

# Лекція 3. Штучні нейронні мережі

## Питання лекції

### Вступ

1. Аналогія з мозком людини
2. Штучний нейрон
3. Архітектура з'єднань штучних нейронів
4. Навчання штучної нейронної мережі
5. Основні етапи розв'язання задач за допомогою нейромереж

### ВСТУП

В останні роки разом з нечіткими системами великий інтерес викликає проблематика нейронних мереж і генетичних алгоритмів. Ці напрямки відносяться до наукової області, обумовленої в англійській літературі терміном Computational Intelligence. На рис. 4.1. видно, що задачі нейронних мереж, генетичних алгоритмів і нечітких систем можуть розглядатися поза зв'язком між собою, однак їхня взаємозалежність виявляється надзвичайно важливою. Зокрема, генетичні алгоритми можна застосовувати для підбору ваг і топології нейронної мережі, а також для формування бази правил і функцій приналежності нечіткої системи



Рис. 4.1. Взаємозв'язки між нейронними мережами, генетичними алгоритмами і нечіткими системами

У свою чергу, нейронні мережі дозволяють вибирати відповідні параметри для самих генетичних алгоритмів (параметри схрещування і мутації); саму філософію нейронних мереж можна закласти у фундамент нечітких систем, що у результаті знаходять здатність до навчання.

Крім того, методи теорії нечітких множин дозволяють підбирати як згадані вище параметри генетичних алгоритмів, так і коефіцієнти, що визначають швидкість навчання нейронних мереж. Одним з популярних напрямків Artificial Intelligence є теорія нейронних мереж (neuron nets).

Людей завжди цікавило їхнє власне мислення. Це самопитання, думання мозку самого про себе є, можливо, відмінною рисою людини. Нейробіологи і нейроанатоми досягли в цій області значного прогресу. Ретельно вивчаючи структуру і функції нервової системи людини, вони багато чого зрозуміли в «електропроводці» мозку, але мало довідалися про його функціонування. У процесі нагромадження ними знань з'ясувалося, що мозок має приголомшуючу складність. Сотні мільярдів нейронів, кожний з яких з'єднаний із сотнями або тисячами інших, утворюють систему, що далеко перевершує наші самі сміливі мрії про суперкомп'ютери. На сьогоднішній день існують дві взаємно збагачуючі одна одну мети нейронного моделювання: перша – зрозуміти функціонування нервової системи людини на рівні фізіології і психології і друга – створити обчислювальні системи (штучні нейронні мережі), що виконують функції, подібні до функцій мозку.

Нейроні мережі можна розглядати як сучасні обчислювальні системи, що перетворюють інформацію з образу процесів, що відбуваються в мозку людини. Оброблювана інформація має числовий характер, що дозволяє використовувати нейронну мережу, наприклад, як модель об'єкта з зовсім невідомими характеристиками. Інші типи додатки нейронних мереж охоплюють задачі розпізнавання, класифікації, аналізу і стиску образів.

Штучні нейронні мережі є моделями нейронної структури мозку, який здатен сприймати, обробляти, зберігати та продукувати інформацію. Особливістю мозку також є навчання та самонавчання на власному досвіді. Адаптивні системи на основі штучних нейронних мереж дозволяють з успіхом вирішувати проблеми розпізнавання образів, виконання прогнозів, оптимізації, асоціативної пам'яті і керування.

У порівнянні з традиційними комп'ютерами з архітектурою фон Неймана, мозок людини має багато інших якостей:

Низьке енергоспоживання.

Толерантність до помилок.

Адаптивність.

Здатність до навчання й узагальнення.

Розподілене представлення інформації і паралельні обчислення.

Механізм природного мислення базується на збереженні інформації у вигляді образів. Штучні нейронні мережі дозволяють створення паралельних мереж, їх навчання та вирішення інтелектуальних завдань, не використовуючи традиційного програмування. В лексиконі розробників та користувачів нейромереж присутні слова "поводити себе", "реагувати", "самоорганізовувати", "навчати", "узагальнювати" та "забувати".

## *1. Аналогія з мозком людини*

Розвиток штучних нейронних мереж надихається біологією. Тобто, розглядаючи мережні конфігурації й алгоритми, дослідники роблять це, використовуючи терміни, які характерні для опису організації мозкової діяльності. Але на цьому аналогія, мабуть, закінчується. Наші знання про роботу мозку настільки обмежені, що мало б знайшлося орієнтирів для тих, хто став би йому наслідувати. Тому розроблювачам мереж приходиться виходити за межі сучасних біологічних знань у пошуках структур, здатних виконувати корисні функції.

Базовим елементом мозку людини є нейрони - специфічні клітини, що здатні запам'ятовувати і застосовувати попередній досвід до наступних дій, що докорінно відрізняє їх від решта клітин тіла.

Кора головного мозку людини є протяжною, утвореною нейронами поверхнею товщиною від 2 до 3 мм із площею близько 2200 см<sup>2</sup>, що вдвічі перевищує площу поверхні стандартної клавіатури. Кора головного мозку містить близько 10<sup>11</sup> нейронів, що приблизно дорівнює числу зірок Чумацького шляху. Кожен нейрон зв'язаний з 10<sup>3</sup> - 10<sup>4</sup> іншими нейронами. У цілому мозок людини містить приблизно від 10<sup>14</sup> до 10<sup>15</sup> взаємозв'язків.

Потужність людського розуму залежить від числа базових компонентів, різноманіття з'єднань між ними, а також від генетичного програмування й навчання.

Окремий нейрон є складним, має свої складові, підсистеми та механізми керування і передає інформацію через велику кількість електрохімічних зв'язків. Налічують біля сотні різних класів нейронів. Разом нейрони та з'єднання між ними формують недовійковий, нестійкий та несинхронний процес, що відрізняється від процесу обчислень традиційних комп'ютерів. Штучні нейромережі моделюють лише найголовніші елементи природного мозку, але надихає науковців та розробників до пошуку нових шляхів розв'язування проблеми.

## *Біологічний нейрон*

Нейрон (нервова клітка) складається з тіла клітини - соми (soma), і двох типів зовнішніх деревоподібних відгалужень: аксона (axon) і дендритів (dendrites). Тіло клітини вміщує ядро (nucleus), що містить інформацію про властивості нейрона, і плазму, яка продукує необхідні для нейрона матеріали. Нейрон отримує сигнали (імпульси) від інших нейронів через дендрити (приймачі) і передає сигнали, що згенеровані тілом клітки, вздовж аксона (передавач), який наприкінці розгалужується на волокна (strands). На закінченнях волокон знаходяться синапси (synapses).

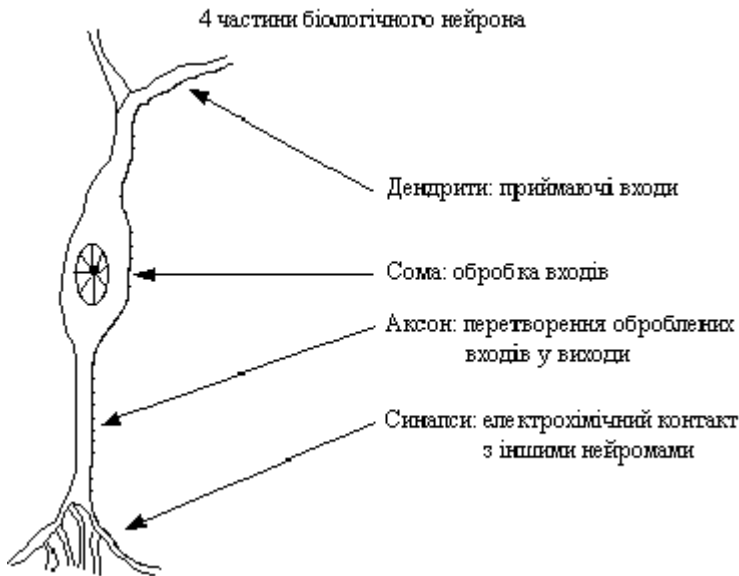


Рис. 1. Схема біологічного нейрона

Синапс є функціональним вузлом між двома нейронами (волокно аксона одного нейрона і дендрит іншого). Коли імпульс досягає синаптичного закінчення, там продукуються хімічні речовини - нейротрансмітери. Нейротрансмітери проходять через синаптичну щілину і в залежності від типу синапсу, збуджують або гальмують здатність нейрона-приймача генерувати електричні імпульси. Результативність синапсу налаштовується сигналами, які проходять скрізь нього, тому синапси навчаються у відповідності до активності процесів, у яких вони приймають участь. Нейрони взаємодіють за допомогою короткої серії імпульсів, повідомлення передається за допомогою частотно-імпульсної модуляції.

Біологічні нейрони є структурно складнішими, ніж існуючі штучні нейрони. Оскільки нейрофізіологія надає науковцям розширене розуміння дії нейронів, а технологія обчислень постійно вдосконалюється, розробники мереж мають необмежений простір для вдосконалення моделей біологічного мозку.

## 2. Штучний нейрон

Штучний нейрон є базовим модулем нейронних мереж. Він моделює основні функції природного нейрона (рис. 2).

При функціонуванні нейрон одночасно отримує багато вхідних сигналів. Кожен вхід має свою власну синаптичну вагу, яка надає входу вплив, необхідний для функції суматора елемента обробки. Ваги є мірою сили вхідних зв'язків і моделюють різноманітні синаптичні сили біологічних нейронів. Ваги суттєвого входу підсилюються і, навпаки, вага несуттєвого входу примусово зменшується, що визначає інтенсивність вхідного сигналу. Ваги можуть змінюватись відповідно до навчальних прикладів, топології мережі та навчальних правил.

Вхідні сигнали  $x_n$  зважені ваговими коефіцієнтами з'єднання  $w_n$  додаються, проходять через передатну функцію, генерують результат і виводяться.

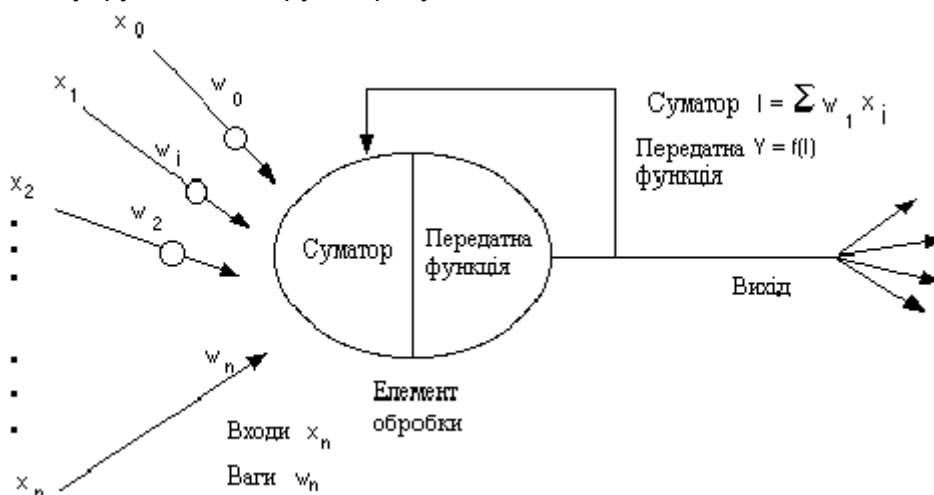


Рис. 2. Базовий штучний нейрон

В програмних реалізаціях штучні нейрони називають «елементами обробки» або «процесорами» і вкладають в них більше можливостей, ніж в базовому штучному нейроні, що описаний вище.

На рис. 3 зображена детальна схема штучного нейрону.

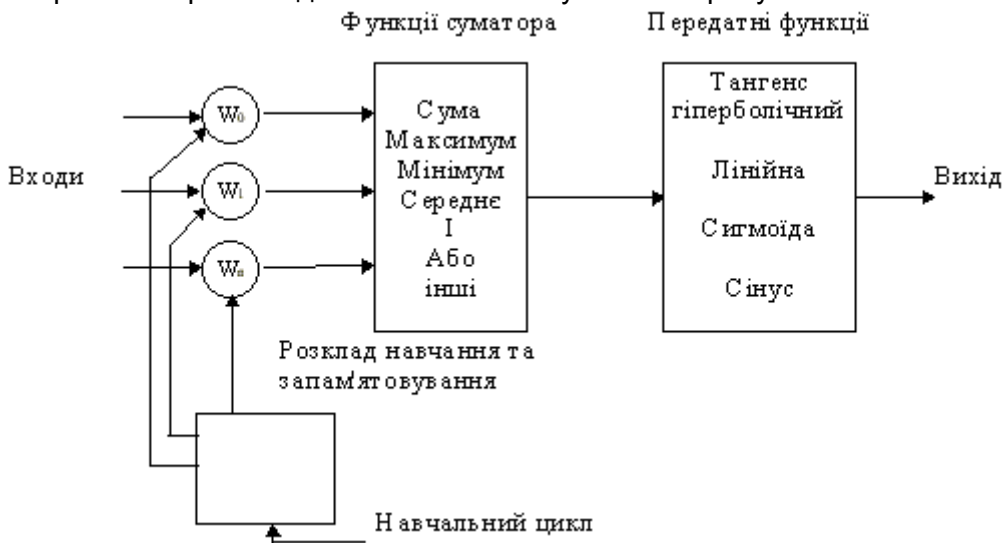


Рис. 3. Модель "елементу обробки"

Функція суматора може бути складнішою, наприклад, вибір мінімуму, максимуму, середнього арифметичного, добутку або обчислюватися за іншим алгоритмом. Багато програмних реалізацій використовують власні функції суматора, що запрограмовані на мові вищого рівня (C, C++, TurboPascal).

Перед надходженням до передатної функції вхідні сигнали та вагові коефіцієнти можуть комбінуватись багатьма способами. Алгоритми для комбінування входів нейронів визначають відповідно до мережної архітектури та парадигми.

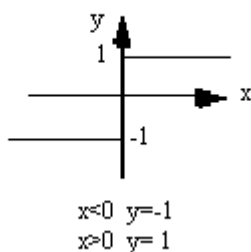
В деяких нейромережах суматор виконує додаткову обробку, так звану функцію активації, яка зміщує вихід функції суматора в часі. Цю функцію найкраще використовувати як компоненту мережі в цілому, ніж як компоненту окремого нейрона. Часто, ця функція є відсутньою.

Результат функції суматора перетворюється у вихідний сигнал через передатну функцію. В передатній функції для визначення виходу нейрона загальна сума порівнюється з деяким порогом (зазвичай, це діапазон [0, 1] або [-1, 1] або інше) за допомогою певного алгоритму.

Переважають застосовують нелінійну передатну функцію, оскільки лінійні (прямолінійні) функції є обмеженими і вихід є пропорційним до входу. Застосування лінійних передатних функцій було проблемою у ранніх моделях мереж, і їх обмеженість та недоцільність була доведена в книзі Мінські та Пейперта "Перцептрони".

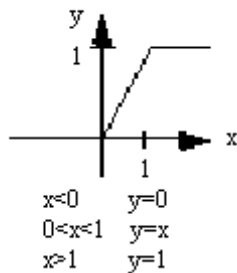
В існуючих нейромережах як передатну функцію використовують сигмоїду, синус, гіперболічний тангенс тощо. На рис. 4 зображені типові передатні функції.

Жорстка порогова функція

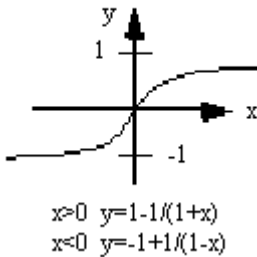


Для простої передатної функції нейромережа може видавати 0 чи 1, 1 чи -1 або інші числові комбінації. Передатна функція в таких випадках є пороговою або «жорстким обмежувачем» (рис. 4а).

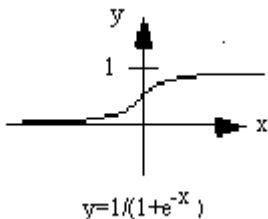
Лінійна з насиченням



**Передатна функція лінійна з насиченням** віддзеркалює вхід всередині заданого діапазону і діє як жорсткий обмежувач поза межами цього діапазону. Це лінійна функція, яка відсікається до мінімальних та максимальних значень, роблячи її нелінійною (рис. 4б).



Сигмоїда або S-подібна крива наближує мінімальне та максимальне значення у асимптотах. Вона називається **сигмоїдою** (рис. 4в), коли її діапазон  $[0, 1]$ , або **гіперболічним тангенсом** (рис. 4г), при діапазоні  $[-1, 1]$ . Важливою рисою сигмоїд є неперервність функцій та їх похідних. Застосування сигмоїдних функцій надає добрі результати і має широке застосування.



Для різних нейромереж можуть вибиратись інші передатні функції.

Після обробки сигналу, нейрон на виході має результат передатної функції, який надходить на входи інших нейронів або до зовнішнього з'єднання, як це передбачається структурою нейромережі.

### 3. Архітектура з'єднань штучних нейронів

Штучні нейромережі конструюються з базового блоку - штучного нейрону. Іншою властивістю нейромереж є величезна кількість зв'язків, які пов'язують окремі нейрони. Групування нейронів у мозку людини забезпечує обробку інформації динамічним, інтерактивним та самоорганізуючим шляхом.

Біологічні нейронні мережі з мікроскопічних компонентів існують у тривимірному просторі і здатні до різноманітних з'єднань. Але для реалізації штучних мереж присутні фізичні обмеження.

Об'єднуючись у мережі, штучні нейрони утворюють систему обробки інформації, яка забезпечує ефективну адаптацію моделі до постійних змін з боку зовнішнього середовища. В процесі функціонування мережі відбувається перетворення вхідного вектора сигналів у вихідний. Конкретний вид перетворення визначається архітектурою нейромережі, характеристиками нейронних елементів, засобами керування та синхронізації інформаційних потоків між нейронами.

Важливим фактором ефективності мережі є встановлення оптимальної кількості нейронів та типів зв'язків між ними.

Для опису нейромереж використовують кілька усталених термінів, які в різних джерелах можуть мати різне трактування, зокрема:

**Структура нейромережі** - спосіб зв'язків нейронів у нейромережі.

**Архітектура нейромережі** - структура нейромережі та типи нейронів.

**Парадигма нейромережі** - спосіб навчання та використання, іноді містить поняття архітектури.

На базі однієї архітектури може бути реалізовано різні парадигми нейромережі і навпаки.

Серед відомих архітектурних рішень виділяють групу **слабозв'язаних нейронних мереж**, у випадку, коли кожний нейрон мережі зв'язаний лише із сусідніми. В **повнозв'язаних нейромережах** входи кожного нейрона зв'язані з виходами всіх решти нейронів.

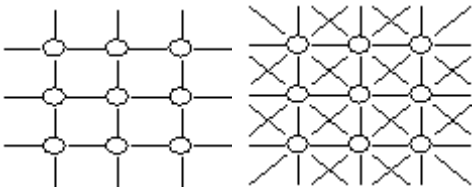


Рис. 5а. Слабозв'язані нейромережі

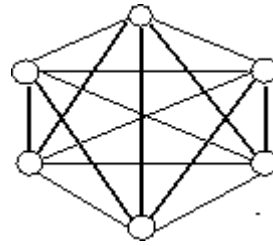


Рис. 5б. Повнозв'язані нейромережі

Самим поширеним варіантом архітектури є багатoshарові мережі. Нейрони в даному випадку об'єднуються у прошарки з єдиним вектором вхідних сигналів. Зовнішній вхідний вектор подається на вхідний прошарок нейронної мережі (рецептори). Виходами нейронної мережі є вихідні сигнали останнього прошарку (ефектори). Окрім вхідного та вихідного прошарків, нейромережа має один або кілька прихованих прошарків нейронів, які не мають контактів із зовнішнім середовищем.

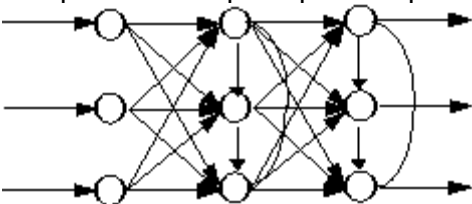


Рис. 6. Багатoshаровий тип з'єднання нейронів

Зв'язки між нейронами різних прошарків називають **проективними**.

Зв'язки між нейронами одного прошарку називають **бічними** (латеральними).

На рис. 6 показана типова структура штучних нейромереж. Хоча існують мережі, які містять лише один прошарок, або навіть один елемент, більшість застосувань вимагають мережі, які містять як мінімум три типи прошарків - вхідний, прихований та вихідний. Прошарок вхідних нейронів отримує дані або з вхідних файлів, або безпосередньо з електронних датчиків. Вихідний прошарок пересилає інформацію безпосередньо до зовнішнього середовища, до вторинного комп'ютерного процесу, або до інших пристроїв. Між цими двома прошарками може бути багато прихованих прошарків, які містять багато нейронів в різноманітних зв'язаних структурах. Входи та виходи кожного з прихованих нейронів сполучені з іншими нейронами.

Важливим аспектом нейромереж є напрямок зв'язку від одного нейрону до іншого:

Зв'язки скеровані від вхідних прошарків до вихідних називаються **аферентними**,

Зв'язки в зворотному напрямку називаються **еферентними**.

В більшості мереж кожен нейрон прихованого прошарку отримує сигнали від всіх нейронів попереднього прошарку чи від нейронів вхідного прошарку. Після виконання операцій над сигналами, нейрон передає свій вихід до всіх нейронів наступних прошарків, забезпечуючи передачу вперед (*feedforward*) на вихід.

При зворотному зв'язку, вихід нейронів прошарку скеровується до нейронів попереднього прошарку (рис. 7).

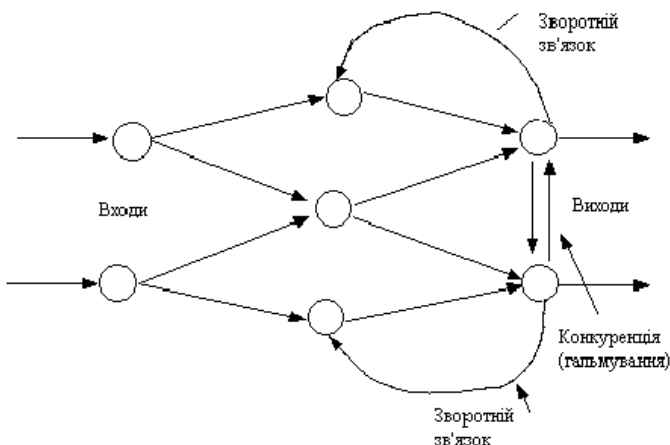


Рис.7

Напрямок зв'язків нейронів має значний вплив на роботу мережі. Більшість програмних нейромереж дозволяють користувачу додавати, вилучати та керувати з'єднаннями як завгодно. Корегуєючи параметри, можна налаштувати зв'язки як на посилення так і на послаблення величини сигналів.

За архітектурою зв'язків, більшість відомих нейромереж можна згрупувати у два великих класи: Мережі прямого поширення (з односкерованими послідовними зв'язками). Мережі зворотного поширення (з рекурентними зв'язками).

### Типові архітектури нейронних мереж

#### Мережі прямого поширення

Перцептрони  
Мережа Back Propagation  
Мережа зустрічного поширення  
Карта Кохонена

#### Рекурентні мережі

Мережа Хопфілда  
Мережа Хемінга  
Мережа адаптивної резонансної теорії  
Двоскерована асоціативна пам'ять

Мережі прямого поширення відносять до статичних, тут на входи нейронів надходять вхідні сигнали, які не залежать від попереднього стану мережі.

Рекурентні мережі вважаються динамічними, оскільки за рахунок зворотних зв'язків (петель) входи нейронів модифікуються в часі, що призводить до зміни станів мережі.

## 4. Навчання штучної нейронної мережі

Оригінальність нейромереж, як аналога біологічного мозку, полягає у здібності до навчання за прикладами, що складають навчальну множину. Процес навчання нейромереж розглядається як налаштування архітектури та вагових коефіцієнтів синаптичних зв'язків відповідно до даних навчальної множини для ефективного вирішення поставленої задачі.

Для навчання нейромереж можливо:

- Навчання з вчителем (контрольоване навчання)
- Навчання без вчителя (неконтрольоване навчання)

### Контрольоване навчання

Більшість реалізацій нейромереж використовують контрольоване навчання, де вихід, що змінюється, постійно порівнюється з бажаним виходом. Вагові коефіцієнти зв'язків на початку встановлюються випадково (*ініціалізація мережі*), але під час наступних ітерацій коректуються, щоб досягти близької відповідності між бажаним та біжучим виходами. Такі методи навчання націлені на мінімізації біжучих похибок всіх елементів обробки, що відбувається завдяки неперервній зміні синаптичних ваг до досягнення прийнятної точності мережі.

Перед використанням, нейромережа з контрольованим навчанням повинна бути навченою. Фаза навчання займає певний час. Навчання вважається закінченим при досягненні нейромережею визначеного користувачем рівня ефективності і бажаної статистичної точності. Після навчання вагові коефіцієнти зв'язків фіксуються для подальшого застосування. Деякі типи мереж дозволяють під час використання продовжувати навчання, і це допомагає мережі адаптуватись до змінних умов.

Навчальні множини повинні бути достатньо великими, щоб містити всю необхідну інформацію для виявлення важливих особливостей і зв'язків. Навчальні приклади повинні містити широке різноманіття даних. Якщо мережа навчається лише для одного прикладу, вагові коефіцієнти, що старанно встановлено для цього прикладу, радикально змінюються у навчанні для наступного прикладу. Попередні приклади при навчанні наступних просто забуваються. В результаті система повинна навчатись всьому разом, знаходячи найкращі вагові коефіцієнти для загальної множини прикладів.

Наприклад, у навчанні системи розпізнавання піксельних образів для десяти цифр, які представлені двадцятьма прикладами кожної цифри, всі приклади цифри "сім" не доцільно представляти послідовно. Краще надати мережі спочатку один тип представлення всіх цифр, потім другий тип і так далі.

Головною компонентою для успішної роботи мережі є представлення і кодування вхідних і вихідних даних. Штучні мережі працюють лише з числовими вхідними даними, отже, необроблені дані, що надходять із зовнішнього середовища повинні перетворюватись. Важливою є нормалізація даних,

тобто приведення всіх значень даних до єдиного діапазону. Нормалізація виконується шляхом ділення кожної компоненти вхідного вектора на довжину вектора, що перетворює вхідний вектор в одиничний. Попередня обробка зовнішніх даних, отриманих за допомогою сенсорів, у машинний формат є спільною і легко доступною для стандартних комп'ютерів.

Якщо після контрольованого навчання нейромережа ефективно опрацьовує дані навчальної множини, важливим стає її ефективність при роботі з даними, які не використовувались для навчання. У випадку отримання незадовільних результатів для тестової множини, навчання продовжується. Тестування використовується для забезпечення запам'ятовування не лише даних заданої навчальної множини, але і створення загальних образів, що можуть міститись в даних.

### **Неконтрольоване навчання**

Неконтрольоване навчання може бути великим надбанням у майбутньому. Воно проголошує, що комп'ютери можуть самонавчатись у справжньому роботизованому сенсі. На даний час, неконтрольоване навчання використовується в мережах відомих, як самоорганізовані карти (self organizing maps). Мережі не використовують зовнішніх впливів для коректування своїх ваг і внутрішньо контролюють свою ефективність, шукаючи регулярність або тенденції у вхідних сигналах та здійснюють адаптацію відповідно до навчальної функції. Навіть без повідомлення правильності чи неправильності дій, мережа повинна мати інформацію відносно власної організації, яка закладена у топологію мережі та навчальні правила.

Алгоритм неконтрольованого навчання скеровано на знаходження близькості між групами нейронів, які працюють разом. Якщо зовнішній сигнал активує будь-який вузол в групі нейронів, дія всієї групи в цілому збільшується. Аналогічно, якщо зовнішній сигнал в групі зменшується, це приводить до гальмуючого ефекту на всю групу.

Основу для навчання формує *конкуренція* між нейронами. Навчання конкуруючих нейронів підсилює відгуки певних груп на певні сигнали. Це пов'язує групи між собою та відгуком. При конкуренції змінюються ваги лише нейрона-переможця.

### **Оцінки навчання**

Оцінка ефективності навчання нейромережі залежить від кількох керованих факторів, важливими з яких є: ємність, складність зразків і обчислювальна складність.

**Ємність** показує, скільки зразків може запам'ятати мережа, і які межі прийняття рішень можуть бути на ній сформовані.

**Складність зразків** визначає кількість навчальних прикладів, необхідних для досягнення здатності мережі до узагальнення.

**Обчислювальна складність** напряму пов'язана з потужністю комп'ютера.

## **5. Основні етапи розв'язання задач за допомогою нейромереж**

Основні етапи розв'язання задач за допомогою нейромереж

- Збір даних для навчання;
- Підготовка і нормалізація даних;
- Вибір топології мережі;
- Експериментальний підбір характеристик мережі;
- Експериментальний підбір параметрів навчання;
- Власне навчання;
- Перевірка адекватності навчання;
- Коректування параметрів, остаточне навчання;
- Вербалізація мережі з метою подальшого використання.

Розглянемо докладніше деякі з цих етапів.

### **Збір даних для навчання**

Вибір даних для навчання мережі і їхня обробка є самим складним етапом розв'язання задачі. Набір даних для навчання повинний задовольняти декільком критеріям:

- Репрезентативність — дані повинні ілюструвати дійсне положення речей у предметній області;



- Несуперечність — суперечливі дані в навчальній вибірці призведуть до поганої якості навчання мережі;
- Обсяг — як правило, число записів у вибірці повинне на кілька порядків перевершувати кількість зв'язків між нейронами в мережі. У протилежному випадку мережа просто «запам'ятає» усю навчальну вибірку і не зможе виконати узагальнення.

### **Підготовка і нормалізація даних**

Вихідні дані перетворюються до виду, у якому їх можна подати на входи мережі. Кожен запис у файлі даних називається навчальною парою або навчальним вектором. Навчальний вектор містить по одному значенню на кожен вхід мережі і, у залежності від типу навчання (із вчителем або без), по одному значенню для кожного виходу мережі. Навчання мережі на «сирому» наборі, як правило, не дає якісних результатів. Існує ряд способів поліпшити «сприйняття» мережі.

- Нормування виконується, коли на різні входи подаються дані різної розмірності. Наприклад, на перший вхід мережі подаються величини зі значеннями від нуля до одиниці, а на другий — від ста до тисячі. При відсутності нормування значення на другому вході будуть завжди робити істотно більший вплив на вихід мережі, чим значення на першому вході. При нормуванні розмірності усіх вхідних і вихідних даних зводяться воедино;
- Квантування виконується над безперервними величинами, для яких виділяється кінцевий набір дискретних значень. Наприклад, квантування використовують для завдання частот звукових сигналів при розпізнаванні мови;
- Фільтрація виконується для «зашумлених» даних.

Крім того, велику роль грає саме представлення як вхідних, так і вихідних даних. Припустимо, мережа навчається розпізнаванню букв на зображеннях і має один числовий вихід — номер букви в алфавіті. У цьому випадку мережа одержить неправильне уявлення про те, що букви з номерами 1 і 2 більш схожі, чим букви з номерами 1 і 3, що, загалом, невірно. Для того, щоб уникнути такої ситуації, використовують топологію мережі з великим числом виходів, коли кожен вихід має свій зміст. Чим більше виходів у мережі, тим більша відстань між класами і тим складніше їх поплутати.

### **Вибір топології мережі**

Вибирати тип мережі необхідно виходячи з постановки задачі і наявних даних для навчання. Для навчання з учителем потрібна наявність для кожного елемента вибірки «експертної» оцінки. Іноді одержання такої оцінки для великого масиву даних просто неможливо. У цих випадках природним вибором є мережа, що навчається без учителя, наприклад, така як самоорганізуюча карта Кохонена або нейрона мережа Хопфілда. При розв'язанні інших задач, таких як прогнозування тимчасових рядів, експертна оцінка вже утримується у вихідних даних і може бути виділена при їхній обробці. У цьому випадку можна використовувати многошаровий перцептрон або мережу Ворда.

### **Експериментальний підбір характеристик мережі**

Після вибору загальної структури потрібно експериментально підібрати параметри мережі. Для мереж, подібних перцептроні, це буде число шарів, число блоків у схованих шарах (для мереж Ворда), наявність або відсутність обхідних з'єднань, передатні функції нейронів. При виборі кількості шарів і нейронів у них варто виходити з того, що здатності мережі до узагальнення тим вище, чим більше сумарне число зв'язків між нейронами. З іншого боку, число зв'язків обмежене зверху кількістю записів у навчальних даних. Експериментальний підбір параметрів навчання Після вибору конкретної топології, необхідно вибрати параметри навчання нейронної мережі. Цей етап особливо важливий для мереж, які навчаються з учителем. Від правильного вибору параметрів залежать не тільки те, наскільки швидко відповіді мережі будуть сходитися до правильних відповідей. Наприклад, вибір низької швидкості навчання збільшить час сходження, однак іноді дозволяє уникнути паралічу мережі. Збільшення моменту навчання може привести як до збільшення, так і до зменшення часу збіжності, у залежності від форми поверхні похибки. Виходячи з такого суперечливого впливу параметрів, можна зробити висновок, що їх значення потрібно вибирати експериментально, керуючись при цьому критерієм завершення навчання (наприклад, мінімізація похибки або обмеження за часом навчання).

### **Власне навчання мережі**

У процесі навчання мережа у визначеному порядку переглядає навчальну вибірку. Порядок перегляду може бути послідовним, випадковим і т.д. Мережі, які навчаються без учителя, переглядають вибірку тільки один раз. При навчанні з учителем мережа переглядає вибірку множини разів, при цьому один повний прохід по вибірці називається епохою навчання.

Зазвичай набір вихідних даних поділяють на дві частини — власне навчальну вибірку і тестові дані; принцип поділу може бути довільним. Навчальні дані подаються мережі для навчання, а перевірочні використовуються для розрахунку похибки мережі (перевірочні дані ніколи для навчання мережі не застосовуються). Таким чином, якщо на перевірочних даних похибка зменшується, то мережа дійсно виконує узагальнення.

Якщо похибка на навчальних даних продовжує зменшуватися, а похибка на тестових даних збільшується, виходить, мережа перестала виконувати узагальнення і просто «запам'ятовує» навчальні дані. Це явище називається перенавчанням мережі або оверфітінгом. У таких випадках навчання зазвичай припиняють. У процесі навчання можуть проявитися інші проблеми, такі як параліч або влучення мережі в локальний мінімум поверхні похибок. Неможливо заздалегідь пророчити прояв тієї або іншої проблеми, так само як і дати однозначні рекомендації до їх розв'язання.

### Перевірка адекватності навчання

Навіть у випадку успішного, на перший погляд, навчання мережа не завжди навчається саме тому, чого від неї хотів творець. Відомий випадок, коли мережа навчалася розпізнаванню зображень танків по фотографіях, однак пізніше з'ясувалося, що всі танки були сфотографовані на тому самому тлі. У результаті мережа «навчилася» розпізнавати цей тип ландшафту, замість того, щоб «навчитися» розпізнавати танки.

Таким чином, мережа «розуміє» не те, що від неї було потрібно, а те, що найпростіше узагальнити.

## Обґрунтованість застосування нейромереж

Нейромережі не можна вважати доцільним рішенням для всіх обчислювальних проблем. Для багатьох застосувань ідеальними є традиційні комп'ютери та обчислювальні методи. Сучасні цифрові обчислювальні машини перевершують людину за здатністю робити числові й символічні обчислення. Однак людина може без зусиль вирішувати складні задачі сприйняття зовнішніх даних (наприклад, впізнання людини в юрбі по його обличчю) з такою швидкістю і точністю, що значно перевищує можливості потужних комп'ютерів.

### Порівняння традиційного комп'ютера з біологічної нейронною системою

**	Комп'ютер з архітектурою фон Неймана	Нейрона система
<b>Процесор</b>	Складний Високошвидкісний Один чи декілька	Простий Низькошвидкісний Велика кількість
<b>Пам'ять</b>	Відділена від процесора Локалізована Адресація за адресою	Інтегрована в процесор Розподілена Адресація по змісту
<b>Обчислення</b>	Централізовані Послідовні Збережені програми	Розподілені Паралельні Самонавчання
<b>Надійність</b>	Висока вразливість	Живучість
<b>Спеціалізація</b>	Числові й символічні операції	Проблеми сприйняття
<b>Середовище функціонування</b>	Строго визначене Строго обмежене	Погано визначене Без обмежень
<b>Функції</b>	Логічно, через концепції, обчислення	Через зображення, рисунки, керування
<b>Метод навчання</b>	За правилами (дидактично)	За прикладами (сократично)
<b>Застосування</b>	Числова та символічна обробка інформації	Розпізнавання мови, розпізнавання образів,

## *Завдання, які вирішують за допомогою нейромодельовання.*

**Класифікація образів.** Визначення належності вхідного образу (наприклад, мовного сигналу чи рукописного символу), який представлено вектором ознак, до одного чи кількох попередньо визначених класів. До відомих застосувань відносяться розпізнавання букв, розпізнавання мови, класифікація сигналу електрокардіограми.

**Кластеризація/категоризація.** При вирішенні задачі кластеризації, що відома також як класифікація образів "без вчителя", навчальна множина з визначеними класами відсутня. Алгоритм кластеризації засновано на подібності образів і розміщенні близьких образів в один кластер. Застосовують кластеризацію для видобутку знань, стиснення даних і дослідження властивостей даних.

**Апроксимація функцій.** Припустимо, що є навчальна вибірка  $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$  (пари даних вхід-вихід), яка генерується невідомою функцією  $F$ , яка спотворена шумом. Завдання апроксимації полягає в знаходженні невідомої функції  $F$ . Апроксимація функцій необхідна при рішенні численних інженерних і наукових задач моделювання.

**Передбачення/прогноз.** Нехай задані  $n$  дискретних відліків  $\{y(t_1), y(t_2), \dots, y(t_n)\}$  у послідовні моменти часу  $t_1, t_2, \dots, t_n$ . Завдання полягає в передбаченні значення  $y(t_{n+1})$  у деякий майбутній момент часу  $t_{n+1}$ . Передбачення/прогноз мають значний вплив на прийняття рішень у бізнесі, науці й техніці (передбачення цін на фондовій біржі, прогноз погоди).

**Оптимізація.** Численні проблеми в математиці, статистиці, техніці, науці, медицині й економіці можуть розглядатися як проблеми оптимізації. Задачею алгоритму оптимізації є знаходження такого рішення, яке задовольняє системі обмежень і максимізує чи мінімізує цільову функцію.

**Пам'ять, що адресується за змістом.** В традиційних комп'ютерах звертання до пам'яті доступно лише за допомогою адреси, що не залежить від змісту пам'яті. Більш того, якщо допущено помилку в обчисленні адреси, то може бути знайдена зовсім інша інформація. Асоціативна пам'ять, чи пам'ять, що адресується за змістом, буде доступною за вказівкою заданого змісту. Вміст пам'яті може бути викликано навіть за частковим фрагментом або спотвореним змістом. Асоціативна пам'ять є потрібною при створенні мультимедійних інформаційних баз даних.

**Керування.** Розглянемо динамічну систему, задану сукупністю  $\{u(t), y(t)\}$ , де  $u(t)$  є вхідним керуючим впливом, а  $y(t)$  - виходом системи в момент часу  $t$ . В системах керування з еталонною моделлю метою керування є розрахунок такого вхідного впливу  $u(t)$ , при якому система діє за бажаною траєкторією, яка задана еталонною моделлю. Прикладом є оптимальне керування двигуном.

Незважаючи на переваги нейронних мереж в певних областях над традиційними обчисленнями, існуючі нейромережі не є досконалими рішеннями. Вони навчаються і можуть робити "помилки". Окрім того, не можна гарантувати, що розроблена мережа є оптимальною мережею.

Застосування нейромереж вимагає від розробника виконання ряду умов:

Наявність репрезентативної та достатньої за розміром множини даних для навчання й тестування мережі.

Розуміння базової природи проблеми, яка буде вирішена.

Вибір функції суматора, передатної функції та методів навчання.

Розуміння інструментальних засобів розробника.

Відповідна потужність обробки.

Робота з нейромоделями вимагає вміння розробника поза межами традиційних обчислень.

Спочатку, обчислення були лише апаратними й інженери робили їх працюючими.

Потім, були спеціалісти з програмного забезпечення: програмісти, системні інженери, спеціалісти по базах даних та проєктувальники.

Тепер є нейронні архітектори. Новий професіонал повинен мати високу кваліфікацію і різнобічні знання. Наприклад, він повинен знати статистику для вибору і оцінювання навчальних і тестувальних множин. Логічне мислення сучасних інженерів, їх емпіричне вміння та інтуїтивне відчуття гарантує створення ефективних нейромереж.