Student ID: 40189233

6630 - Assignment - 1

Name: Sai Chandra Sekhar Reddy Dwarampudi

Q1

① a) $f(z) = \frac{1}{1+e^{-z}}$

where $z = w_1x_1 + w_2x_2 + w_3x_3 \cdots w_nx_n + b$

$y = f(z)$

$w = (w_1, w_2, w_3) = (0.5, -0.25, 0.1)$

bias $b = 0.5$

input $x = (0,0,0) \Rightarrow y = f(z)$

$\Rightarrow y = f(w_1x_1 + w_2x_2 + w_3x_3 + b)$

$y = f(0.5\times0 + -0.25\times0 + 0.1\times0 + 0.5)$

$y = f(0.5)$

$y = \frac{1}{1+e^{-0.5}} = \frac{1}{1+0.6065} = \frac{1}{1.6065}$

$= \underline{0.6224}$

The value of the output $\underline{y = 0.6224}$

ⓑ Rectified Linear unit (ReLU) $f(z) = \begin{cases} z & \text{if } z \geq 0 \\ 0 & \text{otherwise} \end{cases}$

given

weights $w = (0.5, -0.25, 0.1)$

bias $b = 0.5$

input $x = (0,0,0)$

output $y = f(z)$

$y = f(w_1x_1 + w_2x_2 + w_3x_3 + b)$

$= f(0.5\times0 - 0.25\times0 + 0.1\times0 + 0.5)$

$= f(0.5)$

$\underline{y = z = 0.5}$ $\because \underline{0.5 \geq 0}$

$\therefore$ The value of output $\underline{y = 0.5}$

ⓒ f Sigmoid function $f(z) = \frac{1}{1+e^{-z}}$

given values

$w = (0.5, -0.25, 0.1)$; $b = 0.5$

$x = (1,1,1)$

output $y = f(z) = f(w_1x_1 + w_2x_2 + w_3x_3 + b)$

$y = f(0.5\times1 - 0.25\times1 + 0.1\times1 + 0.5)$

$y = f(0.5 - 0.25 + 0.6)$

$y = f(0.85)$

$y = \frac{1}{1+e^{-0.85}} = \frac{1}{1+0.427}$

value of output $\underline{\underline{y = 0.7007}}$ (since $y \geq 0.5$) $\Rightarrow$ Class 1

ⓓ True class label $t = 1$

output $y = 0.55$

Squared loss $E = \frac{1}{2}(t-y)^2$

Cross-entropy loss $E = -[t\log y + (1-t)\log(1-y)]$

Squared loss $E = \frac{1}{2}(t-y)^2$

$= \frac{1}{2}(1-0.55)^2$

$= \frac{1}{2}(0.2025)$

$= \underline{0.1012}$

Cross Entropy loss $E = -[1\log 0.55 + (1-1)\log(1-0.55)]$

$= -[\log 0.55]$

$= -[-0.2596]$

$= \underline{\underline{0.2596}}$

(e). If $\eta = 0.1$ & $t = 1$

what are the updated weights after performing backpropagation.

$$w \leftarrow w - \eta \frac{\partial E}{\partial w}$$

gradient $\frac{\partial E}{\partial w} = \frac{\partial E}{\partial y} \times \frac{\partial y}{\partial z} \times \frac{\partial z}{\partial w}$

Sigmoid function $f = \frac{1}{1+e^{-z}}$ & $E = \frac{1}{2}(t-y)^2$

Current weights $w = (0.5, -0.25, 1)$, bias $b = 0.5$, input $x = (1,1,1)$

$\Rightarrow$ gradient $\frac{\partial E}{\partial w} = \frac{\partial (t-y)^2/2}{\partial y} \cdot \frac{\partial f(z)}{f(z)} \cdot \frac{\partial x \cdot w}{w}$

$$= -(t-y) f(x^T z)(1 - f(z)) x$$

$y = f(z)$

$z = 0.5 \times 1 - 0.25 \times 1 + 0.1 \times 1 + 0.5$

$= 0.85$

$y = \frac{1}{1+e^{-0.85}} = \underline{0.7007}$

① $\rightarrow = -(1-0.7007) f(0.85)(1 - f(0.85))$

$= -(0.2993)(0.7007)(0.2993)$

$= \underline{-0.0627}$

② $\rightarrow = -(1-0.7007) f(0.85)(1 - f(0.85))$

$= -0.0627$

③ $\rightarrow = -0.0627$

$w_1 = 0.5 - (0.1 \times -0.0627) = 0.5 + 0.00627 = 0.5062$

$w_2 = -0.25 - (0.1 \times -0.0627) = -0.25 + 0.00627 = -0.243$

$w_3 = 0.1 - (0.1 \times -0.0627) = 0.1 + 0.00627 = 0.10627$

updated weights $w = \underline{(0.5062, -0.243, 0.10627)}$

Q2

Given

input feature map = $\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$

Convolution filter = $\begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix}$

(a) Complete output feature map of the Convolution layer = $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$

$2+0+8+3 = 13$

$A = \begin{bmatrix} 1\times2 & 0\times1 & 3 \\ 2\times4 & 1\times3 & 0 \\ 0 & 2 & 1 \end{bmatrix}$

$C = \begin{bmatrix} 1 & 0 & 3 \\ 2\times2 & 1\times1 & 0 \\ 0\times4 & 2\times3 & 1 \end{bmatrix}$

$0+3+4+0 = 7$

$4+1+0+6 = 11$

$B = \begin{bmatrix} 1 & 0\times2 & 3\times1 \\ 2 & 1\times4 & 3\times0 \\ 0 & 2 & 1 \end{bmatrix}$

$D = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1\times2 & 0\times1 \\ 0 & 2\times4 & 1\times3 \end{bmatrix}$

$2+0+8+3 = 13$

Output feature map = $\begin{bmatrix} 13 & 7 \\ 11 & 13 \end{bmatrix}$

(b) Output of the max pooling layer where a 2x2 input is down sampled to one output. = 13

Output feature map $\begin{bmatrix} 13 & 7 \\ 11 & 13 \end{bmatrix}$

Q3 Given the values

$h_0 = 0$

time $t = 3 \rightarrow (x_1, x_2, x_3) = (2, 1, 4)$

weight b/w i/p & Hidden layer $\rightarrow w_x(1), w_x(2), w_x(3) = (3, 2, 2)$

weight b/w hidden & output layer $\rightarrow w_h(1), w_h(2), w_h(3) = (4, 2, 3)$

weight b/w previous Hidden layer output & the current hidden layer input $\rightarrow w_y(1), w_y(2), w_y(3) = (2, 1, 4)$

$$h_t = \begin{cases} z_t & \text{if } z_t \geq 0 \\ 0 & ; \text{ otherwise} \end{cases} \quad \& \quad z_t = h_{t-1} w_h + x_t w_x$$

ⓐ value of $h_1 \Rightarrow z_1 = h_0 w_h + x_1 w_x$

$= 0 \times 4 + 2 \times 3$

$= 0 + 6$

$= \underline{6}$

$z_1 > 0$ so implies $h_1 = z_1$

$\underline{\underline{h_1 = 6}}$

value of $h_2 \Rightarrow z_2 = h_1 w_h + x_2 w_x$

$= 6 \times 2 + 1 \times 2$

$= 12 + 2$

$= \underline{14}$

$z_2 > 0$ implies $h_2 = z_2$

$\underline{\underline{h_2 = 14}}$

value of $h_3 \Rightarrow z_3 = h_2 w_h + x_3 w_x$

$= 14 \times 3 + 4 \times 2$

$= 42 + 8$

$= \underline{50}$

$z_3 > 0$ implies $h_3 = z_3$

$\underline{\underline{h_3 = 50}}$

ⓑ value of $y_1 = h_1 \times w_y = 6 \times 2 = 12$

value of $y_2 = h_2 \times w_y = 14 \times 1 = 14$

value of $y_3 = h_3 \times w_y = 50 \times 4 = 200$

$\therefore$ The output $y \ (y_1, y_2, y_3) = (12, 14, 200)$

Q4 Case 1:

$TP = 482$

$FP = 72$

$FN = 13$

$TN = 422$

$P = 495$

$N = 494$

① Accuracy $= \frac{TP+TN}{P+N}$

$= \frac{482+422}{495+494} = \frac{904}{985} = 0.917 \approx 91.7\%$

② Balanced accuracy $= \frac{TPR+TNR}{2}$

$= \frac{0.97+0.85}{2} = 0.91 \approx 91\%$

③ $TPR = \frac{TP}{P}$

$= \frac{482}{495} = 0.97 = 97\%$

④ $TNR = \frac{TN}{N}$

$= \frac{422}{494} = 0.85 \approx 85\%$

⑤ $FPR = \frac{FP}{N}$

$= \frac{72}{494} = 0.145 \approx 14.5\%$

⑥ $FNR = \frac{FN}{P}$

$= \frac{13}{495} = 0.026 \approx 2.6\%$

⑦ $PPV = \frac{TP}{TP+FP}$

$= \frac{482}{482+72} = \frac{482}{554} = 0.87 \approx 87\%$

⑧ F_1 Score $= 2 \times \frac{PPV \times TPR}{PPV + TPR}$

$= 2 \times \frac{0.87 \times 0.97}{0.87 + 0.97}$

$= 2 \times \frac{0.843}{1.84}$

$= 2 \times 0.458 = 0.916$

$= 91.6\%$

Case 2:

TP = 803

FP = 72

FN = 13

TN = 101

P = 816

N = 173

① Accuracy $= \frac{803+101}{816+173}$

$= \frac{904}{989} = 0.914 \approx \underline{91.4\%}$

② Balanced Accuracy $= \frac{0.984+0.583}{2}$

$= 0.783 \approx \underline{78.3\%}$

③ TPR $= \frac{803}{816} = 0.984 \approx \underline{98.4\%}$

④ TNR $= \frac{101}{173} = 0.583 \approx \underline{58.3\%}$

⑤ FPR $= \frac{72}{173} = 0.416 \approx \underline{41.6\%}$

⑥ FNR $= \frac{13}{816} = 0.0159 \approx \underline{1.5\%}$

⑦ PPV $= \frac{803}{803+72} = \frac{803}{875} = 0.917 \approx \underline{91.7\%}$

⑧ F score $= 2 \times \frac{0.917 \times 0.984}{0.917+0.984} = 2\times\frac{0.902}{1.901} = 0.948 = \underline{94.8\%}$

Observation:

In Case 1: accuracy = 91.7 %

Balanced accuracy = 91 %

whereas in Case 2:

accuracy = 91.4 %

Balanced accuracy = 78.3 %

Case1 represents balanced data with high precision (PPR) and low recall value (FPR). In case 2, since there is a huge variation between the accuracy which represents imbalanced data with high precision and moderate recall.