

Тема 15. Архітектура нейрокомп'ютерів

15.1 Визначення поняття «нейрокомп'ютер»

В середині 80–х років минулого століття в США, а потім в Японії і країнах ЄЕС були розгорнуті широкомасштабні національні і міжнародні програми досліджень і розробок, які були направлені на створення нейрокомп'ютерів – ЕОМ на основі штучних нейронних мереж, які володіють розвиненим інтелектом і програмуються шляхом навчання на прикладах розв'язання задач.

У 1995 році була завершена розробка першого нейрокомп'ютера на стандартній мікропроцесорній елементній базі.

В останні роки, у зв'язку з бурхливим розвитком обчислювальної техніки, теорії хаосу і теорії самоорганізації, а також на підставі досягнень синергетики і теорії диссипативних структур (структур, фазовий обсяг яких зменшується з часом) спостерігається якісний бум у розвитку нейрокомп'ютерних технологій.

У світі існує декілька десятків спеціалізованих фірм, які випускають продукцію в області нейроінформатики, крім того, багато спеціалізованих комп'ютерних фірм – IBM, Siemens Nocsdorff, Mitsubisi – ведуть дослідження і мають власні розробки в даній області. Фірма Siemens в останні роки випускає спеціальні нейрочипи. Дані пристрої складаються з великого числа нейропроцесорів, здатних, на відміну від звичайних процесорів, робити послідовно–паралельні обчислення. Така схема обчислень пов'язана з особливістю роботи головного мозку людини, аналогом котрого і є нейрочипи.

В основу побудови нейрокомп'ютерів ліг штучний нейрон (перцептрон). Кожний нейрон отримує сигнали від сусідніх нейронів за допомогою спеціальних нервових волокон. Ці сигнали можуть бути збудливими або гальмуючими. Їх сума складає електричний потенціал у середині тіла нейрона. Коли потенціал перевищує деякий поріг, нейрон переходить у збуджений стан і посилає сигнал по вихідному нервовому волокну. Окремі штучні нейрони з'єднуються один з одним різноманітними методами. Це дозволяє створювати

різноманітні нейронні мережі з різною архітектурою, правилами навчання і можливостями.

Нейрокомп'ютер – це обчислювальна система з архітектурою MSIMD (паралельно-векторна модифікація), в якій реалізовані два принципових технічних рішення: спрощено до рівня нейрона процесорний елемент однорідної структури і різко ускладнені зв'язки між елементами; програмування обчислювальної структури перенесено на зміну вагових зв'язків між процесорними елементами.

Загальне визначення нейрокомп'ютера може бути представлено таким чином.

Нейрокомп'ютер – це обчислювальна система з архітектурою апаратного і програмного забезпечення, адекватною виконанню алгоритмів, представлених у нейромережному логічному базисі.

15.2 Архітектурні особливості й апаратне забезпечення нейрокомп'ютерів

Ідея побудови автомата на основі порогових елементів, подібних до нейронів (нервових клітин), які здатні виконувати логічні функції, була сформульована більше як півстоліття тому Мак–Каллоком і Піттсом. Однак задача проектування систем на основі порогових елементів викликала великі труднощі і її рішення було знайдене лише 20 років по тому. Це було настільки складним, що практично виключало можливість синтезу автоматів, які склалися більш як з десятків нейронів.

Системи на основі порогових елементів отримали назву штучних нейронних мереж (ШНМ) [23]. Перші роботоздатні штучні нейронні системи (ШНС) були створені вже в кінці 50–х років минулого століття – перцептрон Ф. Розенблатта, система "Альфа" А.Г. Івахненко. Перші великі перцептрони на основі аналогової і цифрової техніки ("Адам–А" і "Адам–Д") за межами США були створені в 1969–71 рр. в одному з київських НДІ.

Схема перцептрона Розенблатта приведена на рис. 15.1.

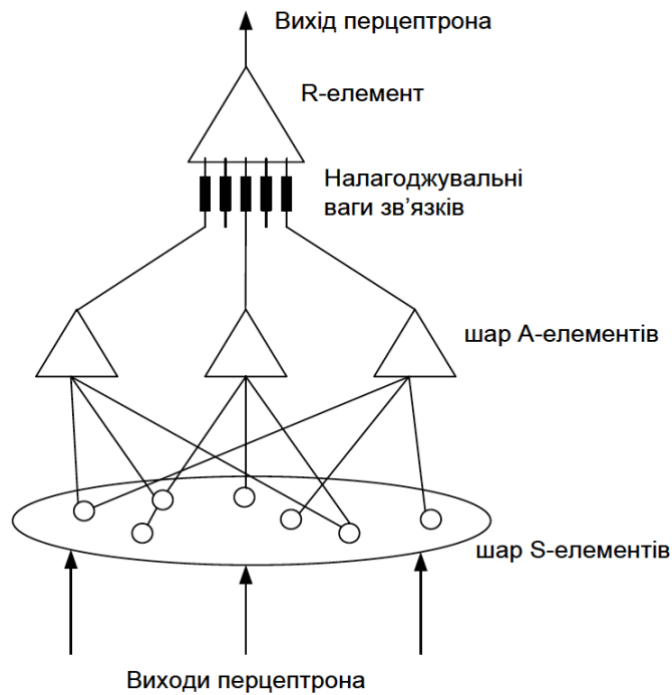


Рисунок 15.1 – Перцептрон Розенблатта

На рис. 15.2 приведена модель штучного нейрона. Штучний нейрон імітує в першому наближенні властивості біологічного нейрона.

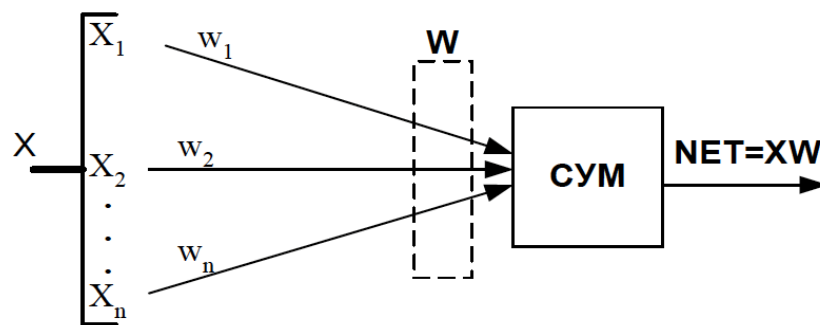


Рисунок 15.2 – Модель штучного нейрона

Множина вхідних сигналів x_1, x_2, \dots, x_n надходить на штучний нейрон. Дані вхідні сигнали, у сукупності позначені вектором X , відповідають сигналам, які приходять у синапси біологічного нейрона. Кожний сигнал збільшується на відповідну вагу w_i ($i=1,2,\dots, n$) і надходить у підсумовуючий блок СУМ (адаптивний суматор). Кожна вага відповідає «силі» одного синаптичного біологічного зв'язку. Множина ваг у сукупності утворює вектор ваг W .

Підсумовуючий блок, що відповідає тілу біологічного елемента, підсумовує зважені входи алгебраїчно, утворюючи вихід NET. У векторних позначеннях це може бути виражено таким чином: $NET=XW$. Сигнал NET у подальшому, як правило, перетворюється активаційною функцією F і дає вихідний нейронний сигнал OUT. Активаційна функція може бути звичайною лінійною функцією

$$OUT=K(NET),$$

де K – постійна граничної функції

$$OUT=1, \text{ якщо } NET > T$$

$$OUT=0 \text{ в інших випадках,}$$

де T– деяка постійна гранична величина.

На рис. 15.3 приведена структура штучного нейрона з активаційною функцією.

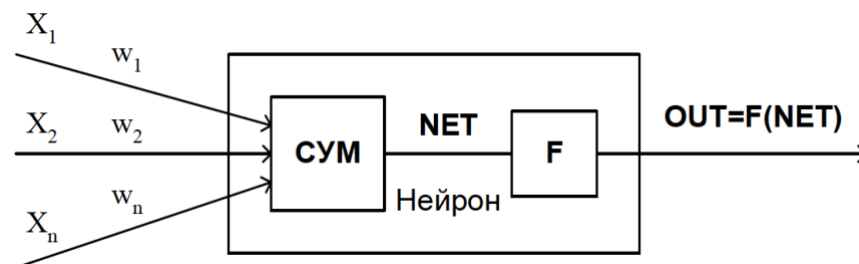


Рисунок 15.3 – Структура штучного нейрона з активаційною функцією

Блок F приймає сигнал NET і видає сигнал OUT. Якщо блок F звужує діапазон зміни величини NET так, що за будь-яких значень NET значення OUT належать деякому кінцевому інтервалу, то F називається стискаючою функцією. Як стискаюча функція часто використовується логістична чи сигмоїдальна (S-подібна) функція, що математично виражається так:

$$F(x)= 1/(1+e^{-x}).$$

Таким чином,

$$OUT= 1/(1+e^{-NET}).$$

За аналогією з електронними системами активаційну функцію можна вважати нелінійною підсилювальною характеристикою штучного нейрона.

Коефіцієнт підсилення обчислюється як відношення збільшення величини OUT до його невеликого збільшення, що викликало величини NET.

Розглянута модель штучного нейрона ігнорує багато властивостей свого біологічного аналога. Наприклад, вона не бере до уваги затримки в часі, що впливають на динаміку системи. Вхідні сигнали відразу ж породжують вихідний сигнал.

Починаючи з перцептрона основна увага розробників ШНС приділялась розробці і вдосконаленню методів їх навчання. Через відсутність надійної теорії навчання дані розробки носили в основному евристичний характер і отримали назву нейропарадигм.

На сьогоднішній день кількість проданих в світі нейрокомп'ютерів обчислюється десятками, а можливо і сотнями тисяч. В основному це нейрокомп'ютерні програми для персональних комп'ютерів, які призначені для розв'язання задач апроксимації і прогнозування числових даних. Близько 5% нейрокомп'ютерів відносяться до пристроїв професійного рівня, котрі орієнтовані на застосування потужних робочих станцій і апаратних нейроакселераторів. Програмне забезпечення таких систем, як правило, вміщує бібліотека нейропарадигм, що дозволяє під час розв'язання задач використовувати різні типи нейронних мереж.

Типовим прикладом може бути система Brain Maker фірми CSS (США). Система може працювати на будь-якому комп'ютері, де встановлено Windows. Базова версія орієнтована на широке коло користувачів. Її застосування не потребує спеціальних знань. Для розширення можливостей системи служить набір додаткових програм Toolkit Option, які дозволяють прискорити процес навчання і покращити подання графічних даних.

15.3 Нейрокомп'ютерні мережі

Штучна нейронна мережа може розглядатися як направлений граф зі зваженими зв'язками, в якому штучні нейрони є вузлами. За архітектурою зв'язків ШНМ можуть бути згруповані у два класи (рис. 15.4): мережі прямого

поширення, в яких графи не мають петель, і рекурсивні мережі, або мережі із зворотними зв'язками. У найбільш розповсюдженному сімействі мереж першого класу, багат шарових перцептронів, нейрони розташовані шарами і мають односпрямовані зв'язки між шарами.

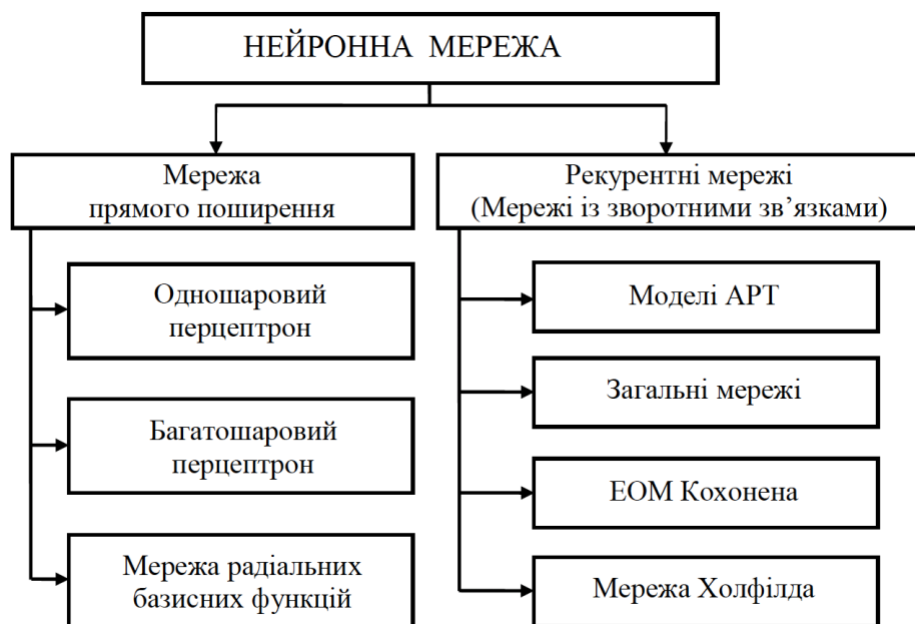


Рисунок 15.4 – Базові архітектури нейронних мереж

Мережі прямого поширення є статичними в тому сенсі, що на заданий вхід вони виробляють одну сукупність вихідних значень, які не залежать від попереднього стану мережі.

Рекурентні мережі є динамічними мережами в тому плані, що в силу зворотних зв'язків у них модифікуються входи нейронів, що приводить до зміни стану мережі.

Навчання штучних нейронних мереж. Здатність до навчання є фундаментальною властивістю мозку людини. У контексті ШНМ процес навчання може розглядатися як настроювання архітектури мережі і ваг зв'язків для ефективного виконання спеціальної задачі. Нейронна мережа повинна налагодити ваги зв'язків за наявною навчальною вибіркою. Функціонування мережі поліпшується в міру ітеративного настроювання вагових коефіцієнтів.

Властивість мережі навчатися на прикладах робить їх більш привабливими. Алгоритм навчання означає процедуру, в якій використовуються правила навчання для настроювання ваг.

Існують три парадигми навчання:

- з "учителем";
- без "учителя" (самонавчання);
- змішана (комбінована).

У першому випадку нейронна мережа має в своєму розпорядженні правильні відповіді на кожний вхідний приклад. Ваги налагоджуються так, щоб мережа робила відповіді найбільш близькі до відомих правильних відповідей. Посилений варіант навчання з "учителем" припускає, що відома тільки критична оцінка правильності виходу нейронної мережі, але не самі правильні значення виходу.

Навчання без "учителя" не вимагає знання правильних відповідей на кожний приклад навчальної вибірки. В даному випадку розкривається внутрішня структура даних чи кореляції між зв'язками в системі даних, що дозволяє розподілити зв'язки за категоріями.

Під час змішаного навчання частина ваг визначається за допомогою навчання з "учителем", в той час як інша за допомогою самонавчання.

Відомі чотири основних типи правил навчання: корекція помилково, машина Больцмана, правило Хебба і навчання методом змагання [23].

Правило корекції помилково. Під час навчання з "учителем" для кожного вхідного прикладу заданий бажаний вихід d . Реальний вихід мережі u може не збігатися з бажаним. Принцип корекції помилково під час навчання полягає у використанні сигналу $(d-u)$ для модифікації ваг, що забезпечує поступове зменшення помилки. Навчання має місце тільки у випадку, коли перцептрон помиляється.

Навчання Больцмана являє собою стохастичне правило навчання, котре впливає з інформаційних теоретичних і термодинамічних принципів. Метою навчання Больцмана є таке настроювання вагових коефіцієнтів, за якого стани

видимих нейронів задовольняють бажаний розподіл ймовірностей. Навчання Больцмана може розглядатися як спеціальний випадок корекції помилково, в якому під помилкою розуміється розбіжність кореляції станів у двох режимах.

Правило Хебба. Найстаршим навчальним правилом є постулат Хебба. Хебб спирався на такі нейрофізіологічні спостереження: якщо нейрони по обидва боки синапса активізуються одночасно регулярно, то сила синаптичного зв'язку зростає. Важливою особливістю цього правила є те, що зміна синаптичної ваги залежить тільки від активності нейронів, що зв'язані даним синапсом. Це істотно спрощує ланцюг навчання в реалізації VLSI.

Метод змагання. Під час навчання методом змагання вихідні нейрони змагаються між собою за активізацію. Дане явище відоме як правило "переможець бере все". Навчання за допомогою змагання дозволяє кластеризувати вхідні дані: подібні приклади групуються мережею відповідно до кореляцій і представляються одним елементом. Під час навчання модифікуються тільки ваги нейрона "переможця". Ефект даного правила досягається за рахунок такої зміни збереженого в мережі зразка (вектора ваг зв'язків, що переміг нейрона), за якого він стає трохи ближче до вхідного прикладу.

У порівнянні із звичайними комп'ютерами нейрокомп'ютери мають ряд переваг:

- висока швидкодія, котра пов'язана з тим, що алгоритми нейроінформатики мають високий ступінь паралельності;
- нейросистеми стійкі до завад;
- стійкі і надійні нейросистеми можуть створюватися з ненадійних елементів, які мають значний розкид параметрів.

Крім переваг нейросистеми мають і ряд недоліків:

- вони створюються спеціально для розв'язання конкретних задач, пов'язаних з нелінійною логікою і теорією самоорганізації. Розв'язання подібних задач на звичайних комп'ютерах можливе тільки чисельними методами.
- через свою унікальність дані пристрої достатньо дорогі.

Незважаючи на недоліки, нейрокомп'ютери можуть успішно використовуватися в різноманітних областях народного господарства:

- керування в режимі реального часу (літаками, ракетами, технологічними процесами беззупинного виробництва);
- розпізнавання об'єктів (букв і ієрогліфів, сигналів радара і сонара, відбитків пальців у криміналістиці, захворювань за симптомами у медицині, місцевостей під час пошуку корисних копалин і т.п.);
- прогнози: погоди, курсу акцій (та інших фінансових показників), політичних подій, поведження супротивників у військових конфліктах і в економічній конкуренції;
- оптимізація і пошук найкращих варіантів: під час конструювання технічних пристроїв, вибору економічної стратегії і т.п.