

2. Структура і властивості штучного нейрона

Штучний нейрон – це складова частина штучної нейронної мережі (рис. 2.1).

До складу нейрона входять помножувачі (синапси), суматор і нелінійний перетворювач. Синапси здійснюють зв'язок між нейронами і перемножують вхідний сигнал на число, яке характеризує силу зв'язку – вагу синапсу.

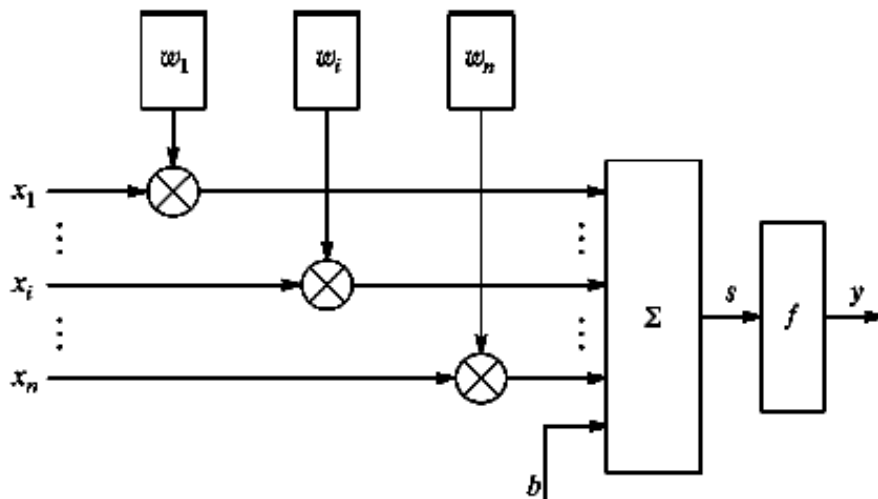


Рис. 2.2. Структура штучного нейрона

Суматор складає сигнали, що поступають по синаптичних зв'язках від інших нейронів, і зовнішніх вхідних сигналів. Нелінійний перетворювач реалізує нелінійну функцію одного аргументу – виходу суматора. Ця функція називається «функція активації» або «передавальна функція» нейрона. Нейрон в цілому реалізує скалярну функцію векторного аргументу. Математична модель нейрона описується співвідношеннями:

$$s = \sum_{i=1}^n \omega_i \cdot x_i + b, \quad (2.1)$$

де: ω_i – вага синапсу ($i = 1, \dots, n$); b – значення зміщення;

x_i – компонента вхідного вектору (вхідний сигнал) ($i = 1, \dots, n$);

s - результат додавання; y – вихідний сигнал нейрона;

n - число входів нейрона;

f - нелінійне перетворення (функція активації або передаточна функція).

У загальному випадку вхідний сигнал, вагові коефіцієнти і значення зсуву можуть приймати дійсні значення. Вихід (y) визначається видом функції активації і може бути як дійсним, так і цілим (табл.2.1). У багатьох практичних задачах вага і зсуви можуть приймати лише деякі фіксовані значення.

Синаптичні зв'язки із позитивними вагами називають збуджуючими, з негативними масами – гальмуючими.

Таким чином, нейрон повністю описується своїми вагами ω_i і передаточною функцією $f(s)$. Одержавши набір чисел (вектор) x_i , нейрон видає деяке число y на виході.

Таблиця 2.1. Перелік функцій активації нейронів

Назва	Формула	Область значень
Порогова	$f(s) = \begin{cases} 0, & s < \theta \\ 1, & s \geq \theta \end{cases}$	(0, 1)
Знакова (сигнатурна)	$f(s) = \begin{cases} 1, & s > 0 \\ -1, & s \leq 0 \end{cases}$	(-1, 1)
Сигмоїдальна (логістична)	$f(s) = \frac{1}{1 + e^{-s}}$	(0, 1)
Напівлінійна	$f(s) = \begin{cases} s, & s > 0 \\ 0, & s \leq 0 \end{cases}$	(0, ∞)
Лінійна	$f(s) = s$	($-\infty, \infty$)
Радіальна базисна (гаусівська)	$f(s) = e^{-s^2}$	(0, 1)
Гіперболічний тангенс	$f(s) = \frac{e^s - e^{-s}}{e^s + e^{-s}}$	(-1, 1)

Описаний обчислювальний елемент (2.1) можна вважати спрощеною математичною моделлю біологічних нейронів – клітин, з яких складається нервова система людини і тварин.

Щоб підкреслити відмінність нейронів біологічних і математичних, їх іноді називають нейроноподібними елементами або формальними нейронами.

На вхідний сигнал (s) нелінійний перетворювач відповідає вихідним сигналом $f(s)$, який є виходом нейрона y . Приклади активаційних функцій представлені у таблиці 2.1 і на рисунку. 2.3.

Однією з найпоширеніших є нелінійна функція з насиченням, так звана логістична функція або сигмоїд (тобто функція S-подібного вигляду):

$$f(s) = \frac{1}{1 + e^{-s}}. \quad (2.2)$$

При зменшенні a сигмоїд стає більш пологим, в межах при $a=0$ вироджуючись в горизонтальну лінію на рівні 0,5, при збільшенні a сигмоїд наближається до вигляду функції одиничного стрибка з порогом θ в точці $s=0$. Із виразу для сигмоїда очевидно, що вихідне значення нейрона лежить у діапазоні $[0, 1]$. Одна з цінних властивостей сигмоїдної функції – простий вираз для її похідної:

$$f'(s) = af(s)(1 - f(s)). \quad (2.3)$$

Слід зазначити, що сигмоїдна функція диференційована на всій осі абсцис. Крім того вона має властивість посилювати слабкі сигнали краще, ніж великі, і запобігає насиченню від великих сигналів, оскільки вони відповідають областям аргументів, де сигмоїд має пологий нахил.

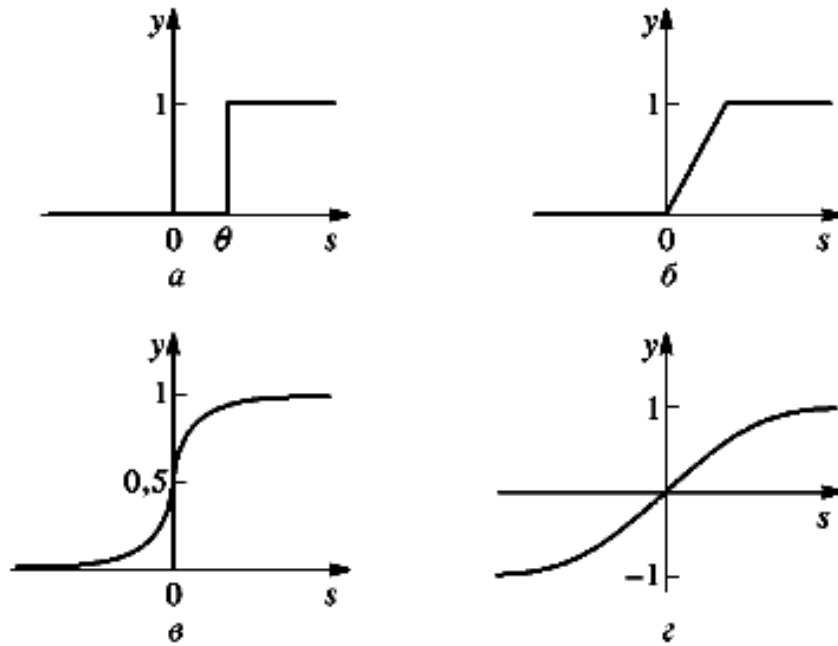


Рис. 2.3. Приклади активаційних функцій: *a* – функція одиничного стрибка; *б* – лінійний поріг (гістерезис); *в* – сигмоїд (гіперболічний тангенс); *г* – сигмоїд (логістична)

Повертаючись до загальних рис, властивих всім НМ, відзначимо принцип паралельної обробки сигналів, який досягається шляхом об'єднання великого числа нейронів у так звані шари і з'єднання певним чином нейронів різних шарів, а також, в деяких конфігураціях, і нейронів одного шару між собою, причому обробка взаємодії всіх нейронів ведеться пошарово.