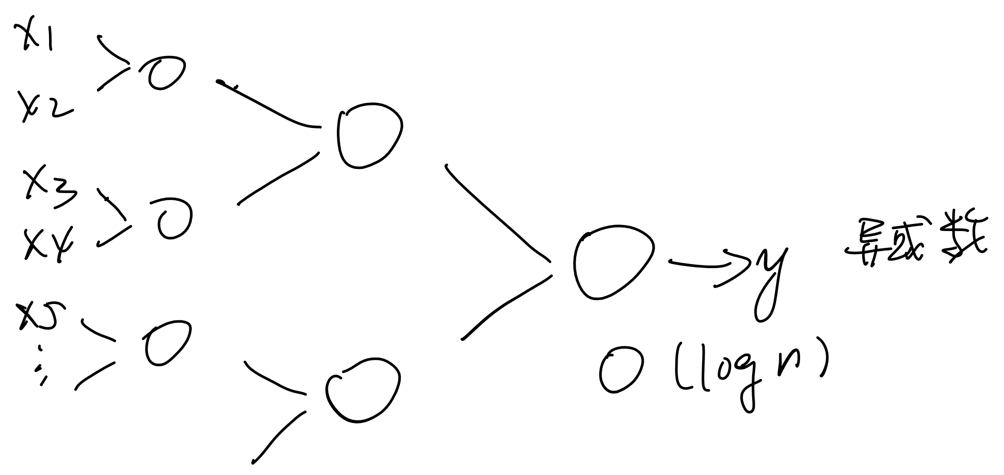
矩阵的维度判别（略）

Deep representation



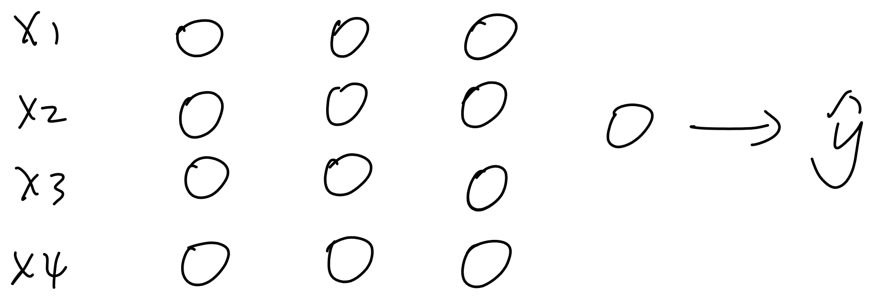
Circuit theory and deep learning.

$y = x_1 \text{ XOR } x_2 \text{ XOR } x_3 \text{ XOR } \ldots \text{ XOR } x_n$



shallow 判别: $O(2^n)$

构造 Neural Networks



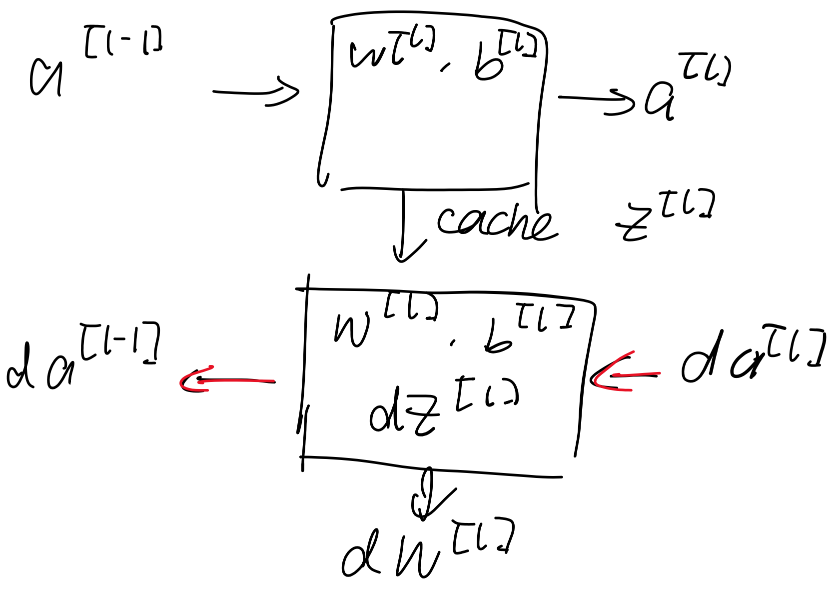
l: $W^{[l]}$, $b^{[l]}$

Forward: Input $a^{[l-1]}$, output $a^{[l]}$

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

$a^{[l]} = g^{[l]}(z^{[l]})$

Backward: Input $da^{[l]}$, output $da^{[l-1]}$



Vectorizal.

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

$H^{[l]} = g^{[l]} \cdot (Z^{[l]})$

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

$dW^{[l]} = dz^{[l]} \cdot a^{[l-1]T}$

$db^{[l]} = dz^{[l]}$

$da^{[l-1]} = W^{[l]T} \cdot dz^{[l]}$

$dz^{[l]} = W^{[l+1]T} dz^{[l+1]} * g^{[l]\prime} \cdot (z^{[l]})$

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

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

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

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

Hyperparameters & Parameters.

超参数: α, I, $n^{[l]}$ (自己设定的参数)

$W^{[1]}, b^{[1]}, W^{[2]}, b^{[2]}, W^{[3]}, b^{[3]} \ldots\ldots$