

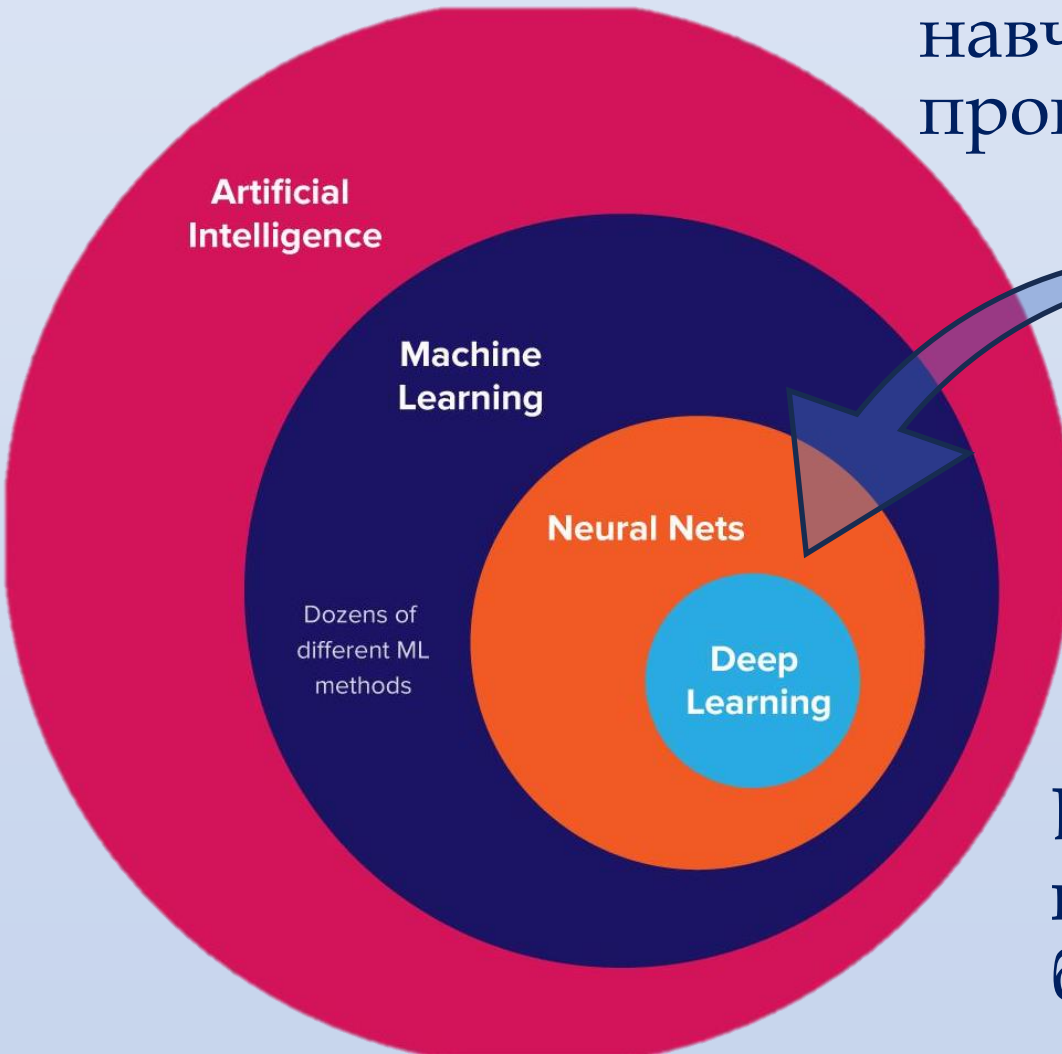
ОСНОВИ СИСТЕМ ШТУЧНОГО ІНТЕЛЕКТУ, НЕЙРОННИХ МЕРЕЖ та ГЛИБОКОГО НАВЧАННЯ

Модуль 5. Глибоке навчання

**Лекція 5.1. Природний нейрон. Правило
Хебба. Штучний нейрон. Функція активації.**

Ареал ШІ

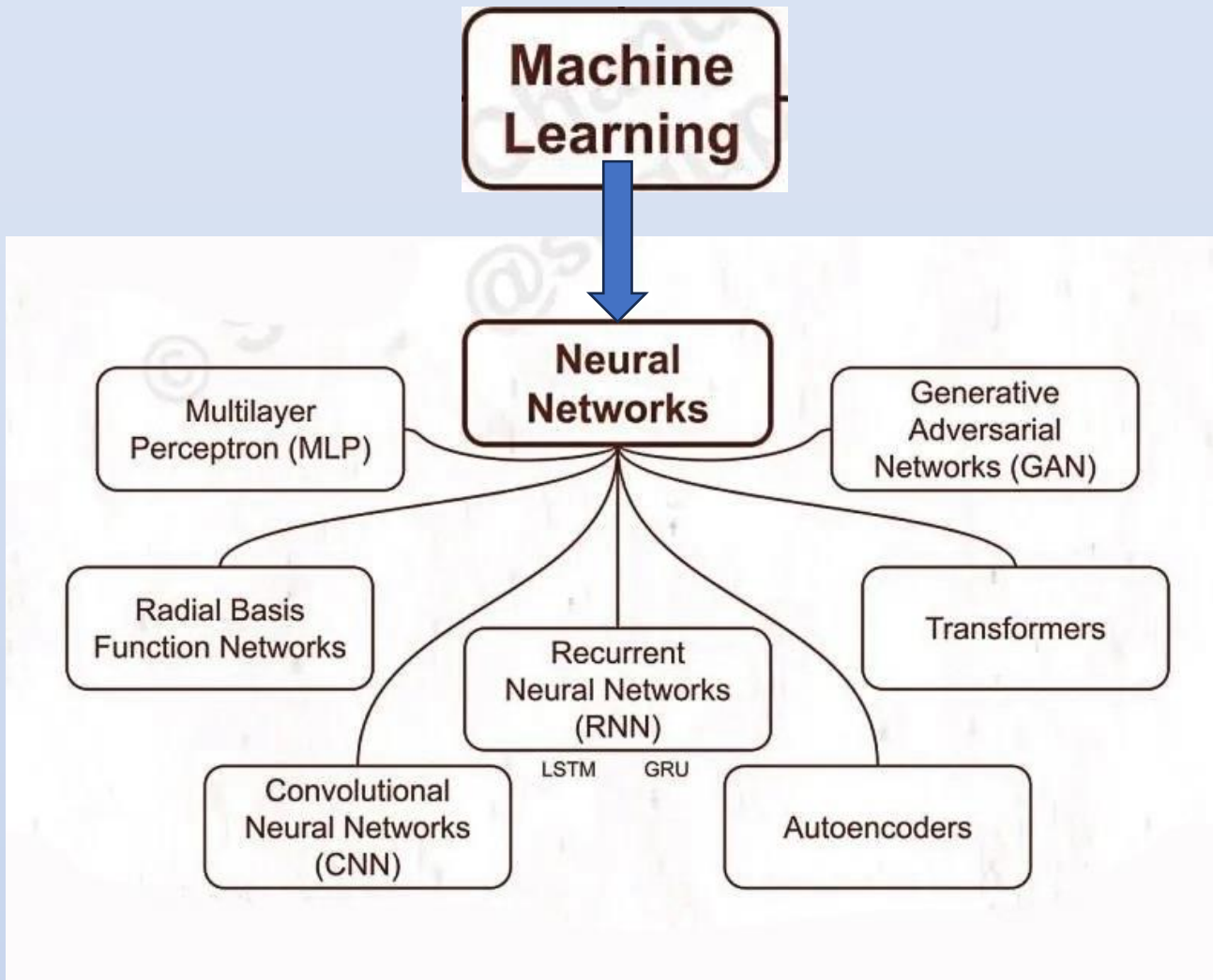
ML – машинне навчання –
підрозділ ШІ, де системи
навчаються без явного
програмування



**NN – нейронна
мережа –**
математична модель,
що імітує роботу
людського мозку

**DL – глибоке
навчання –** навчання
багатошарових NN

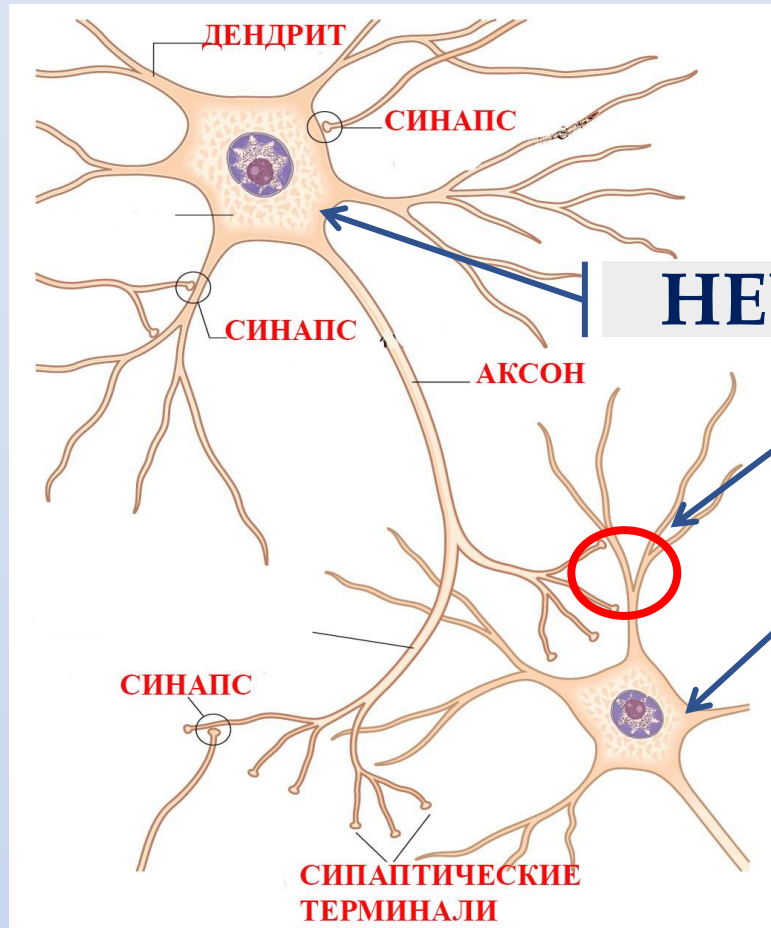
Моделі ML



!!! Вирішують всі (?) завдання ML

Природна нейрона мережа

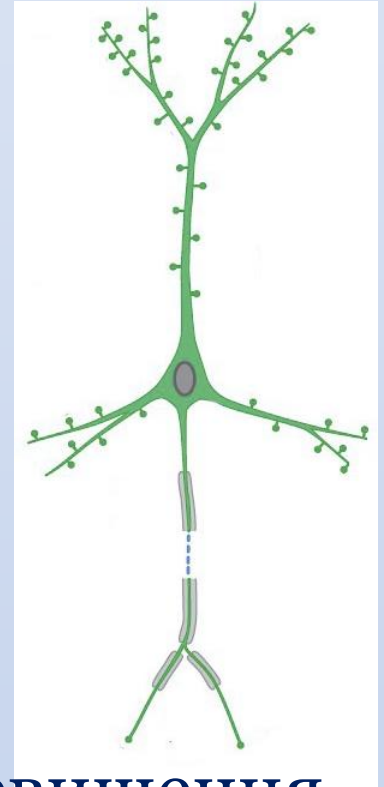
Нейрон – декілька десятків тисяч вхідних синапсів та біля десяти вихідних



НЕЙРОН 1

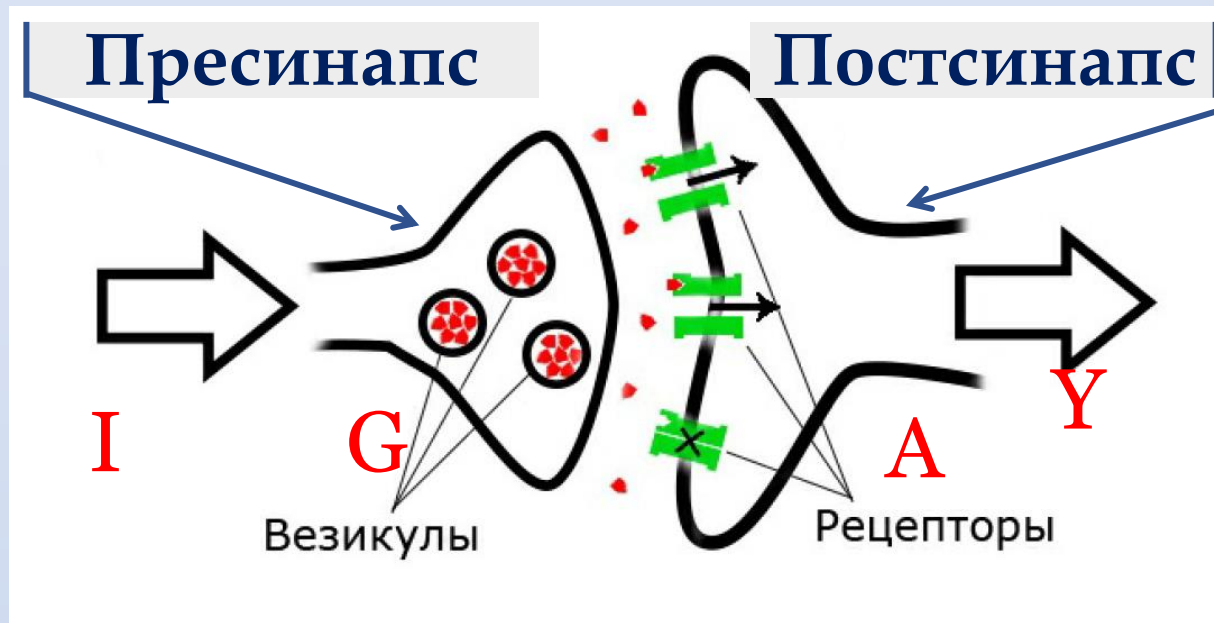
Синаптический
звязок

НЕЙРОН 2



Нейрон накопичує потенціал і після перевищення порога – по аксону проходить хвиля імпульсів, що досягає синаптичних терміналів (хвиля потенціалу дії).

Синапс



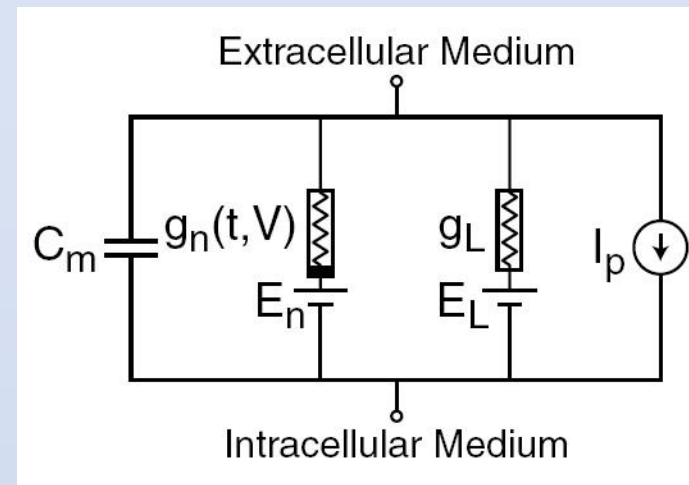
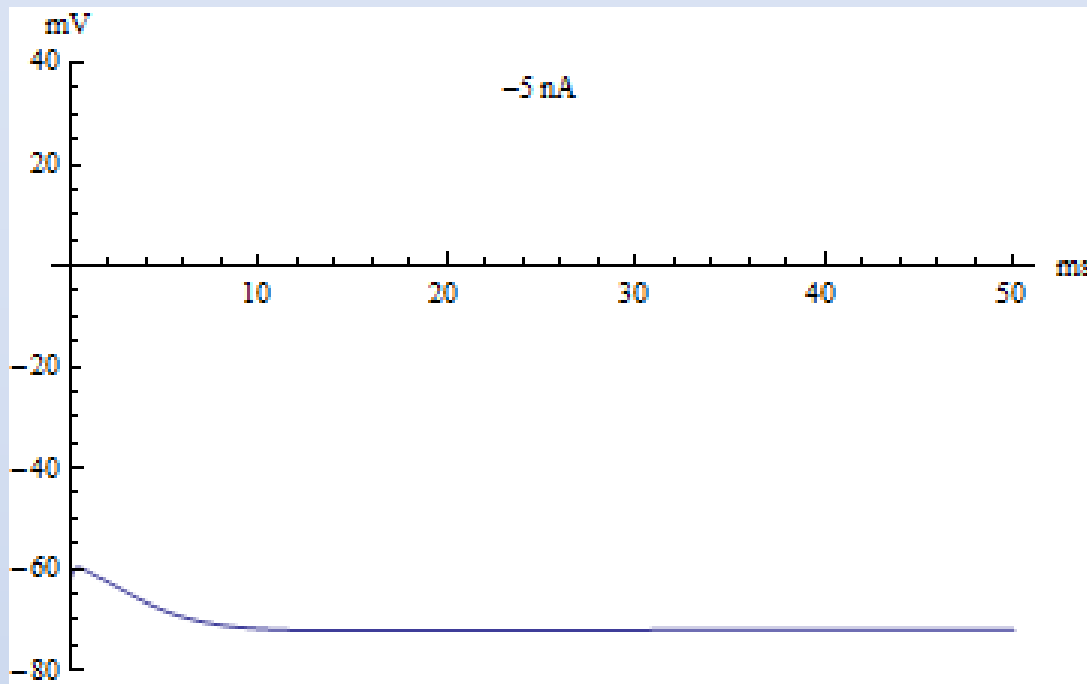
Плазматична мембрана пре-синаптичного нейрона вступає у взаємодію з мембраною постсинаптичного нейрона

Везикула - бульбашка з порцією речовини нейромедіатора.
Нехай **g** кількість везикул (різне).

Рецептор – реагує на нейромедіатор.
Нехай **a** – кількість (вага), що визначає чутливість постсинапсу.

“СИЛА СИНАПСУ” $V = g * a$

Модель збудження Ходжкіна – Хакслі



**Істотна
нелінійність**

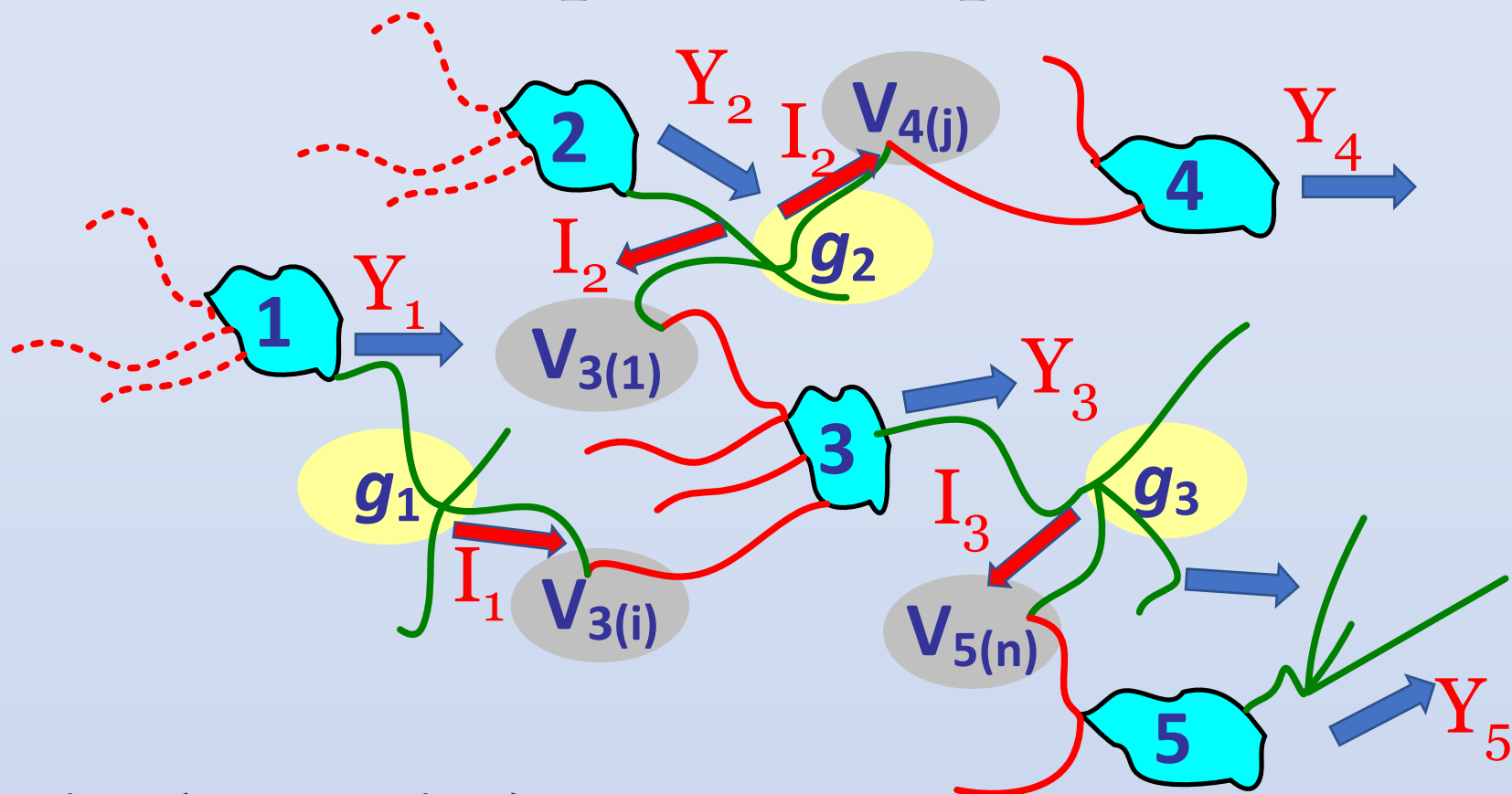
Якщо вихідний потенціал недостатній, модель залишається в рівноважному стані. Якщо потенціал переходить через певний поріг, модель відповідає одним імпульсом. Якщо потенціал значно перевищує даний поріг, то модель відповідає серією імпульсів.

Правило ХЕББА

Якщо аксон нейрона 1 знаходиться досить близько, щоб збуджувати нейрон 2, і неодноразово або постійно бере участь у її збудженні, то спостерігається деякий процес, що веде до підвищення ефективності 1, як однієї з збуджуючих клітин 2»

- **Причинно-наслідковий зв'язок.** Зв'язок між нейронами має тенденцію до посилення.
- **Розташування змін.** Посилення зв'язаності відбувається або за рахунок зміни провідності синапсу, або за рахунок зміни метаболічних особливостей самих клітин.
- **Сукупне збудження.** Збудження постсинаптичного нейрона може бути здійснено лише з допомогою одного пресинаптичного стимулу.

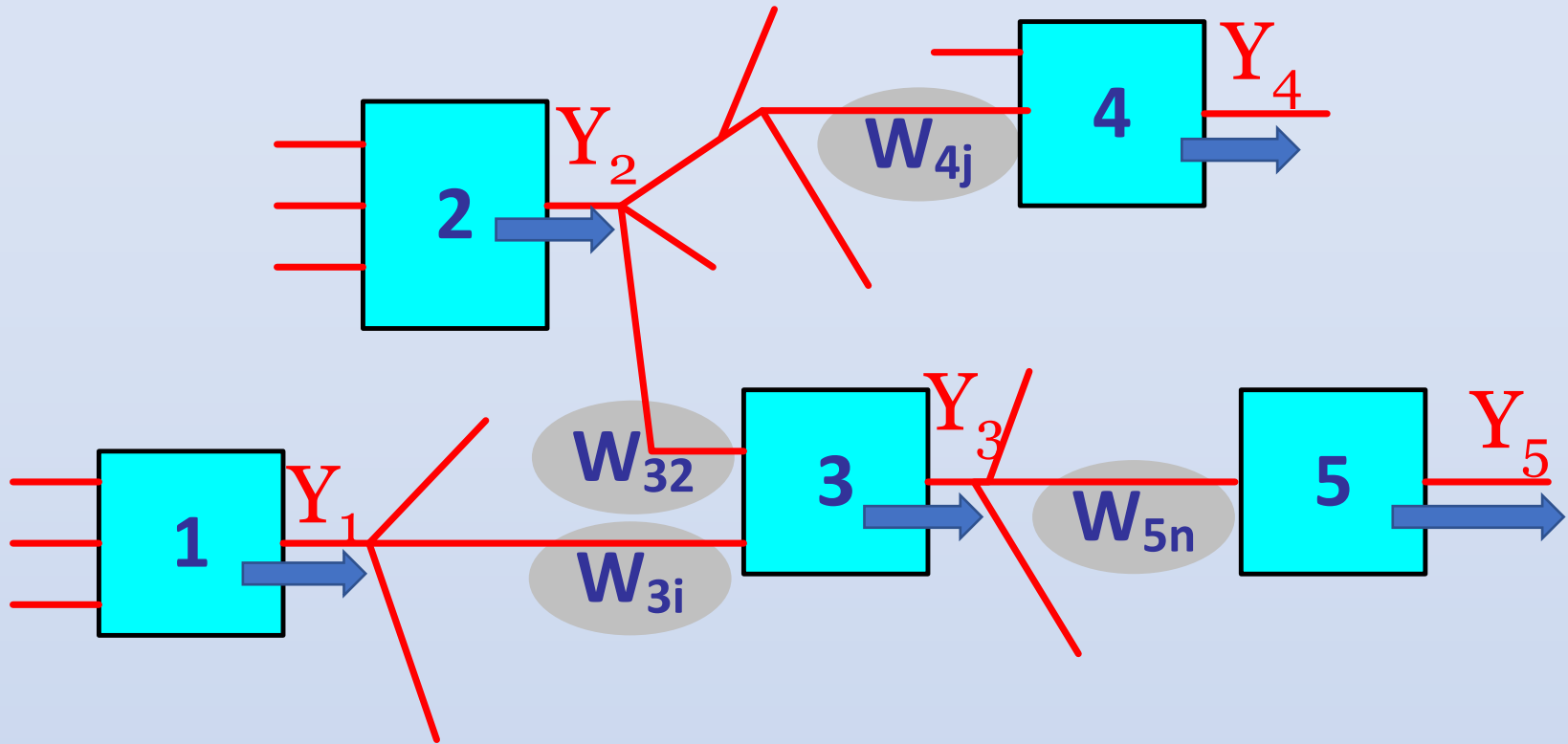
Мережа нейронів



Вхід (потенціал) синапсу I - деяка частка g виходу пресинаптічного аксона Y .

Вихід нейрона Y – складна (нелінійна, нестатична) функція входів синапсів та його ваг W .

Мережа нейронів

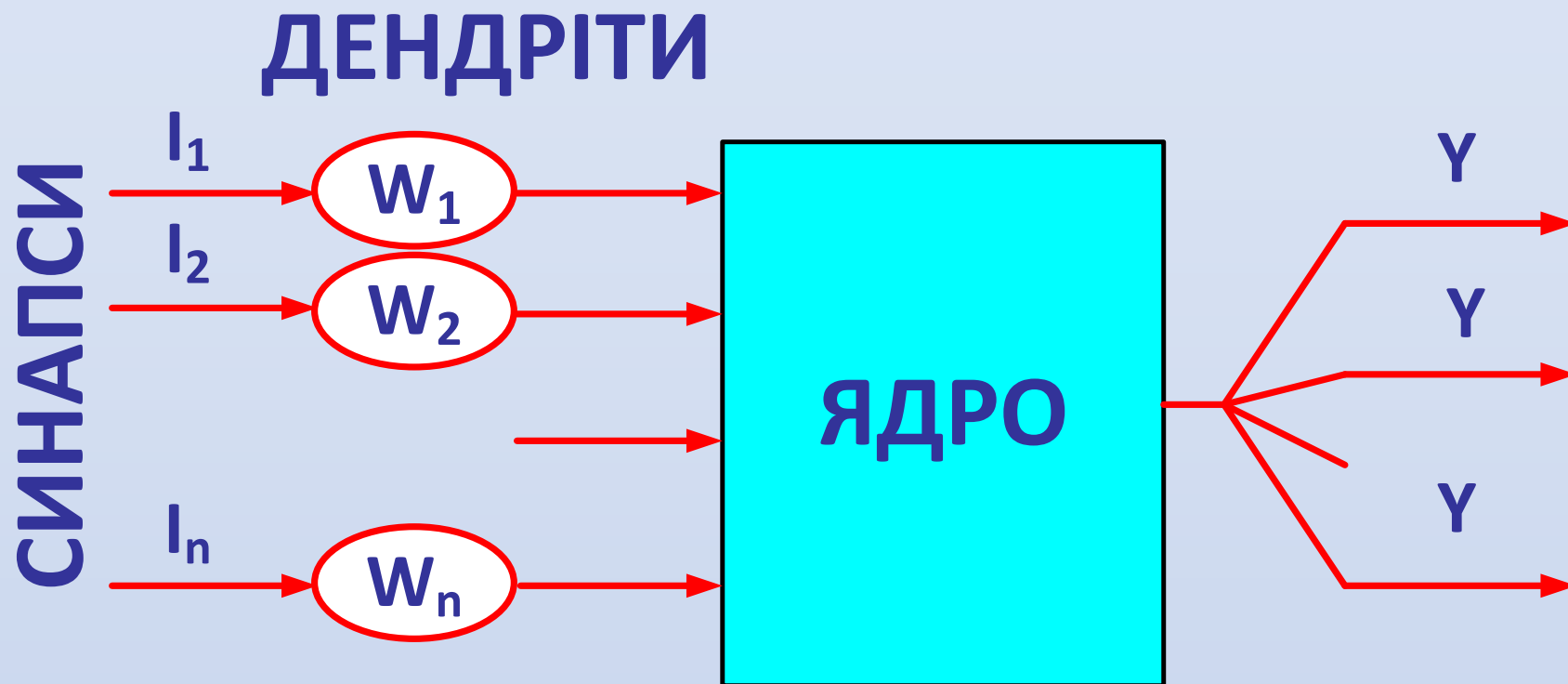


СПРОЩЕНО:

$I = Y$ – вихідний сигнал аксона = вхідний сигнал (через синапс) дендриту.

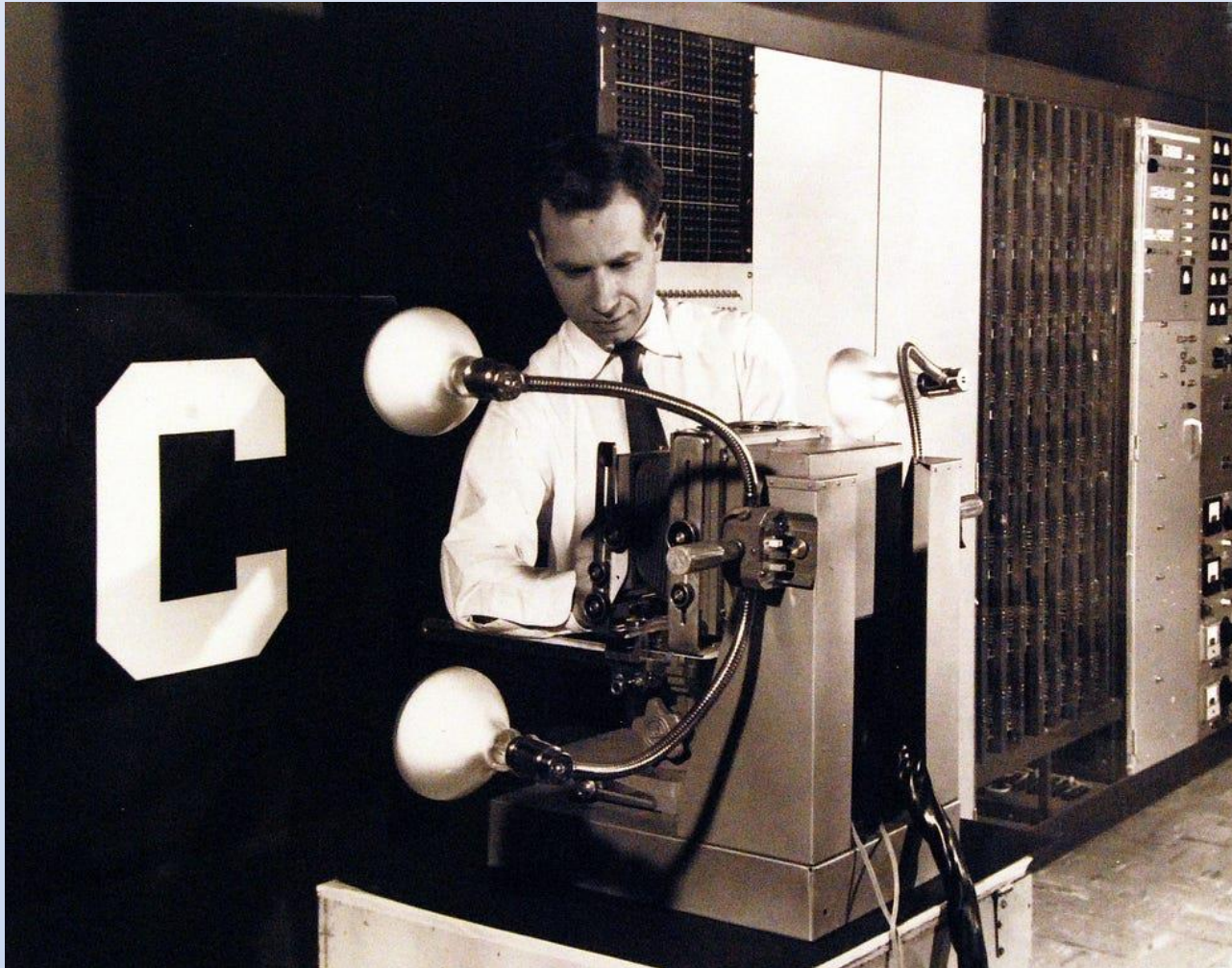
$W = g * V$ - вага зв'язку синапсу.

Нейрон McCulloch-Pitts (1943)



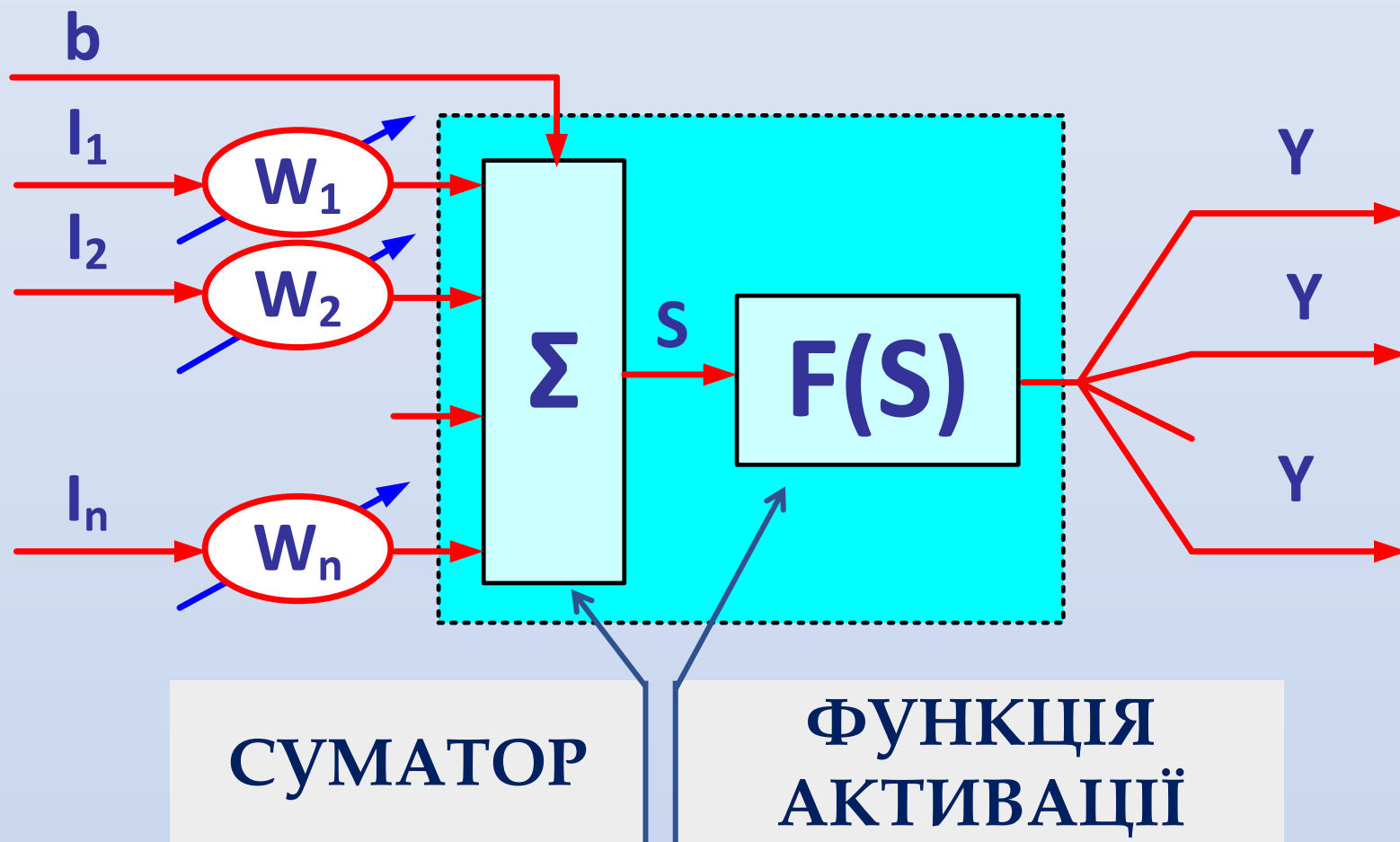
Функціонування синапсу – зв'язок аксона та дендрита моделюється **вагою** зв'язку **W** .

Штучний нейрон (персептрон)



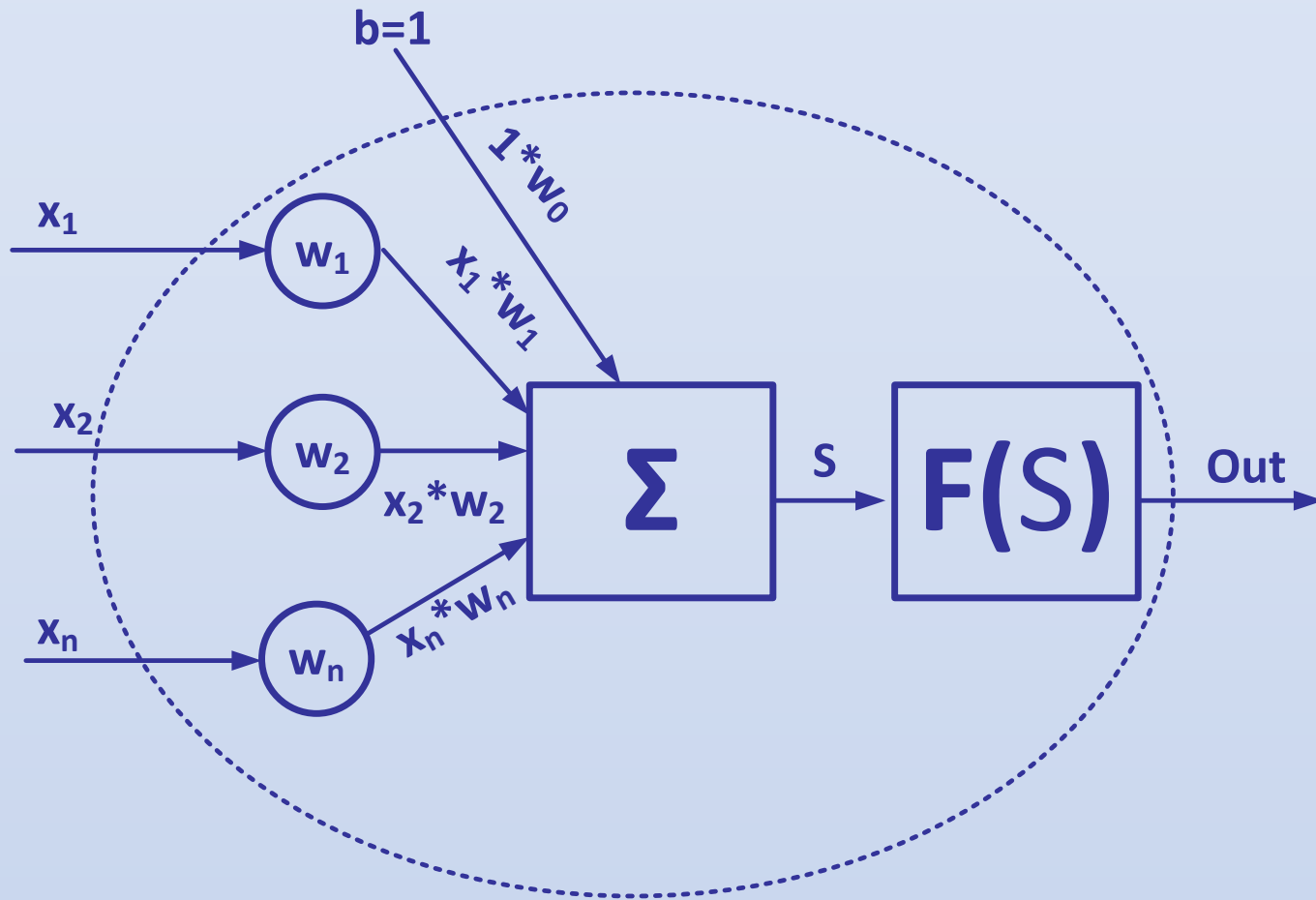
Scientists Warren McCulloch and Walter Pitts invented the perceptron in **1943**. The first hardware (Mark I perceptron) for perceptrons was developed in the **1950s** and 1960s by the scientist Frank Rosenblatt.

Штучний нейрон



$$S = b + \sum_{j=1}^n I_j * w_j \quad Y = F(S)$$

Штучний нейрон (формально)



$$\text{out} = F(x_0, x_1, x_2, \dots, x_n) = \\ F(1 * w_0 + x_1 w_1 + \dots + x_n w_n) = F(\langle x, w \rangle)$$

Функція активації нейрону

Функція активації нейрона $F(S)$ моделює збудження нейрона.

Формує вихідний сигнал Y нейрона в залежності від виваженої суми вхідних сигналів та деяких параметрів збудження.

У природних нейронах – складна залежність (див. модель Ходжкіна-Хакслі), що визначає силу та швидкість формування потенціалу дії аксона нейрона.

У штучних нейронах моделюється нелінійною функцією однієї змінної. Прийнято (?), що вихідний сигнал штучного нейрона може змінюватися в межах $0 \leftrightarrow 1$ або $-1 \leftrightarrow +1$.

Функція активації | Activation function

Важливі властивості функції активації:

Нелінійність

для нелінійної функції показано, що дворівнева нейронна мережа буде універсальним апроксиматором функцій.

Безперервна диференційованість

- бажана властивість (методи градієнтного спуску).

Гладкість з монотонною похідною – вищий ступінь спільності.

Монотонність - гарантує опуклість поверхні помилок (одношарові архітектури).

Наближення до тотожної функції $F(S) = S$ в початку координат.

Функція активації | Activation function

Рідж-функції — функції від лінійної комбінації вхідних змінних.

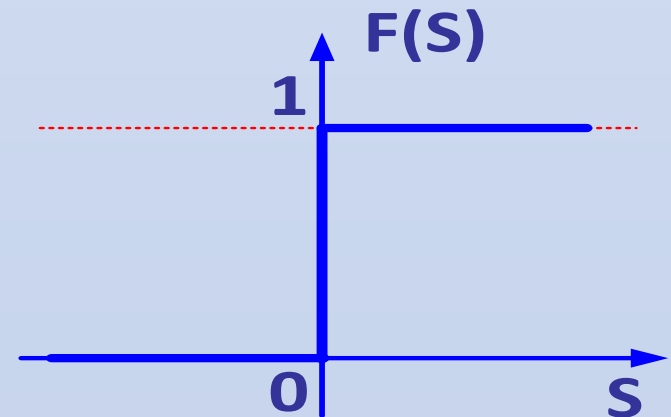
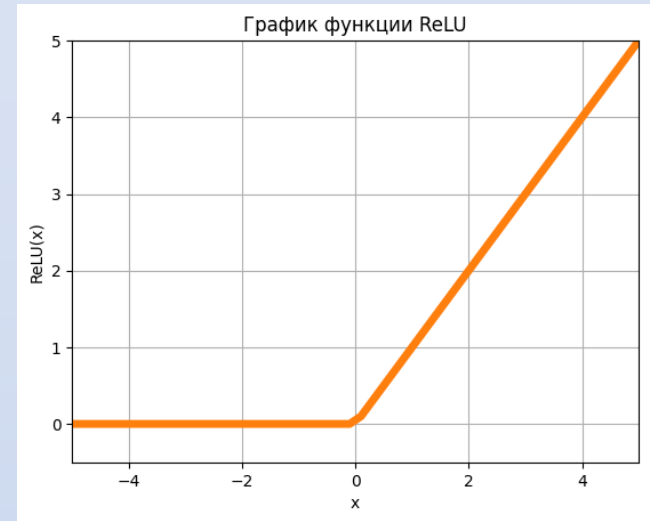
$$\text{RelU} \quad F(s) = \begin{cases} 0, & s < 0 \\ s, & s \geq 0 \end{cases}$$

$$F'(s) = \begin{cases} 0, & s < 0 \\ 1, & s \geq 0 \end{cases}$$

Двійковий крок (Step)

$$F(s) = \begin{cases} 0, & s < 0 \\ 1, & s \geq 0 \end{cases}$$

$$F'(s) = \begin{cases} 0, & s \neq 0 \\ ?, & s = 0 \end{cases}$$



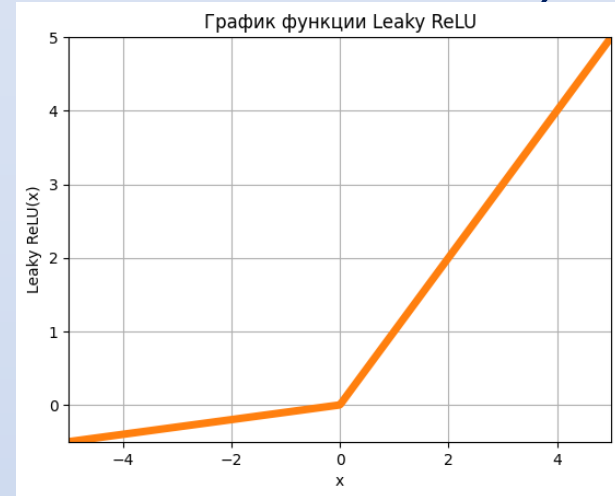
Функція активації | Activation function

ReLU модифікації

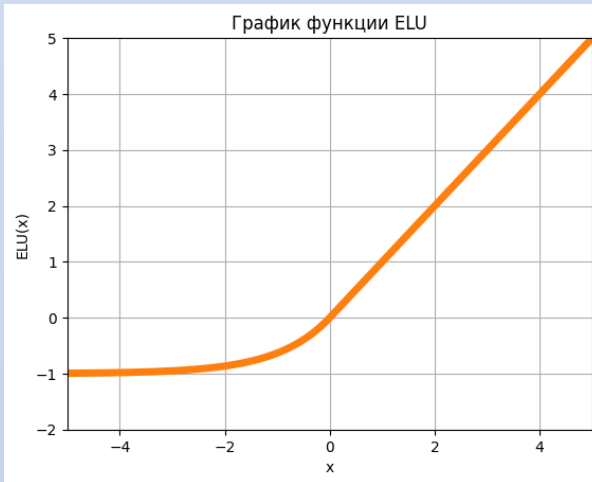
Leaky ReLU (LReLU) ! (коли $\alpha \rightarrow$ вага PReLU)

$$F(s) = \begin{cases} a * s, & s < 0 \\ s, & s \geq 0 \end{cases}$$

$$F'(S) = \begin{cases} 0, & s < 0 \\ a, & s \geq 0 \end{cases}$$



ELU



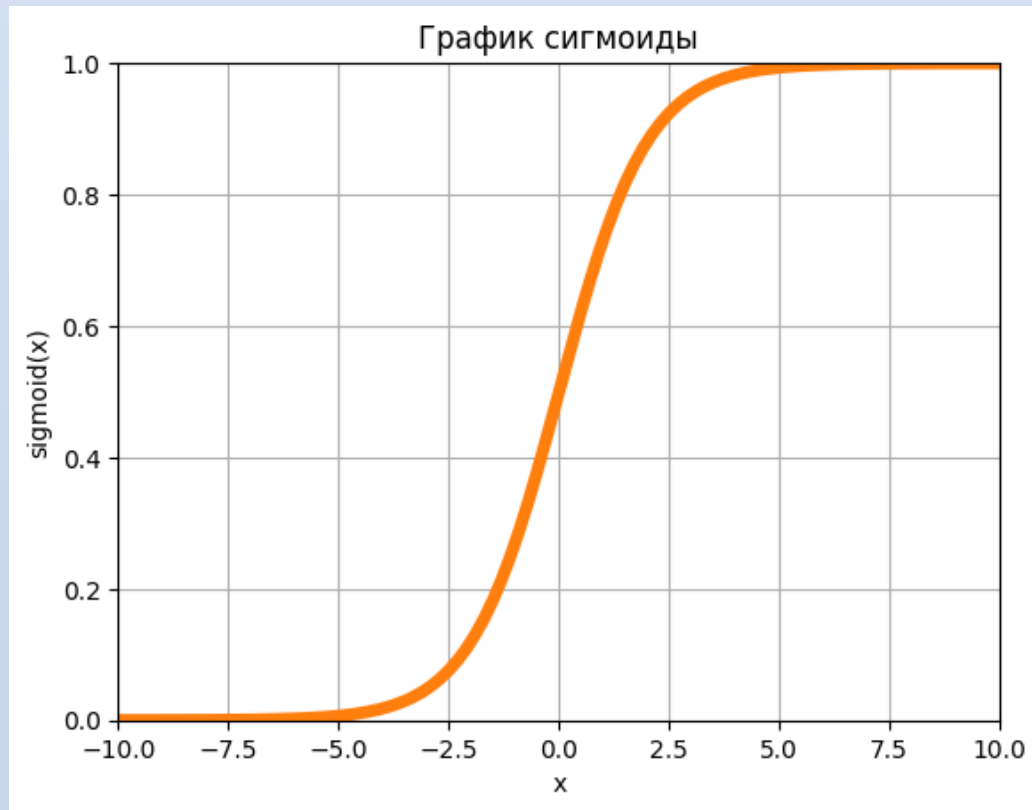
$$F(s) = \begin{cases} a * (e^s - 1), & s < 0 \\ s, & s \geq 0 \end{cases}$$

$$F'(S) = \begin{cases} F(s) + a, & s < 0 \\ 1, & s \geq 0 \end{cases}$$

Функція активації | Activation function

Рідж-функції — функції від лінійної комбінації вхідних змінних.

Логістична (сігмоїда) $F(s) = \frac{1}{1 + e^{-s}}$
 $F'(s) = F(s)(1 - F(s))$



Функція активації | Activation function

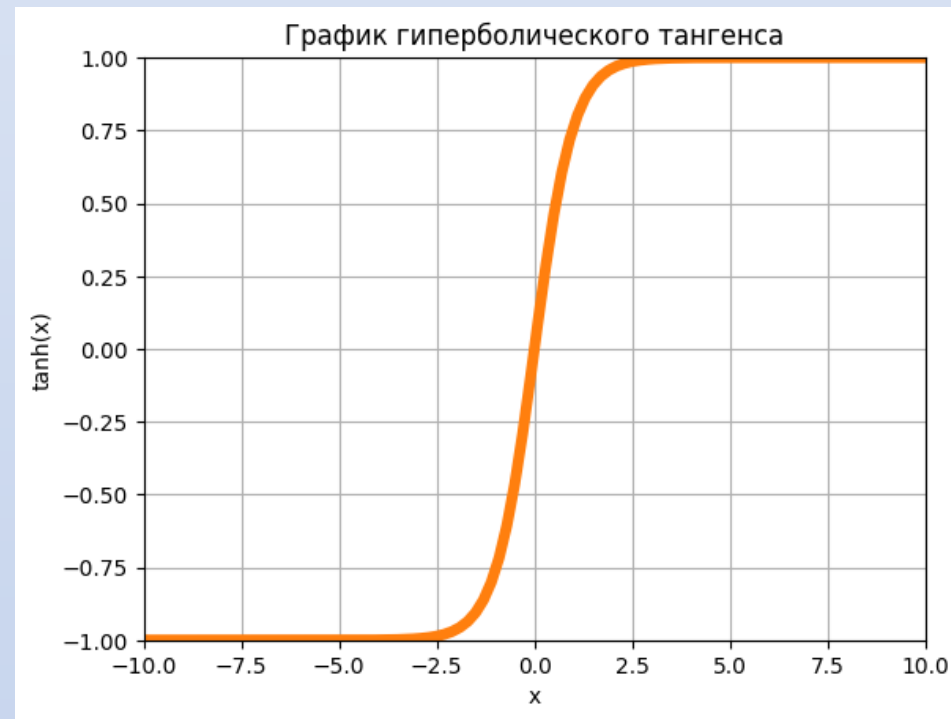
Рідж-функції — функції від лінійної комбінації вхідних змінних.

Гіперболічний тангенс

$$F(s) = \tanh(s)$$

$$= \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$F'(s) = 1 - F(s)^2$$



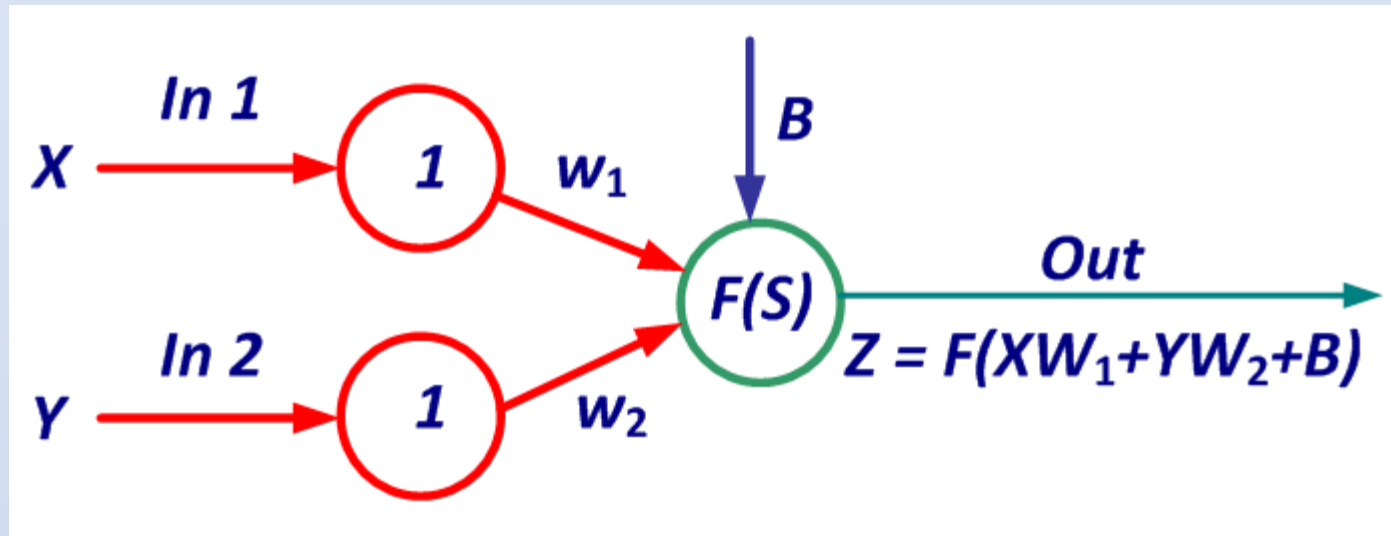
Функція активації | Activation function

Softmax - нормована експоненційна функція – узагальнення логістичної функції, що «стискує» K -вимірний вектор із довільним значеннями компонент до K -вимірного вектора з дійсними значеннями компонент в області $[0, 1]$ що в сумі дають одиницю.

$$F(s) = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}; \text{ for } j = 1, \dots, K$$

Перетворює набори чисел у ймовірності, причому сума останніх дорівнює 1. Функція виводить як результат вектор, що представляє розподіл ймовірностей списку потенційних результатів.

Персептрон як модель логічного елемента



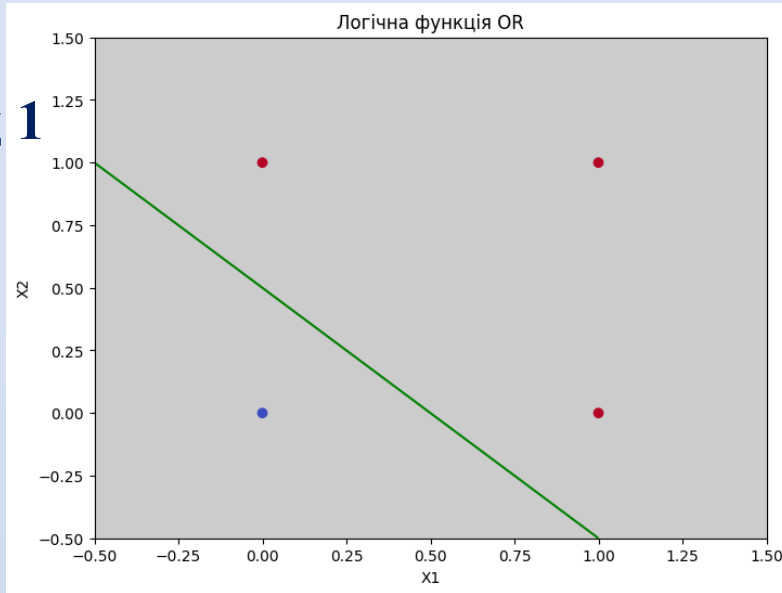
X	Y	Z
0	0	0
0	1	1
1	0	1
1	1	1

X	Y	Z
0	0	0
0	1	0
1	0	0
1	1	1

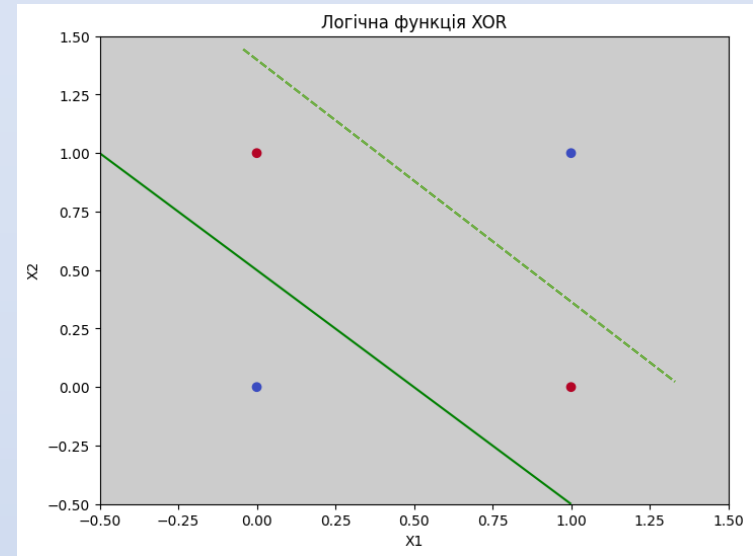
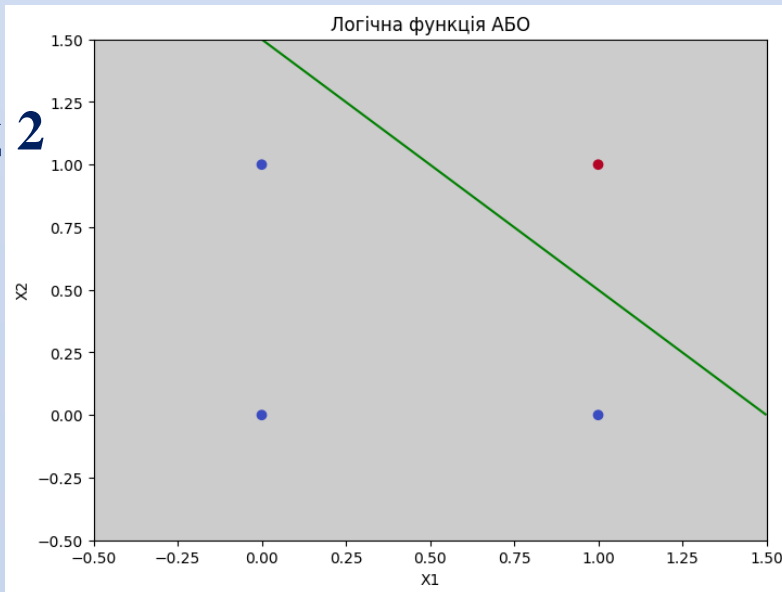
X	Y	Z
0	0	0
0	1	1
1	0	1
1	1	0

Персептрон як модель логічного елемента

OR
приклад 1



AND
приклад 2



XOR
приклад 3
! Класифікація
неможлива

Рекомендована ЛІТЕРАТУРА

Литвин В. В., Пелешак Р. М., Висоцька В. А.
Глибинне навчання : навч. посіб. – Львів : Вид-во
Львівської Політехніки, 2011. – 264 с.

Тимощук П.В., Лобур М. В. Principles of Artificial
Neural Networks and Their Applications :: Принципи
штучних нейронних мереж та їх застосування : навч.
посіб. – Львів : Вид-во Львівської Політехніки, 2011. –
292 с.

Тимощук, П.В. Штучні нейронні мережі : навч.
посіб. – Львів : Вид-во Львівської Політехніки, 2011. –
444 с.

Рекомендована ЛІТЕРАТУРА

Beyeler M. Machine Learning for OpenCV . — Packt Publishing Ltd., 2017 . — 350 p.

Sarkar D., Bali R., Sharma T. Practical Machine Learning with Python . — APress, 2018. — 530p.

Raschka S., Mirjalili V. Python Machine Learning. Machine Learning and Deep Learning with Python, scikit-learn, and TensorFlow 2 .- 3rd Edition, Packt Publishing, 2019 .- 859 p.

The END

Модуль 5. Лекція 01.