N. KAUSIK
CS21M037

Page No.
Date

Tutorial 8

1) i) Input nodes – A, B, C
ii) Hidden nodes – D, E
iii) Output nodes – F, G
iv) # Params = $3\times2 + 2\times2 = 10$
with bias $\Rightarrow 3\times2 + 2\times2 + 1\times2 + 1\times2 = 14$

2.1) $D \rightarrow 0.1\times4 + 0.2\times3 + 0.2\times1 = 1.2$
$E \rightarrow 0.1\times2 + 0.2\times5 + 0.2\times6 = 2.4$

2.2) $D_o \rightarrow \frac{1}{1+e^{-1.2}} = 0.77$

$E_o \rightarrow \frac{1}{1+e^{-2.4}} = 0.92$

2.3) $F \rightarrow 0.77\times1 + 0.92\times2 = 2.61$
$G \rightarrow 0.77\times3 + 0.92\times4 = 5.99$

2.4) $F_o \rightarrow \frac{e^{2.61}}{e^{2.61} + e^{5.99}} = 0.03$

$G_o \rightarrow \frac{e^{5.99}}{e^{2.61} + e^{5.99}} = 0.97$

2.5) Loss $\Rightarrow$ Cross Entropy Loss
$= -\sum y \log \hat{y}$

$\Rightarrow -(0\times\log F_o + 1\times\log G_o) = -\log_2(0.967)$

$= 0.0146$

3) Yes, input $X = \begin{bmatrix} 0.1 \\ 0.2 \\ 0.2 \end{bmatrix}$ output $Y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

Weights,
$W_1 = \begin{bmatrix} 4 & 2 \\ 3 & 5 \\ 1 & 6 \end{bmatrix}$ $W_2 = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$

Page No.
Date

To get out, $\Rightarrow$ Softmax $\left(W_2^T \sigma\left(W_1^T x \right) \right)$

4) If we dont use non-linear functions, the neural network simply does a linear operation on the inputs and so cant model a non-linear function.

5) $\sigma(x) = \frac{1}{1+e^{-x}}$ $\quad \frac{d\sigma(x)}{dx} = \frac{-1}{(1+e^{-x})^2} \cdot -e^{-x}$

$$\frac{d\sigma(x)}{dx} = \frac{1+e^{-x}-1}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}} - \left(\frac{1}{1+e^{-x}}\right)^2$$

$$= \sigma(x) - (\sigma(x))^2 = \sigma(x)\left(1-\sigma(x)\right)$$

6) $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x}}{e^{2x}} \frac{1-e^{-2x}}{1+e^{-2x}}$

$$\frac{d\tanh(x)}{dx} = \frac{2e^{-2x}}{1+e^{-2x}} + (1-e^{-2x}) \cdot \frac{-1}{(1+e^{-2x})^2} \cdot -2e^{-2x}$$

$$= \frac{2e^{-2x}}{1+e^{-2x}} + \frac{2e^{-2x}(1-e^{-2x})}{(1+e^{-2x})^2}$$

$$= \frac{2e^{-2x}}{1+e^{-2x}} \left(1 + \frac{1-e^{-2x}}{1+e^{-2x}}\right)$$

$$= \left(1 - \frac{(1-e^{-2x})}{1+e^{-2x}}\right)\left(1 + \frac{1-e^{-2x}}{1+e^{-2x}}\right)$$

$$= \left(1-\tanh(x)\right)\left(1+\tanh(x)\right)$$

$$= 1 - (\tanh(x))^2$$

Page No.
Date

7) $\sigma(x) = \frac{1}{1+e^{-x}}$

$\tanh(x) = \frac{e^{+x} - e^{-x}}{e^{x} + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}}$

$= \frac{1}{1+e^{-2x}} - \frac{(1+e^{-2x} - 1)}{1+e^{-2x}}$

$= \frac{1}{1+e^{-2x}} - 1 + \frac{1}{1+e^{-2x}}$

$= 2 \cdot \frac{1}{1+e^{-2x}} - 1$

$= (2\sigma(2x) - 1)$ //