

Виконав студент групи КМ-01: Романецький Микита Сергійович

## Питання № 16: Функції активації в шарах згортки та в повнозв'язних шарах.

Шари нейронних мереж грають ключову роль у виконанні обчислень та виборі інформації, що передається наступному шару. Функції активації керують тим, як нейрони реагують на вхідні сигнали та передають їх наступним шарам.

### Шари згортки:

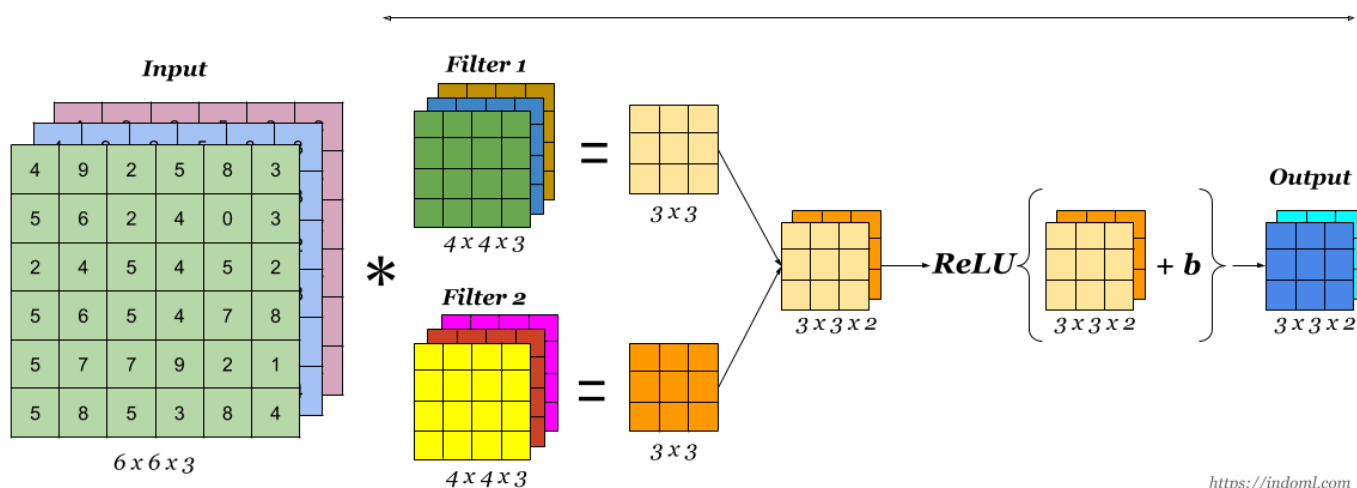
Шари згортки є основними компонентами згорткових нейронних мереж і використовуються для виявлення локальних ознак у вхідних даних, таких як зображення. Кожен нейрон у шарі згортки виконує згортку (convolution) на вхідних даних з використанням фільтра або ядра (kernel). Функція активації в шарі згортки допомагає нейронам визначити, які функції або ознаки має виявити фільтр. Однією з найпоширеніших функцій активації в цьому типі шарів є ReLU (Rectified Linear Unit).

Функція активації ReLU визначається наступним чином:

$$f(x) = \max(0, x),$$

де  $x$  – вхідний сигнал для нейрона. Якщо  $x$  більше або дорівнює нулю, то вихід нейрона буде  $x$ . В іншому випадку, вихід буде нулем. ReLU допомагає вирізняти найважливіші ознаки в даних, зберігаючи значення позитивних сигналів та обнуляючи негативні.

### A Convolution Layer



<https://indoml.com>

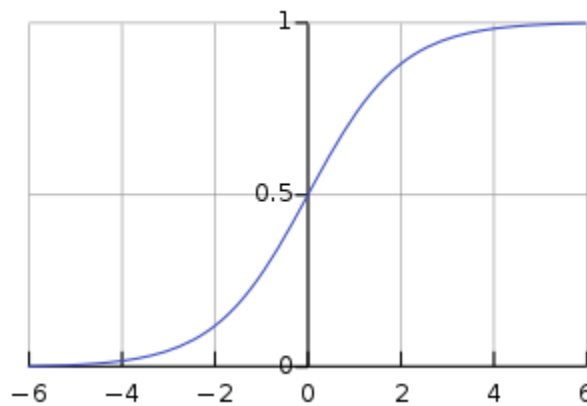
## Повнозв'язні шари:

Повнозв'язні шари є стандартними шарами нейронних мереж, де кожен нейрон підключений до кожного нейрона попереднього шару. Функції активації в повнозв'язних шарах допомагають моделі вивчати складні нелінійні залежності в даних. Однією з таких функцій активації є сигмоїдальна функція.

Сигмоїдальна функція активації визначається наступним чином:

$$f(x) = \frac{1}{1 + e^{-x}}$$

Ця функція перетворює вхідні значення  $x$  в діапазон від 0 до 1, що може інтерпретуватися як ймовірність. Сигмоїдальна функція використовується у завданнях бінарної класифікації, де ми хочемо передбачити ймовірність належності до одного з двох класів.

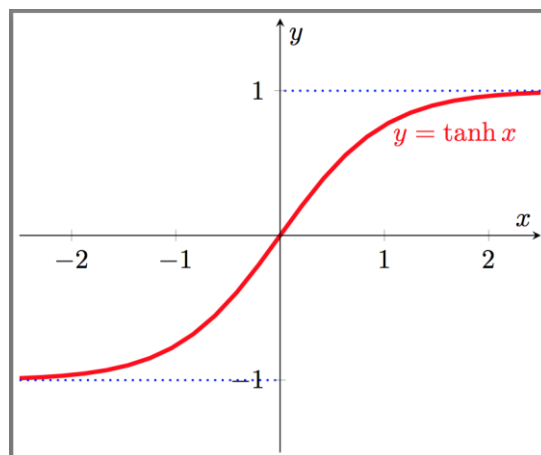


<https://upload.wikimedia.org/wikipedia/commons/thumb/8/88/Logistic-curve.svg/320px-Logistic-curve.svg.png>

Іншою поширеною функцією активації у повнозв'язних шарах є гіперболічний тангенс ( $\tanh$ ):

$$f(x) = \tanh(x)$$

Гіперболічний тангенс також обмежує вихід у діапазоні від -1 до 1 і допомагає моделі вивчати нелінійні залежності в даних.



<https://i.stack.imgur.com/vxfdW.png>

Функції активації визначають, як нейрони реагують на вхідні дані та впливають на вихід моделі. Використання відповідних функцій активації важливо для успішної навчання та використання нейронних мереж в різних завданнях.

## **Питання № 18: Порівняння згорткових нейронних мереж та мереж з прямим розповсюдженням сигналу.**

Згорткові нейронні мережі – (Convolutional Neural Networks, CNN)

Мережі з прямим розп. сигналу – (Feedforward Neural Networks, FNN)

### **1. Архітектура мережі:**

- Згорткова нейронна мережа (CNN):

Згорткові нейронні мережі спеціалізовані на обробці зображень, відео та використовують шари згортки та пулінгу. Ці мережі використовують згорткові фільтри для виділення важливих ознак із зображень та пулінг для зменшення розмірності даних.

- Мережа з прямим розповсюдженням сигналу (FNN):

Мережі FNN, відомі як перцептрони, складаються з вхідного шару, одного або декількох прихованих шарів і вихідного шару. Ці мережі не мають концепції згортки і пулінгу, і кожен нейрон з'єднаний з кожним нейроном у сусідніх шарах.

### **2. Застосування:**

- Згорткова нейронна мережа (CNN):

CNN надзвичайно ефективні для завдань, пов'язаних із зображеннями, такими як розпізнавання об'єктів, класифікація зображень, визначення руху і сегментація об'єктів. Вони зберігають просторову ієрархію ознак у зображеннях, що робить їх ідеальними для роботи з багатовимірними даними.

- Мережа з прямим розповсюдженням сигналу (FNN):

FNN зазвичай використовуються для завдань класифікації тексту, аналізу числових даних та інших завдань, де важливо з'єднати кожен нейрон у кожному шарі. Вони менш ефективні у роботі з багатовимірними даними, такими як зображення.

### **3. Ваги та параметри:**

- Згорткова нейронна мережа (CNN):

У CNN використовуються спеціальні згорткові фільтри, які спільно навчаються під час тренування. Це дозволяє мережі виявляти важливі ознаки вхідних даних.

- Мережа з прямим розповсюдженням сигналу (FNN):

В FNN кожен нейрон з'єднаний з кожним нейроном у сусідньому шарі, і ваги між нейронами розглядаються окремо.

### **4. Метод навчання:**

- CNN та FNN навчаються за допомогою методу зворотного поширення помилки. Метод зворотного поширення помилки є алгоритмом навчання, який використовує похибку між прогнозованим та фактичним значенням для оновлення параметрів мережі.

### **5. Переваги та недоліки:**

- Переваги CNN перед FNN:
  - Вони є більш стійкими до шуму в даних.
  - Вони вимагають меншої кількості даних для навчання, ніж FNN.
- Недоліки CNN:
  - Вони можуть бути складні для розуміння та пояснення.
  - Вони можуть бути менш ефективними для обробки даних, які не мають просторової структури.
- Переваги FNN перед CNN:
  - Вони є більш універсальними, ніж CNN.
  - Вони є більш простими для розуміння та пояснення.
- Недоліки FNN:
  - Вони можуть бути менш ефективними для обробки даних, які мають просторову структуру.
  - Вони вимагають більшої кількості даних для навчання, ніж CNN