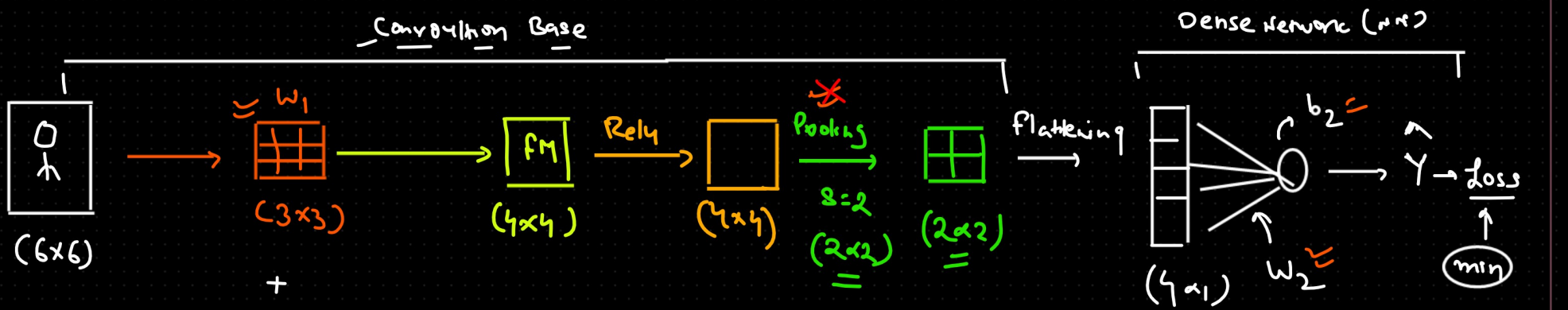



Loss = binary classification

$$L = -Y_i \log(\hat{Y}_i) - (1-Y_i)\log(1-\hat{Y}_i)$$

Trainable Parameter

$W_1 = (3\times3) = 9$ $\quad W_2 = 4 = 4$

$b_1 = (1\times1) = 1$ $\quad b_2 = 1 = 1$

$\Rightarrow 15$ ↓ trainable parameter

$$\Rightarrow \frac{n+2p-f}{s} + 1$$

$$\Rightarrow \frac{6+2\times0-3}{1} + 1$$

$$\Rightarrow 6-3+1 = \boxed{4}$$

$$I/P \rightarrow (*) \rightarrow Z_1 \xrightarrow{Relu} A_1 \xrightarrow[Pooling]{Max} P_1 \xrightarrow{Flatten} F \rightarrow Z_2 \xrightarrow{Sigmoid} A_2 \rightarrow L$$

$W_1, b_1$

## Forward Prop

$$Z_1 = Conv(x, w_1) + b_1$$
$$A_1 = Relu(Z_1)$$
$$P_1 = MaxPool(A_1)$$

$$F = Flatten(P_1)$$
$$Z_2 = W_2 F + b_2$$
$$A_2 = \sigma(Z_2)$$

Loss = Binary Cross entropy

$$\mathcal{L} = -y_i \log(\hat{y}_i) - (1-y_i)\log(1-\hat{y}_i)$$

optimizer ⇒ GD ← Gradient Decent opt.

$$w_n' = w_0' - \eta \frac{\partial \mathcal{L}}{\partial w_1} \qquad b_1 = b_1 - \eta \frac{\partial \mathcal{L}}{\partial b_1}$$

$$w_n^2 = w_0^2 - \eta \frac{\partial \mathcal{L}}{\partial w_2} \qquad b_2 = b_2 - \eta \frac{\partial \mathcal{L}}{\partial b_2}$$

$\frac{\partial \mathcal{L}}{\partial w_2}$ ← Partial derivation

$\left\{ \frac{\Delta \mathcal{L}}{\Delta w} \right\}_{\Delta \to 0}$ = rate of change

$\left\{ \frac{d\mathcal{L}}{dw} \right\}$ rate of change at the particular point

$$\left[\frac{\partial L}{\partial w_2} = \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial w_2}\right] \Rightarrow \text{Chain rule}$$

$$w_1 = \text{mat}(3\times3)$$
$$b_1 = (1\times1)$$

$$\left[\frac{\partial L}{\partial b_2} = \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial b_2}\right] \Rightarrow \text{Chain rule}$$

$$\left[\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial f} \times \frac{\partial f}{\partial p_1} \times \frac{\partial p_1}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial w_1}\right] \Rightarrow \text{Chain rule}$$

$$\left[\frac{\partial L}{\partial b_1} = \frac{\partial L}{\partial a_2} \times \frac{\partial a_2}{\partial z_2} \times \frac{\partial z_2}{\partial f} \times \frac{\partial f}{\partial p_1} \times \frac{\partial p_1}{\partial a_1} \times \frac{\partial a_1}{\partial z_1} \times \frac{\partial z_1}{\partial b_1}\right] \Rightarrow \text{Chain rule}$$

( Chain rule eq. wrt to the weight which will take during convolution )

- { Open Google Colab
  Functional API from keras }

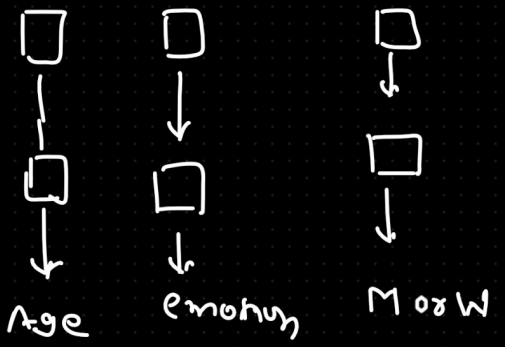
## Sequential API


⇒ Linear topology


⇒ Age, emotion, Men or women

First approach ⇒ Naive approach

3 network ⇒ 1st Network Age
2nd Network emotion
3rd Network Men or Women




Sequential model | CNN

2nd approach ⇒ Functional Network