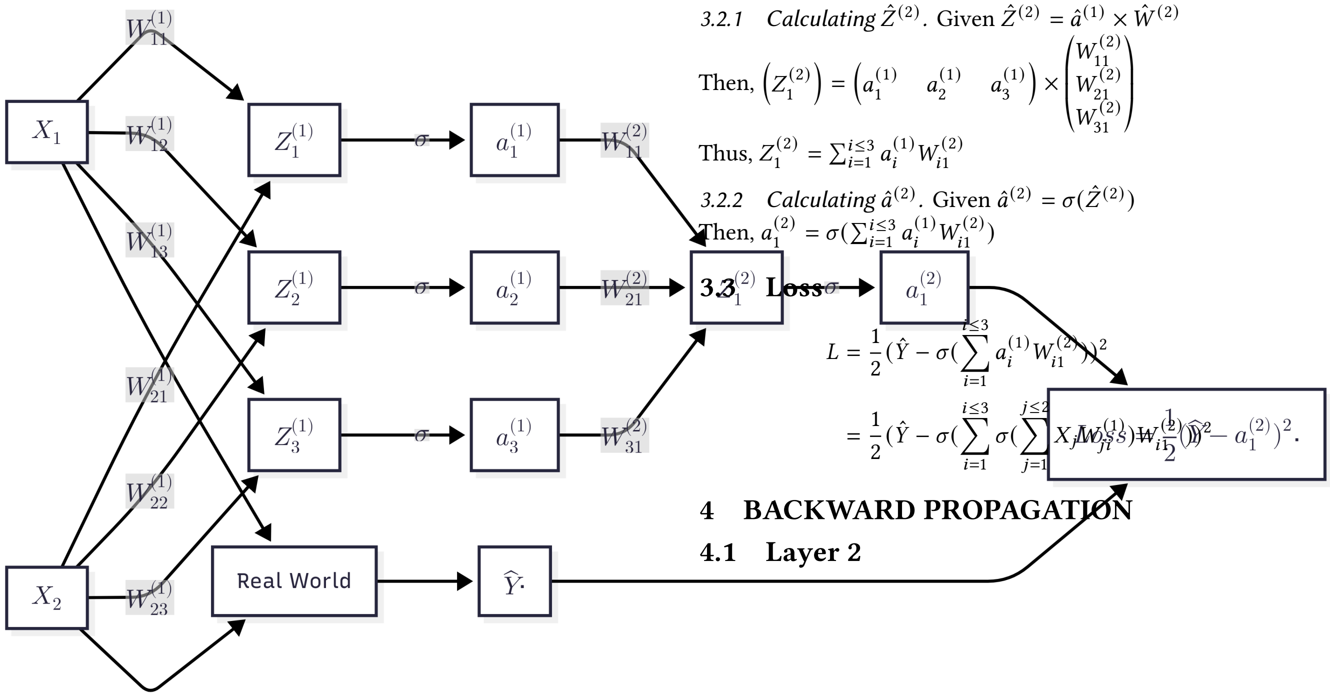
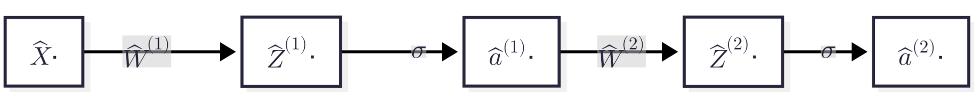
# CogSci 131: Back Propagation


Chanwut (Mick) Kittivorawong
SSID: 3037251249


## 1 NEURAL NETWORK DIAGRAM




## 2 ALL VARIABLES

Below are all the variables in the neural networks.

$$\hat{X} = \begin{pmatrix} X_1 & X_2 \end{pmatrix} \qquad \hat{W}^{(1)} = \begin{pmatrix} W^{(1)}_{11} & W^{(1)}_{12} & W^{(1)}_{13} \\ W^{(1)}_{21} & W^{(1)}_{22} & W^{(1)}_{23} \end{pmatrix} \qquad \hat{Z}^{(1)} = \begin{pmatrix} Z^{(1)}_1 & Z^{(1)}_2 & Z^{(1)}_3 \end{pmatrix} \qquad \hat{a}^{(1)} = \begin{pmatrix} a^{(1)}_1 & a^{(1)}_2 & a^{(1)}_3 \end{pmatrix}$$

$$\hat{W}^{(1)} = \begin{pmatrix} W^{(1)}_{11} \\ W^{(1)}_{21} \\ W^{(1)}_{31} \end{pmatrix} \qquad \hat{Z}^{(2)} = \begin{pmatrix} Z^{(2)}_1 \end{pmatrix} \qquad \hat{a}^{(2)} = \begin{pmatrix} a^{(2)}_1 \end{pmatrix} \qquad L = \tfrac{1}{2}(\hat{Y} - a^{(2)}_1)^2$$

## 3 FORWARD PROPAGATION

### 3.1 Layer 1

#### 3.1.1 *Calculating* $\hat{Z}^{(1)}$. Given $\hat{Z}^{(1)} = \hat{X} \times \hat{W}^{(1)}$

Then, $\begin{pmatrix} Z^{(1)}_1 & Z^{(1)}_2 & Z^{(1)}_3 \end{pmatrix} = \begin{pmatrix} X_1 & X_2 \end{pmatrix} \times \begin{pmatrix} W^{(1)}_{11} & W^{(1)}_{12} & W^{(1)}_{13} \\ W^{(1)}_{21} & W^{(1)}_{22} & W^{(1)}_{23} \end{pmatrix}$

Thus, $Z^{(1)}_i = \sum_{j=1}^{j\le 2} X_j W^{(1)}_{ji}$

#### 3.1.2 *Calculating* $\hat{a}^{(1)}$. Given $\hat{a}^{(1)} = \sigma(\hat{Z}^{(1)})$

Then, $a^{(1)}_i = \sigma(\sum_{j=1}^{j\le 2} X_j W^{(1)}_{ji})$

### 3.2 Layer 2

#### 3.2.1 *Calculating* $\hat{Z}^{(2)}$. Given $\hat{Z}^{(2)} = \hat{a}^{(1)} \times \hat{W}^{(2)}$

Then, $\begin{pmatrix} Z^{(2)}_1 \end{pmatrix} = \begin{pmatrix} a^{(1)}_1 & a^{(1)}_2 & a^{(1)}_3 \end{pmatrix} \times \begin{pmatrix} W^{(2)}_{11} \\ W^{(2)}_{21} \\ W^{(2)}_{31} \end{pmatrix}$

Thus, $Z^{(2)}_1 = \sum_{i=1}^{i\le 3} a^{(1)}_i W^{(2)}_{i1}$

#### 3.2.2 *Calculating* $\hat{a}^{(2)}$. Given $\hat{a}^{(2)} = \sigma(\hat{Z}^{(2)})$

Then, $a^{(2)}_1 = \sigma(\sum_{i=1}^{i\le 3} a^{(1)}_i W^{(2)}_{i1})$

### 3.3 Loss

$$\begin{aligned} L &= \frac{1}{2}(\hat{Y} - \sigma(\sum_{i=1}^{i\le 3} a^{(1)}_i W^{(2)}_{i1}))^2 \\ &= \frac{1}{2}(\hat{Y} - \sigma(\sum_{i=1}^{i\le 3} \sigma(\sum_{j=1}^{j\le 2} X_j W^{(1)}_{ji}) W^{(2)}_{i1}))^2 \end{aligned}$$

## 4 BACKWARD PROPAGATION

### 4.1 Layer 2