

$$a_1 = \cancel{f(z_1)} = \frac{e^{z_1}}{\sum e^{z_i}} \quad 0 + 1 \quad \underline{\underline{0 + \infty}}$$

$$a_2 = \cancel{f(z_2)} = \frac{e^{z_2}}{\sum e^{z_i}} \quad 0 + 1$$

$$a_3 = \cancel{f(z_3)}$$

$$a_4 = \cancel{f(z_4)}$$

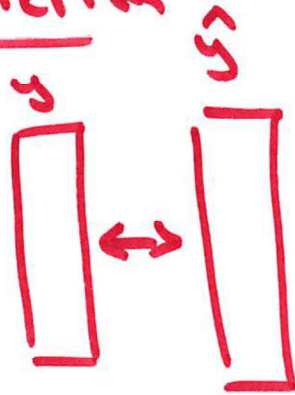
$$a_5 = \cancel{f(z_5)}$$

⋮  
⋮  
⋮  
⋮

$$\checkmark \frac{\cancel{e^{z_1}}}{\cancel{\sum e^{z_i}}}$$

$$e^{z_1} \quad e^{z_2} \quad \dots \quad e^{z_5}$$

# Loss Function



Loss Func  $\rightarrow$  distance measure

$\downarrow$   
a 'norm'

Cross entropy

binary classification

$$\frac{1}{N} \sum |y_i - \hat{y}_i| \quad \left| \quad \frac{1}{N} \sum (y_i - \hat{y}_i)^2 \right.$$

Below the first formula (MAE):

- $\rightarrow$  MAE
- SAE
- $\rightarrow$  MAD

Below the second formula (MSE):

- SSE
- MSE
- RMSE

Key

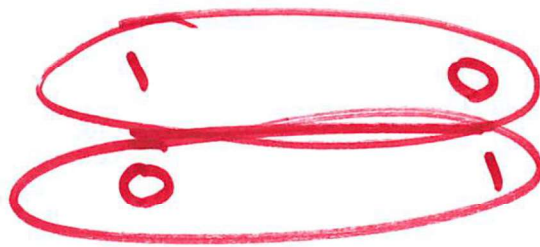
Cross  
entropy

$$L = - (y_i \ln(\hat{y}_i) + (1-y_i) \ln(1-\hat{y}_i))$$

$$L = \begin{cases} -\ln \hat{y}_i & \text{if } y_i = 1 \\ -\ln(1-\hat{y}_i) & \text{if } y_i = 0 \end{cases}$$

truth

pred



$$(\hat{y} - y)^2$$

→ Reg → not a big deal  
cause you are  
1 off.  $(1-0)^2$

bin

→ Clam →  $-\ln(0) \rightarrow \infty$

$-\ln(1-1) \rightarrow \infty$

1  
①

0.9  
⑥.4

~~$(\hat{y} - y)^2$~~