# Тема 9. Штучні нейронні мережі

Штучні нейромережі є моделями нейронної структури мозку, який здатен сприймати, обробляти, зберігати та продукувати інформацію, що представлена образами. 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, керування та інші інтелектуальні завдання, не використовуючи традиційного програмування.

У порівнянні з традиційними комп'ютерами з архітектурою фон Неймана, штучні нейронні мережі мають інші принципи обробки інформації та нові якості:

* Розподілене представлення інформації і паралельні обчислення.
* Адаптивність до змін у середовищі.
* Здатність до навчання й узагальнення.
* Толерантність до помилок.
* Низьке енергоспоживання.

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

## 9.1. Моделі штучного нейрону

Всі штучні нейромережі конструюються з базового формуючого блоку - штучного нейрону, що моделює основні функції природного нейрона.

При функціонуванні штучний нейрон одночасно отримує багато вхідних сигналів. Кожен вхід має власний коефіцієнт, що називається синаптичною вагою і моделює різноманітні синаптичні з’єднання біологічних нейронів. Ваговий коефіцієнт збільшує або зменшує значення входу, тому, 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 xn помножені на вагові коефіцієнти з'єднання wn додаються (блок суматора), проходять через передатну функцію та генерують результат (рис.9.1).





Рис. 9.1. Базовий штучний нейрон

Поглиблені знання щодо будови біологічного нейрона, як ефективного перетворюючого інструмента, можна розглядати як джерело базових ідей та концепцій по створенню нових типів нейромереж.

В наявних на цей час нейронних мережах штучні нейрони називаються «процесорами» або «елементами обробки» і вони мають значно більше можливостей, ніж базовий нейрон, що описано вище. На рис. 9.2 зображено детальну схему штучного нейрону.



Рис. 9.2. Розширена модель штучного нейрону

Модифіковані входи (зважені синаптичними коефіцієнтами) надходять до блока суматора, де, окрім додавання входів можна виконувати інші операції, наприклад, обрати середнє, найбільше чи найменше значення входів, здійснити логічні операції AND чи OR, або виконати інші функції. Інколи до блоку суматора додається функції активації, яка дозволяє функції суматора зміщуватися в часі.

Значення, що спродуковано в блоці суматора надсилається до передатної функції, яка за допомогою певного алгоритму, обмежує вихід в певному діапазоні, наприклад, [0 ÷ 1] чи [-1÷1]. В існуючих нейромережах в якості передатних функцій можуть бути використані сигмоїда, синус, гіперболічний тангенс та поліноми. Приклад передатної функції показано на рис. 9.3.



Рис. 9.3. Сигмоїдна передатна функція

Результат передатної функції надходить до входів інших нейронів або до зовнішнього з'єднання, відповідно до структури нейромережі.

## 9.2. Компоненти штучного нейрона

Незалежно від розташування та функціонального призначення, штучні нейрони мають спільні компоненти. Нижче представлено сім основних компонентів.

#### Компонента 1. Нормалізація входів

Нормалізація вхідних даних - це процес, при якому всі вхідні дані проходять процес "вирівнювання", тобто приведення до певного інтервалу, наприклад [0,1] або [-1,1]. Нормалізація дозволяє значно підвищити швидкість збіжності алгоритму навчання нейронної мережі, оскільки не можна порівнювати величини різних порядків. Якщо не провести нормалізацію, то вхідні дані будуть надавати додатковий вплив на нейрон, що може привести до невірних рішень.

#### Компонента 2. Вагові коефіцієнти

До нейрону одночасно надходять сигнали від багатьох входів. Кожен вхід змінює своє значення відповідно до коефіцієнта синаптичної ваги. Ваги суттєвого входу підсилюються і навпаки вага несуттєвого входу примусово зменшується, що визначає інтенсивність вхідного сигналу. Ваги можуть змінюватись відповідно до алгоритму навчання. Після проходження навчання ваги фіксуються і в подальшому використовуються.

#### Компонента 3. Функція суматора

Першим кроком дії нейрону є обчислення зваженої суми всіх входів. Математично, вхідні сигнали та відповідні їм ваги представлено векторами (х10, х20 ... хn0) та (w10, w20 . . . wn0). Добуток цих векторів є загальним вхідним сигналом.

Спрощеною функцією суматора є множення кожної компоненти вектора х на відповідну компоненту вектора w: вхід1 = х1 \* w1, вхід2 = х2 \* w2, і знаходження суми всіх добутків: вхід1 + вхід2 + . . . + вхідn. Результатом є не багатоелементний вектор, а єдине число.

Функція суматора може бути складнішою, наприклад, вибір мінімуму, максимуму, середнього арифметичного, добутку або інший нормалізуючий алгоритм. Вхідні сигнали та вагові коефіцієнти можуть комбінуватись багатьма способами перед надходженням до передатної функції. Особливі алгоритми для комбінування входів нейронів визначаються обраними мережною архітектурою та парадигмою (алгоритмом навчання).

В деяких нейромережах функції суматора виконують додаткову обробку, так звану функцію активації, яка зміщує вихід функції суматора відносно часу. Функції активації на сьогодні час використовуються обмежено і більшість сучасних нейронних реалізацій використовують функцію активації "тотожності", яка еквівалентна її відсутності. Цю функцію доцільніше використовувати як компоненту мережі в цілому, ніж як компоненту окремого нейрона.

#### Компонента 4. Передатна функція

Результат функції суматора перетворюється у вихідний сигнал через алгоритмічний процес відомий як передатна функція. У передатній функції для визначення виходу нейрона загальна сума порівнюється з деяким порогом. Якщо сума є більшою за значення порогу, елемент обробки генерує сигнал, в противному випадку сигнал не генерується або генерується гальмуючий сигнал.

Переважно застосовують нелінійну передатну функцію, оскільки лінійні функції є обмеженими, а вихід є пропорційним до входу. На рис. 9.4 зображено типові передатні функції.

|  |  |
| --- | --- |
|  | Для простої передатної функції нейромережа може видавати 0 та 1, 1 та -1 або інші числові комбінації. Передатна функція в таких випадках є "жорстким обмежувачем" або пороговою функцією (рис. 9.4 а). |
|  | Передатна функція лінійна з насиченням віддзеркалює вхід всередині заданого діапазону і діє як жорсткий обмежувач за межами цього діапазону. Це лінійна функція, яка відсікається до мінімальних та максимальних значень, роблячи її нелінійною (рис. 9.4б). |
|  | S-подібна передатна функція наближує мінімальне та максимальне значення у асимптотах і називається сигмоїдою (рис. 9.4в), коли її діапазон [0, 1]. |
|  | S-подібна передатна функція називається гіперболічним тангенсом (рис. 9.4г), при діапазоні [-1, 1]. Важливою рисою цих кривих є неперервність функцій та їх похідних. Застосування сигмоїдних функцій надає добрі результати і має широке застосування. |

Рис.9.4. Типові передатні функції

Для різних нейромереж можуть вибиратись інші передатні функції, зокрема поліноміальні.

Перед надходженням до передатної функції до вхідного сигналу деколи додають однорідно розподілений випадковий шум, джерело та кількість якого визначається режимом навчання. Цей шум називається "температурою" штучних нейронів, що надає математичній моделі елемент реальності.

#### Компонента 5. Вихідна функція (змагання)

По аналогії з біологічним нейроном, кожний штучний нейрон має один вихідний сигнал, який передається до сотень інших нейронів. Переважно, вихід є прямо пропорційним до результату передатної функції. В деяких мережних топологіях результати передатної функції змінюються для створення змагання між сусідніми нейронами. Нейронам дозволяється змагатися між собою, блокуючи дії нейронів, що мають слабший сигнал. Змагання (конкуренція) може відбуватись між нейронами, які знаходяться на одному або різних прошарках. По-перше, конкуренція визначає, який штучний нейрон буде активним і забезпечить вихідний сигнал. По-друге, конкуруючі виходи допомагають визначити, який нейрон буде брати участь у процесі навчання.

#### Компонента 6. Функція навчання

Метою функції навчання є налаштування вагових коефіцієнтів на входах кожного елемента обробки відповідно до певного алгоритму навчання для досягнення бажаного результату. Існує два типи навчання: контрольоване та неконтрольоване. Контрольоване навчання вимагає навчальної множини даних або вчителя, що ранжирує ефективність результатів мережі. У випадку неконтрольованого навчання система самоорганізовується за внутрішнім критерієм, закладеним в алгоритм навчання.

#### Компонента 7. Функція похибки та поширюване назад значення

У більшості мереж, що застосовують контрольоване навчання обчислюється різниця між спродукованим та бажаним виходом. Похибка відхилення (біжуча похибка) обробляється функцією похибки відповідно до заданої мережної архітектури. Після проходження всіх прошарків поточна похибка поширюється назад до попереднього прошарку, алгоритм навчання корегує вагові коефіцієнти, що враховується в наступному циклі навчання.

## 9.3. Архітектура з'єднань штучних нейронів

Об'єднуючись у мережі, нейрони утворюють системи обробки інформації, які забезпечують ефективну адаптацію моделі до постійних змін з боку зовнішнього середовища. В процесі функціонування мережі відбувається перетворення вхідного вектору сигналів у вихідний.

Конкретний вид перетворення визначається як архітектурою нейромережі так і характеристиками нейронних елементів, засобами керування та синхронізації інформаційних потоків між нейронами. Важливим фактором ефективності мережі є встановлення оптимальної кількості нейронів, типів зв'язків між ними та відповідних правил передачі інформації.

При описі 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 і поняття архітектури.

На основі однієї архітектури можуть бути реалізовані 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 (рис.9.5). Якщо входи кожного нейрона зв'язані з виходами всіх решта нейронів, тоді мова йде про повнозв'язані нейромережі (рис.9.6).

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Рис. 9.5. Слабозв'язані нейромережі | Рис. 9.6. Повнозв'язана мережа |

Зрозуміло, що такий поділ носить дещо теоретичний характер. Самим поширеним варіантом архітектури є багатошарові мережі. Нейрони в даному випадку об'єднуються у з'єднані між собою прошарки з єдиним вхідним вектором 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 декілька прихованих прошарків нейронів, які не мають контактів із зовнішнім середовищем.



Рис.9.7. Багатошаровий тип з'єднання нейронів

На рис.9.7 показано типову структуру 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 давачів. Вихідний прошарок надсилає інформацію безпосередньо до зовнішнього середовища, до вторинного комп'ютерного процесу або до інших пристроїв. Між цими двома прошарками може бути кілька прихованих прошарків, які містять багато нейронів у різноманітних зв'язаних структурах. Входи та виходи кожного з прихованих нейронів надходять до інших нейронів.

Зв'язки між нейронами різних прошарків називають **проективними**. Зв'язки, що скеровані від вхідних прошарків до вихідних називаються **аферентними**, в інакшому випадку, при зворотному напрямку зв'язків - **еферентними**. Зв'язки між нейронами одного прошарку відносять до бічних (**латеральних**).

Важливим аспектом нейромереж є напрямок зв'язку від одного нейрону до іншого. У більшості мереж кожен нейрон прихованого прошарку отримує сигнали від нейронів попереднього прошарку та зазвичай від нейронів вхідного прошарку. Після виконання операцій над сигналами, нейрон передає свій вихід до всіх нейронів наступних прошарків, забезпечуючи шлях передачі вперед (Feedforward) на вихідний прошарок (рис.9.8).



Рис. 9.8. Мережа з прямими і бічними зв’язками (прямого поширення)

При зворотному зв'язку, вихід нейронів прошарку скеровується до нейронів попереднього прошарку (рис. 9.9).



Рис.9.9. Мережа зі зворотнім зв'язком (зворотного поширення)

Спосіб, яким нейрони з'єднуються між собою має значний вплив на роботу мережі. Більшість програмних моделей дозволяють користувачу додавати, видаляти та керувати з'єднаннями як завгодно. Корегуючи параметри, зв'язки можна робити як збуджуючими так і гальмуючими.

За архітектурою зв'язків, більшість відомих нейромереж, що знайшли практичне застосування, можна згрупувати у два великих класи:

* Мережі прямого поширення (з односкерованими послідовними зв'язками).
* Мережі зворотного поширення (з рекурентними зворотними зв'язками).

Мережі прямого поширення відносять до статичних, оскільки на задані входи нейронів надходить вектор вхідних сигналів, що не залежить від попереднього стану мережі. Рекурентні мережі вважаються динамічними, оскільки за рахунок зворотних зв'язків (петель) входи нейронів модифікуються в часі, що приводить до змін станів мережі.

## 9.4. Навчання нейронних мереж

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 так, щоб ефективно вирішити поставлену задачу.

Після визначення архітектури нейронної мережі її потрібно «навчити», тобто підібрати такі значення її ваг, щоб вона працювала належним чином. Навчити нейронну мережу - значить, повідомити їй, чого від неї домагаються. Цей процес дуже схожий на навчання дитини алфавіту. Після того, як показали дитині зображення літери "А", слідує запитання: "Яка це літера?" Якщо відповідь невірна, то дитині підказують вірну відповідь: "Це літера А".

Дитина запам'ятовує приклад разом з вірною відповіддю, тобто в її пам'яті відбуваються деякі зміни в потрібному напрямку. Процес пред'явлення літер повторюється, поки всі літери будуть твердо запам'ятовані. Такий процес називають "навчання з вчителем". Навчання нейронної мережі відбувається аналогічно.

Процес навчання нейронної мережі полягає в налаштуванні її внутрішніх параметрів під конкретну задачу (рис.9.10).



Рис. 9.10. Процес навчання нейромережі

Алгоритм роботи нейронної мережі є ітеративним, його кроки називають епохами або циклами. **Епоха** - одна ітерація в процесі навчання, що містить пред'явлення всіх прикладів з навчальної множини і, можливо, перевірку якості навчання на контрольній множині.

Процес навчання здійснюється на **навчальній вибірці (множині)**. Навчальна вибірка містить набір даних про предметну область (об’єкт, явище, процес), що поділяється на вхідні значення і відповідні їм вихідні значення. Наприклад, після пред'явлення зображення літери "А" на вхід нейронної мережі, вона видає деяку відповідь, не обов'язково вірну.

В навчальній множині присутня вірна (бажана) відповідь, і сенсом навчання є те, щоб на виході нейронної мережі з міткою "А" рівень сигналу був максимальний. В ході навчання нейронна мережа знаходить певні залежності вихідних полів від вхідних.

Зазвичай, в якості бажаного виходу в задачі класифікації беруть набір (1, 0, 0, ...), де 1 стоїть на виході з міткою "А", а 0 - на всіх інших виходах. Обчислюючи різницю між бажаною і реальною відповідями мережі, утворюється похибка, що надалі обробляється відповідною функцією. Функція похибки - це цільова функція, що дозволяє оцінити якість роботи нейронної мережі і потребує мінімізації в процесі керованого навчання нейронної мережі.

Алгоритмом навчання є набір формул, який дозволяє за функцією похибки обчислити необхідні зміни для вагових коефіцієнтів зв’язків нейронної мережі.

Одну літеру (або різні зображення однієї літери) можна пред'являти нейронній мережі багато разів. У цьому сенсі навчання швидше нагадує повторення вправ в спорті - тренування. Після багаторазового пред'явлення прикладів вагові коефіцієнти зв’язків нейронної мережі стабілізуються, причому нейронна мережа надає правильні відповіді на всі (або майже всі) приклади з навчальної множини. У такому випадку говорять, що "нейронна мережа вивчила всі приклади", "нейронна мережа навчена", або "нейронна мережа натренована".

Які вхідні поля (ознаки) необхідно використовувати? Спочатку здійснюється евристичний вибір, далі кількість входів може бути змінено. Складність може викликати питання про кількість прикладів в наборі даних. Вся інформація, яку нейронна мережа має про завдання, міститься в наборі прикладів. Тому, якість навчання нейронної мережі безпосередньо залежить від кількості прикладів в навчальній вибірці, а також від того, наскільки повно ці приклади описують це завдання.

Так, наприклад, безглуздо використовувати нейронну мережу для передбачення фінансової кризи, якщо в навчальній вибірці не представлено жодної кризи. Для повноцінного навчання нейронної мережі потрібна репрезентативна і доволі велика вибірка з сотні (а краще тисячі) прикладів. Кількість необхідних прикладів залежить від складності розв'язуваної задачі.

Розробник має надати можливості вибору кількості прошарків у мережі і кількості нейронів в кожному прошарку. Далі необхідно призначити такі значення ваг, які зможуть мінімізувати похибку рішення. Від якості навчання нейронної мережі залежить її здатність вирішувати поставлені перед нею завдання.

### 9.4.1. Перенавчання нейронної мережі

При навчанні нейронних мереж часто виникає проблема перенавчання (overfitting). Перенавчання, або надмірно близька підгонка - надлишкова точна відповідність нейронної мережі до конкретного набору навчальних прикладів, при якому мережа втрачає здатність до узагальнення.

Перенавчання виникає у разі занадто довгого навчання, недостатньої кількості навчальних прикладів або занадто складної структури нейронної мережі.

Перенавчання пов'язано з тим, що вибір навчальної множини є випадковим. З перших кроків навчання відбувається зменшення похибки. На наступних кроках з метою зменшення похибки (цільової функції) параметри підлаштовуються під особливості навчальної множини. Однак при цьому відбувається "підлаштування" не під загальні закономірності ряду, а під особливості його частини - навчальної підмножини. При цьому точність прогнозу зменшується.

Один з варіантів боротьби з перенавчанням мережі - розділення навчальної вибірки на дві множини (навчальну і тестову). На навчальній множині відбувається навчання нейронної мережі. На тестовій множині здійснюється перевірка побудованої моделі. Ці множини не повинні перетинатися.

### 9.4.2. Застосування нейронної мережі

Після того, як нейронна мережа навчена, її можна застосовувати для вирішення різних завдань (рис.9.11).



Рис.9.11. Застосування нейромережі

Найважливіша особливість людського мозку полягає в тому, що, одного разу навчившись певного процесу, він може вірно діяти і в тих ситуаціях, в яких він не бував під час навчання. Наприклад, людина може читати майже будь-який почерк, навіть якщо бачить його вперше в житті.

Добре навчена нейронна мережа може з великою ймовірністю правильно реагувати на нові дані, що не були їй відомі раніше. Наприклад, можна написати літеру "А" іншим почерком, а потім запропонувати нейронній мережі класифікувати нове зображення. Ваги навченої нейронної мережі зберігають багато інформації про схожість і відмінності літер, тому, можна розраховувати на правильну відповідь і для нового варіанту зображення.

### 9.4.3. Контрольоване навчання

Значну кількість рішень отримано від нейромереж з контрольованим навчанням, де поточне значення виходу постійно порівнюється з бажаним. Ваги на початку встановлюються випадково, але під час наступних ітерації корегуються для досягнення близької відповідності між бажаним та поточним значенням на виході. Існуючі алгоритми навчання націлені на мінімізацію поточних похибок всіх елементів обробки, що відбувається за певний час неперервною зміною вагових коефіцієнтів до досягнення прийнятної точності мережі.

Перед використанням, нейромережа з контрольованим навчанням повинна бути навченою. Фаза навчання може тривати довго, зокрема, в прототипах систем з невідповідною процесорною потужністю навчання може займати кілька годин.

Навчання вважається завершеним при досягненні визначеного рівня ефективності нейромережі. Цей рівень означає, що мережа досягла бажаної статистичної точності, оскільки вона видає бажані виходи для заданої послідовності входів. Після навчання коефіцієнти синаптичних ваг фіксуються для подальшого застосування. Деякі типи мереж допускають навчання під час використання, що допомагає мережі адаптуватись до умов, які повільно змінюються.

Навчальні множини повинні бути достатньо великими, щоб містити всю необхідну інформацію для виявлення важливих особливостей і зв'язків. Також, навчальні приклади повинні містити широке різноманіття даних.

Якщо мережа навчається лише для одного прикладу, ваги старанно встановлені 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 десятьох 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.

Головним чинником для успішної роботи мережі є представлення вхідних і вихідних даних. Штучні мережі працюють лише з числовими вхідними даними, отже, необроблені дані, що надходять із зовнішнього середовища проходять процедуру нормалізації даних відповідно до діапазону всіх значень.

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.

### 9.4.4. Неконтрольоване навчання

Неконтрольоване навчання (Unsupervised Learning) - це один із підходів до навчання нейронних мереж, при якому модель навчається на даних, що не мають явних міток або цільових змінних. Натомість модель шукає приховані структури або патерни в даних самостійно.

На відміну від контрольованого навчання, де кожен навчальний приклад має відповідну мітку, у неконтрольованому навчанні дані не містять явних міток або цільових змінних. Модель повинна самостійно виявити приховані структури та патерни у даних. Це може застосовуватися для кластеризації даних, пошуку прихованих змінних або вивчення залежностей між ознаками.

Часто дані у неконтрольованому навчанні мають високу розмірність, тобто містять велику кількість ознак. Це може створювати складнощі в аналізі та обробці даних, а також вимагати застосування методів зниження розмірності. У неконтрольованому навчанні може бути складно інтерпретувати отримані результати. Оскільки немає явних міток, оцінка якості моделі та інтерпретація прихованих структур потребують додаткових методів та експертного аналізу.

Важливим також є проведення попередньої обробки даних, наприклад масштабування, нормалізацію або видалення викидів, щоб забезпечити ефективніше навчання моделі. Неконтрольоване навчання може бути чутливим до якості попередньої обробки даних.

Відносна ефективність неконтрольованого навчання залежить від конкретної задачі та даних. У деяких випадках, наприклад, завдання кластеризації, неконтрольоване навчання може бути дуже ефективним і допоможе виявити приховані групи або кластери в даних. Однак, без явних міток, важко оцінити якість та точність отриманих результатів. Ефективність неконтрольованого навчання значною мірою залежить від правильного вибору моделі, передобробки даних та інтерпретації отриманих результатів.

### 9.4.5. Оцінки навчання

Оцінка ефективності навчання нейромережі залежить від кількох керованих факторів. Розглядають три фундаментальні властивості, що пов'язані з навчанням: ємність, складність зразків і обчислювальна складність.

* **Ємність** вказує, скільки зразків може запам'ятати мережа, і які межі прийняття рішень можуть бути на ній сформовані.
* **Складність зразків** визначає число навчальних прикладів, необхідних для досягнення здатності мережі до узагальнення.
* **Обчислювальна складність** напряму пов'язана з потужністю комп’ютера.

## 9.5. Завдання, які вирішують за допомогою нейромереж

Нейронні мережі широко застосовуються для вирішення різних завдань у різних галузях.

* **Класифікація**. Нейронні мережі можуть бути використані для класифікації об'єктів на різні класи. Наприклад, розпізнавання зображень, де нейронна мережа може визначати, чи є зображення кішкою чи собакою.
* **Регресія.** Нейронні мережі можна використовувати для передбачення чисельних значень. Наприклад, прогноз ціни нерухомості на основі її характеристик або прогнозування часових рядів.
* **Кластеризація**. Нейронні мережі можуть використовуватися для угруповання об'єктів у кластери на основі їх подібності. Наприклад, кластеризація користувачів соціальних мереж для персоналізації рекомендацій.
* **Обробка природної мови.** Нейронні мережі можуть використовуватись для аналізу та розуміння природної мови. Це може включати завдання, такі як визначення тональності тексту, машинний переклад, системи голосового керування або перетворення аудіо на текст.
* **Голосові помічники та чат-боти**. Нейронні мережі використовуються для розробки голосових помічників і чат-ботів, які можуть розуміти і відповідати на природну мову користувачів, виконувати завдання та надавати інформацію.
* **Рекомендаційні системи**. Нейронні мережі можуть використовуватися для створення персональних рекомендацій для користувачів на основі їх переваг та поведінки. Це може містити рекомендації товарів, фільмів, музики тощо.
* **Обробка зображень та відео**. Нейронні мережі можуть використовуватись для аналізу та обробки зображень та відео. Це може включати завдання, такі як розпізнавання об'єктів на зображеннях, сегментація зображень, анотування відео тощо.
* **Розпізнавання жестів та емоцій**. Нейронні мережі можуть використовуватися для розпізнавання жестів та емоцій на основі відео або зображень. Це може бути корисним, наприклад, для розробки систем розпізнавання жестових команд або систем аналізу емоційного стану.
* **Генерація контенту**. Нейронні мережі можуть використовуватися для створення нового контенту, таких як зображення, тексти, музика та ін. Наприклад, створення автоматичних описів зображень або створення музики.
* **Автоматичне керування.** Нейронні мережі можуть бути використані для розробки систем автоматичного керування. Це може включати керування роботами, автомобілями, дронами чи іншими складними системами.
* **Робототехніка**. Нейронні мережі застосовуються в робототехніці для вирішення задач навігації, розпізнавання об'єктів, планування рухів та взаємодії з навколишнім середовищем. Вони дозволяють роботам навчатися та адаптуватися до нових умов та завдань.

Це лише деякі приклади областей, де нейронні мережі знаходять застосування. З появою нових ідей і технологій, застосування нейронних мереж продовжує розширюватися і включати все більше областей.

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, яку буде вирішено.
* Вибір методів навчання та відповідна потужність обробки.

## 9.6. Доцільність використання нейронних мереж

### Переваги перед традиційними обчислювальними методами

**Рішення задач в умовах невизначеності.** Завдяки здатності до навчання нейронна мережа дозволяє вирішувати завдання з невідомими закономірностями і залежностями між вхідними та вихідними даними, що дозволяє працювати з неповними даними.

**Стійкість до шумів у вхідних даних.** Нейронна мережа може самостійно виявляти неінформативні для аналізу параметри і видаляти їх, в зв'язку з чим відпадає необхідність у попередньому аналізі вхідних даних.

**Гнучкість структури нейронних мереж.** Компоненти нейрокомп'ютерів - нейрони і зв'язки між ними - можна комбінувати в різний спосіб. За рахунок цього один нейрокомп'ютер можна застосовувати для вирішення різних завдань, часто не пов'язаних між собою.

**Висока швидкодія.** Вхідні дані обробляються багатьма нейронами одночасно, завдяки чому нейронні мережі вирішують завдання швидше, ніж більшість інших алгоритмів.

**Адаптація до змін навколишнього середовища.** Нейронні мережі, навчаючись на даних, здатні підлаштовуватися під змінне навколишнє середовище (наприклад, зміни ситуації на ринку). Якщо необхідно вирішувати якесь завдання в умовах нестаціонарного середовища, то можуть бути створені нейронні мережі, що перенавчаються в режимі реального часу. Чим вище адаптивні здібності системи, тим більш стійкою буде її робота в нестаціонарному середовищі.

**Відмовостійкість нейронних мереж.** На несприятливу зміну умов нейромережа реагує лише незначним зниженням продуктивності. Ця особливість пояснюється розподіленим характером зберігання інформації в нейронної мережі, тому істотно вплинути на працездатність нейромережі можуть лише серйозні пошкодження структури.

### Недоліки нейронних мереж

**Відповідь, що надає нейронна мережа, є завжди приблизною.** Нейронні мережі не здатні давати точні і однозначні відповіді. Але завдання, в яких треба застосовувати нейромережі і одночасно отримувати точні відповіді, зустрічаються досить рідко.

**Нездатність прийняття рішень в кілька етапів.** Нейронна мережа не може вирішувати завдання, які вимагають послідовного виконання кількох кроків; вона здатна вирішувати завдання тільки "в один захід". Тому нейромережа не може, наприклад, довести математичну теорему.

**Трудомісткість і тривалість навчання.** Для того щоб нейронна мережа могла вірно вирішувати поставлені завдання, потрібно провести її навчання на десятках мільйонів наборів вхідних даних. Розроблені різні технології прискореного навчання та сучасні відеокарти дозволяють навчати нейромережі в сотні разів швидше. З'являються готові, попередньо навчені нейромережі, зокрема, що розпізнають образи. На основі таких нейромереж можна створювати додатки, не займаючись тривалим навчанням.

## Висновок

Нейромережа це машинна інтерпретація мозку людини, в якому знаходяться мільйони нейронів, що передають інформацію у вигляді електричних імпульсів і з'єднані між собою синаптичними зв’язками. Нейронні мережі використовуються для вирішення складних завдань, які вимагають аналітичних обчислень подібно до того, що робить людський мозок.

Структура нейронної мережі прийшла в світ програмування прямо з біології. Завдяки такій структурі, машина отримує можливість аналізувати і обробляти різну інформацію. Нейронні мережі також здатні відтворювати вхідну інформацію зі своєї пам'яті та надавати результати обчислень.

## Контрольні питання

1. Назвіть основні складові базового штучного нейрону, які функції вони виконують?
2. Яким чином працює блок суматора, які функції він може використовувати?
3. Для чого призначена передатна функція штучного нейрону, назвіть поширені нелінійні передатні функції
4. Який сенс містить поняття парадигми нейромережі?
5. Проведіть класифікацію нейромереж за різними видами зв’язків.
6. Які види навчання нейромереж вам відомі? Яка між ними різниця?
7. Назвіть основні етапи навчання з вчителем.
8. Перелічіть завдання, які вирішують за допомогою нейромоделювання.
9. Наведіть переваги використання нейромереж перед традиційними обчислювальними методами.
10. Які слабкі сторони притаманні нейромережам?

## Використані джерела

1. Нейронні мережі для початківців. Частина 1 - <https://habr.com/ru/post/312450/>
2. Нейронні мережі для початківців. Частина 2 - <https://habr.com/ru/post/313216/>
3. Нейронні мережі. Частина 1. Засади штучних нейронних мереж - <https://habr.com/ru/post/40137/>
4. Біологічний нейрон та його кібернетична модель- <https://intellect.icu/biologicheskij-nejron-i-ego-kiberneticheskaya-model-5561>
5. Нейрони та їх моделювання - <https://habr.com/ru/post/492308/>
6. Штучні нейронні мережі - <https://www.it.ua/ru/knowledge-base/technology-innovation/iskusstvennye-nejronnye-seti-ins>
7. Нейрокомп’ютерна техніка - Теорія і практика - <http://bookwu.net/book_nejrokompyuternaya-tehnika-teoriya-i-praktika_1188/3_glava-1.osnovy-iskusstvennyh-nejronnyh-setej>