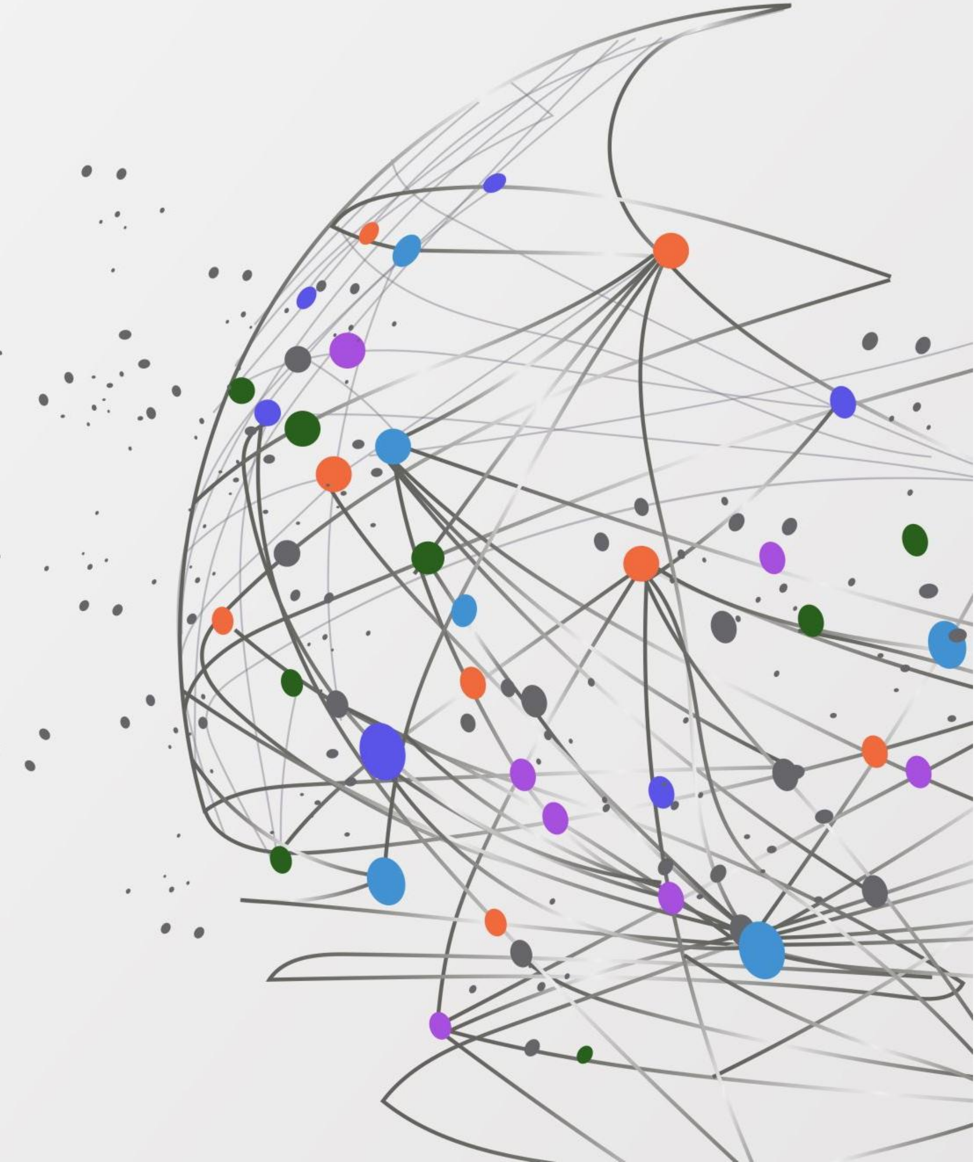


КЛАССИФИКАЦИЯ ОБРАЗЦОВ

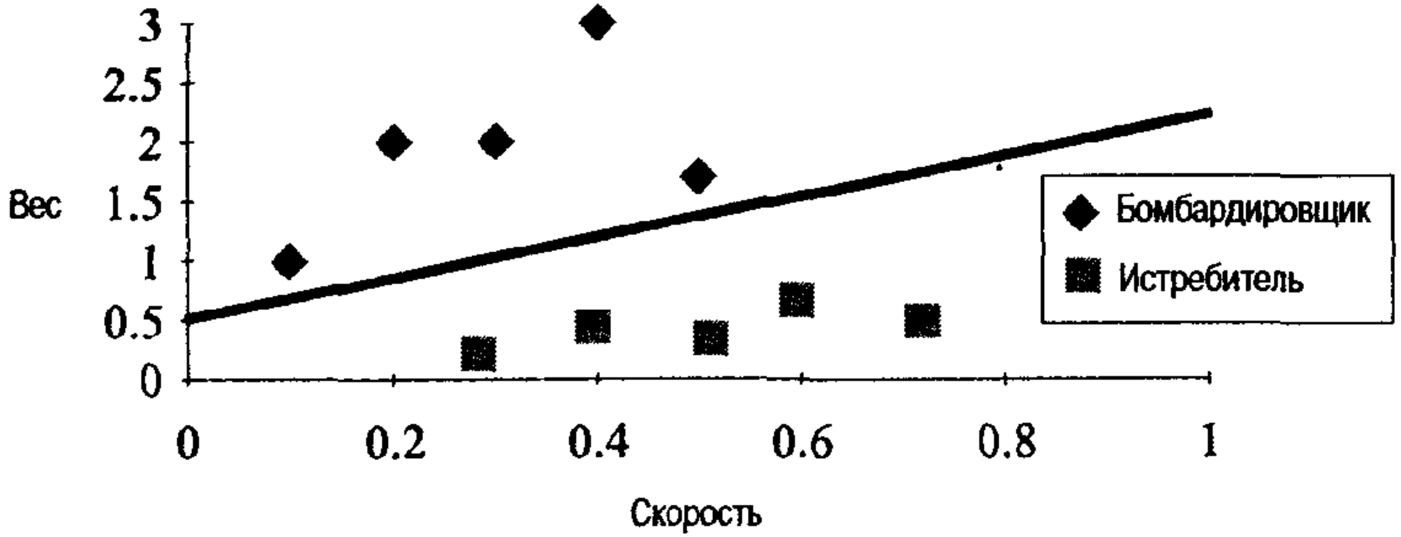
ЛЕКЦИЯ 3

Классификация образцов

<u>**Целью лекции является понять:**</u>

что такое проблема классификации и что такое управляемое обучение;
на основе чего выполняется классификация в сетях с прямой связью;
принципы построения сетей с прямой связью, в которых используется алгоритм обратного распространения ошибок, чтобы иметь возможность реализовать такую нейронную сеть на языке программирования, который вы предпочтете;
разницу между линейной и нелинейной проблемами.

Функция выбора решения



ЕСЛИ вес > 0.80 И скорость < 0.55, ТО бомбардировщик,

ЕСЛИ вес < 0.90 *И* скорость > 0.25, ТО *истребитель.*

Функция выбора решения

Уравнение прямой, разделяющей два типа самолетов, записывается в виде:

$$x_2 = 1.5x_1 + 0.5 ,$$

где x_1 представляет скорость, а x_2 — вес. Это уравнение можно использовать для создания функции выбора решения:

$$f(x_1, x_2) = -x_2 + 1.5x_1 + 0.5 ,$$

$$d = \begin{cases} \text{истребитель}, & \text{если } f(x_1, x_2) \geq 0, \\ \text{бомбардировщик}, & \text{если } f(x_1, x_2) < 0. \end{cases}$$

Например, истребитель, представленный точкой (0.4, 0.5), даст

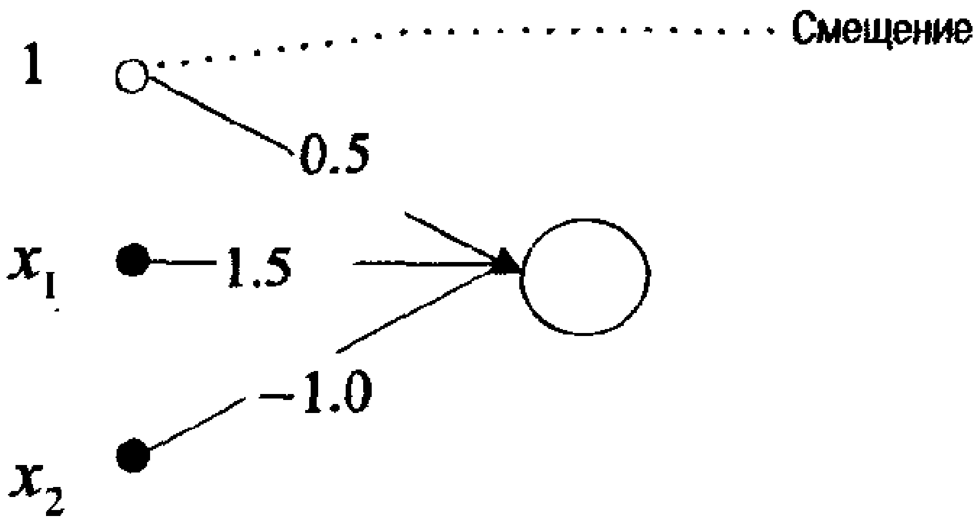
$$f(0.4, 0.5) = -0.5 + 1.5 \times 0.4 + 0.5 = 0.6 ,$$

Функция выбора решения

$$net_j = w_0 + \sum_{i=1}^{n} x_i w_{ij} \ ,$$

где net_j представляет значение комбинированного ввода, w_0 — смещение, связываемое с элементом, значение активности которого считается всегда равным 1, x_i — значение активности i-го элемента, а w_{ij} — вес

$$\text{выход} = \begin{cases} 1, & \text{если комбинированный ввод } \geq 0, \\ 0, & \text{если комбинированный ввод } < 0. \end{cases}$$


Смещение
1
0.5
x_1
1.5
−1.0
x_2

Пример

(а) Вычислите комбинированный сетевой ввод для элемента на **Слайде 6** и соответствующее выходное значение при использовании пороговой функции и входного вектора {0.7 2.5}.

(б) Вычислите выходное значение, использовав в качестве функции активности сигмоидальную функцию. Входной вектор остается таким же, как и в п. (а).

(в) Вычислите комбинированный ввод для сети с архитектурой, показанной на **Слайде 6** но с набором весовых значений [–0,2 0,03 1.2] и таким же входным вектором, как и в п. (а).

Корректировка весов

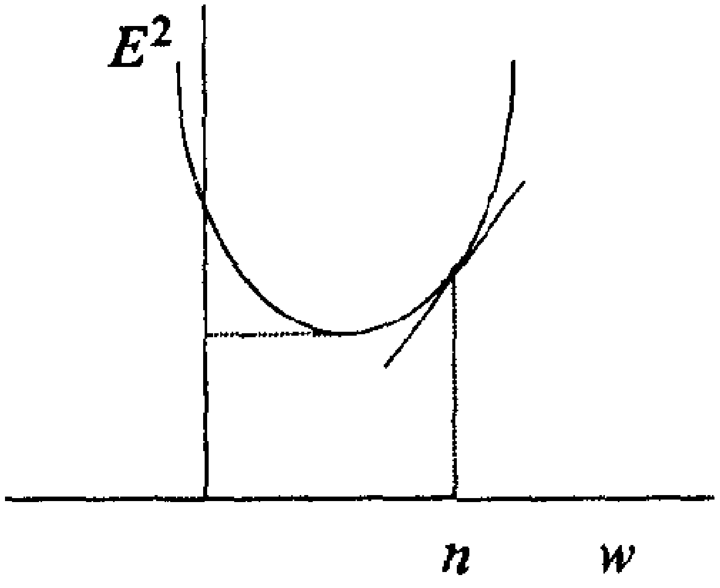
$$E_p = \frac{1}{2}\sum_j \left(t_j - o_j\right)^2 \tag{1}$$

$$E = \frac{1}{2}\left(t - net\right)^2 \tag{2}$$

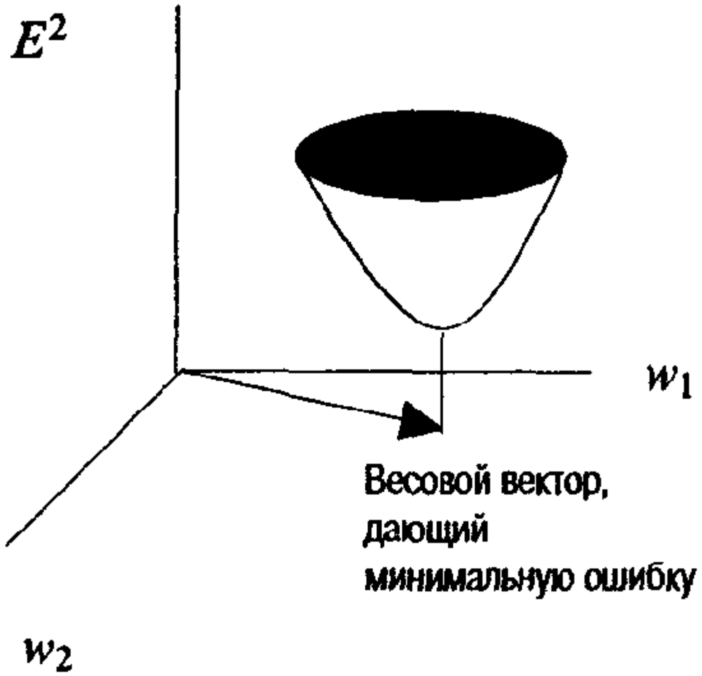
$$\begin{aligned} E &= \frac{1}{2}\left[t^2 - 2t\,net + net^2\right] \\ &= \frac{1}{2}\left[t^2 - 2t\left(x_1 w_1 + x_2 w_2\right) + x_1^2 w_1^2 + 2x_1 w_1 x_2 w_2 + x_2^2 w_2^2\right] \end{aligned} \tag{3}$$

Корректировка весов

$$\frac{\partial E}{\partial w_1} = (-t + x_1 w_1 + x_2 w_2) x_1 \qquad (4)$$



Корректировка весов



Минимизация квадрата ошибки

$$\Delta w_{ij} = \eta \delta_j x_i , \qquad \delta_j = (t_j - o_j)$$

$$o_j = \sum_j x_i w_{ij} \qquad \frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial w_{ij}}$$

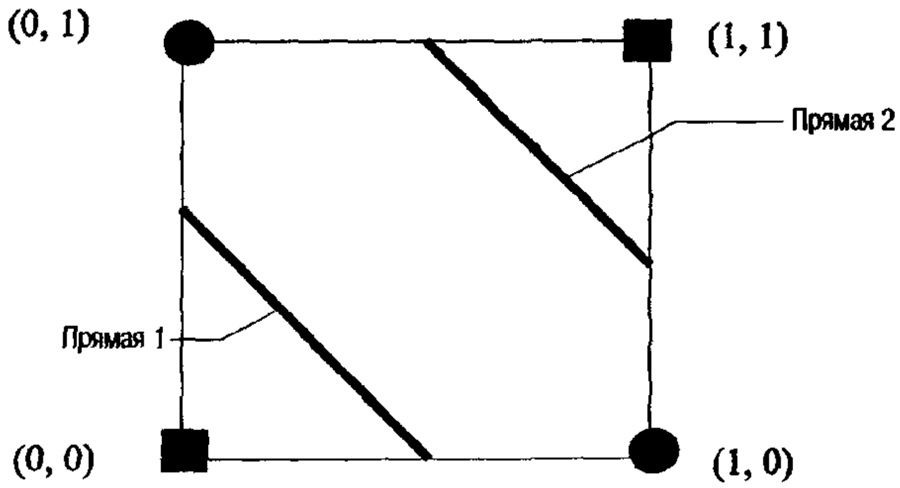
$$\frac{\partial E}{\partial o_j} = -\delta_j$$

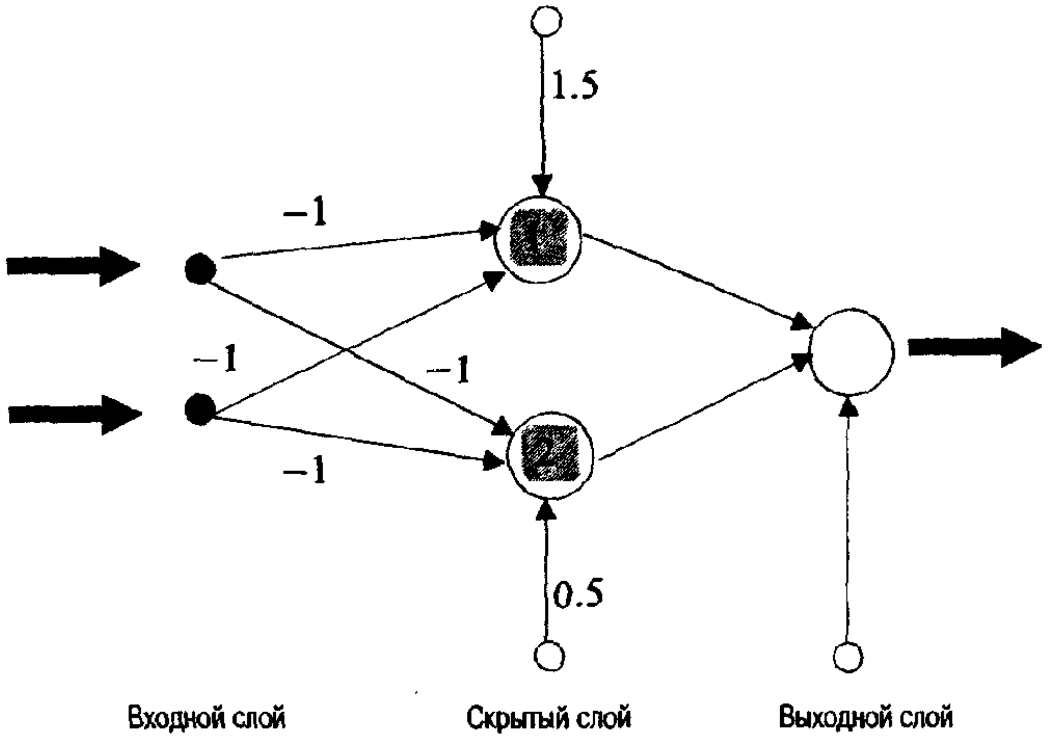
$$\frac{\partial o_j}{\partial w_{ij}} = x_i$$

получаем

$$-\frac{\partial E}{\partial w_{ij}} = \delta_j x_i$$

Линейные и нелинейные проблемы



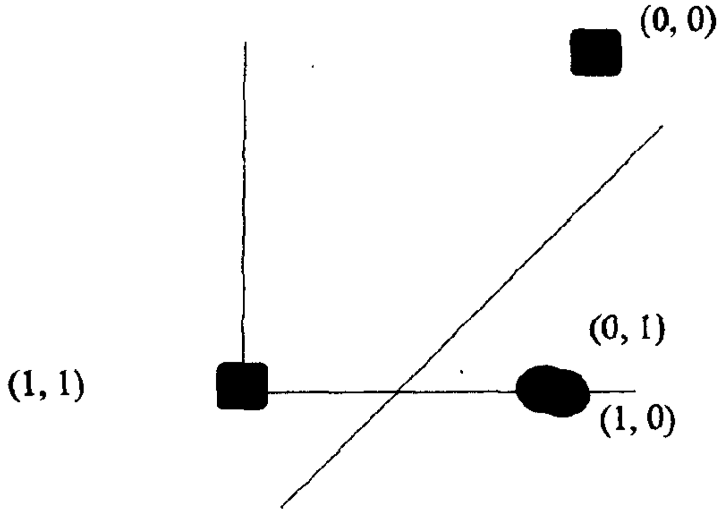


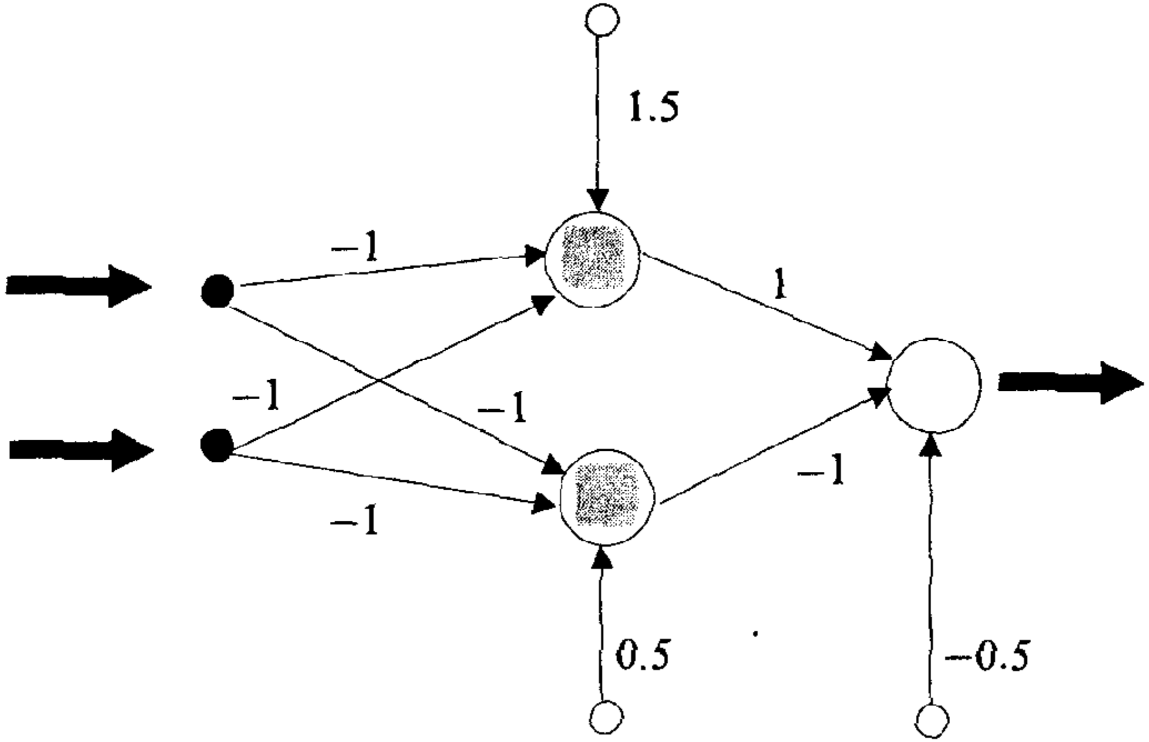
Линейные и нелинейные проблемы

Реакция скрытых элементов сети

		Комбинированный ввод скрытого слоя		Вывод скрытого слоя	
x_1	x_2	*Элемент 1*	*Элемент 2*	*Элемент 1*	*Элемент 2*
1	1	−0.5	−1.5	0	0
1	0	0.5	−0.5	1	0
0	1	0.53	−0.5	1	0
0	0	1.5	0.5	1	1

Линейные и нелинейные проблемы





Линейные и нелинейные проблемы

$$\begin{bmatrix} 1.5 & 0.5 \\ -1.0 & -1.0 \\ -1.0 & -1.0 \end{bmatrix}.$$

Линейные и нелинейные проблемы

Входной вектор, p	Первый слой весов	Реакция скрытого слоя		Второй слой весов	Реакция выходного слоя	
		Комбинированный ввод	Вывод		Ввод	Вывод
$[1 \mid 1 \quad 1]$	$\begin{bmatrix} 1.5 & 0.5 \\ -1.0 & -1.0 \\ -1.0 & -1.0 \end{bmatrix}$	$[-0.5 \quad -1.5]$	$[0 \quad 0]$	$\begin{bmatrix} -0.5 \\ 1.0 \\ -1.0 \end{bmatrix}$	-0.5	0
$[1 \mid 1 \quad 0]$	$\begin{bmatrix} 1.5 & 0.5 \\ -1.0 & -1.0 \\ -1.0 & -1.0 \end{bmatrix}$	$[0.5 \quad -0.5]$	$[1 \quad 0]$	$\begin{bmatrix} -0.5 \\ 1.0 \\ -1.0 \end{bmatrix}$	0.5	1
$[1 \mid 0 \quad 1]$	$\begin{bmatrix} 1.5 & 0.5 \\ -1.0 & -1.0 \\ -1.0 & -1.0 \end{bmatrix}$	$[0.5 \quad -0.5]$	$[1 \quad 0]$	$\begin{bmatrix} -0.5 \\ 1.0 \\ -1.0 \end{bmatrix}$	0.5	1
$[1 \mid 0 \quad 0]$	$\begin{bmatrix} 1.5 & 0.5 \\ -1.0 & -1.0 \\ -1.0 & -1.0 \end{bmatrix}$	$[1.5 \quad 0.5]$	$[1 \quad 1]$	$\begin{bmatrix} -0.5 \\ 1.0 \\ -1.0 \end{bmatrix}$	-0.5	0

Линейные и нелинейные проблемы

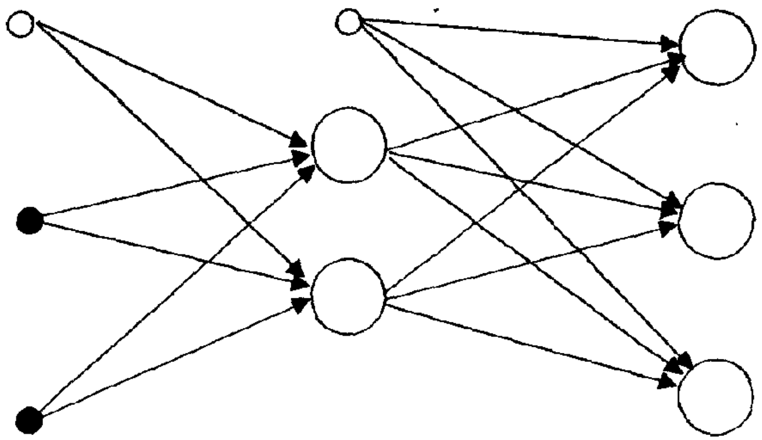
$$\begin{bmatrix} 1.0 & -2.0 \\ 2.0 & 0.5 \\ -3.0 & 1.0 \end{bmatrix}$$

для первого слоя и

$$\begin{bmatrix} 2.0 & 1.0 & 3.0 \\ -1.0 & 5.0 & 4.0 \\ -3.0 & 1.0 & 2.0 \end{bmatrix}$$

для второго.

Линейные и нелинейные проблемы

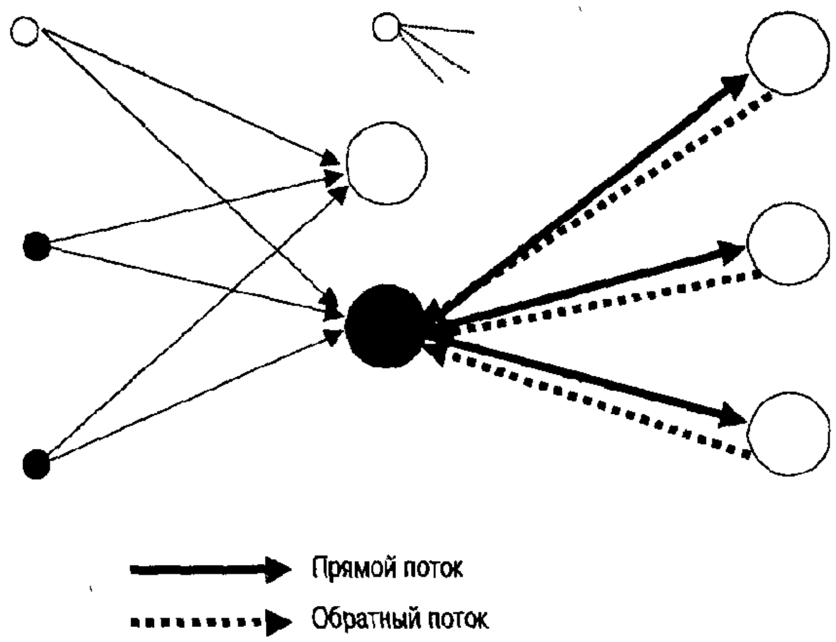


$$[1.0 \quad 0.0 \quad 1.0]\begin{bmatrix} 1.0 & -2.0 \\ 2.0 & 0.5 \\ -3.0 & 1.0 \end{bmatrix} = [-2.0 \quad -1.0].$$

$$[1.0 \quad -2.0 \quad -1.0]\begin{bmatrix} 2.0 & 1.0 & 3.0 \\ -1.0 & 5.0 & 4.0 \\ -3.0 & 1.0 & 2.0 \end{bmatrix} = [7.0 \quad -10.0 \quad -7.0].$$

$$[1.0 \quad 0.0 \quad 1.0]\begin{bmatrix} 5.0 & 3.0 & 3.0 \\ -3.5 & 10.5 & 9.0 \\ 2.0 & -13.0 & -10.0 \end{bmatrix} = [7.0 \quad -10.0 \quad -7.0]$$

Обучение по алгоритму обратного распространения ошибки



Обучение по алгоритму обратного распространения ошибки

для каждого входного вектора и связанного выходного вектора

выполнять, пока not STOP

STOP = TRUE

для каждого входного вектора

выполнить прямой проход и найти реальный выход

получить вектор ошибок путем сравнения реальных и целевых значений

если реальный выход не попадает в допустимые рамки, установить STOP = FALSE

выполнить обратный проход для вектора ошибок

в результате обратного прохода определить величины изменения значений весов

обновить значения весов

Обучение по алгоритму обратного распространения ошибки. Немного теории

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ij}} \qquad (6)$$

$$\delta_j = -\frac{\partial E}{\partial net_j} \qquad (7)$$

$$\delta_j = -\frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial net_j}$$

Поскольку $E_p = \frac{1}{2}\sum_j (t_j - o_j)^2$, **имеем**

$$\frac{\partial E}{\partial o_j} = -(t_j - o_j)$$

Обучение по алгоритму обратного распространения ошибки. Немного теории

Для функции активности f (обычно это логистическая функция) выходом является

$$o_j = f(net_j)$$

и поэтому для производной f' получаем:

$$\frac{\partial o_j}{\partial net_j} = f'(net_j).$$

Таким образом,

$$\delta_j = (t_j - o_j) f'(net_j).$$

Обучение по алгоритму обратного распространения ошибки. Немного теории

Для нахождения комбинированного ввода используется обычное суммирование произведений:

$$net_j = \sum_{i=0} x_i w_{ij} \,.$$

Поэтому

$$\frac{\partial net_j}{\partial w_{ij}} = x_i \,.$$

$$\frac{\partial E}{\partial w_{ij}} = -\left(t_j - o_j\right) f'(net_j) x_i$$

$$\Delta w_{ij} = \eta \delta_j x_i$$

Обучение по алгоритму обратного распространения ошибки. Немного теории

$$\delta_j = f'(net_j)\sum_k \delta_k w_{kj}$$

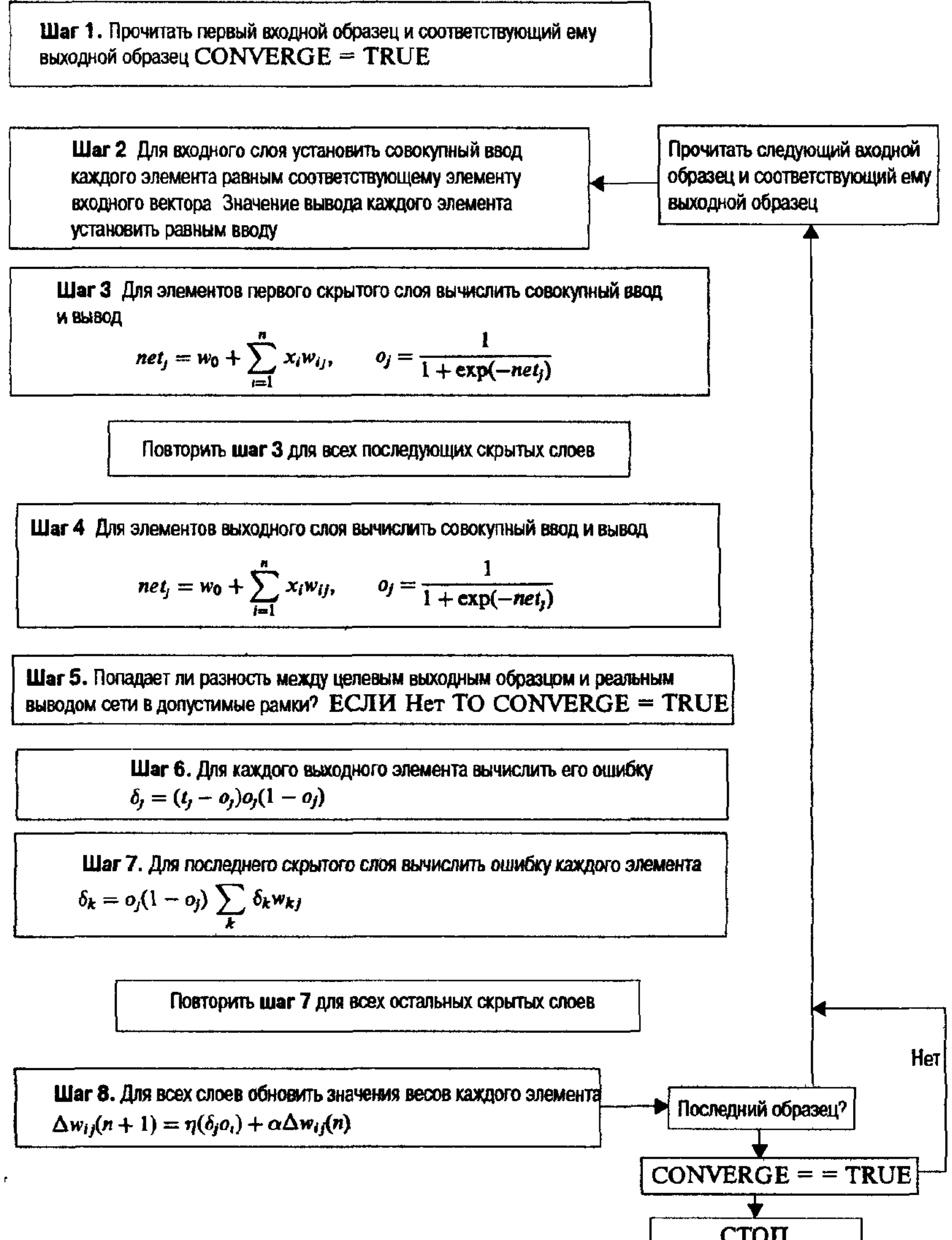
$$f(net_j) = \frac{1}{1+\exp(-net_j)}$$

Производная этой функции активности равна

$$f'(net_j) = \frac{\exp(-net_j)}{\left(1+\exp(-net_j)\right)^2}$$

$$= \frac{1}{1+\exp(-net_j)}\left(1-\frac{1}{1+\exp(-net_j)}\right)$$

$$= f(net_j)\left[1-f(net_j)\right].$$


Шаг 1. Прочитать первый входной образец и соответствующий ему выходной образец CONVERGE = TRUE
Шаг 2 Для входного слоя установить совокупный ввод каждого элемента равным соответствующему элементу входного вектора Значение вывода каждого элемента установить равным вводу
Прочитать следующий входной образец и соответствующий ему выходной образец
Шаг 3 Для элементов первого скрытого слоя вычислить совокупный ввод и вывод
$$net_j = w_0 + \sum_{i=1}^{n} x_i w_{ij}, \qquad o_j = \frac{1}{1+\exp(-net_j)}$$
Повторить **шаг 3** для всех последующих скрытых слоев
Шаг 4 Для элементов выходного слоя вычислить совокупный ввод и вывод
$$net_j = w_0 + \sum_{i=1}^{n} x_i w_{ij}, \qquad o_j = \frac{1}{1+\exp(-net_j)}$$
Шаг 5. Попадает ли разность между целевым выходным образцом и реальным выводом сети в допустимые рамки? ЕСЛИ Нет ТО CONVERGE = TRUE
Шаг 6. Для каждого выходного элемента вычислить его ошибку
$$\delta_j = (t_j - o_j)o_j(1 - o_j)$$
Шаг 7. Для последнего скрытого слоя вычислить ошибку каждого элемента
$$\delta_k = o_j(1 - o_j)\sum_k \delta_k w_{kj}$$
Повторить **шаг 7** для всех остальных скрытых слоев
Шаг 8. Для всех слоев обновить значения весов каждого элемента
$$\Delta w_{ij}(n+1) = \eta(\delta_j o_i) + \alpha \Delta w_{ij}(n)$$
Последний образец?
Нет
CONVERGE = = TRUE
СТОП

Алгоритм обратного распространения ошибки. Пример.

$$\Delta w_{ij}(n+1) = \eta \delta_j o_i + \alpha \Delta w_{ij}(n)$$

Пример. Представьте полностью прямой и обратный проходы в сети с прямой связью, использующей алгоритм обратного распространения ошибок, для входного образца [0 1 0.9] и целевого выходного значения 0 9 в предположении, что сеть имеет архитектуру 2-2-1 (т.е. два входных, два скрытых один выходной элемент) с весовыми коэффициентами

$$\begin{bmatrix} 0\,1 & 0\,1 \\ -0\,2 & -0\,1 \\ 0\,1 & 0\,3 \end{bmatrix}$$

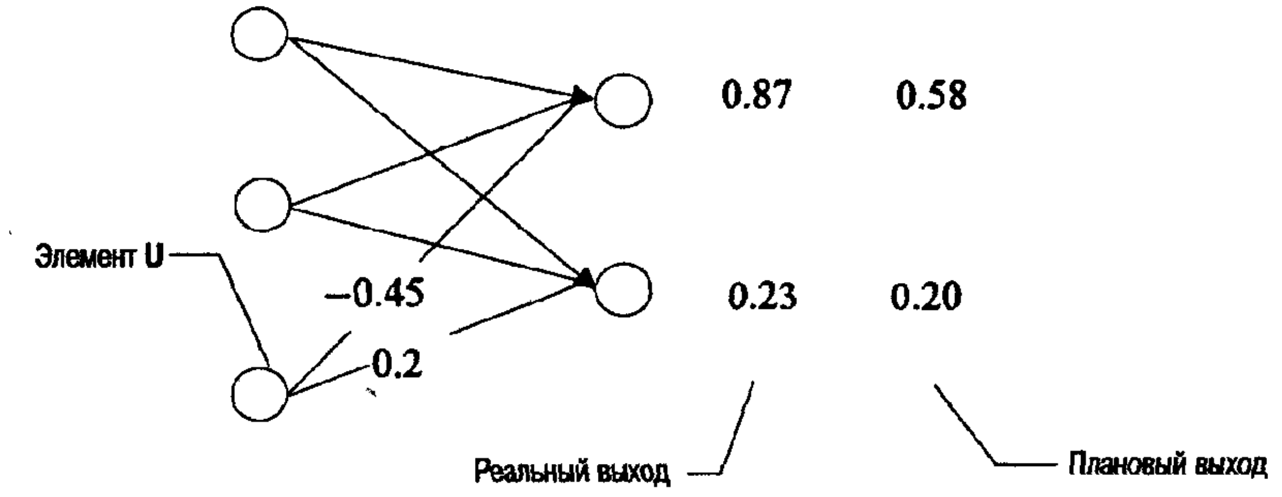
для первого слоя и

$$\begin{bmatrix} 0\,2 \\ 0\,2 \\ 0\,3 \end{bmatrix}$$

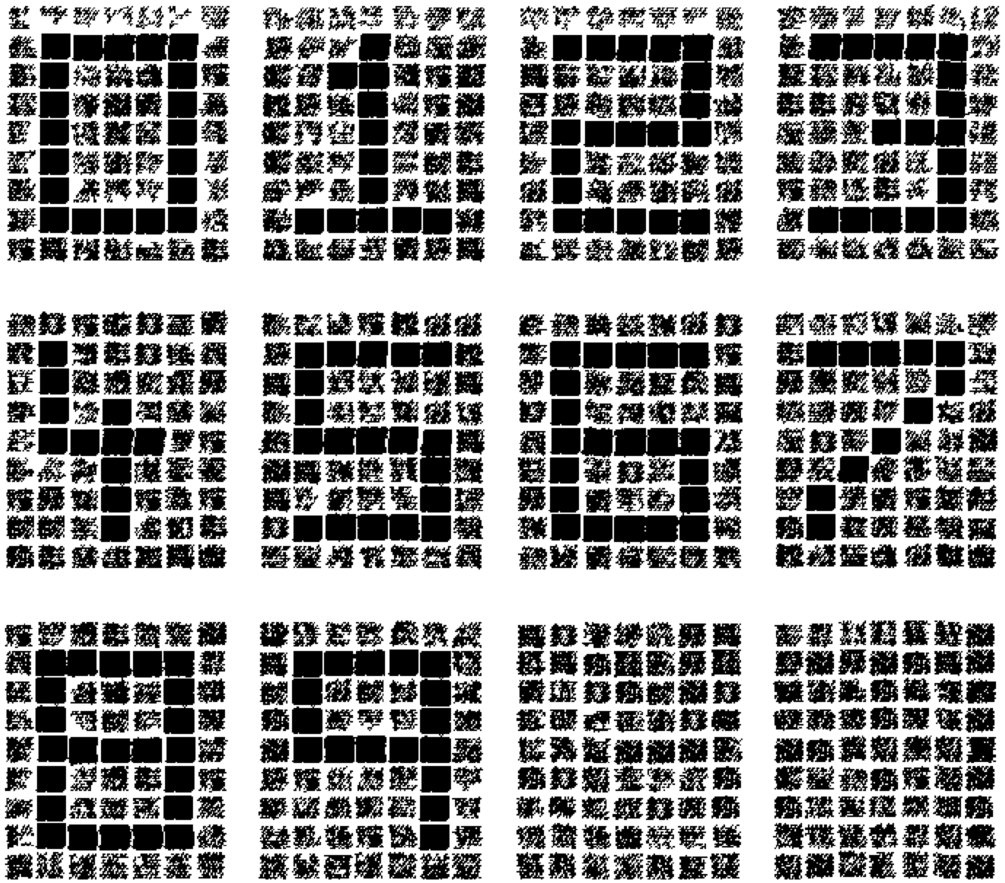
для второго слоя.

Пример. Вычислите ошибку скрытого элемента U

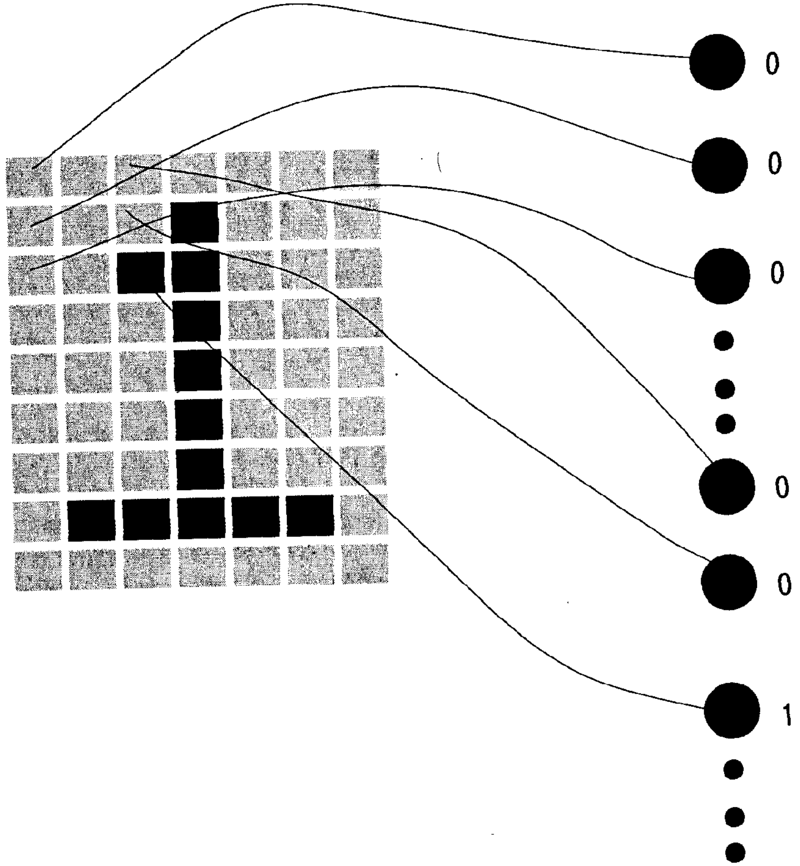
Вычислите ошибку для скрытого элемента U при условии, что значение его активности для обрабатываемого сетью образца равно 0 64.

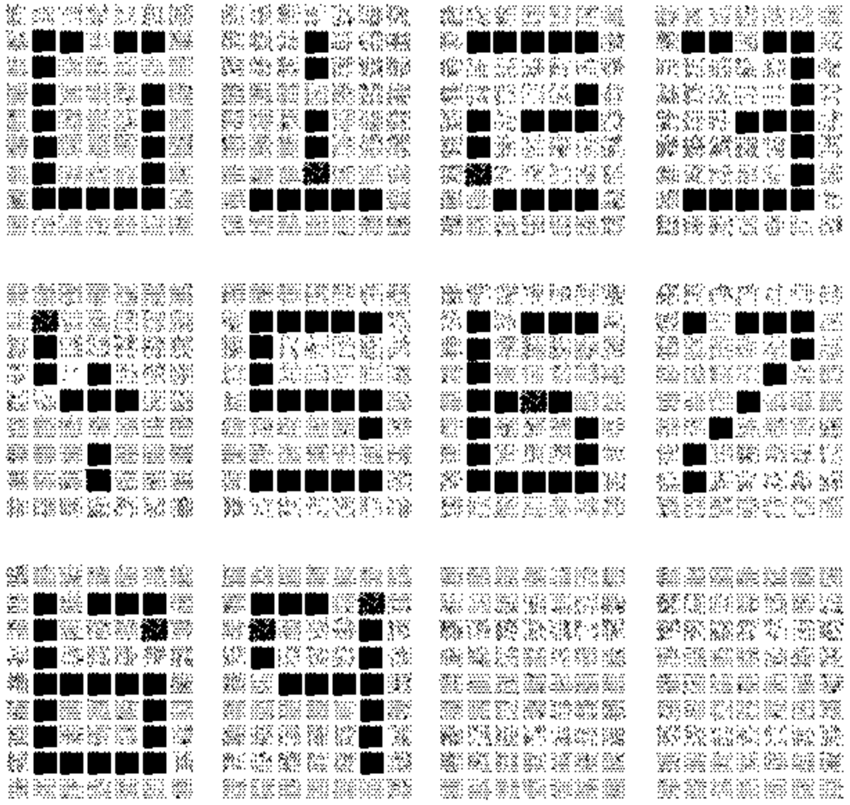


Использование сети с обратным распространением ошибки. Классификация чисел

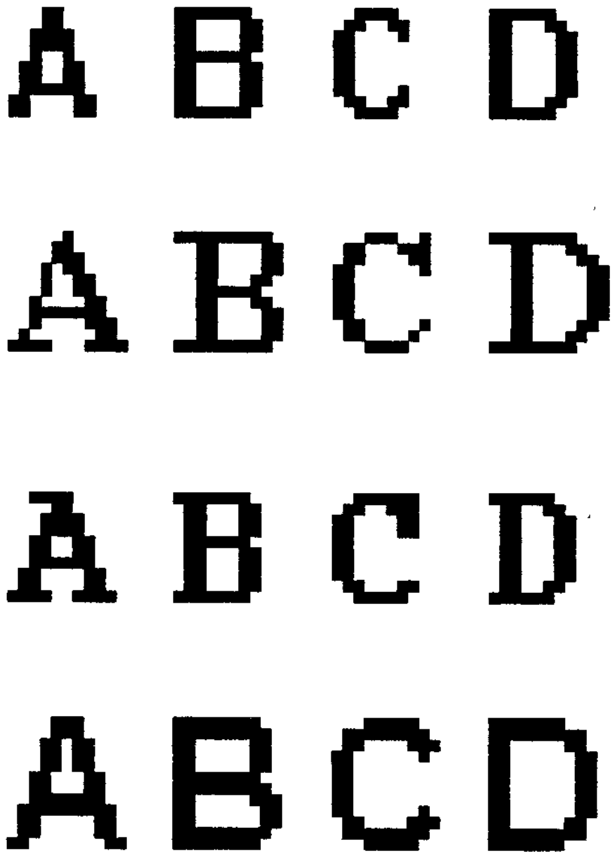


Использование сети с обратным распространением ошибки. Классификация чисел





Использование сети с обратным распространением ошибки. Классификация символов



Использование сети с обратным распространением ошибки. Прогнозирование погоды

	Температура (°C)		Осадки	Давление	Солнце
День	минимум	максимум	(мм)	(мбар)	(часы)
01	−1	4.8	0.7	1011	3.8
02	−1	2 8	0	1024	5.4
03	−5.3	3 6	0	1032	4.8
04	−5	6.6	2.8	1026.5	0
05	3.5	4 7	3	1019.5	0
06	1.5	7 9	0	1018 5	5.2
07	−1.4	9 5	0.2	1034.5	0
08	0.8	10.4	1.4	1028 5	0 2
09	1.8	10 4	0.4	1028.5	0
10	4	12	2	1020.5	0

Сети с радиальными базисными функциями

$$\varphi(\mathbf{x}) = [\varphi(x_1), \varphi(x_2), \quad , \varphi(x_M)]$$

$$net_j = \| \mathbf{x} - \mathbf{w}_j \|$$

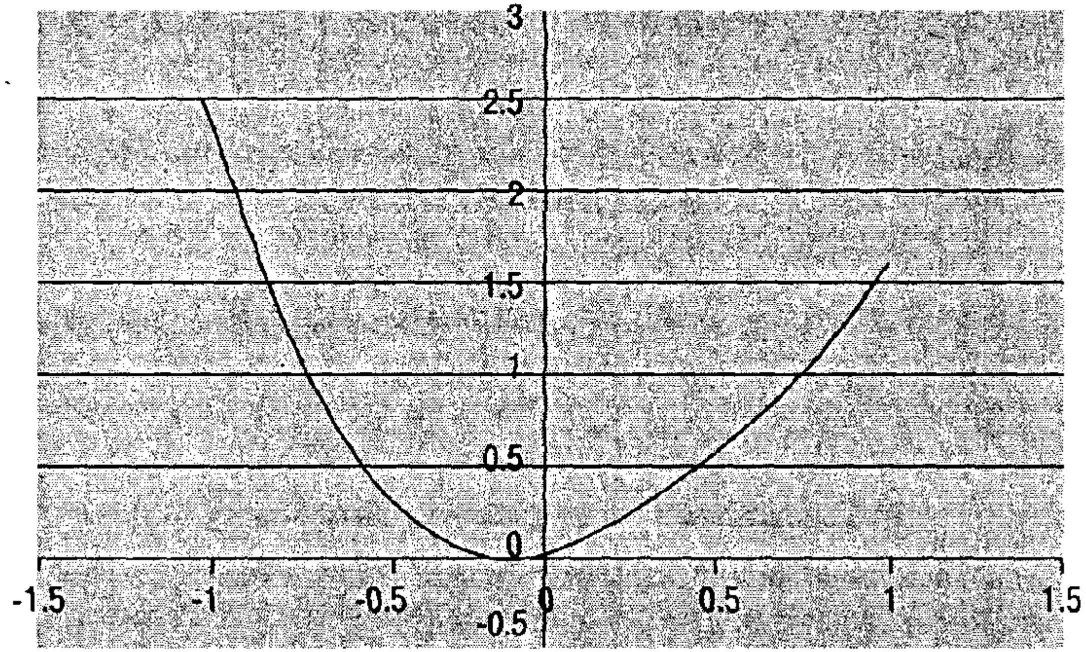
$$= \left[\sum_{i=1}^{n} (x_i - w_{ij})^2 \right]^{1/2}$$

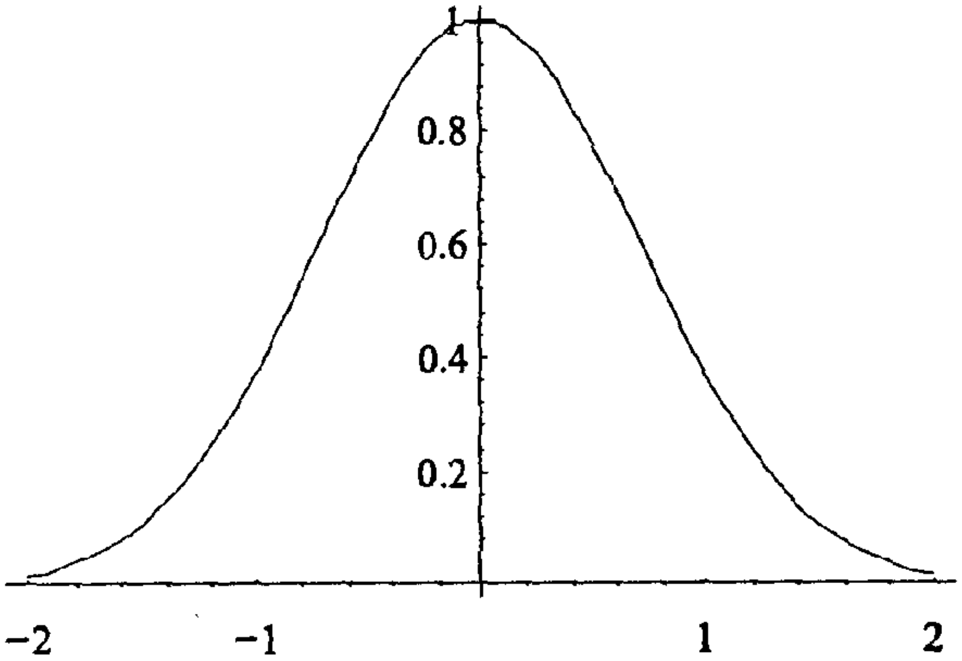
Пример. Сеть типа 2-2-1 с радиальными базисными функциями используется для решения проблемы XOR Первый слой весов задан матрицей

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

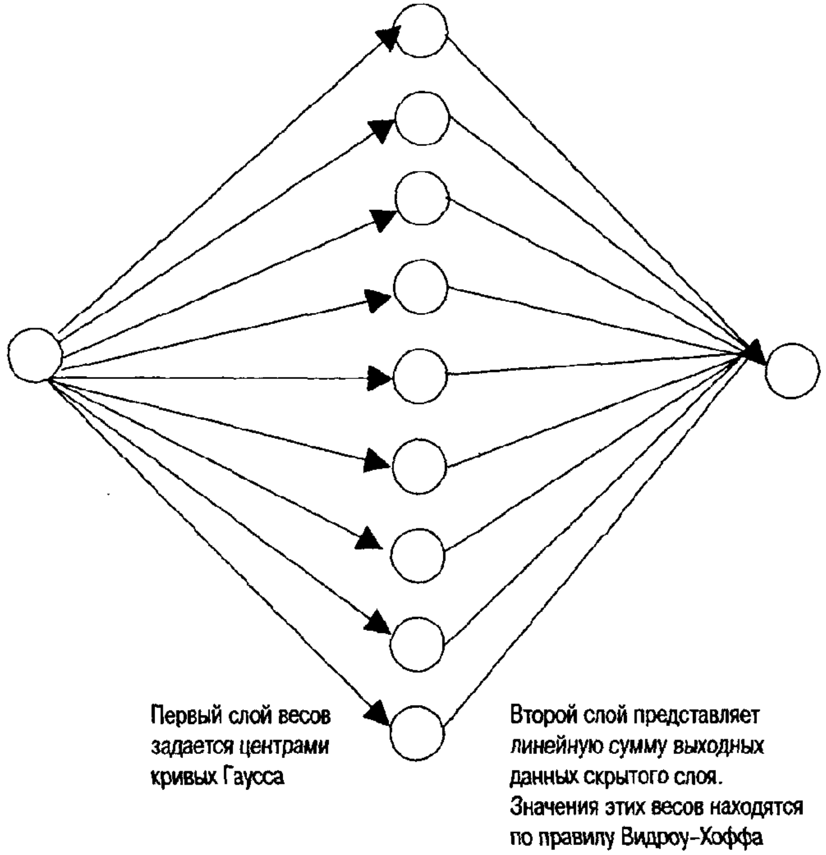
Для каждого вводимого образца XOR вычислите значения активности для всех скрытых элементов, если функция активности имеет вид $\varphi(net) = \exp[-net^2]$, где net является евклидовой нормой

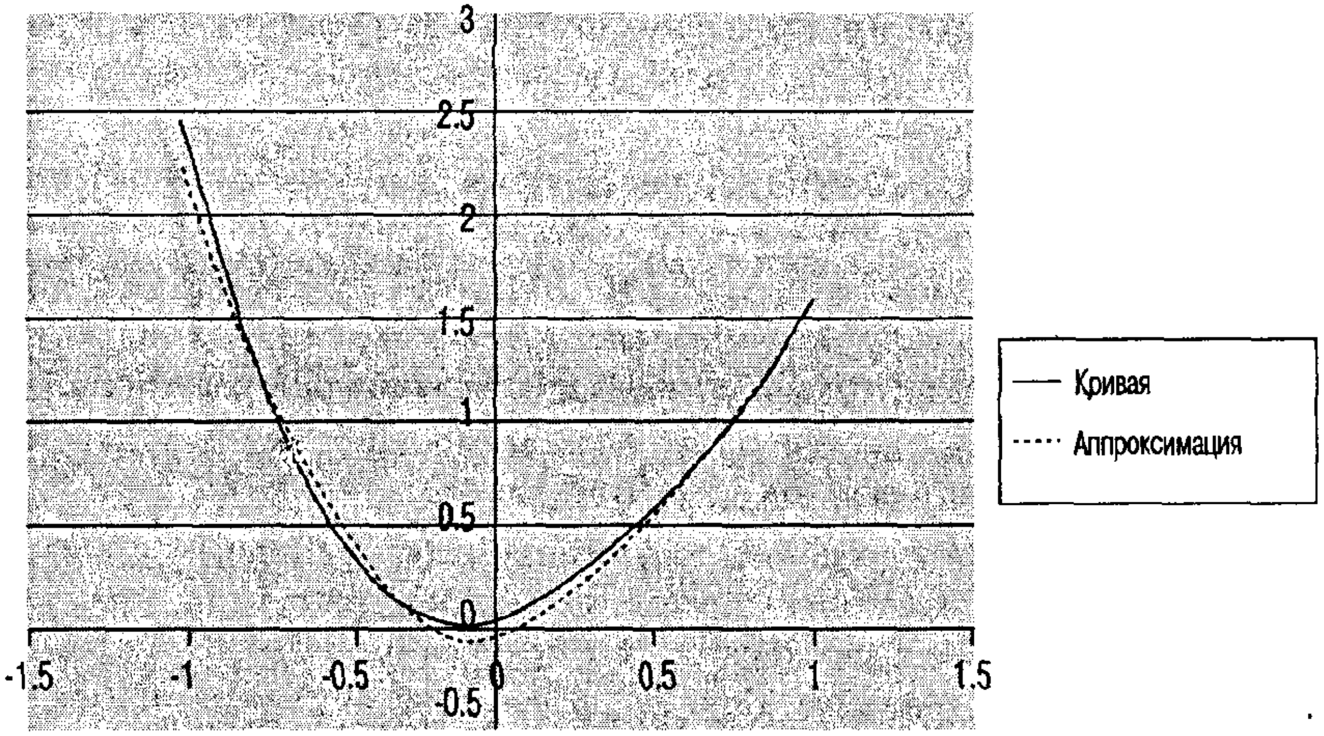
Сети с радиальными базисными функциями





Сети с радиальными базисными функциями





Спасибо за внимание!