

АНАЛІТИЧНИЙ ПІДХІД ДО ВИБОРУ ТОПОЛОГІЇ НЕЙРОСІТОК ПРИ РОЗВ'ЯЗАННІ ПРИКЛАДНИХ ЗАДАЧ

Анотація: На основі розробленого реляційного класифікатора запропонований універсальний підхід до вибору топології нейромереж, що задовольняють критеріям вирішення прикладної проблеми / задачі.

Ключові слова: Перцептрон, Нейротехнології, Штучний нейрон, Штучна нейромережа, Алгоритм зворотного поширення помилки, Навчання нейромережі, Топологія нейромережі, Багатофункціональне середовище, Рейтингове оцінювання, мульти-агентна система

Хронологія та характеристика проблеми. На початку 40-х років минулого століття з'явилися дві фундаментальні роботи, в яких вперше формалізується поняття *штучного нейрону* (ШН) і *штучної нейронної сітки* (ШНС) – це робота У. Мак-Каллоха та У. Піттса [43] про логічне числення ідей та нервової активності та фундаментальна праця Н. Вінера [4] і його колег про подання складних біологічних процесів математичними моделями (в подальшому закріпленому у терміні “кібернетика”). З появою запропонованого Ф. Розенблаттом *одношарового перцептрона* [50] (який використовують для розпізнавання образів, прогнозування погоди тощо), а в подальшому розробленої на основі дельта-правила – формули Уїдрю і побудованого на принципово нових засадах адаптивного лінійного елемента – *Адаліни* Уїдрю Б. і Хоффа М. [54] значно підвищується зацікавленість на комерційному рівні у підтримці розробок в галузі нейротехнологій (НТ). Цьому в значній мірі сприяв запропонований Д. Хеббом перший *алгоритм навчання нейрона* [31].

Проте вже проведені у 1963 році під керівництвом Петрова А.П. в Інституті проблем передачі інформації АН СРСР дослідження задач, “важких” для перцептрона [10], явилися відправною точкою для Бонгарда Н.Н. [3], щоб “порівняно невеликою переробкою алгоритму (перцептрона) виправити його недоліки”. Саме ці дві роботи і стали запорукою того, що перша хвиля ейфорії відносно ШНС була пригальмована в СРСР.

А наприкінці 60-х і майже до середини 70-х років минулого століття спостерігається вже різке спадання активності у розвитку НТ та їх прикладного застосування, яке було викликано спостереженими Мінським М. і Пайпертом С. [44] обмеженнями у роботі (модельованні виключно найпростіших логічних функцій, складності у навчанні та перенавчанні тощо) і пов'язаними з інваріантністю подання (так звана проблема “парності” та “один у блоці”) штучного нейрона (ШН) Мак-Каллоха і Піттса, а також перших моделей одношарового перцептрона Розенблатта Ф.

і адаліни Уїдроу Б. і Хоффа М. І хоча у дослідників НТ було розуміння можливості будувати багат шарові сітки з, очевидно, більш широкими властивостями, але залишалося нез’ясованим питання щодо навчання схованих шарів. Проте, окремі дослідження в галузі НТ продовжувалися, і серед них, в першу чергу, треба відзначити самоорганізуючу НС – *конгнітрон* (та *неокогнітрон*), запропонований Фукушимой К. [24, 25] і призначений для інваріантного розпізнавання образів (досягається при запам’ятовуванні практично усіх станів образу), а також фундаментальні розробки Хопфілда Д. [34, 35] з перспективною і надзвичайно важливою для просування НТ на комерційний ринок ідеєю створення моделі НС з *оберненими зв’язками*, яка уявляє собою мінімізуючу енергетичну систему. Крім того, необхідно відзначити ряд праць Кохонена Т. [37, 38], в яких обґрунтовується можливість створення моделей НС, які *навчаються без учителя* (*самоорганізуючі карти Кохонена*) і використовуються в задачах кластеризації, візуалізації та інших для попередньої обробки даних. Всі ці дослідження набули значного розвитку в реалізації можливості відтворення спотворених або зашумлених даних і базуються на природній властивості мозку використовувати асоціативні зв’язки при пошуку інформації за допомогою ключа – послідовності бітів, яка порівнюється зі всіма ключами інформації, що зберігається. Такий вид пам’яті дістав назву *пам’яті, що адресується за змістом*. Якщо в моделі обчислень фон Неймана звертання до пам’яті є доступним тільки за допомогою адреси, яка не залежить від вмісту пам’яті, і більш того, якщо допущена помилка в обчисленні адреси, то може бути знайдена зовсім інша інформація, то асоціативна пам’ять (що адресується за змістом) є доступною по указанню заданого змісту. Значний внесок у розв’язання проблеми асоціативного навчання і створення *гетеро – авто- і двоспрямованої асоціативної пам’яті*, а також відтворення принципів конкурентного навчання зробили також Андерсен Д.ж. [17], Гроссберг С. [26] і Коско. Б. [39].

Тільки з появою запропонованого в дисертації Вербоса П. [53] і паралельно в роботі Галушкіна О. І. [5], але спочатку залишеного поза уваги дослідниками, а в подальшому розвиненого Румельхартом Д., Хінтоном Г. і Уільямсом Р. [51] та незалежно і одночасно Барцевим С. І. та Охонінім В. О. [2, 18] методу навчання з використанням *алгоритму зворотного поширення похибки* (АЗПП) з’явився потужний теоретичний фундамент для конструювання ефективних багат шарових ШНС. Розпочався період вибухової зацікавленості до ШНС, які навчаються.

Нарешті, великого значення набувають *сітки з радіально-базисними функціями* (РБФ-сітки), що уявляють собою спеціальний тип ШНС з прямими зв’язками і основне призначення яких – апроксимація та інтерполяція багатомірних функцій для розв’язання, зокрема, задач прогнозування. Вони запропоновані Повеллом М. [49] і спираються на математичну основу теорії апроксимації та інтерполяції багатомірних функцій, раніше розробленої Айзерманом М.А., Браверманом Е.М. і Розоноєром Л.І. [1]. Суть математичної основи базується на тому, що скільки

завгодно точна апроксимація функцій досягається шляхом комбінації радіально симетричних функцій.

Сьогодні важко уявити область досліджень, будь то технічне або технологічне проектування, діагностування, генетика, біоніка, авіоніка, економіка, ергономіка, комунальне господарство, транспортні задачі і багато інших застосувань, де би не впроваджувалися НТ. Саме ця зацікавленість з боку користувачів спонукає дослідників в області НТ до створення численних топологій ШНС як спеціального, так і універсального призначення. Проте, багатоваріантність вибору ШНС, в свою чергу, створює утруднення у прийнятті адекватного рішення у користувача, і це, на наш погляд, і складає *основну проблему використання НТ*.

Ускладнення проблеми. Вище наголошувалося на тому, що після “перевідкриття” і суттєвого доопрацювання Румельхартом Д., Хінтоном Г. і Уільямсом Р. і незалежно та одночасно Барцевим С. І. та Охоніним В. О. методу навчання з використанням АЗПП наприкінці минулого – початку теперішнього століть створилися підстави до “ренесансу” НТ із все зростаючою зацікавленістю з боку користувачів і задоволенням попиту новими розробками у топологіях ШНС. Останнє відбувається двома шляхами:

1. За рахунок використання сполучень, модифікацій і розширень вже існуючих моделей ШНС;
2. Створенням моделей ШНС на нових архітектурних, структурних і топологічних принципах побудування і методах навчання або використання їх для обслуговування інших оптимізаційних алгоритмів, зокрема, генетичних.

Яскравим представником *першого* може служити сполучення двох ШНС: Кохонена і Гроссберга з *новим методом навчання – “зустрічного поширення”* (ЗП), запропонованим та розвиненим Хехт-Нільсеном Р. [30]. Такий симбіоз різних ШНС дозволив покращити властивості результуючої сітки (зменшення часу навчання; можливість роботи як з неперервними, так і з двійковими сигналами; блокове побудування сітки шляхом каскадування спеціалізованих модулів тощо) і розширити сфери застосування (апроксимація функцій, розпізнавання, обробка і відновлення зображень, класифікація). Існує два різновиди двошарових ШНС ЗП: *шар Кохонена – шар Гроссберга* та *самоорганізуюча карта Кохонена – шар Гроссберга* (різниця в організації схованого шару і, відповідно, в способі навчання його нейронів) [27].

Іншим прикладом розробок другого шляху розвитку ШНС є створення на базі ідеї сітки Хопфілда з додаванням спеціального модуля нелінійного перетворення нової парадигми, що працює з аналоговими сигналами – *“синергетичного комп'ютера”* (СК), запропонованого засновником синергетики Хакеном Х. [28]. Цей комп'ютер відноситься до асоціаторів з інтерпретацією образів, які в ньому зберігаються, як локальні енергетичні мінімуми. Перевага над сіткою Хопфілда – відсутність не-

стійких станів завдяки наявності динамічного модуля, що збігається до одного з базових векторів.

Суттєве скорочення витрат на пам'ять і об'єм необхідних обчислень досягається в *сітці Хеммінга* [42], в якій замість образу, що зберігається у пам'яті, на виході видається тільки його номер. Тому, на відміну від сітки Хопфілда, місткість сітки Хеммінга визначається не розмірністю вхідного сигналу N , а кількістю нейронів шару L на виході. Крім того, перевагою сітки Хеммінга є те, що залежність між кількістю нейронів і кількістю з'єднань в сітці Хеммінга лінійна, тоді як у Хопфілда – квадратична, і при однаковій ємності з сіткою Хопфілда перша буде містити значно менше сполучень між нейронами.

Ще один приклад створення нової топології за рахунок використання базової моделі багат шарового перцептрона прямого поширення являє собою *ШНС з елементами затримки сигналу* (ШНС ЗС). Основне призначення – краще опрацювання (класифікація) зсунених у часі образів (як правило, мовних сигналів), які надходять на вхідний шар із згасанням у часі. Так, оскільки корисна мовна інформація часто подається згасаючою у часі в середовищі перешкод (паузи між надходячими фонемами, зашумлення тощо), а тривалість окремих фонем може сильно коливатися, ШНС ЗС повинна однаково ефективно опрацювати такі фонемні, тобто бути *робастною* до зовнішнього зсуву надходячих на сітку образів [41, 52]. Відмінність такої сітки полягає в наявності можливості створення за рахунок елементів затримки або у самих нейронах, або між схованими шарами тимчасових копій надходячих сигналів, через що на нейрони наступного шару подаються не тільки самі образи, але й їх копії. Цього принципу дотримуються всі шари ШНС ЗС, завдяки чому і досягається робастність сітки до часових спотворень у вхідних образах.

Представником *другого шляху* може служити ШНС, яка дістала назву *машини Больцмана* [15, 32]. Якщо сітка Хопфілда збігається до локального мінімуму, то в новому різновиді ШНС збіжність має відбуватися у глобальному мінімумі, а оскільки на шляху до цього сітка проходить локальні мінімуми, то існує ймовірність “зависання” в них. Тому по мірі просування до глобального мінімуму на кожному локальному мінімумі алгоритмічно задається збільшення “умовної енергії”, достатньої для подолання локального максимуму. Надалі енергія “розгойдування” сітки поволі зменшується по мірі наближення до глобального мінімуму. Ця ідея і реалізована в машині Больцмана, сполучаючи детерміновану сітку Хопфілда з ймовірнісним/стохастичним правилом навчання, яке дістало назву *відпалу*. Найчастіше цей різновид ШНС застосовується в задачах класифікації/розпізнавання невідомих або спотворених образів. Серед переваг – можливість виходу з локальних екстремумів, серед недоліків – трудомісткість пов'язаних з алгоритмом обчислень і, як наслідок, суттєві обчислювальні та часові витрати.

Трохи зменшує цей недолік *машини Коші* [46], в якій за рахунок використання в алгоритмі розподілу Коші можна скоротити час навчання;

проте у порівнянні з перцептроном цей час залишається занадто великим.

Класом ШНС, які яскраво доводять свою причетність до другого шляху розвитку суттєвими відмінностями від сіток Хопфілда як структурно (може мати декілька шарів з наявністю власних зворотних зв'язків у нейронів), так і особливостями навчання (матриця ваг може бути несиметричною, а навчання є контрольованим шляхом використання АЗПП), слугують неперервні і дискретні *динамічні рекурсивні сітки* (ДРС) [48, 16, 15, 50]. Наявність зворотних зв'язків між нейронами різних шарів і шару на виході включно забезпечує цьому класу ШНС додаткові позитивні властивості, яких неможна досягти в статичних багат шарових ШНС прямого поширення: можливість роботи з образами, параметри яких змінюються в часі. Дискретні ДРС поділяються на:

- *повнозв'язні* [48, 15] (застосування: аналіз та опрацювання послідовностей, ідентифікація нелінійних динамічних об'єктів);
- *частково-рекурсивні* – багат шарові перцептрони, доповнені контекстним шаром, кількість нейронів якого співпадає з кількістю виходів ШНС. Останнє має місце в сітці Джордана М. [36] (застосування: розпізнавання і класифікація образів, прогнозування часових рядів, а також більш ефективно в порівнянні з ШНС прямого поширення асоціювання надходячих на вхід образів з послідовностями на виході – так зване “проковзне” вікно; недолік – повільна збіжність і виникаючі через це проблеми сталості) або в модифікованій версії цієї ШНС – *сітці Елмана Дж.* [22] з сигналами зворотних зв'язків не вихідного шару, а з виходів нейронів контекстного шару з кількістю нейронів, співпадаючою з кількістю нейронів схованого шару;
- *локально-рекурсивні* з лінійною структурою як багат шаровий перцептрон або як сітка радіальної основи, в яких рекурсивність завжди обмежується одним нейроном і інтерпретується як фільтр із скінченною або нескінченною імпульсними характеристиками [45].

Представниками нового класу ШНС прямого поширення з можливістю за необхідності розширювання своєї архітектури в процесі функціонування є *каскадно кореляційні сітки* (ККС) з лінійними або/та нелінійними активаційними функціями нейронів і використанням алгоритму навчання з будь-яким правилом (дельта-правило, АЗПП тощо), запропоновані Фалманом С. і Леб'єром К. [23]. Навчання ККС повторюється до досягнення бажаної точності для всіх пар образів і базується на: покроковому інкрементному введенні нейронів у сховані шари, які з'являються тільки після введення нових нейронів; навчанні знов уведених нейронів при зберіганні ваг нейронів, які вже до того були у сітці. Переваги ККС: відсутність апріорного вибору архітектури сітки, бо як вихідна береться найпростіша з вхідним та вихідним шарами, а далі нарощується, поки сітка не визначить задовольняючу критеріям архітектуру; відсутність конкуренції між нейронами через їх незв'язність

значно скорочує процес навчання в порівнянні з іншими топологіями сіток; з поданням нових образів може не відбуватися перенавчання всієї сітки, а тільки шляхом корекції зв'язків нейронів вхідного та схованого шарів; процес навчання сітки значно спрощується в обчислювальному сенсі, бо коректуються ваги тільки одного шару, і сигнали у сітці проходять тільки в одному напрямі. Призначення: основні переваги ККС проявляються в повній мірі при застосуванні в обчислювальних середовищах з використанням паралельних обчислень.

Розв'язання “дилеми стабільності-пластичності” (за Гроссбергом С.) запропонованою двошаровою нейросіткою *адаптивної резонансної теорії* (АРТ) Карпендером Г. і Гроссбергом С. [20] дозволяє адаптивно навчатися при надходженні нових даних із збереженням *стабільності* щодо гарантії неруйнування попередньо накопичених даних, з одного боку, і достатньої *пластичності* відносно визначення суттєвості нових надходжень, з другого. Основне призначення АРТ – моделювання володіючих високим ступенем паралелізму архітектур самоорганізованих ШНС розпізнавання образів, відтворюючих їх біологічний і поведінковий контексти. Основні властивості сіток АРТ – стабілізація і стійкість процесу навчання (який є скінченим) і сталість процесу пошуку. В залежності від виду вхідних змінних і способу їх опрацювання розрізняють АРТ: з двійковими вхідними векторами (АРТ-1); з двійковими та неперервними вхідними векторами (АРТ-2); для моделювання біологічних процесів (АРТ-3), з двійковими вхідними векторами з використанням нечіткої логіки (Fuzzy-АРТ)[21].

Вважається, що *генетичні алгоритми* (ГА) були запропоновані та досліджені Холландом Дж. в 1975 році [33]. Проте, майже 10-ма роками раніше розроблений Івахненко О.Г. *метод групового урахування аргументів* (МГУА) [7-9] базувався на використанні одно- та багаторядних фільтруючих за певними критеріями алгоритмів саме для визначення “моделей оптимальної складності”, що і являли собою ніщо інше, як процедуру генетичного добору за певними ознаками якості. Відмінності МГУА від ГА полягають лише в тому, що родинні хромосоми останнього у МГУА мають вигляд складових поліноміальних попарних сполучень (“батьки”), а нові хромосоми (“діти”) – відібрані на етапах алгоритмування (кросовер, мутація) проміжні моделі складності в МГУА. Сітки, основані на МГУА [47], при складанні з m шарів можуть реалізовувати поліном степеня 2^m . “Просіюючи” на кожному з етапів багаторядного алгоритму через “решето” критеріїв отримані на попередньому етапі моделі (“особини” за термінологією ГА), можна керувати відбиранням найкращих. Більш того, МГУА є кращим методом для розв'язання задач ідентифікації та короткострокових прогнозувань, а для довгострокового прогнозування є не тільки найкращим, а й, можливо, унікальним методом, який забезпечує точність прогнозування за великим часом його випередження. Таким чином, як ГА, так і алгоритми МГУА є оптимізаційними і в ШНС можуть застосовуватися для пошуку оптимальних синаптичних ваг сіток зворотного поширення обмеженої розмірності [55]

, а також для визначення оптимальної структури ШНС під конкретну задачу [29]. Крім того, заслуговують на увагу дослідження, пов'язані з використанням ГА-підходу до *генетичного програмування* [40]. В загальному вигляді ГА можна подати наступною послідовністю кроків:

1. Подання всіх змінних, які оптимізуються, у вигляді бінарного або дійсного (в залежності від модифікації методу) ланцюжка параметрів – “хромосоми”, і далі випадковим чином створюється “популяція хромосом”;
2. Оцінювання кожної хромосоми з популяції за цільовою функцією;
3. Сортування всі оцінювань. Якщо кількість виконаних ітерацій не перевищує заздалегідь задане значення, то перехід до наступного кроку. Інакше рішенням є найкраща хромосома за результатами сортування на цьому кроці;
4. Відібрані на попередньому кроці найкращі представники популяції складають основу (“хромосом-батьків”) для створення за допомогою генетичних операцій (“кросовера” і “мутації”) нових хромосом (“хромосом-дітей”). Кросовер визначає рекомбінацію, за якою різні найкращі “хромосоми-батьки” даної популяції обмінюються деякими частинами бінарних ланцюжків. При цьому довільні двійкові розряди довільних хромосом можуть інвертуватися з малою ймовірністю цього процесу;
5. Оцінювання “хромосом-дітей” цільовою функцією і розміщення їх в популяцію на місця “гірших” її представників. Далі – перехід до пункту 3.

Позитивна якість ГА полягає в його універсальності – він дозволяє працювати як з неперервними, так і дискретними параметрами, і можливість його застосування не залежить від виду цільової функції; крім того, ГА спроможний знайти глобальний екстремум, тоді як інші алгоритми “зависають” на локальному екстремумі. Нарешті, ГА можна застосовувати на будь-якому етапі синтезу нечітких моделей, особливо у тих випадках, коли використання градієнтних методів неможливе.

До недоліків слід віднести досить низьку швидкість, яка значно знижується із зростанням кількості параметрів, що оптимізуються; крім того, ГА має велику кількість своїх параметрів – розмір популяції, довжина хромосоми, ймовірності кросовера і мутації, кількість “хромосом-дітей” тощо. Істотним недоліком є також явище “стогнації” як наслідок “невдалого” вибору параметрів з виродженням популяції або відсутністю прогресу.

Отже, розширення інструментального “асортименту” різновидів ШНС, дозволяючи більш “тонко” і професійно підходити до вибору моделі/топології ШНС, адекватної умовам обслуговування прикладної задачі, разом з тим ускладнюють цей процес через його багатоваріантність. За недостатньої професійної підготовленості користувача це може призводити до значних додаткових матеріальних (апаратних, обчислювальних) і/або часових витрат.

Постановка задачі. Як і у будь-якій іншій галузі знань, інтелектуальний прорив у широкому використанні НТ став можливим, тільки коли для його реалізації були дотримані *необхідні* і *достатні* умови (рис. 1):

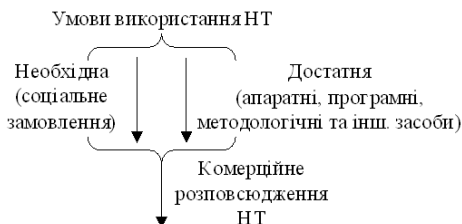


Рис. 1 – Умови комерційної реалізації ШНС

Ця умова виконується як завдяки явно вираженій зацікавленості прикладних користувачів у нових підходах до розуміння функцій природного (біологічного) інтелекту шляхом створення математичної моделі самосвідомості та усвідомлення своїх можливостей у розв’язанні складних інтелектуальних задач (*необхідна умова*), так і досягненням в області інструментальних (апаратних, програмних, методологічних тощо) засобів для побудування ефективних за швидкодією, точністю моделювання і потужності (розмірності) ШНС (*достатня умова*).

Означення 1. *Штучні нейронні сітки* – математичні моделі та їх програмні і/або апаратні реалізації, виконані за принципом організації біологічних нейросіток, уявляють собою систему з’єднаних і взаємодіючих простих процесорів (штучних нейронів).

Роль і уявлення про ШНС, які й визначають зростаючий попит на застосування ШНС в різних сферах життєдіяльності людини, багатогранні [6]. З точки зору:

- *машинного навчання* – ШНС уявляють собою окремий випадок методів розпізнавання образів, кластеризації, дискримінантного аналізу тощо;
 - *математичної* – навчання ШНС є багатопараметричною задачею нелінійної оптимізації;
 - *кібернетики* – ШНС використовуються в задачах адаптивного управління і прогнозування;
- *розвитку обчислювальної техніки* – програмування ШНС є способом розв’язання проблеми ефективного паралелізму;
 - *штучного інтелекту* – ШНС є основою філософської течії конективізму та основним напрямком у структурному підході по вивченню можливостей побудування/моделювання природного інтелекту за допомогою комп’ютерних алгоритмів.

Отже, обчислювальні системи на основі ШНС володіють якостями, яких позбавлені машини з архітектурою фон Неймана, проте притаманних мозку людини: масовий паралелізм; розподілене подання інформації та обчислень; здібність до навчання та узагальнення; адаптивність; властивість контекстуального оброблення інформації; толерантність до похибок; низьке енергоспоживання.

Апаратна реалізація ШНС – *нейрокомп'ютер* – має суттєві позитивні для широкого впровадження відмінності (як за структурою, так і за класами розв'язуваних задач) від обчислювальних машин, виконаних у відповідності з традиційною архітектурою фон Неймана. Порівняльні характеристики нейрокомп'ютерів і традиційних комп'ютерів наведено у табл. 1.

Таблиця 1.

Порівняльні характеристики традиційних ЕОМ і нейрокомп'ютерів

Категорія порівняння	ЕОМ традиційної архітектури	Нейрокомп'ютер
Процесор	Складний Високошвидкісний Одни або декілька	Простий Низькошвидкісний Велика кількість
Пам'ять	Відокремлена від процесора Локалізована Адресація не за змістом	Інтегрована у процесор Розподілена Адресація за змістом
Обчислення	Централізовані Послідовні Програми, що зберігаються	Розподілені Паралельні Самонавчання
Надійність	Висока уразливість	Живучість
Спеціалізація	Чисельні та символічні операції	Проблеми сприйняття
Середовище функціонування	Суворо визначене Суворо обмежене	Без обмежень

Основні тези, які складають підґрунтя використання ШНС і нейромоделювання, полягають у наступному:

- нейросітка відтворює структуру і властивості нервової системи живих організмів – складається з великої кількості простих обчислювальних елементів (нейронів) і володіє більш складною поведінкою по відношенню до можливостей кожного окремого нейрона;
- нейросітка отримує на вході набір вхідних сигналів і видає відповідну до них відповідь (сигнали на виході НС), яка і є розв'язком задачі;
- як і природна біологічна нейросітка, ШНС може навчатися розв'язанню задач через наявність внутрішніх адаптивних параметрів

нейронів і своєї структури та, змінюючи їх, може міняти свою поведінку;

- місце програмування займає навчання, тренування НС, і для розв’язання задачі не треба програмувати алгоритм;
- нейросітка навчається розв’язанню задачі на деякому “асоціаторі” – наборі ситуацій, кожна з яких описує значення вхідних сигналів НС і бажану для них відповідь. При цьому “асоціатор” задає набір еталонних ситуацій з відомими розв’язками, а НС при навчанні сама знаходить залежності між вхідними сигналами та відповідями.

Саме широка варіативність ШНС, наявність тонких відмінностей у принципах побудови і функціонування, відсутність строгих класифікаційних ознак і чітких рекомендацій щодо сфер застосування і розв’язуваних задач вимагають високої усвідомленості в питаннях вибору топологій ШНС у прикладного користувача, що виявляється не завжди виконуваним, перш за все, через його професійну невідповідність у галузі сучасних НТ. Отже, розрив між *об’єктивною доцільністю* використання новітніх досягнень в НТ для розв’язання прикладних задач і *суб’єктивною непоінформованістю* кінцевого користувача обмежує можливість ефективного використання ШНС. Через це ШНС використовуються в двох варіантах: будується нейросітка, яка розв’язує певний клас задач; під кожний екземпляр задачі будується деяка нейросітка, яка знаходить квазіоптимальний розв’язок цієї задачі. Проте в обох випадках прийняття рішення покладається на користувача.

Покращення ситуації нам уявляється в наступному:

- формуванні *набору вирішних класифікаційних ознак* (НВКО) і створенні класифікатора ШНС;
 - побудуванні чіткої *логічної моделі поетапного синтезу* (МПС) ШНС;
- створенні строгої *узагальненої моделі вибору типових топологій* (УМВТТ) ШНС для конкретних прикладних задач, що базується на формалізованих моделях подання знань [2] з використанням НВКО та *агентно-орієнтованого підходу* [19].

Означення 2. *Набір вирішних класифікаційних ознак ШНС* – така їх мінімально допустима сукупність, яка є *необхідною* для формалізації процесу подання основних властивостей і вибору задовольняючих топологій нейросіток і *достатньою* для адекватного обслуговування вимог (критеріїв оцінки) з боку прикладної розв’язуваної задачі.

Означення 3. *Модель поетапного синтезу ШНС* – така послідовність їх перебирання в просторі НВКО, яка, будучи виконувана користувачем і/або *мультиагентною підсистемою автоматизованого вибору* (МАПАВ), відтворює принципи агентно-орієнтованого підходу та автономно дозволяє виокремити топологію/топології ШНС, здатну/здатні задовольнити критерії обслуговування властивостей розв’язуваної задачі.

Означення 4. *Агент / мультиагент* – це програмно-апаратний чи програмно-емуляційний автономний компонент мультиагентної підсистеми, який функціонує за певним сценарієм/алгоритмом на основі *конкретної унікальної послідовності дій* в інтересах досягнення поставлених користувачем перед нею цілей.

Твердження 1. Алгоритм дії агента може мінятися і коректуватися по ходу виконання завдання заради досягнення мети.

Твердження 2. Конкретна унікальна послідовність дій (КУПД), що приводить до мети, шукається агентом кожного разу з урахуванням відповідних критеріїв обслуговування властивостей даної розв’язуваної задачі.

Означення 5. *Мультиагентна підсистема (МАП)* – це складна система, в якій функціонують два чи більше інтелектуальних агента.

Формування набору вирішених класифікаційних ознак. Проведений аналітичний перегляд основних напрямків розвитку ІТ, а також результати численних досліджень в інших джерелах, присвячених даній проблемі, дозволяють сформувати НВКО і запропонувати класифікатор ШНС, які в сукупності сприятимуть послідовності обґрунтування вибору відповідностей “розв’язувана прикладна задача – топологічна модель ШНС”.

Досвід роботи по створенню агентно-орієнтованого інтерпретатора ШНС дозволяє стверджувати, що наведений нижче перелік НВКО є і необхідним, і достатнім для розв’язання поставленої задачі і містить:

- *тип вхідної інформації* – аналогові і двійкові;
- *модель інтелектуального нейрона* – Мак-Каллоха–Питтса, одношаровий перцептрон, Адаліна, Фукушима, сігма-пі-нейрон, Хопфілда, Гроссберга;
- *структуру* – прямоспрямовані, зворотного поширення (с прямими, непрямыми, латеральними зв’язками, повнозв’язні);
- *архітектуру* – одношарові (одношаровий перцептрон, Адаліна, Н-Адаліна), багатошарові (багатошаровий перцептрон, Мадаліна, сітка на основі МГУА тощо);
- *методи навчання* – з вчителем, без вчителя, шляхом самоорганізації, конкурентні, стохастичні, градієнтні;
- *настроювання ваг* – фіксоване і динамічне;
- *критерії оцінки* – простота реалізації, швидкість та ефективність алгоритму навчання, точність апроксимації, потужність (розмірність) при розв’язанні прикладної задачі, кваліфікація користувача, наявність локальних мінімумів, можливість фільтрації та відновлення.

Класифікація штучних нейронних сіток. Структурна схема класифікації ШНС з бажаним НВКО подана на рис. 2.

Запропонований НВКО потребує уточнення деяких тверджень, означень:

Означення 6. Аналогова вхідна інформація – інформація на вході ШНС, подана у формі дійсних чисел;

Означення 7. Двійкова вхідна інформація – інформація на вході ШНС у вигляді нулів та одиниць.

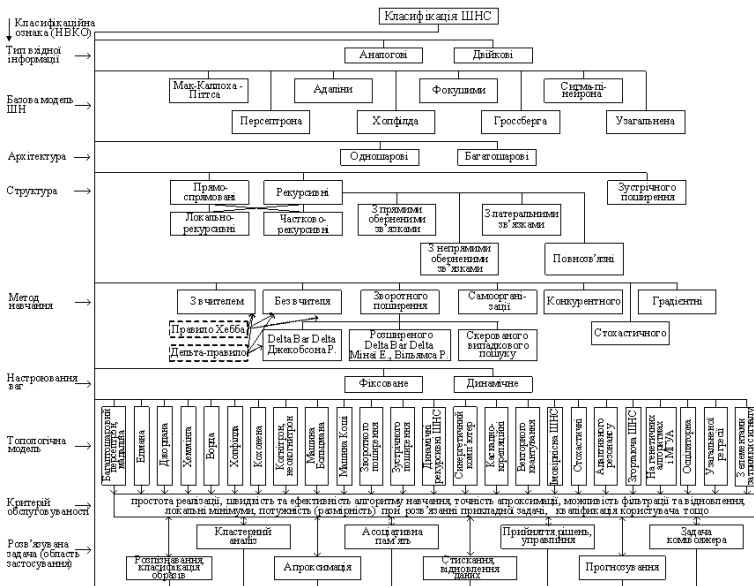


Рис. 2 – Класифікація ШНС з базисним НВКО

Отже, можна виділити два основних підходи до реалізації НС: *цифровий* та *аналоговий*. Перевагою аналогових реалізацій є: *висока швидкодія, надійність та економічність*. Проте сфера можливого масового використання навчуваних аналогових нейронів досить вузка. Це обумовлене великою складністю апаратної реалізації високоефективних навчальних алгоритмів і необхідністю спеціальної підготовки потенційних користувачів для оптимальної організації адаптивного процесу. У той же час, широке розповсюдження можуть отримати аналогові нейрокомп'ютери/нейросітки з фіксованою або незначно підналагоджуваною структурою зв'язків – *нейропроцесори*;

Твердження 3. Задача створення нейропроцесорів зводиться до навчання цифрової нейросіткової моделі необхідній поведінці на звичайному цифровому комп'ютері;

Означення 8. Базова модель ШНС – математична та її програмна і/або апаратна реалізація, на основі яких будується ШНС;

Твердження 4. Архітектура ШНС визначається кількістю шарів і кількістю нейронів у шарі.

З позицій архітектурного побудування ШНС важливу роль відіграє *нелінійність активаційної функції* через те, що, якби вона не володіла даною властивістю або не входила в алгоритм роботи кожного нейрона, результат функціонування будь-якої n -шарової НС зводився б до добутку вхідного вектора \mathbf{X} сигналів x_i на матрицю вагових коефіцієнтів. Тобто фактично така НС була б еквівалентна одношаровій НС з ваговою матрицею \mathbf{W} єдиного шару.

Означення 9. *Активаційна функція* – функція $f_a(*)$, яка визначає правило переходу нейрона, що знаходиться в момент часу τ в стані $g(\tau)$, у новий стан $g(\tau + 1)$ при надходженні вхідних сигналів x_i ;

Твердження 5. Структура ШНС визначається напрямом та охопленням зв'язків між шарами та окремими нейронами, тобто характером зв'язків.

Визначення кількості проміжних шарів і кількості нейронів в них є важливим при моделюванні сітки. Зазвичай, застосовуючи архітектуру до визначених проблем, використовують загальні правила, зокрема:

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

Оскільки всі штучні нейронні сітки базуються на концепції функціонування нейронів, правил та механізмів їх з'єднань, а також передає функцій щодо процедури активації, існує подібність між різними структурами або архітектурами нейронних сіток. Більшість змін походять з різних правил навчання. Таким чином, після того, як визначено кількість шарів і число нейронів в кожному з них, потрібно знайти значення для синаптичних ваг і порогів сітки, які спроможні мінімізувати похибку спродукованого результату. Саме для цього існують *алгоритми навчання*, де відбувається підгонка моделі сітки до наявних навчальних даних;

Твердження 6. Штучні нейронні сітки не програмується у звичайному сенсі – вони навчаються;

Означення 10. *Навчання з вчителем* – таке навчання ШНС, яке припускає існування цільового вектора – бажаного виходу ШНС для

кожного вхідного вектора і виконується із залученням правил навчання: Хебба, дельта-правила/його модифікацій, алгоритму зворотного поширення похибки тощо;

Твердження 7. Разом вхідний та цільовий вектори утворюють навчаючу пару, і зазвичай сітка навчається на деякому сполученні таких пар;

Означення 11. *Навчання без вчителя* – таке навчання, яке відбувається за відсутності цільового вектора, а *навчаючий алгоритм* при цьому підналагоджує ваги таким чином, щоб отримати узгоджені вектори на виході ШНС, причому подання досить близьких вхідних векторів дає однакові вектори на виході;

Означення 12. *Навчання з використанням алгоритму зворотного поширення похибки* – це навчання, яке, використовуючи поширення принципів дельта-правила, реалізує градієнтний метод опуклого функціоналу похибки в багатозарядних ШНС із структурою прямого поширення на моделях ШН з диференційованими функціями активації;

Твердження 8. Похибка для конкретної топології ШНС з варіюваними значеннями синаптичних ваг і порогів сітки (так званих *вільних параметрів топології* ШНС) визначається шляхом проходження через неї всієї навчальної множини і порівняння спродукованих на виході значень з цільовими/бажаними. Множина похибок утворює *функцію похибок*, яку можна розглядати як похибку сітки. В якості функції похибок найчастіше використовують *суму квадратів похибок*.

Твердження 9. Мета навчання ШНС за алгоритмом АЗПП полягає в знаходженні на багатовимірній поверхні найнижчої точки, яка відповідає глобальному мінімуму.

Для кращого розуміння навчання за алгоритмом АЗПП потрібно роз'яснити поняття *поверхні станів*. Кожному значенню синаптичних ваг і порогів сітки (вільних параметрів моделі кількістю N) відповідає один вимір в багатовимірному просторі. Вимір $N + 1$ -ий відповідає похибці сітки. Для різноманітних сполучень ваг відповідну похибку сітки можна зобразити точкою в $N + 1$ -вимірному просторі, всі ці точки утворюють деяку поверхню станів.

Поверхня станів має складну будову і досить неприємні властивості, зокрема, наявність локальних мінімумів (точки, найнижчі в своєму певному околі, але вищі від глобального мінімуму), плоскі ділянки, сідлові точки і довгі вузькі яри. Аналітичними засобами неможливо визначити розташування глобального мінімуму на поверхні станів, тому навчання нейросітки 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 напрямком найкоротшого спуску по поверхні з заданої точки. Якщо трошки просунути по ньому, похибка зменшиться. Зрештою алгоритм зупиняється в нижній точці, що може виявитись лише локальним мінімумом (в ідеальному випадку - глобальним мінімумом). Складність

застосування АЗПП полягає у виборі довжини кроків переходу вектора градієнта поверхні похибок, через що або зростає кількість ітерацій і разом з тим час пошуку (при замалих кроках, або *epochax*), або зростає ймовірність “перестрибування” оптимального розв’язку чи руху в невірному напрямку (при занадто великих кроках). Ітеративний процес навчання припиняється або коли *пройдена визначена кількість epoch*, або коли *похибка досягає визначеного рівня малості*, або коли *похибка перестав зменшуватись* (користувач переважно сам вибирає потрібний критерій зупинення процесу);

Означення 13. *Навчання шляхом самоорганізації* – таке навчання, за яким аналогічні вхідні сигнали збуджують сусідні нейрони, і такий “колективізм” може здійснюватися завдяки оберненим зв’язкам. Ваги змінюються на кожному кроці навчання, і їх змінювання залежить як від сусідніх вхідних образів, так і ймовірного розподілення, за яким пропонуються для навчання допущені вхідні образи;

Означення 14. *Сітки з фіксованими зв’язками* – такі, у яких коефіцієнти ШНС призначаються відразу, виходячи з умов задачі;

Означення 15. *Сітки з динамічними зв’язками* – такі, у яких в процесі навчання відбувається налагодження синаптичних ваг;

Означення 16. *Топологія ШНС* – це модель нейросітки з певними архітектурою, структурою, методами навчання і настроювання ваг і побудовані на основі певних моделей ШН;

Твердження 10. Топологічні властивості (ТВ) ШНС не змінюються при будь-яких структурно-архітектурних перетвореннях без втрати суттєвості певного виду нейросіток. Прикладом ТВ ШНС є розмірність/потужність останніх, що визначається кількісною характеристикою: шарів, нейронів у шарі, а також зворотних зв’язків, входів та виходів сітки тощо.

Означення 17. *Топологічний простір ШНС* – множина моделей сіток будь-якої топології, в якій через НВКО в той чи інший спосіб визначені граничні значення ТВ ШНС.

Логічна схема поетапного синтезу ШНС. Наведена на рис. 3 схема містить, з урахуванням запропонованого НВКО, відображення послідовності їх урахування при ітераційній процедурі синтезу ШНС на основі КУПД інтелектуальних агентів в процесі функціонуванні МА-ПАВ.

Заключна процедура поетапного синтезу ШНС зводиться до перевірки у реляційному перебиранні інтелектуальними агентами умов виконання критеріїв обслуговуваності поточним вектором можливостей конкретної топології ШНС вимог з боку прикладної задачі.

Твердження 11. В мультиагентній системі функції інтелектуалізованих агентів можуть реалізуватися певними топологіями ШНС.

Означення 18. *Критерій обслуговуваності* – показник задоволення топологією ШНС вимог з боку розв’язуваної задачі і професійної відповідності користувача.

Друга складність полягає у кількісному визначенні вагомості реляційних зв'язків між вирішними класифікаційними ознаками і реалізується експертним рейтинговим оцінюванням альтернативних варіантів із застосуванням методів ранжирування та попарних порівнянь [14]. Відповідні дослідження потребують глибоких і ретельних проробок щодо однозначності розв'язання цієї частини задачі адекватного вибору ШНС. На рис. 4 наведені умовні кількісні результати попередніх обробок джерельних даних, які, не претендуючи на строгість експертного оцінювання, можна розглядати як приклад загальної постановки при розв'язанні задачі адекватного вибору ШНС в інтерпретаційній моделі.

Узагальнення та висновки. Можна констатувати, що з найпоширеніших сфер застосування ШНС є *розпізнавання образів*. Це тим важливіше враховувати через те, що у більшості інших застосувань задача розпізнавання, навіть не будучи метою використання у явному вигляді, все ж припускає цю процедуру при надходячих на вхід образах з наступною їх класифікацією, кластеризацією тощо. Під кластеризацією мають на увазі розбиття множини вхідних образів на класи, коли ані кількість, ані ознаки класів не відомі. Сітка може констатувати, що вхідний образ не відноситься до жодного з виділених класів – це свідчить про те, що з'явилися нові образи з відмінними ознаками від тих, що визначають належність до певного класу. Тобто сітка виявляє нові, невідомі раніше класи образів, виконуючи функції *кластеризатора*.

З іншого боку, будь-яке застосування ШНС для розв'язання прикладних задач асоціюється з механізмом *прийняття рішень при керуванні*, рівно як і прийняття рішень, в свою чергу, базується на перебиранні можливих варіантів і виборі задовольняючого заданий показник ефективності. А це й є визначенням такого стану ШНС, яке відповідає “образу” ефективного керування [11].

Прогнозування можна віднести до так званої “імпровізуючої” властивості ШНС у вигляді узагальнення і виокремлення схованих взаємозв'язків між даними на вході і виході. Інакше кажучи, якщо у вхідній виборці даних є сховані закономірності їх впливу на змінювання даних на виході, то можна використовувати ШНС як *прогнозуючий інструмент*. А вже далі, на підставі цієї прогнозуючої функції, ШНС може приймати рішення при керуванні.

В багатьох випадках буває необхідно підібрати для функції, заданої тільки таблично або графіком і відображаючої результати експериментальних даних, аналітичний вираз, відбиваючий ці залежності. Формула ж може виявитися досить складною і некоректною для вимаганої мети (наприклад, функція має бути проінтегрованою, а інтеграл від неї не виражається через елементарні функції). Тоді емпіричні формули для зображення необхідної функції $f(x)$ вибирають наближену функцію $g(x)$ як набір з функцій певного виду, вимагаючи, щоб функція $g(x)$ якомога ближче співпадала з $f(x)$ на деякому інтервалі ($a \leq x \leq b$). Нелінійна характеристика при цьому може бути довільною: від сигмоїдальної до будь-якого хвильового пакету, синуса або багаточлена. Від

вибору нелінійної функції може залежати складність ШНС, але з будь-якою нелінійністю сітка залишається універсальним *апроксиматором* і при коректному виборі структури може досить точно апроксимувати функціонування будь-якого неперервного автомата.

Здатність ШНС до виявлення взаємозв'язків між різними параметрами дозволяє подавати дані великої розмірності більш компактно, якщо останні тісно взаємопов'язані одне з одним. Нейросітки із зворотним процесом – відновлення вихідного образу з його частини або з його спотвореного (зашумленого/пошкодженого) оригіналу – дістали назву (*автоасоціативної пам'яті*). Створення на ШНС *гетероасоціативної пам'яті* дозволяє реалізовувати пам'ять, яка адресується за вмістом.

Отже, наведені в статті класифікатор ШНС, побудований на НВКО, а також запропонована інтерпретаційна модель вибору адекватної топології ШНС є основою для автоматизації цього процесу, що дозволить усунути суб'єктивний фактор некомпетентності користувача і реалізувати в автоматизованому виді логічну схему поетапного синтезу ШНС.

Література

1. Айзерман М.Ф. Метод потенциальных функций в теории обучения машин / Айзерман М.Ф., Браверман Э.М., Розоноэр Л.И. – М.: Наука, 1970
2. Барцев С.И. Адаптивные сети обработки информации / Барцев С.И., Охонин В.А. – Красноярск: Ин-т физики СО АН СССР, 1986. Препринт № 59Б. – 20 с.
3. Бонгард М.М. Проблемы узнавания. – М.: Физматгиз, 1967
4. Винер Н. Кибернетика или управление и связь в животном и машине. – М.: Сов. радио, 1968. – 326 с.
5. Галушкин А.И. Синтез многослойных систем распознавания образов. – М.: Энергия. – 1974.
6. Горбань А.Н. Нейроинформатика: кто мы, куда мы идём, как путь наш измерить? // Вычислительные технологии. – М.: Машиностроение. – 2000. – № 4. – С. 10-14
7. Ивахненко А.Г. Метод группового учета аргументов – конкурент метода стохастической аппроксимации // Автоматика. – 1968. – № 3. – С. 58-72.
8. Ивахненко А.Г. Самообучающиеся системы распознавания и автоматического управления. – К.: “Техніка”, 1969. – 392 с.
9. Ивахненко А.Г. Системы эвристической самоорганизации в технической кибернетике. – К.: “Техніка”, 1971. – 372 с.
10. Петров А.П. О возможностях перцептрона // Известия АН СССР, Техническая кибернетика. – 1964. – №6
11. Чернодуб А.Н. Обзор методов нейроуправления / Чернодуб А.Н., Дзюба Д.А.//Проблемы программирования. – 2011. – № 2. – С. 79-94

12. Ямпольский Л.С. Автоматизированные системы технологической подготовки робототехнического производства / Ямпольский Л.С., Калинин О.М., Ткач М.М. – К.: Вища шк., 1987. – 271 с.
13. Ямпольский Л.С. Объектно-ориентированный выбор топологии нейросетей при решении прикладных задач / Ямпольский Л.С., Лисовиченко О.И. // Стратегия качества в промышленности и образовании. Материалы VIII Международной конференции, 8-15 мая 2012. – Варна, Болгария: Технический ун-т. – 2012. – С. 475 – 478
14. Ямпольский Л.С. Системы искусственного интеллекта в планировании, моделировании и управлении (на укр. яз.) / Ямпольский Л.С., Ткач Б.П., Лисовиченко О.И. – К.: ДП “Вид. Дім “Персонал”, 2011. – 544 с.
15. Ackley D.H., Hinton G.E. and Sejnowski T.J. A Learning Algorithm for Boatman Machines // Cognitive Science. – 1985. – 9. – P. 147-169
16. Almeida L.B. A Learning Rule for Asynchronous Perceptrons with Feedback in a Combinatorial Environment // Proc. of the First IEEE International Conference on Neural Networks, USA, SanDiego, 1987. – Vol. 2. – P. 609-618
17. Anderson J.A. Two Models for Memory Organization // Mathematical Biosciences. – 1970. – 8. – P. 137-160
18. Bartsev S.I., Okhonin V.A. The algorithm of dual functioning (back-propagation): general approach, versions and applications. Krasnoyarsk: Biophysics Institute SB AS USSR of, 1989. Preprint, №107B. – 16 p.
19. Bellifemine F.L., Caire G. and Greenwood D. Developing Multi-Agent Systems with JADE. – Wiley, 2007.
20. Carpenter G.A. and Grossberg S. The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Networks // Computer. – 1988. – March. – P. 77-88
21. Carpenter G.A., Grossberg S. and Rosen D.B. Fuzzy ART: Fast Stable Learning and Categorization of Analog Input Patterns by an Adaptive Resonance System // Neural Networks. – 1991. – 4. – P. 759-771
22. Elman J.L. Finding Structure in Time // Cognitive Science. – 1990. – 14. – P. 179-211
23. Fahlman S.E. and Lebiere C. The Cascade-Correlation Learning Architecture / Carnegie Mellon Report. Nr. CMU-CS-88-162, 1990
24. Fukushima K. Neocognitron: A Self-organizing Neural Network for a Mechanism of Pattern Recognition Unaffected by Shift in Position // Biological Cybernetics. – 1980. – 36. – P. 193-202
25. Fukushima K. Cognitron: A Self-organizing Multiplayered Neural Network // Biological Cybernetics. – 1975. – 20. – P. 121-136
26. Grossberg S. Competitive Learning: From Interactive Activation do Adaptive Resonance // Cognitive Science. – 1987. – 11. – P. 23-63

27. *Grossberg S.* Nonlinear Neural Networks: Principles, Mechanism and Architectures // Neural Networks, 1988. – V.1. – № 1. – P. 17-62.
28. *Haken H., Fuchs A., Banzhatt W.* Mustererkennung durch synergetische Computer. Teil 1. Und 2 // Design and Elektronik, 1989
29. *Harp S. and Samad T.* Genetic Optimization of Neural Networks Architectures for Electric Utility Applications / Final Report. Electric Power Research Institute, Research Project № 8016-04, Palo Alto, CA. March. – 1994
30. *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
31. *Hebb D.* The Organization of Behavior. – New York: Willey Publications, 1949
32. *Hinton G.E.* Connectionist Learning Procedures // Artificial Intelligence. – 1989. – 40. – P. 185-234
33. *Holland J.L.* Adaptation in Natural and Artificial Systems: An Introductory Analysis with Application to Biology, Control and Artificial Systems. – The University of Michigan Press, Ann Arbor. – 1975
34. *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
35. *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
36. *Jordan M.I.* Attractor Dynamics and Parallelism in a Connectionist Sequential Machine // Proc. of the Eight Annual Conference of the Cognitive Science Society, Erlbaum, Hillsdale NJ, 1986. – P. 531-546
37. *Kohonen T.* Associative Memory: A System Theoretic Approach. – Berlin: Springer, 1977
38. *Kohonen T.* Self-Organized Formation of Topologically Correct Feature Maps // Biological Cybernetics. – 1982. – 43. – P. 59-69
39. *Kosko B.* Adaptive Bidirectional Associative Memories // Appl. Optics. – 1987. – 26. – № 33. – P. 4947-4960
40. *Koza J.P.* Genetic Programming: On the Programming of Computers by Means of Natural Selection. – Cambridge, MA, MIT Press, 1992
41. *Lang K.J., Waibel A.H. and Hinton G.E.* A Time-Delay Neural Network Architecture for Isolated Word Recognition // Neural Network. – 1990. – 3. – № 1. – P. 23-43
42. *Lippman R.P.* An Introduction to Computing with Neural Nets // IEEE ASSP Magazine. – 1987. – № 4. – P. 4-22

43. *McCulloch W.S. and Pitts W.* A Logical Calculus of the Ideas Immanent in Nervous Activity // Bulletin of Mathematical Biophysics. – 1943. – 5. – P. 115-133
44. *Minsky M.L. and Papert S.A.* Perceptrons // Cambridge (Mass): MIT press, 1969. – 321 p.
45. *Nelles O., Ernst S. and Isermann R.* Neuronale Netze sur Identifikation nichtlinearer dynamischer Systeme: Ein Überblick // Automatisierungstechnik. – 1997. – 45. – № 6. – S. 251-262
46. *Patterson D.* Artificial Neural Networks: Theory and Application. – Singapore: Prentice Hall Inc., 1996
47. *Pham D.T. and Liu X.* Modelling and Prediction using GMDH Networks of Adalines with Nonlinear Preprocessors // Intern. Journal System Science. 1994. – 25. – № 11. – P. 1743-1759
48. *Pineda F.J.* Dynamic and Architectures for Neural Computation // Journal of Complexity. – 1988. – 4. – P 216-245
49. *Powell V.J.D.* Radial Basis Functions for Multivariable Interpolation: A review / Proc. of IMA Conf. on Algorithms for the Approximation of Functions and Data, Shrivenham, UK. – 1985. – P. 143-167
50. *Rosenblatt F.* The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain // Psychological Review. – 1958. – 65. – P. 386-408
51. *Rumelhart D.E., Hilton G.E. and 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: MIT Press, 1986. – Vol. 1. – Chapt. 8. – P. 318-364
52. *Waibel A., Hanazawa T., Hinton G., Shikano K. And Lang K.J.* Phoneme Recognition using Time-Delay Neural Network // IEEE Trans. on Acoustics, Speech and Signal Processing, 1989. – 37. – № 3. – P. 328-339
53. *Werbos P.J.* Beyond regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D thesis. – Cambridge, MA, Harvard University, 1974
54. *Widrow B. and Hoff M.E.* Adaptive Switching Circuits / IRE WESCON Convention Record – New York, IRE, 1960. – P. 96-104
55. *Whitley D., Dominic S. and Das R.* Genetic Reinforcement Learning with Multilayer Neural Networks / In Belew and Booker. – 1991. – P. 562-570

Отримано 10.02.2012 р.