ВСТУП

Уже ні в кого не викликає здивування проникнення комп'ютерів практично в усі сфери людської діяльності. Удосконалювання елементної бази, що визначає архітектуру комп'ютера, і розпаралелювання обчислень дозволяють швидко й ефективно вирішувати задачі все зростаючої складності. Вирішення багатьох проблем немислиме без застосування комп'ютерів. Однак, маючи величезну швидкодію, комп'ютер часто не в змозі впоратися з поставленим перед ним завданням так, як це робить людина. Прикладами подібних завдань є розпізнавання (наприклад, знайоме обличчя людина впізнає за 100-120 мс, а найсучасніший комп'ютер — за хвилини або години), розуміння мови й тексту, написаного від руки тощо. Таким чином, мережа нейронів, що створює мозок людини, будучи, як і комп'ютерна мережа, системою паралельної обробки інформації, у багатьох випадках є більш ефективною. Ідея переходу від обробки закладеним у комп'ютер алгоритмом деяких формалізованих знань до реалізації в ньому властивих людині прийомів обробки інформації (розумової діяльності) призвели до появи штучних нейронних мереж (ШНМ).

Характерною рисою біологічних систем є адаптація, завдяки якій такі системи в процесі навчання розвиваються й здобувають нові властивості. Як і біологічні нейронні мережі, ШНМ складаються із з'єднаних між собою елементів, штучних нейронів, функціональні можливості яких тією чи іншою мірою відповідають елементарним функціям біологічного нейрона. Як і біологічний прототип, ШНМ має такі властивості:

— *адаптивність* – здатність до адаптації (адаптація – сукупність пристосувальних, реакцій на зміну зовнішнього середовища, що протікають на різних ієрархічних рівнях, та організації сигналів зовнішнього середовища, що впливають на організм. НМ є адаптивними системами, оскільки можуть

пристосовуватися до змін зовнішнього середовища за допомогою зміни своєї структури і значень параметрів);

— *здатність до навчання*: ШНМ здатні удосконалювати свою роботу (навчатися шляхом адаптації), використовуючи приклади для налаштування на вирішення певної задачі;

— *узагальнення* — здатність інтегрувати окремі дані для визначення закономірностей та пролонгації результатів, що дозволяє після навчання на одних даних застосовувати отримані знання для інших даних. Для ШНМ під властивістю узагальнення розуміється здатність генерувати правильні виходи для вхідних сигналів, які не були враховані в процесі навчання. Однак ШНМ не завжди може узагальнювати дані. Якщо НМ була перетренована, тобто вона запам'ятала тренувальні дані як таблицю і видає коректні дані лише при точній вказівці «адреси» в ній, то в цьому випадку мережа не зможе коректно інтерполювати вхідні дані між тренувальними. При низькому рівні узагальнення перетворення, здійснюване ШНМ, недостатньо гладко. Гладкість перетворення в НМ безпосередньо пов'язана з необхідністю вибрати найпростішу модель в умовах відсутності апріорної інформації. У даному випадку під найпростішою мається на увазі найбільш гладка функція, що найкращим чином апроксимує перетворення, яке розглядається. Вимога гладкості забезпечує гарні інтерполяційні властивості ШНМ і гарантує їм мінімальну складність структури, що прямо впливає на обсяг обчислень, які виконуються при навчанні. Іншою вимогою є достатня репрезентативність тренувального набору (навчальної вибірки) даних: ШНМ тим краще узагальнює, чим щільніше і рівномірніше розташовані тренувальні дані у вхідному просторі. Якщо тестові дані завжди будуть надаватися між близько розташованими тренувальними шаблонами, то мережа зможе проводити коректне узагальнення, інтерполюючи вхідні дані;

— *самоорганізація* являє собою процес динамічної перебудови системи з метою адаптації до зовнішнього середовища. Самоорганізація на рівні організму відбувається за допомогою навчання, в результаті якого динамічно перебудовуються нейронні структури головного мозку. Самоорганізація ШНМ називається також здатністю до навчання:НМ можуть автономно «вивчати» статичні і динамічні властивості модельованого об'єкта на основі результатів вимірювань, які проводилися

в минулому, а потім діяти таким чином, щоб прийняти найкраще рішення при невідомому стані зовнішнього середовища. Звичайні комп'ютери повинні бути попередньо запрограмовані, щоб мати можливість обробляти дані, вони не можуть працювати за межами рішень, що задаються програмою. Таким чином, інженерія знань не може бути повною мірою реалізована на звичайних комп'ютерах, так як вони не можуть приймати рішення в нових зовнішніх умовах;

— *обчислення в реальному часі*: нейромережеві обчислення можуть здійснюватися паралельно в часі, що істотно збільшує швидкодію ШНМ;
— *стійкість до перебоїв*: часткове руйнування мережі призводить до втрати якості, однак деякі її властивості зберігаються навіть у випадку руйнування більшої частини мережі.

Незважаючи на таку подібність, ШНМ ще дуже далекі від дублювання властивостей мозку людини. Однак дивна подібність функціонування деяких ШНМ з їхніми біологічними прототипами наводить на думку про можливості проникнення в людський інтелект уже в близькому майбутньому.

Переваги ШНМ і нейрокомп'ютерів полягають у тому, що вони дають стандартний спосіб рішення багатьох нестандартних завдань (неважливо, що спеціалізована машина краще вирішить один клас задач, важливіше, що один нейрокомп'ютер вирішить і цю задачу, і другу, і третю — і не треба кожен раз проектувати спеціалізовану ЕОМ — нейрокомп'ютер зробить все сам і майже не гірше); замість програмування використовується навчання (нейрокомп'ютер вчиться — потрібно лише формувати навчальні задачники, праця програміста, розпорядчого машині всі деталі роботи, заміщається працею вчителя, що створює «освітнє середовище», до якого пристосовується нейрокомп'ютер); нейрокомп'ютери особливо ефективні там, де необхідно подобу людської інтуїції і важко створити явний метод.

Основними ***властивостями нейромереж*** є [1]:

— *однорідність* нейроелементів і базових операцій, а також технологічна простота різних способів їх фізичної реалізації: НМ за аналогією з мозком будуються з множини простих уніфікованих типових елементів (нейронів), що виконують елементарні дії (множення, додавання, обчислення найпростішої нелінійної функції) і з'єднаних між собою різними зв'язками;
— *можливість реалізації нелінійних відображень* шляхом використання нелінійних функцій активації нейронів (це важ-

ливо для вирішення завдань управління з істотними нелінійностями, для яких традиційні підходи поки не дають практично реалізованих рішень);

— *пластичність* — обумовлює складність поведінки ШНМ, яка розглядається як результат взаємодії багатьох елементів, кожен з яких обмежує дію інших і сам обмежується іншими на шляху до формування глобальної спостережуваною поведінки. Розрізняють *нейронну пластичність* (як пластичні елементи розглядаються нейрони), а також *синаптичну пластичність* (модифікація сили синаптичного зв'язку між нейронами). Оскільки кількість синапсів, як правило, на кілька порядків перевищує кількість нейронів, то мережу з пластичними синапсами буде володіти більшими можливостями, ніж мережа еквівалентного розміру з пластичними нейронами;

— *універсальна апроксимація*: ШНМ здійснюють наближення функцій багатьох змінних за допомогою лінійних операцій і суперпозицій функцій однієї змінної. Радянські вчені О. М. Колмогоров і В. І. Арнольд показали, що будь-яку неперервну функцію N змінних можна отримати за допомогою операцій додавання, множення і суперпозиції з неперервних функцій однієї змінної. На основі цього доведено ряд теорем про апроксимації неперервних функцій багатьох змінних ШНМ з використанням довільної неперервної функції однієї змінної. ШНМ дозволяють з будь-якою точністю обчислювати довільну неперервну функцію. Отже, з їх допомогою можна як завгодно точно апроксимувати функціонування будь-якої неперервної системи. У 1987 році Р. Хехт-Нільсеном була запропонована теорема, яка доводить зображуваність функції багатьох змінних досить загального вигляду за допомогою ШНМ з прямими повними зв'язками, N нейронами вхідного шару, $(2N+1)$ нейронами прихованого шару з наперед відомими обмеженими функціями активації (наприклад, сигмоїдальними) і NM нейронами вихідного шару з невідомими функціями активації. З теореми Хехт-Нільсена випливає зображуваність будь-якої багатовимірної функції декількох змінних за допомогою НМ фіксованого розмірності. Таким чином, ШНМ є універсальними структурами, що дозволяють реалізувати будь-який обчислювальний метод;

— *асоціативна пам'ять*: ШНМ розглядається як мультистабільна система, стійкий стан якої (атрактор) відповідає за-

пам'ятованим станам. Якщо зовнішній вплив встановлює мережу у стан, близький до одного з атракторів, то, будучи надана сама собі, мережа за одну або декілька ітерацій конвергенції перейде у стан атрактора. Таким чином, якщо початковий стан відображає неповний або спотворений набір раніше запам'ятованих даних, то після конвергенції весь цей набір даних буде відновлений. На цьому базується властивість асоціативної пам'яті, що дозволяє застосовувати нейрокомп'ютери для розпізнавання образів, прогнозування складних процесів і т. п. Обсяг асоціативної пам'яті НМ залежить від кількості стійких станів, положення яких визначається значенням матриці зв'язків, отриманих в результаті навчання. Число таких станів не може бути як завгодно великим. Якщо число запам'ятованих станів перевищує деяку межу, відбувається насичення пам'яті і з'являються неправдиві стійкі стани, що спотворюють вміст асоціативної пам'яті;

— *розподіленість пам'яті* — інформація зберігається за багатьма адресами, розподілених таким чином, що кожен елемент даних представляється шаблоном активності, розподіленим по багатьом обчислювальним елементам, і кожен обчислювальний елемент бере участь в уявленні багатьох різних елементів даних. У звичайних комп'ютерах реалізується локальна пам'ять, або локальне подання, в якому використовується один обчислювальний елемент для кожного елемента даних. На основі розподіленої архітектури подання інформація в ШНМ може дробитися і оброблятися по частинах;

— *ізоморфізм* ШНМ топології навчальних даних і структурній інформації про модельовану залежність. В основі принципів побудови штучних НМ закладені: наївний ізоморфізм Кьолера, який вимагає, щоб подання події в НМ, мало просторову структуру, що нагадує структуру відображуваного реального об'єкта, а також ізоморфізм Гріна, що допускає, щоб подання події в ШНМ мало таку логічну структуру (не обов'язково фізично реалізовану в просторовій формі), яка дозволяла б розбивати це подання, розчленовувати його і знову відновлювати за допомогою відповідних процедур або процесів направленої уваги так, що встановлюється зв'язок між частинами, поверхні або аспектами явищ реального світу. Відомо, що існують ділянки мозку —«сенсорні проекційні поля», — в яких справді зберігається просторова структура образів. Однак залишаєть-

ся невідомим, наскільки далеко має поширюватися код уявлень в сенсі просторового ізоморфізму, щоб у ньому враховувалися організуючі властивості досвіду. Показано, що в ряді завдань необхідне структурне подання може бути отримано за допомогою адаптивних процесів і не потрібно, щоб уявлення були повністю ізоморфні піднаглядним явищам;

— *паралельна архітектура*, *масовий паралелізм і розподіленість обчислень* — обробка інформації в ШНМ виконується кількома процесорними елементами, в той час як у звичайних комп'ютерах, що мають тільки один центральний процесор, інформація обробляється послідовно, крок за кроком;

— *ієрархічна організація структури, цілісність і дробимість елементів ШНМ*: багато завдань штучного інтелекту неможливо вирішити без використання ієрархічних структур, що дозволяють будувати моделі складних об'єктів з простіших; робота ієрархічної структури вимагає, щоб інформаційний елемент у кожному ієрархічному рівні вів себе як єдине ціле, але при переході з рівня на рівень допускав дроблення, причому при переході з верхнього ієрархічного рівня на нижній це дроблення відповідає виділенню складових його елементів, а при переході з нижнього рівня на верхній воно відповідає включенню певної частини цього елемента в більш складний об'єкт;

— *попередня організація в навчанні*: незважаючи на те, що ШНМ мають більш широкі можливості до адаптації і навчання на прикладах порівняно з традиційними методами штучного інтелекту, вирішити на їх основі складні задачі штучного інтелекту за рахунок одного лише навчання, мабуть, неможливо. Необхідно, щоб вихідна структура ШНМ була попередньо організована. Для виконання цієї умови необхідно мати достатньо розвинені інструментальні засоби, що дозволяють формувати структуру ШНМ до початку її навчання;

— *функціональна блочність* — полягає у побудові архітектури НМ на стандартизованих функціональних блоках. Для вирішення задач штучного інтелекту необхідно будувати ШНМ, які мають велику кількість нейронів, тому базовими елементами для інструментальних засобів передорганізації ШНМ повинні бути великі функціональні блоки, внутрішня організація і властивості яких визначені заздалегідь і відомі розробнику;

— *поліалгоритмічність* — властивість, що дозволяє одній і тій самій ШНМ одночасно обробляти вхідну інформацію за різними методами. Поліалгоритмічність пов'язана з наявністю в мережі різних вихідних нейронів, з'єднаних з нейронами інших ієрархічних рівнів. При цьому ланцюжки, по яких збудження передається від входу до вихідних нейронів, мають різну структурну конфігурацію, яка залежно від сполучень схем проміжних нейронів, що включають збуджуючі або гальмують синапси, відповідає різним математичним формулам;
— *логічна прозорість*: логічно прозорою називають НМ, яка вирішує задачу зрозумілим для людини способом, для якої легко сформулювати словесний опис у вигляді явного методу. Така мережа повинна мати мінімальну структурну складністю і при цьому задовольняти вимогам (перевагам) користувача і (або) вимогам методу автоматизованого здобуття знань до виду результуючої мережі;
— *здатність отримувати знання з даних*: в процесі навчання ШНМ засвоює (апроксимує) найбільш загальні закономірності, присутні в навчальних даних, витягуючи тим самим неявні для людини знання з даних. Труднощі використання таких мереж складаються в нашій обмеженій здатності пояснити, як ШНМ вирішує завдання: внутрішні уявлення НМ, що виходять в результаті навчання, настільки складні, що їх неможливо проаналізувати, за винятком найпростіших випадків. Спрощення ШНМ, що досягається контрастуванням, дозволяє отримувати логічно прозорі ШНМ, на основі структури яких можна побудувати вербальний опис методу отримання відповіді за наявною нейромережевої моделі – витягти явні знання з даних;
— *точність вирішення задач* ШНМ визначається помилкою розпізнавання (апроксимації залежності) для навчальних даних, а також помилкою розпізнавання (апроксимації залежності) для тестових даних;
— *ефективність (якість) вирішення завдань* ШНМ визначається точністю (помилкою) рішення задачі для навчальних і тестових даних, простотою, логічної прозорістю і швидкодією отриманої нейромережевої моделі, а також витратами на побудову нейромережевої моделі (вимоги до апаратних засобів, ітераційність і витрати часу методу навчання);

— *варіативність моделей апроксимованої залежності*: для однієї і тієї ж таблиці даних і різних мереж або однієї мережі, але з різною початковою генерацією вихідних значень набору параметрів, що настроюються, після навчання, спрощення за єдиною схемою та вербалізацією можуть вийти різні ШНМ (нейромережеві моделі) і, відповідно, кілька методів вирішення однієї і тієї ж задачі. За кінцевою таблицею даних завжди будується кілька методів (моделей) рішення. Далі методи починають перевірятися і конкурувати між собою. Комбінуючи фрагменти декількох методів (моделей), можна сконструювати нову теорію. У силу цього неєдиність одержуваного знання (варіативність, неєдиність нейромережевої моделі) не є недоліком;

— *складність ШНМ* визначається кількістю і топологією міжнейронних зв'язків, складністю виконуваних перетворень і кількістю нейроелементів;

Новітня історія ШНМ бере свій початок з 1943 р. — з моменту виходу праці *У. Маккаллоха* та *У. Піттса* [2], у якій досліджено властивості найпростішої моделі нейрона із двома стійкими станами (модель Маккаллоха — Піттса), на вхід якої надходить приско- рювальний і гальмуючий сигнали. Вони довели, що за допомогою такої моделі можуть бути реалізовані різні логічні функції, наприклад, І, АБО, НЕ та ін.

У 1949 р. *Д. Гебб* [3] не тільки встановив, що пам'ять у біологічних системах викликається процесами, що змінюють зв'язки між нейронами, але й запропонував правило зміни цього зв'язку — правило навчання нейрона, назване його ім'ям. І сьогодні правило Гебба в тій або іншій формі присутнє в багатьох алгоритмах навчання ШНМ.

Підбадьорюючими були й результати першого комп'ютерного моделювання ШНМ, проведеного в 1956 р. під керівництвом *Н. Рочестера* [4]. Цей рік вважається роком народження наукового напрямку, що має назву «штучний інтелект». В основу модельованої ШНМ покладено вдосконалену модель Гебба. Дана робота дала поштовх численним роботам з моделювання нейронних мереж.

Найбільш важливі результати для розвитку ШНМ були отримано в цей період у Массачусетському технологічному інституті групою вчених під керівництвом Ф. *Розенблатта* [5]. Ними не тільки досліджено різноманітні варіанти мереж, названі персептронами, та вивчено різні способи їх навчання, але й здійснено технічну реалізацію персептрона. Синтезований ними персептрон використав граничну логіку, складався із трьох шарів і був здатен вирішувати прості завдання розпізнавання

цифр. Більш докладно властивості персептрона ми розглянемо нижче. Зазначимо тільки, що робота *Ф. Розенблатта* [6], у якій наведено різні варіанти персептрона й доведено теорему збіжності для персептрона, викликала величезний інтерес до ШНМ.

Водночас під керівництвом Б. Уідроу розвивався напрямок, пов'язаний із практичним застосуванням ШНМ, побудованих на елементах ADALINE (ADAptive Linear NEuron), аналогічних персептрону. Мережа, що містить багато таких елементів, була названа MADALINE (Multiple ADALINE's). Для навчання такої мережі було використано алгоритм, що мінімізував квадратичну помилку навчання — алгоритм *Уідроу — Гоффа* [7,8]. Слід зазначити, що згодом Б. Уідроу заснував фірму, що займається розробкою електронних компонентів для нейрокомп'ютерів.

Райдужним надіям щодо систем штучного інтелекту, компонентами яких були ШНМ, поклала кінець монографія *М. Мінські* й *С. Пейперта* [9], присвячена серйозному аналітичному дослідженню властивостей персептрона, що з'явилася у 1969 р. Автори показали, що можливості персептрона (мається на увазі одношаровий) обмежуються реалізацією тільки найпростіших логічних функцій, і то не всіх. Наприклад, з його допомогою неможливо організувати функцію «що виключає АБО» тощо. Хоча дослідники й розуміли, що можна будувати персептрони/ що містять більше одного шару й мають, вочевидь, ширші можливості, зовсім незрозуміло було, як можна навчити приховані шари. Тому монографія М. Мінські й С. Пейперта викликала справжній шок серед учених і стала настільки сильним ударом по дослідженнях у цій галузі, що такі роботи практично припинилися.

Увага дослідників переключилася на інші види мереж. У 80-ті роки XX ст. досягнуто значних успіхів у такій галузі ШНМ, як асоціативна пам'ять. Піонерськими тут є роботи *Т. Когонена* [10—14] і *Дж. Андерсона* [15, 16]. Як модель Т. Когонен використав так званий одношаровий лінійний асоціатор, а навчання здійснював за правилом Гебба, при якому деякий пропонований мережі образ асоціювався з іншими. Дж. Андерсон, вирішуючи завдання побудови асоціативної пам'яті, що здійснює пошук інформації, яка зберігається, не за її адресою, а за змістом, також користувався правилом навчання Гебба для побудови матриць, що служать для подання збережених у пам'яті й виклику необхідних образів.

Важливою подією в теорії й практиці ШНМ стала поява алгоритму навчання багатошарових мереж, який отримав назву *алгоритму зворотного поширення помилки*. Вперше запропонований 1974 року

П. Вербосом у його дисертації [17] метод виявився непоміченим. Непоміченою виявилася й праця *Д. Паркера* [18], що вийшла в 1985 р. Лише після появи в 1986 р. праці *Д. Румельхарта*, *Г. Гінтона* й *Р. Уільямса* [19] цей алгоритм отримав широке розповсюдження. Алгоритм забезпечує мінімізацію вихідної помилки мережі шляхом настроювання вагових коефіцієнтів усіх шарів ШНМ, починаючи з останнього, переходячи до передостаннього й т. д., аж до вхідного шару. Таким чином, під час його роботи помилка немовби поширюється від виходу мережі на її вхід, чим і пояснюється назва алгоритму. Розробка цього алгоритму дозволила не тільки перебороти обмеженість одношарового персептрона й конструювати багатошарові мережі, що мають значно більші можливості, але й дала новий імпульс розвитку ШНМ. Поступово з'явився необхідний для конструювання потужних багатошарових мереж теоретичний фундамент. Оцінка М. Мінські виявилася надто песимістичною, і багато хто з поставлених у його книзі завдань вирішуються в цей час мережами за допомогою стандартних процедур.

Якщо перші структури ШНМ були простими й використовували лінійні статичні моделі нейронів, то з часом сконструйовано мережі, зокрема динамічні, що мають більш складні структури й здатні вирішувати значно складніші завдання. Особливий інтерес викликали праці С. *Ґроссберґа* [20-22]. Так, займаючись дослідженням динамічних властивостей мереж, він разом з М. Коеном довів фундаментальну теорему про глобальну збіжність динамічних мереж — теорему Коєна — Ґроссберґа. Дослідження біологічних об'єктів дозволили Ґроссберґу і його колегам установити дилему стабільності — пластичності (запам'ятовування нової інформації без руйнування тієї, що вже зберігається у пам'яті) і сконструювати мережу, що володіє такими властивостями. Ця мережа отримала назву ART (*Adaptive Resonance Theory*). Робота [22] поклала початок цілій низці публікацій С. Ґроссберґа і його колег, присвячених розробці й дослідженню різних видів ART: ART-1, ART-2, ART-3, ART MAP, Fuzzy-ART.

Величезну роль у розвитку ШНМ відіграли й праці фізика *Дж. Гопфілда* [23-25], у яких на основі функцій Ляпунова досліджувалася проблема стійкості симетричних рекурсивних мереж. У цих працях мережа вперше характеризувалася деякою енергетичною функцією, мінімуми якої відповідають збереженим образам. Сконструйована Дж. Гопфілдом мережа, що носить його ім'я, здатна вирішувати дуже складні оптимізаційні завдання, наприклад відома задача про комівояжера, що й було продемонстровано в його спільній з *Д. Танком* роботі [25]. Слід зазначити, що підкреслена практична спрямованість робіт

Дж. Гопфілда послужила основою того, що вже в 1987 р. під його керівництвом був створений *нейрочип*.

Перша переконлива перевага використання ШНМ під час вирішення складних практичних завдань продемонстрована роботою *Т. Сейновського* й *Ч. Розенберґа* [26] на прикладі автоматичного навчання англійській мові.

Сьогодні існує вже достатньо типів ШНМ, які дозволяють успішно вирішувати завдання розпізнавання образів і мови, класифікації, побудови математичних моделей і керування, оптимізації й прогнозування.

Останнім часом виникле нове покоління штучних нейронних мереж [27, 28], які отримали назва «спайк-нейронні мережі» та які є (за характером свого функціонування) близькими аналогами біологічних нейросистем. Вирішуючи в принципі ті ж проблеми, що й традиційні ШНМ, спайк-нейронні мережі володіють істотно більшою швидкодією та обчислювальною потужністю. При цьому при вирішенні одного і того ж завдання для вони мають простішу архітектуру, тобто мають значно менше нейронів, ніж «класичні» ШНМ.

Особливо слід відзначити появу та розвиток згорткових мереж, які є вдалою серединою між біологічно правдоподібними мережами та звичайним багатошаровим персептроном. На сьогоднішній день найкращі результати у розпізнаванні зображень отримують за їх допомогою. У середньому точність розпізнавання таких мереж перевищує стандартні ШНМ на 10–15 %. Згорткові мережі – це ключова технологія Deep Learning.

Галузі застосування ШНМ

ШНМ знаходять сьогодні широке застосування у будь-яких галузях, починаючи від завдань апроксимації функцій і закінчуючи створенням нейрокомп'ютерів. Важко описати всі можливі сфери застосування нейронних мереж, тому ми коротко зупинимося лише на деяких з них.

Апроксимація функцій

Установлення універсальних апроксимуючих властивостей ШНМ стало досить важливим етапом у становленні загальної теорії й стимулювало дослідження в даній галузі. Багатьма дослідниками було встановлено, що нейронна мережа з одним прихованим й одним вихідним шаром здатна апроксимувати з будь-якою наперед заданою точністю на компактній множині будь-яку неперервну функцію. Так, *Р. Гехт-Нільсен* [29, 30], спираючись на теорему Колмогорова — Арнольда [31, 32] про подання неперервних функцій декількох змінних у вигляді суперпозицій неперервних функцій одного змінного, довів можливість

апроксимації функцій багатьох змінних досить загального виду за допомогою двошарової нейронної мережі із прямими повними зв'язками з фіксованою кількістю нейронів з заздалегідь відомими обмеженими функціями активації. *Дж. Цибенко* й *К. Хорник* [33, 34], використовуючи теорему Хана-Банаха, довели, що кінцева лінійна комбінація фіксованих одновимірних функцій може однозначно апроксимувати будь-яку неперервну функцію n дійсних змінних на заданому гіперкубі при досить м'яких припущеннях щодо функцій одного змінного.

Асоціативна пам'ять

Існують архітектури ШНМ, які запам'ятовують образи, що надходять на них, з чимось їх асоціюючи, а при пред'явленні деякого «асоціативного» образу витягують їх з пам'яті. Ця властивість дозволяй організовувати пошук інформації не за адресою, а за її змістом. Навіть якщо пропонована асоціація, що відновлюється, буде спотворена перешкодами, мережа може видати правильний результат. Залежно від того, чи збігається шуканий образ зі збереженим у пам'яті, чи ні, розрізнюють авто- і гетероасоціативну пам'ять.

Дослідженню асоціативної пам'яті присвячена монографія [10].

Стиснення даних

Деякі типи ШНМ мають властивості, що дозволяють використовувати ці мережі для стиснення даних, наприклад перед їхньою передачею, зменшуючи тим самим кількість переданих бітів інформації. Подібні завдання виникають і в кластерному аналізі, коли різні, схожі за певними ознаками образи об'єднуються в деякі групи або кластери, тобто здійснюється перехід від вихідного n-вимірного простору образів до m-вимірного простору кластерів, де $m < n$. Подальша робота в просторі меншої розмірності призводить до економії обчислювальних ресурсів і зменшення обсягу необхідної пам'яті.

Застосування ШНМ для стиснення даних у системах розпізнавання мови й зображень забезпечує стиснення даних у 100 і більше разів [35].

Розпізнавання та класифікація

Розпізнавання образів (зображень, зокрема текстів, друкованих і рукописних, звуку, мови тощо) є тією галуззю, де найбільш яскраво виявляються переваги ШНМ. Вирішення багатьох задач розпізнавання образів ускладнено внаслідок їхньої високої розмірності. Як ми вже зазначали раніше, використання ШНМ шляхом стиснення даних дозволяє знизити розмірність задачі, зберігаючи властивості подільності розподілів, що відповідають різним класам.

Багато важливих застосувань теорії розпізнавання образів відносяться до задач класифікації кривих і геометричних фігур. Такими є, наприклад, завдання діагностики, виявлення несправностей тощо. І в цьому випадку ШНМ дають ефективне вирішення завдання незалежно від того, існує навчальна множина вже класифікованих об'єктів чи не існує. Однак наявність такої інформації прискорює процес пошуку вирішення [36,37].

Оптимізаційні задачі

Більшість практично важливих задач можуть бути сформульовані як оптимізаційні, що доставляють екстремум деякому заздалегідь обраному критерію. Часто число варіантів рішень швидко зростає з числом структурних елементів N у завданні (наприклад, як N!), І пошук точного рішення для практично корисних значень N стає свідомо неприйнятно дорогим.

Прикладом такої задачі є задача про комівояжера, що має величезне практичне значення й вирішення якої вимагає значних обчислювальних витрат. Ця задача полягає в тому, що комівояжер має відвідати задану кількість міст, вибравши для цього найкоротший маршрут, причому в кожному місті він повинен побувати не більше одного разу. Така задача виникає, наприклад, під час створення автоматичних технологічних ліній (свердлення отворів у друкованих платах, трасування друкованих плат та ін.), визначення оптимальних маршрутів перевезення тощо.

Такі завдання називають неполіноміально складними чи NP-повними. Якщо вдається сформулювати таке завдання термінах оптимізації функції енергії (інформації, Ляпунова), то нейронна мережа дає дуже потужний інструмент пошуку наближеного рішення.

Доведено, що задача задачу про комівояжера відноситься до класу «NP-повних», кращим методом вирішення яких є повний перебір можливих варіантів. Оскільки у випадку *N* міст існує *N!* варіантів обходу, очевидно, що із збільшенням кількості міст складність вирішення завдання різко зростає. Одне з можливих ефективних вирішень цієї задачі за допомогою ШНМ, що вимагає значно менших обчислювальних витрат і при цьому збігається з розв'язком, отриманим повним перебором, було запропоновано в роботі [25].

Керування складними процесами

Проблема синтезу ефективної системи керування є досить складною, оскільки реальні процеси характеризуються, як правило, нелінійними залежностями, високим рівнем шумів та їх корельованістю, які змінюються умовами функціонування, що обумовлюють зміну характе-

ристик досліджуваних об'єктів тощо. Необхідність вирішення задач керування в реальному часі висуває певні вимоги як до самих алгоритмів керування, що входять до складу математичного забезпечення проектованої системи, так і до технічних засобів, що їх реалізують. У цих умовах найбільш ефективними виявляються методи й алгоритми, що базуються на теорії адаптації. Однак, як і при будь-якому підході, ці методи вимагають розробки математичних моделей досліджуваних об'єктів.

Слід зазначити, що отримані математичні моделі використовуються не тільки з метою безпосередньо керування, але й для упередження поведінки об'єкта, що дозволяє підвищити ефективність керування шляхом завчасної корекції керованих параметрів. Якщо одержання математичної моделі ускладнено або вимагає істотних зусиль, доцільне застосування ШНМ. Використанню ШНМ у системах керування присвячені, наприклад, монографії [38—44].

Прогнозування

Будь-яке прогнозування (екстраполяція) спирається на формалізоване уявлення про існуючий зв'язок між причинами й наслідком. Багато процесів формуються під впливом великої кількості факторів, що діють у різних напрямках і нерідко невідомих. Статистичний аналіз цих процесів містить дослідження взаємозв'язків факторів як у статичному стані, так і в часі. Інформацією для вивчення взаємозв'язків служать часові ряди показників, що характеризують розвиток об'єктів. Найпоширенішим підходом до вирішення задачі прогнозування є екстраполяція чинних у цей час зв'язків і закономірностей на майбутнє. Побудовані відповідно до цього принципу моделі прогнозування відрізняються одна від іншої лише гіпотезами про конкретні види збережених зв'язків. Чим більш загальні припущення закладені у форму моделі й чим більший клас процесів можна описати з її допомогою, тим ширше її можливості під час дослідження окремої реалізації.

Таким чином, вибір і основа математичної моделі є центральним моментом прогнозування. На практиці ж нерідко виявляється, що внаслідок тих чи інших причин отримати математичну модель,, яка адекватно відбиває властивості досліджуваного об'єкта, надзвичайно складно. І в цих випадках ефективним виявляється використання ШНМ.

У роботах [39, 45-48] наведено численні приклади успішного застосування ШНМ для вирішення завдань економічного, зокрема, фінансового прогнозування.

Нейрокомп'ютери

Один з сучасних напрямків створення розумних машин — це розробка нейрокомп'ютерів. Нейрокомп'ютер — це програмно—технічна

система (спеціалізована ЕОМ), яка реалізує деяку формальну модель природної мережі нейронів. В основу машин п'ятого покоління покладено ідею паралельної обробки інформації в нейроподібних системах. Не зважаючи на те, що електронний процесор працює в тисячі разів швидше, ніж його нейронний еквівалент у мозку, мережі нейронів вирішують багато задач (особливо нечислових) в тисячі раз швидше, ніж електронний процесор.

Причини цього наступні:

1) Характер взаємозв'язків між нейронами дозволяє розв'язувати багато задач на основі паралельної обробки;

2) У нейронній мережі пам'ять не локалізована в одному місці (як в послідовних машинах), а розподілена по всій структурі. В біологічних системах пам'ять реалізується підсиленням або послабленням зв'язків між нейронами, а не зберіганням двійкових символів;

3) Біологічні мережі реагують не на всі, а тільки на визначені зовнішні подразнення. Кожний нейрон виступає як елемент прийняття рішення і як елемент зберігання інформації. Перевага такої структури — «життєздатність» (вихід з ладу декількох нейронів не приводить до значної зміни даних, що зберігаються, або ж до руйнування всієї системи).

4) Можливість адресації за вмістом (асоціативної пам'яті) є ще однією важливою характеристикою систем з розподіленою пам'яттю (кожний елемент відшукується за його вмістом, а не зберігається в комірці пам'яті з визначеним номером).

В основу зв'язків в нейрокомп'ютерах покладено принцип асоціацій. Асоціативні зв'язки пронизують все мислення людини. Існує думка, що процеси мислення є не що інше, як розповсюдження певного збудження, як деяка ланцюгова реакція. Навіть найбільш примітивні процеси навчання принципово залежать від послідовності подій в часі. Це й закладено в природу нейронних систем. Тому їм притаманне реагування тільки на жорстко визначені зовнішні подразнення. Наприклад, домашні тварини «навчаються» ігнорувати повторні несуттєві зовнішні подразнення («цокання» годинника), але посилюють сприйняття подразнень, які можуть мати серйозні наслідки (звук автомобільних гальм).

Багато дослідників вважає, що майбутнє належить комп'ютерам, які базуються на аналізі зв'язків, а не обробці символів. Мінський говорив, що якщо комп'ютер повинен діяти подібно мозку, тоді й його конструкція повинна бути також подібна до мозку.

Моделі ШНМ і схеми з адресацією за вмістом мають і недоліки. Внаслідок нефіксованої організації вони можуть плутати різні об'єкти.

Але це аналогічно звиканню, посиленню чуттєвості до асоціацій, які лежать в основі психічних особливостей людини. Іншими словами, будь-який комп'ютер, який претендує на «розумність», повинен мати такі особливості.

Нейрокомп'ютери є обчислювачами нового класу, що забезпечують більш швидке, більш дешеве та якісне вирішення конкретних завдань. Нейрокомп'ютер реалізує ідею створення аналого-цифрової ЕОМ, у якій аналогова частина виконує багатовимірні операції в граничному базисі, а цифрова реалізує алгоритми настроювання параметрів нейронних мереж. Деякі порівняльні характеристики комп'ютера й ШНМ наведено в табл. В.1.

Таблиця В.1

	Комп'ютер	**ШНМ**
Підхід до вирішення задачі	Дедуктивний (використовує відомі правила перетворення вхідних даних у вихідні)	Індуктивний (задані вхідні й вихідні дані використовуються для конструювання правил)
Обчислення	Централізовані, синхронні, послідовні	Розподілені, асинхронні, паралельні
Пам'ять	Пакетована, зосереджена й адресована за розташуванням	Розподілена, адресована за змістом
Стійкість до перебоїв	Нестійкий (вихід з ладу одного елемента призводить до виходу з ладу всього пристрою)	Стійка за рахунок надлишковості й поділу функцій
Швидкодія	Висока (млн оп./с)	Низька (тис. оп./с)
Точність	Висока	Невисока
Архітектура	Фіксована	Що змінюється

Висока швидкодія нейрокомп'ютерів досягається за рахунок розпаралелювання обчислень. Кардинальним напрямком розвитку нейрокомп'ютерів як загального призначення, так і проблемно- орієнтованих є розробка нейрочипів. Досягнення в мікроелектро- ніці вселяють упевненість у тому, що незабаром будуть створені могутніші й водночас більш дешеві нейронні ЕОМ

Література до вступу

1. Субботін С. О. Нейронні мережі : теорія та практика: навч. посіб./ С. О. Субботін. – Житомир : Вид. О. О. Євенок, 2020. – 184 с.
2. Mc Culloch W.S., Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity // Bulletin of Mathematical Biophysics. – 1943. – 5. – P. 115-133 .
3. Hebb D. The Organization of Behavior. – New York: Wiley Publications, 1949.
4. Rochester N.J., Holland J.H., Haibt L.H., Duda W.L. Test on a Cell Assembly Theory on the Action of the Brain, Using a Large Digital Computer // IRE Trans. on Inf. Theory. –1956. – IT-2. – P. 80 – 93.

5. Rosenblatt F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain // Psychological Review. – 1958. – 65. – P. 386-408.
6. Розенблатт Ф. Принципы нейродинамики. Персептрон и теория механизмов мозга / Пер. с англ. – М.: Мир. - 1965.
7. Widrow B., Hoff M.E. Adaptive Switching Circuits / IRE WESCON Convention Record – New York, IRE, 1960. – P.96-104.
8. Widrow B., Winter R. Neural Nets for Adaptive Filtering and Adaptive Pattern Recognition // IEEE Computer. – 1988. – 21. – №3. – P.25-39.
9. Минский М.Л., Пейперт С. Персептроны / Пер. с англ.. – М.: Мир. - 1971.
10. Kohonen T. Correlation Matrix Memories // IEEE Trans. on Computers. – 1972. – 21. – P. 253-259.
11. Kohonen T. Associative Memory: A System Theoretic Approach. –Berlin: Springer, 1977.
12. Kohonen T. Self-Organized Formation of Topologically Correct Feature Maps // Biological Cybernetics. – 1982. – 43. – P. 59-69.
13. Kohonen T. The Self-Organizing Map // Proc. of IEEE. – 1990. – 78 – №9. – P. 1464-1480.
14. Kohonen T. The "Neural" Phonetic Typewriter // IEEE Computer. – 1988. –21. – №3. – P. 11-22.
15. Kohonen T. Self-Organization and Assotiative Memory. 3rd ed. – New York: Springer Verlag, 1989.
16. Anderson J.A. A Memory Model Using Spatial Correlation Functions // Kybernetik. – 1968. – 5. – P. 113-119.
17. Anderson J.A. Two Models for Memory Organization // Mathematical Biosciences. – 1970. – 8. – P. 137-160.
18. Werbos P.J. Bejond regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. thesis. – Cambridge, MA, Harvard University, 1974.
19. Parker D. Learning Logic // Technical Report TR-87, Center for Computational Research in Economics and Management Sciences – Cambridge, MA: MIT Press, 1985.
20. Rumelhart D.E., Hinton G.E., Williams R.J. Learning Internal Representations by Error Propagation / In Parallel Distributed Processing: Explorations in the Microstructure of Cognition. D.E. Rumelhart, J.L. McClelland (eds). – Cambridge: MitPress, 1986. – Vol. 1. – Chapt. 8. – p. 318-364.
21. Grossberg S. Embedding Fields: A Theory of Learning with Physiological Implications // J. of Math. Psychology. – 1969. – 6. – p. 209-239.
22. Grossberg S Competitive Learning: From Interactive Activation do Adaptive Resonance // Cognitive Science. – 1987. – 11. – p. 23-63.
23. Cohen M.A., Grossberg S. Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks // IEEE Tr. on Systems, Man and Cybernetics. – 1983. – 13. – p. 815-826.
24. Hopfield J.J. Neural Networks and Physical Systems with Emergent Collective Computational Abilities // Proc. of the National Academy of Science. – 1982. – 79. – p. 2554- 2558.
25. Hopfield J.J. Neurons with Graded Response Have Collective Computational Properties Like Those of Two-State Neurons // Proc. of the National Academy of Science. – 1982. – 81. – p. 3088-3092.
26. Hopfield J.J., Tank D.W. Neural Computation of Decisions in Optimization Problems // Biol. Cybernetics. – 1985. – 52. – p. 141-152.
27. Sejnowski T.J., Rosenberg Ch. NETTalk: A Parallel Network that Learns to Read Aloud. John Hopkins University Electrical Engineering and Computer Science Technical Report JHU / EECS – 86/01. – 1986. – 32 p.

28. . Maass W. Networks of spiking neurons: the third generation of neural network
29. models // Neural Networks. – 1997. – 10. –pp. 1659-1671.
30. Natschlager T. Networks of spiking neurons: a new generation of neural networks models. – 1998. – http://www.igi.tugraz.at/tnatschl/online/3rd_gen_eng/3rd_gen_eng.html.
31. Hecht-Nielsen R. Kolmogorov's Mapping Neural Network Existence Theorem / Proc. of IEEE First Ann. Int. Conf. on Neural Networks. – San Diego, 1987. – 3. – P. 11-13.
32. Hecht-Nielsen R. Theory of the Backpropagation Neural Network / Proc. of Int. Joint. Conf. on Neural Networks. – Washington, D.C., 1989. – 1. – P. 593-606.
33. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных суперпозициями функций меньшего числа переменных // Докл. АН СССР.–1956. – 108. – №2. – С.179-182.
34. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных в виде суперпозиции непрерывных функций одного переменного и сложения // Докл. АН СССР, 1957. – 114. – С.953-956.
35. Cybenko G. Approximation by Superposition of a Sigmoidal Function // Math. Contr. Sign. Syst. – 1989. – 2. – P. 303-314.
36. Hornik K., Stinchcombe M., White H. Multilayer Feedforward Networks are Universal Approximators // Neural Networks. – 1989. – 2. – P. 359-366.
37. Hornik K. Approximation Capabilities of Multilayer Feedforward Networks // Neural Networks. – 1991. – 4. – P. 251-257.
38. Patterson D. Artifical Neural Networks, Theory and Application. – Singapur: Prentice Hall Inc., 1996.
39. Bishop C.M. Neural Networks for Pattern Recognition – Oxford: Clarendon Press, 1995.
40. Image Processing and Pattern Recognition / T. Leondes (ed) – London: Academic Press, 1998.
41. Pham D.T., Liu X. Neural Networks for Identification, Prediction and Control. – London: Springer Verlag, 1997.
42. Control and Dynamic Systems / T. Leondes (ed) – London: Academic Press, 1998.
43. Industrial and Manufacturing Systems / T. Leondes (ed) – London: Academic Press, 1998.
44. Омату С., Халид М., Юсоф Р. Нейроуправление и его приложения / Пер. с англ. – М.: ИПРЖР, 2000.
45. Neural Networks for Control / Miller W.T., Satton R.S., Werbos P.J. (eds) – Cambridge: MIT Press, 1991.
46. Zbikowski R., Hunt K.J. Neural Adaptive Control – Berlin: Springer – Verlag, 1996.
47. Masters T. Neural, Novel and Hybrid Algorithms for Time Series Prediction. – N.Y.: John Willey and Sons, Inc., 1995.
48. Beltratti A., Margarita S., Terna P. Neural Networks for Economic and Financial Modelling – London: Int. Thomson Computer Press, 1996.
49. Zirilli J. Financial Prediction Using Neural Network London: Int. Thomson Computer Press, 1997.
50. Pham D.T., Liu X. Modelling and Prediction using GMDH Networks of Adalines with Nonlinear Preprocessors // Int. J. System Science. – 1994. – 25. – №11. – P. 1743 – 1759.
51. Субботін С. О. Нейронні мережі : теорія та практика: навч. посіб./ С. О. Субботін. – Житомир : Вид. О. О. Євенок, 2020. – 184 с.