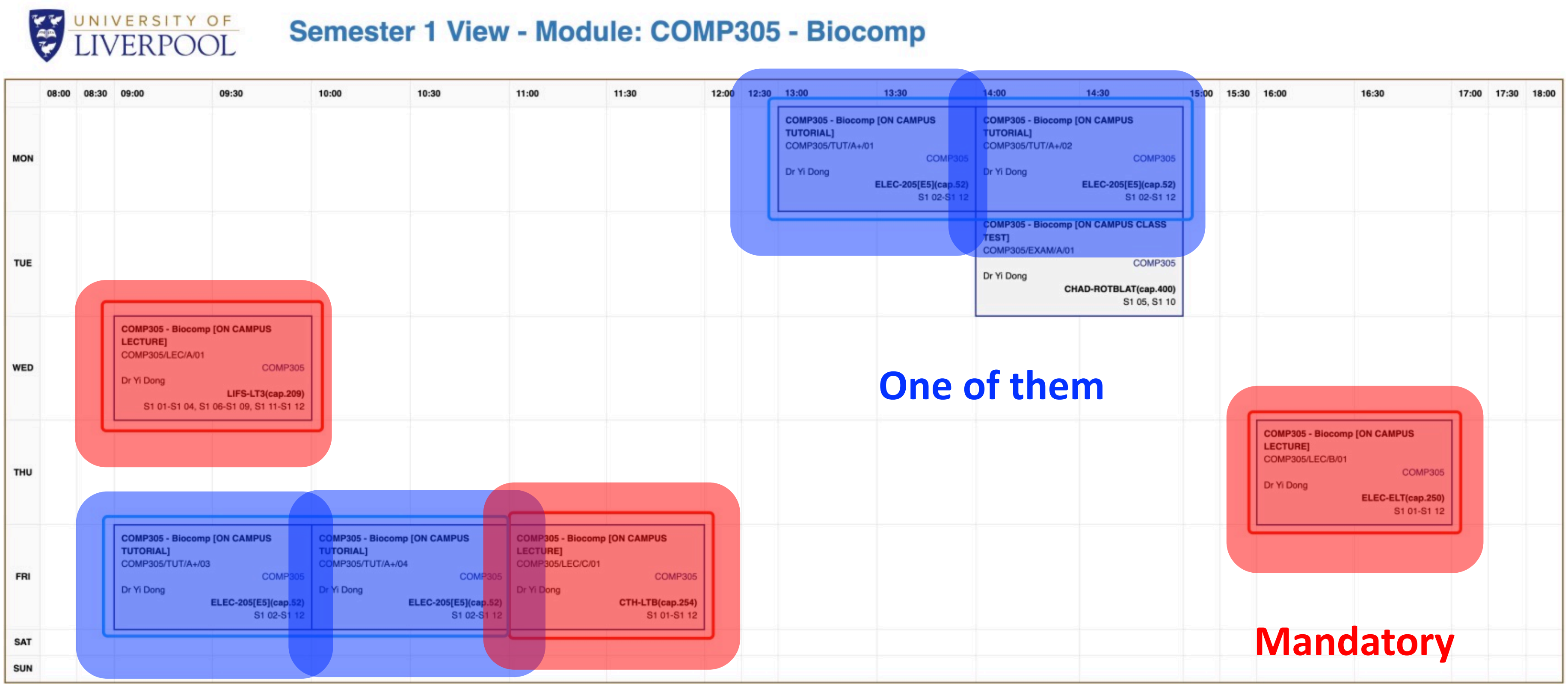
Comp305

Biocomputation

Lecturer: Yi Dong

Comp305 Module Timetable



There will be **26-30** lectures, thee per week. The lecture slides will appear on Canvas. Please use Canvas to access the lecture information. There will be **9** tutorials, one per week.

Lecture/Tutorial Rules

Questions are welcome as soon as they arise, because

1. Questions give feedback to the lecturer;

2. Questions help your understanding;

3. Your questions help your classmates, who might experience difficulties with formulating the same problems/doubts in the form of a question.

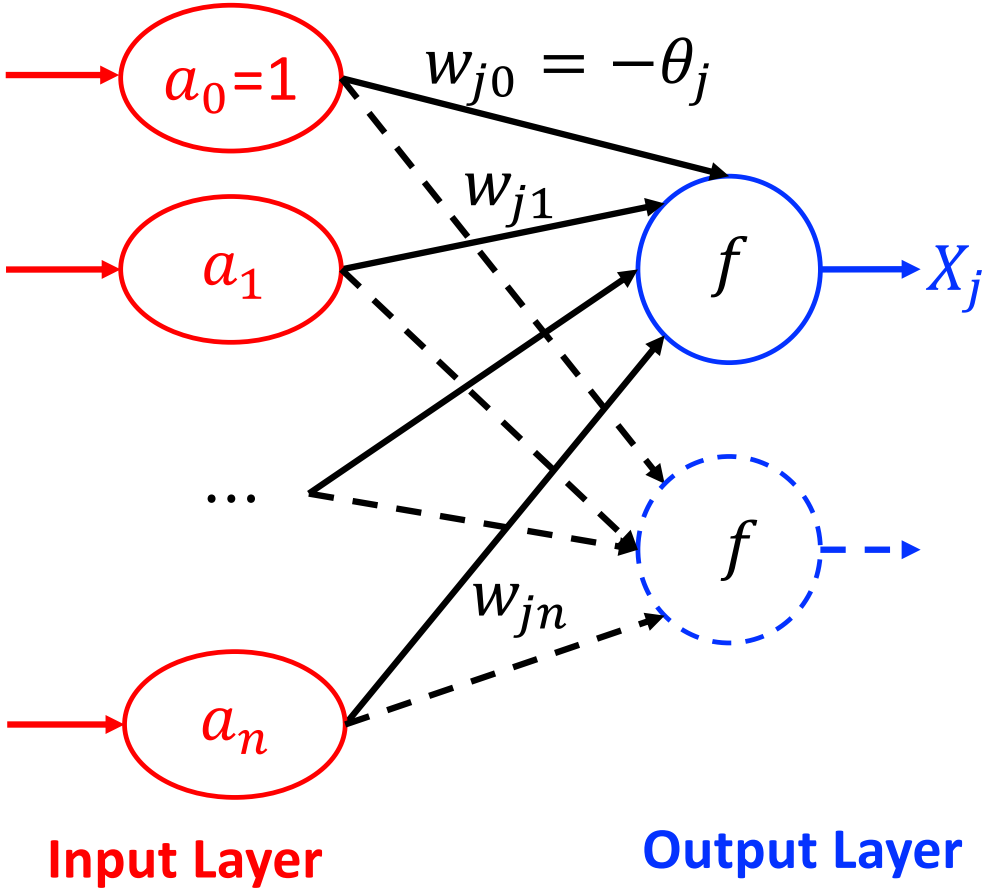
Comp305 Part I.

Artificial Neural Networks

Topic 6.

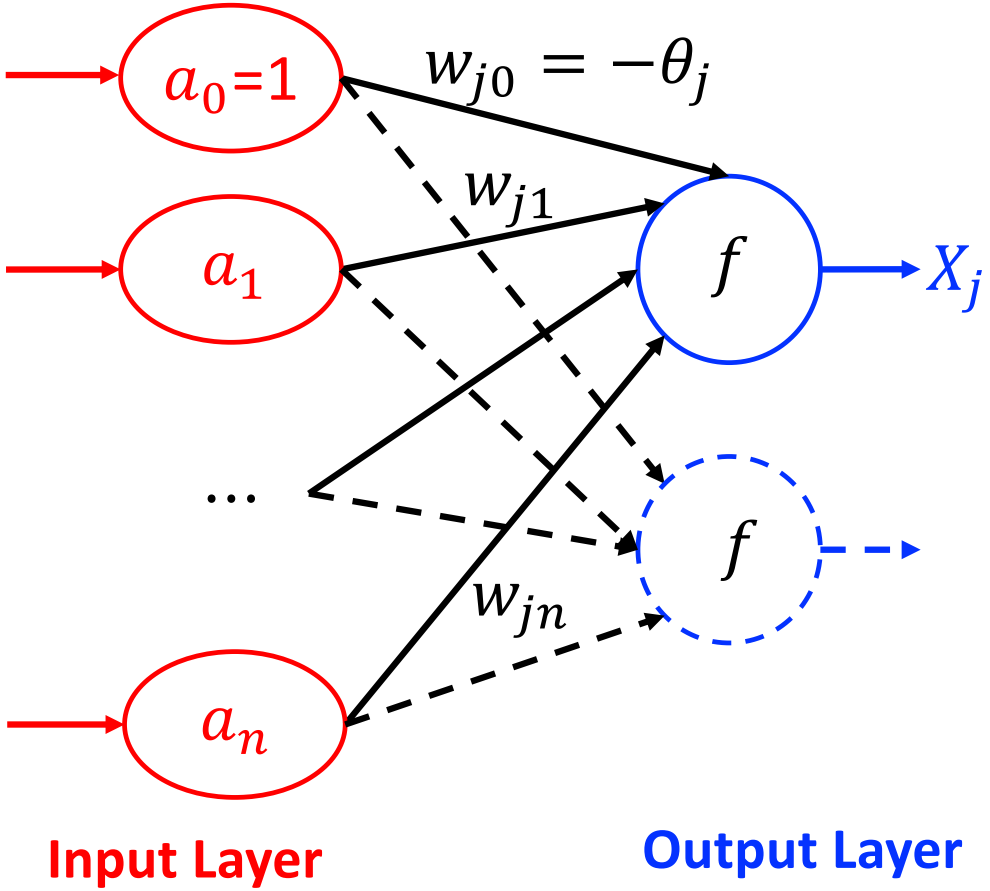
Perceptron

Perceptron (1958)



Weights w_{ji} of connections between two layers are changed according to perceptron learning rule, so the network is more likely to produce the desired output in response to certain inputs.

Weight Update



1. Compute "error" of very connection for every output neuron:

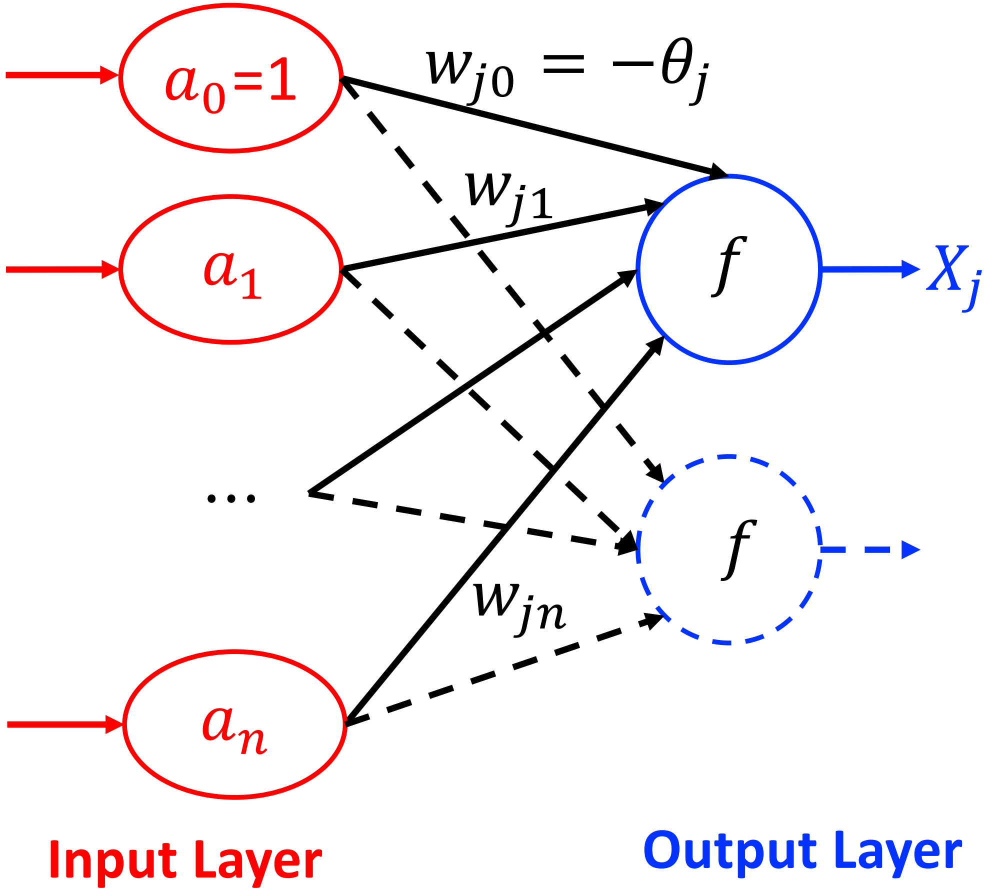
$$e_j^k = \left(t_j^k - X_j^k\right)$$

where

t_j^k: the target value for the j-th output neuron for the k-th input pattern in the data set,

X_j^k: the instant value for the j-th output neuron for the k-th input pattern.

Weight Update



2. Update the weights:

$$w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k$$

where

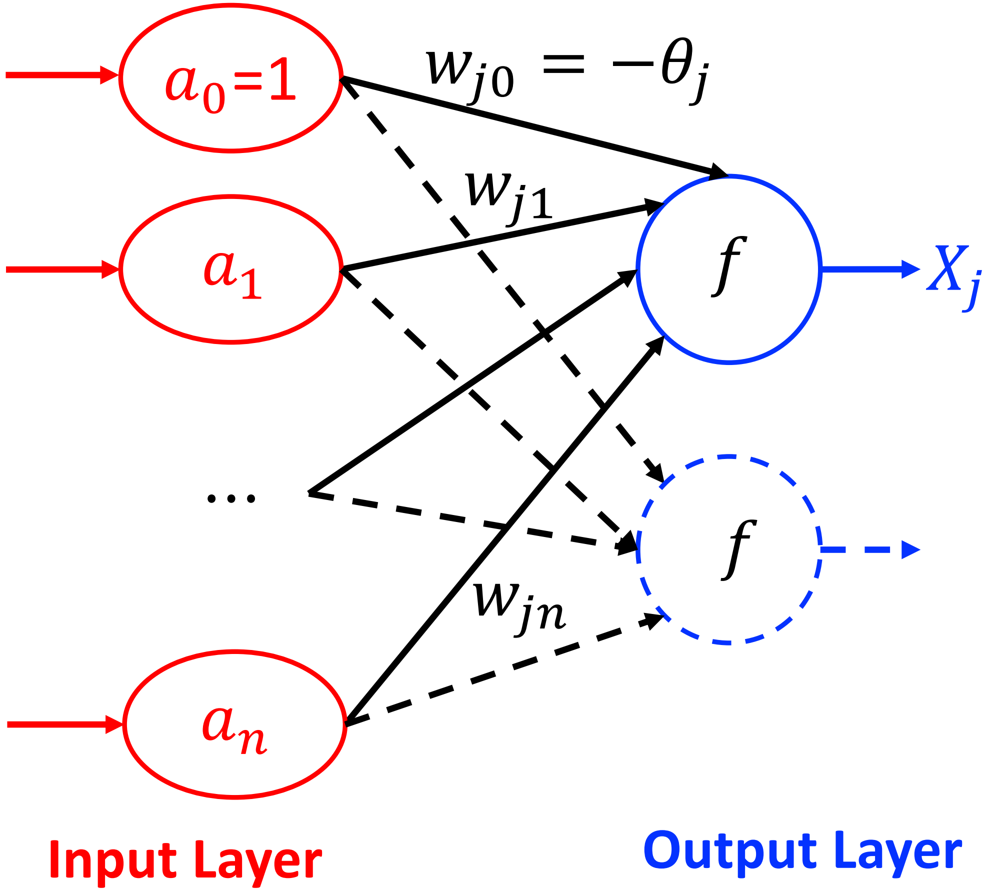
$$\Delta w_{ji}^k = C e_j^k a_i^k$$

Perceptron Learning Rule!

According to the learning rule, a weight of connection changes If and only if both the input value and the error of the output are not equal to 0.

Q: Difference with Hebb's rule?

Network Performance

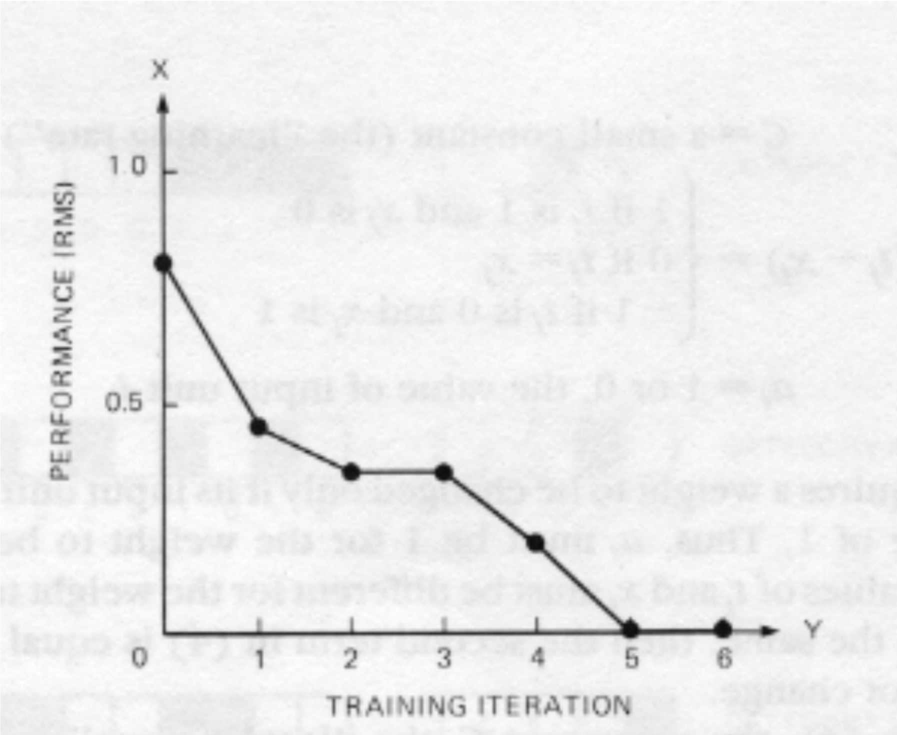


$$\text{RMS} = \sqrt{\frac{\sum_{k=1}^{r}\sum_{j=1}^{m}\left(e_j^k\right)^2}{rm}}$$
$$= \sqrt{\frac{\sum_{k=1}^{r}\sum_{j=1}^{m}\left(t_j^k - X_j^k\right)^2}{rm}}$$

where
r: the number of patterns in the data set,
m: the number of output neurons.

*The **network performance*** during training session is measured by a root-mean-square (RMS) error value.

Network Performance



Thus, RMS error can be represented as:

$$\mathrm{RMS} = F_D(w)$$

where w is network weight, D is data set.

Learning curve: dependency of the RMS error on the number of iterations.

- ***Initially***, the adaptable ***weights are all set to small random values***, and the network does not perform very well;
- Performance improves **during training**;
- Finally, the error gets close to zero, training stops. We say the network has **converged**.

Perceptron Learning Algorithm

Algorithm 1: Perceptron Learning Algorithm

Data: Labelled data set D: r n-dimensional input points, each of which has m labels. Small positive real δ. Learning rate C.

Result: Weight matrix $w = [w_1, \cdots, w_m]$

Initialize weights w randomly;
while *!convergence (RMS $\leq \delta$)* **do**
    Pick random $a' \in D$;
    $a \longleftarrow [1, a']$;
    for $j = 1, \cdots, m$ **do**
        `/* We represent the learning rule in the vector form */`
        $w_j = w_j + C(t_j - X_j)a$;
return w;

Perceptron Learning Algorithm

Algorithm 1: Perceptron Learning Algorithm

Data: Labelled data set D: r n-dimensional input points, each of which has m labels. Small positive real δ. Learning rate C.

Result: Weight matrix $w = [w_1, \cdots, w_m]$

Initialize weights w randomly;

while *!convergence* $(RMS \leq \delta)$ **do**

    Pick random $a' \in D$;

    $a \longleftarrow [1, a']$;

    for $j = 1, \cdots, m$ **do**

        `/* We represent the learning rule in the vector form */`

        $w_j = w_j + C(t_j - X_j)a$;

return w;

A common way is to enumerate all the patterns in *D* sequentially. An epoch means training the neural network with all the training data for one cycle.

Perceptron Learning Algorithm

Algorithm 1: Perceptron Learning Algorithm

Data: Labelled data set D: r n-dimensional input points, each of which has m labels. Small positive real δ. Learning rate C.

Result: Weight matrix $w = [w_1, \cdots, w_m]$

Initialize weights w randomly;

while $!convergence\ (RMS \leq \delta)$ **do**

  Pick random $a' \in D$;

  $a \longleftarrow [1, a']$;

  for $j = 1, \cdots, m$ **do**

    `/* We represent the learning rule in the vector form */`

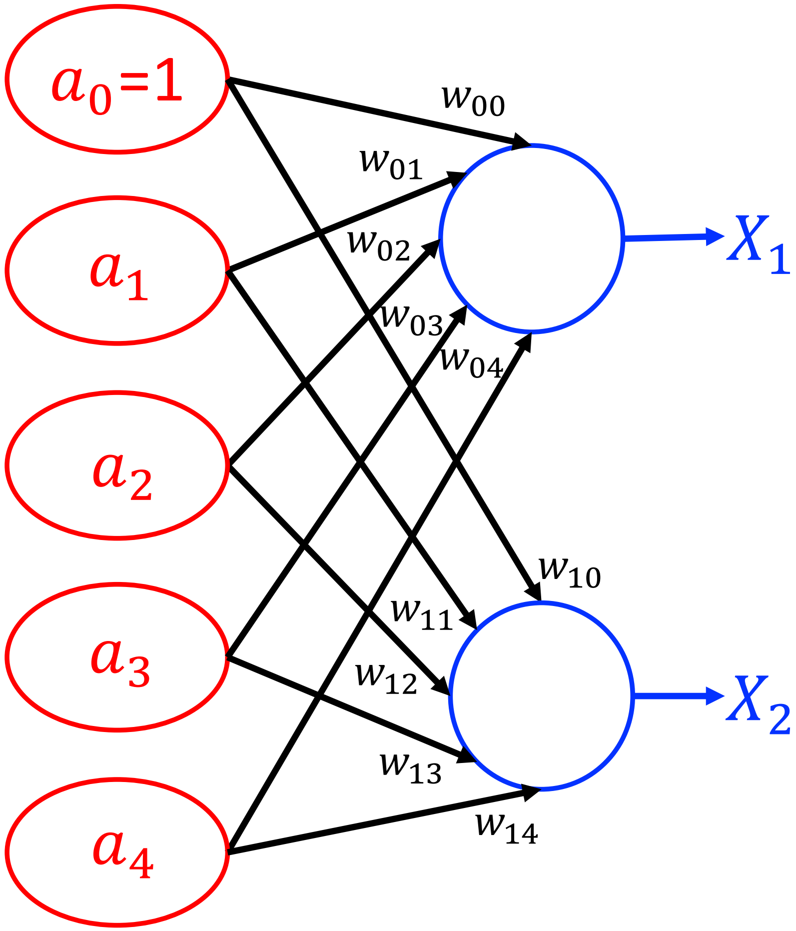
    $w_j = w_j + C(t_j - X_j)a$;

return w;

Then the convergence checking is only done after one epoch.

A common way is to enumerate all the patterns in *D* sequentially. An epoch means training the neural network with all the training data for one cycle.

A Running Example

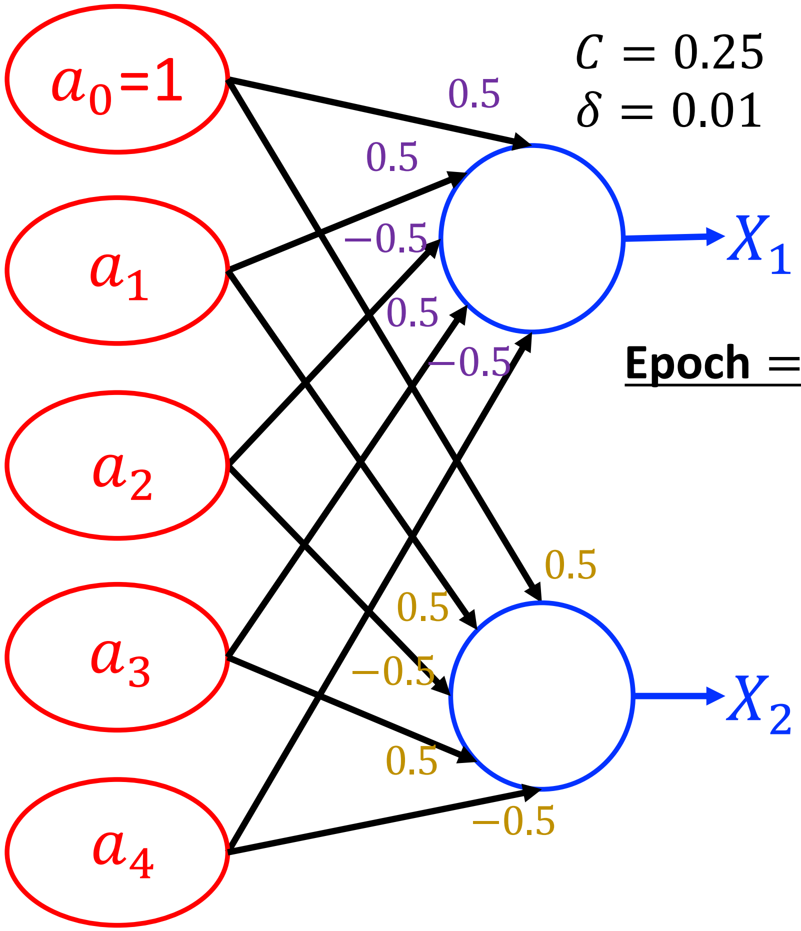


Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

A Running Example



$C = 0.25$
$\delta = 0.01$

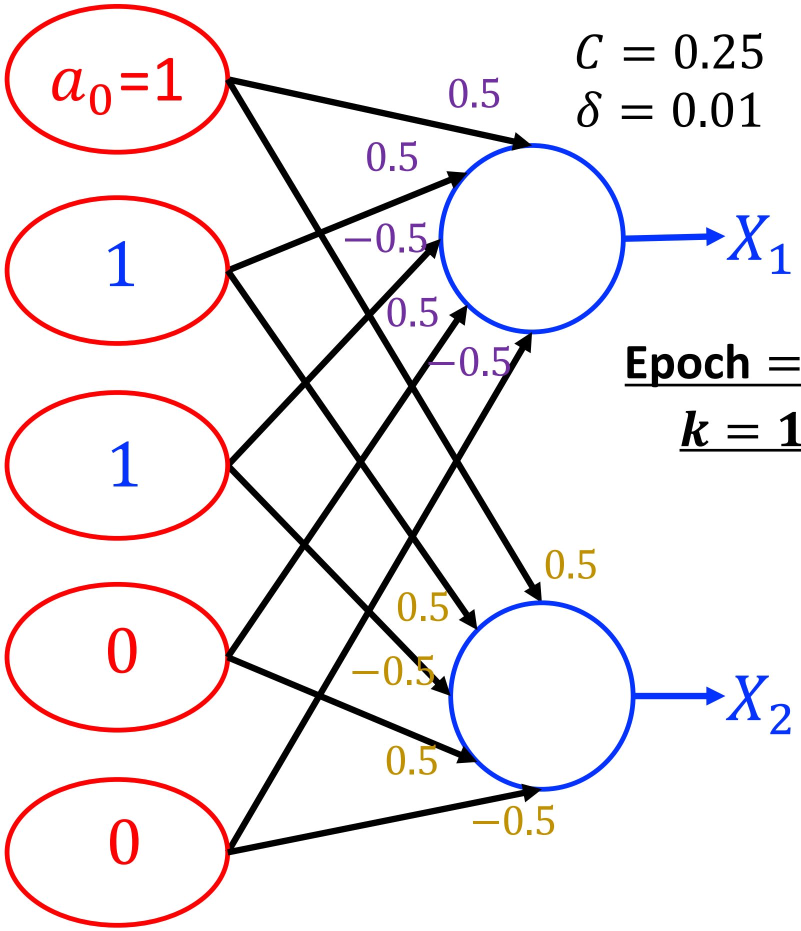
Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

Epoch = 1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 1}$

Training set: 2 patterns

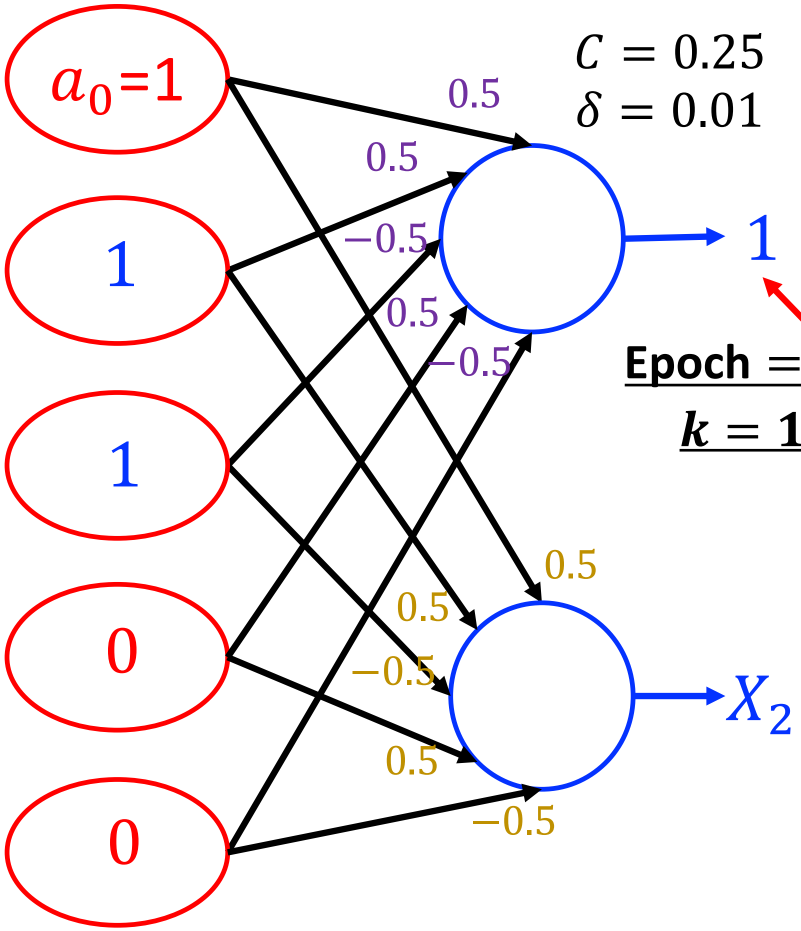
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

$$S_1^1 = \sum_{i=0}^{4} w_{1i}^1 a_i^1$$
$$= 0.5\times1 + 0.5\times1 + (-0.5)\times1 + 0.5\times0 + (-0.5)\times0 = 0.5$$
$$\geq 0$$

$$X_1^1 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

<u>**Epoch = 1**</u>

<u>$\boldsymbol{k} = \mathbf{1}$</u>

Training set: 2 patterns

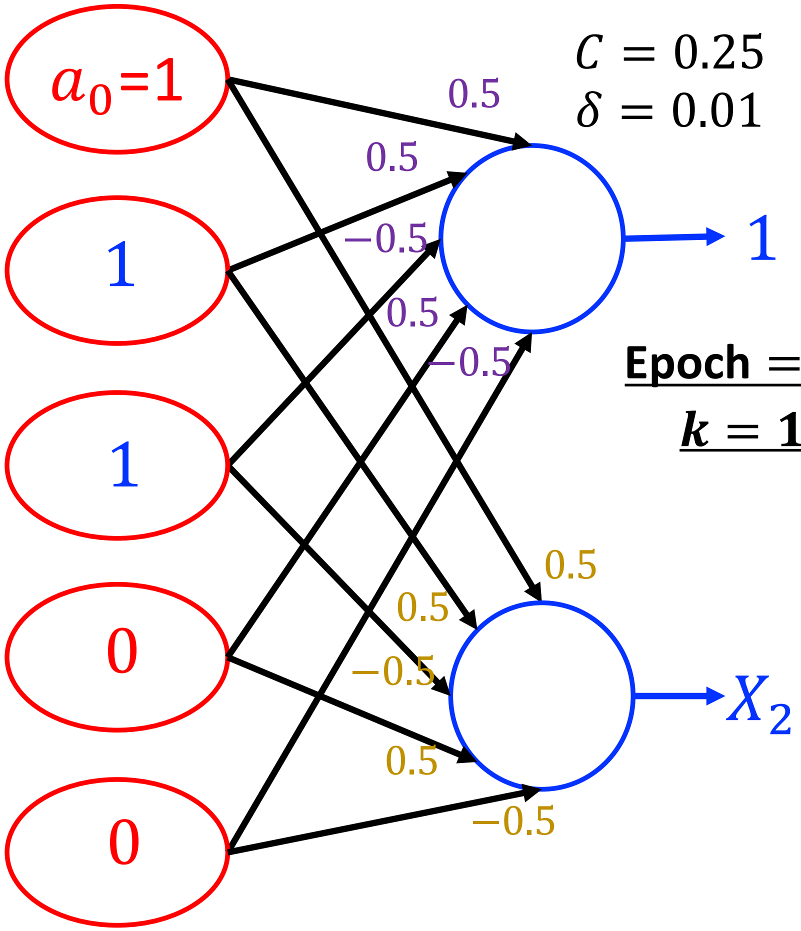
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

$$S_1^1 = \sum_{i=0}^{4} w_{1i}^1 a_i^1$$
$$= 0.5\times1 + 0.5\times1 + (-0.5)\times1 + 0.5\times0 + (-0.5)\times0 = 0.5$$
$$\geq 0$$

$X_1^1 = 1$

A Running Example



$C = 0.25$
$\delta = 0.01$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

<u>**Epoch = 1**</u>

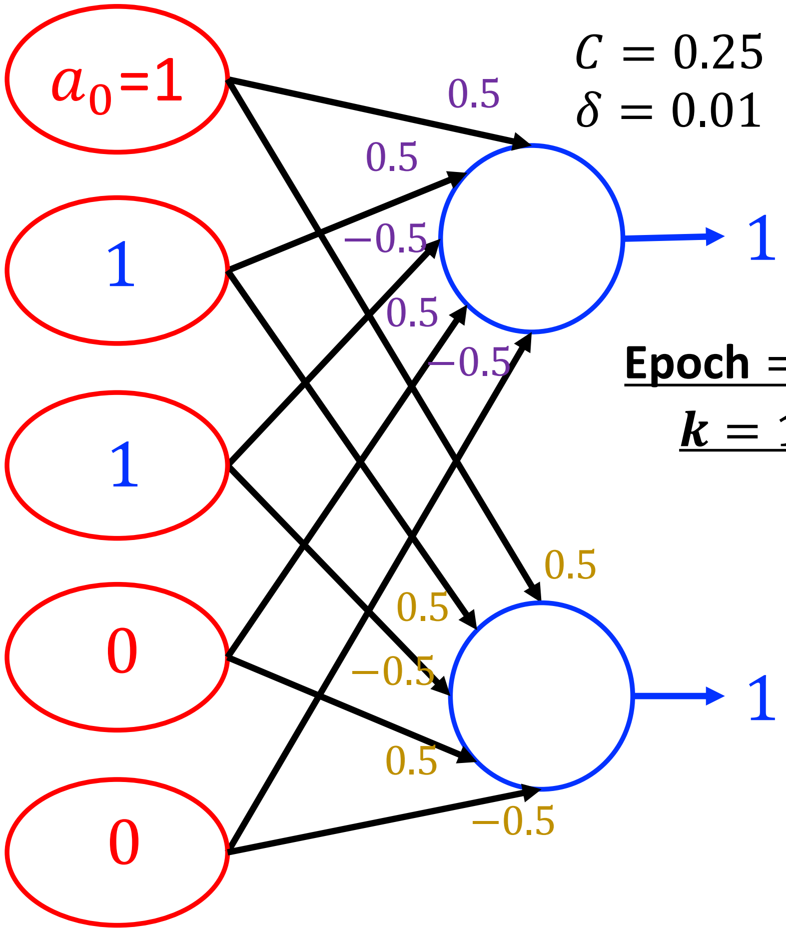
<u>$\boldsymbol{k} = \mathbf{1}$</u>

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

$$S_2^1 = \sum_{i=0}^{4} w_{2i}^1 a_i^1$$
$$= 0.5\times1 + 0.5\times1 + (-0.5)\times1 + 0.5\times0 + (-0.5)\times0 = 0.5$$
$$\geq 0$$

$$X_2^1 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

Epoch = 1

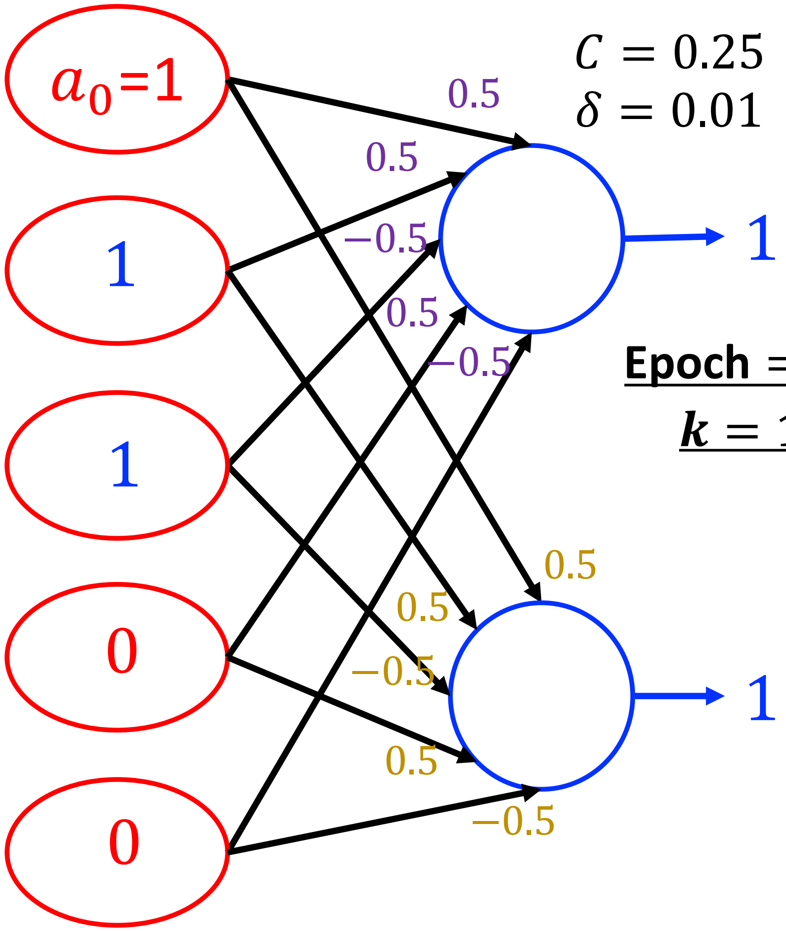
***k* = 1**

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

$$S_2^1 = \sum_{i=0}^{4} w_{2i}^1 a_i^1$$
$$= 0.5\times1 + 0.5\times1 + (-0.5)\times1 + 0.5\times0 + (-0.5)\times0 = 0.5$$
$$\geq 0$$

$X_2^1 = 1$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

***k* = 1**

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

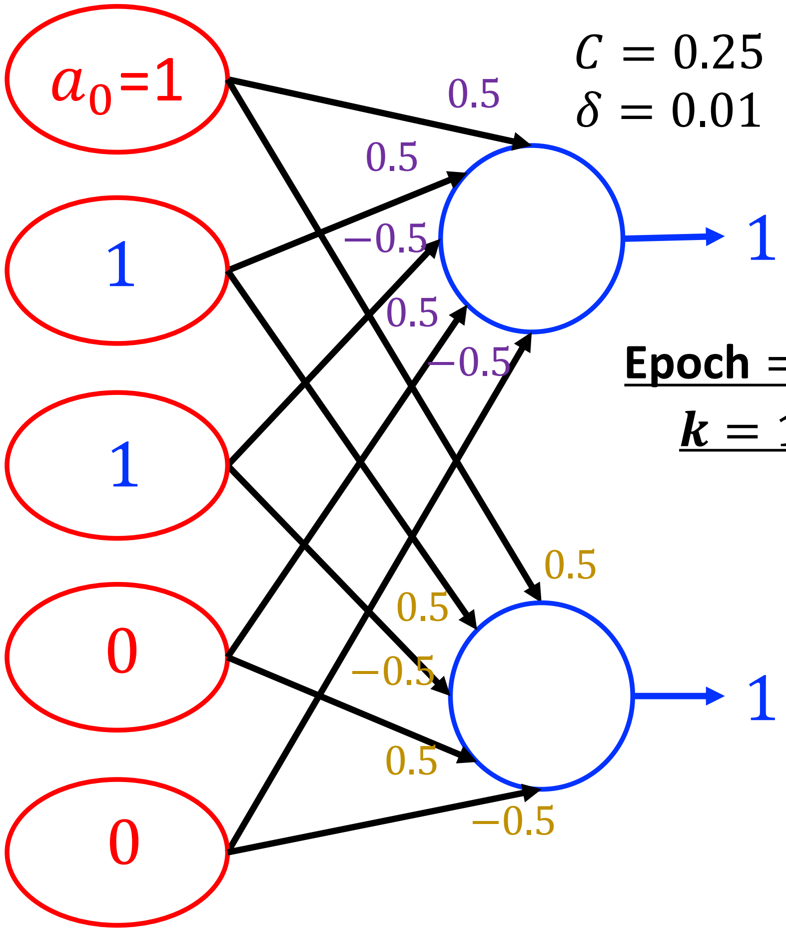
w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

$e_1^1 = (t_1^1 - X_1^1) = (1 - 1) = 0$

$\Delta w_{1i}^1 = Ce_1^1 a_i^1$

$$\Delta w_{10}^1 = Ce_1^1 a_0^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{11}^1 = Ce_1^1 a_1^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{12}^1 = Ce_1^1 a_2^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{13}^1 = Ce_1^1 a_3^1 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{14}^1 = Ce_1^1 a_4^1 = 0.25 \times 0 \times 0 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 1}$

Training set: 2 patterns

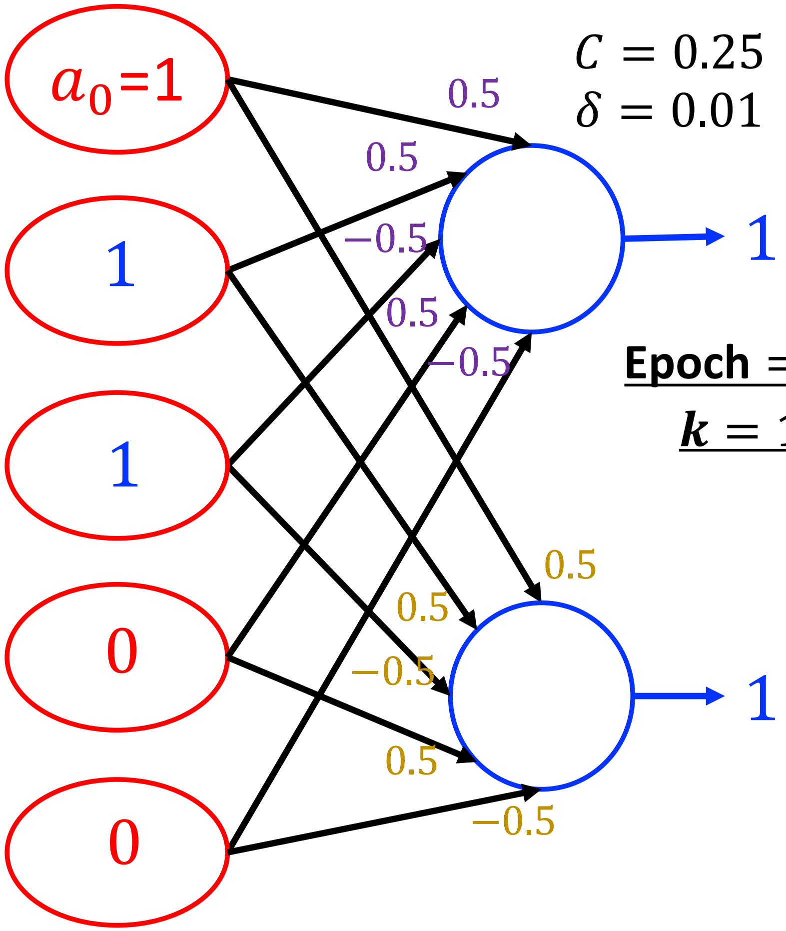
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

$e_2^1 = (t_2^1 - X_2^1) = (0 - 1) = -1$ $\quad \Delta w_{2i}^1 = Ce_2^1 a_i^1$

$$\Delta w_{20}^1 = Ce_2^1 a_0^1 = 0.25\times(-1)\times 1 = -0.25$$
$$\Delta w_{21}^1 = Ce_2^1 a_1^1 = 0.25\times(-1)\times 1 = -0.25$$
$$\Delta w_{22}^1 = Ce_2^1 a_2^1 = 0.25\times(-1)\times 1 = -0.25$$
$$\Delta w_{23}^1 = Ce_2^1 a_3^1 = 0.25\times(-1)\times 0 = 0$$
$$\Delta w_{24}^1 = Ce_2^1 a_4^1 = 0.25\times(-1)\times 0 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

Epoch = 1

$\boldsymbol{k = 1}$

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

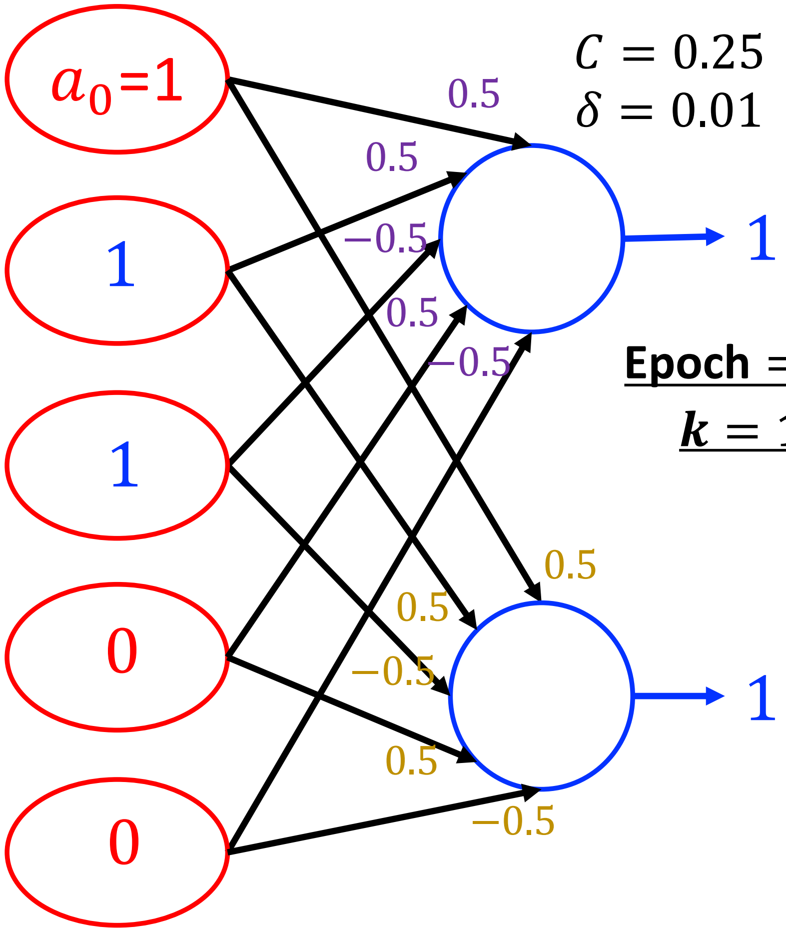
$e_1^1 = 0$

$\Delta w_{10}^1 = 0$
$\Delta w_{11}^1 = 0$
$\Delta w_{12}^1 = 0$
$\Delta w_{13}^1 = 0$
$\Delta w_{14}^1 = 0$

$e_2^1 = -1$

$\Delta w_{20}^1 = -0.25$
$\Delta w_{21}^1 = -0.25$
$\Delta w_{22}^1 = -0.25$
$\Delta w_{23}^1 = 0$
$\Delta w_{24}^1 = 0$

A Running Example



$C = 0.25$
$\delta = 0.01$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

Epoch = 1

$\boldsymbol{k = 1}$

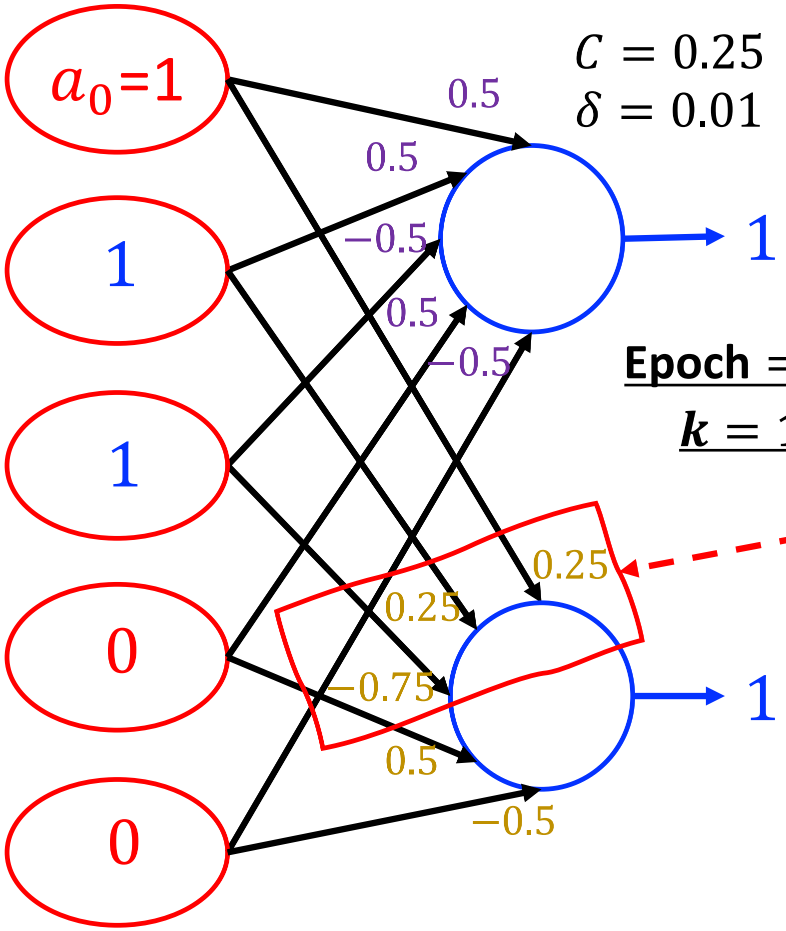
w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.5	0.5	-0.5	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$\Delta w_{10}^1 = 0$
$\Delta w_{11}^1 = 0$
$\Delta w_{12}^1 = 0$
$\Delta w_{13}^1 = 0$
$\Delta w_{14}^1 = 0$

$\Delta w_{20}^1 = -0.25$
$\Delta w_{21}^1 = -0.25$
$\Delta w_{22}^1 = -0.25$
$\Delta w_{23}^1 = 0$
$\Delta w_{24}^1 = 0$

A Running Example



$C = 0.25$
$\delta = 0.01$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

Epoch = 1

$\boldsymbol{k = 1}$

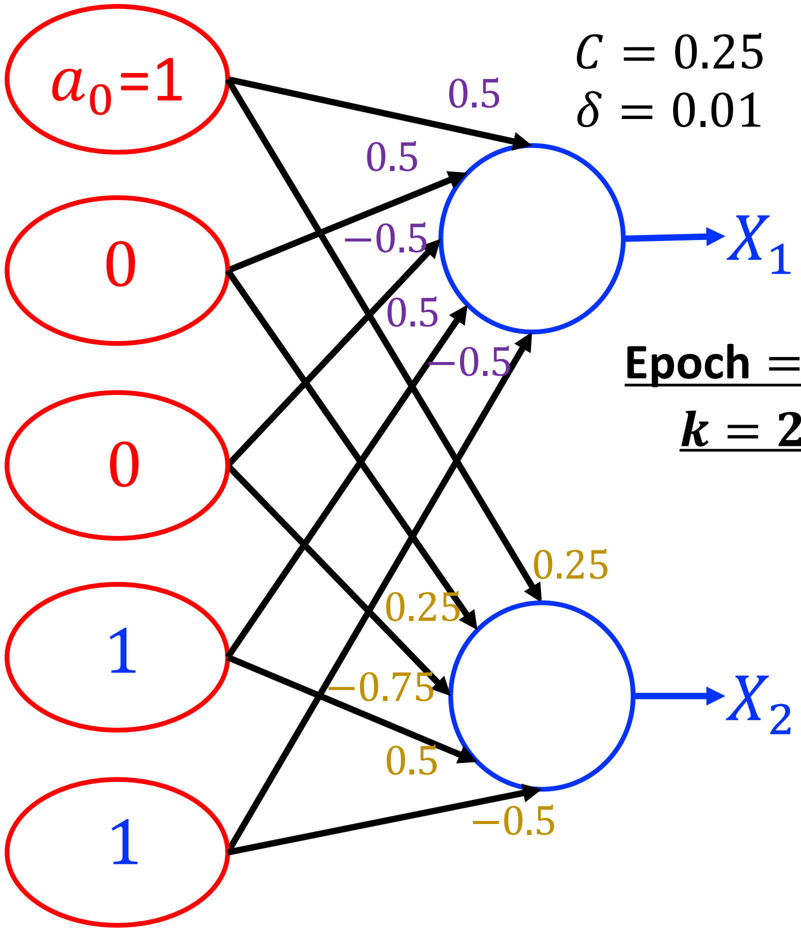
w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$$w_{10}^1 = w_{10}^1 + \Delta w_{10}^1 = 0.5$$
$$w_{11}^1 = w_{11}^1 + \Delta w_{11}^1 = 0.5$$
$$w_{12}^1 = w_{12}^1 + \Delta w_{12}^1 = -0.5$$
$$w_{13}^1 = w_{13}^1 + \Delta w_{13}^1 = 0.5$$
$$w_{13}^1 = w_{13}^1 + \Delta w_{14}^1 = -0.5$$

$$w_{20}^1 = w_{20}^1 + \Delta w_{20}^1 = 0.25$$
$$w_{21}^1 = w_{21}^1 + \Delta w_{21}^1 = 0.25$$
$$w_{22}^1 = w_{22}^1 + \Delta w_{22}^1 = -0.75$$
$$w_{23}^1 = w_{23}^1 + \Delta w_{23}^1 = 0.5$$
$$w_{23}^1 = w_{13}^1 + \Delta w_{24}^1 = -0.5$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

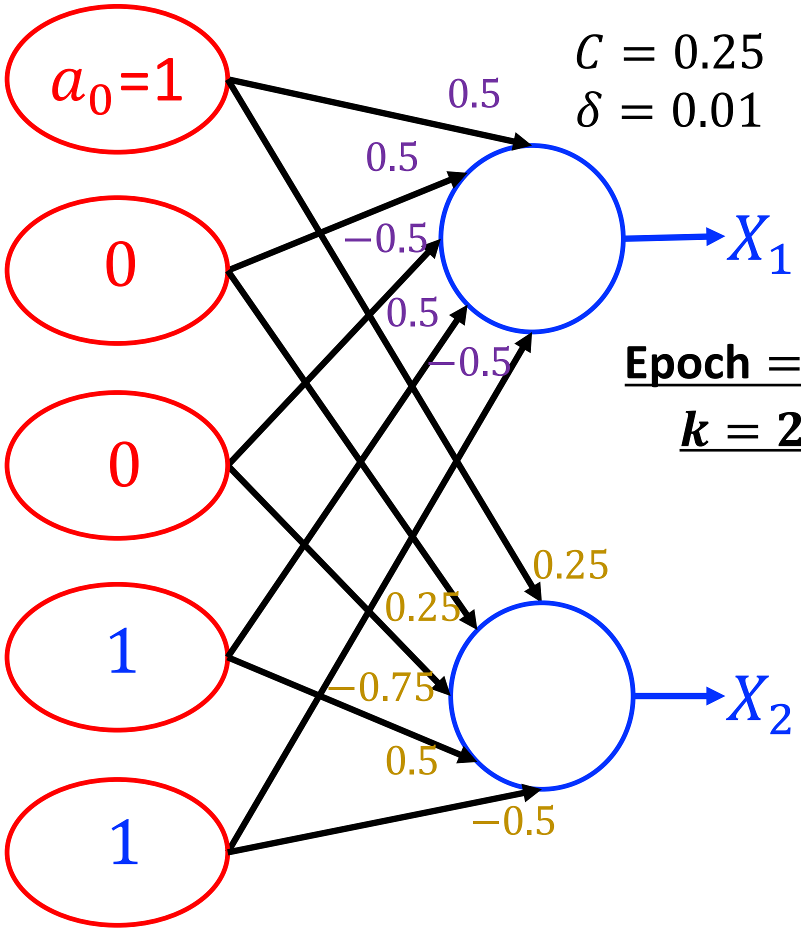
$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 2}$

Training set: 2 patterns

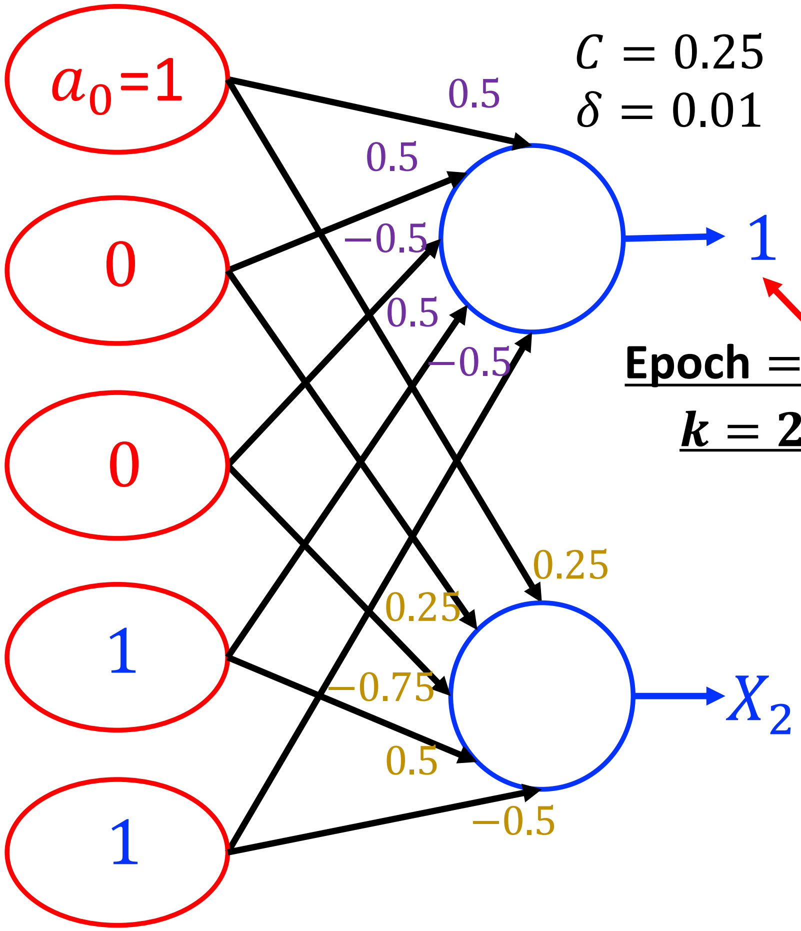
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$$S_1^2 = \sum_{i=0}^{4} w_{1i}^2 a_i^2$$
$$= 0.5\times1 + 0.5\times0 + (-0.5)\times0 + 0.5\times1 + (-0.5)\times1 = 0.5$$
$$\geq 0$$

$$X_1^2 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

***k* = 2**

Training set: 2 patterns

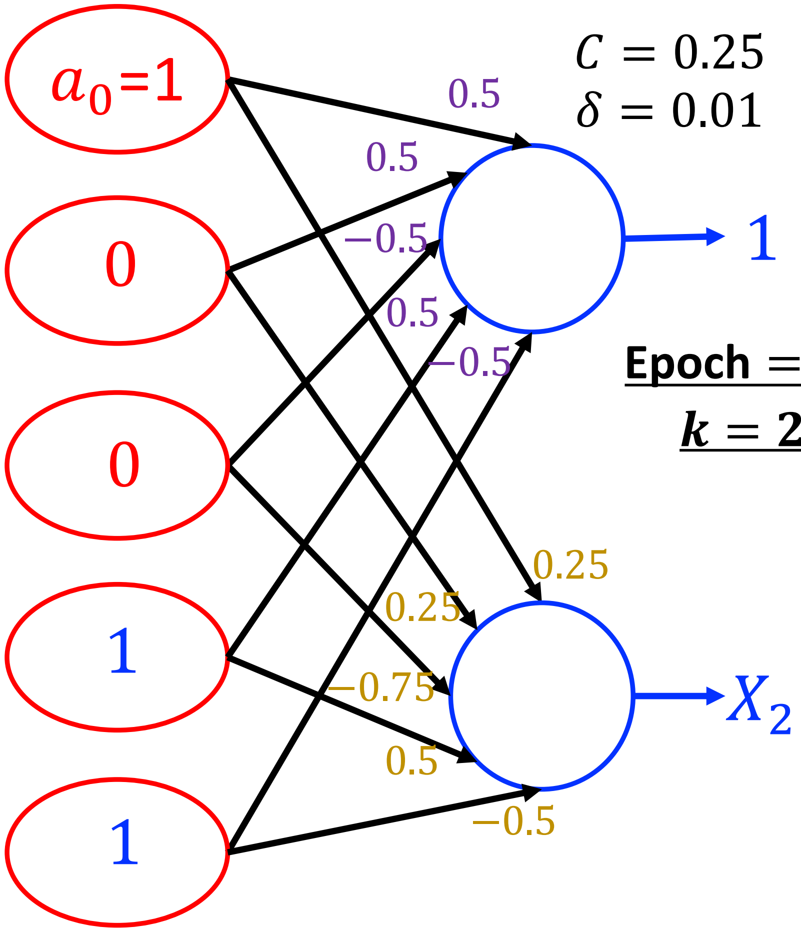
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$$S_1^2 = \sum_{i=0}^{4} w_{1i}^2 a_i^2$$
$$= 0.5\times1 + 0.5\times0 + (-0.5)\times0 + 0.5\times1 + (-0.5)\times1 = 0.5$$
$$\geq 0$$

$$X_1^2 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 2}$

Training set: 2 patterns

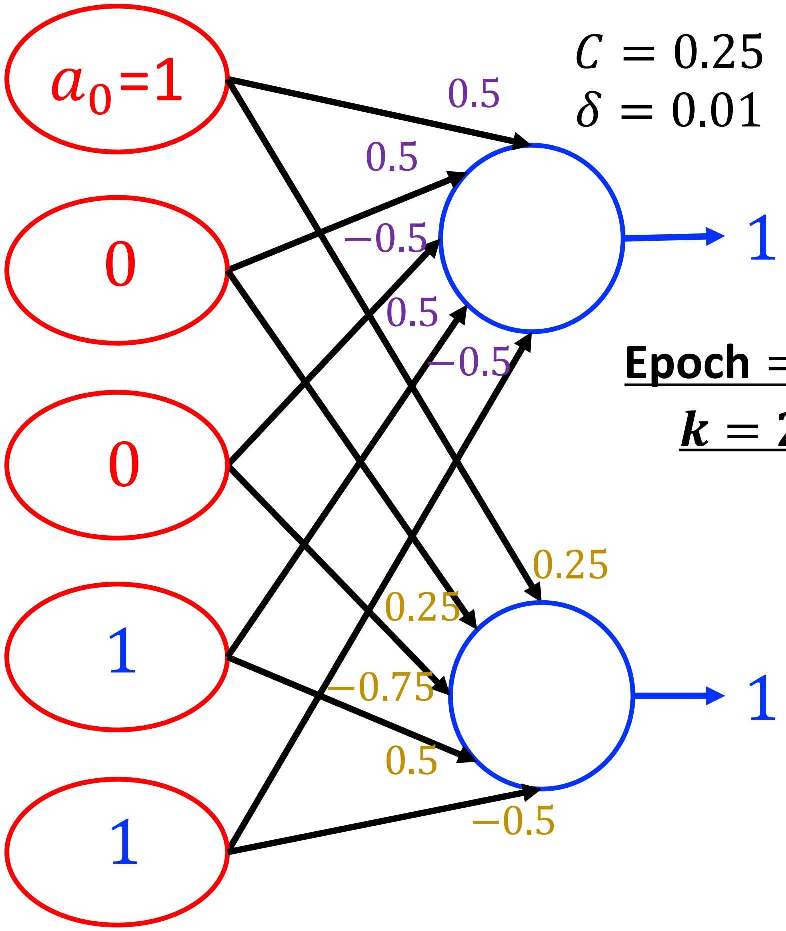
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$$S_2^2 = \sum_{i=0}^{4} w_{2i}^2 a_i^2$$
$$= 0.25\times1 + 0.25\times0 + (-0.75)\times0 + 0.5\times1 + (-0.5)\times1$$
$$= 0.25 \geq 0$$

$$X_2^2 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

<u>**Epoch = 1**</u>

<u>***k* = 2**</u>

Training set: 2 patterns

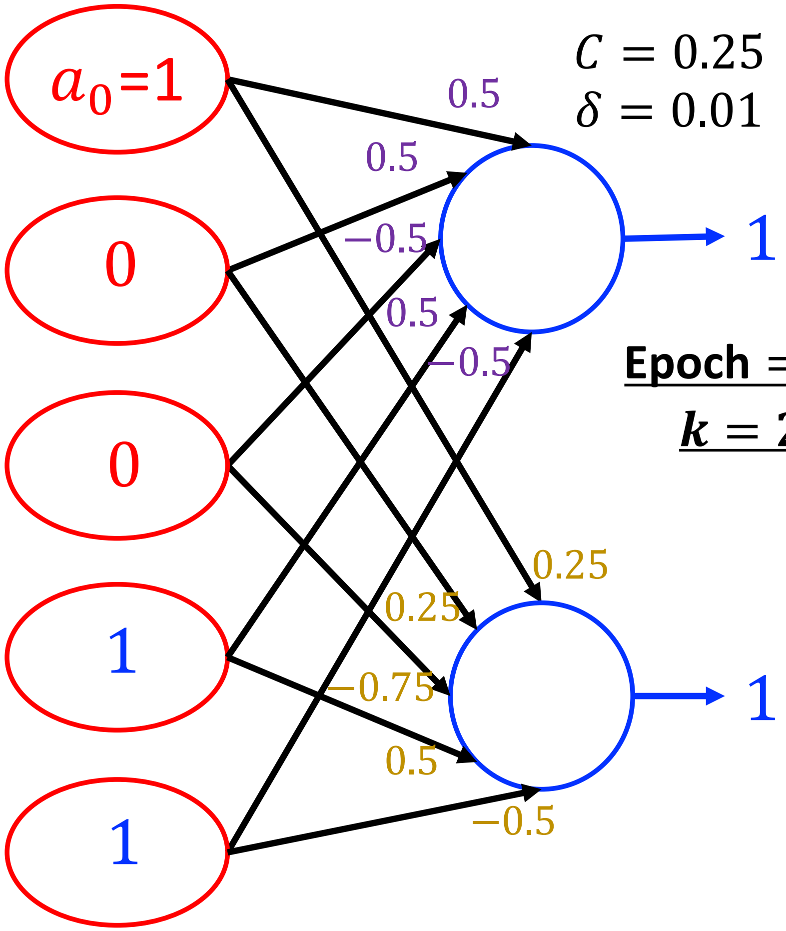
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$$S_2^2 = \sum_{i=0}^{4} w_{2i}^2 a_i^2$$
$$= 0.25\times1 + 0.25\times0 + (-0.75)\times0 + 0.5\times1 + (-0.5)\times1$$
$$= 0.25 \geq 0$$

$$X_2^2 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 2}$

Training set: 2 patterns

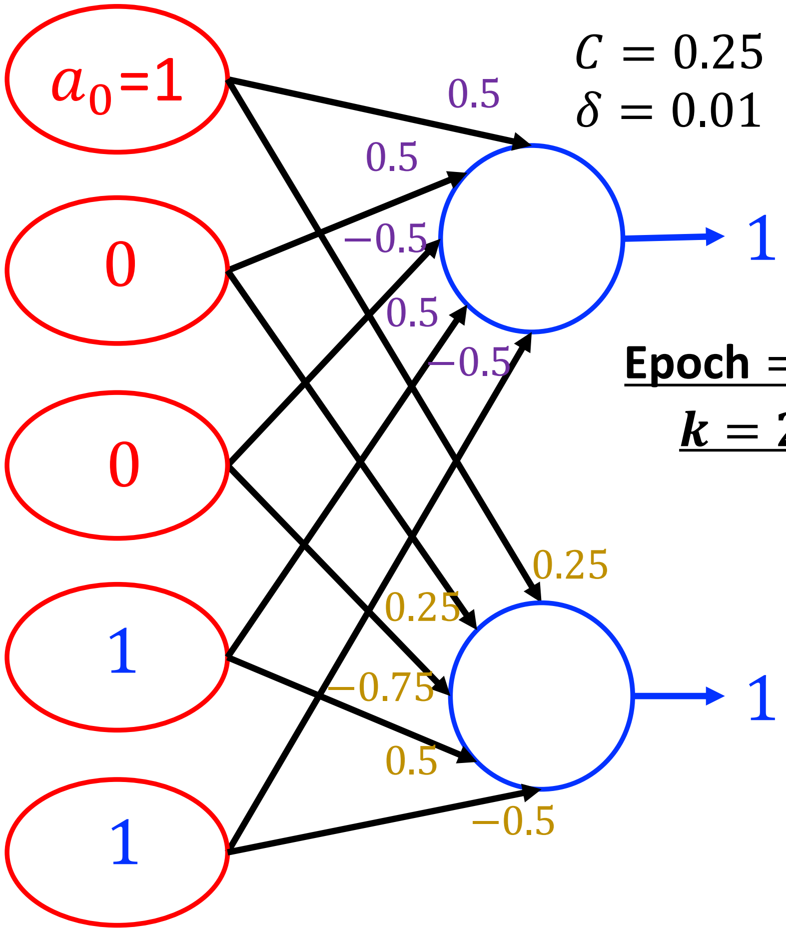
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$e_1^2 = (t_1^2 - X_1^2) = (0 - 1) = -1$ $\quad \Delta w_{1i}^2 = Ce_1^2 a_i^2$

$$\Delta w_{10}^2 = Ce_1^2 a_0^2 = 0.25 \times (-1) \times 1 = -0.25$$
$$\Delta w_{11}^2 = Ce_1^2 a_1^2 = 0.25 \times (-1) \times 0 = 0$$
$$\Delta w_{12}^2 = Ce_1^2 a_2^2 = 0.25 \times (-1) \times 0 = 0$$
$$\Delta w_{13}^2 = Ce_1^2 a_3^2 = 0.25 \times (-1) \times 1 = -0.25$$
$$\Delta w_{14}^2 = Ce_1^2 a_4^2 = 0.25 \times (-1) \times 1 = -0.25$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

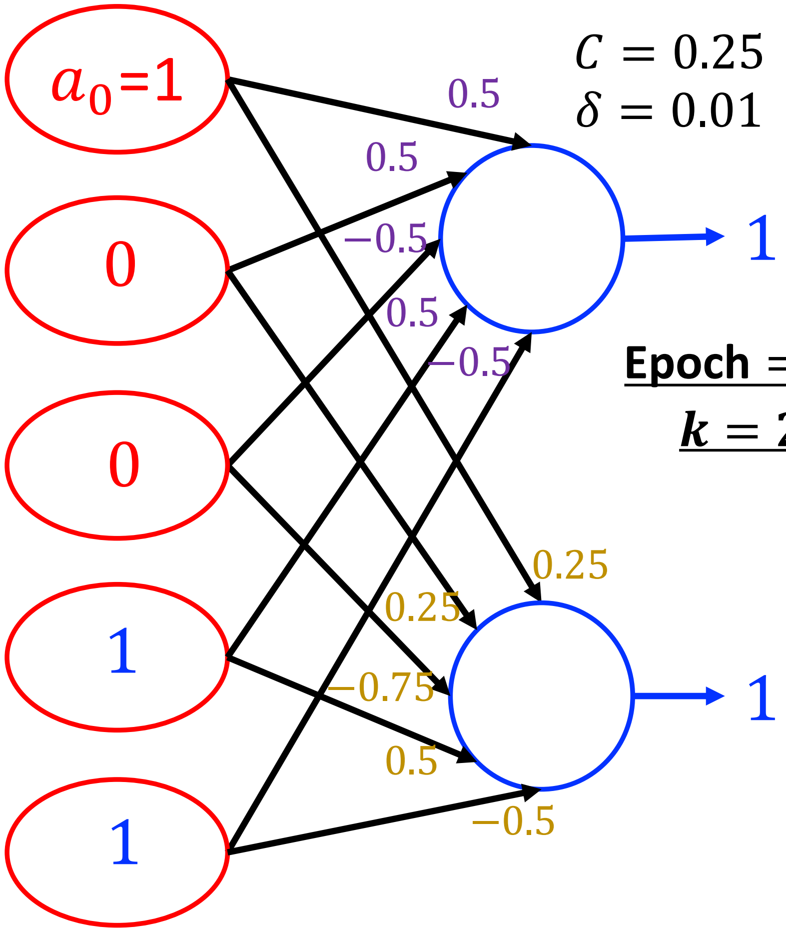
w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$e_2^2 = (t_2^2 - X_2^2) = (1 - 1) = 0$

$\Delta w_{2i}^2 = C e_2^2 a_i^2$

$$\Delta w_{20}^2 = C e_2^2 a_0^2 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{21}^2 = C e_2^2 a_1^2 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{22}^2 = C e_2^2 a_2^2 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{23}^2 = C e_2^2 a_3^2 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{24}^2 = C e_2^2 a_4^2 = 0.25 \times 0 \times 1 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$e_1^2 = -1$

$\Delta w_{10}^2 = -0.25$

$\Delta w_{11}^2 = 0$

$\Delta w_{12}^2 = 0$

$\Delta w_{13}^2 = -0.25$

$\Delta w_{14}^2 = -0.25$

$e_2^2 = 0$

$\Delta w_{20}^2 = 0$

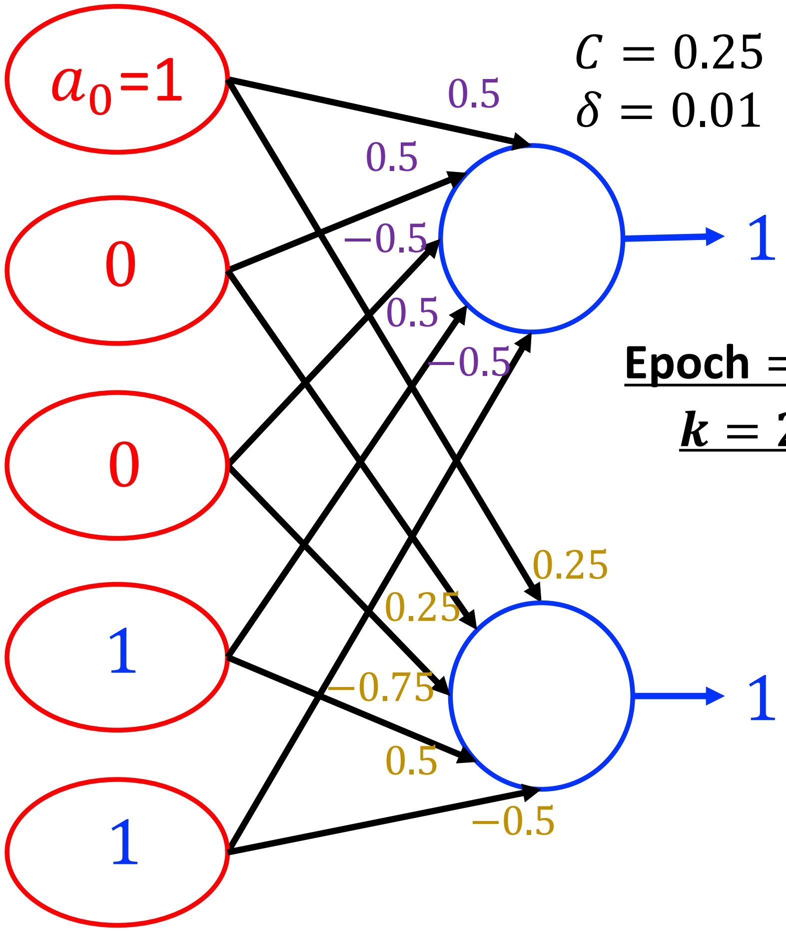
$\Delta w_{21}^2 = 0$

$\Delta w_{22}^2 = 0$

$\Delta w_{23}^2 = 0$

$\Delta w_{24}^2 = 0$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.5	0.5	-0.5	0.5	-0.5	0.25	0.25	-0.75	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$\Delta w_{10}^2 = -0.25$

$\Delta w_{11}^2 = 0$

$\Delta w_{12}^2 = 0$

$\Delta w_{13}^2 = -0.25$

$\Delta w_{14}^2 = -0.25$

$\Delta w_{20}^2 = 0$

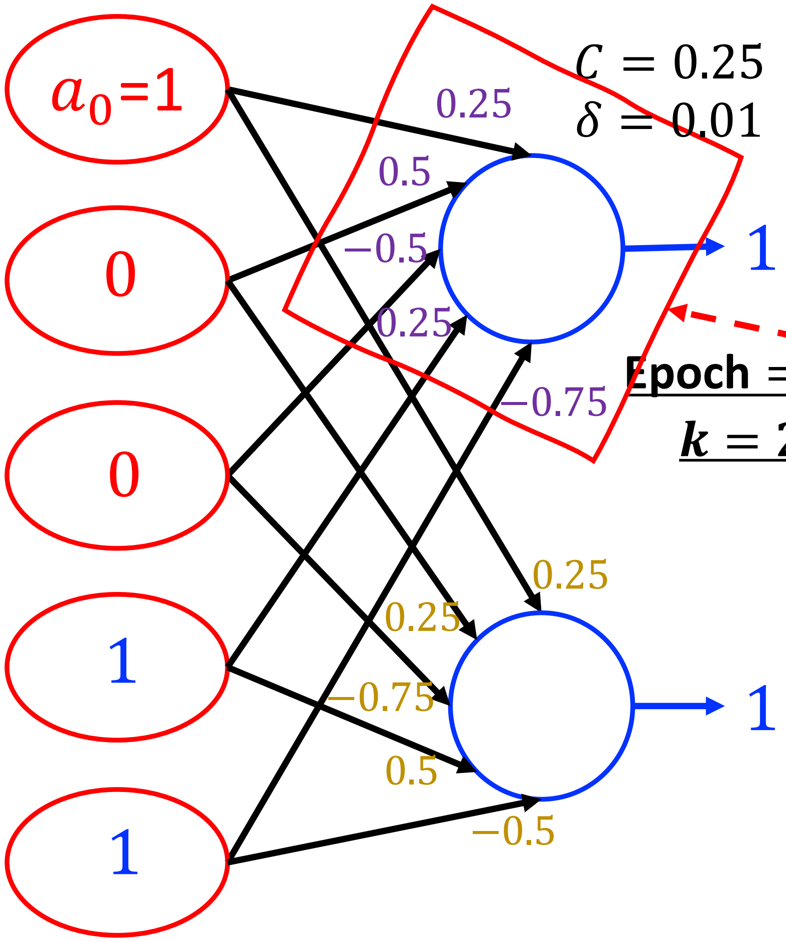
$\Delta w_{21}^2 = 0$

$\Delta w_{22}^2 = 0$

$\Delta w_{23}^2 = 0$

$\Delta w_{24}^2 = 0$

A Running Example



$C = 0.25$

$\delta = 0.01$

Epoch = 1

$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

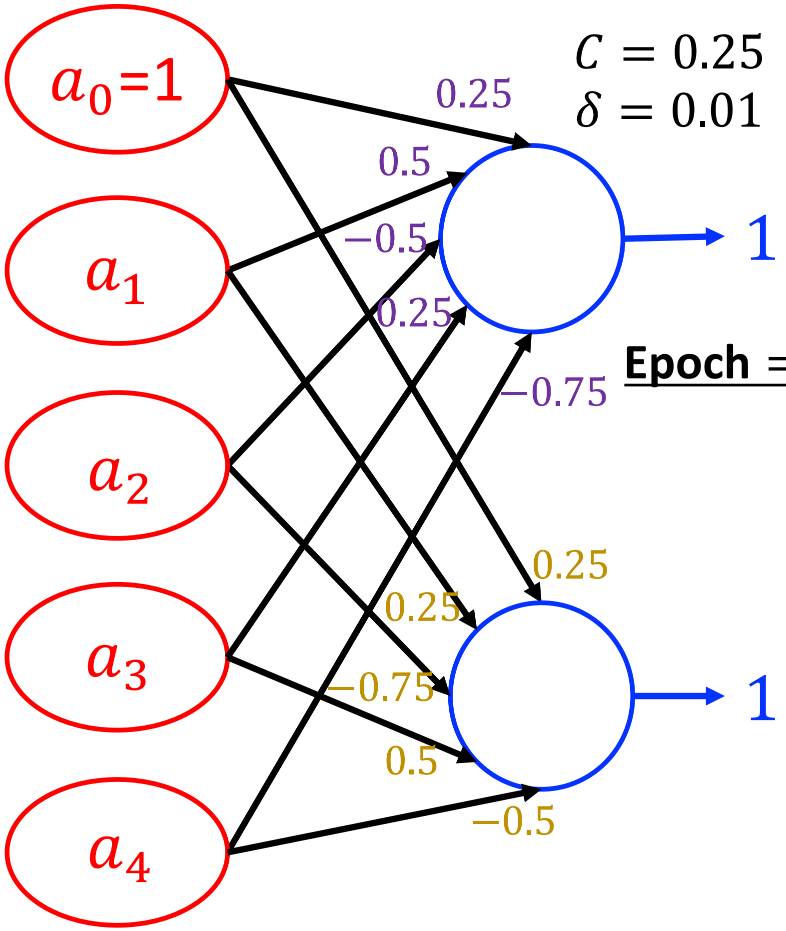
w_{10}^2	w_{11}^2	w_{12}^2	w_{13}^2	w_{14}^2	w_{20}^2	w_{21}^2	w_{22}^2	w_{23}^2	w_{24}^2
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$$w_{10}^2 = w_{10}^2 + \Delta w_{10}^2 = 0.25$$
$$w_{11}^2 = w_{11}^2 + \Delta w_{11}^2 = 0.5$$
$$w_{12}^2 = w_{12}^2 + \Delta w_{12}^2 = -0.5$$
$$w_{13}^2 = w_{13}^2 + \Delta w_{13}^2 = 0.25$$
$$w_{13}^2 = w_{13}^2 + \Delta w_{14}^2 = -0.75$$

$$w_{20}^2 = w_{20}^2 + \Delta w_{20}^2 = 0.25$$
$$w_{21}^2 = w_{21}^2 + \Delta w_{21}^2 = 0.25$$
$$w_{22}^2 = w_{22}^2 + \Delta w_{22}^2 = -0.75$$
$$w_{23}^2 = w_{23}^2 + \Delta w_{23}^2 = 0.5$$
$$w_{23}^2 = w_{13}^2 + \Delta w_{24}^2 = -0.5$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 1

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

w_{10}^2	w_{11}^2	w_{12}^2	w_{13}^2	w_{14}^2	w_{20}^2	w_{21}^2	w_{22}^2	w_{23}^2	w_{24}^2
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

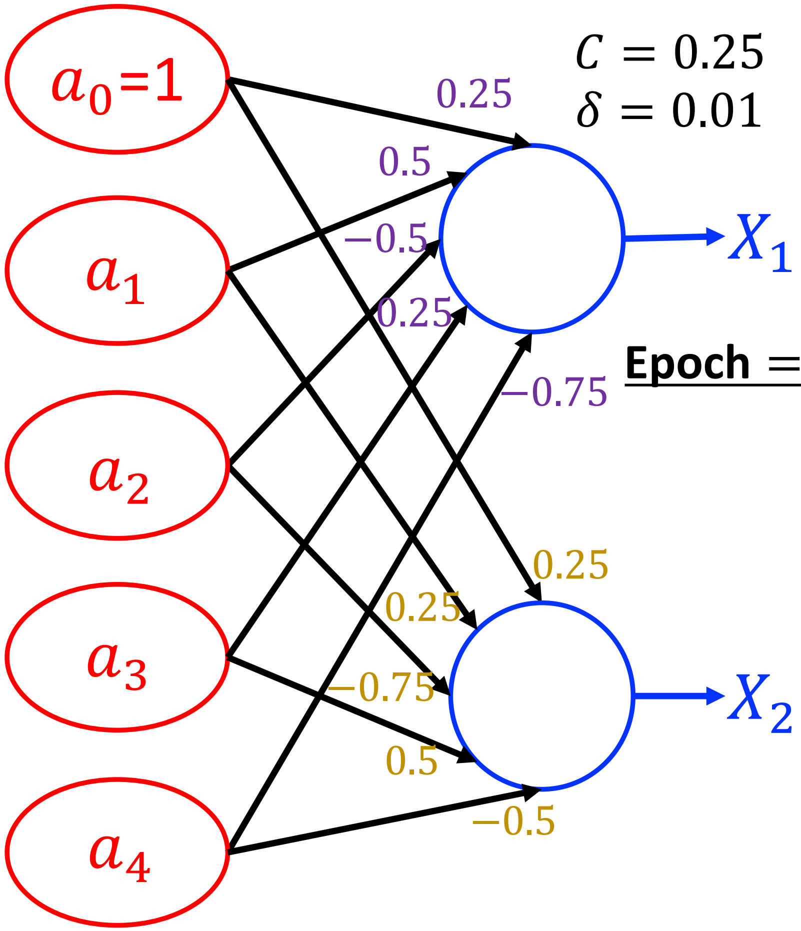
Epoch 1 finishes. Evaluate performance.

$$\mathrm{RMS} = \sqrt{\frac{\sum_{k=1}^{r}\sum_{j=1}^{m}\left(e_j^k\right)^2}{rm}} = \sqrt{\frac{\sum_{k=1}^{2}\sum_{j=1}^{2}\left(e_j^k\right)^2}{2\times 2}}$$

$$= \sqrt{\frac{(e_1^1)^2 + (e_2^1)^2 + (e_1^2)^2 + (e_2^2)^2}{2\times 2}} = \frac{\sqrt{2}}{2} > \delta = 0.01$$

Not good. Continue.

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

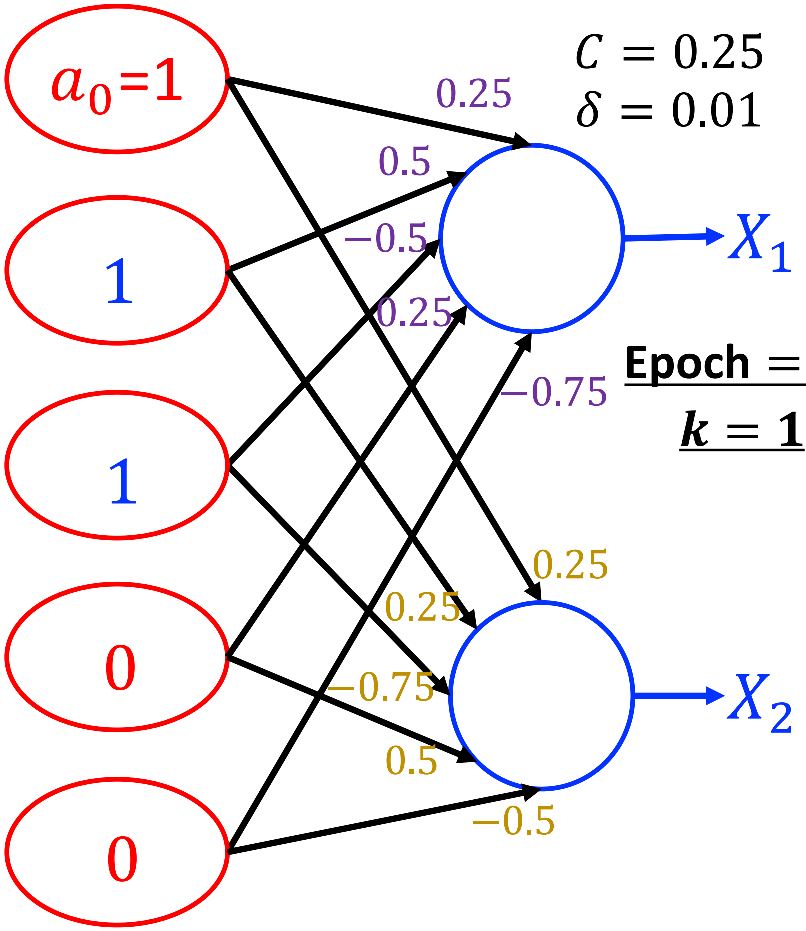
Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

***k* = 1**

Training set: 2 patterns

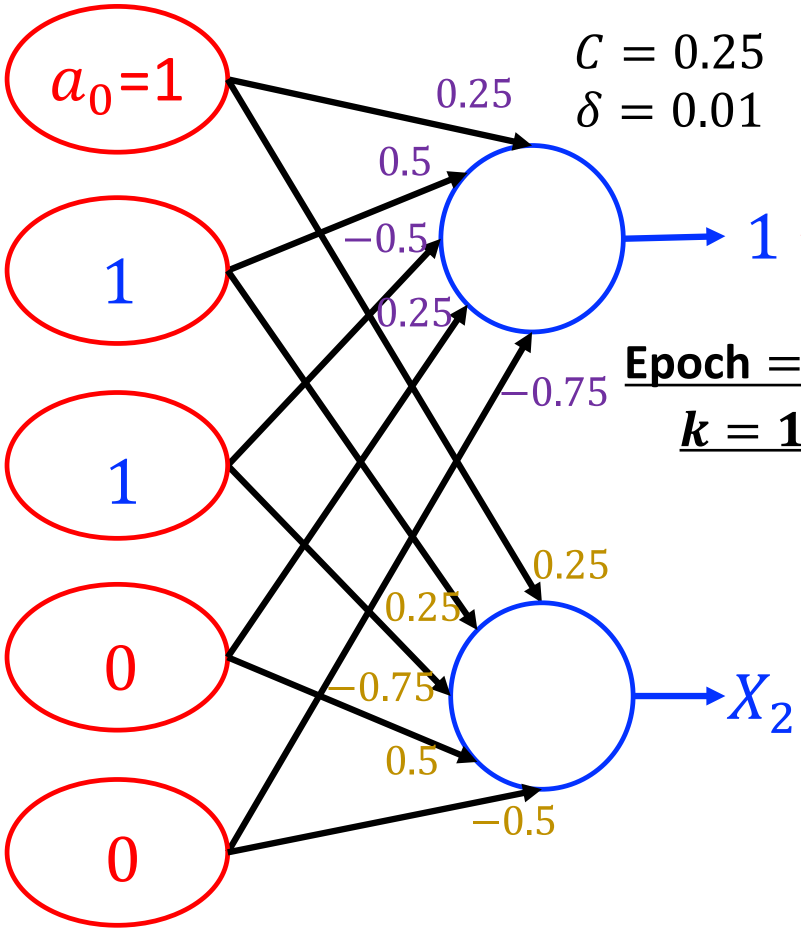
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_1^1 = \sum_{i=0}^{4} w_{1i}^1 a_i^1$$
$$= 0.25\times1 + 0.5\times1 + (-0.5)\times1 + 0.25\times0 + (-0.75)\times0$$
$$= 0.25 \geq 0$$

$$X_1^1 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2
***k* = 1**

Training set: 2 patterns

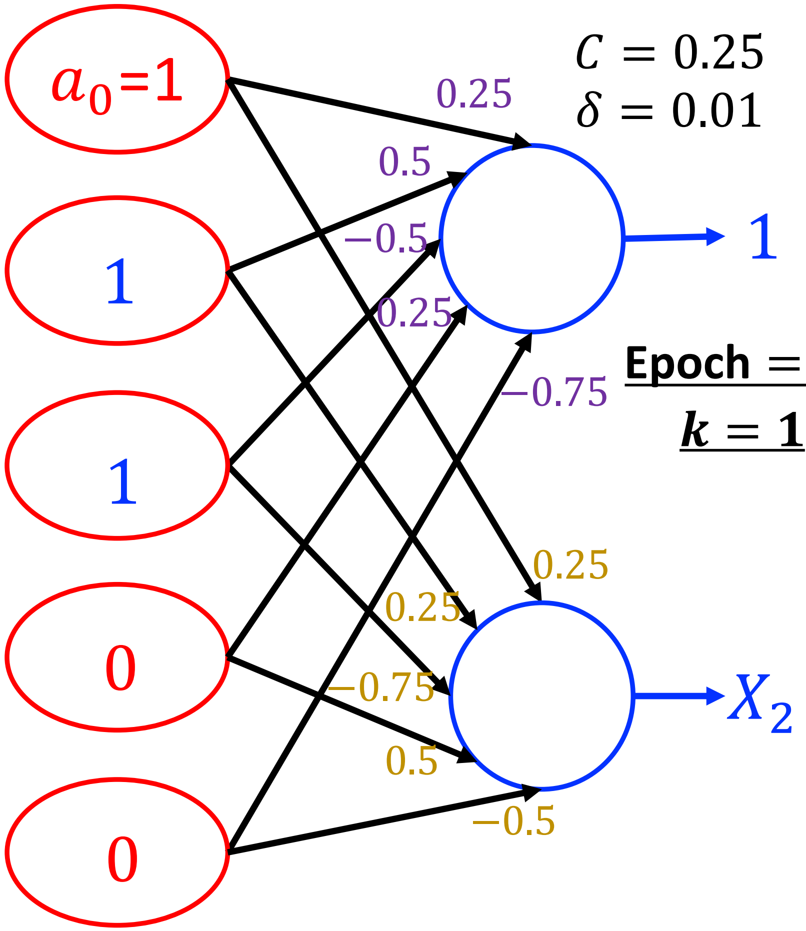
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_1^1 = \sum_{i=0}^{4} w_{1i}^1 a_i^1$$
$$= 0.25\times1 + 0.5\times1 + (-0.5)\times1 + 0.25\times0 + (-0.75)\times0$$
$$= 0.25 \geq 0$$

$$X_1^1 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

$\boldsymbol{k = 1}$

Training set: 2 patterns

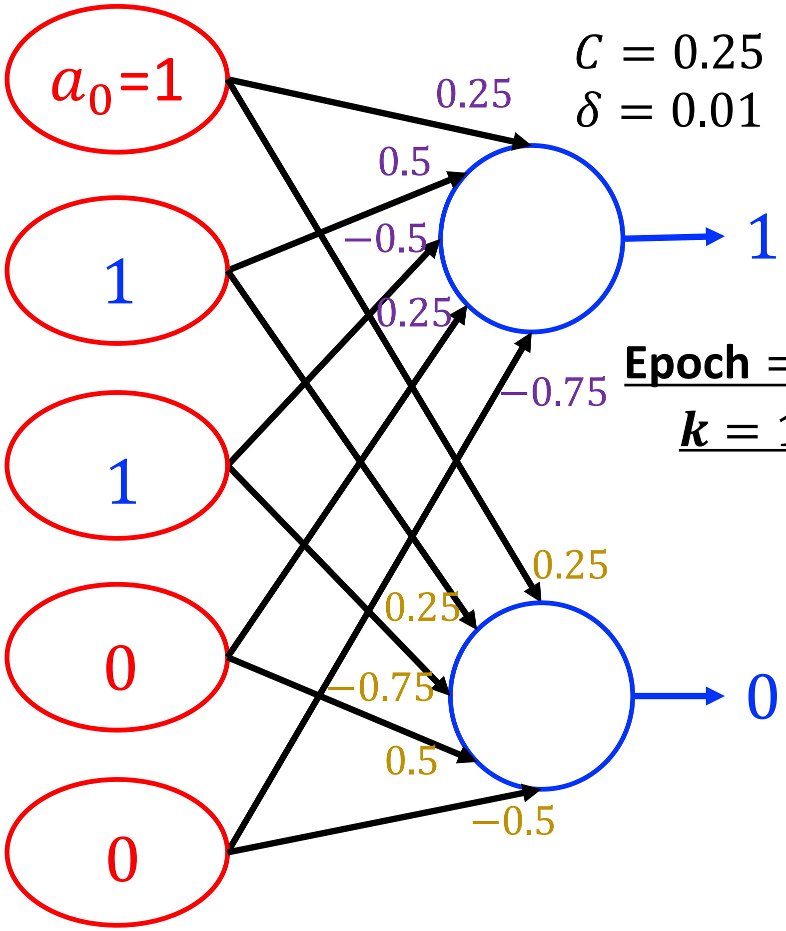
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_2^1 = \sum_{i=0}^{4} w_{2i}^1 a_i^1$$
$$= 0.25\times1 + 0.25\times1 + (-0.75)\times1 + 0.5\times0 + (-0.5)\times0$$
$$= -0.25 < 0$$

$$X_2^1 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

$\boldsymbol{k = 1}$

Training set: 2 patterns

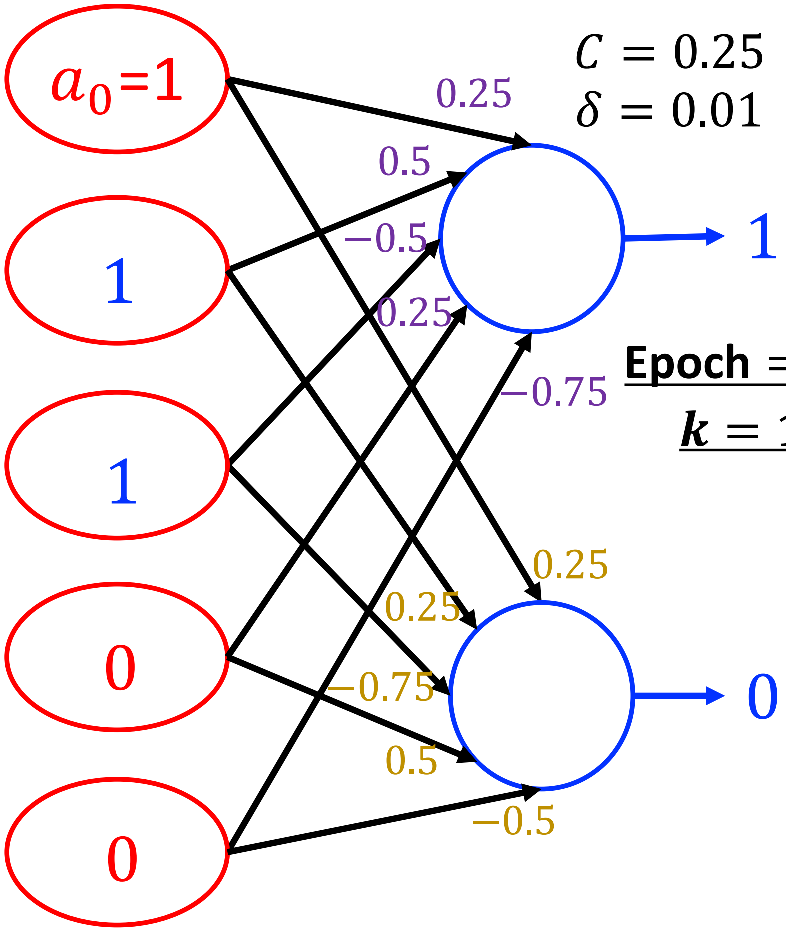
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_2^1 = \sum_{i=0}^{4} w_{2i}^1 a_i^1$$
$$= 0.25\times1 + 0.25\times1 + (-0.75)\times1 + 0.5\times0 + (-0.5)\times0$$
$$= -0.25 < 0$$

$$X_2^1 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

***k* = 1**

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

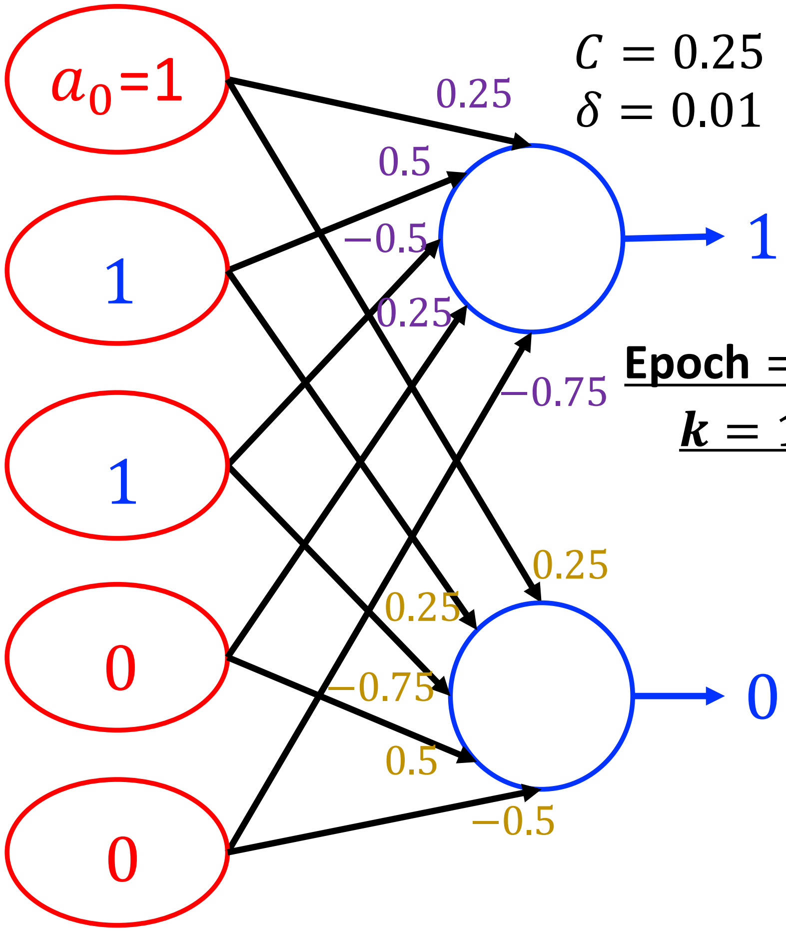
$e_1^1 = (t_1^1 - X_1^1) = (1 - 1) = 0$

$\Delta w_{1i}^1 = Ce_1^1 a_i^1$

$$\Delta w_{10}^1 = Ce_1^1 a_0^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{11}^1 = Ce_1^1 a_1^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{12}^1 = Ce_1^1 a_2^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{13}^1 = Ce_1^1 a_3^1 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{14}^1 = Ce_1^1 a_4^1 = 0.25 \times 0 \times 0 = 0$$

A Running Example

$C = 0.25$
$\delta = 0.01$



Epoch = 2

$\boldsymbol{k = 1}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

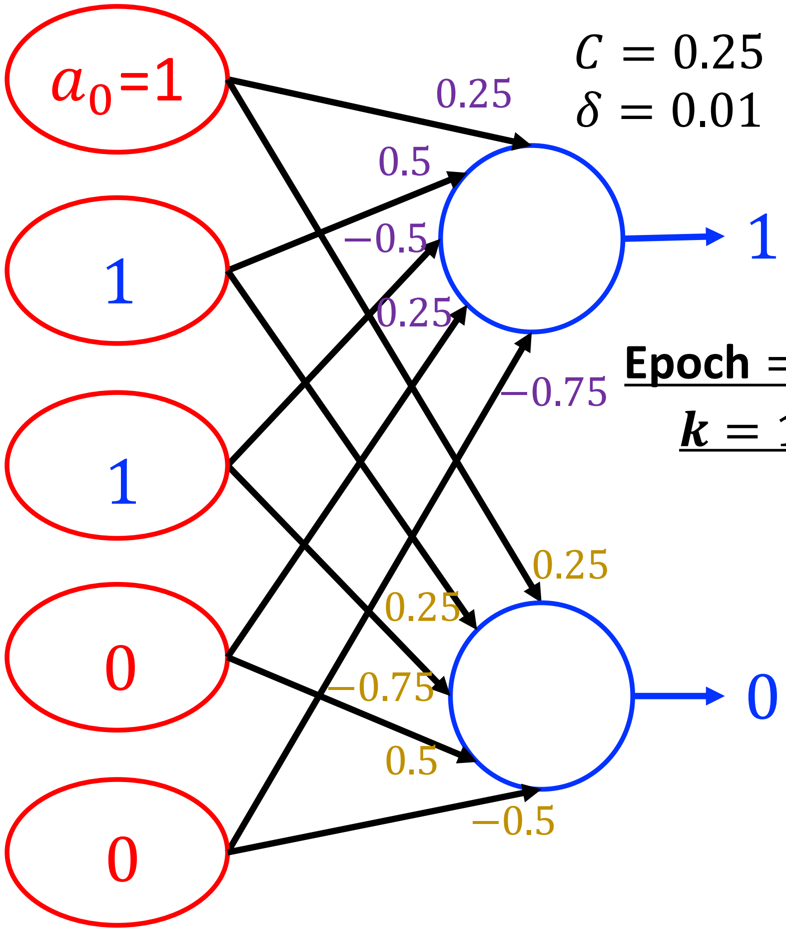
w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$e_2^1 = (t_2^1 - X_2^1) = (0 - 0) = 0$

$\Delta w_{2i}^1 = C e_2^1 a_i^1$

$$\Delta w_{20}^1 = C e_2^1 a_0^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{21}^1 = C e_2^1 a_1^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{22}^1 = C e_2^1 a_2^1 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{23}^1 = C e_2^1 a_3^1 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{24}^1 = C e_2^1 a_4^1 = 0.25 \times 0 \times 0 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

Epoch = 2

$\boldsymbol{k = 1}$

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

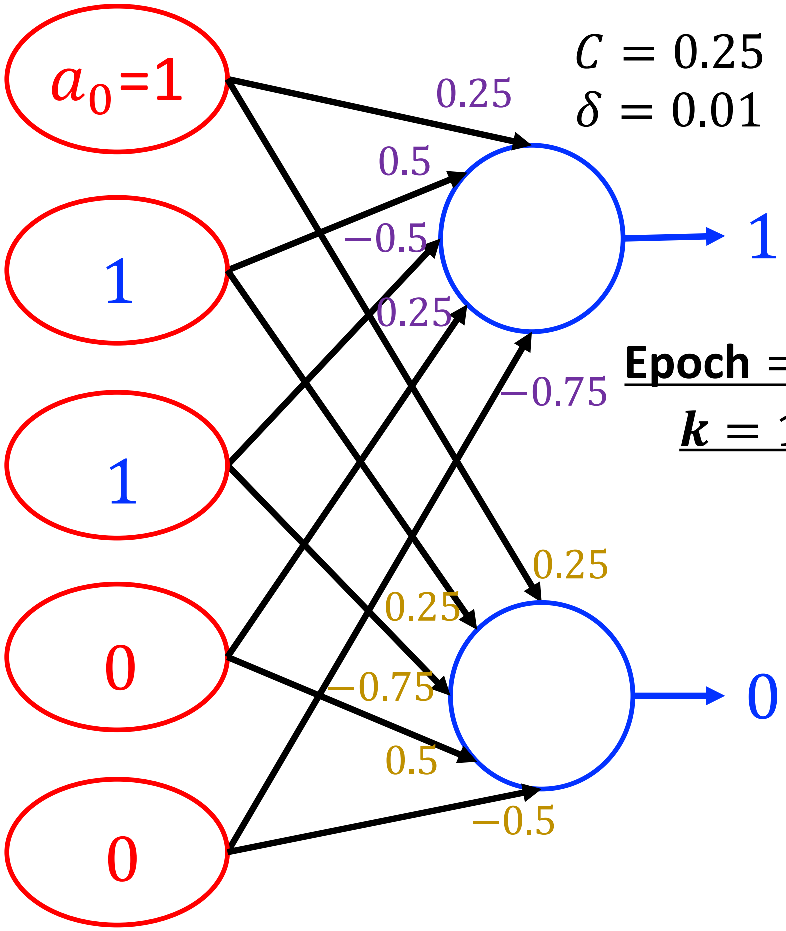
$e_1^1 = 0$

$\Delta w_{10}^1 = 0$
$\Delta w_{11}^1 = 0$
$\Delta w_{12}^1 = 0$
$\Delta w_{13}^1 = 0$
$\Delta w_{14}^1 = 0$

$e_2^1 = -1$

$\Delta w_{20}^1 = 0$
$\Delta w_{21}^1 = 0$
$\Delta w_{22}^1 = 0$
$\Delta w_{23}^1 = 0$
$\Delta w_{24}^1 = 0$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

$\boldsymbol{k = 1}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$\Delta w_{10}^1 = 0$

$\Delta w_{11}^1 = 0$

$\Delta w_{12}^1 = 0$

$\Delta w_{13}^1 = 0$

$\Delta w_{14}^1 = 0$

$\Delta w_{20}^1 = 0$

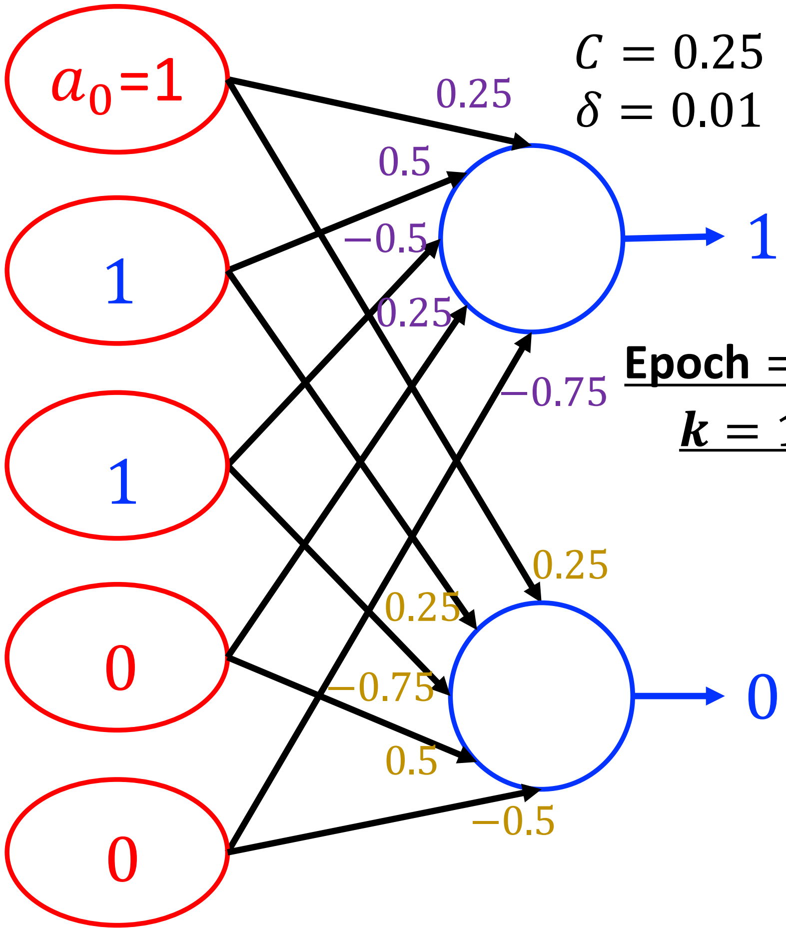
$\Delta w_{21}^1 = 0$

$\Delta w_{22}^1 = 0$

$\Delta w_{23}^1 = 0$

$\Delta w_{24}^1 = 0$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

$\boldsymbol{k = 1}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$w_{10}^1 = w_{10}^1 + \Delta w_{10}^1 = 0.25$

$w_{11}^1 = w_{11}^1 + \Delta w_{11}^1 = 0.5$

$w_{12}^1 = w_{12}^1 + \Delta w_{12}^1 = -0.5$

$w_{13}^1 = w_{13}^1 + \Delta w_{13}^1 = 0.25$

$w_{13}^1 = w_{13}^1 + \Delta w_{14}^1 = -0.75$

$w_{20}^1 = w_{20}^1 + \Delta w_{20}^1 = 0.25$

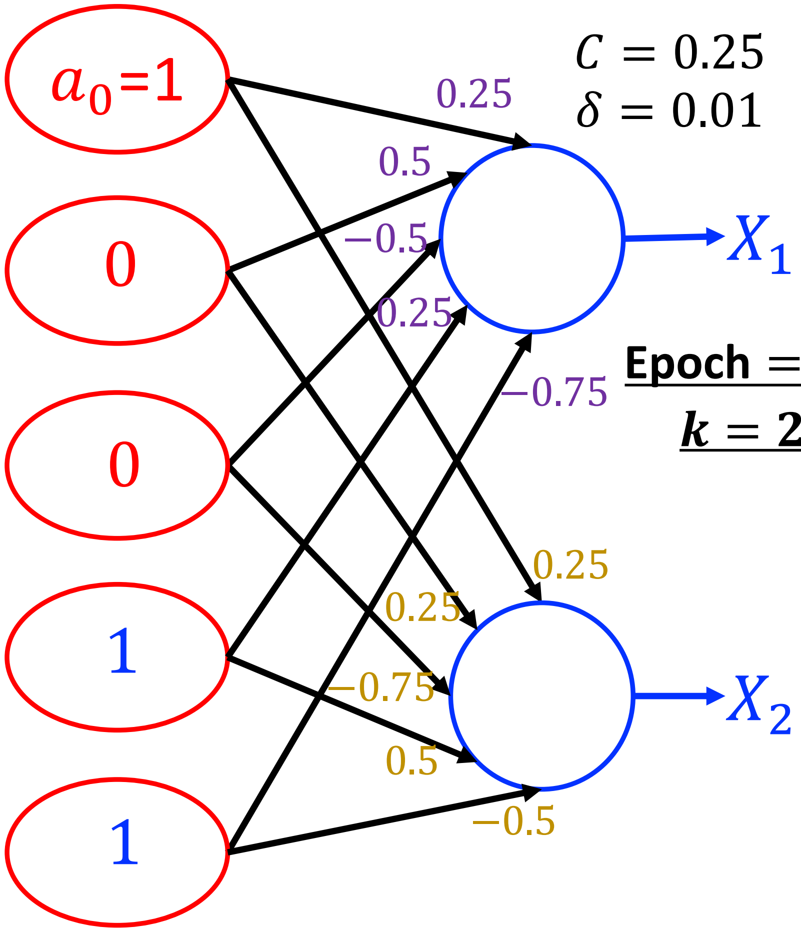
$w_{21}^1 = w_{21}^1 + \Delta w_{21}^1 = 0.25$

$w_{22}^1 = w_{22}^1 + \Delta w_{22}^1 = -0.75$

$w_{23}^1 = w_{23}^1 + \Delta w_{23}^1 = 0.5$

$w_{23}^1 = w_{13}^1 + \Delta w_{24}^1 = -0.5$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

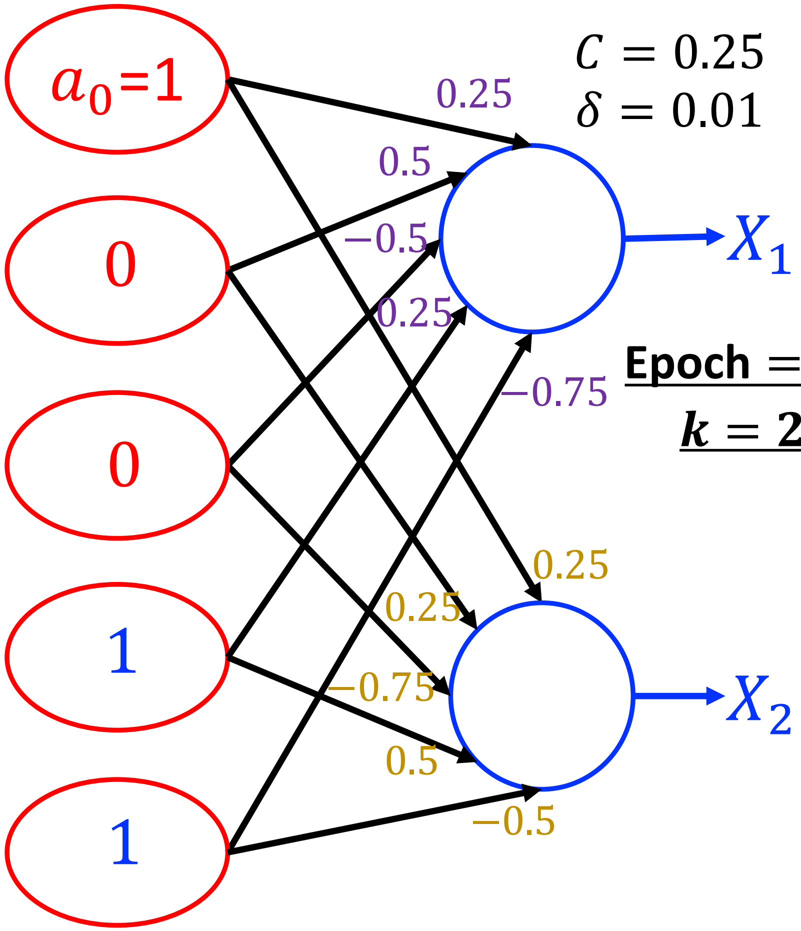
$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2
$\boldsymbol{k = 2}$

Training set: 2 patterns

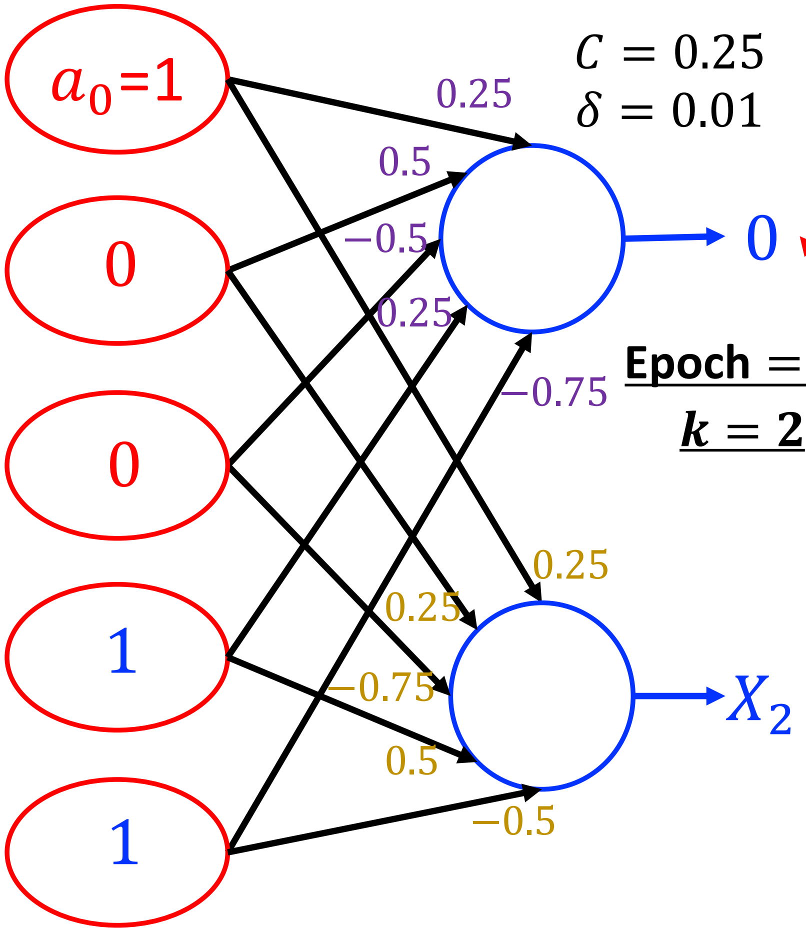
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_1^2 = \sum_{i=0}^{4} w_{1i}^2 a_i^2$$
$$= 0.25\times1 + 0.5\times0 + (-0.5)\times0 + 0.25\times1 + (-0.75)\times1$$
$$= -0.25 < 0$$

$$X_1^2 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

$\boldsymbol{k = 2}$

Training set: 2 patterns

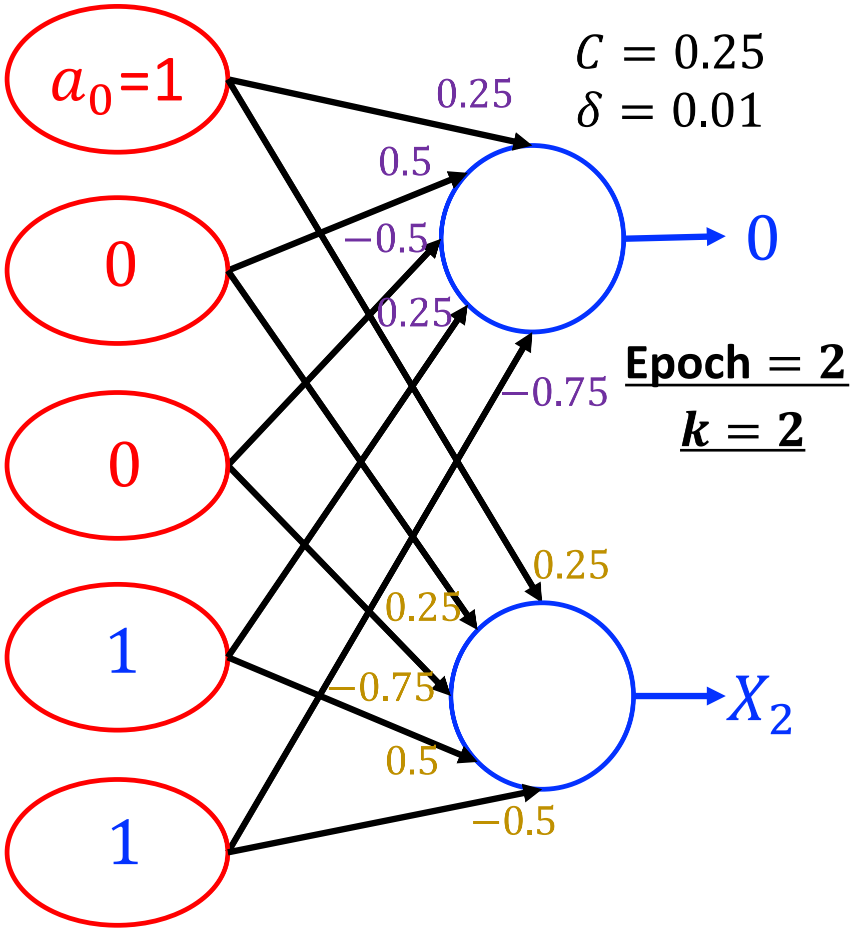
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_1^2 = \sum_{i=0}^{4} w_{1i}^2 a_i^2$$
$$= 0.25\times1 + 0.5\times0 + (-0.5)\times0 + 0.25\times1 + (-0.75)\times1$$
$$= -0.25 < 0$$

$$X_1^2 = 0$$

A Running Example



Training set: 2 patterns

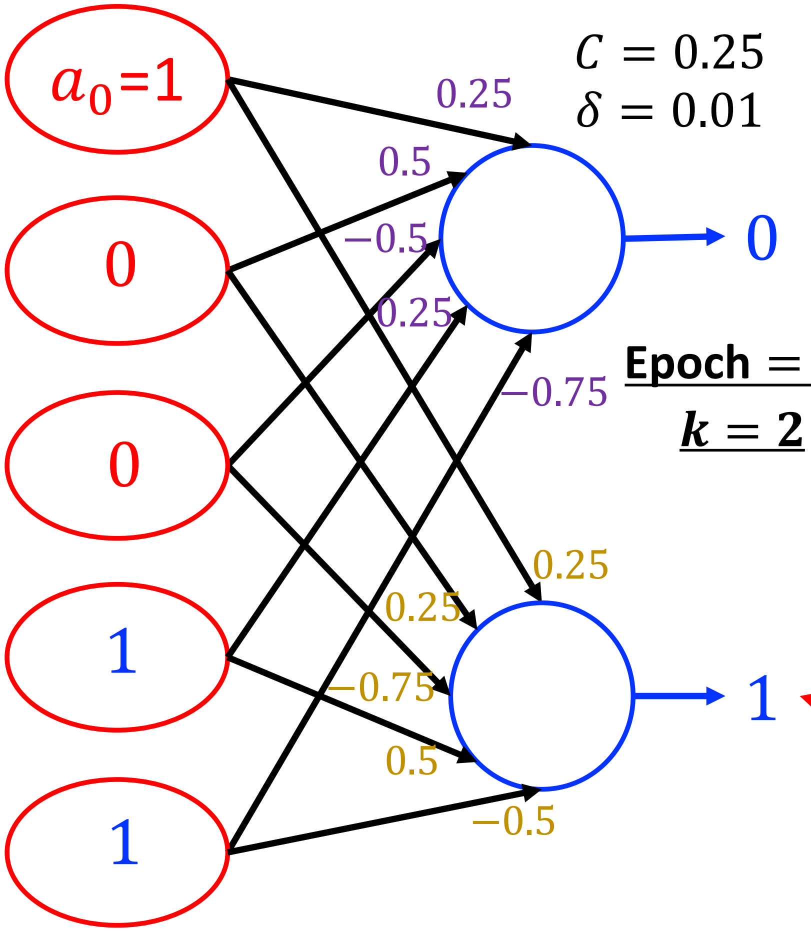
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_2^2 = \sum_{i=0}^{4} w_{2i}^2 a_i^2$$
$$= 0.25\times1 + 0.25\times0 + (-0.75)\times0 + 0.5\times1 + (-0.5)\times1$$
$$= 0.25 \geq 0$$

$$X_2^2 = 1$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2
***k* = 2**

Training set: 2 patterns

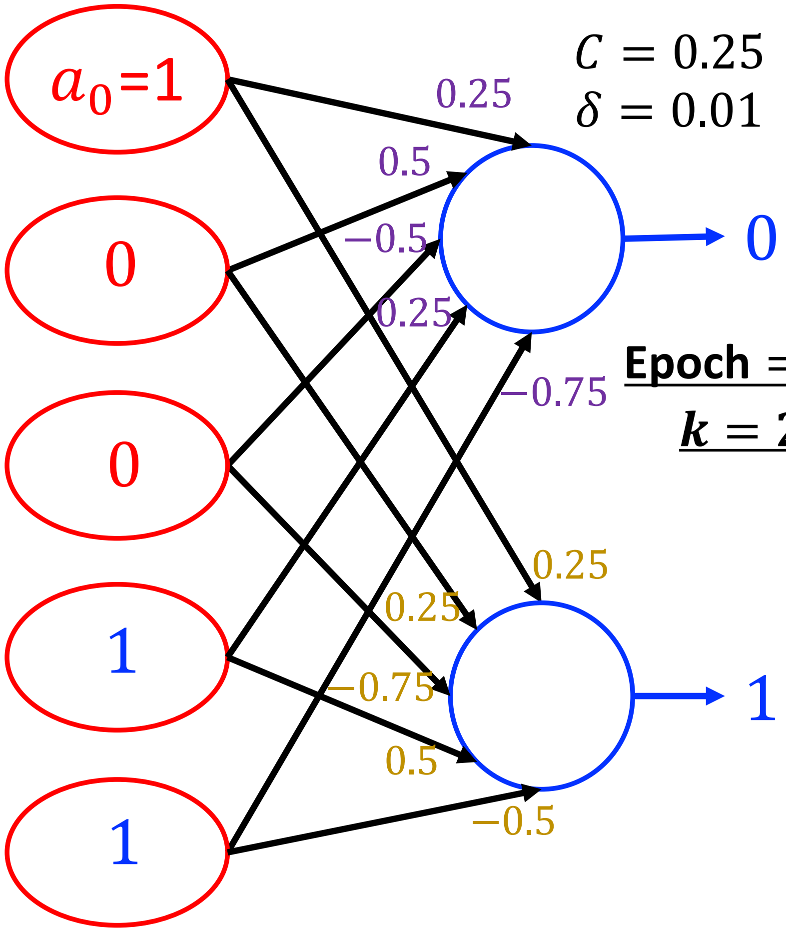
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$S_2^2 = \sum_{i=0}^{4} w_{2i}^2 a_i^2$$
$$= 0.25\times1 + 0.25\times0 + (-0.75)\times0 + 0.5\times1 + (-0.5)\times1$$
$$= 0.25 \geq 0$$

$X_2^2 = 1$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

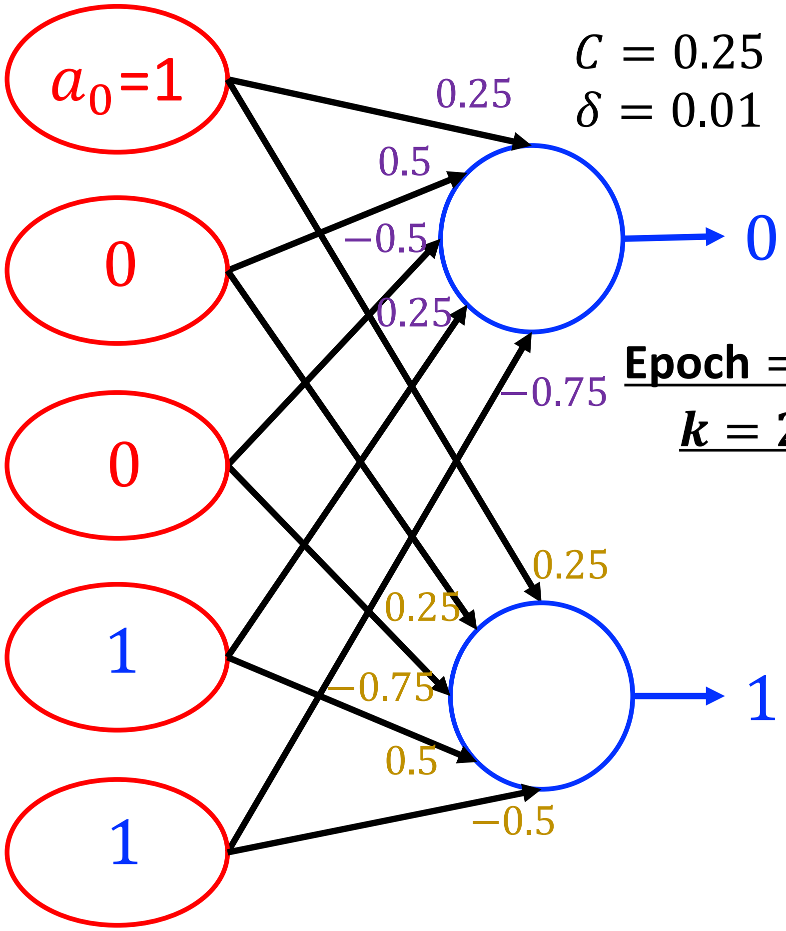
w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$e_1^2 = (t_1^2 - X_1^2) = (0 - 0) = 0$

$\Delta w_{1i}^2 = C e_1^2 a_i^2$

$$\Delta w_{10}^2 = C e_1^2 a_0^2 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{11}^2 = C e_1^2 a_1^2 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{12}^2 = C e_1^2 a_2^2 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{13}^2 = C e_1^2 a_3^2 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{14}^2 = C e_1^2 a_4^2 = 0.25 \times 0 \times 1 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2
***k* = 2**

Training set: 2 patterns

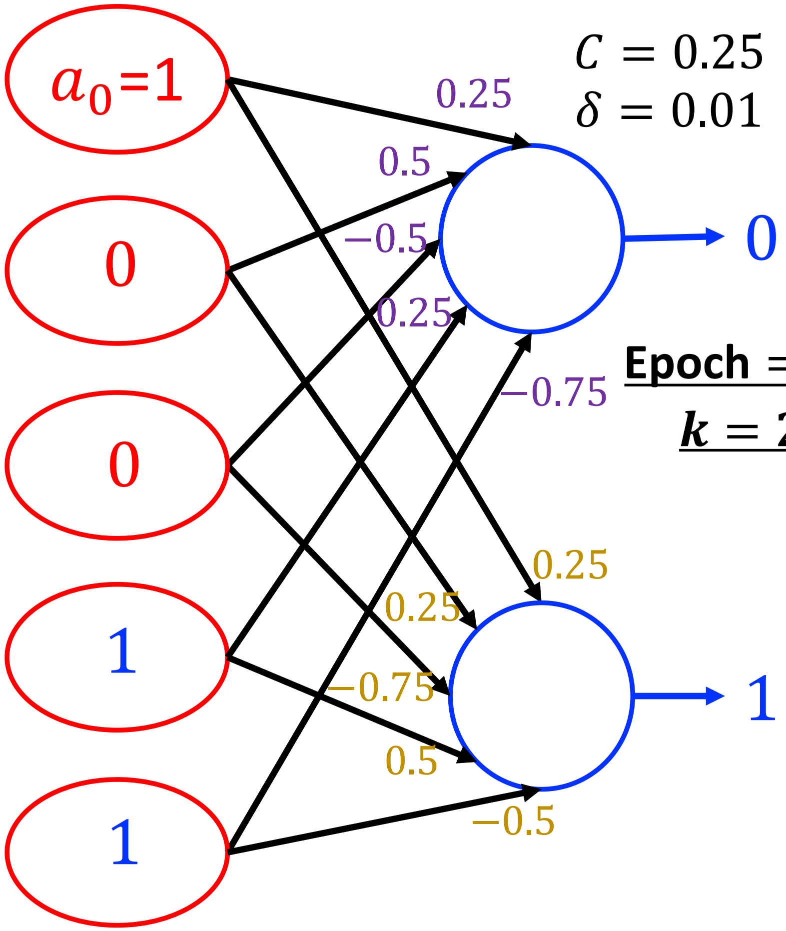
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$e_2^2 = (t_2^2 - X_2^2) = (1 - 1) = 0$ $\quad \Delta w_{2i}^2 = Ce_2^2 a_i^2$

$$\Delta w_{20}^2 = Ce_2^2 a_0^2 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{21}^2 = Ce_2^2 a_1^2 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{22}^2 = Ce_2^2 a_2^2 = 0.25 \times 0 \times 0 = 0$$
$$\Delta w_{23}^2 = Ce_2^2 a_3^2 = 0.25 \times 0 \times 1 = 0$$
$$\Delta w_{24}^2 = Ce_2^2 a_4^2 = 0.25 \times 0 \times 1 = 0$$

A Running Example



$C = 0.25$
$\delta = 0.01$

<u>**Epoch = 2**</u>

<u>$\boldsymbol{k = 2}$</u>

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

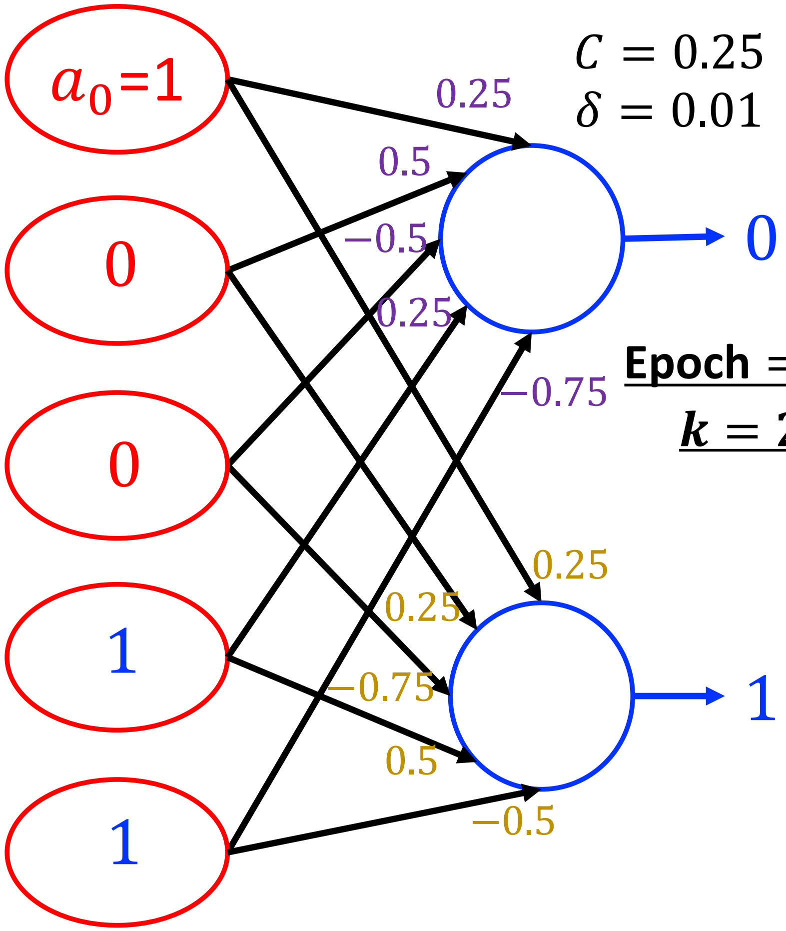
$e_1^2 = -1$

$\Delta w_{10}^2 = 0$
$\Delta w_{11}^2 = 0$
$\Delta w_{12}^2 = 0$
$\Delta w_{13}^2 = 0$
$\Delta w_{14}^2 = 0$

$e_2^2 = 0$

$\Delta w_{20}^2 = 0$
$\Delta w_{21}^2 = 0$
$\Delta w_{22}^2 = 0$
$\Delta w_{23}^2 = 0$
$\Delta w_{24}^2 = 0$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2
***k* = 2**

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$\Delta w_{10}^2 = 0$

$\Delta w_{11}^2 = 0$

$\Delta w_{12}^2 = 0$

$\Delta w_{13}^2 = 0$

$\Delta w_{14}^2 = 0$

$\Delta w_{20}^2 = 0$

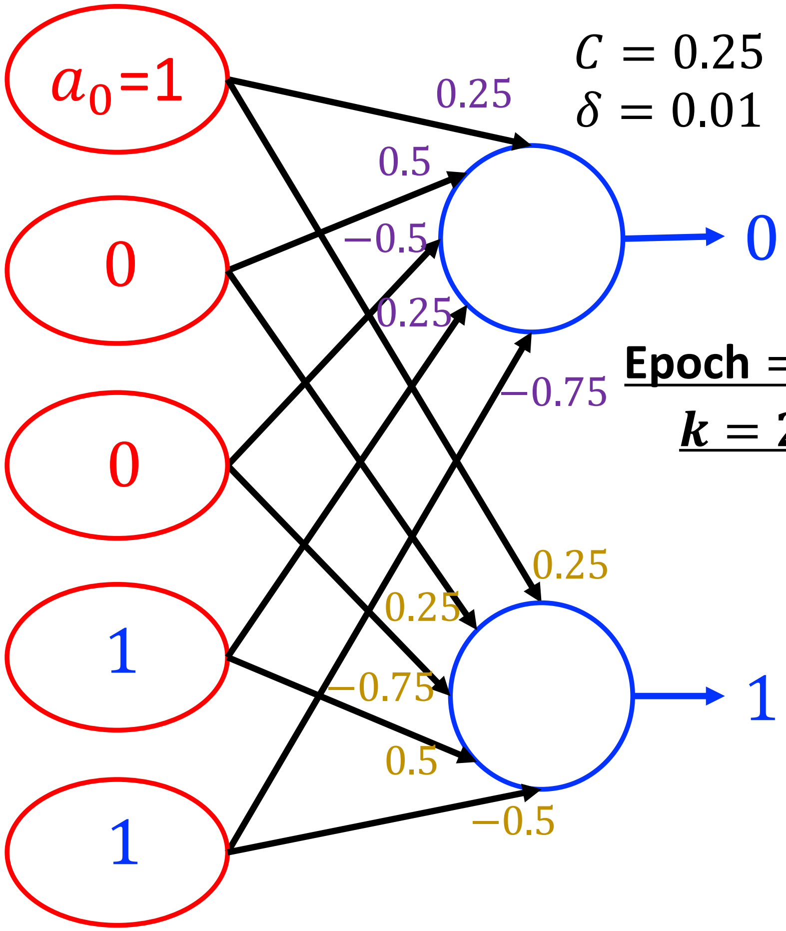
$\Delta w_{21}^2 = 0$

$\Delta w_{22}^2 = 0$

$\Delta w_{23}^2 = 0$

$\Delta w_{24}^2 = 0$

A Running Example



$C = 0.25$
$\delta = 0.01$

Epoch = 2

$\boldsymbol{k = 2}$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1	a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
1	1	0	0	1	0	0	0	1	1	0	1

w_{10}^1	w_{11}^1	w_{12}^1	w_{13}^1	w_{14}^1	w_{20}^1	w_{21}^1	w_{22}^1	w_{23}^1	w_{24}^1
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

$$\boxed{w_{ji}^k = w_{ji}^k + \Delta w_{ji}^k}$$

$w_{10}^2 = w_{10}^2 + \Delta w_{10}^2 = 0.25$

$w_{11}^2 = w_{11}^2 + \Delta w_{11}^2 = 0.5$

$w_{12}^2 = w_{12}^2 + \Delta w_{12}^2 = -0.5$

$w_{13}^2 = w_{13}^2 + \Delta w_{13}^2 = 0.25$

$w_{13}^2 = w_{13}^2 + \Delta w_{14}^2 = -0.75$

$w_{20}^2 = w_{20}^2 + \Delta w_{20}^2 = 0.25$

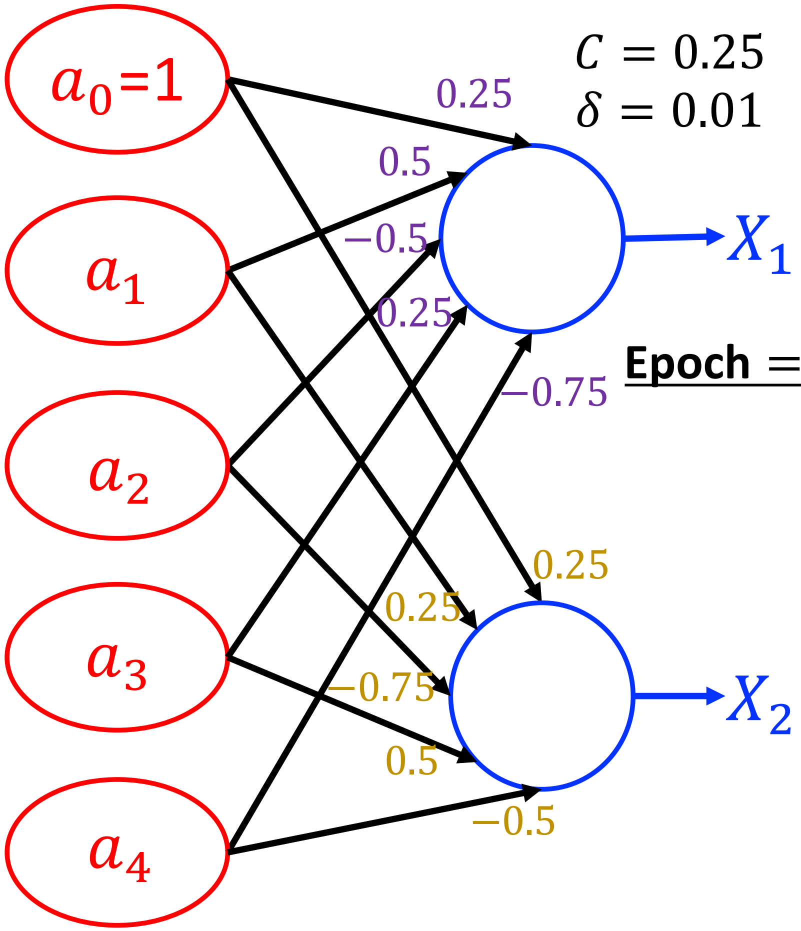
$w_{21}^2 = w_{21}^2 + \Delta w_{21}^2 = 0.25$

$w_{22}^2 = w_{22}^2 + \Delta w_{22}^2 = -0.75$

$w_{23}^2 = w_{23}^2 + \Delta w_{23}^2 = 0.5$

$w_{23}^2 = w_{13}^2 + \Delta w_{24}^2 = -0.5$

A Running Example



$C = 0.25$
$\delta = 0.01$

Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

Epoch = 2

w_{10}^2	w_{11}^2	w_{12}^2	w_{13}^2	w_{14}^2	w_{20}^2	w_{21}^2	w_{22}^2	w_{23}^2	w_{24}^2
0.25	0.5	-0.5	0.25	-0.75	0.25	0.25	-0.75	0.5	-0.5

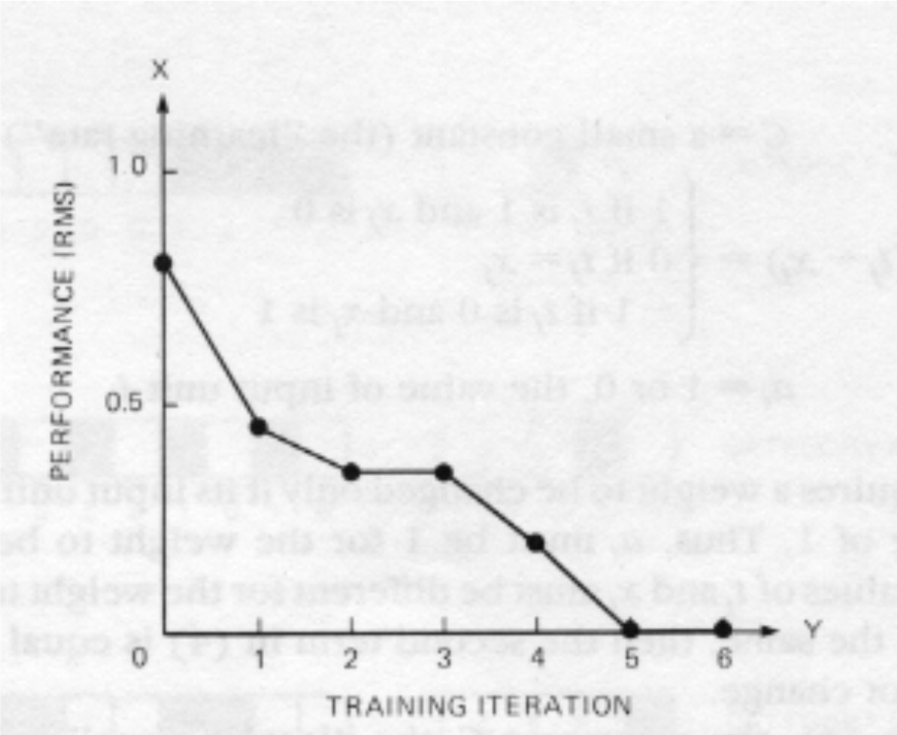
Epoch 2 finishes. Evaluate performance.

$$\text{RMS} = \sqrt{\frac{\sum_{k=1}^{r}\sum_{j=1}^{m}\left(e_j^k\right)^2}{rm}} = \sqrt{\frac{\sum_{k=1}^{2}\sum_{j=1}^{2}\left(e_j^k\right)^2}{2\times 2}}$$

$$= \sqrt{\frac{(e_1^1)^2 + (e_2^1)^2 + (e_1^2)^2 + (e_2^2)^2}{2\times 2}} = 0 < \delta = 0.01$$

Great! Stop.

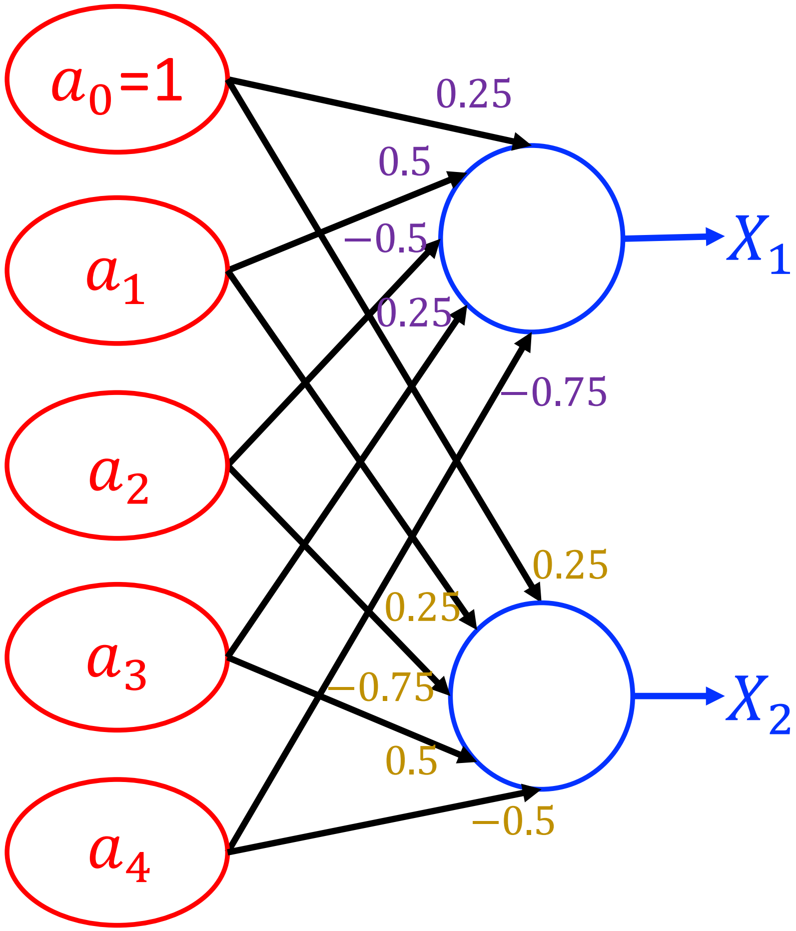
Network Performance



Learning curve: dependency of the RMS error on the number of iterations.

- ***Initially***, the adaptable ***weights are all set to small random values***, and the network does not perform very well;
- Performance improves **during training**;
- Finally, the error gets close to zero, training stops. We say the network has **converged**.

A Running Example

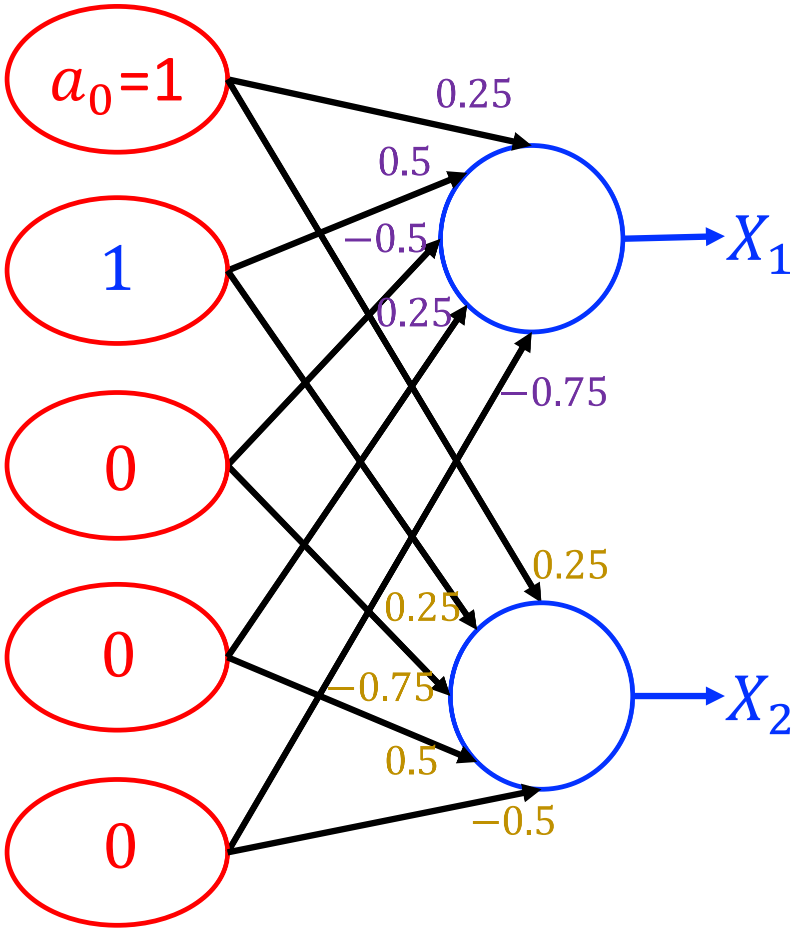


Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

A Running Example



Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

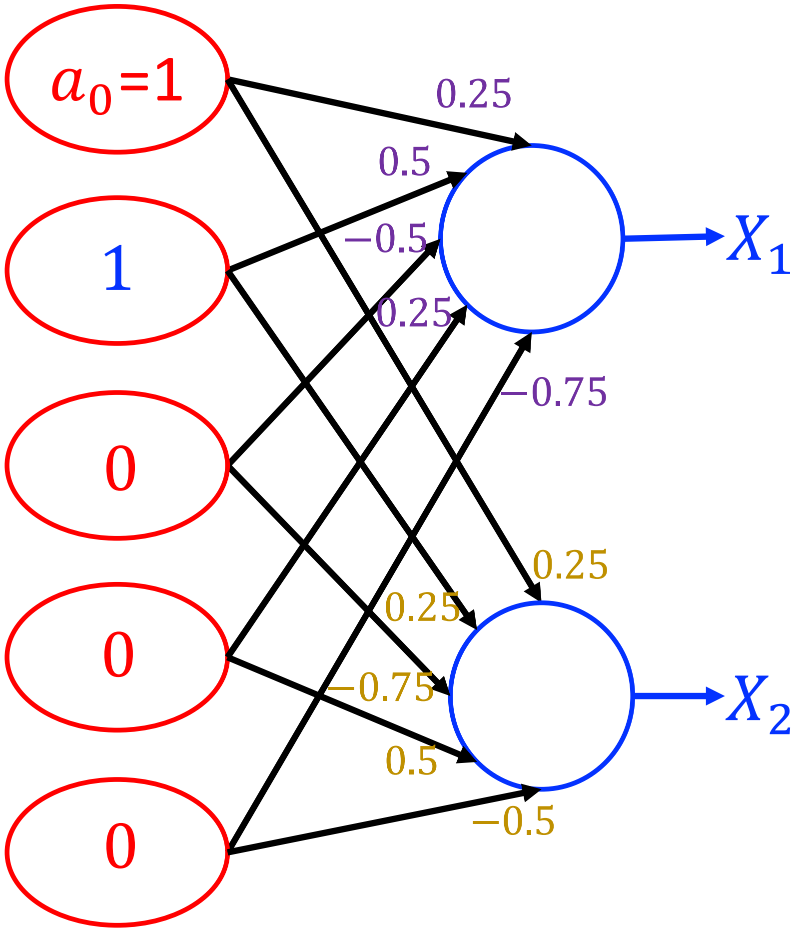
a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

Test set: 1 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	0	0	0	1	0

TEST

A Running Example



Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

Test set: 1 patterns

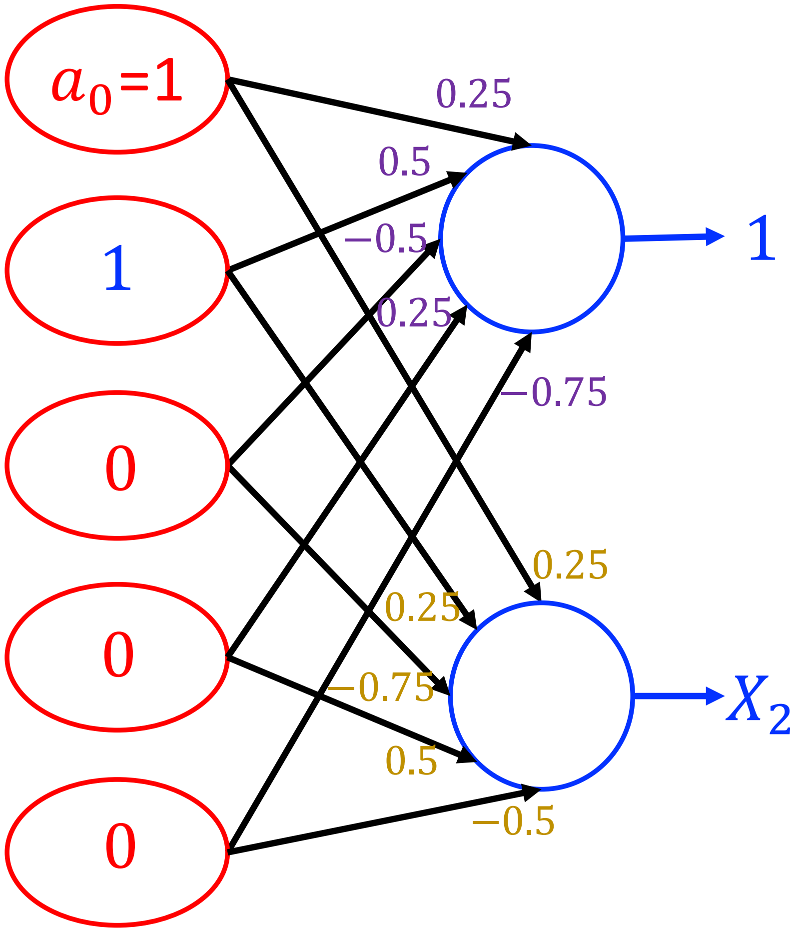
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	0	0	0	1	0

TEST

$$S_1 = \sum_{i=0}^{4} w_{1i} a_i$$
$$= 0.25\times1 + 0.5\times1 + (-0.5)\times1 + 0.25\times0 + (-0.75)\times0$$
$$= 0.75 > 0$$

$$X_1 = 1$$

A Running Example



Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

Test set: 1 patterns

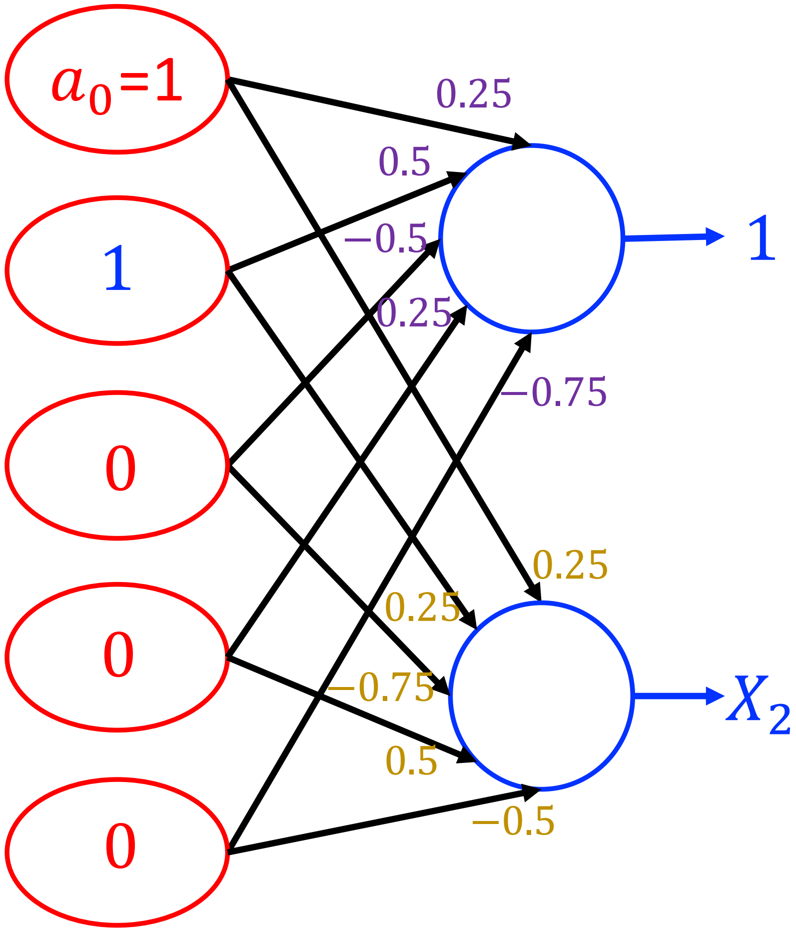
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	0	0	0	1	0

TEST

$$S_1 = \sum_{i=0}^{4} w_{1i} a_i$$
$$= 0.25\times1 + 0.5\times1 + (-0.5)\times1 + 0.25\times0 + (-0.75)\times0$$
$$= 0.75 > 0$$

$$X_1 = 1$$

A Running Example



Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

Test set: 1 patterns

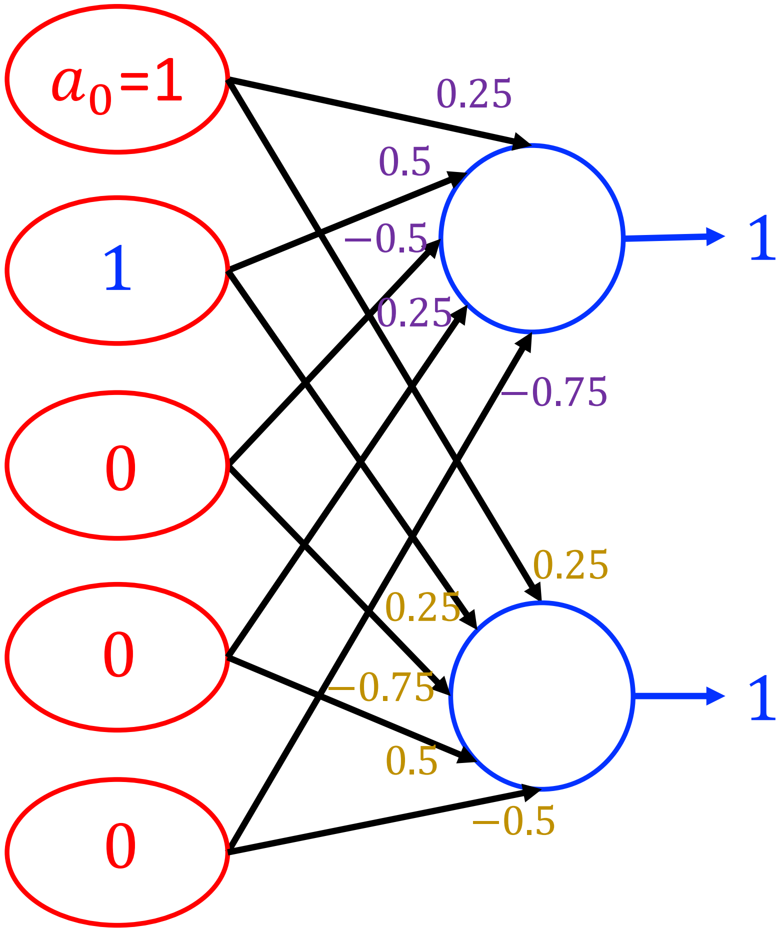
a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	0	0	0	1	0

TEST

$$S_2 = \sum_{i=0}^{4} w_{2i} a_i$$
$$= 0.25\times1 + 0.25\times1 + (-0.75)\times0 + 0.5\times0 + (-0.5)\times0$$
$$= 0.5 > 0$$

$$X_2 = 1$$

A Running Example



Training set: 2 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	1	0	0	1	0

a_1^2	a_2^2	a_3^2	a_4^2	t_1^2	t_2^2
0	0	1	1	0	1

Test set: 1 patterns

a_1^1	a_2^1	a_3^1	a_4^1	t_1^1	t_2^1
1	0	0	0	1	0

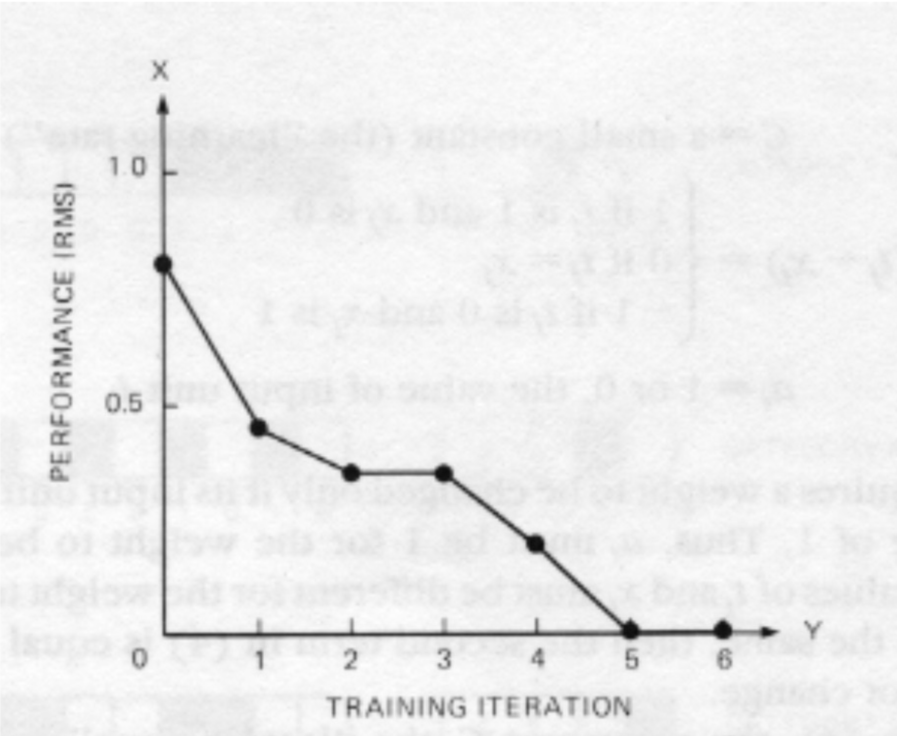
TEST

The network after training does not necessarily work well on a test set.

$$S_2 = \sum_{i=0}^{4} w_{2i} a_i$$
$$= 0.25\times1 + 0.25\times1 + (-0.75)\times0 + 0.5\times0 + (-0.5)\times0$$
$$= 0.5 > 0$$

$$X_2 = 1$$

Network Performance



Q: Does the learning algorithm always converge?

Learning curve: dependency of the RMS error on the number of iterations.

- ***Initially***, the adaptable ***weights are all set to small random values***, and the network does not perform very well;
- Performance improves **during training**;
- Finally, the error gets close to zero, training stops. We say the network has **converged**.