$$10^{11}$$

$$10^{14}$$

$$10^{15}$$

n

$$a = \sum_{j=1}^{n} w_j x_j + b$$

b

x_j

w_j

$$w_j \begin{cases} > 0 & \text{excitatory input} \\ < 0 & \text{inhibitory input} \\ = 0 & \text{no connection} \end{cases}$$

$$x_0 = 1$$

$$w_0 = b$$

$$a = \sum_{j=1}^{n} w_j x_j + b = \sum_{j=0}^{n} w_j x_j = \mathbf{w}^T \mathbf{x}$$

$$y = g(a) = g\left(\sum_{j=0}^{n} w_j x_j + b\right) = g(\mathbf{w}^T \mathbf{x} + b)$$

$$\mathbf{x} = [x_0 = 1,\ x_1, \cdots, x_n]^T$$

$$\mathbf{w} = [w_0 = b, \ , w_1, \cdots, w_n]^T$$

$$g(x)$$

$$g(x,a) = \frac{1}{1+e^{-ax}} = \frac{e^{ax}}{1+e^{ax}} = \begin{cases} 0 & x = -\infty \\ 1/2 & x = 0 \\ 1 & x = \infty \end{cases} \quad , \qquad \frac{d\,g(x)}{dx} = \frac{a\,e^{-ax}}{(1+e^{-ax})^2}$$

$$g(x,a) = \frac{2}{1+e^{-ax}} - 1 = \frac{e^{ax}-1}{e^{ax}+1} = \begin{cases} -1 & x = -\infty \\ 0 & x = 0 \\ 1 & x = \infty \end{cases} \quad , \qquad \frac{d\,g(x)}{dx} = \frac{2a\,e^{-ax}}{(1+e^{-ax})^2}$$

a

$$a \to 0$$

$$g(x,a)$$

$$a \to \infty$$

$$\lim_{a\to\infty} g(x,a) = \begin{cases} 0 \text{ or } -1 & x<0 \\ 1 & x>0 \end{cases}$$

$$g(x) = \max(0,\, x) = \begin{cases} 0 & x < 0 \\ x & x > 0 \end{cases}$$

N

$$\{(\mathbf{x}_n, \mathbf{y}_n),\ n = 1, \cdots, N\}$$

$$f : \mathbf{x} \in \mathcal{R}^d \Longrightarrow \mathbf{y} \in \mathcal{R}^m$$

$$\mathbf{x}_n$$

$$\mathbf{y}_n$$

$$E = mc^2$$

$$\{\mathbf{x}_1, \cdots, \mathbf{x}_N\}$$

$$\mathbf{x} \in \mathcal{R}^d$$

$y \in \mathcal{R}$

$$y = f(\mathbf{x})$$

$$\{(\mathbf{x}_n,\ y_n),\ n = 1, \cdots, N\}$$

X

$$\{C_1, \cdots, C_K\}$$

$$f : \mathbf{x} \in \mathcal{R}^d \Longrightarrow y \in \{C_1, \cdots, C_K\}$$

$$F \to S$$

$$F \cap B \to S$$

$$B \to S$$

$$\{\mathbf{x}_1, \cdots, \mathbf{x}_K\}$$

$$\{\mathbf{y}_1, \cdots, \mathbf{y}_K\}$$

$$\mathbf{x} = [x_1, \cdots, x_n]^T$$

m

$$\mathbf{y} = [y_1, \cdots, y_m]^T$$

$$y_i = \sum_{j=1}^{n} w_{ij} x_j = \mathbf{w}_i^T \mathbf{x} \quad (i = 1, \cdots, m),$$

$$\mathbf{y} = \mathbf{W}\mathbf{x}$$

$$\mathbf{w}_i = [w_{i1}, \cdots, w_{in}]^T$$

$$\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_m]^T$$

$$w_{ij}^{new} = w_{ij}^{old} + \eta \, x_j y_i \quad (i = 1, \cdots, m, \ j = 1, \cdots, n)$$

$$\mathbf{W}^{new} = \mathbf{W}^{old} + \eta\ \mathbf{y}\mathbf{x}^T$$

η

y_i

w_{ij}

$$w_{ij} = 0$$

$$\eta = 1$$

$$\{(\mathbf{x}_k, \mathbf{y}_k), \quad k = 1, \cdots, N\}$$

$$w_{ij} = \sum_{k=1}^{N} x_j^{(k)} y_i^{(k)} \quad (i = 1, \cdots, m, \ \ j = 1, \cdots, n)$$

K

$$\mathbf{W}_{m\times n}=\sum_{k=1}^{N}\mathbf{y}_k\mathbf{x}_k^T=\sum_{k=1}^{N}\begin{bmatrix} y_1^{(k)} \\ \vdots \\ y_m^{(k)} \end{bmatrix}[x_1^{(k)},\cdots,x_n^{(k)}]$$

$\mathbf{x}_l$

$$\mathbf{y} = \mathbf{W}\mathbf{x}_l = \left(\sum_{k=1}^{K} \mathbf{y}_k \mathbf{x}_k^T \right) \ \mathbf{x}_l = \mathbf{y}_l(\mathbf{x}_l^T \mathbf{x}_l) + \sum_{k \neq l} \mathbf{y}_k(\mathbf{x}_k^T \mathbf{x}_l)$$

y

$$\mathbf{x}_i^T \mathbf{x}_j = ||\mathbf{x}_i||\ ||\mathbf{x}_j|| \cos\phi = \delta_{ij} = \left\{ \begin{array}{cc} 1 & i = j \\ 0 & i \neq j \end{array} \right.$$

ϕ

$\mathbf{x}_i$

$$\mathbf{x}_j$$

$$\cos\phi = \frac{\mathbf{x}^T\mathbf{y}}{||\mathbf{x}||\,||\mathbf{y}||} = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2}\sqrt{\sum_i y_i^2}}$$

$$r_{xy} = \sigma_{xy}^2 / \sigma_x \sigma_y$$

$$0 < r_{xy} \leq 1$$

$$r_{xy} = 1$$

$$\mathbf{x} = \mathbf{y}$$

$$r_{xy} = 0$$

$$-1 \leq r_{xy} \leq 0$$

$$r_{xy} = -1$$

$$\mathbf{x} = -\mathbf{y}$$

$$n \geq N$$

$$\mathbf{y} = \mathbf{y}_l(\mathbf{x}_l^T\mathbf{x}_l) + \sum_{k \neq l} \mathbf{y}_k(\mathbf{x}_k^T\mathbf{x}_l) = \mathbf{y}_l$$

$k \neq l$

$$\mathbf{x}_k^T \mathbf{x}_l = 0$$

$$\mathbf{x}_k$$

$\mathbf{y}_k$

$$k = 1, \cdots, N$$

$$\mathbf{y} - \mathbf{y}_l \neq 0$$

$$\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_K]$$

$$\mathbf{x}_k = [x_1, \cdots, x_d]^T$$

$$x_i \in \{-1,\ 1\},\ (i = 1, \cdots, d)$$

W

d

$$\mathbf{W}_{d\times d} = \frac{1}{d}\sum_{k=1}^{K}\mathbf{x}_k\mathbf{x}_k^T = \frac{1}{d}\sum_{k=1}^{K}\left[\begin{array}{c} x_1^{(k)} \\ \vdots \\ x_d^{(k)} \end{array}\right][x_1^{(k)},\cdots,x_d^{(k)}]$$

$$w_{ij} = \frac{1}{d} \sum_{k=1}^{K} x_i^{(k)} x_j^{(k)} = w_{ji}$$

$$w_{ii} = 0 \ (i = 1, \cdots, d)$$

$$x_i^{(n+1)} = sgn\left(\sum_{j=1}^{d} w_{ij} x_j^{(n)}\right) = \begin{cases} +1, & \text{if} \quad \sum_{j=1}^{d} w_{ij} x_j^{(n)} \geq 0 \\ -1, & \text{if} \quad \sum_{j=1}^{d} w_{ij} x_j^{(n)} < 0 \end{cases}$$

$$x_i^{(n)}$$

$$x_i^{(n+1)}$$

x_i

$$e_{ij} = -w_{ij} x_i x_j$$

$$\mathcal{E}(\mathbf{x}) = \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} e_{ij} = -\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} w_{ij}x_i x_j = -\frac{1}{2}\mathbf{x}^T\mathbf{W}\mathbf{x}$$

	x_j	x_i	$w_{ij} > 0$	$w_{ij} < 0$
1	-1	-1	$e_{ij} < 0$	$e_{ij} > 0$
2	-1	1	$e_{ij} > 0$	$e_{ij} < 0$
3	1	-1	$e_{ij} > 0$	$e_{ij} < 0$
4	1	1	$e_{ij} < 0$	$e_{ij} > 0$

$e_{ij} < 0$

$w_{ij} > 0$

$$x_j = \mp 1$$

干1

$w_{ij} < 0$

± 1

$$e_{ij} > 0$$

e_{ij}

$\mathcal{E}(\mathbf{x})$

x_k

$$x_k^{(n+1)} \neq x_k^{(n)}$$

$$x_k = \pm 1$$

$$x_k^{(n+1)} = \mp 1$$

$$x_{l\neq k}^{(n+1)} = x_{l\neq k}^{(n)}$$

$$\mathcal{E}^{(n)}(\mathbf{x})$$

$=$

$$-\frac{1}{2}\left[\sum_{i\neq k}\sum_{j\neq k} w_{ij}x_i^{(n)}x_j^{(n)} + \sum_i w_{ik}x_i^{(n)}x_k^{(n)} + \sum_j w_{kj}x_k^{(n)}x_j^{(n)}\right]$$

$$-\frac{1}{2}\sum_{i\neq k}\sum_{j\neq k} w_{ij}x_i x_j - \sum_i w_{ik}x_i^{(n)}x_k^{(n)}$$

$$\mathcal{E}^{(n+1)}(\mathbf{x}) = -\frac{1}{2}\sum_{i \neq k}\sum_{j \neq k} w_{ij} x_i x_j - \sum_i w_{ik} x_i^{(n+1)} x_k^{(n+1)}$$

$$\triangle\mathcal{E} = (x_k^{(n)} - x_k^{(n+1)}) \sum_i w_{ik} x_i$$

$$x_k^{(n)} = -1$$

$$\sum_i w_{ik} x_i \geq 0$$

$$x_k^{(n+1)} = 1$$

$$(x_k^{(n)} - x_k^{(n+1)}) = -2 \leq 0$$

$$\triangle\mathcal{E} \leq 0$$

$$x_k^{(n)} = 1$$

$$\sum_i w_{ik} x_i < 0$$

$$x_k^{(n+1)} = -1$$

$$(x_k^{(n)} - x_k^{(n+1)}) = 2 \geq 0$$

$$\triangle\mathcal{E}(\mathbf{x}) \leq 0$$

$$\mathcal{E}(\mathbf{x}) = -\frac{1}{2}\mathbf{x}^T\mathbf{W}\mathbf{x} = -\frac{1}{2}\mathbf{x}^T\left[\frac{1}{d}\sum_{k=1}^{K}\mathbf{x}_k\mathbf{x}_k^T\right]\mathbf{x} = -\frac{1}{2d}\sum_{k=1}^{K}\mathbf{x}^T\mathbf{x}_k\mathbf{x}_k^T\mathbf{x} = -\frac{1}{2d}\sum_{k=1}^{K}(\mathbf{x}^T\mathbf{x}_k)^2$$

$$\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N]$$

$$\mathbf{y} = [y_1, \cdots, y_N]^T$$

$$m = 1$$

C_+

C_-

$m > 1$

$$\mathbf{x} = [x_1, \cdots, x_d]^T$$

$$y = 1$$

$$\mathbf{x} \in C_+$$

$$y = -1$$

$\mathbf{x} \in C_{-}$

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b$$

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b \begin{cases} > 0 \\ < 0 \end{cases},$$

$$\hat{y} = \operatorname{sign}(f(\mathbf{x})) = \begin{cases} 1 & \text{for } \mathbf{x} \in C_+ \\ -1 & \text{for } \mathbf{x} \in C_- \end{cases}$$

$$\hat{y} = 1$$

$$\hat{y} = -1$$

$$\delta = y - \hat{y}$$

y

$\hat{y}$

W

$$||\mathbf{w}||$$

$$p_{\mathbf{w}}(\mathbf{x}) = \frac{\mathbf{w}^T\mathbf{x}}{||\mathbf{w}||} > -\frac{b}{||\mathbf{w}||} = b' \qquad p_{\mathbf{w}}(\mathbf{x}) = \frac{\mathbf{w}^T\mathbf{x}}{||\mathbf{w}||} < -\frac{b}{||\mathbf{w}||} = b'$$

$$p_{\mathbf{w}}(\mathbf{x})$$

$$b' = -b/||\mathbf{w}||$$

$$\mathbf{w}^T\mathbf{x} + b = 0$$

$$r = \delta = y - \hat{y}$$

δ

$$\mathbf{x} = [x_0 = 1, x_1, \cdots, x_n]^T$$

$$\mathbf{w} = [w_0 = b, w_1, \cdots, w_n]^T$$

$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x}$$

$$\mathbf{w}^{new} = \mathbf{w}^{old} + \eta(y - \hat{y})\,\mathbf{x} = \mathbf{w}^{old} + \eta\,\delta\mathbf{x}$$

$\eta > 0$

$y \pm 1$

$$\hat{y} = \mathrm{sign}(f(\mathbf{x}))$$

	y	$f(\mathbf{x}) = \mathbf{w}^T\mathbf{x},\ \hat{y} = \text{sign}(f(\mathbf{x}))$	$\delta = y - \hat{y}$
1	$y = 1$	$f(\mathbf{x}) > 0,\ \hat{y} = 1$	$\delta = 0$
2	$y = -1$	$f(\mathbf{x}) > 0,\ \hat{y} = 1$	$\delta = -2$
3	$y = 1$	$f(\mathbf{x}) < 0,\ \hat{y} = -1$	$\delta = 2$
4	$y = -1$	$f(\mathbf{x}) < 0,\ \hat{y} = -1$	$\delta = 0$

$$y f(\mathbf{x}) > 0$$

$$\delta = y - \hat{y} = 0$$

$$\mathbf{w}^{new} = \mathbf{w}^{old} + \delta\mathbf{x} = \mathbf{w}^{old}$$

$$y f(\mathbf{x}) < 0$$

$$\delta = y - \hat{y} = 1$$

$$y\, f(\mathbf{x}) < 0$$

$$\delta = -2$$

$$\mathbf{w}^{new} = \mathbf{w}^{old} - 2\mathbf{x}$$

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{w}^{new} = \mathbf{x}^T \mathbf{w}^{old} - 2\mathbf{x}^T \mathbf{x} < \mathbf{x}^T \mathbf{w}^{old}$$

$$\delta = 2$$

$$\mathbf{w}^{new} = \mathbf{w}^{old} + 2\mathbf{x}$$

$$f(\mathbf{x}) = \mathbf{x}^T \mathbf{w}^{new} = \mathbf{x}^T \mathbf{x}^{old} + 2\mathbf{x}^T \mathbf{x} > \mathbf{x}^T \mathbf{w}^{old}$$

$$\mathbf{w}^{new} = \mathbf{w}^{old} + 2y\mathbf{x}$$

2

$$y\, f(\mathbf{x}) \begin{cases} > 0 \\ < 0 \end{cases}$$

$$\begin{cases} \mathbf{w}^{new} = \mathbf{w}^{old} \\ \mathbf{w}^{new} = \mathbf{w}^{old} + y\,\mathbf{x} \end{cases}$$

$$\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_m]$$

$$y_i \in \{-1,\, 1\},\ (i = 1, \cdots, m)$$

$$K = m$$

$$y_k = 1$$

$$y_l = -1$$

$$l \neq k$$

$$K = 2^m$$

$$K = 4$$

4

binary output	$\mathbf{y}_0$	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$
y_1	1	-1	-1	-1
y_2	-1	1	-1	-1
y_3	-1	-1	1	-1
y_4	-1	-1	-1	1

binary output	$\mathbf{y}_0$	$\mathbf{y}_1$	$\mathbf{y}_2$	$\mathbf{y}_3$
y_1	-1	-1	1	1
y_2	-1	1	-1	1

$$\hat{\mathbf{y}} = [\hat{y}_1, \cdots, \hat{y}_m]^T$$

$$\delta = ||\mathbf{y} - \hat{\mathbf{y}}||$$

$$\hat{\mathbf{y}} = [-1\ 1\ \ -1\ 1]^T$$

X

W

$d + 1$

$$n = 2$$

$m = 2$

$$2^2 = 4$$

$$\mathbf{y} = [y_1,\, y_2]$$

$$\mathbf{w}_1$$

$$\mathbf{w}_2$$

$$\delta = y - \hat{y} \neq 0$$

$\delta \neq 0$

$n = 3$

$m = 3$

$$2^3 = 8$$

$$\mathbf{w}_3$$

$$\mathbf{w}^{new} = \mathbf{w}^{old} + y_n \mathbf{x}_n$$

y_n

$$\delta = y_n - \hat{y}_n \neq 0$$

$$\mathbf{w} = 0$$

$$\mathbf{w} = \sum_{n=1}^{N} \alpha_n y_n \mathbf{x}_n$$

α_n

y_l

$$f(\mathbf{x}_l) = \mathbf{w}^T \mathbf{x}_l = \sum_{n=1}^{N} \alpha_n y_n \mathbf{x}_n^T \mathbf{x}_l,$$

$$\hat{y}_l = \mathrm{sign}(f(\mathbf{x}_l))$$

$$\delta = y_l - \hat{y}_l \neq 0$$

$$\mathbf{w}^{new} = \mathbf{w}^{old} + y_l\, \mathbf{x}_l = \sum_{n=1}^{N} \alpha_n y_n \mathbf{x}_n + y_l\, \mathbf{x}_l$$

$$\alpha_l^{new} = \alpha_l^{old} + 1$$

$\{\alpha_1, \cdots, \alpha_N\}$

$$\mathbf{w}^T\mathbf{x} = \sum_{n=1}^{N} \alpha_n y_n \mathbf{x}_n^T \mathbf{x} \left\{ \begin{array}{l} > 0 \\ < 0 \end{array} \right. ,$$

$$\mathbf{x} \in \begin{cases} C_- \\ C_+ \end{cases}$$

$$\mathbf{x}_n^T \mathbf{x}_m$$

$$k(\mathbf{x}_n, \mathbf{x}_m)$$

$$\{(\mathbf{x}_i, \mathbf{y}_i),\ i = 1, ..., N\}$$

$\mathbf{y}_i$

l

$\hat{\mathbf{y}}$

$$||\mathbf{y} - \hat{\mathbf{y}}||$$

$\mathcal{E}$

$$z_j = g(a_j^h) = g\left(\sum_{k=1}^{d} w_{jk}^h x_k + b_j^h\right) = g\left(\sum_{k=0}^{d} w_{jk}^h x_k\right) = g(\mathbf{x}^T \mathbf{w}_j^h) \qquad (j = 1, \cdots, l)$$

$$\mathbf{x} = [x_0 = 1, x_1, \cdots, x_d]^T$$

$$\mathbf{w}_j^h = [w_{j0}^h = b_j^h, w_{j1}^h, \cdots, w_{jd}^h]^T$$

$$a_j^h = \mathbf{x}^T \mathbf{w}_j^h$$

$$\mathbf{z} = \mathbf{g}\left(\mathbf{W}_h^T \mathbf{x}\right)$$

$$\mathbf{z} = [z_1, \cdots, z_l]^T$$

$$\mathbf{W}_h = [\mathbf{w}_1^h, \cdots, \mathbf{w}_l^h]$$

$$\hat{y}_i = g(a_i^o) = g\left(\sum_{j=1}^{l} w_{ij}^o z_j + b_i^o\right) = g\left(\sum_{j=0}^{l} w_{ij}^o z_j\right) = g(\mathbf{z}^T \mathbf{w}_i^o) \qquad (i = 1, \cdots, m)$$

$$\mathbf{z} = [z_0 = 1, z_1, \cdots, z_l]^T$$

$$\mathbf{w}_i^o = [w_{i0}^o = b_i^o, w_{i1}^o, \cdots, w_{il}^o]^T$$

$$a_i^o = \mathbf{z}^T \mathbf{w}_i^o$$

$$\hat{\mathbf{y}} = \mathbf{g}\left(\mathbf{W}_o^T \mathbf{z}\right)$$

$$\mathbf{W}_o = [\mathbf{w}_1^o, \cdots, \mathbf{w}_m^o]$$

$$\hat{\mathbf{y}} = [\hat{y}_1, \cdots, y_m]^T$$

$$w_{ij}^{o} \ (i = 1, \cdots, m, \ j = 0, 1, \cdots, l)$$

$$w_{jk}^{h} \ (j = 1, \cdots, l, \ k = 0, 1, \cdots, d)$$

$\mathcal{E}$

$$\frac{1}{2}||\mathbf{y}-\hat{\mathbf{y}}||^2 = \frac{1}{2}\sum_{i=1}^{m}(\hat{y}_i - y_i)^2 = \frac{1}{2}\sum_{i=1}^{m}[g(a_i^o) - y_i]^2 = \frac{1}{2}\sum_{i=1}^{m}\left[g\left(\sum_{j=0}^{l} w_{ij}^o z_j\right) - y_i\right]^2$$

$$\frac{1}{2}\sum_{i=1}^{m}\left[g\left(w_{i0}^{o}+\sum_{j=1}^{l}w_{ij}^{o}\,g(a_j^h)\right)-y_i\right]^2=\frac{1}{2}\sum_{i=1}^{m}\left[g\left(w_{i0}^{o}+\sum_{j=1}^{l}w_{ij}^{o}\,g\left(\sum_{k=0}^{n}w_{jk}^{h}x_k\right)\right)-y_i\right]^2$$

$$w_{ij}^{o} \quad (i = 1, \cdots, m,\ j = 0, 1, \cdots, l)$$

$$\frac{\partial\,\varepsilon}{\partial\,w_{ij}^{o}} = \frac{\partial\,\varepsilon}{\partial\,\hat{y}_i}\;\frac{\partial\,\hat{y}_i}{\partial\,a_i^{o}}\;\frac{\partial\,a_i^{o}}{\partial\,w_{ij}^{o}} = (\hat{y}_i - y_i)\;g'(a_i^{o})\;z_j = -\delta_i\;g'(a_i^{o})\;z_j$$

$$\delta_i = y_i - \hat{y}_i$$

$w_{ij}^{o} \ (i=1,\cdots,m,\ j=0,1,\cdots,l)$

$$w_{ij}^{o(new)} = w_{ij}^{o(old)} - \eta \, \frac{\partial e}{\partial w_{ij}^{o}} = w_{ij}^{o(old)} + \eta \; \delta_i \; g'(a_i^o) \; z_j$$

$$\mathbf{W}^{o(new)}_{m\times(l+1)} = \mathbf{W}^{o(old)}_{m\times(l+1)} + \eta\ \mathbf{d}^{o}_{m\times 1}(\mathbf{z}^{T})_{1\times(l+1)}$$

$$\mathbf{d}_{m\times 1}^{o} = [\delta_1 g'(a_1^o), \cdots, \delta_m g'(a_m^o)]^T$$

$$[\delta_1, \cdots, \delta_m]^T$$

$$[g'(a_1^o), \cdots, g'(a_m^o)]^T$$

Z

$$w_{jk}^h \quad (j = 1, \cdots, l,\ k = 0, 1, \cdots, d)$$

$$\frac{\partial \varepsilon}{\partial w^h_{jk}}$$

$$\frac{\partial \varepsilon}{\partial \hat{y}_i} \frac{\partial \hat{y}_i}{\partial a_i^o} \frac{\partial a_i^o}{\partial z_j} \frac{\partial z_j}{\partial a_j^h} \frac{\partial a_j^h}{\partial w_{jk}^h} = \sum_{i=1}^{m} (\hat{y}_i - y_i) \frac{\partial \hat{y}_i}{\partial a_i^o} \frac{\partial a_i^o}{\partial z_j} \frac{\partial z_j}{\partial a_j^h} \frac{\partial a_j^h}{\partial w_{jk}^h}$$

$$-\sum_{i=1}^{m} \delta_i g'(a_i^o) w_{ij}^o \; g'(a_j^h) x_k = -\delta_j^h \; g'(a_j^h) x_k$$

$$\delta_j^h = \sum_{i=1}^{m} \delta_i^o g'(a_i^o) w_{ij}^o = \sum_{i=1}^{m} d_i^o w_{ij}^o \quad (j = 1, \cdots, l)$$

$$d_i^o = \delta_i^o g'(a_i^o)$$

$$\begin{bmatrix} \delta_1^h \\ \vdots \\ \delta_l^h \end{bmatrix} = \begin{bmatrix} w_{11}^o & \cdots & w_{1m}^o \\ \vdots & \ddots & \vdots \\ w_{l1}^o & \cdots & w_{lm}^o \end{bmatrix} \begin{bmatrix} \delta_1^o g'(a_1^o) \\ \vdots \\ \delta_m^o g'(a_m^o) \end{bmatrix} = \mathbf{W}_{l\times m}^{oT} \mathbf{d}_{m\times 1}^o$$

$\mathbf{W}^{o}$

$l \times m$

b_j

$$\mathbf{d}^o = [d_1, \cdots, d_m]^T$$

$$\mathbf{y} - \hat{\mathbf{y}}$$

$$\mathbf{g}'(\mathbf{a}^o)$$

$$\mathbf{d}^o = (\mathbf{y} - \hat{\mathbf{y}}) \odot \mathbf{g}'(\mathbf{a}^o)$$

$$w_{jk}^{h} \quad (j=1,\cdots,l,\ k=0,1,\cdots,d)$$

$$w_{jk}^{h(new)} = w_{jk}^{h(old)} - \eta \; \frac{\partial \varepsilon}{\partial w_{jk}} = w_{jk}^{h(old)} + \eta \; \delta_j^h \; g'(a_j^h) x_k = w_{jk}^{h(old)} + \eta \; d_j^h x_k$$

$$d_j^h = \delta_j^h g'(a_j^h)$$

$$\mathbf{W}_{l\times(d+1)}^{h(new)} = \mathbf{W}_{l\times(d+1)}^{h(old)} + \eta\ \mathbf{d}_{l\times 1}^{h}\mathbf{x}_{1\times(d+1)}^{T}$$

$$\mathbf{d}^h = \mathbf{W}^{oT}_{l\times m}\mathbf{d}^o_{m\times 1} \odot \mathbf{g}'(\mathbf{a}^h)$$

$$\mathbf{d}_{l\times 1}^{h}$$

$$\mathbf{x}_{n \times 1}$$

$$[x_1, \cdots, x_n]^T$$

$$\mathbf{x} = [1, x_1, \cdots, x_d]^T$$

$$\mathbf{z} = \mathbf{g}(\mathbf{W}^h \mathbf{x})$$

$$l + 1$$

$$\mathbf{z} \leftarrow [1, \mathbf{z}]$$

$$\hat{\mathbf{y}} = \mathbf{g}(\mathbf{W}^o \mathbf{z});$$

$$\mathbf{d}^h = \mathbf{W}_{l \times m}^{oT} \mathbf{d}^o \odot \mathbf{g}'(\mathbf{a}^h)$$

$$\mathbf{W}^{o}_{m \times l}$$

$$\mathbf{W}^o \leftarrow \mathbf{W}^o + \eta\ \mathbf{d}^o \mathbf{z}^T$$

$$\mathbf{W}^h \leftarrow \mathbf{W}^h + \eta\, \mathbf{d}^h \mathbf{x}^T$$

$$\varepsilon = ||\mathbf{y} - \hat{\mathbf{y}}||^2/2$$

C

Y

$C * K$

$$N + 1$$

M

$$\{(\mathbf{x}_i,\, \mathbf{y}_i),\ (i = 1, \cdots, N)\}$$

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, \mathbf{W}^h, \mathbf{W}^o)$$

$$\mathbf{W}_h$$

$$\mathbf{r} = \mathbf{y} - \hat{\mathbf{y}}$$

13%

11.5%

$$\begin{bmatrix} 185 & 14 & 1 \\ 5 & 181 & 14 \\ 11 & 33 & 156 \end{bmatrix} \qquad \begin{bmatrix} 176 & 24 \\ 22 & 178 \end{bmatrix}$$

$$\begin{bmatrix} 216 & 0 & 2 & 1 & 0 & 0 & 4 & 0 & 1 & 0 \\ 0 & 212 & 0 & 1 & 0 & 1 & 3 & 0 & 7 & 0 \\ 0 & 0 & 221 & 0 & 0 & 0 & 1 & 0 & 2 & 0 \\ 1 & 0 & 1 & 203 & 0 & 0 & 0 & 0 & 14 & 5 \\ 0 & 0 & 0 & 0 & 221 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 & 2 & 214 & 0 & 0 & 6 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 221 & 0 & 2 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 214 & 4 & 2 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 221 & 0 \\ 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 4 & 217 \end{bmatrix}$$

$$\mathbf{y} = [y_1, \cdots, y_K]^T$$

$$y_i \in \{0,\ 1\}$$

$$\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N\}$$

$\mathbf{w}_i$

$$a_i = \sum_{j=1}^{n} w_{ij} x_j = \mathbf{w}_i^T \mathbf{x} = ||\mathbf{w}_i|| \; ||\mathbf{x}|| \; \cos\theta \qquad (i = 1, \cdots, K)$$

θ

$$||\mathbf{x}||^2 = 1$$

$$||\mathbf{w}_i|| = 1$$

$$y_i = \begin{cases} 1 & \text{if } \; a_i = \mathbf{w}_i^T \mathbf{x} = \max_j \mathbf{w}_j^T \mathbf{x} \\ 0 & \text{otherwise} \end{cases} \qquad (i = 1, \cdots, K)$$

$$\mathbf{w}_i^T \mathbf{x}$$

$$d(\mathbf{x},\ \mathbf{w}) = ||\mathbf{w}_i - \mathbf{x}||$$

$$d^2(\mathbf{w}_i, \mathbf{x})$$

$$||\mathbf{w}_i - \mathbf{x}||^2 = (\mathbf{w}_i - \mathbf{x})^T(\mathbf{w}_i - \mathbf{x}) = \mathbf{w}_i^T\mathbf{w} - 2\mathbf{w}_i^T\mathbf{x} + \mathbf{x}^T\mathbf{x}$$

$$||\mathbf{w}_i||^2 + ||\mathbf{x}||^2 - 2\mathbf{w}_i^T\mathbf{x} = ||\mathbf{w}_i||^2 + ||\mathbf{x}||^2 - 2||\mathbf{w}_i||\,||\mathbf{x}||\ \cos\theta$$

$$a_i = \mathbf{w}_i^T \mathbf{x}$$

$$||\mathbf{w}_i - \mathbf{x}||$$

$$y_i = \begin{cases} 1 & \text{if } ||\mathbf{w}_i - \mathbf{x}|| = \min_j ||\mathbf{w}_j - \mathbf{x}|| \\ 0 & \text{otherwise} \end{cases} \qquad (i = 1, \cdots, K)$$

$$\mathbf{w}_i^{new} = \mathbf{w}_i^{old} + y_i\,\eta\,(\mathbf{x} - \mathbf{w}_i^{old}) = \begin{cases} (1-\eta)\mathbf{w}_i^{old} + \eta\mathbf{x} & \text{if } y_i = 1 \\ \mathbf{w}_i^{old} & \text{if } y_i = 0 \end{cases}$$

$0 < \eta < 1$

$$\eta_0 = 0.9$$

α

$\alpha = 0.99$

$$\eta_l = \alpha\eta_{l-1} = \cdots = \alpha^l \eta_0$$

$$\mathbf{w}_i^{new}$$

$$\mathbf{w}_j^{old}$$

$$\mathbf{w}_1^{(0)},\ \mathbf{w}_2^{(0)}, \cdots, \mathbf{w}_m^{(0)}$$

$$l = 1$$

$$a_i = \mathbf{w}_i^T \mathbf{x} \qquad (i = 1, \cdots, m)$$

$$a_j \geq a_i \quad (i = 1, \cdots, n)$$

$$\mathbf{w}_j^{(l+1)} = \mathbf{w}_j^{(l)} + \eta(\mathbf{x} - \mathbf{w}_i^{(l)})$$

$$\eta \leftarrow \alpha\eta$$

$$\mathbf{w}_j^{(l+1)}$$

η_0

$$l \leftarrow l + 1$$

$$a_i = \mathbf{w}_i^T \mathbf{x} + b_i$$

b_i

$1/m$

$$b_i = c\ \left(\frac{1}{K} - \frac{\text{number of winnings of node } i}{\text{total number iterations so far}}\right)$$

$b_i < 0$

$b_i > 0$

c

$$d = 2$$

$$tr\mathbf{S}_T^{-1}\mathbf{S}_B$$

$$K = 5$$

n_k

$$\mathbf{w}_i^{new} = \mathbf{w}_i^{old} + u_{ik}\eta\left(\mathbf{x} - \mathbf{w}_i^{old}\right) \qquad i = 1, \cdots, K$$

u_{ik}

$$u_{ik} = exp\left(-\frac{d_{ik}^2}{2\sigma^2}\right)\begin{cases} = 1 & \text{if } i = k \text{ and } d_{ik} = 0 \\ < 1 & \text{else} \end{cases}$$

σ

d_{ik}

n_i

$$d_{kk} = 0$$

$$u_{kk} = 1$$

$$y_i = \mathbf{w}_i^T \mathbf{x}$$

$$y_k \geq y_i$$

T

$$\mathbf{w} = [w_1,\ w_2]^T$$

$$\mathbf{w}_i^T \mathbf{x}$$

$$d = 3$$

$$N \times N \times 3$$