Assignment: 02

Course Code: CSE427

Date of Submission: 13-01-2026

Name: YASIR

Department: Computer Science & Engineering

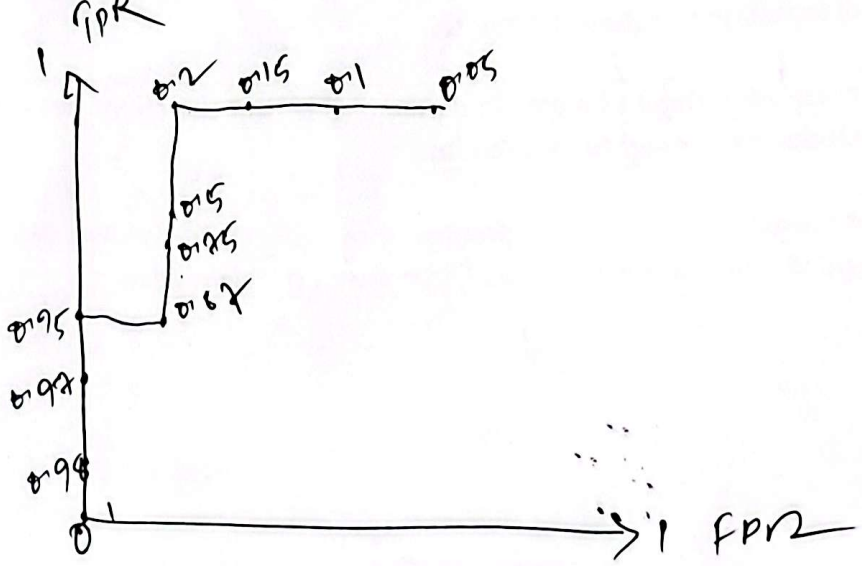
Student ID: 23101122

Section: 04

3/

Th	TPR	FPR
1	$\frac{0}{6} = 0$	$\frac{0}{4} = 0$
0.925	$\frac{1}{6} = 0.17$	$\frac{0}{4} = 0$
0.965	$\frac{2}{6} = 0.33$	0
0.95	0.5	0
0.87	0.5	0.25
0.75	0.67	0.25
0.5	0.67	0.25
0.2	1	0.25
0.13	1	0.5
0.1	1	0.75
0.05	1	1
0	1	1

ROC curve;



Threshold = 100th, TPR = 0.20, as it is giving 100th TPR with new FPR.

Also, to achieve 0th FP, the threshold is 0.95, as it's giving 0th FP, as well as highest TPR among the other rates.

2

Here,

$x_1 = x_2 = 1$, $w_{all} = 0.5$, $b_{all} = 0.5$

$\therefore z_1 = 0.5 \times 1 + 0.5 \times 1 + 0.5$

$= 1.5$

$a_1 = \sigma(z_1) = 0.82$

$\therefore z_2 = 1.5$, $a_2 = 0.82$, $z_3 = 1.5$, $a_3 = 0.82$

$z_5 = 0.82 \times 0.5 + 0.82 \times 0.5 + 0.82 \times 0.5 + 0.5$

$= 1.73$

$\hat{y}_6 = \theta(z_5) = 0.85$

$\therefore z_6 = 1.73$, $\hat{y}_7 = 0.85$

And,

$w_{13} = \delta_3 x_1$

$\delta_6 = L \times b y_6'$

$= (1 - 0.85) \{ \sigma(z_5)(1 - \sigma(z_5)) \}$

$= 0.15 \times 0.85 \times 0.15 = 0.0191$

$$\delta_7 = err_7 \cdot y_7'$$

$$= (0 - 0.85) \times \{0.85 \times 0.15\}$$

$$= -0.1084$$

$$\delta_3 = (w_{36} \times \delta_6 + w_{37}\delta_7) \times a_i'$$

$$= \{(0.5 \times 0.191) + (0.5 \times (-0.1084))\} \times (0.82 \times (1 - 0.82))$$

$$= -0.00659$$

$$\rightarrow w_{13}^{new} = w_{13}^{old} + \eta \delta_3 x_1$$

$$= 0.5 + 0.1 \times (-0.00659) \times 1$$

$$= 0.4993$$

$$w_{23}^{new} = 0.4993$$

$$b_3^{new} = b_3^{old} + \eta \delta_3$$

$$= 0.4993$$

$$\frac{dL}{dw_{14}} = \left\{ \frac{dL}{dy_6} \times \frac{dy_6}{dz_5} \times \frac{dz_5}{da_2} \times \frac{da_2}{dz_2} \times \frac{dz_2}{dw_{14}} \right\} +$$

$$\left\{ \frac{dL}{dy_7} \times \frac{dy_7}{dz_6} \times \frac{dz_6}{da_2} \times \frac{da_2}{dz_2} \times \frac{dz_2}{w_{14}} \right\}$$

$$= \left\{ (\hat{y}_6 - y_6) \{ \sigma(z_5)(1-\sigma(z_5)) \} \cdot w_{46} \{ \sigma(z_2)(1-\sigma(z_2)) \} \times 1 \right\} +$$

$$\left\{ (\hat{y}_7 - y_7) \{ \sigma(z_6)(1-\sigma(z_6)) \} \cdot w_{47} \cdot \{ \sigma(z_2)(1-\sigma(z_2)) \} \times 1 \right]$$

$$= -1.4114 \times 10^{-3} + 7.998 \times 10^{-3}$$

$$= 6.59 \times 10^{-3}$$

$$w_{14}^{new} = 0.5 - [0.1 \times 6.59 \times 10^{-3}] = 0.4993$$

$$w_{24}^{new} = 0.4993$$

$$b_4^{new} = 0.4993$$

$$\delta_5 = \left(w_{56}\delta_6 + w_{57}\delta_7\right) a_3'$$

$$= \left\{(0.5 \times 0.19) + (0.5 \times (-0.108))\right\} \times \left(\sigma(z_3) \times (1-\sigma(z_3))\right)$$

$$= 6.096 \times 10^{-3}$$

$$\therefore b_5^{new} = 0.5 + 0.1 \times \left(6.096 \times 10^{-3}\right) = 0.5006$$

4/

Class A:

$$\text{Sensitivity} \,|\, \text{recall} = \frac{800}{800+100+100} = 0.8$$

$$\text{Specipicity} = \frac{1750}{1750+50+200} = 0.88$$

$$\text{Precision} = \frac{800}{800+50+200} = 0.76$$

Class B:

$$\text{Sensitivity} \,|\, \text{recall} = \frac{900}{900+50+50} = 0.9$$

$$\text{Specificity} = \frac{1800}{1800+100+100} = 0.9$$

$$\text{Precision} = \frac{900}{900+100+100} = 0.82$$

Class - C:

$$\text{Sensitivity / recall} = \frac{700}{700+200+100} = 0.7$$

$$\text{Specificity} = \frac{1850}{1850+100+50} = 0.93$$

$$\text{precision} = \frac{700}{700+100+50} = 0.82$$

9

$$w = (1.87, 0.09) \rightarrow \text{pos} = 0$$

$$b_0 = (-0.36, 0.27) \rightarrow \text{pos} = 1$$

Positional encoding:

$$PE(0,0) = \sin\left(\frac{0}{10000^{\frac{0}{2}}}\right) = 0$$

$$PE(0,1) = \cos\left(\frac{0}{10000^{\frac{1}{2}}}\right) = 1$$

$$PE(1,0) = \sin\left(\frac{1}{10000^{\frac{0}{2}}}\right) = 0.84$$

$$PE(1,1) = \cos\left(\frac{1}{10000^{\frac{1}{2}}}\right) = 0.54$$

$$we = (1.87, 1.09)$$

$$PE = (0.06, 0.81)$$

$$query(we) = |1.87, 1.09| \begin{vmatrix} 1.1 & 0.6 \\ -2.6 & 2.4 \end{vmatrix}$$

$$= (-0.99, 3.73)$$

$$(PE) = |0.06, 0.81| \times \begin{vmatrix} 1.1 & 0.6 \\ -2.6 & 2.4 \end{vmatrix}$$

$$= (-2.2, 1.98)$$

$$key(we) = |1.87, 1.09| \times 4$$

$$= -4.82, 1.9|$$

$$(go) = (0.06, 0.81) \times 4$$

$$= -1.31, 0.76$$

$$value(we) = |1.87, 1.09| \times v$$

$$= (2.48, -2.09)$$

$$(go) = |0.06, 0.47| \times 4 \times v$$

$$= (-0.15, -0.22)$$

WC → WC = (−0.99 × −4.82) · (3.73 × 1.91)

= 11.89 = soft = 0.99

WC → GO = (−0.99 × −4.82) · (3.73 × 0.76)

= 4.15 = soft = 0.004

v(WC) × 0.99 = (0.99 × 2.48), (0.99 × −2.09)

= 2.46, −2.069

v(GO) × 0.004 = −0.0006, −0.00088

WC = (2.47, −2.08)

GO → GO = GOk × GOq

= (−2.28, −1.31) · (1.98, 0.76)

= [illegible] 3.12 = soft = 0.007

GO → WC = (−2.2, −4.6) · (1.98, 1.91)

= 10.17 = soft = 0.99

output(G10) = 0.99 × (2.48, -2.09) + 0.001 (-0.15, -0.22)

= (2.48, -2.09)

Decoder

EOS = |2.7, -1.34|

query (EOS) = |2.7, -1.34| × Q

= |3.7, 0.53|

~~key (EOS) = Q~~

key output of Encoder:

We = |2.47, -2.06| × $\begin{vmatrix} -1.1 & 0.3 \\ -1.5 & -0.4 \end{vmatrix}$

= |0.4, 2.4|

So = |2.48, -2.09| $\begin{vmatrix} 1.1 & 0.3 \\ -1.5 & -0.4 \end{vmatrix}$.

= |0.41, 2.4|

Value: We = |2.47, -2.06| × V

= |5.22, 2.52|

$$GTO = |2.48, -2.09| \times V$$

$$= |5.24, 2.54|$$

$$EOS \rightarrow We = (3.65, 0.4) \cdot (0.53, 2.4)$$

$$= 1.93 = \text{soft} = 0.43$$

$$EOS \rightarrow GTO = (3.65, 0.41) \cdot (0.53, 2.41)$$

$$= 1.97 = \text{soft} = 0.57$$

$$EOS \text{ embedding} = 0.49 \times (5.22, 2.52) + 0.51 (5.24, 2.54)$$

$$= |5.23, 2.53|$$