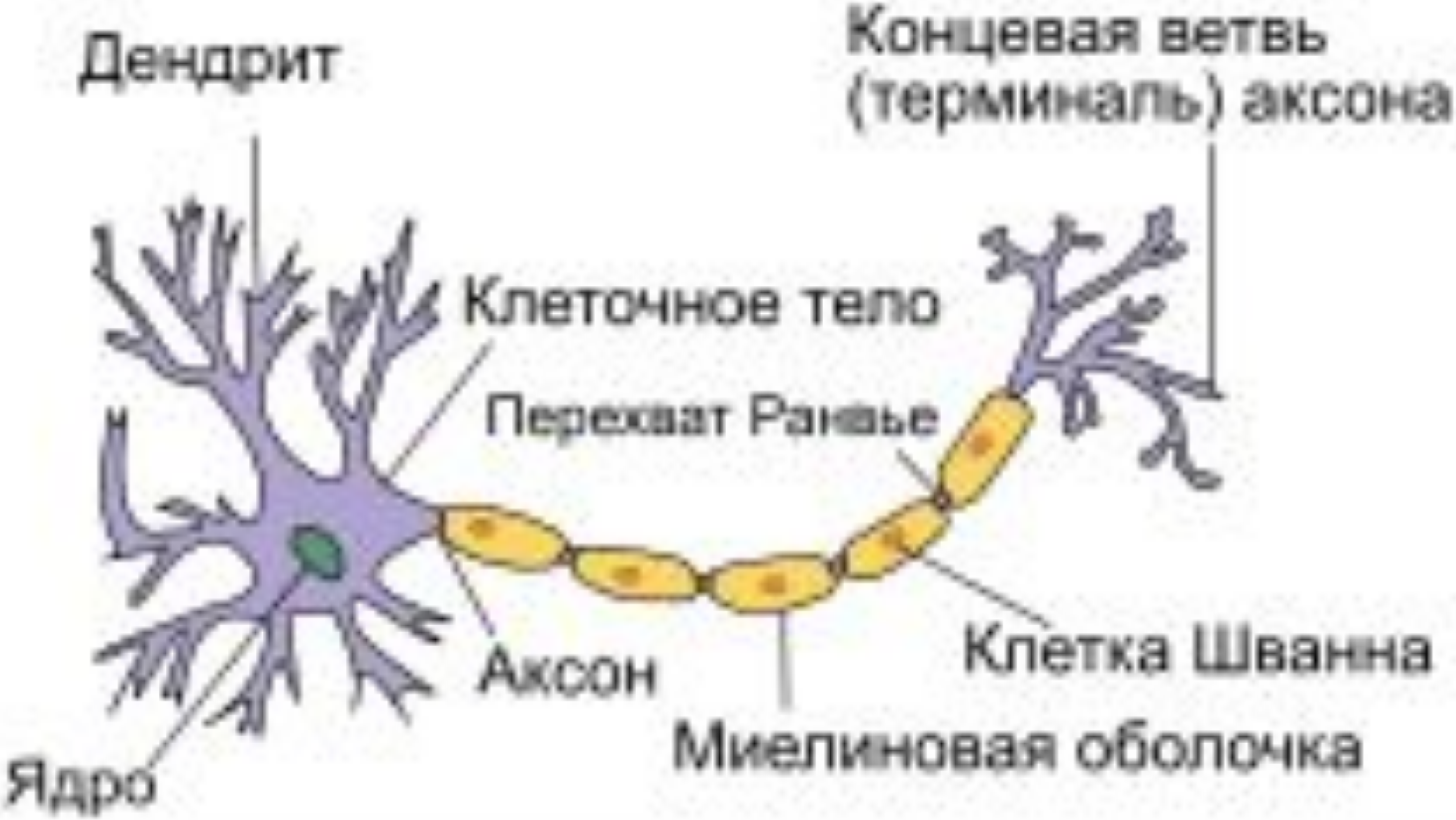
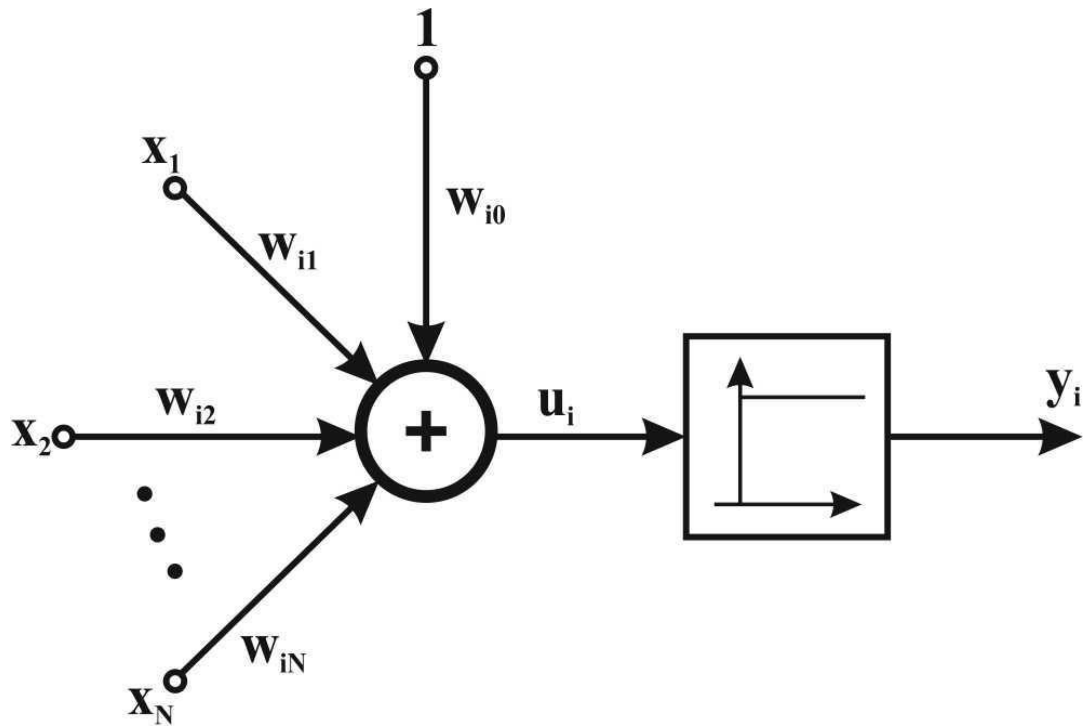
Нейронные сети

Типичная структура нейрона



Структура искусственного нейрона



$x_1, x_2, \ldots x_N$ – входной вектор сигналов;

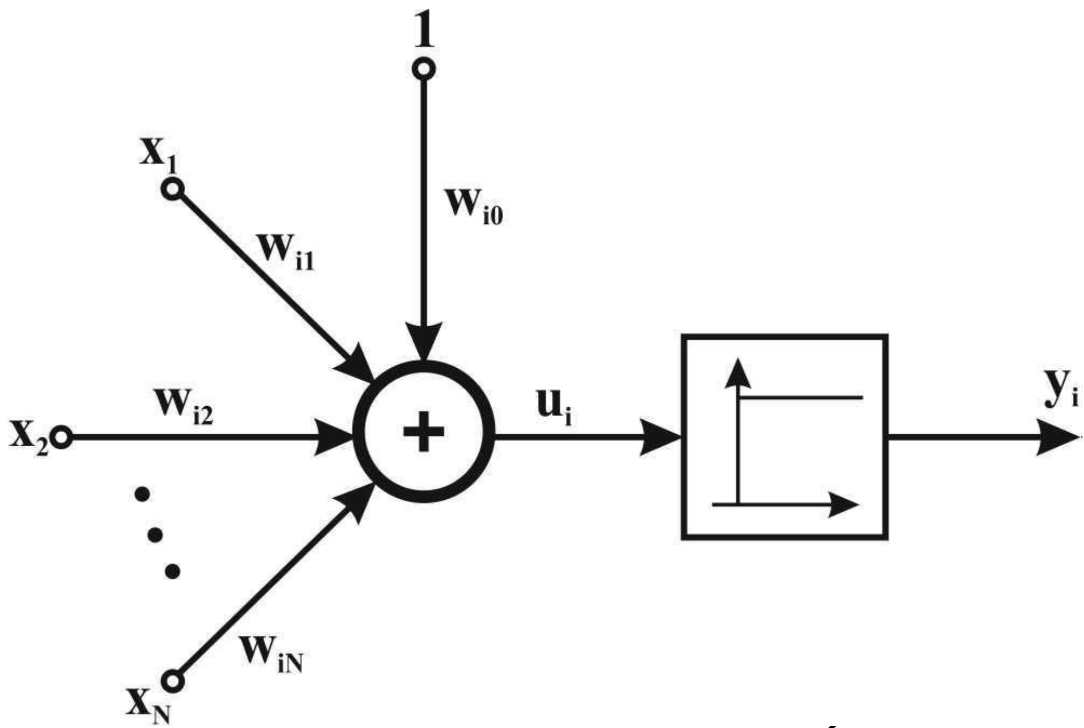
x_0 – входной сигнал поляризатора;

$w_0, w_1, w_2, \ldots w_N$ – вектор весовых коэффициентов;

$u = \sum_{j=0}^{N} x_j * w_j$ - выходное значение сумматора;

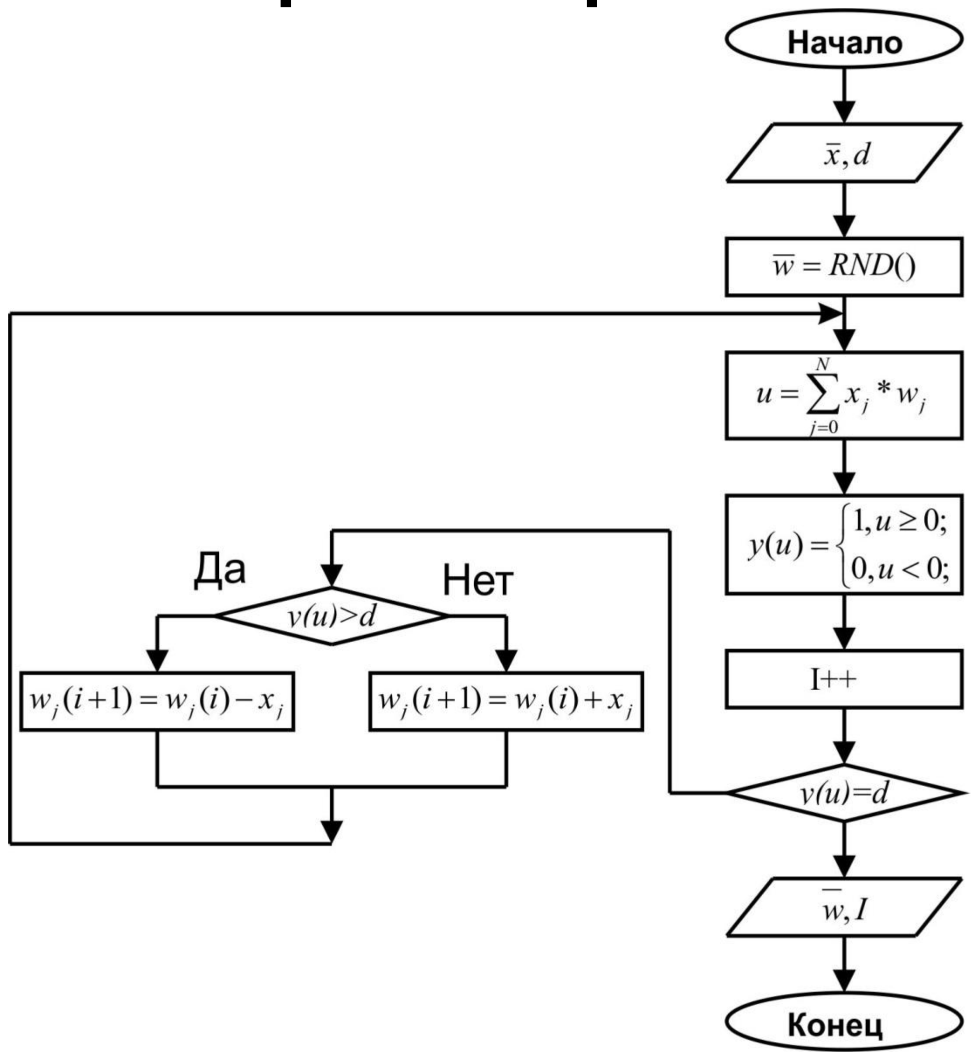
$y=f(u)$ - функция активации нейрона;

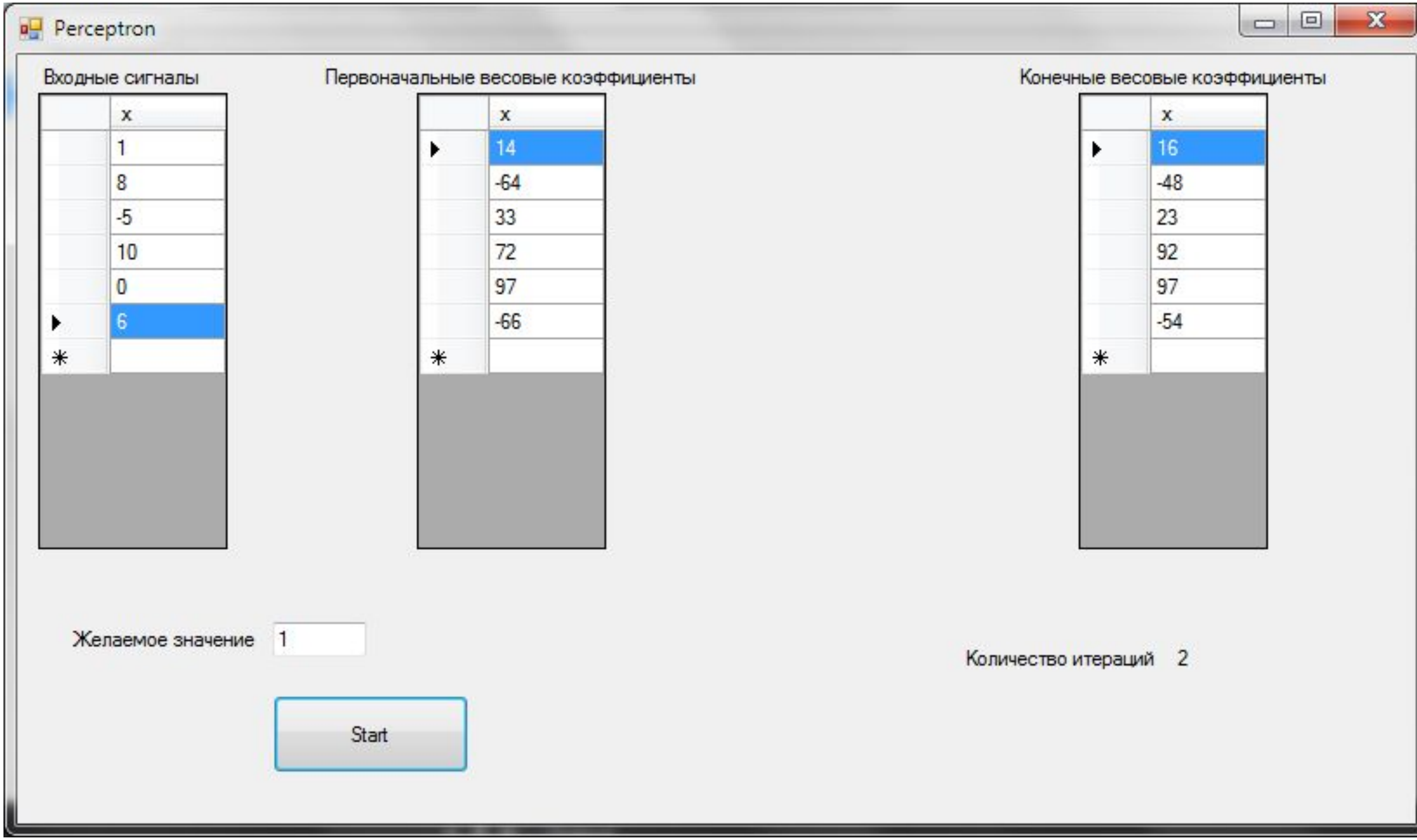
Модель нейронной клетки по МакКаллоку-Питсу



$$y(u) = \begin{cases} 1, u \geq 0; \\ 0, u < 0; \end{cases}$$

Блок-схема обучения персептрона

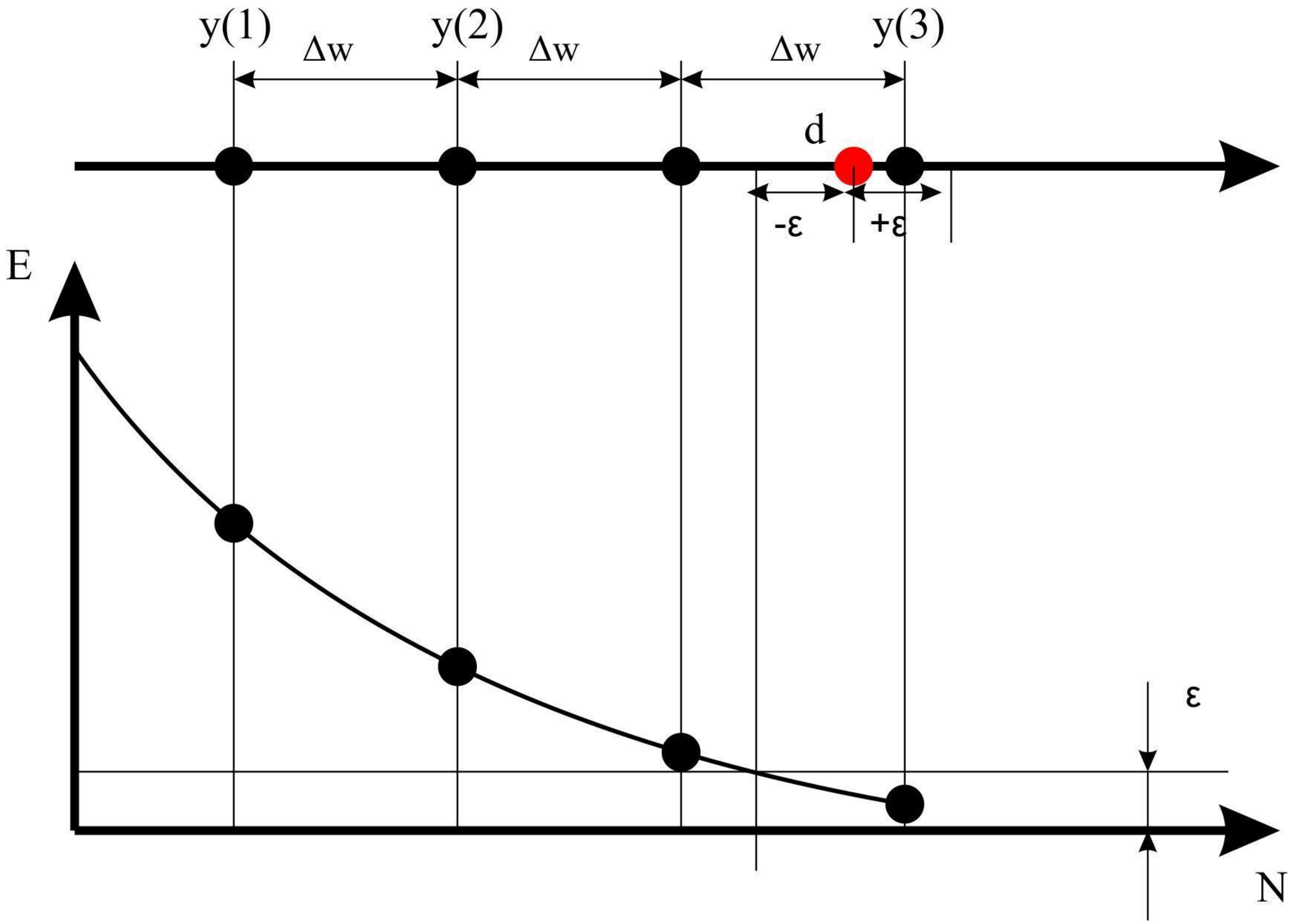



Perceptron
Входные сигналы
x
1
8
-5
10
0
6
Первоначальные весовые коэффициенты
x
14
-64
33
72
97
-66
Конечные весовые коэффициенты
x
16
-48
23
92
97
-54
Желаемое значение
1
Количество итераций 2
Start

Правило Видроу-Хоффа

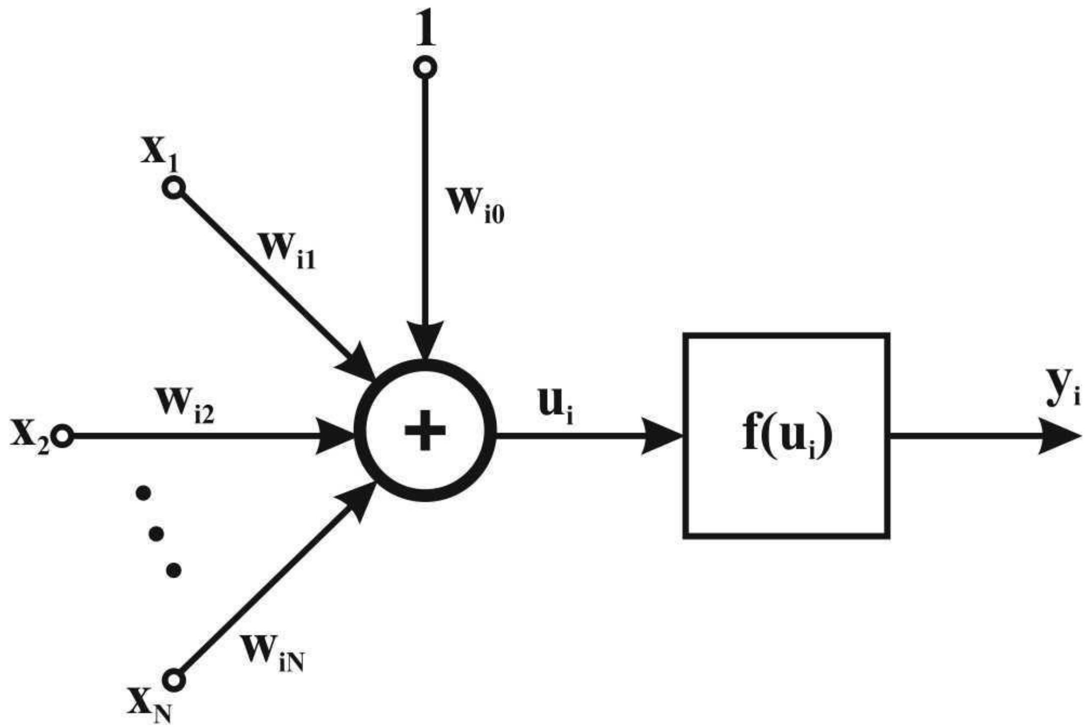
$$w_{i\,j}(t+1) = w_{i\,j}(i) + \Delta w_{ij};$$

$$\Delta w_{ij} = x_j(d_i - y_i);$$



$$E = \sum_{k=1}^{p} (y_i^{(k)} - d_i^{(k)})^2$$

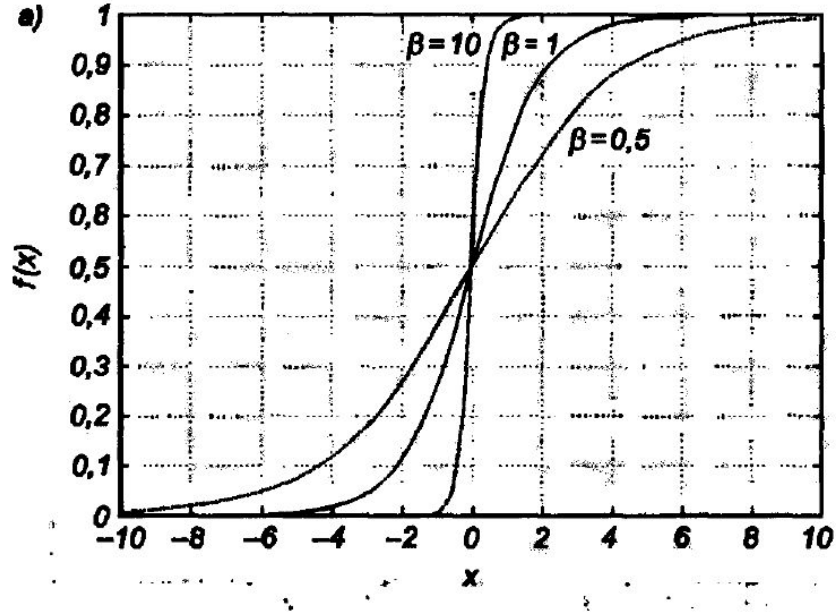
Сигмоидальный нейрон



Функция активации сигмоидального нейрона бывает двух типов:
- униполярная;
- биполярная.

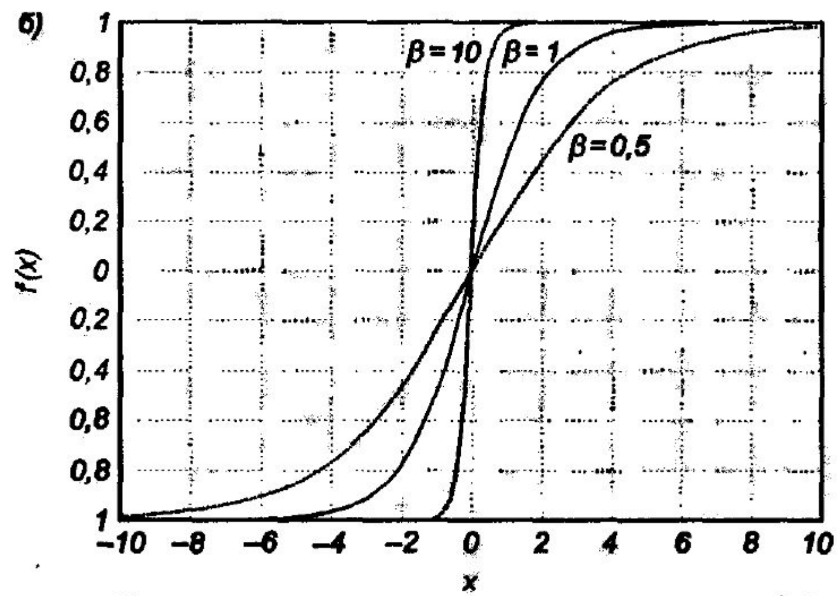
а) униполярная функция

$$f(x) = \frac{1}{1+e^{-\beta x}}$$



б) биполярная функция:

$$f(x) = \tanh(\beta x)$$



а) Дифференциал униполярной функции:

$$\frac{df(x)}{dx} = \beta f(x)(1 - f(x))$$

б) Дифференциал биполярной функции:

$$\frac{df(x)}{dx} = \beta(1 - f^2(x))$$

Сигмоидальный нейрон обучается с учителем путем минимизации целевой функции:

$$E = \frac{1}{2}(y_i - d_i)^2$$

где

$$y_i = f(u_i) = f(\sum_{j=0}^{N} w_{ij} x_j)$$

Градиент целевой функции имеет следующий вид:

$$\nabla_j E = \frac{dE}{dw_{ij}} = e_i x_j \frac{df(u_i)}{du_i}$$

где

$$e_i = y_i - d_i$$

Если ввести обозначение: $\delta_i = e_i \dfrac{df(u_i)}{du_i}$

то градиент целевой функции примет следующий вид:

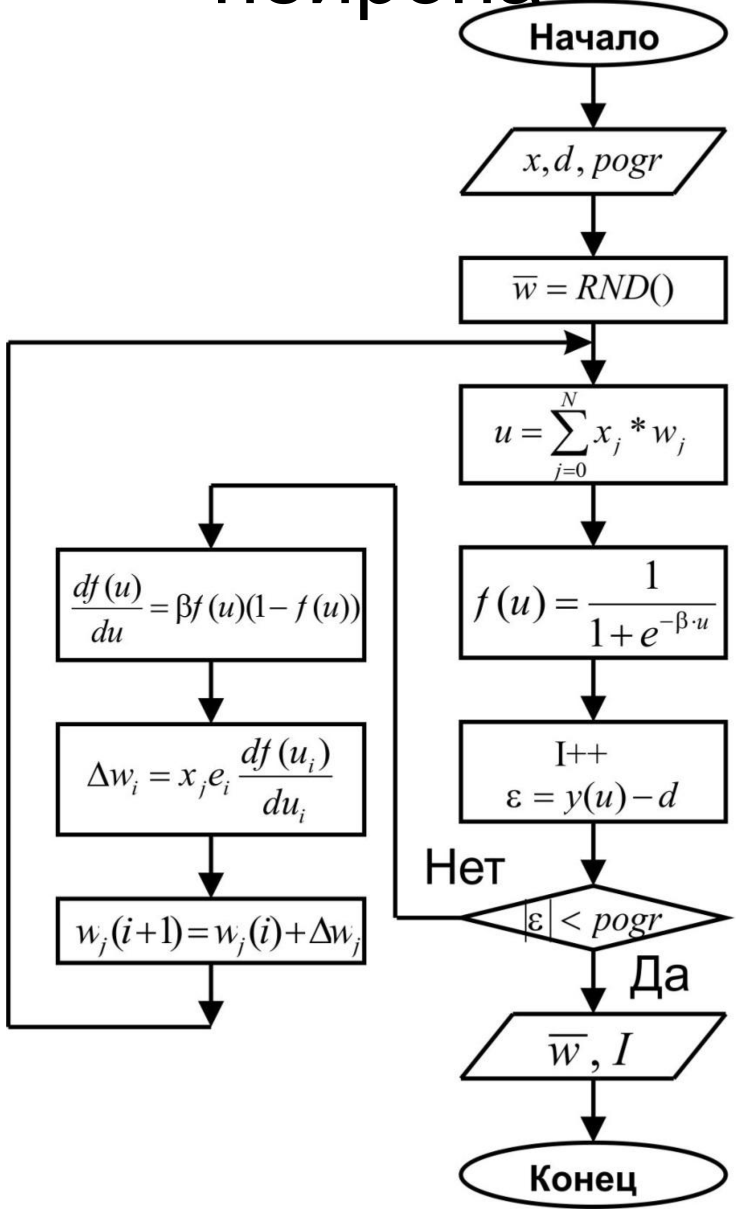
$$\nabla_j E = x_j \delta_i$$

а уточнение весовых коэффициентов производиться по следующей формуле:

$$w_{ij}(t+1) = w_{ij}(i) - \eta \delta_i x_j;$$

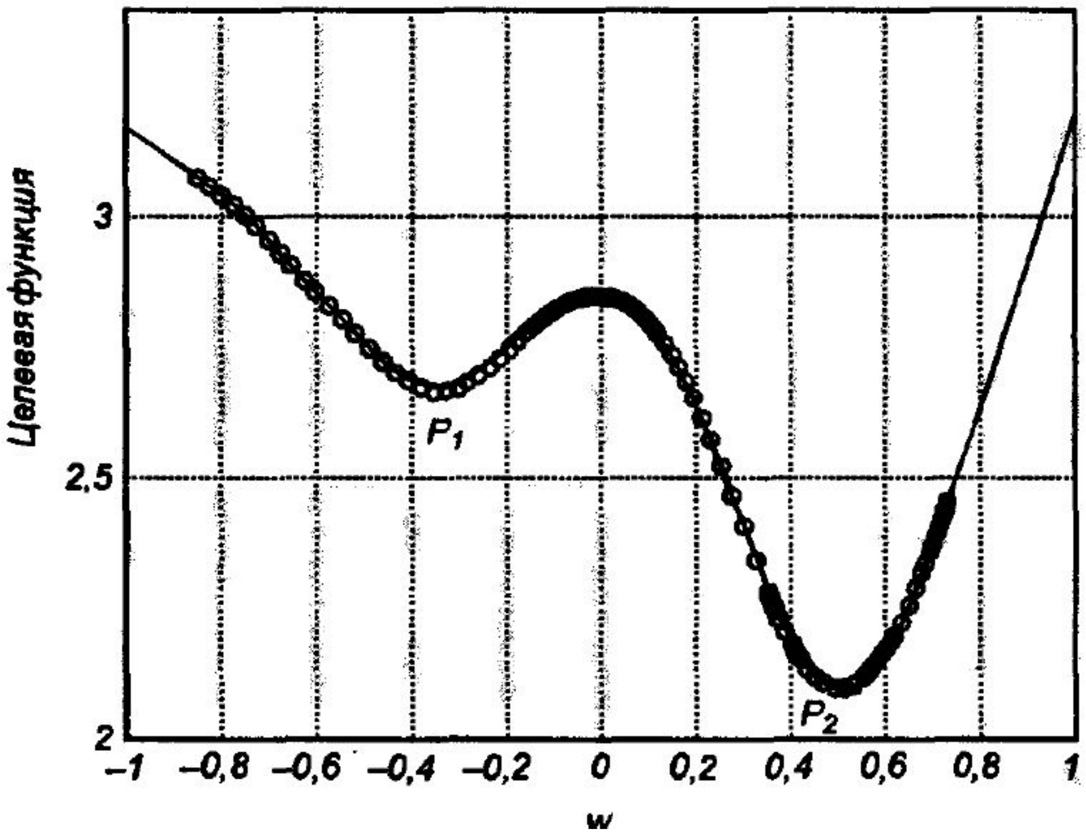
Где η - коэффициент обучения, в диапазоне от 0 до 1.

Блок-схема обучения сигмоидального нейрона

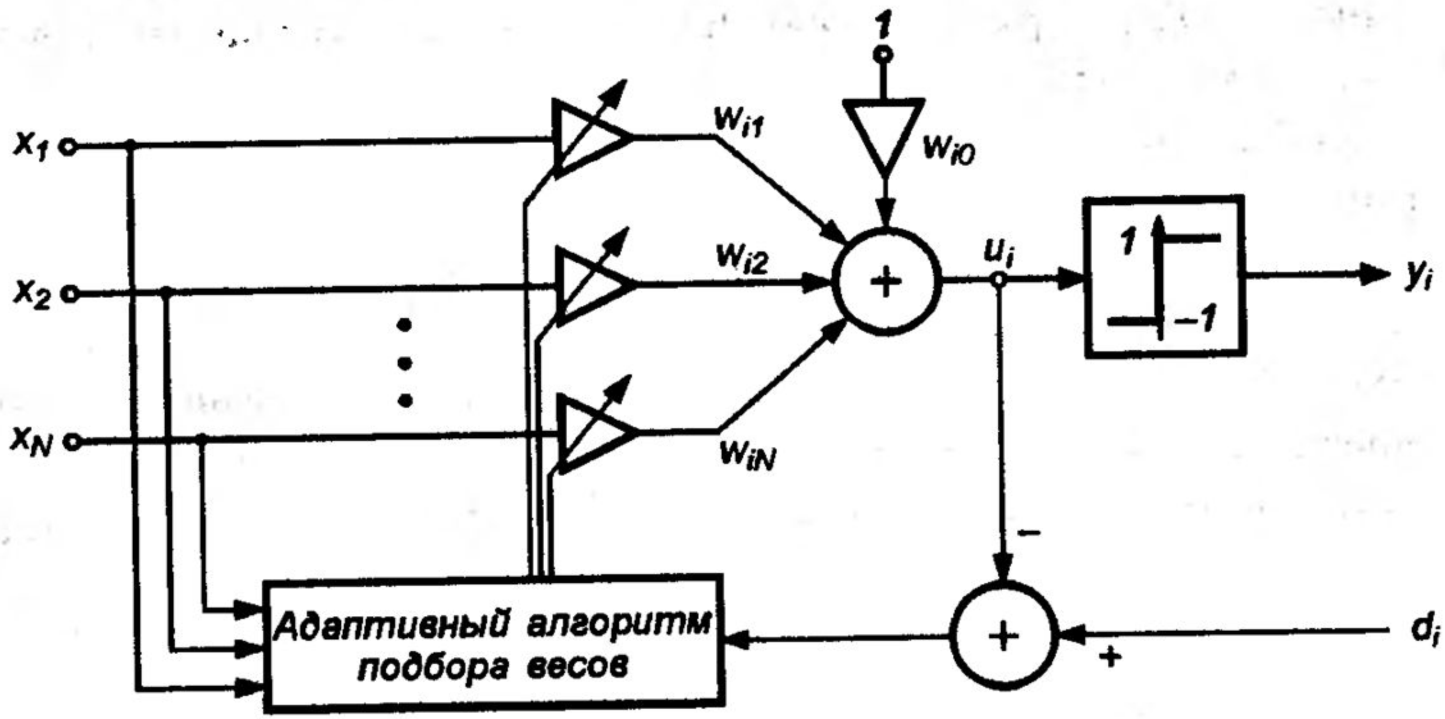


Обучение градиентным методом обеспечивает нахождение, только локального минимума, для нахождения глобального минимума вводиться понятие момента. $\Delta w_{ij}(t+1) = -\eta\delta_i x_j + \alpha\Delta w_{ij}(i);$

Где α - коэффициент момента.



Нейроны типа «Адалайн»



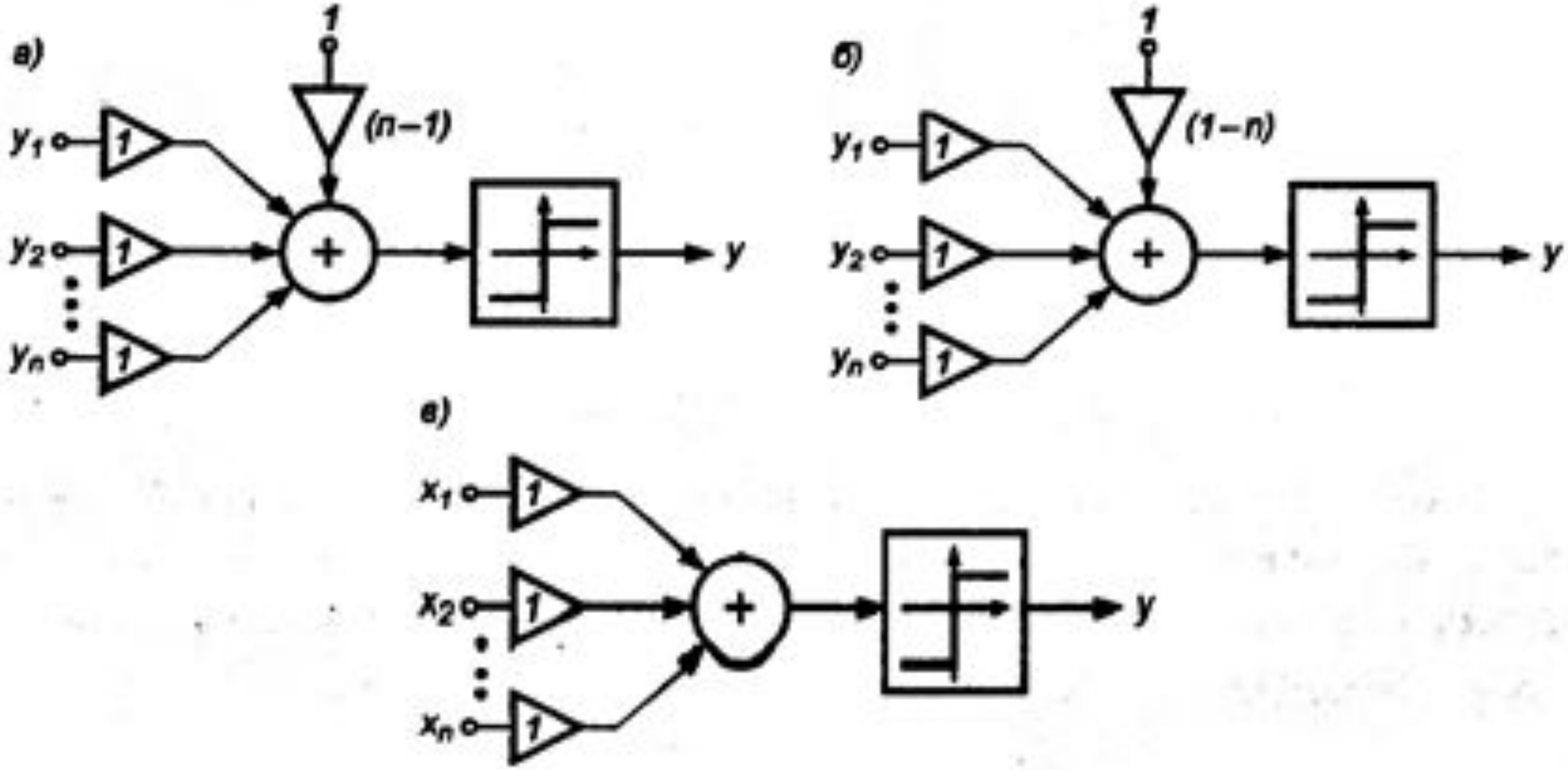
$$y_i(u_i) = \begin{cases} 1 & \text{для} \;\; u_i > 0 \\ -1 & \text{для} \;\; u_i \le 0 \end{cases}.$$

$$E(w) = \frac{1}{2} e_i^2 = \frac{1}{2}\left[d_i - \left(\sum_{j=0}^{N} w_{ij} x_j \right) \right]^2$$

$$w_{ij}(t+1) = w_{ij}(t) + \eta e_i x_j$$

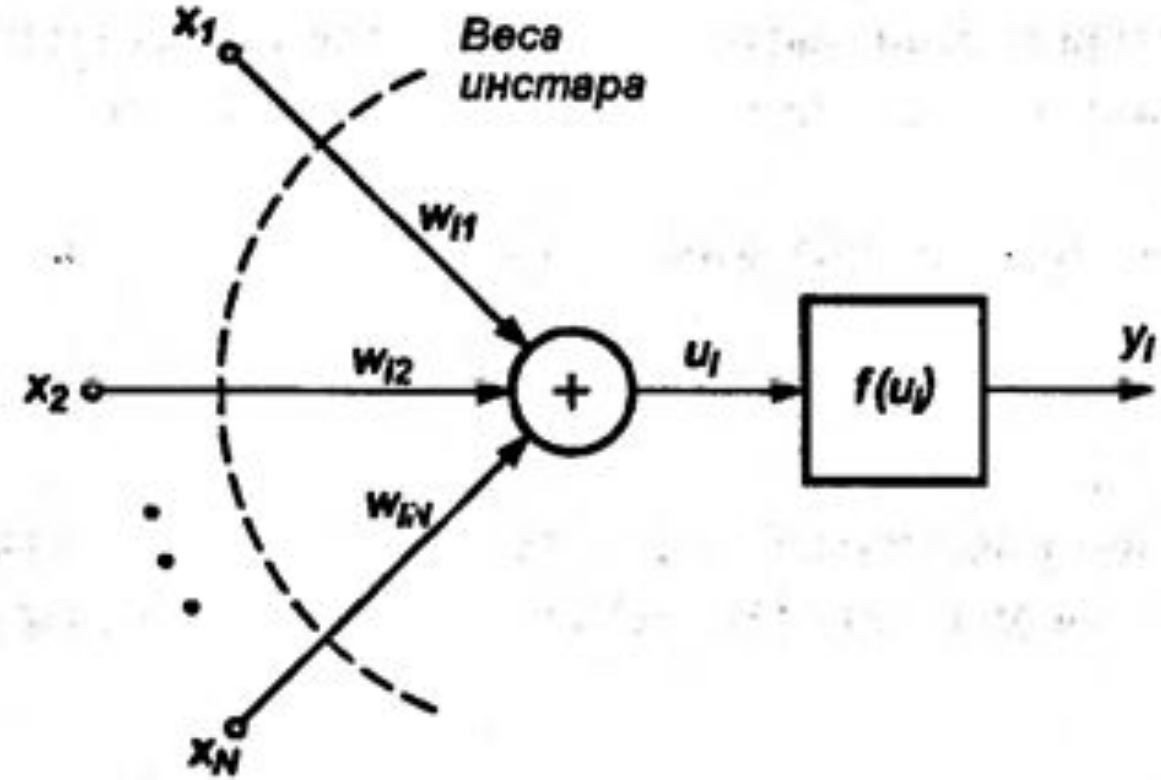
$$\frac{dw_{ij}}{dt} = \mu e_i x_j \,,$$

$$e_i = \left(d_i - \sum_{j=0}^{N} w_{ij} x_j \right).$$



Сеть мадалайн с выходами типа: *а*) OR; *б*) AND; *в*) мажоритарный

Структурная схема инстара

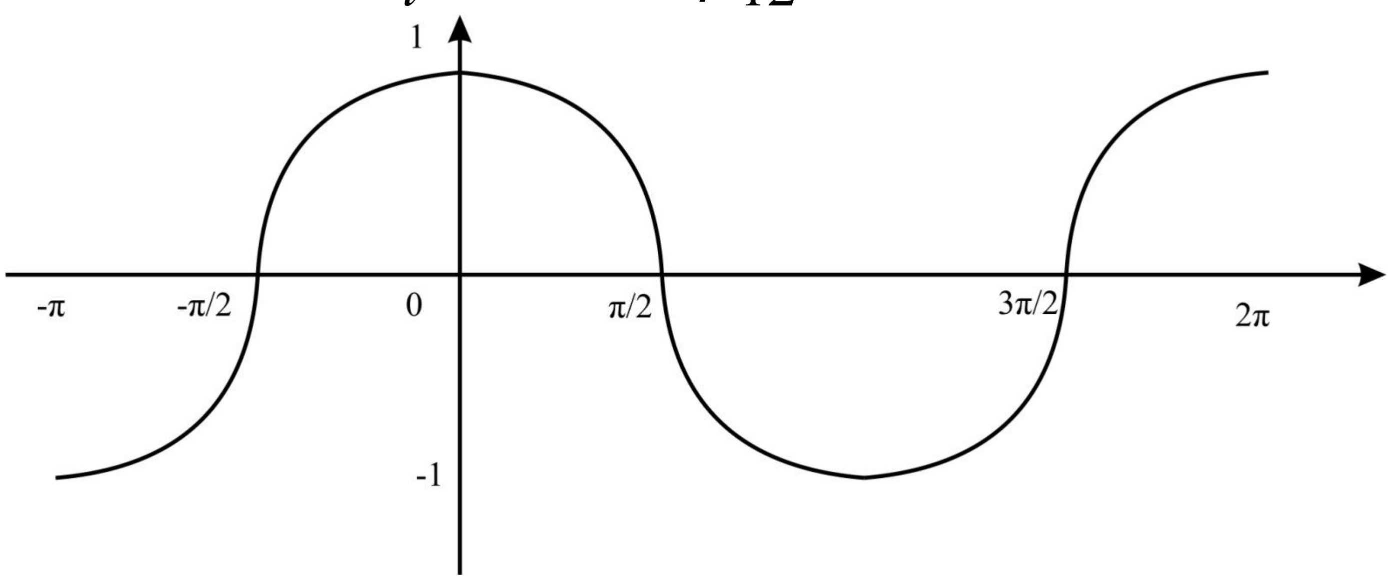


$$x_j \leftarrow \frac{x_j}{\sqrt{x_1^2 + x_2^2 + \dots x_n^2}}$$

$$w = \left[w_{i1}, w_{i2}, \dots w_{iN}\right]^T = x_1$$

$$u_i = w^T x_2 = x_1^T x_2 = \|x_1\| \|x_2\| \cos\varphi_{12}$$

$$u_i = \cos\varphi_{12}$$

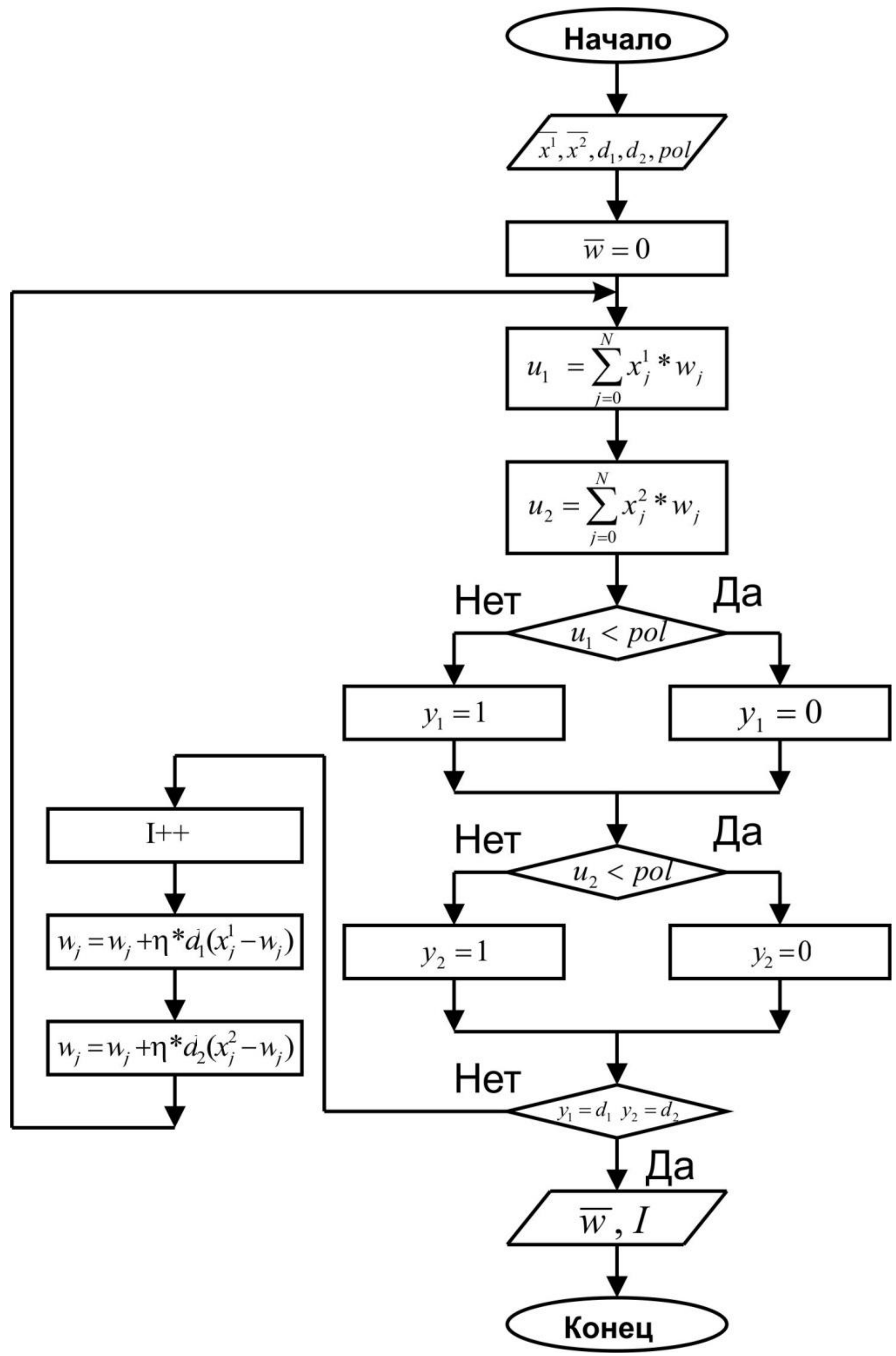


$$|x| = \sqrt{x_1^2 + x_2^2 + \dots x_n^2} = \sqrt{\left(\frac{x_1}{\sqrt{x_1^2 + x_2^2 + \dots x_n^2}}\right)^2 + \left(\frac{x_2}{\sqrt{x_1^2 + x_2^2 + \dots x_n^2}}\right)^2 + \dots + \left(\frac{x_n}{\sqrt{x_1^2 + x_2^2 + \dots x_n^2}}\right)^2} =$$

$$= \sqrt{\frac{{x_1}^2}{x_1^2 + x_2^2 + \dots x_n^2} + \frac{{x_2}^2}{x_1^2 + x_2^2 + \dots x_n^2} + \dots + \frac{{x_n}^2}{x_1^2 + x_2^2 + \dots x_n^2}} = \sqrt{\frac{x_1^2 + x_2^2 + \dots x_n^2}{x_1^2 + x_2^2 + \dots x_n^2}} = 1$$

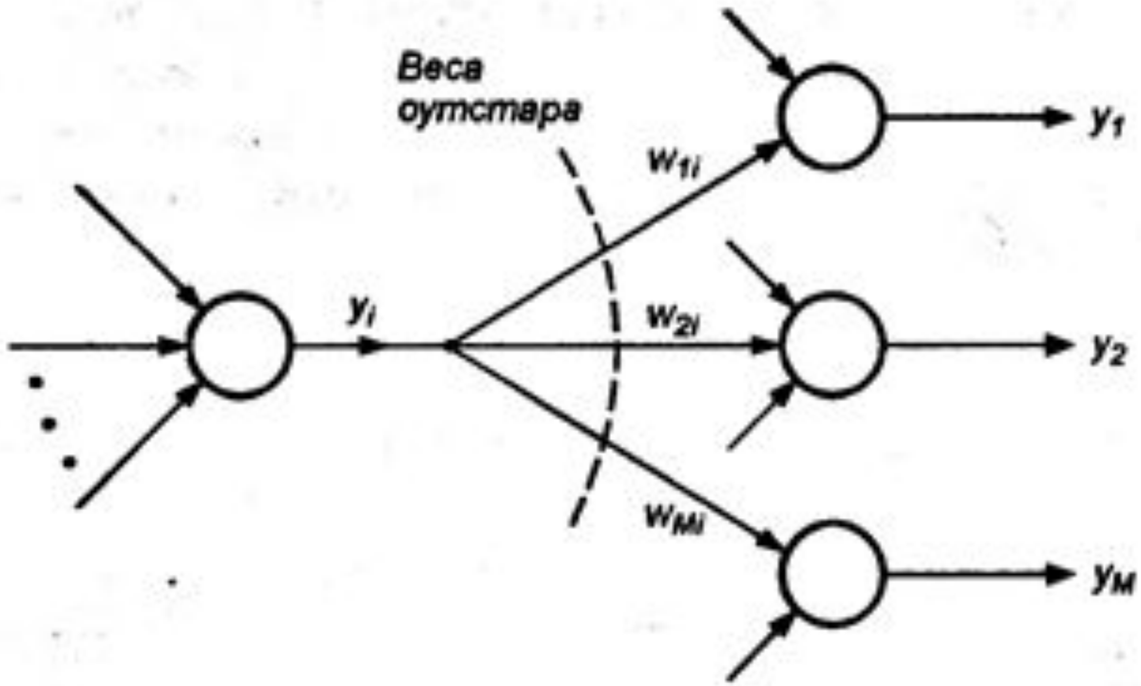
Правило Гроссберга

$$w_{ij}(t+1) = w_{ij}(t) + \eta y_i\left[x_j + w_{ij}(t)\right]$$



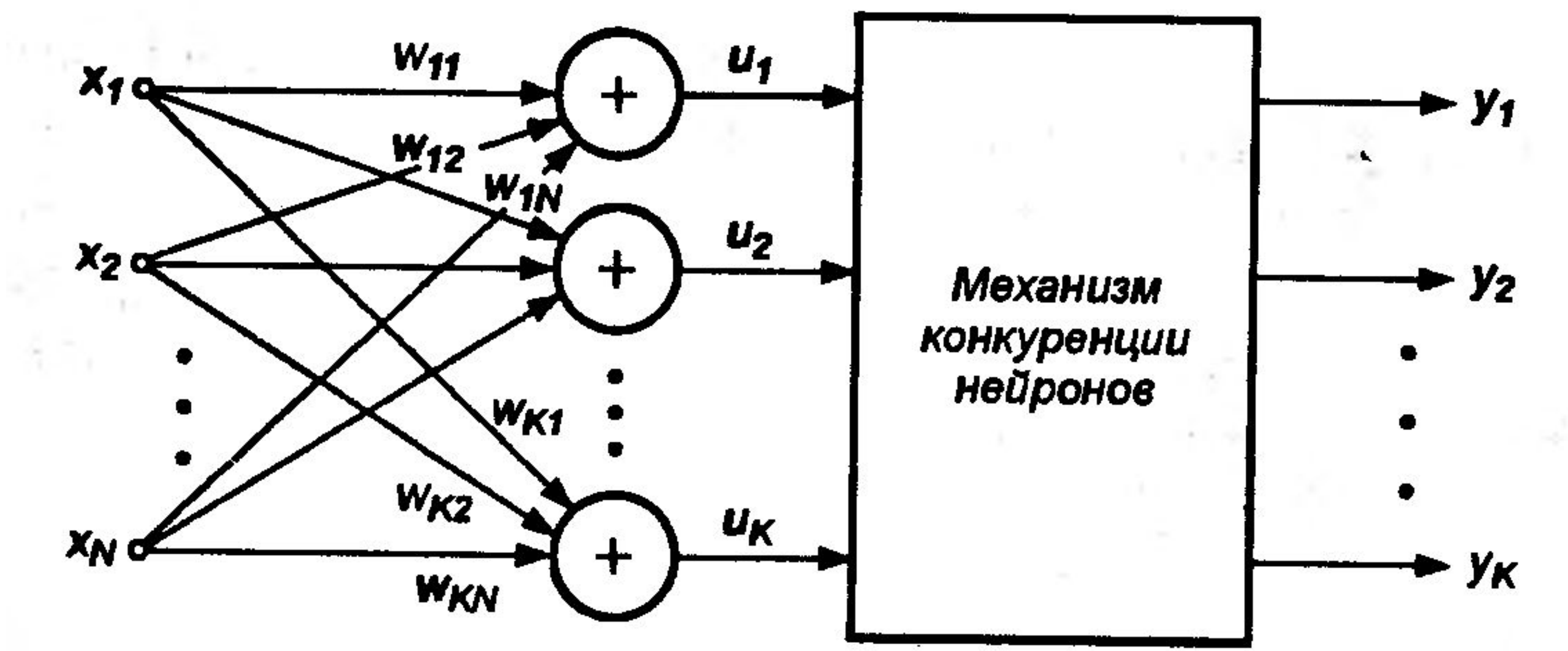
Начало
$\overline{x^1}, \overline{x^2}, d_1, d_2, pol$
$\overline{w} = 0$
$u_1 = \sum_{j=0}^{N} x_j^1 * w_j$
$u_2 = \sum_{j=0}^{N} x_j^2 * w_j$
Нет
Да
$u_1 < pol$
$y_1 = 1$
$y_1 = 0$
Нет
Да
$u_2 < pol$
$y_2 = 1$
$y_2 = 0$
I++
$w_j = w_j + \eta * d_1(x_j^1 - w_j)$
$w_j = w_j + \eta * d_2(x_j^2 - w_j)$
Нет
$y_1 = d_1 \; y_2 = d_2$
Да
$\overline{w}, I$
Конец

Структурная схема оутстара



$$w_{ji}(t+1) = w_{ji}(t) + \eta y_i (y_j - w_{ji}(t))$$

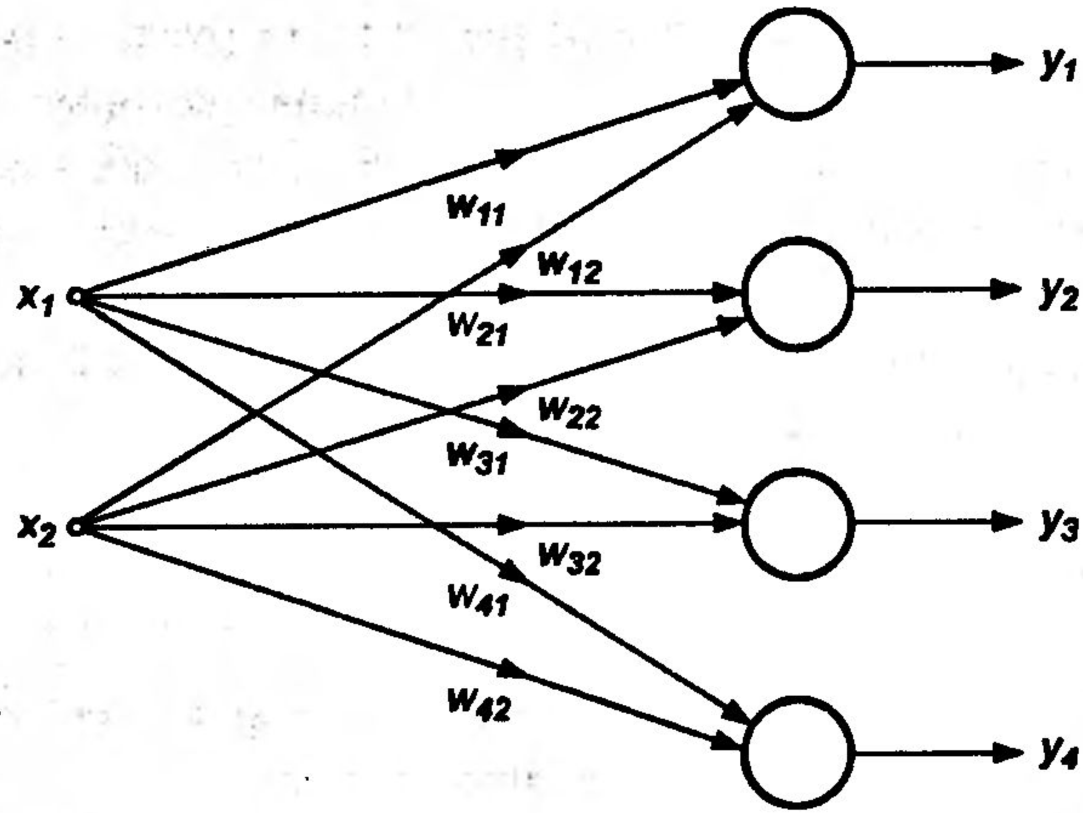
Нейроны типа WTA



$$w_{ij}(t+1) = w_{ij}(t) + \eta[x_j - w_{ij}(t)].$$

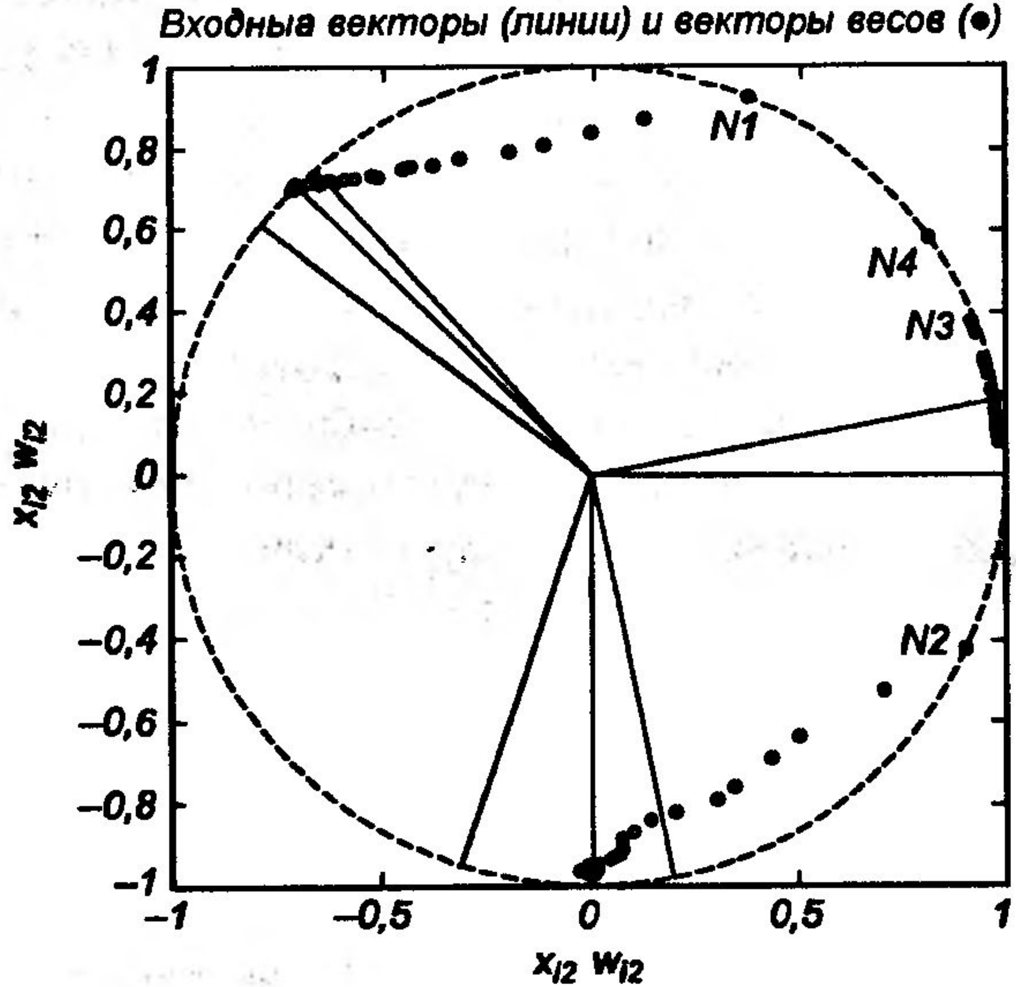
$$u_i = w^T x = \|w\| \, \|x\| \cos\varphi_i .$$

$$u_i = \cos\varphi_i.$$



$$x_1=\begin{bmatrix}0,97\\0,20\end{bmatrix},\ x_2=\begin{bmatrix}1,00\\0,00\end{bmatrix},\ x_3=\begin{bmatrix}-0,72\\0,70\end{bmatrix},\ x_4=\begin{bmatrix}-0,67\\0,74\end{bmatrix},$$

$$x_5=\begin{bmatrix}-0,80\\0,60\end{bmatrix},\ x_6=\begin{bmatrix}0,00\\-1,00\end{bmatrix},\ x_7=\begin{bmatrix}0,20\\-0,97\end{bmatrix},\ x_8=\begin{bmatrix}-0,30\\-0,95\end{bmatrix}.$$

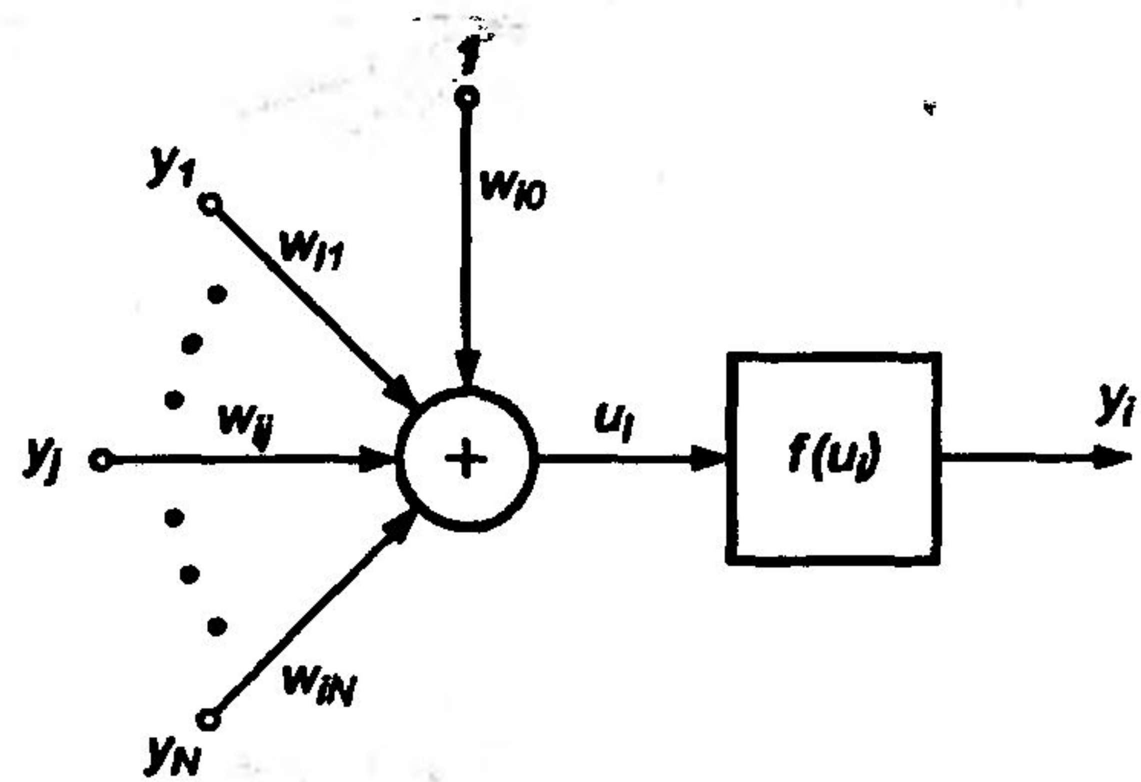


$$w_1 = \begin{bmatrix} -0,7314 \\ 0,6786 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 0,0276 \\ -0,9790 \end{bmatrix}, \quad w_3 = \begin{bmatrix} 0,9904 \\ -0,0656 \end{bmatrix}.$$

Модель нейрона Хебба

y_j $\quad y_i$ $\quad w_{ij}$

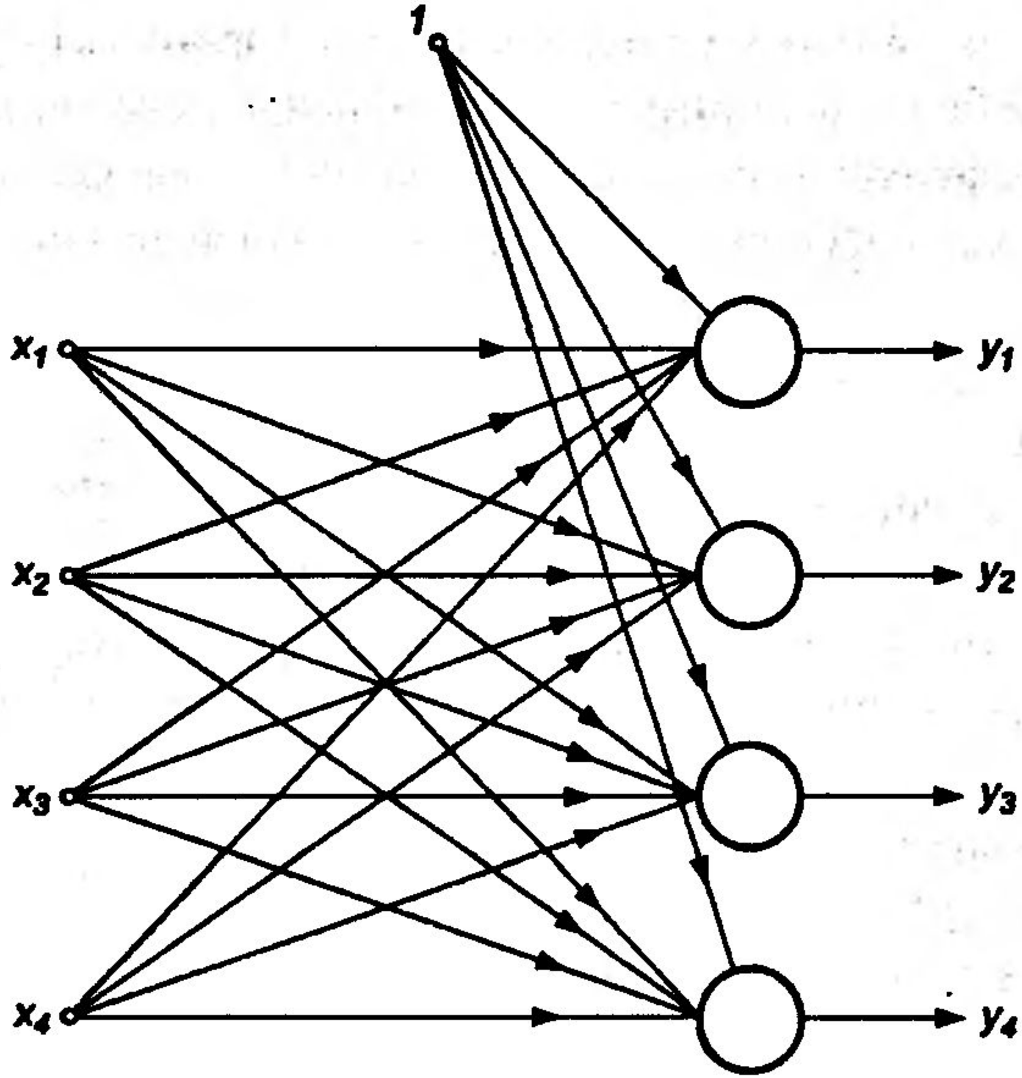
$$\Delta w_{ij} = \eta y_j y_i \,,$$



$$\Delta w_{ij} = \eta\, y_j\, d_i$$

$$w_{ij}(t+1) = w_{ij}(t) + \Delta w_{ij}$$

$$w_{ij}(t+1) = w_{ij}(t)(1-\gamma) + \Delta w_{ij}$$


1
x1
x2
x3
x4
y1
y2
y3
y4

$$x_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad x_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}. \qquad \mathbf{w} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

$\eta = 0{,}1$ и $\gamma = \dfrac{\eta}{3}$

$$\mathbf{w} = \begin{bmatrix} 0{,}971 & 0 & 0 & 0 \\ 0 & 1{,}029 & 0 & 0{,}899 \\ 0 & 0 & 0{,}999 & 0 \\ 0 & 0 & 0 & 1{,}029 \end{bmatrix}.$$