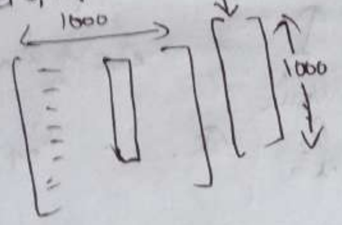
[negative log likelihood = nll]

$\hat{y}_i \rightarrow nll_i(\,)$ of
Average

$$\hat{y}_i = \frac{1}{K}\sum_{j=1}^{K} nll_j(\,)$$

$$S_i = \sigma(W_s S_{i-1} + W_x x_i + b_s)$$

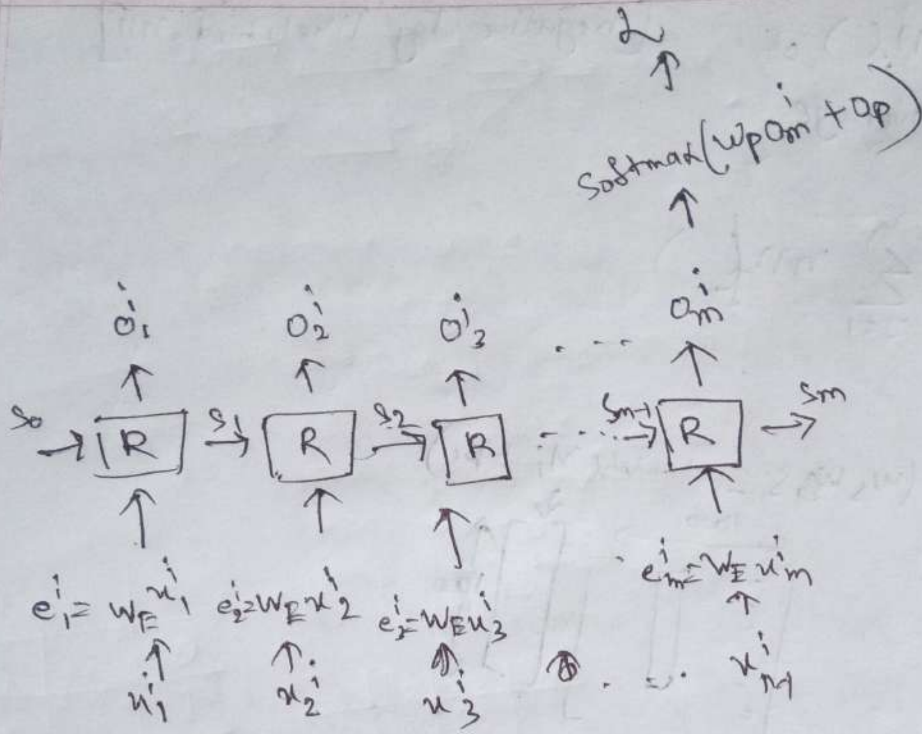


Example Your results are awesomely bad
Your results are not bad

$(\vec{X}_1 = (x_1, x_2 \cdots x_K), \quad \vec{P} = \begin{bmatrix} pos \\ neg \\ not \end{bmatrix}$

$x_1^{(i)} \quad x_2^{(i)} \quad x_3^{i} \quad \cdots \quad x_M^{i}$

<Pad>

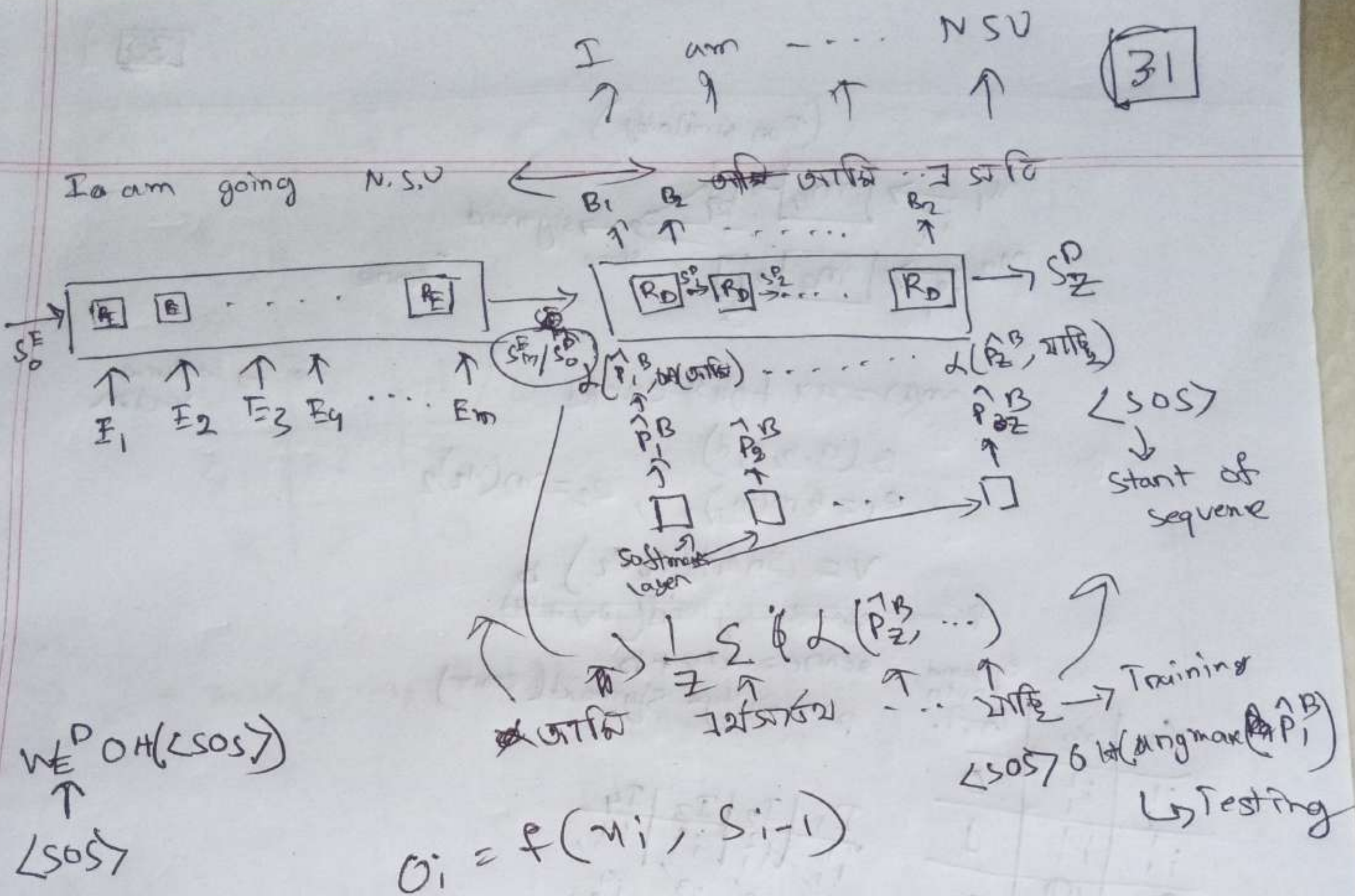


Many to One

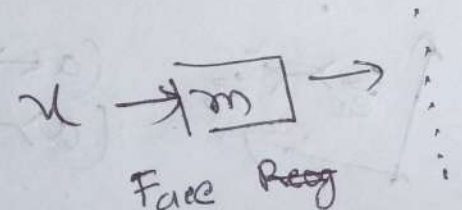
Encoder - I am going to North South University
(E_1 E_2 E_3 E_4 E_5 E_6 E_7)

Decoder - আমি নর্থসাউথ বিশ্ববিদ্যালয়ে যাচ্ছি।
(B_1 B_2 B_3 B_4)

$B_1 \downarrow$ I ✗ $\quad$ $B_2 \downarrow$ am ✗ $\quad$ $B_3 \downarrow$ going ✗



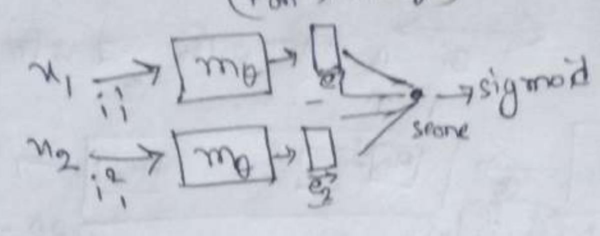
$$O_i = f(x_i, S_{i-1})$$



Face ~~Reog~~ Recognition

Open set Problem

(For similarity)



score

$\vec{e} \rightarrow$ feature vector

$m(x) = ax + bx^2 + cx^3 + d$

$\theta(a, b, c, d)$

$e_1 = m(x_1)_1$, $e_2 = m(x_2)_2$

$V = \text{concat}(e_1, e_2)$

~~score = sigmoid(wv + b)~~

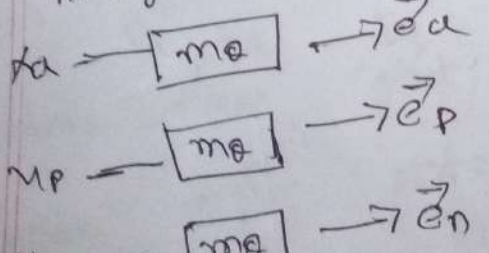
$\text{score} = wv + b$

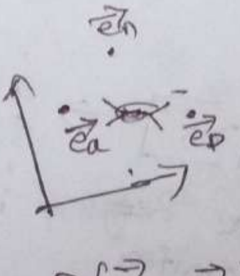
probability = sigmoid(score)

x_1	x_2	Ground truth G.T.
i_1^1	i_1^2	0
i_1^1	i_1^1	1
i_1^1	i_3^1	1

I_1	I_2	I_3	I_4
i_1^1,	i_1^2,	i_1^3,	i_1^4,
i_2^1,	i_2^2,	i_2^3,	i_2^4,
i_3^1	i_3^2	i_3^3	i_3^4

a - anchor
p - positive (same as a)
n - negative (different as a)





$(a \leftrightarrow p)$

$(a \rightarrow\leftarrow p)$

$D(\vec{e}_a, \vec{e}_p) = d_{a,p}$

$D(\vec{e}_a, \vec{e}_n) = d_{a,n}$

$d_{a,p} - d_{a,n} > 0$

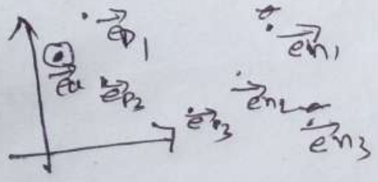
$d_{a,n} - \cancel{d_{a,p} + m} = 0$

$d_{a,p} - d_{a,n} + m = 0$

$g=5$

$d_{a,n}$	$d_{a,p}$	$d_{a,p}-d_{a,n}+5$
3	4	6
7	3	1
8	2	-1 / 0
3	3	5
7	2	0

$\vec{e}_a$, $\vec{e}_p$, $\vec{e}_n$ $\rightarrow$ $\alpha = \max(0, d_{ap} - d_{an} + g)$ $\rightarrow$ Triplet loss

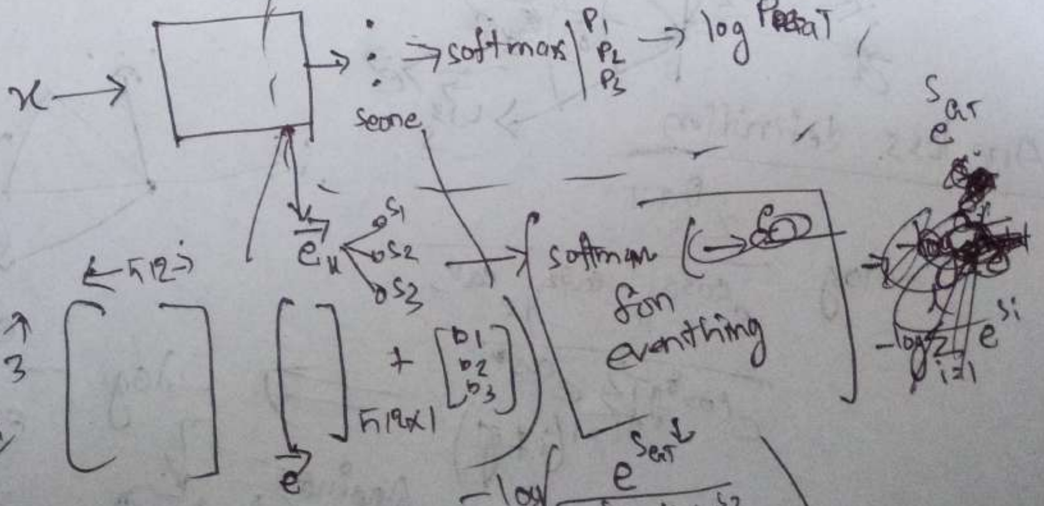


$\vec{r} = m_\theta$ (reference)

$\vec{q} = m_\theta$ (query)

$d_{r,q} = D(\vec{r}, \vec{q})$ $\rightarrow$ if ($d_{r,q}$ < threshold) similar end if.

Angular Margin loss



$$\begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = (\text{softmax})\, n\ell \left(\begin{bmatrix} \quad \end{bmatrix}_{3 \times 512} \begin{bmatrix} \quad \end{bmatrix}_{\vec{e}} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}_{512 \times 1} \right)$$

$$-\log\left(\frac{e^{s_{ar}}}{e^{s_1}+e^{s_2}+e^{s_3}}\right)$$

$s_1 = dot(\vec{w}_s^1, \vec{e})$ [Equivalent to before]

$$s_i = dot(\vec{w}_s^i, \vec{e})$$

$$P_1 = \frac{e^{dot(\vec{w}_s^1, \vec{e})}}{e^{dot(\vec{w}_s^1, \vec{e})} + e^{dot(\vec{w}_s^2, \vec{e})} + e^{dot(\vec{w}_s^3, \vec{e})}}$$

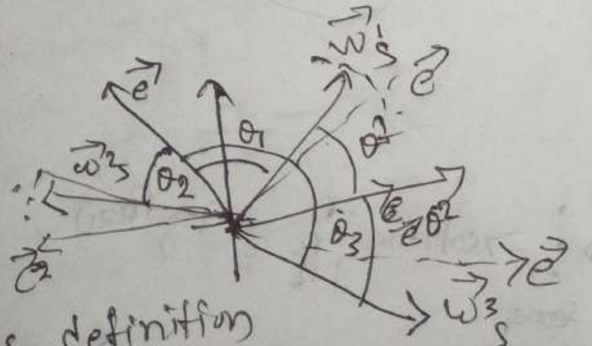
$$s_i = \|\vec{w}\| \, \|\vec{e}\| \cos\theta_i$$

$\|w_s^i\| = 1$ for $i = 1, 2, 3$

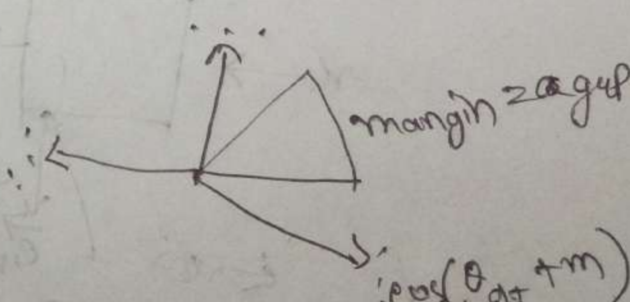
$\wedge \ \|\vec{e}\| = 1$

$$s_i = \cos\theta_i$$

$$P_1 = \frac{e^{\cos\theta_1}}{e^{\cos\theta_1} + e^{\cos\theta_2} + e^{\cos\theta_3}}$$



$\theta^1 ? \theta^2$ = for $\vec{w}_s^1 ? \vec{w}_s^2$



Arc loss definition

$$-\log \frac{e^{\theta_{g.T}}}{e^{\cos\theta_1} + e^{\cos\theta_2} + e^{\cos\theta_3}}$$

$$\frac{}{e^{\cos\theta_{gT}} \sum_{i=1 (i \neq gT)} e^{\cos\theta_i}} \Longrightarrow -\log \frac{e^{\cos(\theta_{gT} + m)}}{e^{\cos(\theta_{gT} + m)} + \sum_{j=1 (j \neq gT)} e^{\cos\theta_j}}$$

Arcface

Arcface