

Міністерство освіти і науки України
Чернівецький національний університет
імені Юрія Федьковича

НЕЙРОННІ МЕРЕЖІ
Методичні рекомендації та
завдання для лабораторних робіт

Для студентів спеціальностей:

«Комп'ютерні науки»,
«Програмна інженерія»,
«Комп'ютерна інженерія»

Чернівці
Чернівецький національний університет
2022

УДК 621.312, 004.032.026

Друкується за ухвалою Вченої ради Навчально-наукового інституту фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича, протокол № 8 від 22 вересня 2022 р.

Нейронні мережі: Методичні вказівки та завдання для лабораторних робіт, Укл. В.Г.Фратавчан, Т.М.Фратавчан. – Чернівці: ЧНУ, 2022 – 35 с.

Посібник містить завдання для лабораторних робіт та теоретичні відомості, необхідні для їх виконання, з дисципліни «Нейронні мережі».

Для студентів, що здобувають освіту в галузі знань „Інформатика і обчислювальна техніка” та інших спеціальностей з поглибленим вивченням інформаційних технологій.

Укладачі:

Фратавчан Валерій Григорович, кандидат фізико-математичних наук, доцент кафедри математичних проблем управління і кібернетики,

Фратавчан Тоня Михайлівна, кандидат фізико-математичних наук, доцент кафедри математичного моделювання Чернівецького національного університету ім.Ю.Федьковича

ВСТУП

Дисципліна «Нейронні мережі» є складовою частиною якісної підготовки фахівців в області ІТ технологій і ознайомлює студентів з новітніми та сучасними підходами до розв'язування складних задач обробки даних, виявлення закономірностей в числових рядах, розпізнавання образів.

Метою викладання навчальної дисципліни «Нейронні мережі» є

- ознайомлення студентів з новітніми та перспективними комп'ютерними технологіями обробки даних;
- вивчення основних архітектур штучних нейронних мереж;
- засвоєння методики розв'язування задач обробки даних нейромережевими засобами.

Основними завданнями вивчення дисципліни «Нейронні мережі» є:

- ознайомлення з поняттям штучного нейрону та штучної нейронної мережі;
- засвоєння схеми та методики роботи нейронного пристрою перцептронного типу;
- засвоєння схеми та принципу роботи одношарової асоціативної нейронної мережі;
- засвоєння схеми та принципів роботи двошарової нейронної мережі-класифікатору;
- ознайомлення зі схемою, методом навчання та роботи багатошарового перцептрону;
- ознайомлення з поняттям само організованої нейронної мережі;
- ознайомлення зі схемою та принципом роботи нейронних мереж когнітронного типу.

Згідно з вимогами освітньо-професійної програми студенти повинні:

знати :

- загальні схеми і конфігурації класичних штучних нейронних мереж;
- класифікацію нейронних мереж за призначенням та принципом роботи;
- класифікацію алгоритмів навчання нейронних мереж;
- математичні моделі розпізнавання образів на базі нейронних мереж;
- схему перцептрона;
- схеми основних нейронних мереж;
- схему генетичного алгоритму та його застосування для навчання;
- виявлення закономірностей в числових рядах із застосуванням само організованих нейронних мереж.

вміти :

- створювати модель задачі розпізнавання образів;
- реалізувати етапи попередньої обробки та побудови ознакових векторів;
- реалізувати алгоритми навчання та класифікації в задачах розпізнавання образів із застосуванням нейронних мереж;

- реалізувати схеми перцептронів та простих нейронних мереж в задачах розпізнавання образів;
- реалізувати системи розпізнавання за принципом асоціативної нейронної мережі;
- організувати системи класифікації на базі нейронної мережі Хеммінга;
- організувати систему розпізнавання з використанням неокогнітрону;
- проаналізувати статистичні набори даних з використанням карт Кохонена.

Зокрема, під час вивчення дисципліни формуються наступні компетентності (відповідно до галузевого стандарту):

Загальні компетентності:

- ЗК1. Здатність до абстрактного мислення, аналізу та синтезу.
- ЗК2. Здатність застосовувати знання у практичних ситуаціях.
- ЗК3. Знання та розуміння предметної області та розуміння професійної діяльності.
- ЗК6. Здатність вчитися й оволодівати сучасними знаннями.
- ЗК8. Здатність генерувати нові ідеї (креативність).
- ЗК12. Здатність оцінювати та забезпечувати якість виконуваних робіт.
- ЗК13. Здатність діяти на основі етичних міркувань.

Спеціальні (фахові, предметні) компетентності:

СК1. Здатність до математичного формулювання та досліджування неперервних та дискретних математичних моделей, обґрунтування вибору методів і підходів для розв'язування теоретичних і прикладних задач у галузі комп'ютерних наук, аналізу та інтерпретування.

СК2. Здатність до виявлення статистичних закономірностей недетермінованих явищ, застосування методів обчислювального інтелекту, зокрема статистичної, нейромережевої та нечіткої обробки даних, методів машинного навчання та генетичного програмування тощо.

СК3. Здатність до логічного мислення, побудови логічних висновків, використання формальних мов і моделей алгоритмічних обчислень, проектування, розроблення й аналізу алгоритмів, оцінювання їх ефективності та складності, розв'язності та нерозв'язності алгоритмічних проблем для адекватного моделювання предметних областей і створення програмних та інформаційних систем.

СК7. Здатність застосовувати теоретичні та практичні основи методології та технології моделювання для дослідження характеристик і поведінки складних об'єктів і систем, проводити обчислювальні експерименти з обробкою й аналізом результатів.

СК11. Здатність до інтелектуального аналізу даних на основі методів обчислювального інтелекту включно з великими та погано структурованими даними, їхньої оперативної обробки та візуалізації результатів аналізу в процесі розв'язування прикладних задач.

СК17. Здатність класифікувати задачі, створювати та реалізовувати

алгоритми синтезу оптимального керування процесів у кіберфізичних, економічних та соціальних динамічних системах.

Нормативний зміст підготовки здобувачів вищої освіти, сформульований у термінах результатів навчання:

ПР1. Застосовувати знання основних форм і законів абстрактно-логічного мислення, основ методології наукового пізнання, форм і методів вилучення, аналізу, обробки та синтезу інформації в предметній області комп'ютерних наук.

ПР3. Використовувати знання закономірностей випадкових явищ, їх властивостей та операцій над ними, моделей випадкових процесів та сучасних програмних середовищ для розв'язування задач статистичної обробки даних і побудови прогнозних моделей.

ПР4. Використовувати методи обчислювального інтелекту, машинного навчання, нейромережевої та нечіткої обробки даних, генетичного та еволюційного програмування для розв'язання задач розпізнавання, прогнозування, класифікації, ідентифікації об'єктів керування тощо.

ПР12. Застосовувати методи та алгоритми обчислювального інтелекту та інтелектуального аналізу даних в задачах класифікації, прогнозування, кластерного аналізу, пошуку асоціативних правил з використанням програмних інструментів підтримки багатовимірного аналізу даних на основі технологій DataMining, TextMining, WebMining.

Навчальна тема № 1: Біологічна та математична модель нейрону. Штучна нейронна мережа.

Зміст теми:

1. Структура нервової клітини. Функції елементів.
2. Математична модель нейрона. Вагові коефіцієнти. Передаточна функція.

Теоретичні відомості.

1. Структура нервової клітини. Функції елементів.

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 нейрону, і плазму, яка продукує необхідні для нейрону матеріали. Нейрон отримує сигнали (імпульси) від інших нейронів через дендрити (приймачі) і передає сигнали, згенеровані тілом клітки, вздовж аксона (передавач), що наприкінці розгалужується на волокна (*strands*). На закінченнях волокон знаходяться синапси (*synapses*).

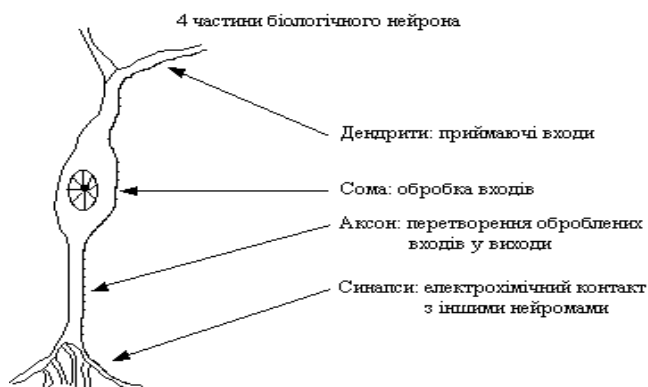


Рис. 1. Схема біологічного нейрона.

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

Останні експериментальні дослідження доводять, що біологічні нейрони структурно складніші, ніж спрощене пояснення, наведене вище і значно

складніші, ніж існуючі штучні нейрони, які є елементами сучасних штучних нейронних мереж. Оскільки нейрофізіологія надає науковцям розширене розуміння дії нейронів, а технологія обчислень постійно вдосконалюється, розробники мереж необмежений простір для вдосконалення моделей біологічного мозку.

2. Математична модель нейрона. Вагові коефіцієнти. Передаточна функція.

Базовий модуль нейронних мереж штучний нейрон моделює чотири основні функції природного нейрона (рис. 2).

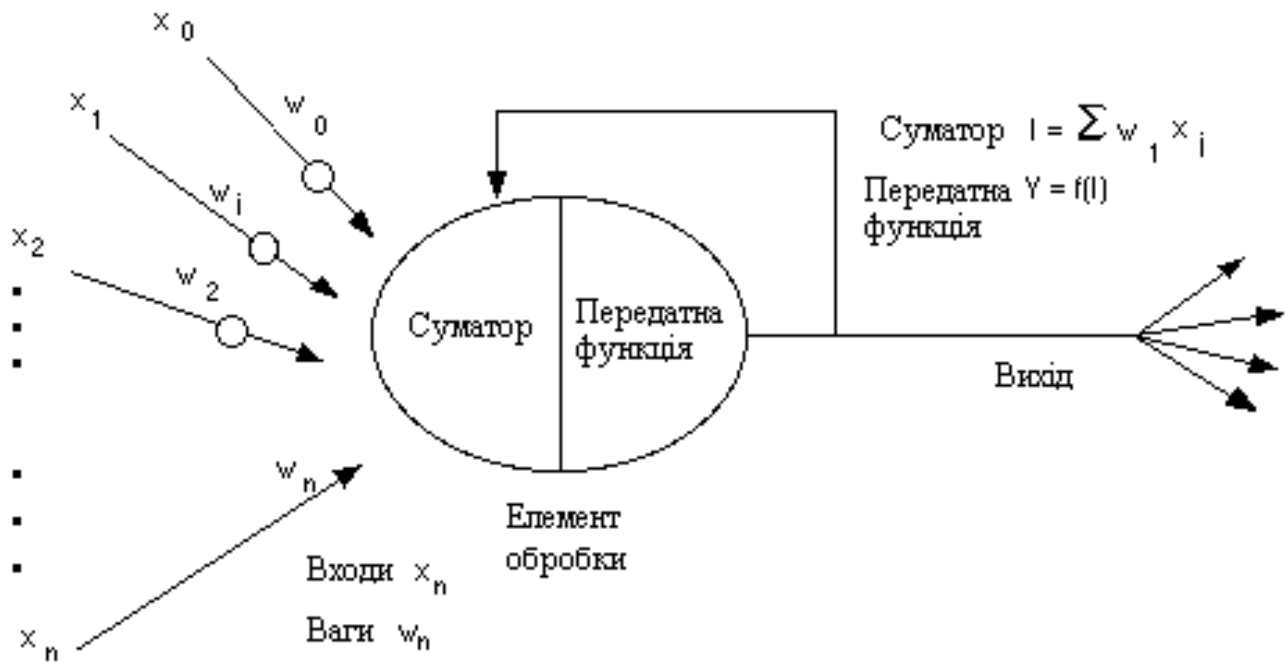


Рис. 2. Базовий штучний нейрон.

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

Лабораторна робота № 1: Створення середовища для попередньої обробки образів та формування ознакових векторів в системах розпізнавання.

У візуальному середовищі програмування розробити програмний комплекс для побудови абсолютних та нормованих векторів ознак розпізнавання стандартизованих графічних зображень (цифр, букв, логотипів).

Середовище повинно виконувати наступні функції:

- Послідовне введення графічних зображень з графічних файлів. Розміри зображень – змінні, але образ займає весь прямокутник зображення і не потребує додаткової локалізації у графічному компоненті. Зображення чорно-білі і не потребують додаткової соляризації та фільтрації;
- Відображення образу у графічному компоненті;
- Проведення сегментації за варіантом. Схема сегментації відображається на зображенні;
- Підрахунок точок образу на фоні зображення у кожному секторі сегментації та обчислення абсолютного вектора ознак. Значення вектора ознак відобразити у відповідному компоненті;
- Проведення нормування ознакових векторів за вказаним варіантом та відображення нормованих значень.

Варіанти сегментації:

- 1) Задана кількість комірок прямокутної табличної схеми;
- 2) Задана кількість однакових за розміром вертикальних смуг;
- 3) Задана кількість однакових за розміром горизонтальних смуг;
- 4) Задана кількість кутових секторів з центром у геометричному центрі зображення. Нульовий кут – вправо. Порядок нумерації – проти годинникової стрілки.
- 5) Задана кількість кутових секторів з центром у верхньому лівому куті зображення. Порядок нумерації – за годинникової стрілки.
- 6) Задана кількість кутових секторів з центром у середині верхньої межі зображення. Порядок нумерації – проти годинникової стрілки.
- 7) Задана кількість кутових секторів з центром у верхньому лівому куті зображення. Порядок нумерації – за годинникової стрілки.
- 8) Задана кількість кутових секторів з центром у нижньому лівому куті зображення. Порядок нумерації – за годинникової стрілки.

Варіанти нормування:

- 1) За сумою 1 (діленням всіх елементів абсолютного вектора ознак на суму ознакових значень).
- 2) За модулем 1 (діленням всіх елементів абсолютного вектору ознак на максимальний елемент вектора).

Навчальна тема № 2: Прості нейронні архітектури. Перцептрон Розенблата. Алгоритм навчання та класифікації.

Зміст теми:

1. Поняття перцептрону. Структура перцептрону Розенблата для двох класів.
2. Схема навчання. Оцінка функціональних можливостей.
3. Модель штучного нейрона як «елемента обробки».

Теоретичні відомості.

1. Перцептрон Розенблата

Першою моделлю нейромереж вважають перцептрон Розенблата. Теорія перцептронів є основою для багатьох типів штучних нейромереж прямого поширення і вони є класикою для вивчення.

Одношаровий перцептрон здатний розпізнавати найпростіші образи. Окремий нейрон обчислює зважену суму елементів вхідного сигналу, віднімає значення зсуву і пропускає результат через жорстку порогову функцію, вихід якої дорівнює +1 чи -1. В залежності від значення вихідного сигналу приймається рішення:

- +1 - вхідний сигнал належить класу А,
- -1 - вхідний сигнал належить класу В.

На рис. 1 показана схема нейронів, використовуваних в одношарових перцептронах, графік передатної функції і схема вирішальних областей, створених у багатовимірному просторі вхідних сигналів. Вирішальні області визначають, які вхідні образи будуть віднесені до класу А, які - до класу В. Перцептрон, що складається з одного нейрона, формує дві вирішальні області, розділені гіперплощиною. На рисунку показаний випадок, коли розмірність вихідного сигналу дорівнює 2. При цьому поділяюча поверхня уявляє собою пряму лінію на площині. Рівняння, що задає поділяючу пряму, залежить від значень синапсичних ваг і зсуву.

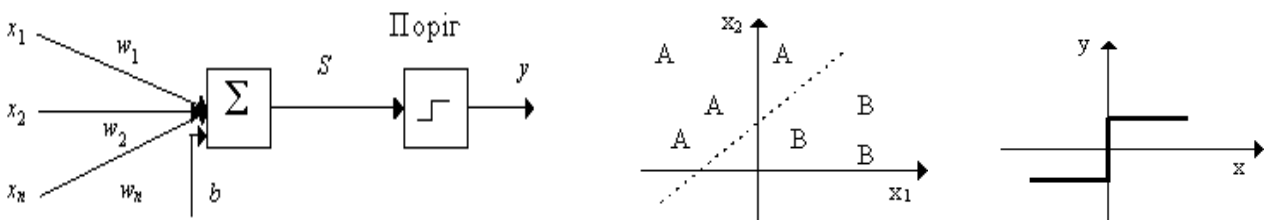


Рис. 1. Схема нейрона, графік передатної функції і поділяюча поверхня.

2. Алгоритм навчання одношарового перцептрону.

Алгоритм навчання одношарового перцептрона:

1. Ініціалізація синапсичних ваг і зсуву: синапсичні ваги приймають малі випадкові значення.
2. Пред'явлення мережі нового вхідного і бажаного вихідного сигналів: вхідний сигнал $x=(x_1, x_2, \dots, x_n)$ пред'являється нейрону разом з бажаним вихідним сигналом d .
3. Обчислення вихідного сигналу нейрона:

$$y(t) = f\left(\sum_{i=1}^N w_i(t)x_i(t) - b\right)$$

4. Налаштування значень ваг:

$$W_i(t+1) = w_i(t) + r[d(t) - y(t)]x_i(t), \quad i=1, \dots, N, \text{ тут}$$

$$d(t) = \begin{cases} +1, & \text{вихідний клас А} \\ -1, & \text{вихідний клас В} \end{cases},$$

де $w_i(t)$ - вага зв'язку від i -го елемента вхідного сигналу до нейрона в момент часу t , r - швидкість навчання (менше 1); $d(t)$ - бажаний вихідний сигнал.

Якщо мережа приймає правильне рішення, синаптичні ваги не модифікуються. В протилежному випадку:

5. Перехід до кроку 2.

Тип вхідних сигналів: бінарні чи аналогові (дійсні).

Розмірності входу і виходу обмежені при програмній реалізації тільки можливостями обчислювальної системи, на якій моделюється нейронна мережа, при апаратній реалізації - технологічними можливостями.

Області застосування: розпізнавання образів, класифікація.

Недоліки. Примітивні поділяючі поверхні (гіперплощини) дають можливість вирішувати лише найпростіші задачі розпізнавання.

Переваги. Програмні та апаратні реалізації моделі дуже прості. Простий і швидкий алгоритм навчання.

Модифікації. Багатошарові перцептрони дають можливість будувати більш складні поділяючі поверхні і тому більш поширені.

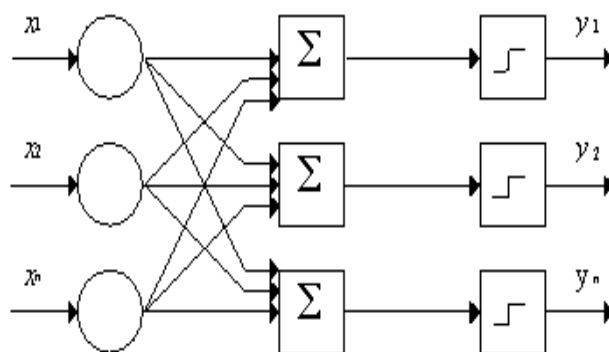


Рис. 2. Перцептрон із багатьма виходами.

3. Модель штучного нейрона як «елемента обробки».

У наявних на цей час пакетах програм штучні нейрони називаються "елементами обробки" і мають набагато більше можливостей, ніж простий штучний нейрон, описаний вище. На рис. 3 зображена детальна схема спрощеного штучного нейрону.

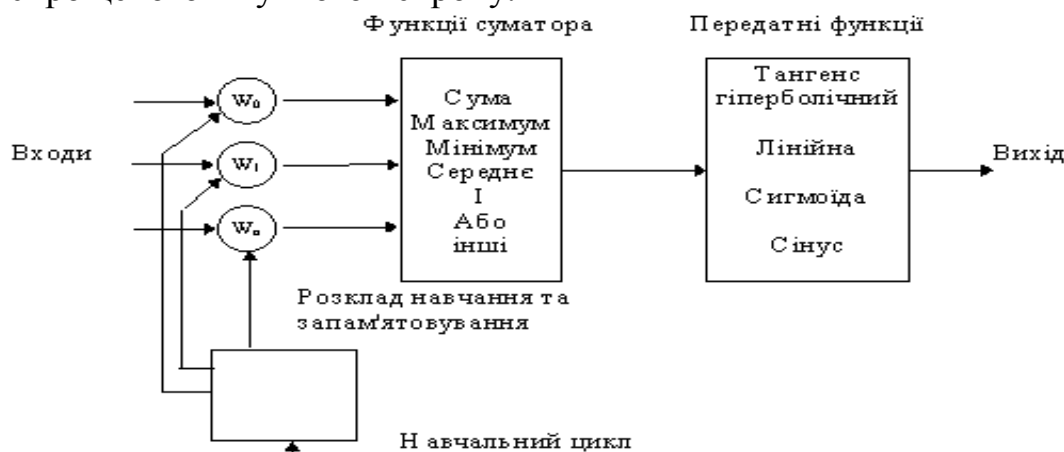


Рис. 3. Модель "елементу обробки".

Модифіковані входи передаються на функцію сумування, яка переважно тільки сумує добутки. Проте можна обрати багато різних операцій, такі як середнє, найбільше, найменше, *OR*, *AND*, тощо, які могли б виробляти деяку кількість різних значень. Окрім того, більшість комерційних програм дозволяють інженерам-програмістам створювати власні функції сумування за допомогою підпрограм, закодованих на мові високого рівня (C, C++, turborascal). Інколи функція сумування ускладнюється додаванням функції активації, яка дозволяє функції сумування оперувати в часі.

В будь-якому з цих випадків, вихід функції сумування надсилається у передатну функцію і скеровує весь ряд на дійсний вихід (0 або 1, -1 або 1, або яке-небудь інше число) за допомогою певного алгоритму. В існуючих нейромережах в якості передатних функцій можуть бути використані сигмоїда, синус,

гіперболічний тангенс та ін. Приклад того, як працює передатна функція показаний на рис. 4.



Рис. 4. Сигмоїдна передатна функція.

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.

Всі штучні нейромережі конструюються з базового формуючого блоку - штучного нейрону. Існуючі різноманітності і фундаментальні відмінності, є підставою мистецтва талановитих розробників для реалізації ефективних нейромереж.

Лабораторна робота № 2: Реалізація системи розпізнавання перцептронного типу для випадку двох класів.

Використовуючи основні можливості першої лабораторної роботи, створити систему розпізнавання перцептронного типу з нормованими векторами ознак числового типу. Система повинна забезпечити наступні дії:

- Ввести послідовно з графічних файлів образи тестової послідовності для двох класів та відобразити їх у відповідному графічному компоненті;
- Створення абсолютних та нормованих ознакових векторів для кожного зразка навчальної послідовності з відображенням абсолютних та нормованих ознакових значень;
- Задати початкові значення коефіцієнтів перцептрона Розенблата та поточний крок навчання;
- Обчислити вектор нев'язки та провести класифікацію за поточним станом перцептрона;
- В режимі навчання з вчителем підтвердити або спростувати результат класифікації;
- Провести корекцію перцептрона та відобразити його стан після корекції.

Оцінити збіжність процесу навчання.

Навчальна тема № 3: Нейронні мережі асоціативного та циклічного типу. ШНМ Хопфілда та Хеммінга.

Зміст теми:

1. Мережа Хопфілда.
2. Алгоритм функціонування мережі Хопфілда.
3. Модель двошарової нейронної мережі Хеммінга.

Теоретичні відомості.

1. Мережа Хопфілда.

Джон Хопфілд вперше представив свою асоціативну мережу у 1982 р. У Національній Академії Наук. На честь Хопфілда та нового підходу до моделювання, ця мережна парадигма згадується як мережа Хопфілда. Мережа базується на аналогії фізики динамічних систем. Початкові застосування для цього виду мережі включали асоціативну, або адресовану за змістом пам'ять та вирішували задачі оптимізації.

Мережа Хопфілда використовує три прошарки: вхідний, прошарок Хопфілда та вихідний прошарок. Кожен прошарок має однакову кількість нейронів. Входи прошарку Хопфілда під'єднані до виходів відповідних нейронів вхідного прошарку через змінні ваги з'єднань. Виходи прошарку Хопфілда під'єднуються до входів всіх нейронів прошарку Хопфілда, за винятком самого себе, а також до відповідних елементів у вихідному прошарку. В режимі функціонування, мережа скеровує дані з вхідного прошарку через фіксовані ваги з'єднань до прошарку Хопфілда. Прошарок Хопфілда коливається, поки не буде завершена певна кількість циклів, і біжучий стан прошарку передається на вихідний прошарок. Цей стан відповідає образу, вже запрограмованому у мережу.

Навчання мережі Хопфілда вимагає, щоб навчальний образ був представлений на вхідному та вихідному прошарках одночасно. Рекурсивний характер прошарку Хопфілда забезпечує засоби корекції всіх ваг з'єднань. Недвійкова реалізація мережі повинна мати пороговий механізм у передатній функції. Для правильного навчання мережі відповідні пари "вхід-вихід" мають відрізнятися між собою.

Якщо мережа Хопфілда використовується як пам'ять, що адресується за змістом вона має два головних обмеження. По-перше, число образів, що можуть бути збережені та точно відтворені є строго обмеженим. Якщо зберігається занадто багато параметрів, мережа може збігатись до нового неіснуючого образу, відмінному від всіх запрограмованих образів, або не збігатись взагалі. Межа ємності пам'яті для мережі приблизно 15% від числа нейронів у прошарку Хопфілда. Другим обмеженням парадигми є те, що прошарок Хопфілда може стати нестабільним, якщо навчальні приклади є

занадто подібними. Зразок образу вважається нестабільним, якщо він застосовується за нульовий час і мережа збігається до деякого іншого образу з навчальної множини. Ця проблема може бути вирішена вибором навчальних прикладів більш ортогональних між собою.

Структурна схема мережі Хопфілда приведена на рис. 1.

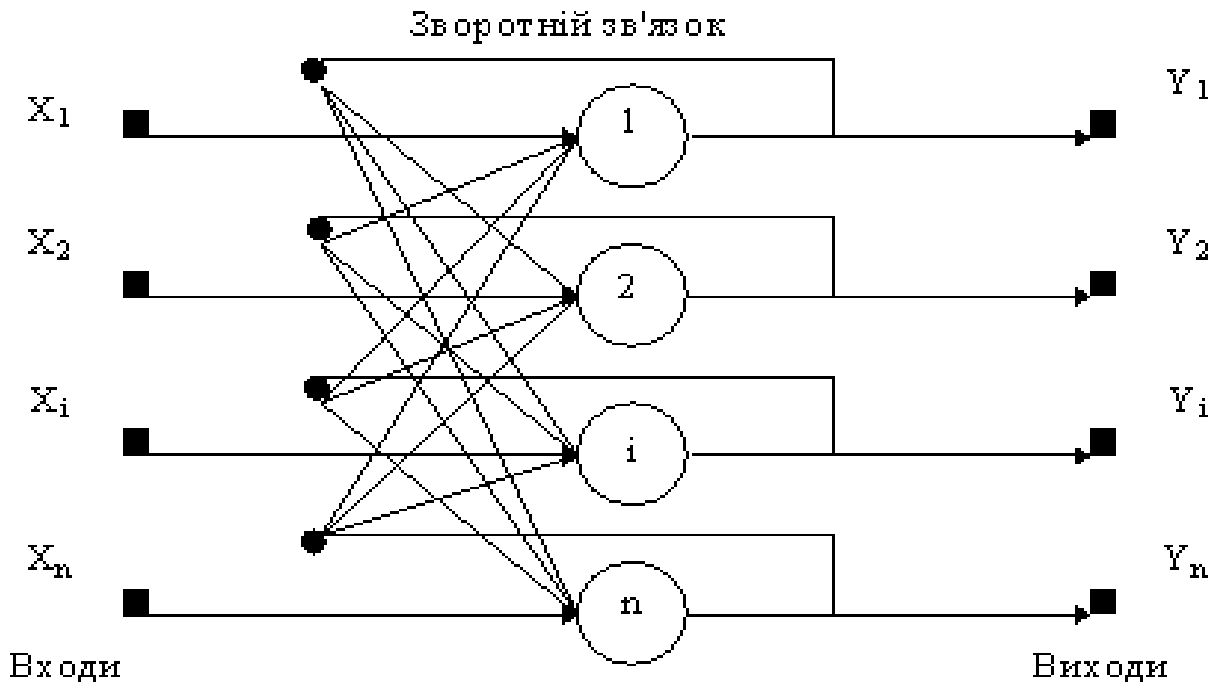


Рис. 1. Структурна схема мережі Хопфілда.

Задача, розв'язувана даною мережею в якості асоціативної пам'яті, як правило, формулюється таким чином. Відомий деякий набір двійкових сигналів (зображень, звукових оцифровок, інших даних, що описують якийсь об'єкти або характеристики процесів), вважають зразковим. Мережа повинна вміти з зашумленого сигналу, поданого на її вхід, виділити ("пригадати" по частковій інформації) відповідний зразок або "дати висновок" про те, що вхідні дані не відповідають жодному із зразків. У загальному випадку, будь-який сигнал може бути описаний вектором x_1, x_i, x_n, \dots , n - число нейронів у мережі і величина вхідних і вихідних векторів. Кожний елемент x_i дорівнює або $+1$, або -1 . Позначимо вектор, що описує k -ий зразок, через X_k , а його компоненти, відповідно, - x_{ik} , $k=0, \dots, m-1$, m - число зразків. Якщо мережа розпізнає (або "пригадає") якийсь зразок на основі пред'явлених їй даних, її виходи будуть містити саме його, тобто $Y = X_k$, де Y - вектор вихідних значень мережі: y_1, y_i, y_n . У протилежному випадку, вихідний вектор не співпаде з жодним зразковим. Якщо, наприклад, сигнали являють собою якесь зображення, то, відобразивши у графічному виді дані з виходу мережі, можна буде побачити картинку, що цілком збігається з однією зі зразкових (у випадку успіху) або ж "вільну імпровізацію" мережі (у випадку невдачі).

2. Алгоритм функціонування мережи Хопфілда.

Алгоритм функціонування мережі

1. На стадії ініціалізації мережи вагові коефіцієнти синапсів встановлюються таким чином:

$$w_{ij} = \begin{cases} \frac{\sum_{k=1}^m x_i^k x_j^k}{m}, & i \neq j \\ 0, & i = j \end{cases}$$

Тут i і j - індекси, відповідно, предсинаптичного і постсинаптичного нейронів; x_{ik} , x_{jk} - i -ий і j -ий елементи вектора k -ого зразка.

2. На входи мережі подається невідомий сигнал (t - номер ітерації). Його поширення безпосередньо встановлює значення виходів:

$$Y_i(0) = x_i, \quad i = 0 \dots n-1,$$

Тому позначення на схемі мережі вхідних сигналів у явному виді носить чисто умовний характер. Нуль у скобці справа від y_i означає нульову ітерацію в циклі роботи мережі.

3. Розраховується новий стан нейронів :

$$S(t+1) = \sum_{j=1}^n w_{ij} y_j(t), \quad j=0 \dots n-1.$$

І нові значення виходів:

$$y_j(t+1) = f|S_j(t+1)|$$

Де f - передатна функція у виді порогової, приведена на рис. 10.

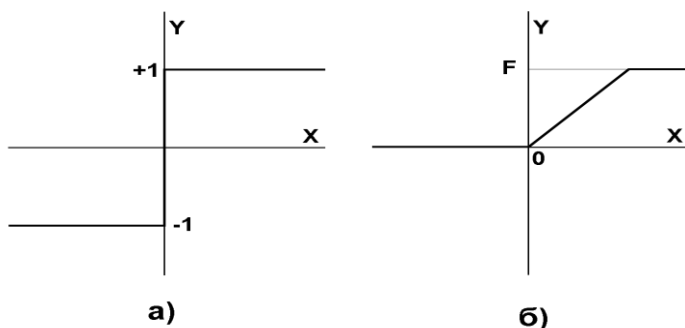


Рис. 10. Передаточні функції.

4. Перевіряємо чи змінилися вихідні значення виходів за останню ітерацію. Якщо так - перехід до пункту 2, інакше (якщо виходи стабілізувались) - кінець. При цьому вихідний вектор являє собою зразок, що найкраще відповідає вхідним даним.

Іноді мережа не може провести розпізнавання і видає на виході неіснуючий образ. Це пов'язано з проблемою обмеженості можливостей мережі. Для мережі Хопфілда число запам'ятованих образів m не повинно перевищувати величини, приблизно рівної $0.15 \cdot n$. Крім того, якщо два образи А і Б сильно схожі, вони, можливо, будуть викликати в мережі перехресні асоціації, тобто пред'явлення на входи мережі вектора А призведе до появи на її виходах вектори Б і навпаки.

4. Модель двошарової нейронної мережі Хеммінга.

Мережа Хеммінга (*Hamming*) є розширенням мережі Хопфілда. Ця мережа була розроблена Річардом Ліппманом (*Richard Lippman*) у середині 80-х рр. Мережа Хеммінга реалізує класифікатор, що базується на найменшій похибці для векторів двійкових входів, де похибка визначається відстанню Хеммінга. Відстань Хеммінга визначається як число бітів, які відрізняються між двома відповідними вхідними векторами фіксованої довжини. Один вхідний вектор є незашумленим прикладом образу, інший є спотвореним образом. Вектор виходів навчальної множини є вектором класів, до яких належать образи. У режимі навчання вхідні вектори розподіляються до категорій для яких відстань між зразковими вхідними векторами та біжучим вхідним вектором є мінімальною.

Мережа Хеммінга має три прошарки: вхідний прошарок з кількістю вузлів, скільки є окремих двійкових ознак; прошарок категорій (прошарок Хопфілда), з кількістю вузлів, скільки є категорій або класів; вихідний прошарок, який відповідає числу вузлів у прошарку категорій.

Мережа є простою архітектурою прямого поширення з вхідним рівнем, повністю під'єднаним до прошарку категорій. Кожен елемент обробки у прошарку категорій є зворотно під'єднаним до кожного нейрона у тому ж самому прошарку і прямо під'єднаним до вихідного нейрону. Вихід з прошарку категорій до вихідного прошарку формується через конкуренцію.

Навчання мережі Хеммінга є подібним до методології Хопфілда. На вхідний прошарок надходить бажаний навчальний образ, а на виході вихідного прошарку надходить значення бажаного класу, до якого належить вектор. Вихід містить лише значення класу до якої належить вхідний вектор. Рекурсивний характер прошарку Хопфілда забезпечує засоби корекції всіх ваг з'єднань.

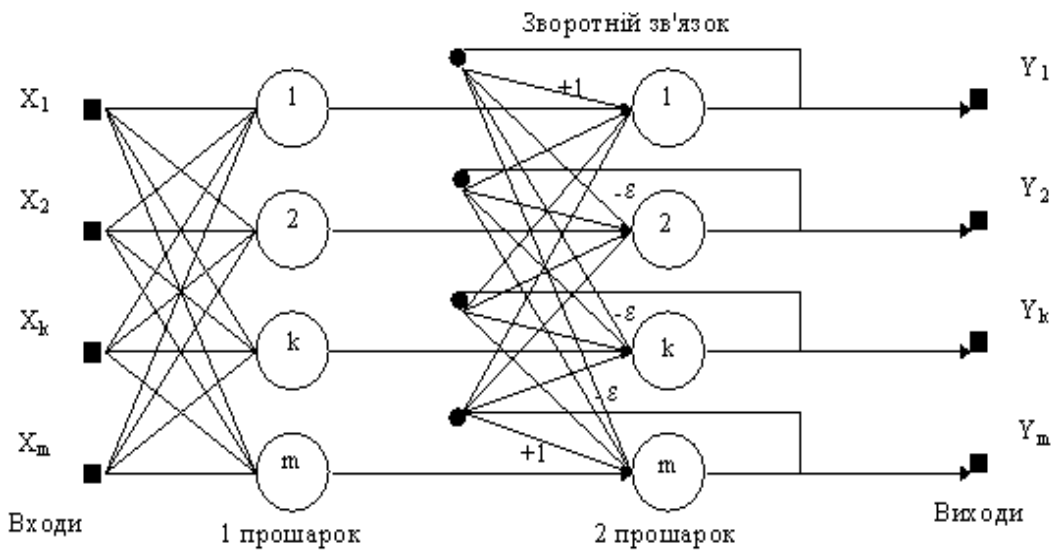


Рис. 11. Структурна схема мережі Хеммінга.

Алгоритм функціонування мережі Хеммінга:

1. На стадії ініціалізації ваговим коефіцієнтам першого прошарку і порогу передатної функції присвоюються такі значення:

$$W_{ik} = x_i^k / 2, \quad i=0 \dots n-1, \quad k=0 \dots m-1$$

$$B_k = n / 2, \quad k = 0 \dots m-1 .$$

Тут x_i^k - i -ий елемент k -ого зразка.

Вагові коефіцієнти гальмуючих синапсів у другому прошарку беруть рівними деякій величині $0 < \nu < 1/m$. Синапс нейрона, пов'язаний із його ж виходом має вагу $+1$.

2. На входи мережі подається невідомий вектор $x_1, x_i, x_n \dots$. Розраховуються стани нейронів першого прошарку (верхній індекс у скобках указує номер прошарку):

$$y_j^{(1)} = S_j^{(1)} = \sum_{i=1}^n w_{ij} x_i + b_j, \quad j=0 \dots m-1.$$

Після цього отримання значення ініціалізують значення виходів другого прошарку:

$$Y_j^{(2)} = y_j^{(1)}, \quad j = 0 \dots m-1.$$

3. Обчислюються нові стани нейронів другого прошарку:

$$S_j^{(2)}(t+1) = y_j(t) - \varepsilon \sum_{k=1}^m y_k^{(2)}(t), k \neq j, j = 1 \dots m$$

І значення їх виходів:

$$y_j^{(1)}(t+1) = f|S_j^{(2)}(t+1)|.$$

Передатна функція f має вид порога, причому величина b повинна бути достатньо великою, щоб будь-які можливі значення аргументу не призводили до насичення.

4. Перевіряється, чи змінилися виходи нейронів другого прошарку за останню ітерацію. Якщо так - перейти до кроку 3. Інакше - кінець.

Роль першого прошарку є умовною: скориставшись один раз на першому кроці 1 значеннями його вагових коефіцієнтів, мережа більше не вертається до нього, тому перший прошарок може бути взагалі виключений із мережі. Мережа Хеммінга має ряд переваг над мережею Хопфілда. Вона реалізує оптимальний класифікатор мінімуму похибки, якщо похибки вхідних бітів є випадковими та незалежними. Для функціонування мережі Хеммінга потрібна менша кількість нейронів, оскільки середній прошарок вимагає лише один нейрон на клас, замість нейрону на кожен вхідний вузол. І, нарешті, мережа Хеммінга не страждає від неправильних класифікацій, які можуть трапитись у мережі Хопфілда. В цілому, мережа Хеммінга є як швидшою, так і точнішою за мережу Хопфілда.

Лабораторна робота № 3: Реалізація системи розпізнавання на базі нейронної мережі асоціативного та рекурентно-асимптотичного типу.

Розробити систему розпізнавання на базі штучної нейронної мережі Хопфілда з наступними діями (за варіантом лабораторної роботи №1):

- введення зразкових зображень образів трьох класів та відображення у відповідних графічних компонентах;
- обчислення абсолютних ознакових векторів та відображення ознакових значень;
- для штучної нейронної мережі Хопфілда провести бінаризацію еталонних ознакових векторів зі значеннями -1 та 1; для мережі Хеммінга провести нормування ознакових еталонних векторів;
- обчислити матрицю коефіцієнтів ШНМ Хопфілда або вагові коефіцієнти першого шару зв'язків у мережі Хеммінга. Відобразити відповідні матриці;
- ввести зображення невідомого вектора, обчислити його ознакові вектори відповідно до типу мережі;
- провести класифікацію за методом Хопфілда або Хеммінга;
- проаналізувати процес класифікації.

Навчальна тема № 4: Однопрохідні багат шарові ШНМ. Алгоритми навчання. Алгоритм оберненого розповсюдження помилки.

Зміст теми:

1. Нейромережа зворотного поширення похибки (Back Propagation)
2. Реалізація алгоритмів навчання методом оберненого розповсюдження помилки. Оптимізаційний підхід до навчання ШНМ.
3. Оцінка функціональних можливостей.

Теоретичні відомості.

1. Нейромережа зворотного поширення похибки (Back Propagation)

Архітектура *feedforward backpropagation* була розроблена на початку 1970-х років декількома незалежними авторами: Вербор (*Werbos*); Паркер (*Parker*); Румельгарт (*Rumelhart*), Хінтон (*Hinton*) та Вільямс (*Williams*). На даний час, парадигма *backpropagation* найбільш популярна, ефективна та легка модель навчання для складних, багат шарових мереж. Вона використовується у різних типах застосувань і породила великий клас нейромереж з різними структурами та методами навчання.

Типова мережа *backpropagation* має вхідний прошарок, вихідний прошарок та принаймні один прихований прошарок. Теоретично, обмежень відносно числа прихованих прошарків не існує, але практично застосовують один або два (рис. 3).

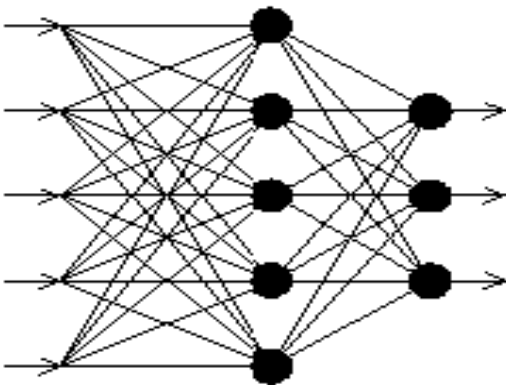


Рис. 3. Багат шаровий перцептрон.

Нейрони організовані в шарову структуру з прямою передачею сигналу. Кожний нейрон мережі продукує зважену суму своїх входів, пропускає цю величину через передатну функцію і видає вихідне значення. Мережа може моделювати функцію практично будь якої складності, причому число прошарків і число нейронів у кожному прошарку визначають складність

функції. Визначення числа проміжних прошарків і числа нейронів в них є важливим при моделюванні мережі. Більшість дослідників та інженерів, застосовуючи архітектуру до визначених проблем використовують загальні правила, зокрема:

1. Кількість входів та виходів мережі визначаються кількістю вхідних та вихідних параметрів досліджуваного об'єкту, явища, процесу, тощо. На відміну від зовнішніх прошарків, число нейронів прихованого прошарку $n_{\text{прих}}$ обирається емпіричним шляхом. В більшості випадків достатня кількість нейронів становить $n_{\text{прих}} = n_{\text{вх}} + n_{\text{вих}}$, де $n_{\text{вх}}$, $n_{\text{вих}}$ - кількість нейронів у вхідному і, відповідно, у вихідному прошарках.
2. Якщо складність у відношенні між отриманими та бажаними даними на виході збільшується, кількість нейронів прихованого прошарку повинна також збільшитись.
3. Якщо процес, що моделюється, може розділятися на багато етапів, потрібен додатковий прихований прошарок (прошарки). Якщо процес не розділяється на етапи, тоді додаткові прошарки можуть допустити перезапам'ятовування і, відповідно, невірне загальне рішення.

Після того, як визначено число прошарків і число нейронів в кожному з них, потрібно знайти значення для синаптичних ваг і порогів мережі, які спроможні мінімізувати похибку спродукованого результату. Саме для цього існують алгоритми навчання, де відбувається підгонка моделі мережі до наявних навчальних даних. Похибка для конкретної моделі мережі визначається шляхом проходження через мережу всіх навчальних прикладів і порівняння спродукованих вихідних значень з бажаними значеннями. Множина похибок створює функцію похибок, значення якої можна розглядати, як похибку мережі. В якості функції похибок найчастіше використовують суму квадратів похибок. Для кращого розуміння алгоритму навчання мережі *Back Propagation* потрібно роз'яснити поняття поверхні станів. Кожному значенню синаптичних ваг і порогів мережі (вільних параметрів моделі кількістю N) відповідає один вимір в багатовимірному просторі. $N+1$ -ий вимір відповідає похибці мережі. Для різноманітних сполучень ваг відповідну похибку мережі можна зобразити точкою в $N+1$ -вимірному просторі, всі ці точки утворюють деяку поверхню - поверхню станів. Мета навчання нейромережі полягає в знаходженні на багатовимірній поверхні найнижчої точки.

Поверхня станів має складну будову і досить неприємні властивості, зокрема, наявність локальних мінімумів (точки, найнижчі в своєму певному околі, але вищі від глобального мінімуму), пласкі ділянки, сідлові точки і довгі вузькі яри. Аналітичними засобами неможливо визначити розташування глобального мінімуму на поверхні станів, тому навчання нейромережі по суті полягає в дослідженні цієї поверхні. Відштовхуючись від початкової конфігурації ваг і порогів (від випадково обраної точки на поверхні), алгоритм навчання поступово відшукує глобальний мінімум. Обчислюється вектор градієнту поверхні похибок, який вказує напрямок найкоротшого спуску по поверхні з заданої точки. Якщо трошки просунутись по ньому, похибка

зменшиться. Зрештою алгоритм зупиняється в нижній точці, що може виявитись лише локальним мінімумом (в ідеальному випадку - глобальним мінімумом). Складність тут полягає у виборі довжини кроків. При великій довжині кроку збіжність буде швидшою, але є небезпека перестрибнути рішення, або піти в неправильному напрямку. При маленькому кроці, правильний напрямок буде виявлений, але зростає кількість ітерацій. На практиці розмір кроку береться пропорційним крутизни схилу з деякою константою - швидкістю навчання. Правильний вибір швидкості навчання залежить від конкретної задачі і здійснюється дослідним шляхом. Ця константа може також залежати від часу, зменшуючись по мірі просування алгоритму.

Алгоритм діє ітеративно, його кроки називаються епохами. На кожній епосі на вхід мережі по черзі подаються всі навчальні приклади, вихідні значення мережі порівнюються з бажаними значеннями і обчислюється похибка. Значення похибки, а також градієнту поверхні станів використовують для корекції ваг, і дії повторюються. Процес навчання припиняється або коли пройдена визначена кількість епох, або коли похибка досягає визначеного рівня малості, або коли похибка перестає зменшуватись (користувач переважно сам вибирає потрібний критерій останову).

2. Реалізація алгоритмів навчання методом оберненого розповсюдження помилки. Оптимізаційний підхід до навчання ШНМ.

За методом найменших квадратів, функція мінімізації загальної помилки мережі подається формулою:

$$E(w) = \frac{1}{2} \sum_{j,p} (y_{j,p}^{(N)} - d_{j,p})^2, \quad (1)$$

де $y_{j,p}^{(N)}$ - реальні вихідні значення j -го нейрона останнього N -го шару нейронної мережі при аналізі p -го образу навчальної послідовності;

$d_{j,p}$ - вірне (бажане) значення вказаного нейрона.

Сумування ведеться по всім нейронам вихідного шару для всіх зразків навчальної послідовності. Зміна вагових значень нейронів проводиться по формулі:

$$\Delta w_{i,j}^{(n)} = -\eta \cdot \frac{\partial E}{\partial w_{i,j}}, \quad (2)$$

де $w_{i,j}$ - вагове значення синаптичного зв'язку, що з'єднує i -ий нейрон шару $n-1$ з j -им нейроном шару n ;

η - крок навчання, вибирається на інтервалі $0 < \eta < 1$.

Частинні похідні обчислюються за формулою:

$$\frac{\partial E}{\partial w_{i,j}} = \frac{\partial E}{\partial y_j} \cdot \frac{dy_j}{ds_j} \cdot \frac{\partial s_j}{\partial w_{i,j}} \quad (3)$$

Тут y_j - вихід j -го нейрона, а s_j - виважена сума його вхідних сигналів.

Для сигмоїдальної функції активації нейрона її похідна здійснюється за формулою:

$$\frac{dy}{ds} = \frac{-b \cdot e^{-bs}}{(1 + e^{-bs})^2} \quad (4)$$

а для обчислення першого множника у формулі (3) використовується вираз

$$\frac{\partial E}{\partial y_j} = \sum_k \frac{\partial E}{\partial y_k} \cdot \frac{dy_k}{ds_k} \cdot w_{j,k}^{(n+1)} \quad (5)$$

В виразі (5) сумування проводиться по ваговим коефіцієнтам наступного, $n+1$ -го шару.

Введенням нової змінної

$$\delta_j^{(n)} = \frac{\partial E}{\partial y_j} \cdot \frac{dy_j}{ds_j} \quad (6)$$

Отримується рекурентна формула для обчислення величин $\delta_j^{(n)}$ шару n з відповідних величин $\delta_j^{(n+1)}$ старшого, наступного шару $n+1$:

$$\delta_j^{(n)} = \left[\sum_k \delta_k^{(n+1)} \cdot w_{j,k}^{(n+1)} \right] \cdot \frac{dy_j}{ds_j} \quad (7)$$

Формула (3.7) застосовується для всіх внутрішніх шарів, а для вихідного шару застосовується формула

$$\delta_j^{(n)} = (y_j^{(N)} - d_j) \cdot \frac{dy_j}{ds_j} \quad (8)$$

В результаті отримується рекурентне відношення для обчислення змін вагових синапсичних коефіцієнтів:

$$\Delta w_{i,j}^{(n)} = -\eta \cdot \delta_j^{(n)} \cdot y_i^{(n-1)} \quad (9)$$

Або (для досягнення інерційності):

$$\Delta w_{i,j}^{(n)}(t) = -\eta \cdot (\mu \cdot \Delta w_{i,j}^{(n)}(t-1) + (1-\mu) \delta_j^{(n)} \cdot y_i^{(n-1)}) \quad (10)$$

де μ - коефіцієнт інертності, $0 < \mu < 1$;

t - крок навчання.

За такою математичною моделлю, алгоритм зворотного розповсюдження похибки має вигляд:

1. Всім ваговим коефіцієнтам надаються малі початкові значення на інтервалі (0,1).

2. Створюється цикл для вибору чергових зразків з навчальної послідовності.
3. На вхід нейронної мережі подається чергових образ навчальної послідовності. Для всіх нейронів мережі обчислюються внутрішні стани за формулою

$$s_j^{(n)} = \sum_{i=0}^M y_i^{(n-1)} \cdot w_{i,j}^{(n)}, \quad (11)$$

де M - розмір вхідного вектору. Нульова компонента вхідного вектора одинична и визначає загальне зміщення суми.

$y_i^{(n-1)}$ - вихідні значення нейронів попереднього шару; для першого шару замінюються компонентами вхідного вектора.

4. Обчислюються вихідні значення всіх нейронів за формулою

$$y_j^{(n)} = f(s_j^{(n)}), \quad (12)$$

де $f(s)$ сигмоїдальна функція.

5. Для останнього шару нейронів вказиваються правильні (бажані) значення вихідних сигналів y_j , за формулою (8) обчислюються величини $\delta_j^{(N)}$.
6. За формулою (9) або (10) обчислюються зміни вагових коефіцієнтів вихідного шару $\Delta w_{i,j}^{(N)}$.
7. За формулою (7) и (9) або (7) и (10) обчислюються величини $\delta_j^{(n)}$ и зміни вагових коефіцієнтів всіх внутрішніх шарів послідовно від передостаннього до першого шару.
8. Вагові коефіцієнти всіх шарів перераховуються за формулою

$$w_{i,j}^{(n)}(t) = w_{i,j}^{(n)}(t-1) + \Delta w_{i,j}^{(n)}(t), \quad (13)$$

де t - крок навчання.

Процес навчання продовжується, поки функція сумарної похибки нейронної мережі не задовольнить наперед задані параметри точності та надійності.

Навчальна тема № 5: Нейронні мережі із самоорганізацією. Карти Кохонена.

Зміст теми:

1. Поняття самоорганізуючої нейронної мережі.
2. Геометрична інтерпретація.
3. Оцінка функціональних можливостей.

Теоретичні відомості.

1. Поняття самоорганізуючої нейронної мережі.

Нейронні мережі Кохонена - клас нейронних мереж, основним елементом яких є *шар Кохонена*. Шар Кохонена складається з адаптивних лінійних суматорів ("лінійних формальних нейронів"). Як правило, вихідні сигнали шару Кохонена обробляються за правилом "переможець забирає все": найбільший сигнал перетворюється в одиничний, решта звертаються в нуль.

За способами настройки входних ваг суматорів і по розв'язуванню завданням розрізняють багато різновидів мереж Кохонена. Найбільш відомі з них:

- Мережі векторного квантування сигналів, тісно пов'язані з найпростішим базовим алгоритмом кластерного аналізу (метод динамічних ядер або К-середніх)
- Карти самоорганізації Кохонена (Self-Organising Maps, SOM)
- Мережі векторного квантування, яких навчають з учителем (Learning Vector Quantization).

Шар Кохонена складається з деякої кількості n паралельно діючих лінійних елементів. Всі вони мають однакову кількість входів m і отримують на свої входи один і той же вектор входних сигналів $\mathbf{x} = (x_1, \dots, x_m)$. На виході j -го лінійного елемента отримуємо сигнал

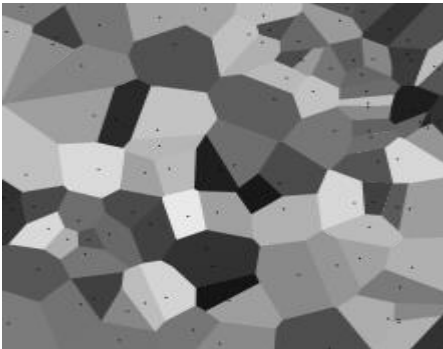
$$y_j = w_{j0} + \sum_{i=1}^m w_{ji}x_i,$$

де w_{ji} - ваговий коефіцієнт i -го входу j -го нейрона, w_{j0} - пороговий коефіцієнт.

Після проходження шару лінійних елементів сигнали посилаються на обробку за правилом "переможець забирає все": серед вихідних сигналів y_j шукається максимальний; його номер $j_{\max} = \arg \max_j \{y_j\}$. Остаточо, на виході сигнал з номером j_{\max} дорівнює одиниці, решта - нулю. Якщо максимум одночасно досягається для декількох j_{\max} , то або приймають всі відповідні сигнали рівними одиниці, або тільки перший у списку (за згодою). "Нейрони Кохонена

можна сприймати як набір електричних лампочок, так що для будь-якого вхідного вектора загоряється одна з них."

2. Геометрична інтерпретація.



Розбиття площині на багатокутники Вороного-Діріхле для випадково вибраних точок (кожна точка вказана в своєму багатокутнику).

Великого поширення набули шари Кохонена, побудовані таким чином: кожному (j -му) нейрону зіставляється точка $W_j = (w_{j1}, \dots, w_{jm})$ в m -мірному просторі (просторі сигналів). Для вхідного вектора $x = (x_1, \dots, x_m)$ обчислюються його евклідові відстані $\rho_j(x)$ до точок W_j "найближчий отримує все" - той нейрон, для якого ця відстань мінімальна, видає одиницю, решта - нулі. Слід зауважити, що для порівняння відстаней достатньо обчислювати лінійну функцію сигналу:

$$\rho_j(x)^2 = \|x - W_j\|^2 = \|W_j\|^2 - 2 \sum_{i=1}^m w_{ji} x_i + \|x\|^2$$

(Тут $\|y\|$ - Евклідова довжина вектора: $\|y\|^2 = \sum_i y_i^2$). Останній доданок $\|x\|^2$ однаково для всіх нейронів, тому для знаходження найближчої точки воно не потрібно. Задача зводиться до пошуку номера найбільшого з значень лінійних функцій:

$$j_{\max} = \arg \max_j \left\{ \sum_{i=1}^m w_{ji} x_i - \frac{1}{2} \|W_j\|^2 \right\}.$$

Таким чином, координати точки $W_j = (w_{j1}, \dots, w_{jm})$ збігаються з вагами лінійного нейрона шару Кохонена (при цьому значення порогового коефіцієнта $w_{j0} = -\|W_j\|^2/2$).

Якщо задані точки $W_j = (w_{j1}, \dots, w_{jm})$, то m -вимірний простір розбивається на відповідні многогранники Вороного-Діріхле V_j : Багатогранник V_j складається з точок, які ближче до W_j , ніж до інших W_k ($k \neq j$).

3. Принцип роботи та алгоритм навчання.

Завдання векторного квантування полягає, по своїй суті, в найкращій апроксимації всієї сукупності векторів даних k кодовими векторами W_j .

Самоорганізовані карти Кохонена також апроксимують дані, проте за наявності додаткової структури в сукупності кодових векторів (англ. codebook). Передбачається, що апіорі задана деяка симетрична таблиця "мір сусідства" (або "мір близькості") вузлів: для кожної пари j, l ($j, l = 1, \dots, k$) визначено число η_{jl} ($0 \leq \eta_{jl} \leq 1$). При цьому діагональні елементи таблиці близькості дорівнюють одиниці ($\eta_{jj} = 1$).

Вектори вхідних сигналів x обробляються по одному, для кожного з них знаходиться найближчий кодовий вектор ("переможець", який "забирає все") $W_{j(x)}$. Після цього всі кодові вектори W_l , для яких $\eta_{j(x)l} \neq 0$, перераховуються за формулою

$$W_l^{\text{new}} = W_l^{\text{old}}(1 - \eta_{j(x)l}\theta) + x\eta_{j(x)l}\theta,$$

де $\theta \in (0, 1)$ - крок навчання. Сусіди кодового вектора - переможця (за апіорно заданою таблицею близькості) зсуваються в ту ж сторону, що і цей вектор, пропорційно до міри близькості.

Найчастіше, таблиця кодових векторів представляється у вигляді фрагмента квадратної решітки на площині, а міра близькості визначається, виходячи з евклідової відстані на площині.

Самоорганізовані карти Кохонена служать, в першу чергу, для візуалізації та первісного ("розвідувального") аналізу даних. Кожна точка даних відображається відповідним кодовим вектором з решітки. Так отримують уявлення даних на площині ("*карту даних*"). На цій карті можливе відображення багатьох шарів: кількість даних, що потрапляють у вузли (тобто "щільність даних"), різні функції даних тощо. При відображенні цих шарів корисний апарат географічних інформаційних систем (ГІС). У ГІС базою для зображення інформаційних шарів служить географічна карта. Карта даних є базою для довільного за своєю природою набору даних. Вона служить заміною географічній карті там, де її просто не існує. Принципова відмінність в наступному: на географічній карті сусідні об'єкти володіють близькими географічними координатами, на карті даних близькі об'єкти володіють близькими властивостями. За допомогою карти даних можна візуалізувати дані, одночасно наносячи на базу супроводжуючу інформацію (підписи, анотації, атрибути, інформаційні розмальовки). Карта служать також інформаційною моделлю даних. З її допомогою можна заповнювати прогалини в даних. Ця здатність використовується, наприклад, для вирішення завдань прогнозування.

Навчальна тема № 5: Специфічні нейронні мережі для розпізнавання образів. Когнітрон.

Зміст теми:

1. Когнітрон. Структура та загальна ідея роботи.
2. Навчання.
3. Оцінка функціональних можливостей.

Теоретичні відомості.

1. Когнітрон. Структура та загальна ідея роботи.

На мікроскопічному рівні виявлено, що нейрони мають як збуджуючі так і гальмуючі синапси. Перші прагнуть до збудження нейрона; останні гальмують його збудження (див. додаток А). Це наводить на думку, що мозок пристосовується або зміною впливу цих синапсів, або створенням або руйнуванням синапсів внаслідок впливу навколишнього середовища. Дане припущення залишається поки гіпотезою з обмеженим фізіологічним підтвердженням. Однак дослідження, проведені в рамках цієї гіпотези, привели до створення цифрових моделей, деякі з яких показують чудові здібності до адаптивного розпізнавання образів.

Засновуючись на поточних знаннях анатомії і фізіології мозку розроблений когнітрон, гіпотетична модель системи сприйняття людини. Комп'ютерні моделі типу когнітрон, продемонстрували вражаючі здібності адаптивного розпізнавання образів, що спонукає фізіологів дослідити відповідні механізми мозку. Ця взаємно посилююча взаємодія між штучними нейронними мережами, фізіологією і психологією може виявитися засобом, за допомогою якого буде згодом досягнуто розуміння механізмів мозку.

Когнітрон конструюється у вигляді прошарків нейронів, сполучених синапсами. Як показано на рис. 10.1, пресинаптичний нейрон в одному прошарку пов'язаний з постсинаптичним нейроном в наступному прошарку. Є два типи нейронів: збуджуючі вузли, які прагнуть спричинити збудження постсинаптичного вузла, і гальмуючі вузли, які гальмують це збудження. Збудження нейрона визначається зваженою сумою його збуджуючих і гальмуючих входів, однак насправді механізм є більш складним, ніж просте підсумовування.

На рис. 10.2 показано, що кожний нейрон пов'язаний тільки з нейронами в сусідній області, званій *областю зв'язку*. Це обмеження області зв'язку узгоджується з анатомією зорової кори, в якій рідко сполучаються між собою нейрони, розташовані один від одного на відстані більше за один міліметр. У моделі, що розглядається нейрони впорядковані у вигляді прошарків зі

зв'язками від одного прошарку до наступного. Це також аналогічно пошаровій структурі зорової кори і інших частин головного мозку.

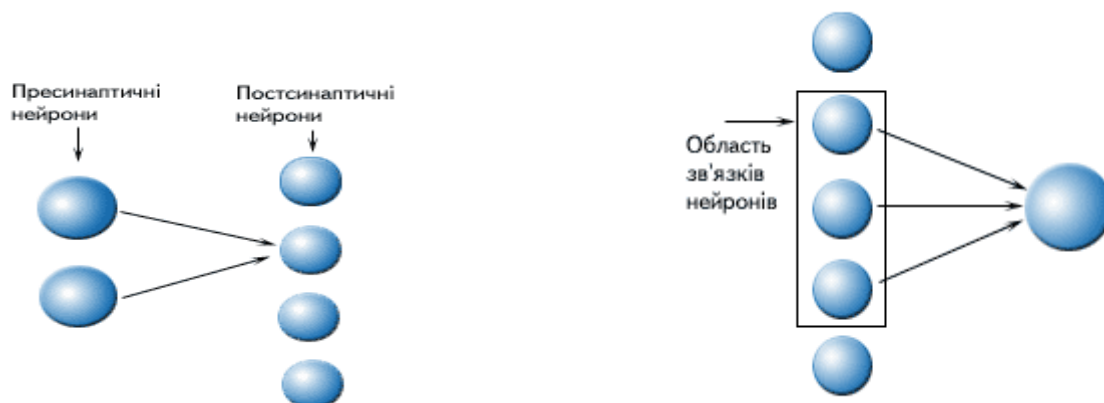


Рис. 10.1, 2. Предсинаптичні і постсинаптичні нейрони та область зв'язків

2. Навчання

Оскільки когнітрон реалізований у вигляді багатошарової мережі, виникають складні проблеми навчання, пов'язані з вибраною структурою. Автор відкинув кероване навчання, як біологічно неправдоподібне, використовуючи замість цього навчання без вчителя. Отримуючи навчальний набір вхідних образів, мережа самоорганізується за допомогою зміни сили синаптичних зв'язків. При цьому відсутні заздалегідь певні вихідні образи, що представляють необхідну реакцію мережі, однак мережа самоналаштується з метою розпізнавання вхідних образів з задовільною точністю.

Алгоритм навчання когнітрону є концептуально привабливим. У заданій області прошарку навчається тільки найбільш сильно збуджений нейрон. Автор порівнює це з "елітним навчанням", при якому навчаються тільки "розумні" елементи. Ті нейрони, які вже добре навчені, що виражається силою їх збудження, отримують приріст сили своїх синапсів з метою подальшого підсилення свого збудження.

На рис. 10.3 показано, що області зв'язку сусідніх вузлів значно перекриваються. Це марнотратне дублювання функцій виправдовується взаємною конкуренцією між найближчими вузлами. Навіть якщо вузли в початковий момент мають абсолютно ідентичний вихід, невеликі відхилення завжди мають місце; один з вузлів завжди буде мати більш сильну реакцію на вхідний образ, ніж сусідні. Його сильне збудження буде надавати стримуючий вплив на збудження сусідніх вузлів, і тільки його синапси будуть посилюватися; синапси сусідніх вузлів залишаються незмінними.

Збуджуючий нейрон. Грубо кажучи, вихід збуджуючого нейрона в когнітроні визначається відношенням його збуджуючих входів до гальмуючих

входів. Ця незвичайна функція має важливі переваги, як практичні, так і теоретичні.

Сумарний збуджуючий вхід в нейрон зваженою сумою входів від збуджуючих з попереднього прошарку. Аналогічно сумарний вхід / є зваженою сумою входів від всіх гальмуючих нейронів. У символічному вигляді

$E = \sum_i a_i u_i$, $I = \sum_j b_j v_j$, де a_i - вага i -го збуджуючого синапсу, u_i - вихід i -го збуджуючого нейрона, b_j - вага j -го гальмуючого синапсу, v_j - вихід j -го гальмуючого нейрона.

Помітимо, що ваги мають тільки позитивні значення. Вихід нейрона обчислюється таким чином:

$$NET = \frac{1 + E}{1 + I} - 1 \quad OUT = NET, \text{ при } NET \gg 0, \quad OUT = 0, \text{ при } NET < 0.$$

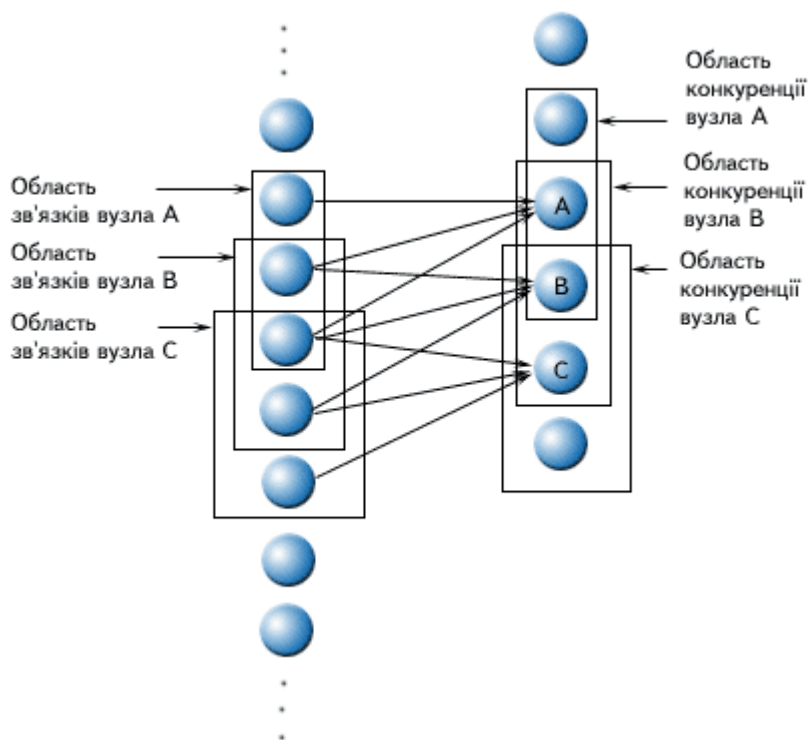


Рис. 10.3. Область зв'язку з областю конкуренції

Передбачаючи, що NET має позитивне значення, це можна записати таким

чином:

$$OUT = \frac{E - I}{1 + I}.$$

Коли гальмуючий вхід малий ($I \ll 1$), OUT можна апроксимувати як

$OUT = E - I$, що відповідає виразу для звичайного лінійного порогового елемента (з нульовим порогом).

Алгоритм навчання когнітрону дозволяє вагам синапсів зростати без обмежень. Завдяки відсутності механізму зменшення ваг вони просто зростають в процесі навчання. У звичайних лінійних порогових елементах це привело б до довільно великого виходу елемента. У когнітроні великі збуджуючі і гальмуючі входи результуються в обмежуючій формулі вигляду:

$$OUT = \frac{E}{I} - 1, \quad \text{якщо } E \gg 1 \text{ і } I \gg 1.$$

У цьому випадку OUT визначається відношенням збуджуючих входів до гальмуючих входів, а не їх різницею. Таким чином, величина OUT обмежується, якщо обидва входи зростають в одному і тому ж діапазоні X . Вважаючи, що це так, E і I можна виразити таким чином:

$E = pX, I = qX, p, q$ - константи, і після деяких перетворень

$$OUT = \frac{p - q}{2q} \cdot \left[1 + \operatorname{th} \left(\frac{\log(pq)}{2} \right) \right]$$

Ця функція зростає згідно із законом Вебера-Фехнера, який часто використовується в нейрофізіології для апроксимації нелінійних співвідношень входу/виходу сенсорних нейронів. При використанні цього співвідношення нейрон когнітрону в точності емулює реакцію біологічних нейронів. Це робить його як могутнім обчислювальним елементом, так і точною моделлю для фізіологічного моделювання.

Гальмуючі нейрони. У когнітроні прошарок складається із збуджуючих і гальмуючих вузлів. Як показано на рис. 10.4, нейрон прошарку 2 має область зв'язку, для якої він має синаптичні з'єднання з набором виходів нейронів в прошарку 1. Аналогічно в прошарку 1 існує гальмуючий нейрон, що має ту ж область зв'язку. Синаптичні ваги гальмуючих вузлів не змінюються в процесі навчання; їх ваги заздалегідь встановлені таким чином, що сума ваги в будь-якому з гальмуючих нейронів рівна одиниці. Відповідно до цих обмежень, вихід гальмуючого вузла INHIB є зваженою сумою його входів, які в цьому випадку являють собою середнє арифметичне виходів збуджуючих нейронів, до яких він приєднаний. Таким чином,

$$INHIB = \sum_i c_i OUT_i \quad \text{де } \sum_i c_i = 1, \quad c_i - \text{збуджуюча вага } i.$$

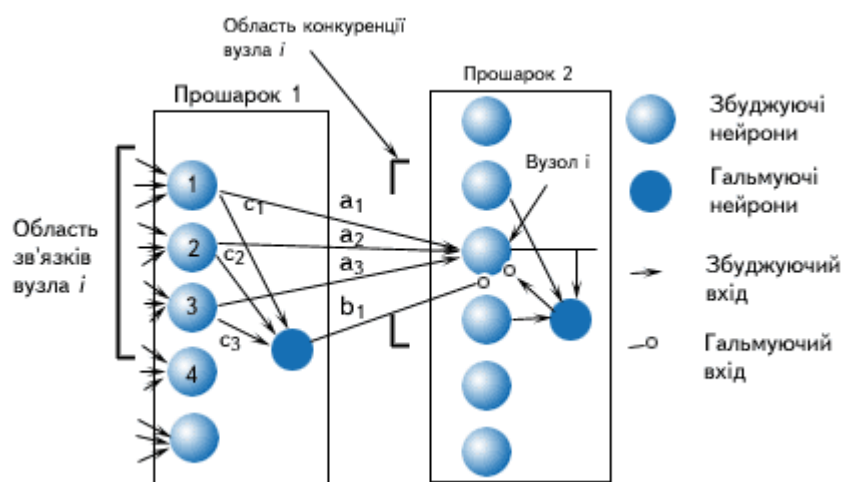


Рис. 10.4. Прошарки когнітрону

Процедура навчання. Як пояснювалося раніше, ваги збуджуючих нейронів змінюються тільки тоді, коли нейрон збуджений сильніше, ніж будь-хто з вузлів в області конкуренції. Якщо це так, зміна в процесі навчання будь-якого з його ваги може бути визначена таким чином:

$$\delta a_i = q c_j u_j,$$

де c_j - гальмуюча вага зв'язку нейрона j в прошарку 1 з гальмуючим нейроном i , u_j - вихід нейрона j в прошарку 1, a_i - збуджуюча вага i , q - нормуючий коефіцієнт навчання.

Зміна гальмуючої ваги нейрона i в прошарку 2 пропорційна відношенню зваженої суми збуджуючих входів до подвоєного гальмуючого входу. Обчислення проводяться по формулі

$$\delta b_i = \frac{q \sum_j a_j u_j}{2 \cdot \text{INHIB}_i}$$

Коли збуджених нейронів в області конкуренції немає, для зміни ваги використовуються інші вирази. Це необхідне, оскільки процес навчання починається з нульових значень ваги; тому спочатку немає збуджених нейронів ні в одній області конкуренції, і навчання проводиться не може. У всіх випадках, коли переможця в області конкуренції нейронів немає, зміна ваги нейронів обчислюється таким чином:

$$\Delta a_i = q' c_j u_j, \delta b_i = q' \text{INHIB}_i,$$

де q' - позитивний навчальний коефіцієнт менший, ніж q .

Приведена стратегія налаштування гарантує, що вузли з великою реакцією примушують збуджуючі синапси, якими вони керують, збільшуватися сильніше, ніж гальмуючі синапси. І навпаки, вузли, що мають малу реакцію, спричиняють мале зростання збуджуючих синапсів, але більший зріст гальмуючих синапсів. Таким чином, якщо вузол 1 в прошарку 1 має більший вихід, синапс a_1 зросте більше, ніж синапс b_1 . І навпаки, вузли, що мають малий вихід, забезпечать малу величину для приросту a_i . Однак інші вузли в області зв'язку будуть збуджуватися, тим самим збільшуючи сигнал INHIB і значення b_i .

У процесі навчання ваги кожного вузла в прошарку 2 налаштовуються таким чином, що разом вони складають шаблон, відповідний образам, які часто пред'являються в процесі навчання. При пред'явленні схожого образу шаблон відповідає йому і вузол виробляє великий вихідний сигнал. Сильно відмінний образ виробляє малий вихід і звичайно гальмується конкуренцією.

Латеральне гальмування. На рис. 10.4 показано, що кожний нейрон прошарку 2 отримує латеральне гальмування від нейронів, розташованих в його області конкуренції. Гальмуючий нейрон підсумовує входи від всіх нейронів в області конкуренції і виробляє сигнал, прагнучий до гальмування цільового нейрона. Цей метод є ефективним, але з обчислювальної точки зору повільним. Він охоплює велику систему із зворотним зв'язком, що включає кожний нейрон в прошарку; для його стабілізації може бути потрібна велика кількість обчислювальних ітерацій.

Для прискорення обчислень в роботі [2] використовується метод прискореного латерального гальмування (рис. 10.5). Тут додатковий вузол латерального гальмування обробляє вихід кожного збуджуючого вузла для моделювання необхідного латерального гальмування. Спочатку він визначає сигнал, рівний сумарному гальмуючому впливу в області конкуренції:

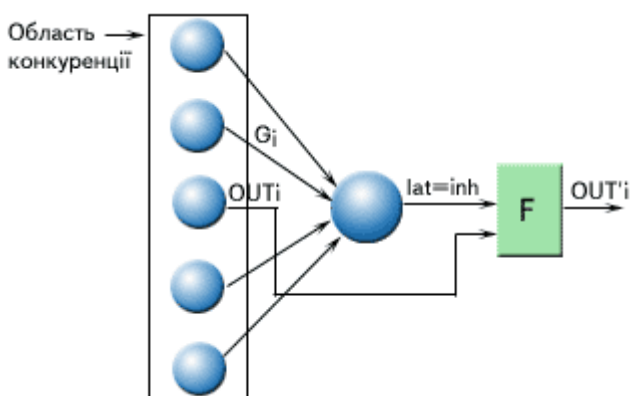


Рис. 10.5. Прискорене гальмування

$$LAT_INHIB = \sum_i g_i OUT_i$$

де OUT_i - вихід i -го нейрона в області конкуренції, g_i - вага зв'язку від цього нейрона до латерально-гальмуючого нейрона; g_i вибрані таким чином, що $\sum_i g_i = 1$.

Вихід гальмуючого нейрона OUT обчислюється таким чином:

$$OUT' = \frac{1 + OUT_i}{1 + LAT_INHIB} - 1$$

Завдяки тому що всі обчислення, пов'язані з таким типом латерального гальмування, є нерекурсивними, вони можуть бути проведені за один прохід для прошарку, тим самим визначаючи ефект у вигляді великої економії в обчисленнях.

Цей метод латерального гальмування вирішує і іншу складну проблему. Передбачимо, що вузол в прошарку 2 збуджується сильно, але збудження сусідніх вузлів меншає поступово із збільшенням відстані. При використанні звичайного латерального гальмування буде навчатися тільки центральний вузол. Інші вузли визначають, що центральний вузол в їх області конкуренції має більш високий вихід. З системою латерального гальмування такої ситуації, що пропонується трапитися не може. Множина вузлів може навчатися одночасно і процес навчання є більш достовірним.

Рецептивна область. Аналіз, що проводиться до цього моменту, був спрощений розглядом тільки одновимірних прошарків. Насправді Когнітрон конструювався як каскад двовимірних прошарків, причому в даному прошарку кожний нейрон отримує входи від набору нейронів на частині двовимірного плану, що складає його область зв'язку в попередньому прошарку.

З цієї точки зору Когнітрон організований подібно зоровій корі людини, що являє собою трьохмірну структуру, що складається з декількох різних прошарків. Виявляється, що кожний прошарок мозку реалізовує різні рівні узагальнення; вхідний прошарок чутливий до простих образів, таких, як лінії, і їх орієнтації в певних областях візуальної області, в той час як реакція інших прошарків є більш складної, абстрактної і незалежної від позиції образу.

Аналогічні функції реалізовані в когнітроні шляхом моделювання організації зорової кори. На рис. 10.6 показано, що нейрони когнітрону в прошарку 2 реагують на певну невелику область вхідного прошарку 1. Нейрон в прошарку 3 пов'язаний з набором нейронів прошарку 2, тим самим реагуючи непрямо на більш широкий набір нейронів прошарку 1. Подібним чином нейрони в подальших прошарках чутливі до більш широких областей вхідного образу доти, поки у вихідному прошарку кожний нейрон не стане реагувати на все вхідне поле.

Якщо область зв'язку нейронів має постійний розмір у всіх шарах, потрібна велика кількість прошарків для перекриття усього вхідного поля вихідними нейронами. Кількість прошарків може бути зменшена шляхом розширення області зв'язку в подальших шарах. На жаль, результатом цього може з'явитися настільки велике перекриття областей зв'язку, що нейрони вихідного прошарку будуть мати однакову реакцію. Для розв'язання цієї проблеми може бути використане розширення області конкуренції. Оскільки в даній області конкуренції може збудитися тільки один вузол, вплив малої різниці в реакціях нейронів вихідного прошарку посилюється.

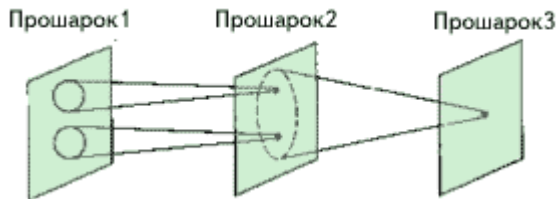


Рис. 10.6. Области зв'язків когнітрону

У альтернативному варіанті зв'язку з попереднім шаром можуть бути розподілені імовірісно з більшістю синаптичних зв'язків в обмеженій області і з більш довгими з'єднаннями, що зустрічаються набагато рідше. Це відображає імовірнісний розподіл нейронів, виявлений в мозку. У когнітроні це дозволяє кожному нейрону вихідного прошарку реагувати на повне вхідне поле при наявності обмеженої кількості прошарків.

Рекомендована література

Базова

1. Нікольський Ю.В., Пасічник В.В., Щербина Ю.М. Системи штучного інтелекту : навчальний посібник. Київ : Магнолія, 2021. 280 с.
2. Гончаров О.А., Васильєва Л.В., Юнда А.М. Чисельні методи розв'язання прикладних задач : навч. посібн. Суми : Сумський державний університет, 2020, 142 с.
3. Шаховська Н.Б., Камінський Р.М., Вовк О.Б. Системи штучного інтелекту : навчальний посібник. Львів : Львівська політехніка. 2018. 392 с.
4. Троцько В.В. Методи штучного інтелекту: навчально-методичний і практичний посібник / В.В. Троцько. К. : Університет "КРОК", 2020. 86 с.
5. Фратавчан В.Г., Фратавчан Т.М. Методи та системи штучного інтелекту: методичні вказівки та завдання для лабораторних робіт. Чернівці: ЧНУ, 2022. 36 с.
6. Савченко А.С., Синельніков О.О. Методи та системи штучного інтелекту: навчальний посібник. Київ : Національний авіаційний інститут. 2017. 188 с.

Допоміжна

1. Зайченко Ю.П. Основи проектування інтелектуальних систем. Навчальний посібник. — Київ, Видавничий Дім "Слово", 2004 р. - 352 с.
2. Wasserman, Philip D, Neural Computing: Theory and Practice, First Edition, Hygiene, Colorado, 1989, 230 p.

Інформаційні ресурси

1. Учбовий курс "Системи штучного інтелекту"
<https://web.archive.org/web/20160907035551/http://victoria.lviv.ua/html/ai/>
2. Earl Hunt, Artificial Intelligence
https://books.google.com.ua/books?hl=uk&lr=&id=9y2jBQAAQBAJ&oi=fnd&pg=PP1&dq=Hunt+Artificial+intelligence&ots=uPXxXi9XqB&sig=0kXUTxTZGJHsiWRRlyBPnG6R59E&redir_esc=y#v=onepage&q=Hunt%20Artificial%20intelligence&f=false
3. Книга Ф. Уосермена «Нейрокомп'ютерна техніка: Теорія і практика»
<https://web.archive.org/web/20090531081836/http://www.victoria.lviv.ua/html/wosserman/index.htm>