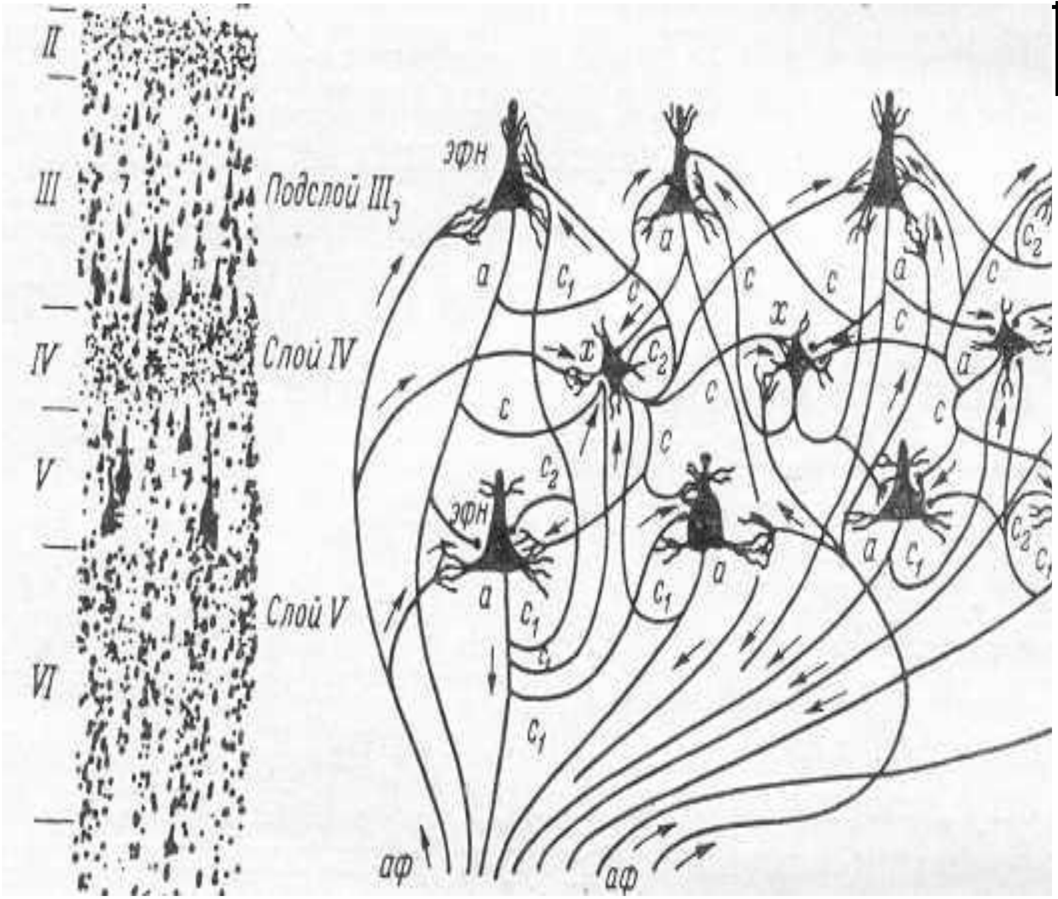
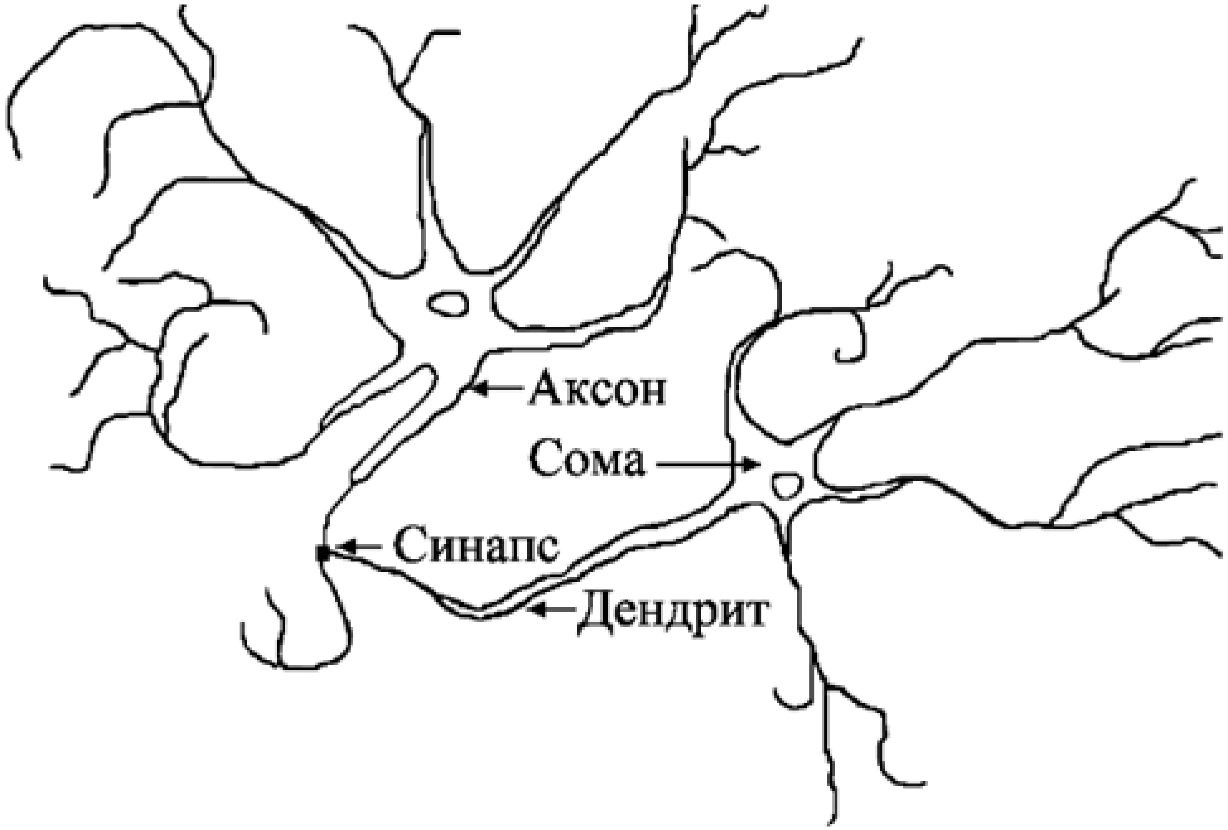
Лекція 1 до модуля 3

Основи нейронних мереж

Фрагмент нервової тканини кори мозку



Біологічний прототип



Теорема Колмогорова-Арнольда

Теорема Колмогорова-Арнольда (про яку часто не підозрюють практики) служить математичною основою нейронних мереж.

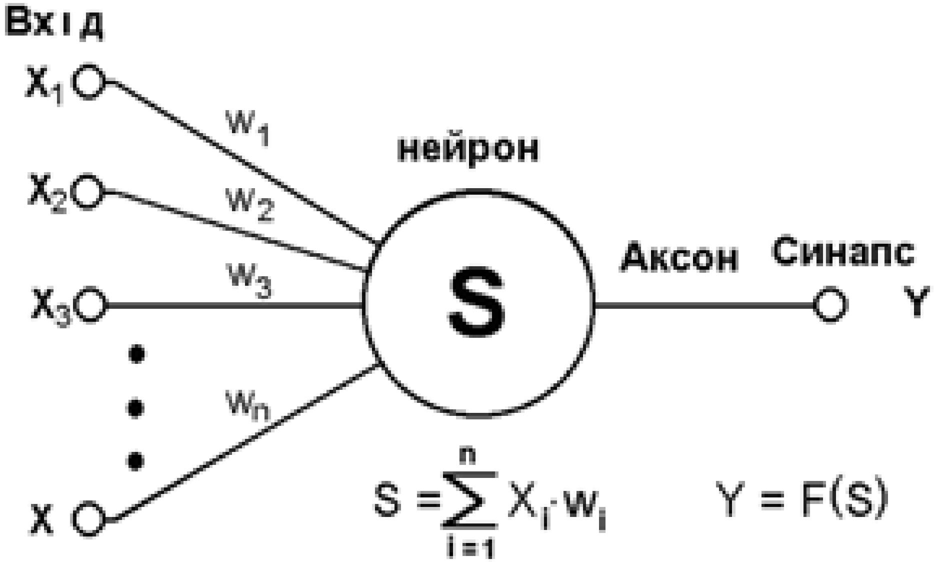
Теорема Колмогорова стверджує, що будь-яка безперервна функція *f,* визначена на n -вимірному одиничному кубі, може бути представлена у вигляді суми 2n+1 суперпозицій безперервних і монотонних відображень одиничних відрізків :

$$f(x_1,\ldots,x_n) = \sum_{q=1}^{2n+1} g_q\left(\sum_{p=1}^{n} \phi_{pq}(x_p)\right)$$

x = (x1, .., xn),   0 ≤ xi ≤ 1

Ліворуч в цій формулі стоїть довільна неперервна функція, визначена на багатовимірному кубі, справа функції визначені на відрізках [0, 1].

Штучний нейрон



Simulating a Neuron

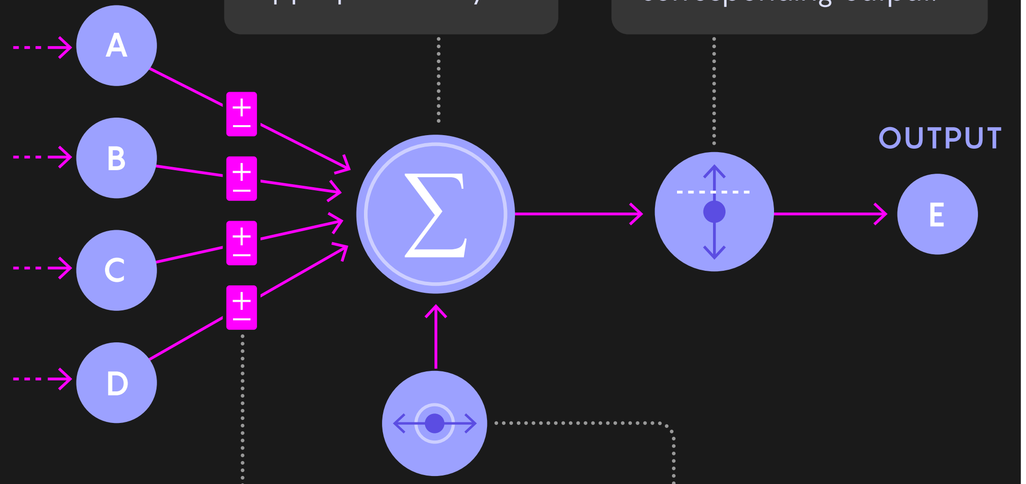
The interconnected devices in artificial neural networks conceptually resemble living neurons: Inputs from their "synapses" are weighted, tallied and translated into an output value.

Net input function

The neuron adds up the weighted inputs it receives in a task-appropriate way.

Activation function

The neuron determines whether the summed inputs meet a threshold value and produces a corresponding output.



Weights

Input values are modified to reflect their relative importance to the determination of an output target.

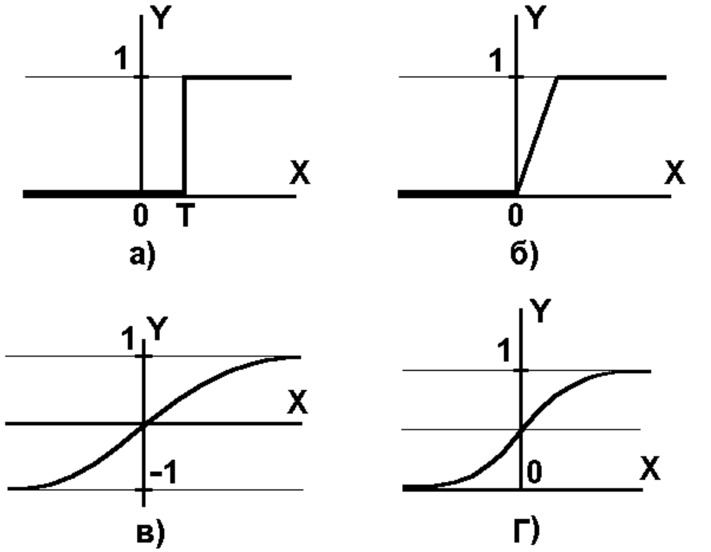
Bias

The neuron "learns" through feedback adjustments to its weights and function values.

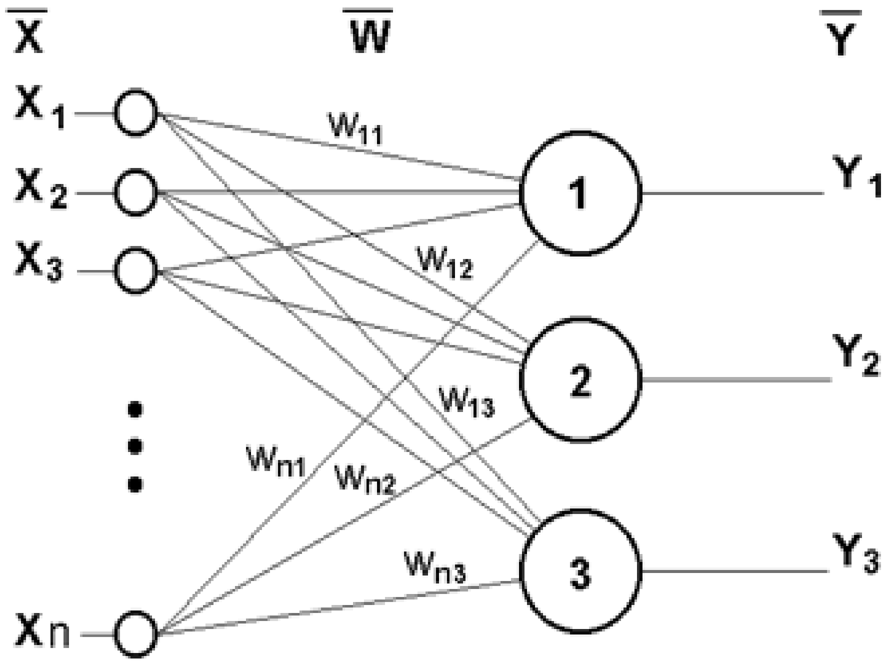
Штучний нейрон

$$S_p = \sum_{i=0}^{n} W_i \overline{X}_{pi}$$

$$F(S) = \frac{2}{1+e^{-aS}} - 1$$



Одношаровий персептрон



$$y_j = F\left[\sum_{i=1}^{n} x_i \cdot w_{ij}\right]$$

Лабораторна робота №5

Завдання: Побудувати бінарний класифікатор з допомогою моделі штучного нейрону.

Алгоритм лабораторної роботи №1

1.Обираємо задачу класифікації образів за неявними ознаками $\mathbf{A_j}$ та готуємо базу прикладів для навчання (навчальну вибірку) у вигляді таблиці:

Xpi	**A1**	**A2**	**...**	**An**	**dp**
P1	X_{11}	X_{12}		X_{1n}	$\mathbf{d_1}$
P2	X_{21}	X_{22}		X_{2n}	$\mathbf{d_2}$
...					
$\mathbf{P_M}$	X_{M1}	X_{M2}		X_{Mn}	$\mathbf{d_M}$

Де X_{pi} - числові значен

d_p - ознака класу, d_p = +1 – I клас, d_p = -1 – II клас.

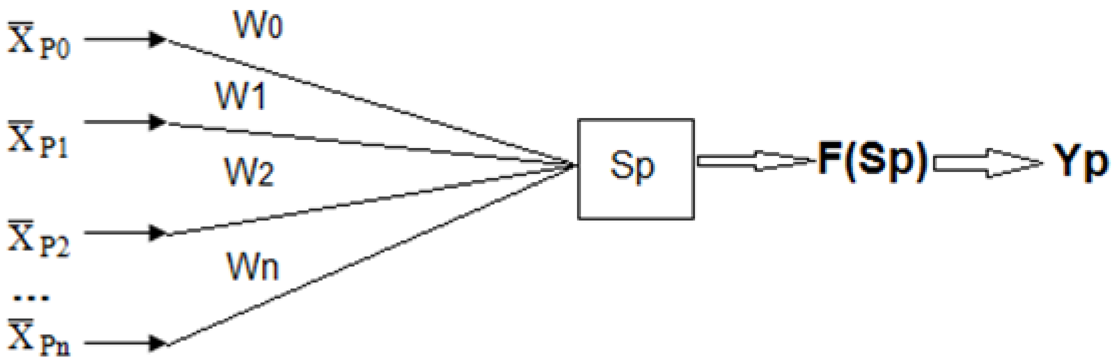
Число ознак повинно бути близько 5, а число прикладів більше 10, контрольна вибірка близько 5.

Лабораторна робота №5

2. Виконуємо природну нормалізацію ознак:

$$\overline{X_{pi}} = \frac{X_{pi} - X_{\min i}}{X_{\max i} - X_{\min i}}$$

3. Ініціалізуємо нейрон:



Лабораторна робота №5

Початкові значення вагових параметрів –невеликі, випадкові:

Wi = 0.01*rand(-1,+1).

Параметр ітерації: t=0.

Функцію активації приймаємо у вигляді:

$$F(S) = \frac{2}{1+e^{-aS}} - 1$$

, де *a=2.*

Задаємо параметр швидкості навчання *η =0.9 і* точність навчання

ε =0.01.

p=1

Лабораторна робота №5

4. Основний цикл навчання за методом Відроу-Хебба:

а) Обираємо випадково новий приклад з навчальної вибірки з ознаками і обчислюємо зважену суму, вихідне значення і помилку:

Yp=F(S_p),

Δp=(Y_p – d_p)

$$S_p = \sum_{i=0}^{n} W_i \overline{X}_{pi}$$

b) Обчислюємо вагові параметри для ітерації (t+1):

$$W_i(t+1) = W_i(t) - \eta \overline{X}_{pi} \Delta_p$$

c) Повторити b) для i=1,2,...,n.

d) Перейти до а) до вичерпання всіх прикладів навчальної вибірки.

e) Обчислюємо сумарну помилку:

f) При умові E > ε^2 продовжуємо ітерації *t=t+1*, переходимо до а),

$$E = \frac{1}{2} \sum_{p=1}^{M} \Delta_p$$

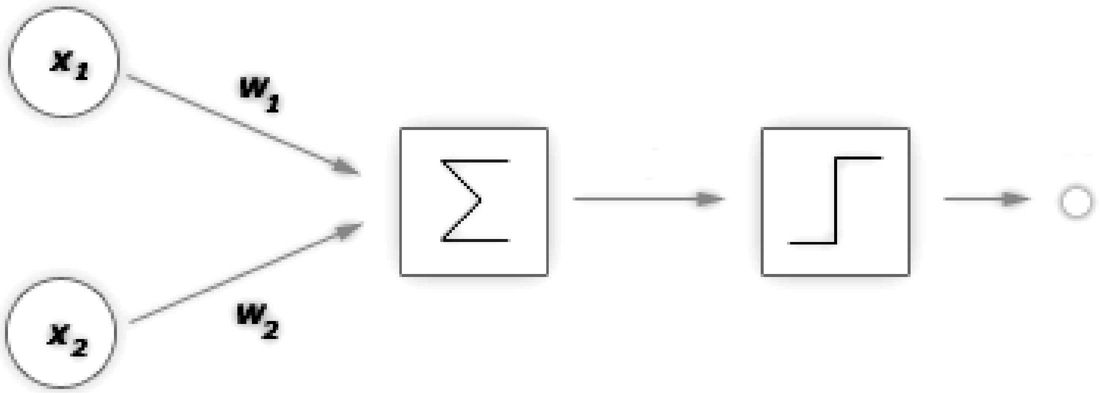
при умові E <= ε^2 - вихід.

Проблема лінійного розподілу в НМ

Часто, для того, щоб продемонструвати обмежені можливості одношарових персептронів при рішенні завдань удаються до розгляду так званої проблеми XOR - що виключає АБО.

Суть завдання полягають в наступному. Дана логічна функція XOR - що виключає АБО. Це функція від двох аргументів, кожен з яких може бути нулем або одиницею. Вона набуває значення, коли один з аргументів дорівнює одиниці, але не обоє, інакше . Проблему можна проілюструвати за допомогою одношарової одно нейронної системи з двома входами, показаної на малюнку нижче.

Проблема лінійного розподілу в НМ



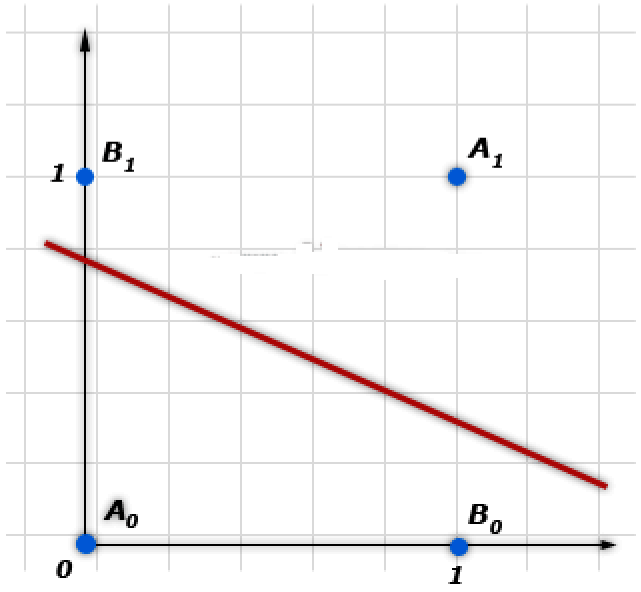
Проблема лінійного розподілу в НМ

Позначимо один вхід через X1, а інший через X2, тоді усі їх можливі комбінації складатимуться з чотирьох точок на площині.

Один нейрон з двома входами може сформувати вирішальну поверхню у вигляді довільної прямої:

$$X1*W1+X2*W2 = T$$

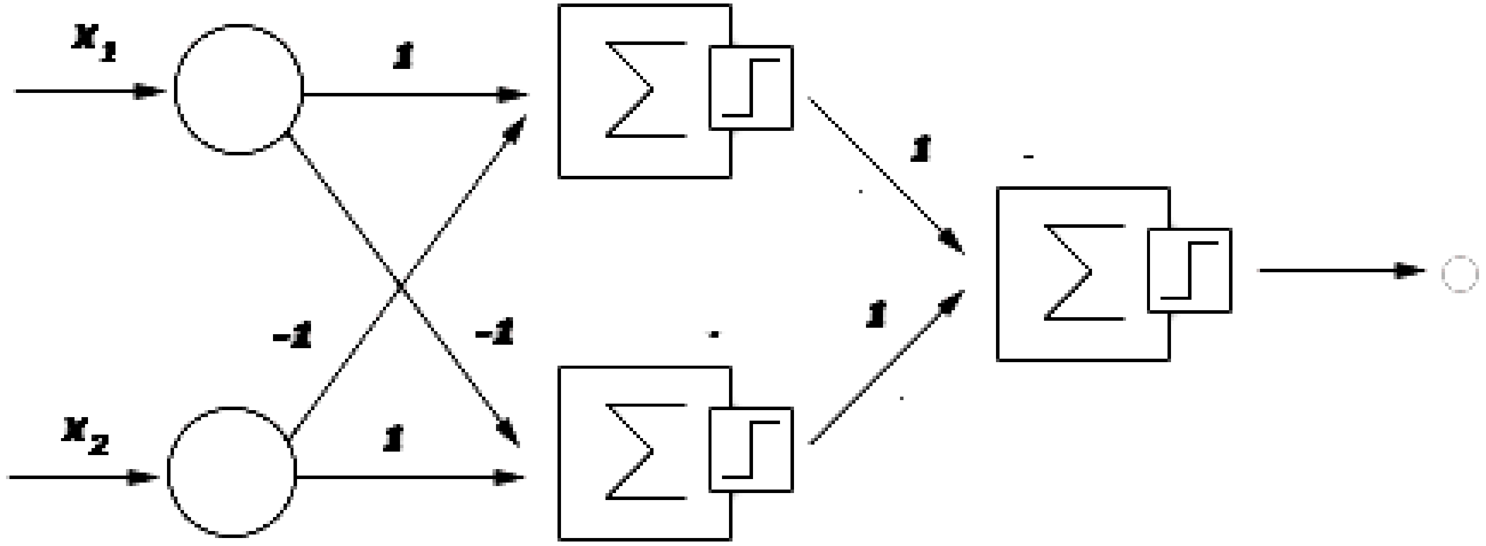
Проблема лінійного розподілу в НМ



Проблема лінійного розподілу в НМ

- Це означає, що які б значення не приписувалися вагам і порогу, одношарова нейронна мережа нездатна відтворити співвідношення між входом і виходом, потрібне для представлення функції XOR.

Проте функція XOR легко формується вже двошаровою мережею, причому багатьма способами. Розглянемо один з таких способів. Модернізуємо мережу, додавши ще один прихований шар нейронів



x1
x2
1
-1
-1
1
1
1

Проблема лінійного розподілу в НМ

- Кожен з двох нейрон першого шару формує вирішальну поверхню у вигляді довільної прямої (ділить площину на дві напівплощини), а нейрон вихідного шару об'єднує ці два рішення, утворюючи вирішальну поверхню у вигляді смуги, утвореної паралельними прямими нейронів першого шару,

Проблема лінійного розподілу в НМ