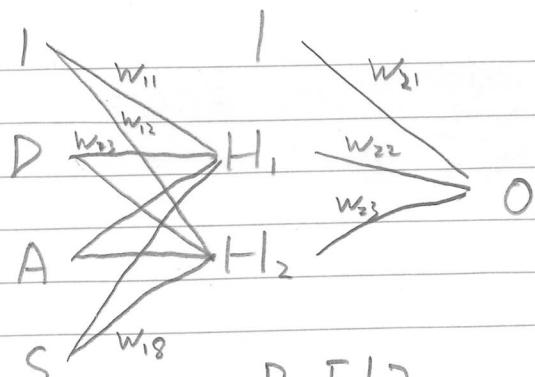


# Week 9 Q2.



$$W_{11} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \\ -2 \\ 2 \\ 3 \\ -2 \end{bmatrix} \quad W_{2j} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

$$G(h) = \frac{1}{1+e^{-h}}$$

$$H_1 = 1 + 1 \cdot 1 + (-2) \cdot 2 + 3 \cdot 2 = 4$$

$$G(H_1) = 0.98$$

$$H_2 = -1 + (-1) \cdot 1 + (2) \cdot 2 + (-2) \cdot 2 = -2$$

$$G(H_2) = 0.12$$

forward  $O = -1 + 0.98 + 0.12 = 0.1$   $G(O) = 0.52 = Z$

$$L(w) = -[y \log(Z) + (1-y) \log(1-Z)]$$

Binary Cross Entropy Loss

$$\frac{\partial L}{\partial z} = -\left[\frac{y}{z} - \frac{1-y}{1-z}\right] = \frac{z-y}{z(1-z)}$$

$$Z = G(h) = \frac{1}{1+e^{-h}} \leftarrow \text{Sigmoid function.}$$

$$\frac{dz}{dh} = z(1-z)$$

$$\frac{\partial O}{\partial} = 1 \text{ for bias}$$

$$\frac{\partial L}{\partial w_{21}} = \frac{\partial L}{\partial z} \cdot \frac{\partial z}{\partial O} = \frac{z-y}{z(1-z)}, Z(1-Z) = 0.52 - 1 = -0.48 \text{ bias}$$

$$\frac{\partial L}{\partial w_{22}} = 0.98 \cdot (-0.48) = -0.48 \leftarrow (\text{current value}) \times (\text{backward gradient})$$

$$\frac{\partial L}{\partial w_{23}} = 0.12 \cdot (-0.48) = -0.06 \leftarrow \text{to } H_2$$

$$\frac{\partial L}{\partial w_{24}} = 0.98(1-0.98) = 0.02$$

$$\begin{aligned} \text{Eg: } & \frac{\partial L}{\partial w_{22}} = \frac{\partial L}{\partial H_1} \cdot \frac{\partial G(H_1)}{\partial H_1} \cdot \frac{\partial L}{\partial G(H_1)} \\ & \frac{\partial L}{\partial w_{22}} = \frac{\partial L}{\partial z} \cdot \frac{\partial z}{\partial O} \cdot \frac{\partial O}{\partial w_{22}} \end{aligned}$$

} derivative of Sigmoid

$$\frac{\partial L}{\partial H_2} = 0.12(1-0.12) = 0.11$$

bias

$$\frac{\partial L}{\partial w_{11}} = \begin{bmatrix} (0.02)(-0.48) \\ (0.11)(-0.06) \\ (1)(-0.0094) \\ (1)(-0.0066) \\ (2)(-0.0094) \\ (2)(-0.0066) \\ (2)(-0.0094) \\ (2)(-0.0066) \end{bmatrix} = \begin{bmatrix} -0.0094 \\ -0.0066 \\ -0.0094 \\ -0.0066 \\ -0.0188 \\ -0.0132 \\ -0.0188 \\ -0.0132 \end{bmatrix}$$

$$\begin{bmatrix} \downarrow \\ W_D \\ \downarrow \\ W_A \\ \downarrow \\ W_S \end{bmatrix}$$

$$w' = w - \ell \frac{\partial L}{\partial w}$$

↑  
Learning rate

choose a LR, then update