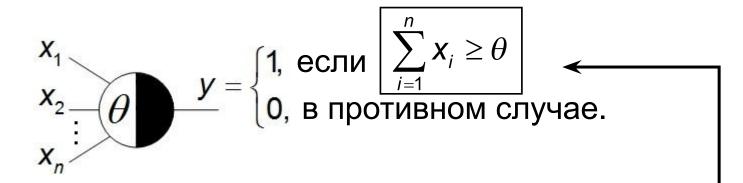


## <u>Лекция 5</u>

Обобщенная модель нейрона. Персептрон (структура, алгоритмы обучения). Решение задачи распознавания образов с помощью НС.

## 1943 – формальный нейрон (МакКаллок, Питтс):



1958 – расширение модели (МакКаллок):

условие возбуждения

 $(\theta$  – порог возбуждения)

$$x_1 w_1 w_2 w_2 \theta = \begin{cases} 1, \text{ если } \sum_{i=1}^n w_i x_i \ge \theta \\ 0, \text{ в противном случае.} \end{cases}$$

Здесь: 
$$x_i \in \{0,1\}$$
;  $y \in \{0,1\}$ ;  $w_i \in \{-1,1\}$ 

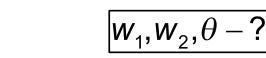
$$w_i = -1$$
 «тюрмозящие» входы; игозбуждающие» входы

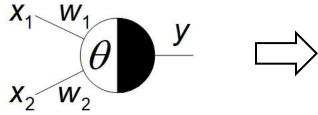
## ПРОБЛЕМА ИСКЛЮЧАЮЩЕГО ИЛИ (Exclusive OR, XOR)

→ [М. Мински, С. Пейперт, 1969]

## Таблица истинности функции XOR

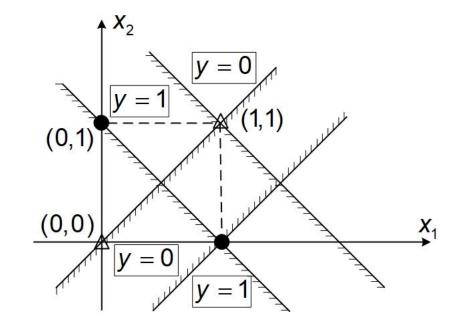
<i>X</i> <sub>1</sub>	<i>X</i> <sub>2</sub>	У	
0	0	0	)
0	1	1	
1	0	1	)
1	1	0	





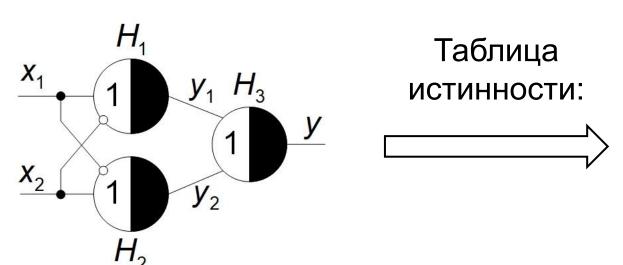
$$W_1X_1 + W_2X_2 \ge \theta$$

$$y = f(x_1, x_2) = x_1 \oplus x_2$$



⇒ XOR = <u>линейно неразделимая</u> функция

## Решение проблемы – <u>2-хслойная</u> HC:



<i>X</i> <sub>1</sub>	<i>X</i> <sub>2</sub>	<i>Y</i> <sub>1</sub>	<i>y</i> <sub>2</sub>	У
0	0	0	0	0
0	1	0	1	1
1	0	1	0	( <u>+</u> )
1	1	0	0	0

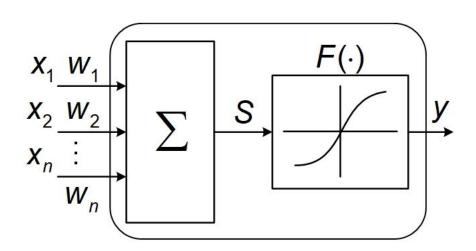
#### Условия возбуждения нейронов:

$$H_1: X_1 - X_2 \ge 1;$$

$$H_2: -x_1 + x_2 \ge 1$$
;

$$H_3: y_1 + y_2 \ge 1.$$

# ОБОБЩЕННАЯ МОДЕЛЬ НЕЙРОНА



Уравнение нейрона:

$$y = F(S)$$
, (1)

где 
$$S = \sum_{i=1}^{n} w_i x_i - \text{суммарное}$$

возбуждение нейрона;

$$x_i \in [-1,1]$$
 – входы нейрона  $(i = 1,2,...,n);$ 

$$y \in [-1,1]$$
 – выход нейрона;

$$W_i$$
 ( $i = 1, 2, ..., n$ ) — вещественные числа, веса синаптических связей.

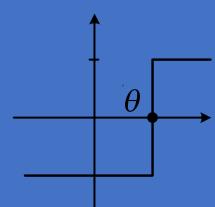
$$F(S)$$
 – функция активации (характеристическая функция), удовлетворяющая 2-м условиям:

a) 
$$|F(S)| \le 1$$
;

б) F(S) – монотонно возрастающая (или неубывающая) функция S.

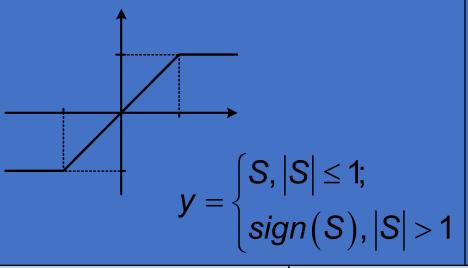
# ВИДЫ ФУНКЦИИ АКТИВАЦИИ

а) логическая

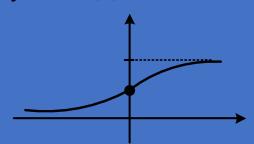


 $y = sign(S - \theta)$ 

б) линейная

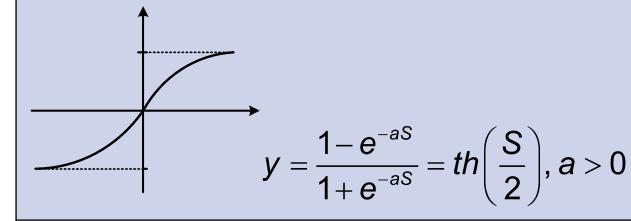


в) сигмоидная

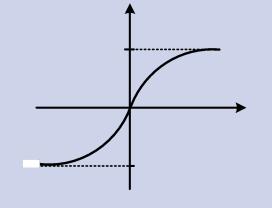


$$y = \frac{1}{1 + e^{-aS}}, a > 0$$

г) 2-хполярная сигмоида (гиперболический тангенс)



д) рациональная сигмоида (функция Элуота)

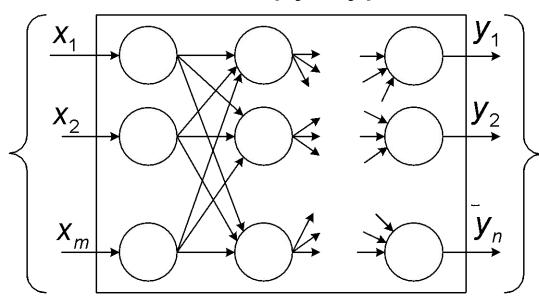


$$y=\frac{S}{c+|S|}, c>0$$

# ПЕРСЕПТРОН (структура)

- 1958 однослойный персептрон (Ф. Розенблат)
  - $\rightarrow$  (от англ. Perception «восприятие»)
- 1958 1-й нейрокомпьютер «Марк-1» (Ф. Розенблат)

## Структура многослойного перспетрона



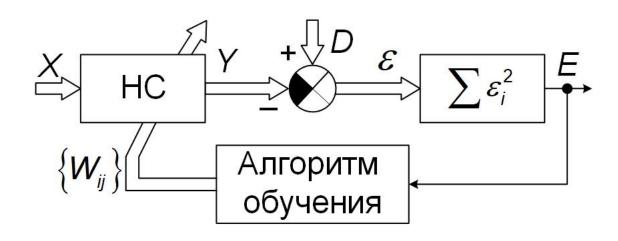
$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$
 — вектор входов,  $x_i \in [-1,1];$ 

$$Y = egin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$
 — вектор выходов  $y_j \in [-1,1].$ 

#### 2 режима работы:

- обучение (настройка весов НС);
- эксплуатация.

## ПЕРСЕПТРОН (процедура обучения)



Цель обучения:  $E \rightarrow \min$ ,

где 
$$E = \left| \mathcal{E} \right|^2$$
сұммарная

квадратичная ошибка обучения (СКО) НС.

$$D = egin{bmatrix} d_1 \ d_2 \ \vdots \ d_n \end{bmatrix}$$
 — вектор эталонов, (желаемых выходов) НС;

$$\mathcal{E} = \begin{bmatrix} arepsilon_1 \\ arepsilon_2 \\ arepsilon_1 \end{bmatrix}$$
 – вектор ошибки обучения НС;

$$\mathcal{E} = D - Y$$

## ОБУЧАЮЩАЯ ВЫБОРКА

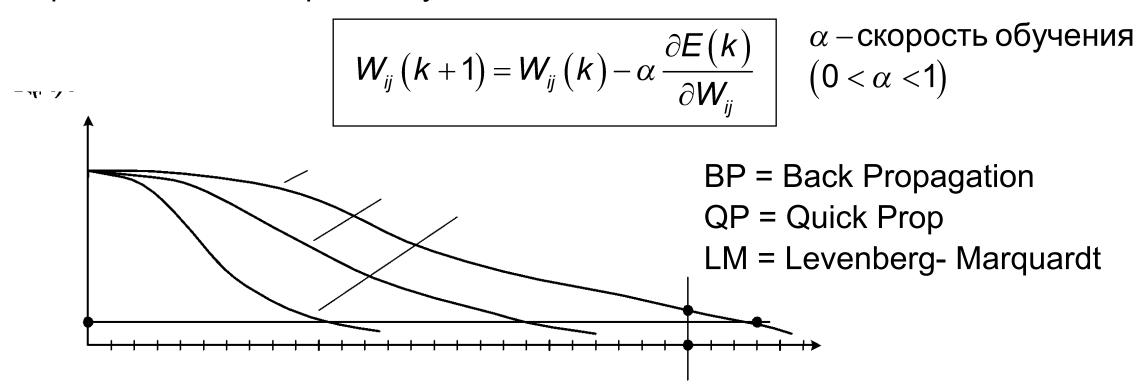
r	Входы НС				Желаемые выходы НС			
1	$X_1$	<i>X</i> <sub>2</sub>		X <sub>m</sub>	$d_1$	$d_2$		$d_n$
1								
2								
R								

#### АЛГОРИТМЫ ОБУЧЕНИЯ ПЕРСЕПТРОНА

$$\underline{\mathsf{CKO}}: \quad E = \sum_{r=1}^{R} \sum_{i=1}^{n} \left[ d_{i}^{(r)} - y_{i}^{(r)} \right]^{2} \rightarrow \mathsf{min},$$

где *n* – число выходов HC; *R* – объем обучающей выборки.

Градиентный алгоритм обучения:



Критерий останова: a)  $k = K_{3aд}$ ; б)  $E = E_{3ad}$ 

# СЛОЖНОСТИ ПРОЦЕССА ОБУЧЕНИЯ

- 1) «зависание» в локальных минимумах
  - → метод «тяжелого» шарика:

$$W_{ij}(k+1) = W_{ij}(k) + \Delta W_{ij}(k) + \beta \cdot \Delta W_{ij}(k-1)$$

где 
$$\Delta W_{ij}(k) = -\alpha \frac{\partial E(k)}{\partial W_{ij}}$$
моФен $\beta$ ;<1 –

2) регуляризация процесса обучения:

$$E = \sum_{r=1}^{n} \sum_{i=1}^{n} \left[ d_i^{(r)} - y_i^{(r)} \right]^2 + \gamma \cdot \sum_{i,j} W_{ij}^2 \longrightarrow \min$$

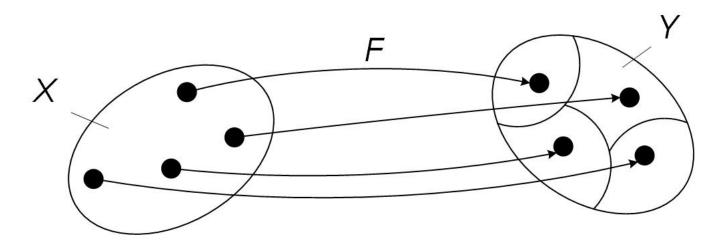
где  $\gamma$  – коэффициент «штрафа» весов НС;

- 3) эффект «переобучения» (over-learning) HC
  - → выбор оптимальной сложности НС (количества нейронов в скрытом слое); тестирование / валидация НС.

# РЕШЕНИЕ ЗАДАЧИ РАСПОЗНАВАНИЯ ОБРАЗОВ С ПОМОЩЬЮ ПЕРСЕПТРОНА

Задача классификации (распознавания образов):

требуется определить, к какому из M известных классов относится предъявляемый объект (образ), представленный вектором  $X = (x_1, x_2, ..., x_m)^T$ .

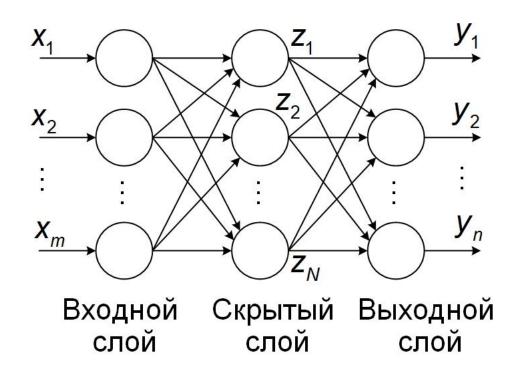


Множество входных векторов (признаков)

Множество решений (классов принадлежности)

#### Решение:

#### Персептрон с 1 скрытым слоем:



$$n = \lceil \log_2 M \rceil$$
;  $R \ge M$ 

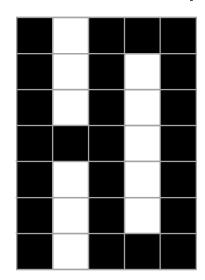
Количество нейронов в скрытом слое:

$$N = \left\lceil \frac{R \cdot n}{m+n} \right\rceil$$

Васильев В.И., Ильясов Б.Г. Интеллектуальные системы управления. Теория и практика, 2009. C. 168-170.

## ПРИМЕР

Дано: 7х5 матрица входных образов (изображений):

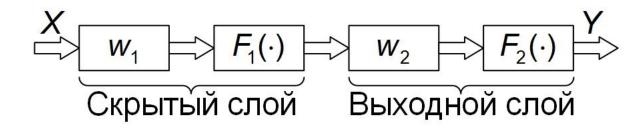


M = 64 (буквы, цифры, символы); входы  $x_i \in \{-1,1\}$ ; вектор желаемых реакций HC:  $D = (d_1, d_2, ..., d_n)^{\text{д}}$  оичный код / номер образа (класса изображений).

<u>Решение</u>:  $m = 7 \cdot 5 = 35$ ; R = M = 64;

$$n = \lceil \log_2 64 \rceil = 6; \ N = \left\lceil \frac{R \cdot n}{m+n} \right\rceil = \left\lceil \frac{64 \cdot 6}{35+6} \right\rceil = 10.$$

Структура HC: <u>35 – 10 – 6</u>:



 $m{w}_1 = \left\| m{w}_{ij}^{(1)} 
ight\|_{35 imes 10}$  ,  $m{w}_2 = \left\| m{w}_{ij}^{(2)} 
ight\|_{10 imes 6}$  — матрицы

весов синаптических связей;

 $F_1(\cdot)$ ,  $F_2(\cdot)$  – сигмоидные функции активации

Количество настраиваемых параметров (весов) НС:  $(K\Pi) = 35 \cdot 10 + 10 \cdot 6 = 410$ .