

# RECURRENT NEURAL NETWORKS

{Grady Kurpasi}

{SSIE 616}

{Prof H. Lewis}



# Multilayer Perceptrons

$$x, y$$

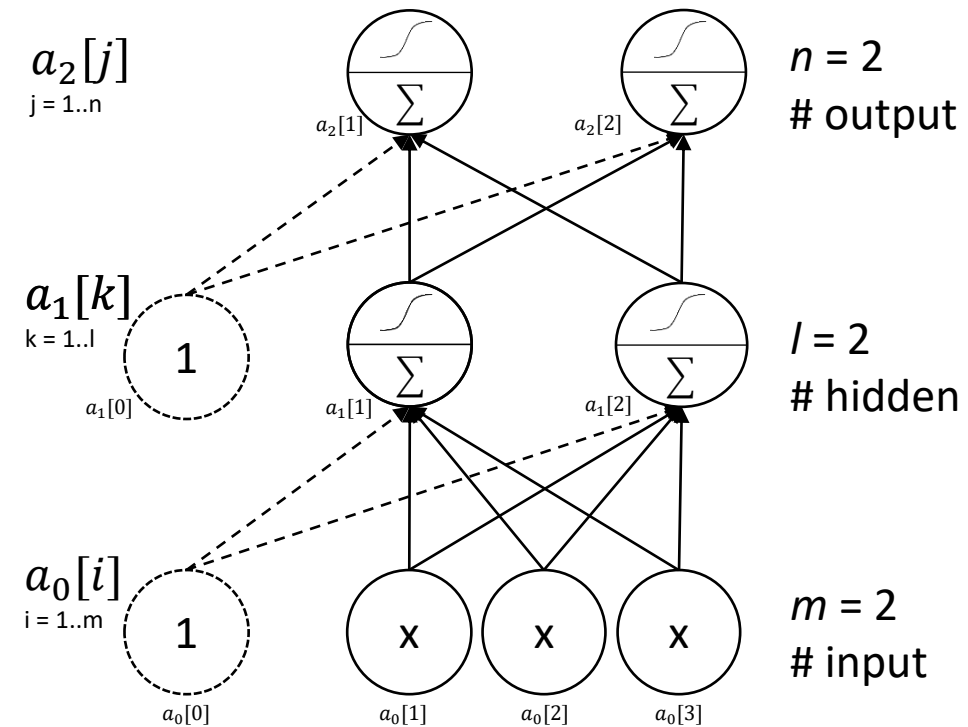
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500
0.1000	-5.0000	3.0000	0.1221	0.0964
6.0000	-5.5420	4.8970	0.1061	0.0702
4.0000	8.0000	9.0000	0.0996	0.0641
12.0000	-2.0000	0.0063	0.1110	0.0732
6.0000	-5.5000	4.8970	0.1060	0.0701

$$w_1[i, k]$$

i	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$$w_2[k, j]$$

k	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1



3-2-2 Multilayer Perceptron

# Multilayer Perceptrons

$$x, y$$

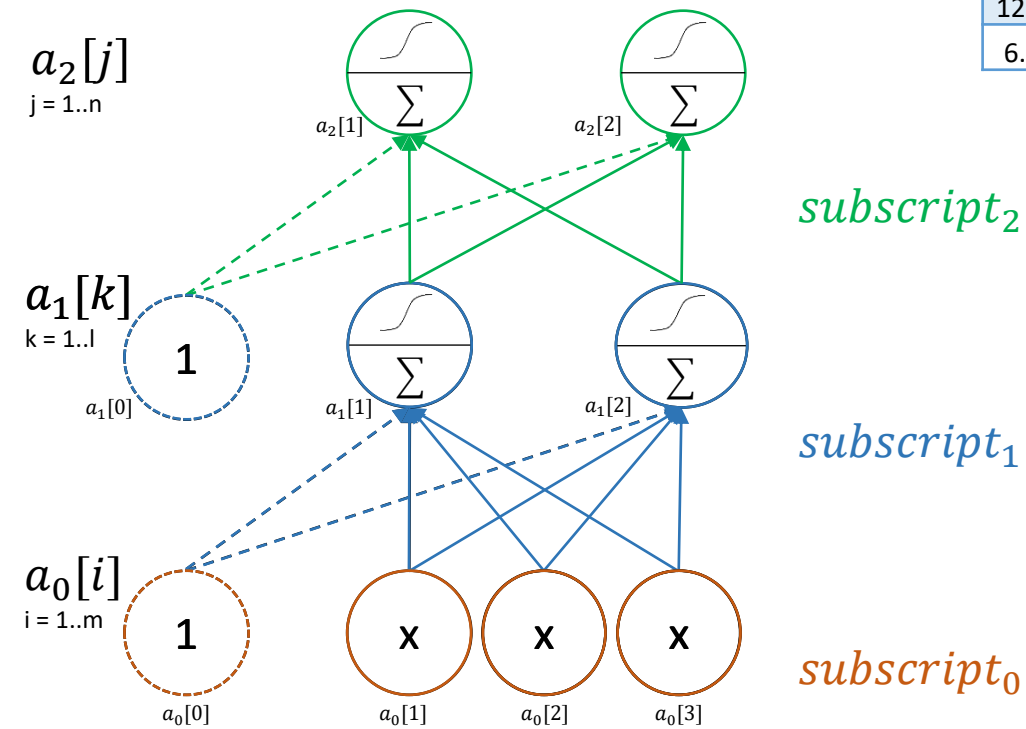
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500
0.1000	-5.0000	3.0000	0.1221	0.0964
6.0000	-5.5420	4.8970	0.1061	0.0702
4.0000	8.0000	9.0000	0.0996	0.0641
12.0000	-2.0000	0.0063	0.1110	0.0732
6.0000	-5.5000	4.8970	0.1060	0.0701

$$w_1[i, k]$$

i	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$$w_2[k, j]$$

k	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1



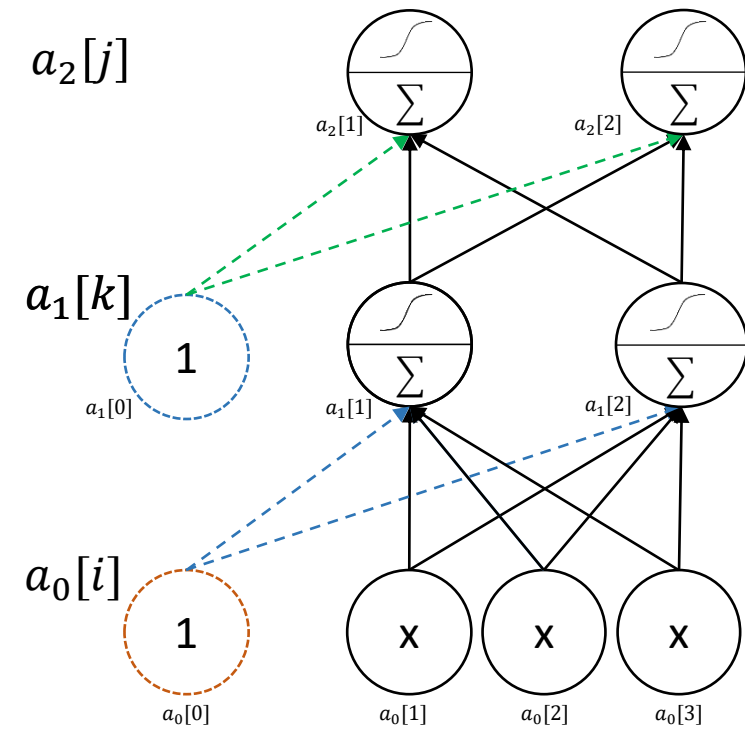
3-2-2 Multilayer Perceptron

# Multilayer Perceptrons

$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500
0.1000	-5.0000	3.0000	0.1221	0.0964
6.0000	-5.5420	4.8970	0.1061	0.0702
4.0000	8.0000	9.0000	0.0996	0.0641
12.0000	-2.0000	0.0063	0.1110	0.0732
6.0000	-5.5000	4.8970	0.1060	0.0701

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

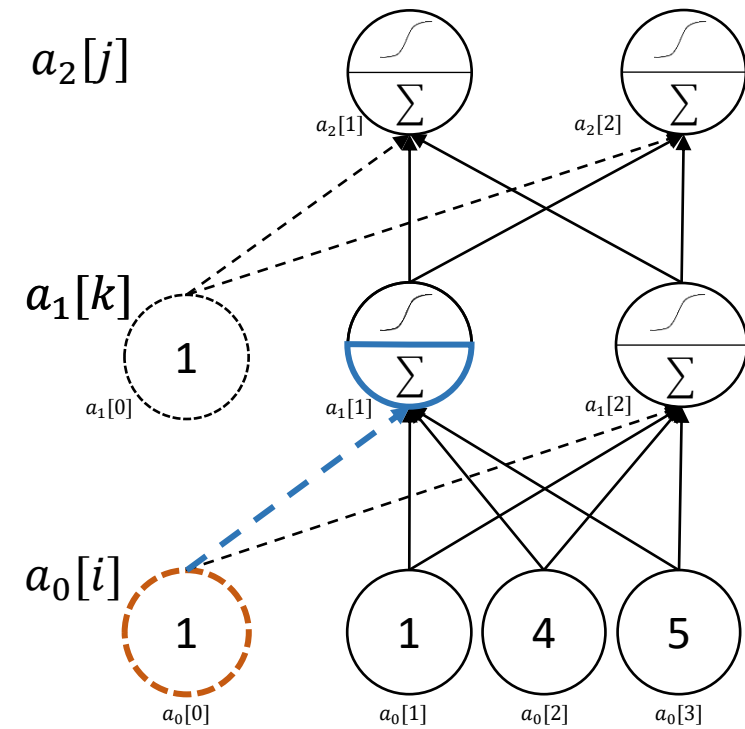


# Forward Propagation

$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	<b>0.5</b>	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1



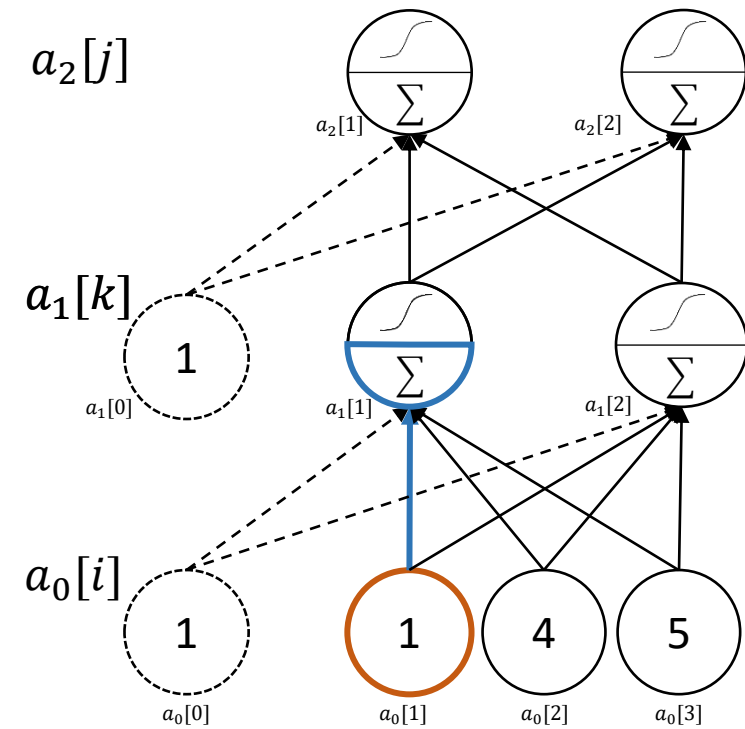
$$s_1[k] = \sum_{i=0}^m w_1[i, k] * a_0[i] \quad \text{for } k = 1..l$$
$$s_1[1] = w_1[0, 1] * a_0[0] = .05 * 1$$

# Forward Propagation

$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	<b>0.1</b>	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1



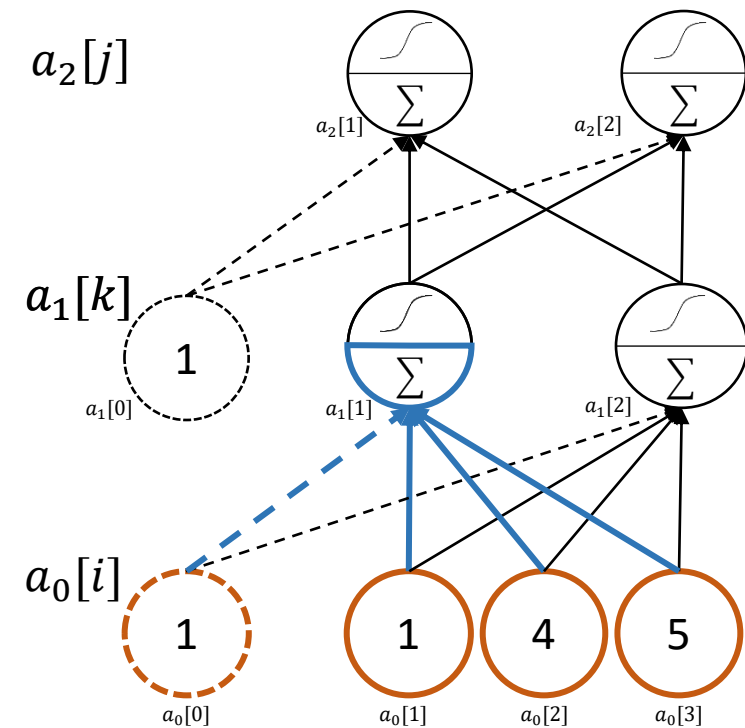
$$s_1[k] = \sum_{i=0}^3 w_1[i, k] * a_0[i] \quad \text{for } k = 1, 2$$
$$s_1[1] = w_1[0, 1] * a_0[0] = .05 * 1$$
$$w_1[1, 1] * a_0[1] = .1 * 1$$

# Forward Propagation

$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	<b>0.5</b>	0.5
1	<b>0.1</b>	0.2
2	<b>0.3</b>	0.4
3	<b>0.5</b>	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1



$$s_1[k] = \sum_{i=0}^3 w_1[i, k] * a_0[i] \quad \text{for } k = 1, 2$$
$$s_1[1] = w_1[0, 1] * a_0[0] + w_1[1, 1] * a_0[1] + w_1[2, 1] * a_0[2] + w_1[3, 1] * a_0[3]$$
$$= 0.05 * 1 + 0.1 * 1 + 0.3 * 4 + 0.5 * 5$$
$$= 4.3$$

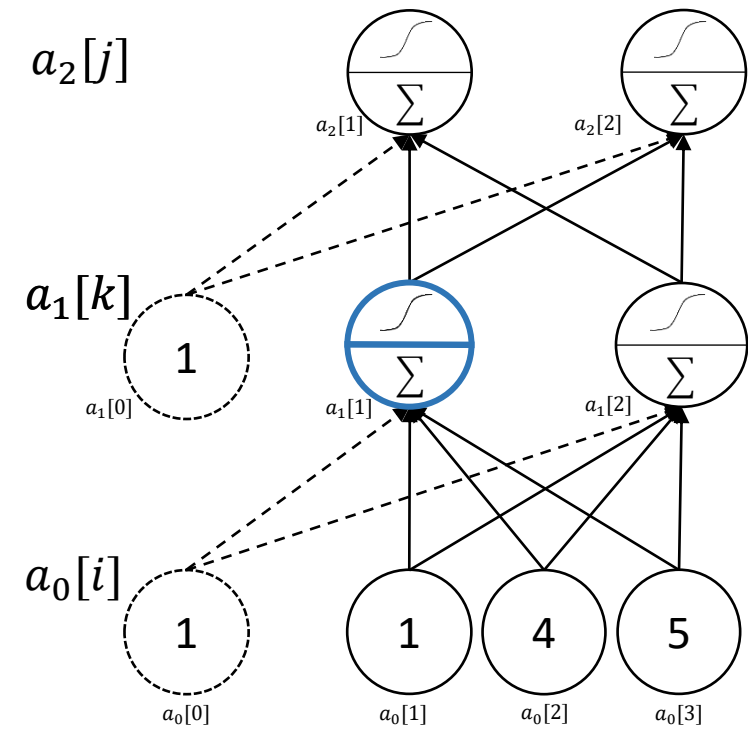
# Forward Propagation

$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$a_1[k] = \frac{1}{1 + e^{-s_1[k]}} \text{ for } k = 1, 2$$
$$a_1[1] = \frac{1}{1 + e^{-4.3}} = .9866$$



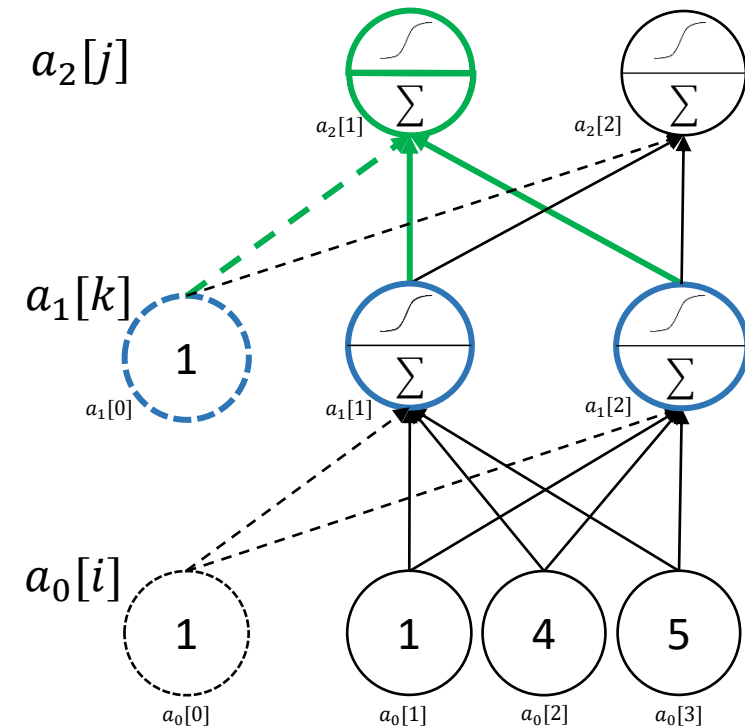


# Forward Propagation

$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1



$$s_2[j] = \sum_{k=0}^l w_2[k, j] * a_1[k] \quad \text{for } j = 1..m$$

$$a_2[j] = \frac{1}{1 + e^{-s_2[j]}} \quad \text{for } j = 1, 2$$

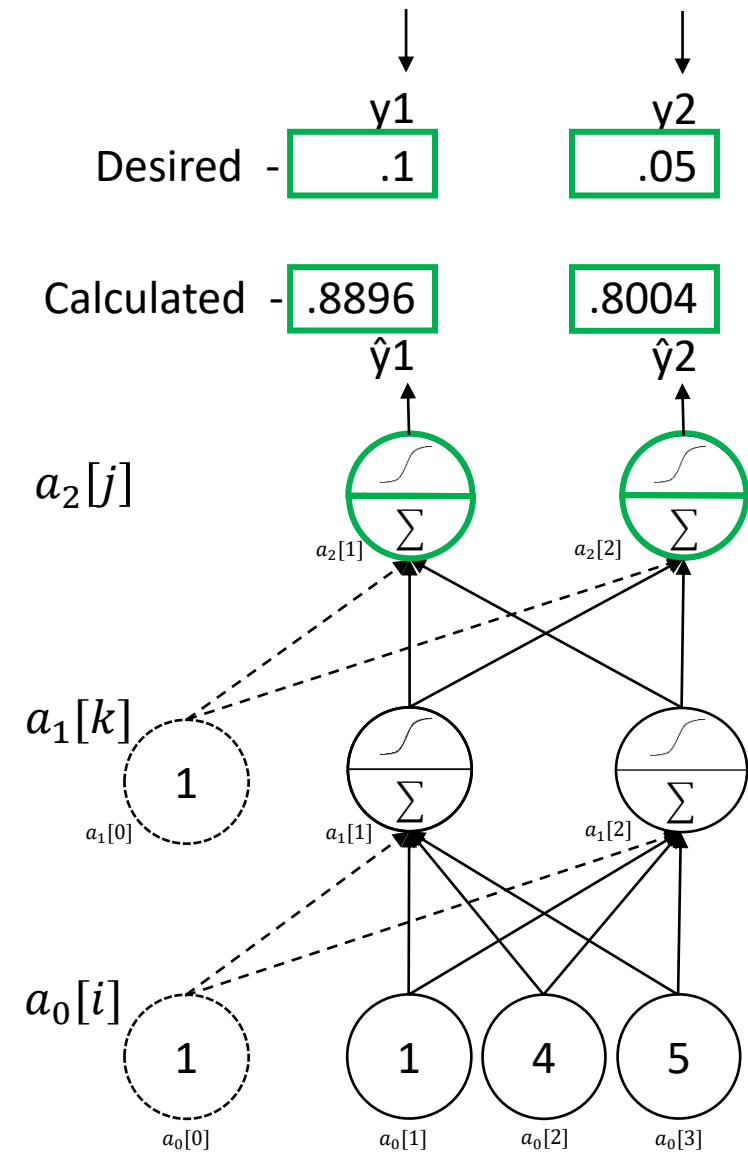
$$s_2[1] = w_2[0, 2] * a_1[0] = .5 * 1$$

$$w_2[1, 2] * a_1[1] = .7 * .9866$$

$$w_2[2, 2] * a_1[2] = .9 * .9950$$

$$= 2.0862$$

$$a_2[1] = \frac{1}{1 + e^{-2.0862}} = .8896 = \hat{y}$$



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i,1]$	$w[i,2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k,1]$	$w[k,2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned}
 E &= \frac{1}{2} \sum_{j=1}^n (a_2[j] - y_r[j])^2 \\
 &= \frac{1}{2} ( (.8896 - .1)^2 + (.8004 - .5)^2 ) \\
 &= .5932
 \end{aligned}$$

# Backpropagation



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

Partial Derivative of Error with respect to  $Weight_2[0,1]$

Functional Dependencies

$$a_1[k] * w_2[0, 1] \rightarrow f$$

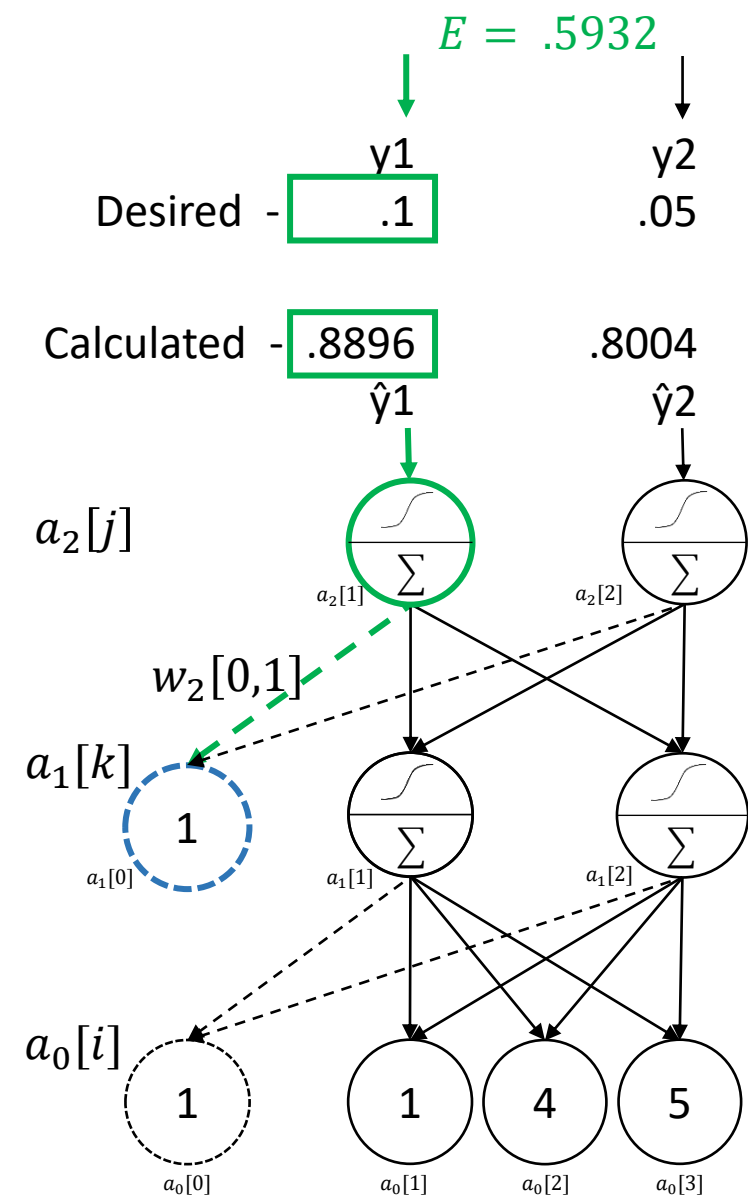
$$\sum f + .6906 + .8955 \rightarrow g$$

$$\sigma(g) \rightarrow h$$

$$(h - y_1)^2 \rightarrow E$$

Chain Rule

$$E(h(g(f(w_{0,1}))))$$

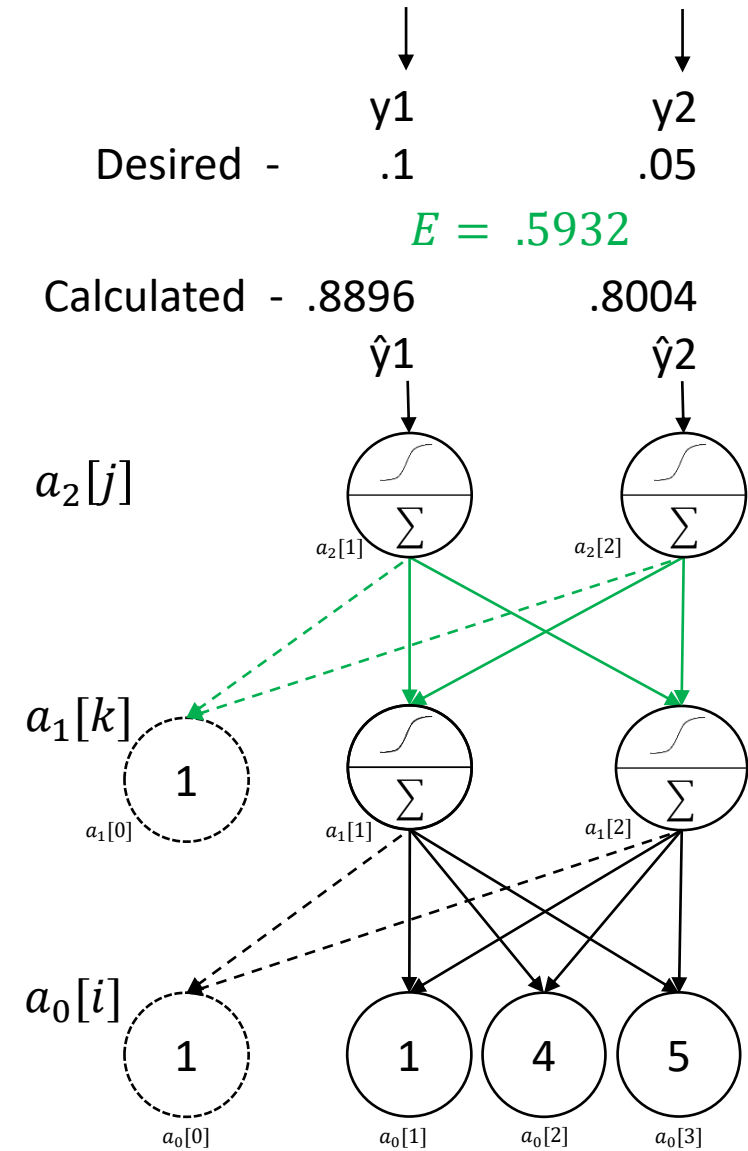


# Backpropagation

Desired -  $y_1 = .1$   $y_2 = .05$

$E = .5932$

Calculated -  $\hat{y}_1 = .8896$   $\hat{y}_2 = .8004$



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
i	$w[i,1]$	$w[i,2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
k	$w[k,1]$	$w[k,2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned} \frac{\partial E}{\partial w_2[k, j]} &= \frac{\partial E}{\partial a_2[j]} * \frac{\partial a_2[j]}{\partial s_2[j]} * \frac{\partial s_2[j]}{\partial w_2[k, j]} \\ &= \{\varepsilon[j]\} * \{a_2[j](1 - a_2[j])\} * \{a_1[k]\} \end{aligned}$$

# Backpropagation

Desired -  $y_1$  .1

$y_2$   
.05

Calculated - .8896

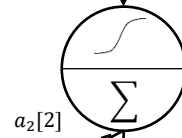
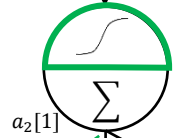
.8004

$\hat{y}_1$

$\hat{y}_2$

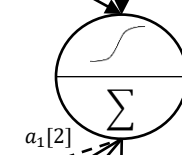
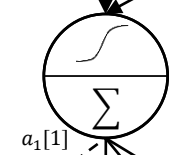
$\epsilon[1]$

$a_2[j]$

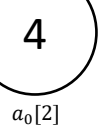
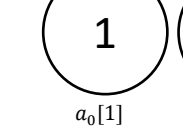


$w_2[0,1]$

$a_1[k]$



$a_0[i]$



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

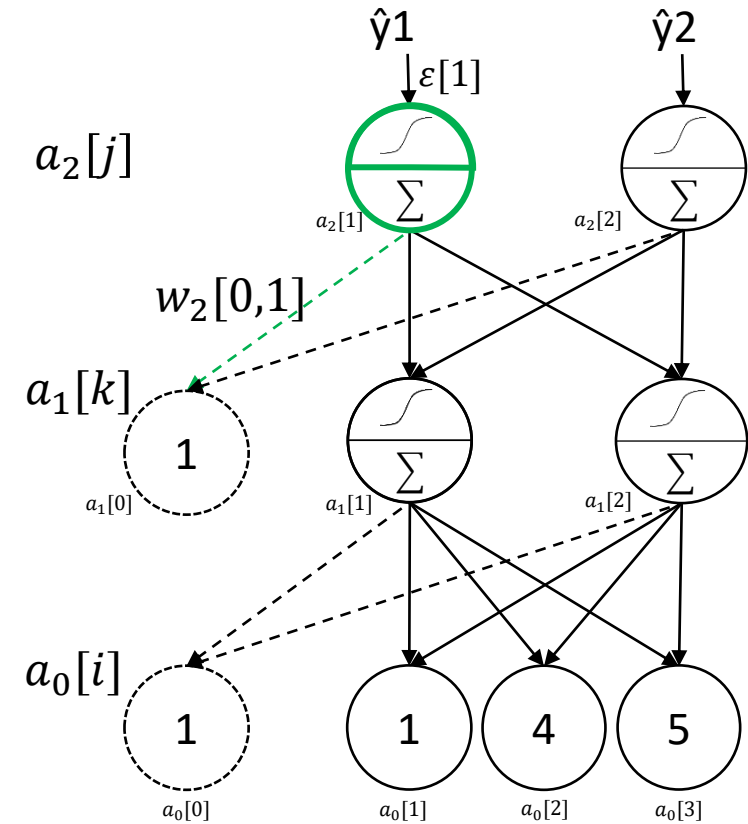
$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned}
 \frac{\partial E}{\partial w_2[0,1]} &= \frac{\partial E}{\partial a_2[1]} * \frac{\partial a_2[1]}{\partial s_2[1]} * \frac{\partial s_2[1]}{\partial w_2[0,1]} \\
 &= \{\epsilon[1]\} * \{a_2[1](1 - a_2[1])\} * \{a_1[0]\} \\
 &= .8896 - .1
 \end{aligned}$$

# Backpropagation

Desired -  $y_1 = .1$   $y_2 = .05$

Calculated -  $\hat{y}_1 = .8896$   $\hat{y}_2 = .8004$



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
i	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

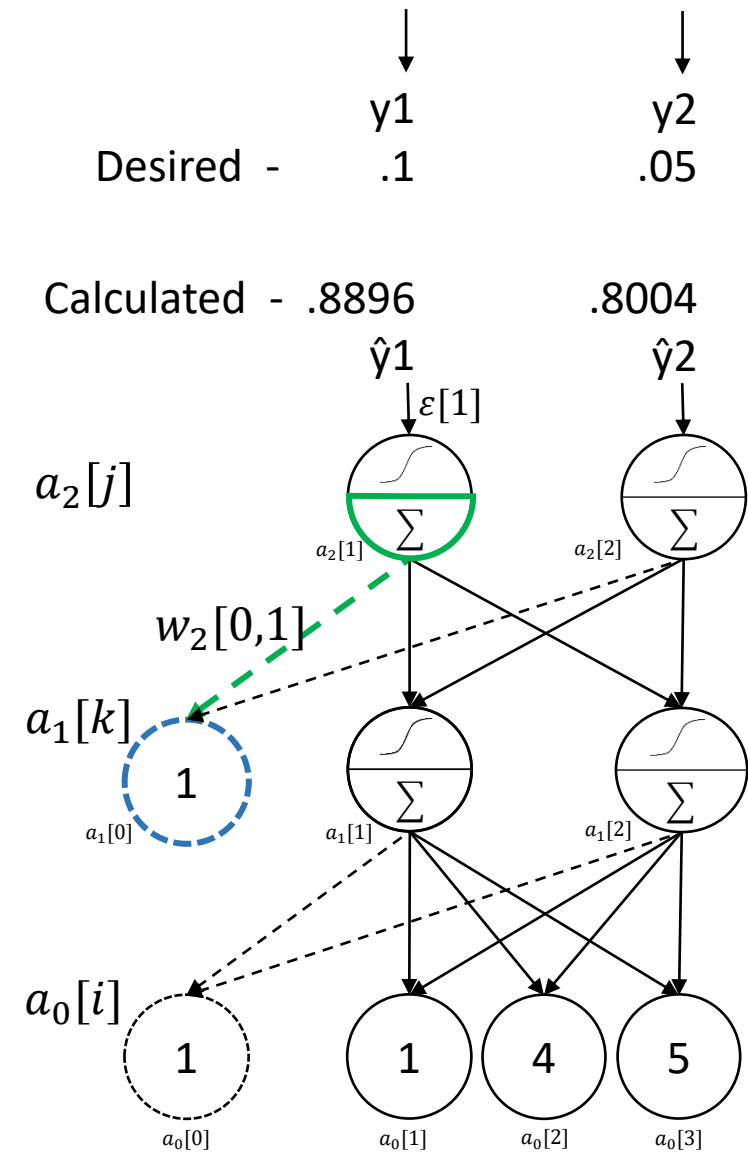
$w_2[k, j]$		
k	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned}
 \frac{\partial E}{\partial w_2[0,1]} &= \frac{\partial E}{\partial a_2[1]} * \frac{\partial a_2[1]}{\partial s_2[1]} * \frac{\partial s_2[1]}{\partial w_2[0,1]} \\
 &= \{\epsilon[1]\} * \{a_2[1](1 - a_2[1])\} * \{a_1[0]\} \\
 &= .8896 - .1 * .8896(1 - .8896)
 \end{aligned}$$

# Backpropagation

Desired -  $y_1 = .1$   $y_2 = .05$

Calculated -  $\hat{y}_1 = .8896$   $\hat{y}_2 = .8004$



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

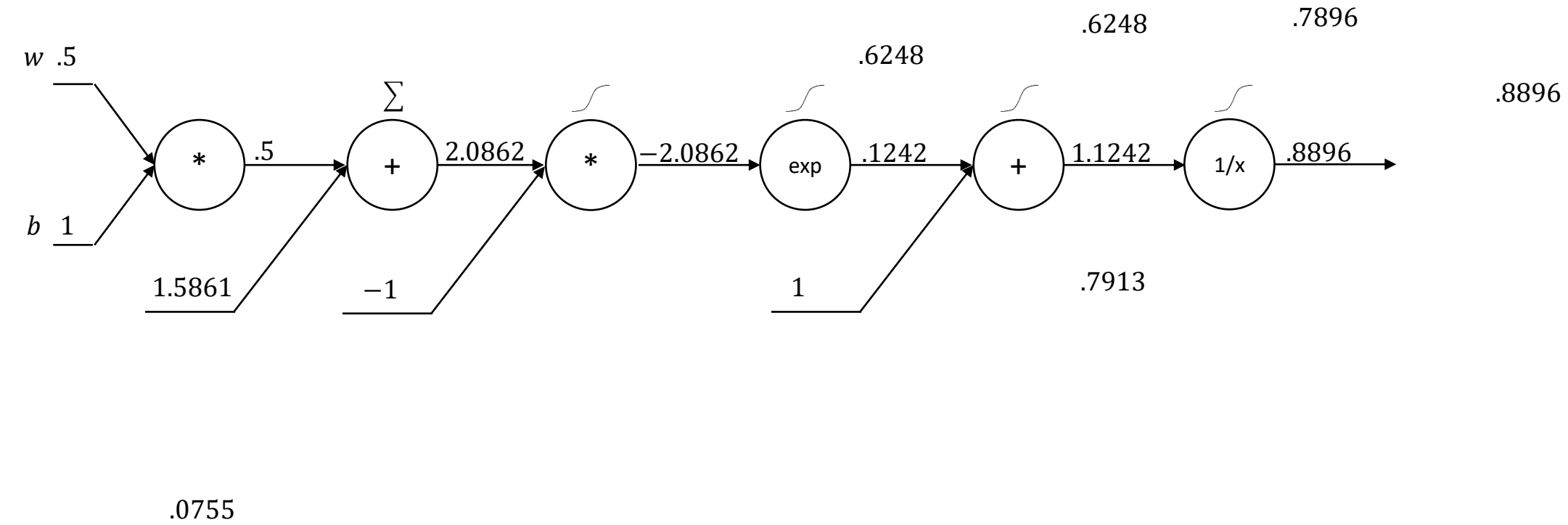
$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned}
 \frac{\partial E}{\partial w_2[0,1]} &= \frac{\partial E}{\partial a_2[1]} * \frac{\partial a_2[1]}{\partial s_2[1]} * \frac{\partial s_2[1]}{\partial w_2[0,1]} \\
 &= \{\varepsilon[1]\} * \{a_2[1](1 - a_2[1])\} * \{a_1[0]\} \\
 &= .8896 - .1 * .8896(1 - .8896) * 1 \\
 &= .0775
 \end{aligned}$$

$c_2[k, j]$		
$k$	$c[k, 1]$	$c[k, 2]$
0	.0775	0
1	0	0
2	0	0

# Backpropagation



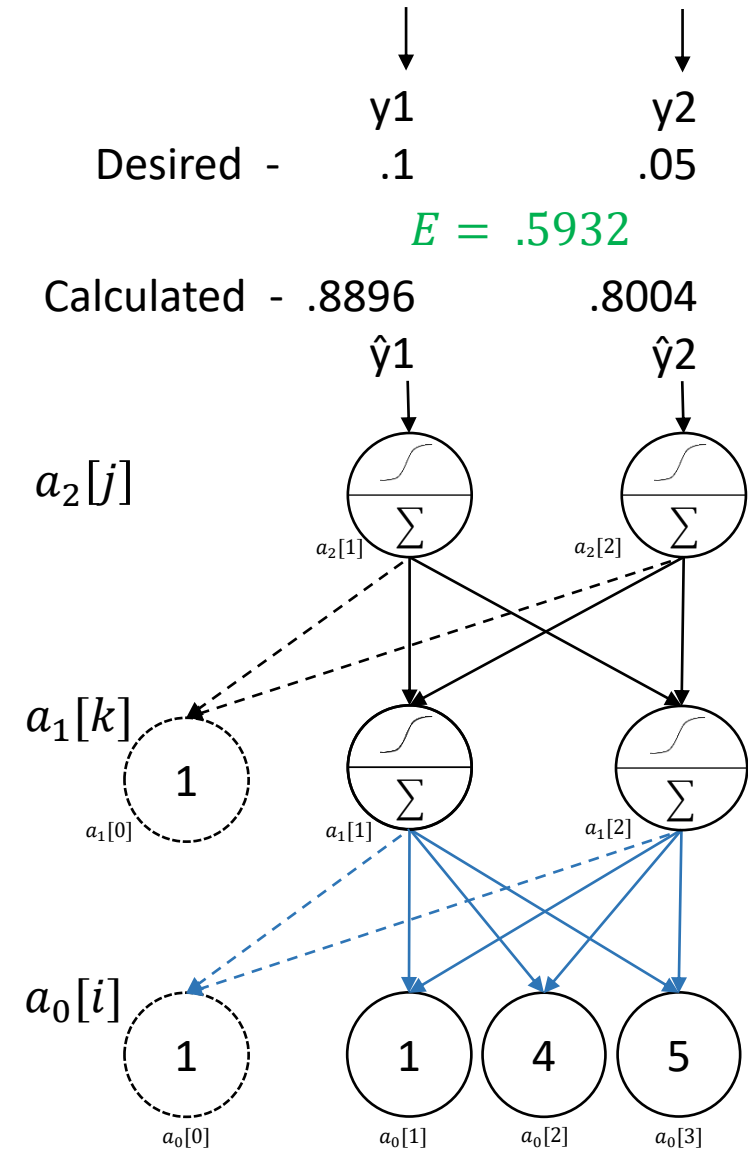


# Backpropagation

Desired -  $y_1 = .1$   $y_2 = .05$

$E = .5932$

Calculated -  $\hat{y}_1 = .8896$   $\hat{y}_2 = .8004$



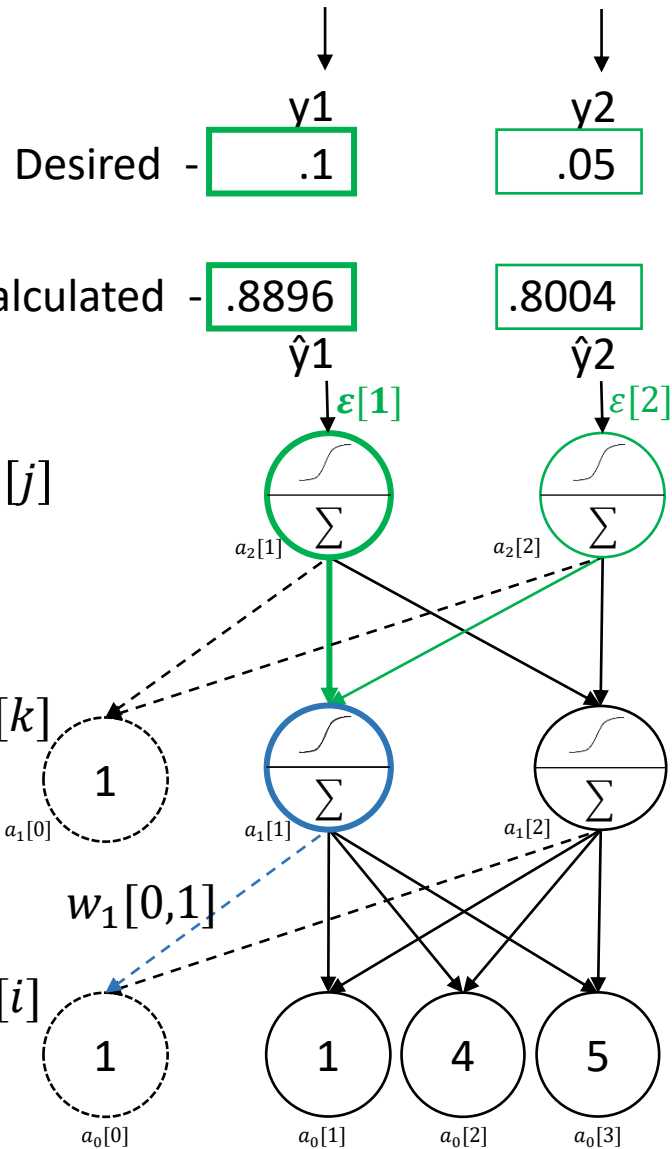
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$i$	$w[i,1]$	$w[i,2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$k$	$w[k,1]$	$w[k,2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned} \frac{\partial E}{\partial w_1[i, k]} &= \frac{\partial E}{\partial a_1[k]} * \frac{\partial a_1[k]}{\partial s_1[k]} * \frac{\partial s_1[k]}{\partial w_1[i, k]} \\ &= \frac{\partial E}{\partial a_1[k]} * \{a_1[j](1 - a_1[j])\} * \{a_0[i]\} \end{aligned}$$

# Backpropagation



$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

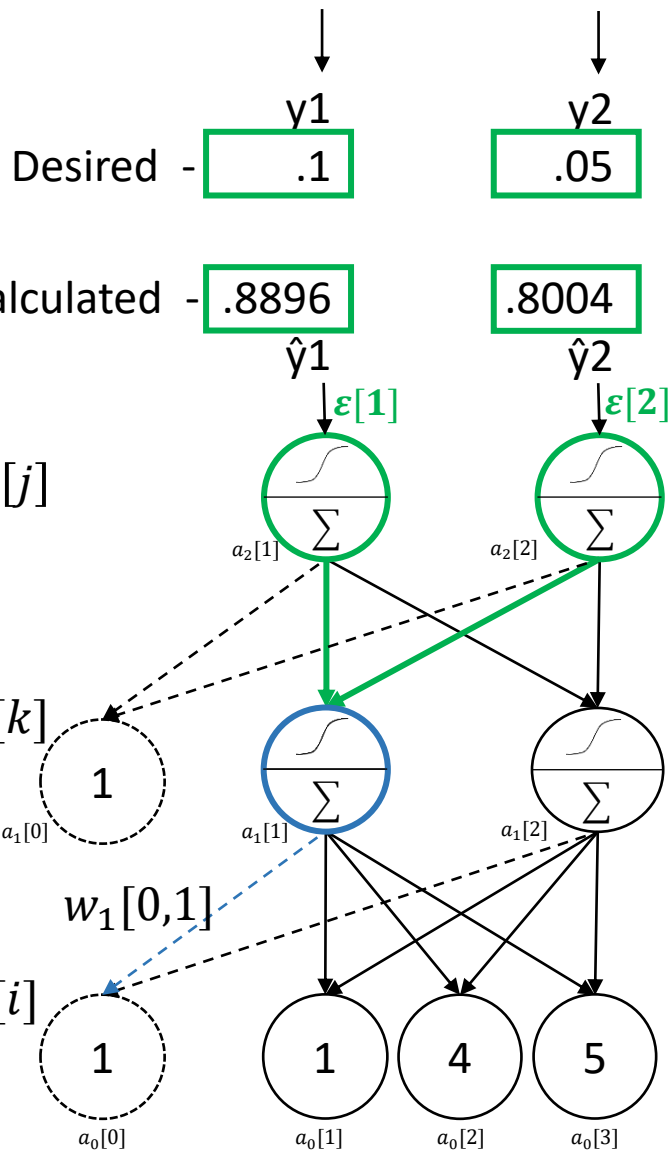
$i$	$w[i,1]$	$w[i,2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$k$	$w[k,1]$	$w[k,2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned}
 \frac{\partial E}{\partial w_1[i, k]} &= \frac{\partial E}{\partial a_1[k]} * \frac{\partial a_1[k]}{\partial s_1[k]} * \frac{\partial s_1[k]}{\partial w_1[i, k]} \\
 &= \frac{\partial E}{\partial a_1[k]} * \{a_1[k](1 - a_1[k])\} * \{a_0[i]\} \\
 \frac{\partial E}{\partial a_1[k]} &= \sum_{j=1}^n \frac{\partial E}{\partial a_2[j]} * \frac{\partial a_2[j]}{\partial s_2[j]} * \frac{\partial s_2[j]}{\partial a_1[k, j]} \\
 &= \sum_{j=1}^n \{\varepsilon[j]\} * \{a_2[j](1 - a_2[j])\} * \{w_2[k, j]\}
 \end{aligned}$$



# Backpropagation



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
$i$	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

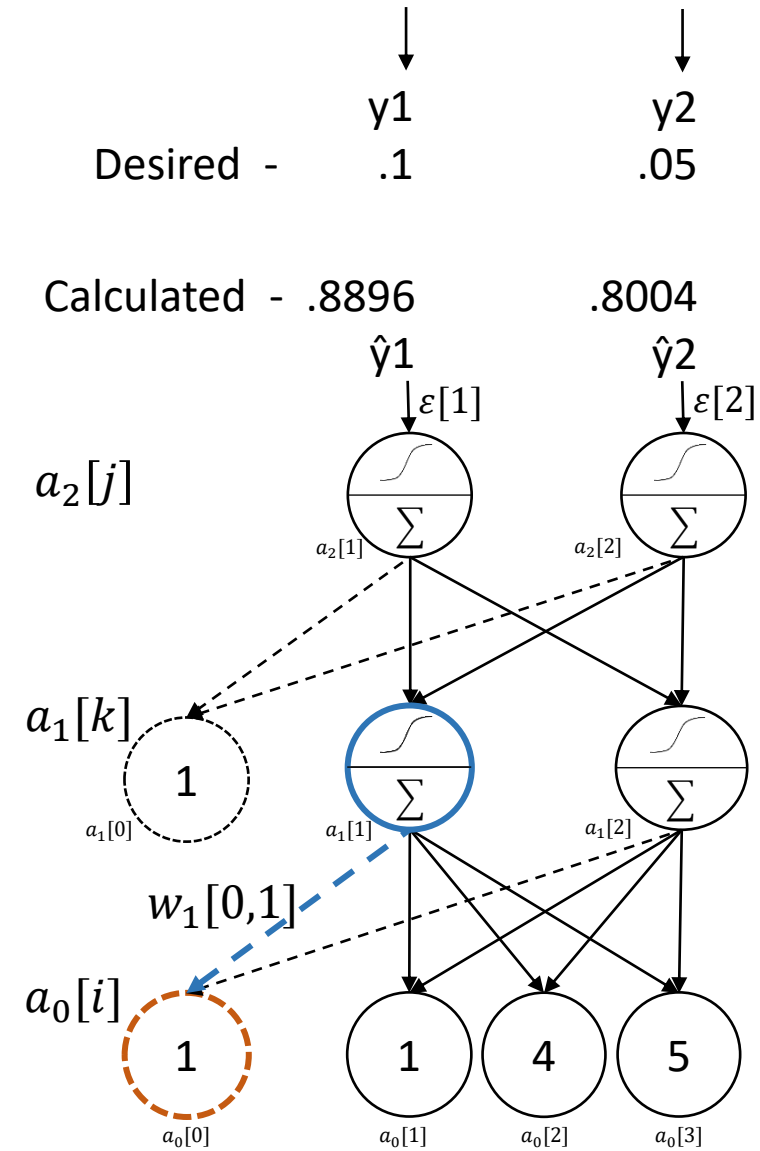
$w_2[k, j]$		
$k$	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned}
 \frac{\partial E}{\partial w_1[0,1]} &= \frac{\partial E}{\partial a_1[1]} * \frac{\partial a_1[1]}{\partial s_1[1]} * \frac{\partial s_1[1]}{\partial w_1[0,1]} \\
 &= \frac{\partial E}{\partial a_1[1]} * \{a_1[1](1 - a_1[1])\} * \{a_0[0]\} \\
 \frac{\partial E}{\partial a_1[k]} &= \sum_{j=1}^n \frac{\partial E}{\partial a_2[j]} * \frac{\partial a_2[j]}{\partial s_2[j]} * \frac{\partial s_2[j]}{\partial a_1[k,j]} \\
 &= \sum_{j=1}^n \{\epsilon[j]\} * \{a_2[j](1 - a_2[j])\} * \{w_2[k, j]\} \\
 &= .8896 - .1 * .8896(1 - .8896) * .7 \\
 &\quad + .8004 - .05 * .8004(1 - .8004) * .8 \\
 &= .1502
 \end{aligned}$$

# Backpropagation

Desired -  $y_1 = .1$   $y_2 = .05$

Calculated -  $\hat{y}_1 = .8896$   $\hat{y}_2 = .8004$



$x, y$				
$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
1.0000	4.0000	5.0000	0.1000	0.0500

$w_1[i, k]$		
i	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$		
k	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$$\begin{aligned}
 \frac{\partial E}{\partial w_1[0,1]} &= \frac{\partial E}{\partial a_1[1]} * \frac{\partial a_1[1]}{\partial s_1[1]} * \frac{\partial s_1[1]}{\partial w_1[0,1]} \\
 &= \frac{\partial E}{\partial a_1[1]} * \{a_1[1](1 - a_1[1])\} * \{a_0[0]\} \\
 &= .1502 * .9866(1 - .9866) * 1 \\
 &= .0020
 \end{aligned}$$

$c_1[i, k]$		
i	$w[i, 1]$	$w[i, 2]$
0	.0020	0
1	0	0
2	0	0
3	0	0

# Backpropagation

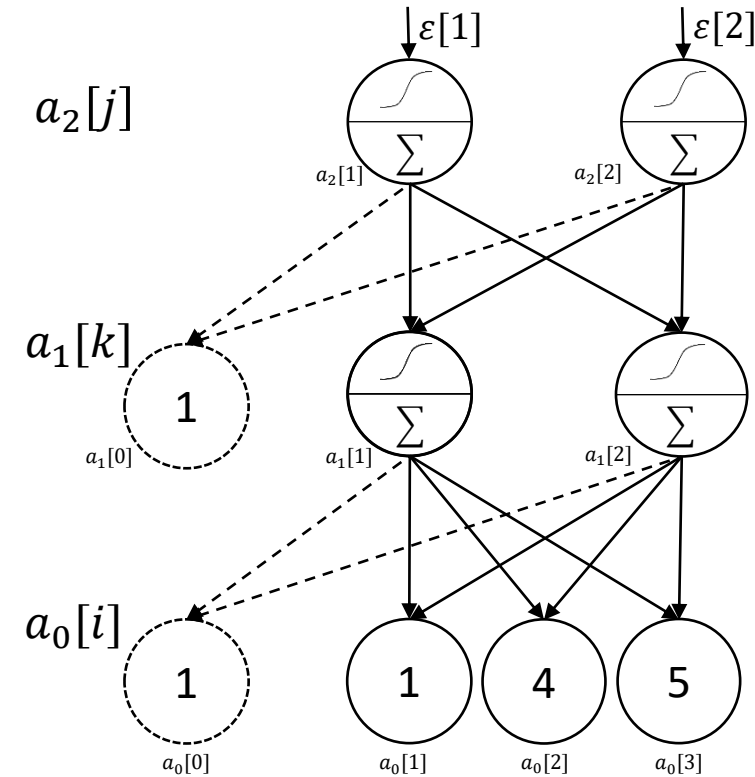


Desired -  $y_1$

$y_2$

Calculated -  $\hat{y}_1$

$\hat{y}_2$



	$x, y$				
	$x_1$	$x_2$	$x_3$	$y_1$	$y_2$
$r=1$	1.0000	4.0000	5.0000	0.1000	0.0500
$\vdots$	0.1000	-5.0000	3.0000	0.1221	0.0964
	6.0000	-5.5420	4.8970	0.1061	0.0702
$r=4$	4.0000	8.0000	9.0000	0.0996	0.0641
	12.0000	-2.0000	0.0063	0.1110	0.0732
	6.0000	-5.5000	4.8970	0.1060	0.0701

Learning rate  $\alpha = .01$

$$w^* = w - \alpha * c$$

$w_1[i, k]$

i	$w[i, 1]$	$w[i, 2]$
0	0.5	0.5
1	0.1	0.2
2	0.3	0.4
3	0.5	0.6

$w_2[k, j]$

k	$w[k, 1]$	$w[k, 2]$
0	0.5	0.5
1	0.7	0.8
2	0.9	0.1

$c_1[i, k]$

i	$w[i, 1]$	$w[i, 2]$
0	.0604	.0332
1	.1134	.0427
2	-.2929	-.1658
3	.2193	.1129

$c_2[k, j]$

k	$c[k, 1]$	$c[k, 2]$
0	.3447	.4816
1	.2929	.4176
2	.2930	.4191

$w_1^*[i, k]$

i	$w[i, 1]$	$w[i, 2]$
0	.4994	.4997
1	.0989	.1996
2	.3029	.4017
3	.4978	.5989

$w_2^*[k, j]$

k	$w[k, 1]$	$w[k, 2]$
0	.4966	.0039
1	.4971	.0038
2	.0021	.0004

