

Вступ до Основних Задач Штучного Інтелекту

Штучний інтелект (ШІ) охоплює низку фундаментальних задач, які дозволяють системам "розуміти" світ, приймати рішення та взаємодіяти з ним. Ці задачі є основою для побудови інтелектуальних систем.



Класифікація

Віднесення об'єкта до класу.



Кластеризація

Пошук прихованих структур у даних.



Апроксимація

Відновлення невідомої функції.



Прогнозування Передбачення майбутніх значень.

Класифікація Образів: Розпізнавання та Рішення

Класифікація образів — це ключова задача ШІ, що полягає у віднесенні вхідного об'єкта (наприклад, зображення чи мовного сигналу), представленого у вигляді **вектора ознак**, до одного або кількох заздалегідь визначених класів.

BILL OF SALE

\$ 849.00 Redlands, Cal., June 14th. 1937

For and in consideration of the sum of

Eight hundred forty nine & no/----- Dollars

Value Received we hereby sell and transfer unto B.F. Martin

of Redlands, Calif One Chevrolet Automobile

Model 6AA-24915 Style, Sedan Motor No. 3190173

Old License No. _____

Guaranteed free and clear of liens and encumbrances of whatever kind or nature.

Signed [Signature]
 SERVICE MOTOR CO. OF REDLANDS

P. O. Address Redlands, Calif. [Signature]

Як це працює?

Система «розпізнає» об'єкти та приймає рішення на основі попереднього досвіду, використовуючи властивості об'єкта, закодовані у векторі ознак.



Це ключове завдання у побудові систем штучного інтелекту.

Сфери Застосування Класифікації

Завдання класифікації образів має широке практичне застосування у багатьох критично важливих галузях.



Розпізнавання Тексту

Автоматичне розпізнавання рукописних або друкованих символів (OCR).



Розпізнавання Мовлення

Перетворення звукових сигналів у текстову форму для голосових помічників.



Медична Діагностика

Класифікація сигналів ЕКГ, МРТ або інших біомедичних даних.



Біометрія

Робота сканерів відбитків пальців, обличчя чи райдужної оболонки ока для ідентифікації.



Ключові Застосування Кластеризації

Кластеризація дозволяє системам знаходити внутрішні структури у даних і формувати «природні» групи без участі людини.

Стиснення Даних

Використовується для зменшення обсягу інформації, наприклад, у зображеннях чи текстах.

Дослідження Структури

Виявлення прихованих закономірностей і взаємозв'язків без попередніх знань про їх природу.

Попередня Обробка

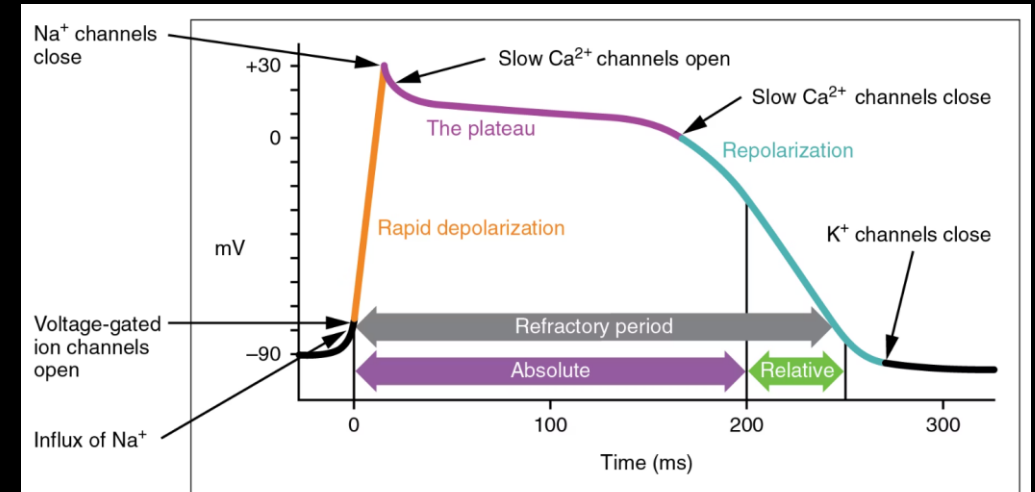
Часто є важливим етапом перед подальшою класифікацією або візуалізацією складних наборів даних.

Апроксимація Функцій: Відновлення Залежностей

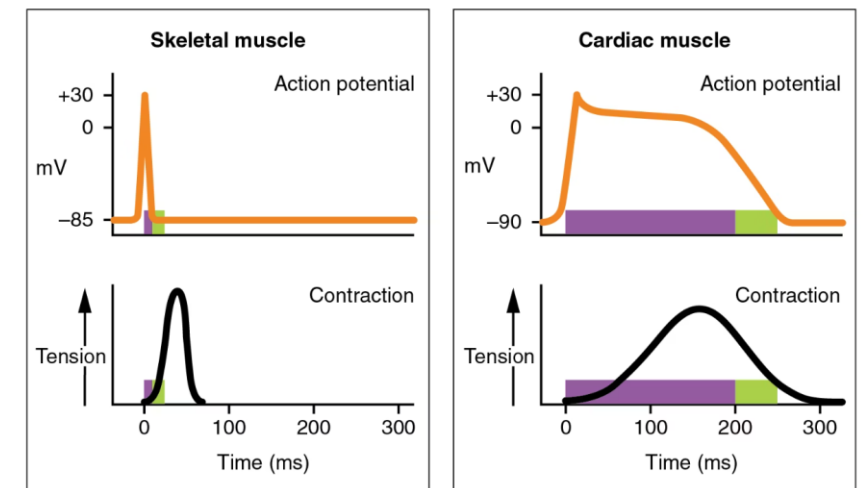
Апроксимація функцій — це процес відновлення невідомої функції на основі навчальної вибірки даних виду «вхід → вихід».

Особливість полягає в тому, що вхідні дані часто **спотворені шумом**. Алгоритм повинен відокремити справжню закономірність від випадкових коливань, щоб відтворити «чисту» залежність.

- Відтворення «чистої» залежності.
- Прогнозування значень для нових параметрів.



(a)



(b)

Практичне Застосування Апроксимації

Апроксимація є ключовою для науки й практики, дозволяючи моделювати складні процеси та очищати дані від небажаних впливів.



Фільтрація Сигналів

Очищення від шуму при прийомі аудіо-, відео- чи радіосигналів.



Моделювання Процесів

Побудова точних математичних моделей у фізиці, техніці та біології.



Інженерні Задачі

Відновлення характеристик систем на основі експериментальних даних.

Прогнозування (Екстраполяція): Погляд у Майбутнє

Прогнозування, або екстраполяція, полягає у передбаченні майбутніх значень процесу на основі відомих даних із минулого. Це ключовий інструмент для роботи з часовими рядами.

Прогнозування має велике практичне значення, оскільки дозволяє формувати очікування щодо майбутнього та підтримувати прийняття рішень.

- Оцінка попиту на продукцію та прогноз цін на біржах.
- Моделювання динаміки природних процесів у науці.
- Прогнозування стану систем для запобігання відмовам обладнання.
- Класичний приклад — **прогноз погоди**.



Оптимізація: Пошук Найкращого Рішення

Оптимізація — це процес пошуку найкращого рішення серед можливих, коли потрібно мінімізувати або максимізувати певну **цільову функцію** за наявності обмежень.



Ефективність

Забезпечення найвищої ефективності системи.



Обмеження

Робота в рамках заданих лімітів і умов.



Цільова Функція

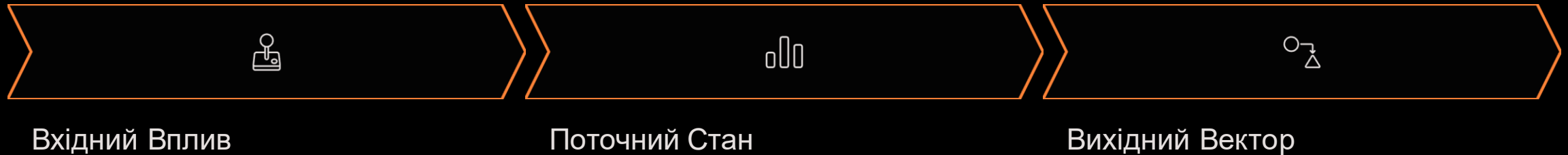
Мінімізація витрат або максимізація прибутку.

Оптимізація є універсальним інструментом у математиці, техніці, економіці та лежить в основі сучасних методів машинного навчання.

Управління та Асоціативна Пам'ять

Управління Динамічними Системами

Управління досліджує динамічні системи та методи впливу на них для досягнення бажаного результату. Класична задача — знайти керуючий вплив, який забезпечить перехід системи в потрібний стан.



Асоціативна Пам'ять

Це тип машинної пам'яті, який дозволяє здійснювати пошук не за адресою, а за **змістом** (пам'ять, що адресується за вмістом). У програмуванні це схоже на **асоціативний масив** (словник), де доступ до елемента здійснюється за ключем, а не за позицією.

Порівняння з Людським Нейроном

Штучні нейронні мережі виникли як спроба наслідувати принципи роботи людського мозку. Дослідники запозичили терміни та базові ідеї з нейробіології, проте ця подібність доволі умовна.

Наші знання про механізми мозкової діяльності ще занадто обмежені, щоб точно відтворити їх у штучних системах. Тому при побудові моделей розробники часто виходять за межі біологічних аналогій, шукаючи власні математичні структури.



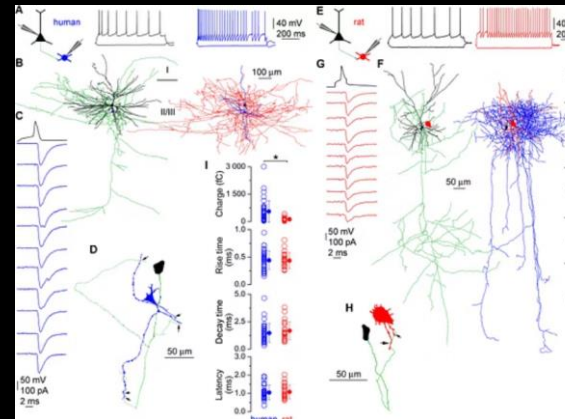
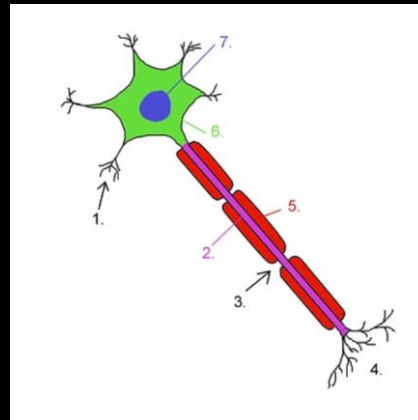
Нейрон: Основна Одиниця Мозку

Визначення

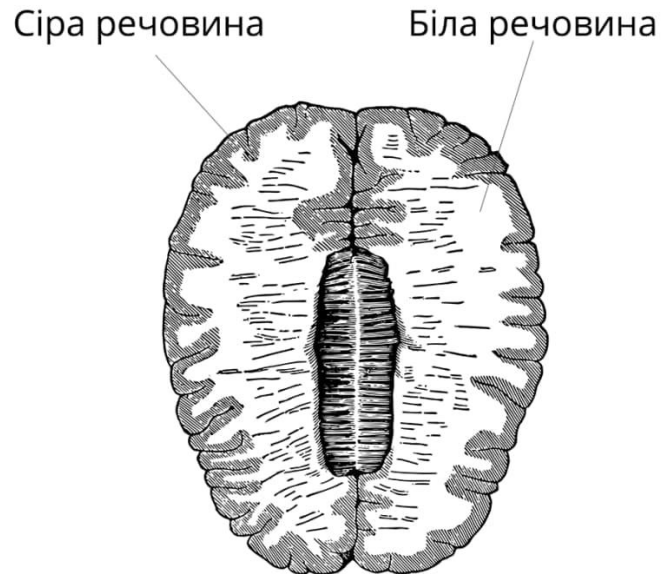
Нейрон — це основна структурна та функціональна одиниця нервової системи, клітина, яка відповідає за передачу й обробку інформації в мозку.

Функція

Відповідає за передачу й обробку інформації, формуючи складну мережу, що лежить в основі мислення, пам'яті та навчання.



Анатомія Мозку: Сіра та Біла Речовина



Сіра Речовина

Тіла нейронів, де відбувається основна обробка сигналів та інформації.

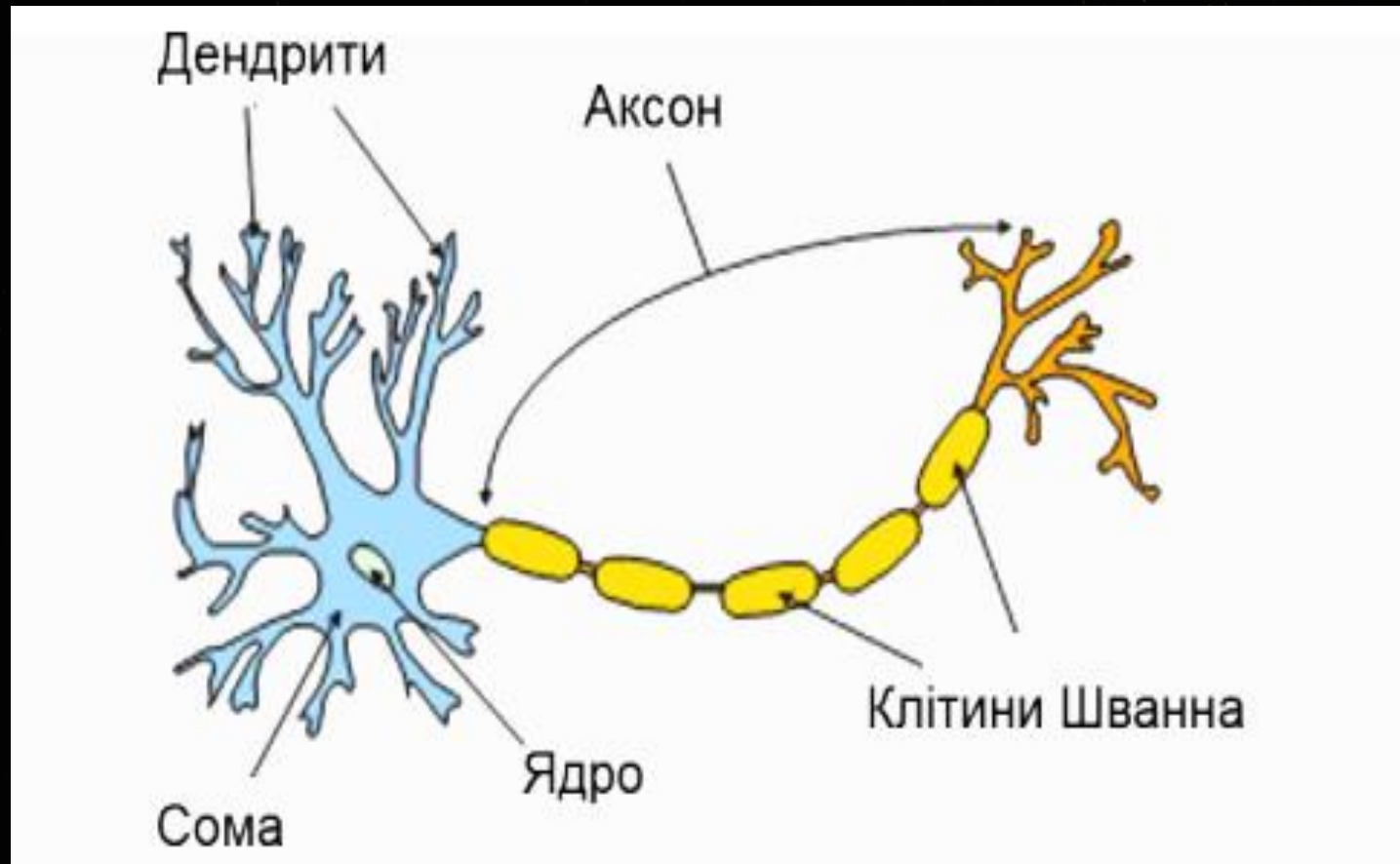


Біла Речовина

Нервові волокна (аксони), які з'єднують нейрони та забезпечують комунікацію між ділянками мозку.

Три Основні Частини Нейрона

Кожен нейрон має спеціалізовані компоненти для отримання, обробки та передачі сигналів.



Тіло Клітини (Сомат)

Центр життєдіяльності нейрона, де відбувається інтеграція сигналів.

Дендрити

Розгалужені відростки, що отримують вхідні сигнали від інших нейронів.

Аксон

Довгий відросток, який передає оброблений сигнал до інших клітин.

Передача Інформації в Нервовій Мережі

Передача сигналів — це динамічний процес, що забезпечує функціонування всієї нервової системи.



Отримання

Інформація надходить через **дендрити** — численні розгалужені відростки, що приймають сигнали.



Обробка

Інформація обробляється в **тілі клітини** (сомі).



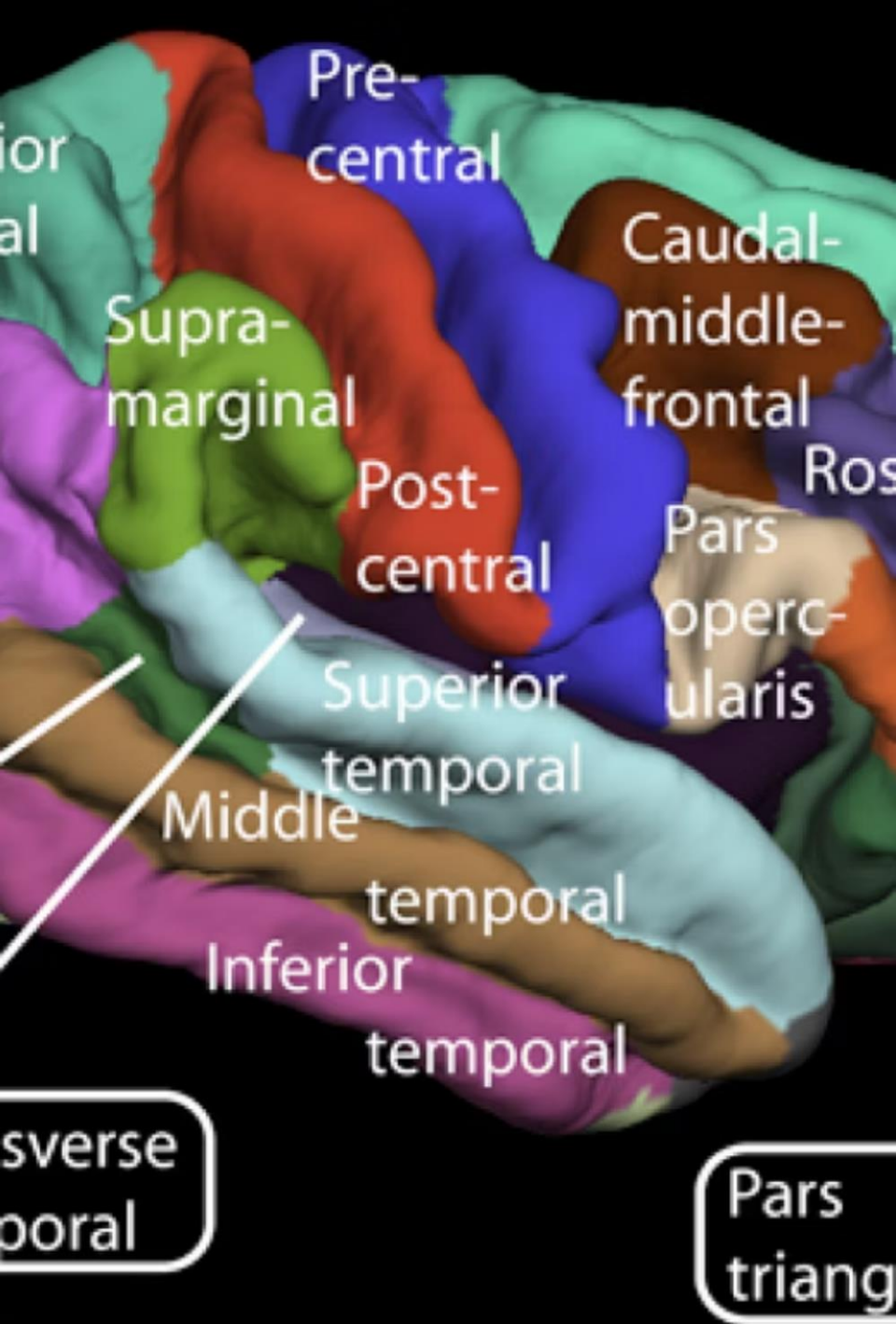
Передача

Сигнал передається по **аксону**.



Зв'язок

Аксон розгалужується на **синапси**, які забезпечують зв'язок з іншими нейронами.



Масштаби Людського Мозку

Людський мозок — це надзвичайно складна і потужна мережа, що забезпечує складність мислення.

2200 cm^2

Площа Кори

Приблизна площа кори головного мозку.

2-3 мм

Товщина Кори

Незважаючи на тонкість, кора містить мільярди нейронів.

10^{11}

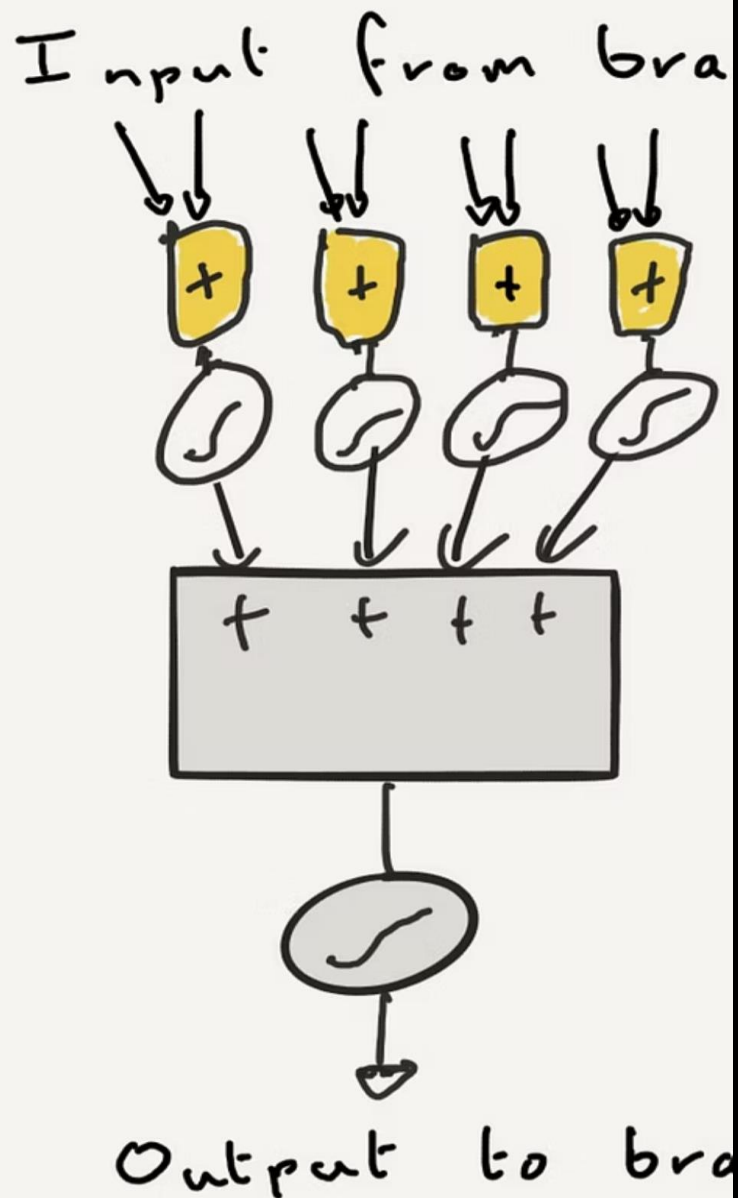
Кількість Нейронів

Приблизна кількість нейронів у мозку.

10^{15}

Міжнейронні Зв'язки

Загальна кількість синаптичних зв'язків, що формуються в мозку.



Принципи Роботи Біологічного Мозку

На відміну від традиційних комп'ютерів, біологічний мозок функціонує за унікальними, нелінійними принципами.



Паралельна Діяльність

Обробка інформації відбувається одночасно у багатьох ділянках, а не послідовно.



Нелінійність

Відносини між вхідними та вихідними даними є складними та непередбачуваними.

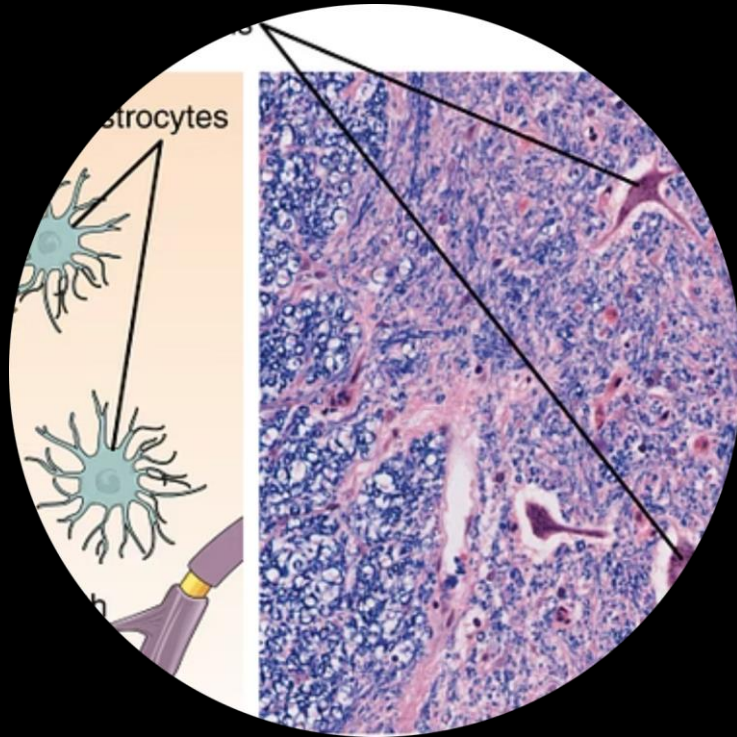


Нестабільність

Мозок постійно адаптується та змінюється (пластичність), що відрізняється від жорсткої логіки комп'ютерів.

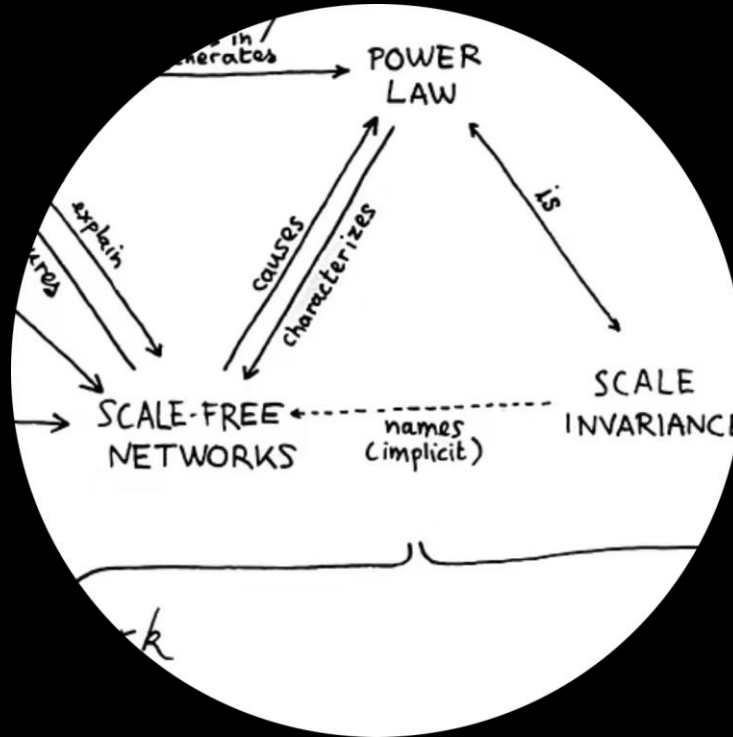
Фактори, що Впливають на Складність Мозку

Складність мислення визначається не лише кількістю нейронів, але й якісними характеристиками їхньої організації.



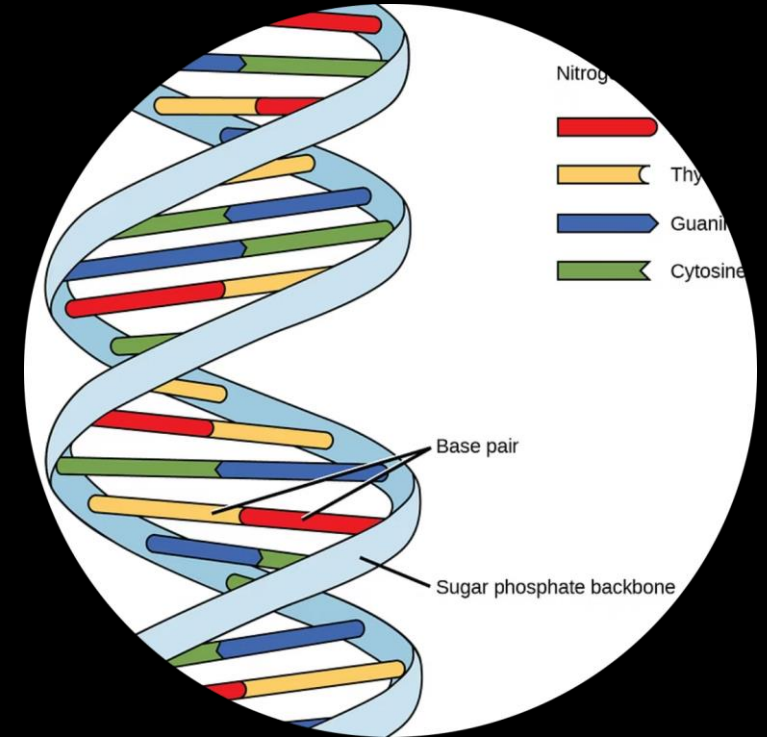
Різноманіття Типів

Існування багатьох типів нейронів зі специфічними функціями.



Структура З'єднань

Складна архітектура міжнейронних зв'язків.



Спадковість і Навчання

Взаємодія генетичних факторів та досвіду, що формує мозок.

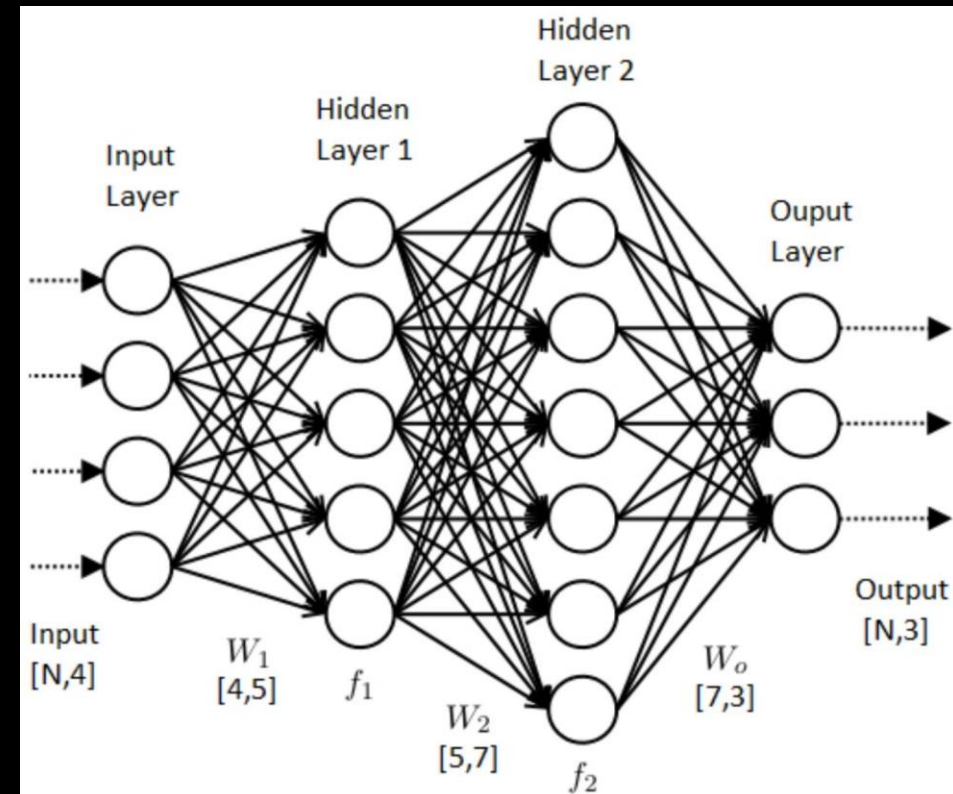
Штучні Нейронні Мережі: Моделювання Базових Принципів

Штучні нейронні мережі (ШНМ) не відтворюють усієї складності біологічного мозку, але моделюють його ключові функції.

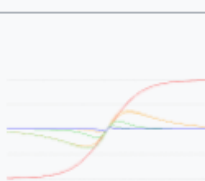
ШНМ моделюють лише базові принципи взаємодії: «вхід – обробка – вихід».

Цього спрощення достатньо, щоб:

- Надихати дослідників на нові відкриття.
- Розв'язувати складні практичні задачі.
- Створювати ефективні математичні структури, що виходять за межі біологічних аналогій.



Функції активації нейронів

Identity		x	1	$(-\infty, \infty)$	C^∞
Binary step		$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$	0	$\{0, 1\}$	C^{-1}
Logistic, sigmoid, or soft step		$\sigma(x) \doteq \frac{1}{1 + e^{-x}}$	$g(x)(1 - g(x))$	$(0, 1)$	C^∞
Hyperbolic tangent (tanh)		$\tanh(x) \doteq \frac{e^x - e^{-x}}{e^x + e^{-x}}$	$1 - g(x)^2$	$(-1, 1)$	C^∞
Soboleva modified hyperbolic tangent (smht)		$\text{smht}(x) \doteq \frac{e^{ax} - e^{-bx}}{e^{cx} + e^{-dx}}$		$(-1, 1)$	C^∞
Softsign		$\frac{x}{1 + x }$	$\frac{1}{(1 + x)^2}$	$(-1, 1)$	C^1
Rectified linear unit (ReLU) ^[13]		$(x)^+ \doteq \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$ $= \max(0, x) = x \mathbf{1}_{x>0}$	$\begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases}$	$[0, \infty)$	C^0

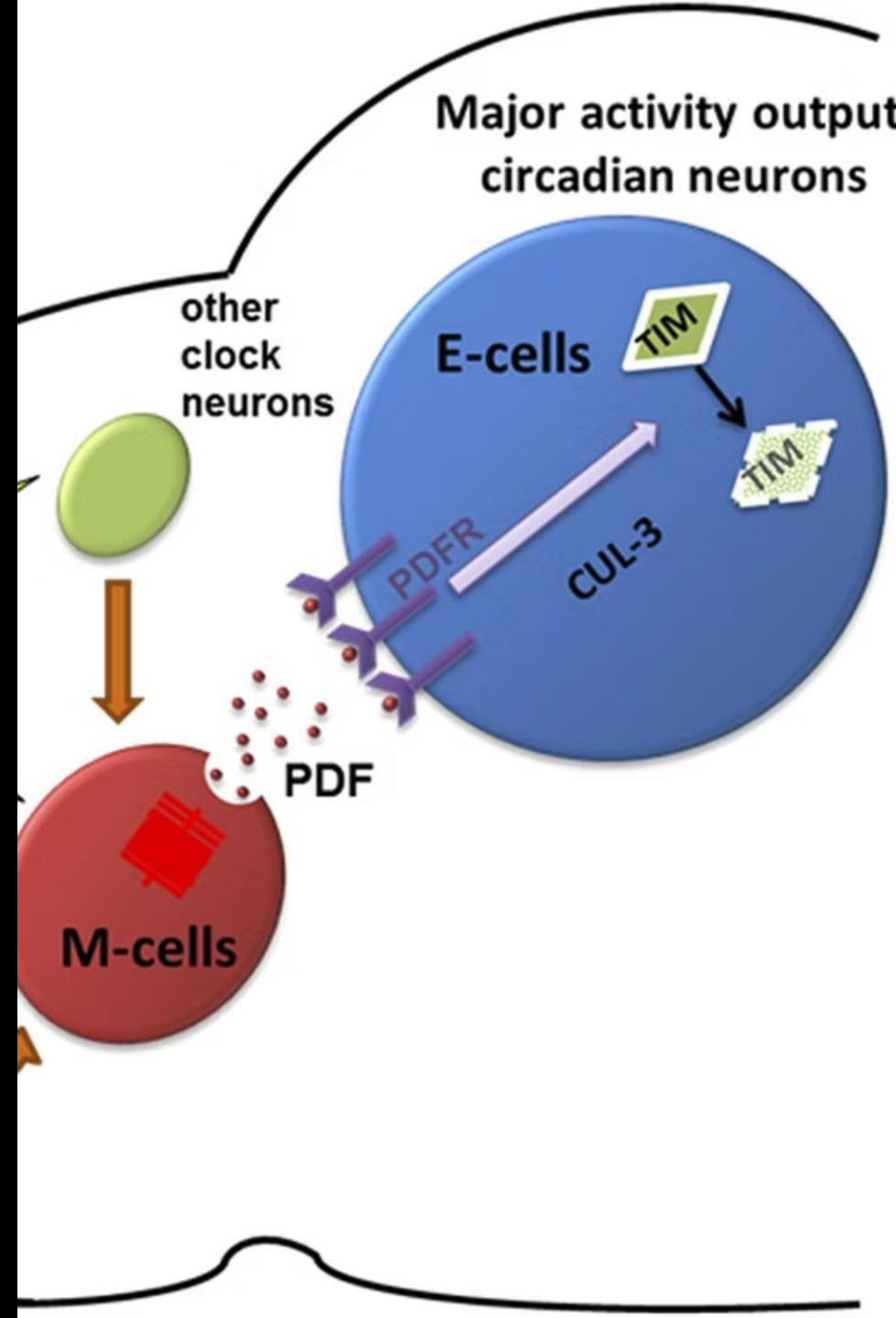
Основи Нейронних Мереж:

Активація та Ваги

Активаційні функції вводять нелінійність у нейронні мережі, допомагаючи їм моделювати складні взаємозв'язки. Вони вирішують, чи активувати нейрон на основі його вхідних даних.

Штучний нейрон можна уявити як порогову систему: він підсумовує вхідні сигнали (з урахуванням ваг). Якщо сума перевищує поріг, нейрон «спрацьовує» й генерує вихідний сигнал; інакше — залишається «мовчазним».

Ваги — це зв'язки між нейронами, що визначають силу сигналу. Вони коригуються під час фази навчання, дозволяючи моделі навчатися та покращувати свою продуктивність.





Ключові Процеси: Пряме та Зворотне Поширення



Пряме Поширення (Forward Propagation)

Основа роботи мереж. Включає передачу вхідних даних через шари нейронів, де кожен нейрон застосовує функцію активації до своєї зваженої суми вхідних даних. Дозволяє мережі робити прогнози та генерувати вихідні дані.



Зворотне Поширення Помилки (Backpropagation)

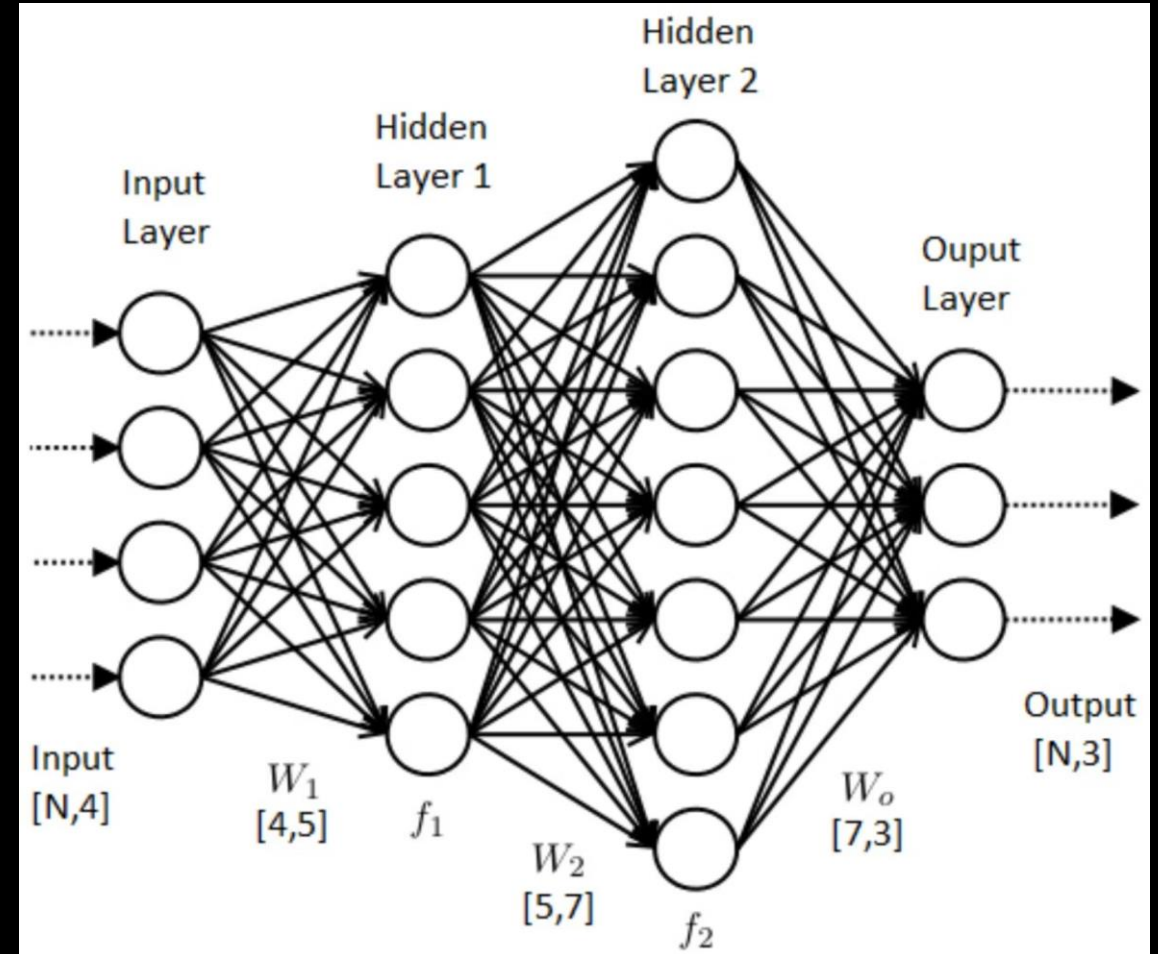
Процес коригування ваг мережі у відповідь на зворотний зв'язок під час навчання. Помилка поширюється з вихідного шару назад, дозволяючи кожному нейрону коригувати ваги. Це ітеративний градієнтний алгоритм для мінімізації середньоквадратичного відхилення.

Архітектура Нейронних Мереж

Штучні нейронні мережі формуються з базового елемента — штучного нейрона. Їм притаманна висока щільність зв'язків, що поєднують окремі вузли. Об'єднані у мережу штучні нейрони створюють систему обробки даних, здатну ефективно пристосовуватися до змін зовнішнього середовища.

Визначення Термінів

- **Структура:** Схема з'єднань між нейронами.
- **Архітектура:** Поєднання структури та типів нейронів (організація шарів, активації).
- **Парадигма:** Спосіб навчання й використання моделі.



Висока ефективність нейромережі залежить від вдалого добору кількості нейронів і конфігурації зв'язків між ними.

Типи Зв'язків та Багатошарова Архітектура

Слабозв'язані Мережі

Нейрон з'єднаний лише з близькими сусідами.

Повнозв'язані Мережі

Кожен нейрон має зв'язки з усіма іншими (або з усіма нейронами суміжного шару).

Рекурентні Мережі

Вважаються динамічними через зворотні зв'язки (петлі), які модифікують входи нейронів у часі, змінюючи стани мережі.

Багатошарова Архітектура

Найпоширеніша архітектура: нейрони згруповані в шари. **Вхідний шар** (рецептори) отримує зовнішній вектор, **останній шар** (ефектори) формує вихідні значення. Між ними — **приховані шари**, що виконують проміжне перетворення сигналів.



Контрольоване Навчання: Мета та Процес

Більшість нейромереж працюють у режимі контрольованого навчання: поточний вихід постійно зіставляють із цільовим. На старті ваги ініціалізують випадково, а далі на кожній ітерації коригують, зменшуючи похибку.

01

Ініціалізація

Ваги ініціалізуються випадково.

02

Коригування Ваг

На кожній ітерації ваги оновлюються для мінімізації помилки.

03

Фіксація Ваг

Навчання завершують, коли досягнуто заданий рівень ефективності та точності, після чого ваги фіксують.

Мета — мінімізувати помилки всіх обчислювальних елементів шляхом безперервного оновлення синаптичних ваг до досягнення прийнятної точності.

Особливості Навчальної Вибірки

Навчальна вибірка має бути достатньо великою й різноманітною, щоб охопити потрібні ознаки та зв'язки. Для ефективного навчання мережу слід учити на всій сукупності даних, а не подавати приклади по одному.



Репрезентативність

Вибірка має відображати реальний стан предметної області.



Несуперечність

Відсутність конфліктних прикладів, інакше якість навчання падає.



Достатній Обсяг

Кількість записів має на кілька порядків перевищувати число параметрів у мережі, щоб уникнути «заучування» вибірки.

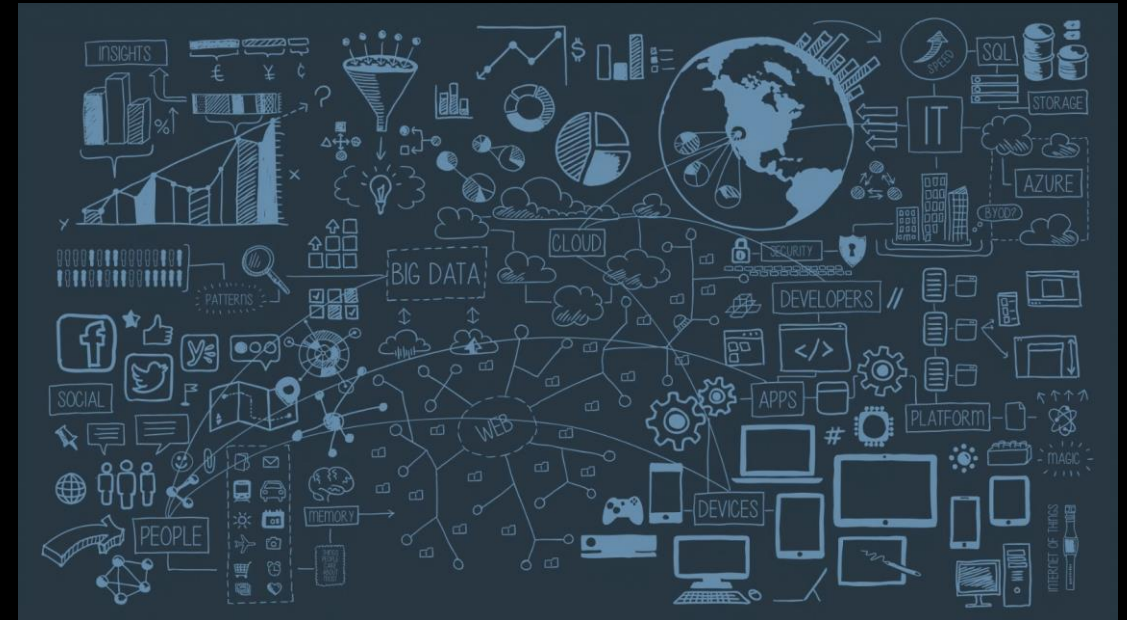
Критично важливими є подання й кодування входів/виходів, оскільки мережі працюють лише з числовими значеннями.

Попередня Обробка Даних та Кодування

Навчання на «сирих» даних зазвичай неефективне, тому застосовують попередню обробку для приведення даних до формату, сумісного з входами мережі.

Методи Попередньої Обробки

- **Нормалізація:** Зведення різновимірних ознак до спільного масштабу, щоб великі діапазони не домінували.
- **Квантування:** Дискретизація безперервних величин (наприклад, частот).
- **Фільтрація:** Зменшення шуму у даних.

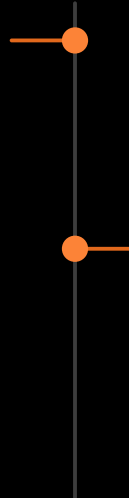


Для подання виходів часто використовують **багатовихідне кодування (one-hot)**: кожен клас має власний вихідний канал. Це збільшує «відстань» між класами та зменшує ризик їх плутати.

Тестування та Узагальнення

Навіть якщо мережа добре працює на тренувальних даних, важливо перевірити її здатність узагальнювати — якість на нових, раніше не використаних прикладах.

Тренувальна Частина
Використовується для оновлення ваг мережі.



Тестова (Перевірочна) Частина

Використовується лише для оцінювання похибки та перевірки узагальнення.

📌 Якщо помилка на тренуванні падає, а на тесті зростає, мережа **перенавчається (overfitting)** — навчання зазвичай зупиняють.

Тестування потрібне для формування узагальнених уявлень (образів), а не лише для «запам'ятовування» конкретних зразків.

Неконтрольоване Навчання та Проблема Адекватності

Неконтрольоване Навчання

Передбачає здатність комп'ютерів навчатися самостійно, шукаючи закономірності у вхідних даних без зовнішніх підказок. Приклад: Самоорганізовані карти (SOM).

Основа — конкуренція між нейронами: на заданий сигнал найсильніше реагує одна група, яка «закріплюється» під цей стимул (winner-takes-all).

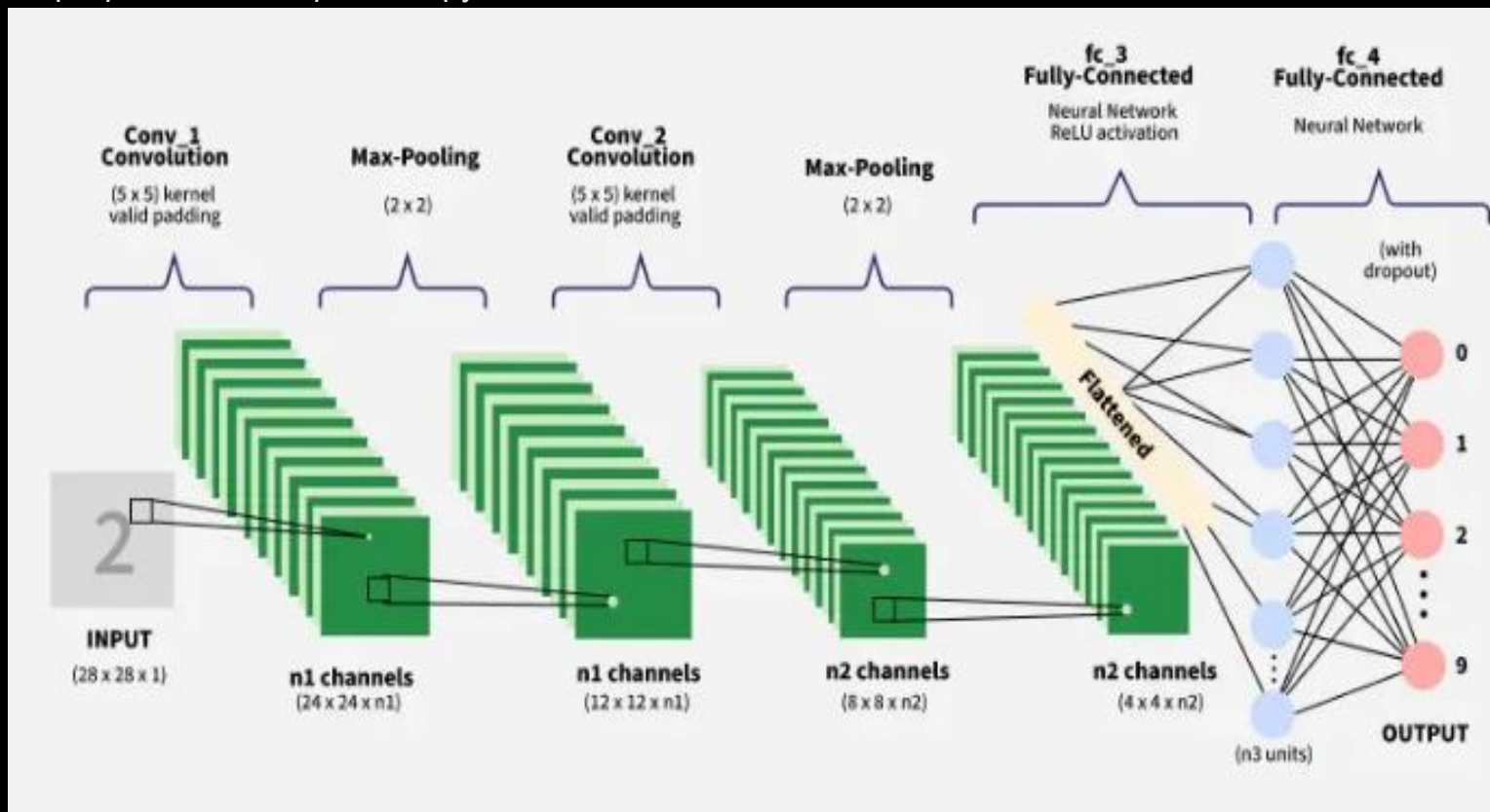
Перевірка Адекватності

Мережа може засвоїти не цільову ознаку, а найпростіший «сурогатний» сигнал. Наприклад, розпізнавати фон замість об'єкта, якщо всі навчальні приклади мали однаковий фон.



Згорткові Нейронні Мережі (CNN)

ЗНМ — це моделі глибокого навчання, призначені для обробки даних із сітчастою топологією, таких як зображення. Вони є основою для більшості сучасних програм комп'ютерного зору.



CNN працює, попередньо обробляючи зображення, вилучаючи ознаки через згортку, зменшуючи розмірність через об'єднання, і, нарешті, використовуючи повністю зв'язані шари для отримання кінцевого прогнозу (наприклад, мітки класифікації).