ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА КОМП'ЮТЕРНА ТЕХНІКА

https://doi.org/10.31649/1997-9266-2024-176-5-48-55

УДК 004.9

С. В. Яковин[1]
С. І. Мельничук[1]

ІМПЛЕМЕНТАЦІЯ БУЛЕВИХ ФУНКЦІЙ В ПЕРЦЕПТРОННИХ СТРУКТУРАХ

[1]Івано-Франківський національний технічний університет нафти і газу

Розглянуто особливості реалізації булевих функцій на основі перцептронних структур. Означено обмеження типової структури перцептрону, зокрема його здатність виконувати лише лінійно сепарабельні функції а також основні підходи до подолання таких обмежень шляхом ускладнення структр.

Також приділено увагу ймовірнісним характеристикам, які можна використати для покращення процесу навчання перцептронів, застосовуючи ймовірнісний алгоритм навчання та ймовірнісні перцептрони вищих порядків з Баєсовими ймовірностями. Розглянуто проблеми реалізації булевих функцій вищих порядків, зокрема необхідність використання поліномів вищих порядків, складність навчання та обчислювальну практичність. Подано підходи до декомпозиції булевих функцій вищих порядків на серії лінійно сепарабельних функцій.

До того ж, проаналізовано можливості одношарових перцептронів щодо класифікації лінійно роздільних об'єктів з бінарними результатами, та їхніх обмежень. Описано варіанти удосконалення одношарових перцептронних структур шляхом застосування степеневих функцій та ймовірнісних індикаторів сигналів для розширення їхніх класифікаційних можливостей.

Отримані результати підтверджують актуальність та перспективність проведення подальших досліджень і розробок нових перцептронних структур для ефективнішого розв'язання задач машинного навчання, включно з розробкою спеціалізованих структур та алгоритмів навчання для роботи з функціями вищих порядків.

Ключові слова: перцептронні структури, бінарні сигнали, булеві функції, опрацювання сигналів, ймовірнісні характеристики.

Вступ

Загалом, структура перцептронів у нейронній мережі є ключовим фактором, який визначає її здатність ефективно розв'язувати різноманітні задачі, адаптуватися до змінних умов та навчатися на основі попередньо ухвалених рішень (досвіду).

Найпоширеніша структура перцептрону включає такі компоненти:

Ваги (*weights*) — визначають важливість кожного вхідного сигналу в перцептроні. Ваги модифікуються під час навчання моделі для досягнення бажаного виходу.

Суматор (*summing function*) — обчислює зважену суму вхідних сигналів.

Функція активації (*activation function*) — визначає, чи буде нейрон активовано або ні, залежно від суми зважених входів і порогу. Функція активації може бути простою функцією порогу (stepfunction) або складнішою, такою як сигмоїда або ReLU.

Поріг (*threshold*) — значення, з яким порівнюється зважена сума для визначення виходу нейрона. Часто використовується зсув (bias), який діє як адаптивний поріг.

В роботі [1] пояснюється, що компоненти перцептрону працюють разом для вирішення завдань класифікації, де ваги і поріг використовуються для визначення, чи є вхідний сигнал позитивним або негативним прикладом. Це базова концепція, що лежить в основі багатьох сучасних методів машинного навчання. Додатково, в роботі [2] розглядається квантовий перцептрон, який імітує функцію активації класичного перцептрону, показуючи, як основні концепції перцептронів можуть бути адаптовані до нових областей наукових досліджень. Загалом, перцептрони є фундаментальними будівельними блоками в нейронних мережах, що демонструє їх важливість у розумінні машинного навчання.

Фактично, кореляція (correlation), згортка (convolution), добутки та додавання є ключовими математичними операціями, які лежать в основі багатьох алгоритмів машинного навчання, зоркрема перцептрони і глибинні нейронні мережі.

Основою перцептрона є лінійна комбінація вхідних сигналів з їхніми відповідними вагами, що є математичною операцією добутку (множення кожного входу на відповідну вагу) та додавання (підсумовування результатів цих множень). Ця сума потім подається на функцію активації для визначення виходу перцептрона.

Використання кореляції дозволяє оцінити ступінь лінійної залежності між двома змінними. У контексті машинного навчання, з перцептронами включно, вона може використовуватися для аналізу зв'язку між вхідними даними та мітками або виходом. Хоча сам перцептрон не виконує операцію кореляції як частину своєї основної функції, розуміння кореляцій між даними може допомогти в налаштуванні та інтерпретації моделей машинного навчання.

Згортка вважається фундаментальною операцією в області опрацювання сигналів та зображень і використовується в нейронних мережах загорткового типу (CNN), які є розширенням ідеї перцептрона. Згортка включає застосування операції фільтрування вхідних даних для виділення необхідних інформативних складових. Хоча перцептрон сам по собі не використовує згортки, ця операція є ключовою для розуміння того, як складніші мережі, що базуються на перцептроні, як-от CNN, вчаться розпізнавати складні шаблони в даних.

До того ж, у перцептронах використовуються різні ймовірнісні характеристики для поліпшення процесу навчання та здатності ухвалювати рішення на основі невизначеності даних, зокрема:

– ймовірнісний алгоритм навчання перцептрона, який зменшує час обчислень в процесі навчання, пропонуючи вирішення з компромісом між швидкістю обчислень та точністю. Це демонструє, що ймовірнісний підхід може бути корисним для оптимізації процесу навчання перцептронів, особливо коли час обчислень є критичним фактором [3].

– ймовірнісні перцептрони вищих порядків, які використовують структури Баєсових висновків, можуть навчатися статистиці задачі класифікації та відповідати на апостеріорні ймовірності за допомогою Баєсового оптимального висновку. Це демонструє здатність перцептронів до виконання складних обчислень, пов'язаних з ймовірністю [4].

Більше того, дослідження в галузі сприйняття людини вказують на те, що мозок представляє і опрацьовує інформацію в термінах ймовірностей, що може надати інсайти для розробки ймовірнісних перцептронних структур та інших моделей машинного навчання, які імітують цей аспект людського сприйняття [5].

В технічному аспекті, у разі імплементації булевих функцій вищих порядків за допомогою перцептронних структур виникає кілька проблемних задач, зокрема:

1. Обмежена експресивна здатність: стандартний перцептрон може виконувати лише лінійно сепарабельні функції, що серйозно обмежує його здатність представляти булеві функції вищих порядків. Булеві функції, які не є лінійно сепарабельними, не можуть бути безпосередньо реалізовані за допомогою одного перцептрона.

2. Необхідність у використанні вищих порядків ваг: для реалізації булевих функцій вищих порядків може знадобитися використовувати поліноми вищих порядків, що передбачає включення вищих порядків взаємодій між входами. Це може викликати експоненційне зростання кількості ваг, необхідних для моделювання таких функцій.

3. Складність навчання: навчання перцептронів для реалізації булевих функцій вищих порядків може бути значно складнішим у порівнянні з функціями нижчих порядків через збільшену кількість параметрів та складність оптимізаційної поверхні.

4. Практичність імплементації: імплементація великої кількості поліноміальних термів вищих порядків може бути непрактичною з погляду обчислювальних ресурсів та ефективності виконання, зокрема для великих систем з багатьма входами.

5. Підхід до декомпозиції: передбачає декомпозицію булевих функцій вищих порядків на серію лінійно сепарабельних функцій, які реалізують окремими перцептронами.

Таким чином, об'єктом дослідження є процеси реалізації булевих функцій на основі перцептронних структур у системах цифрового опрацювання даних.

Предметом дослідження є методи структурної реалізації перцептронних структур для спрощених задач, що зводиться до опрацювання бінарних векторів.

Метою дослідження є означення переваг та недоліків одношарових перцептроних структур а також отримання аналітичної міри оцінювання структурних і, як наслідок, обчислювальних затрат за реалізації пецептроної структури, що моделює логічні функцій першого та другого порядку.

Особливості реалізації найпоширеніших перцепронних структур для опрацювання бінарних сигналів

Реалізація перцептроних структур ґрунтується на використанні типової архітектури: синапси (вхідні сигнали) — агрегація (найчастіше зважена сума вхідних сигналів) — реакція (формування вихідного сигналу на основі функції активації). Основна мета моделі однорівневого перцептрона (ОРП) полягає в аналізі лінійно роздільних об'єктів з бінарними результатами. Тобто, однорівневі перцептрони можуть розпізнавати лише лінійно роздільні шаблони.

Для задачі класифікації з деякою функцією активації один вузол матиме одну лінію, яка розділяє точки даних, що формують шаблони. Як результат, одношарові нейронні мережі виявляються обмеженими в класифікаційних задачах, де необхідно геометрично розділити набори точок. У випадку двох входів такий роздільник є прямою лінією, у випадку трьох входів — це площина в тривимірному просторі, а з чотирма і більше входами розділення відбувається за допомогою «гіперплощини» в просторі вищої розмірності, яку складно візуалізувати. З урахуванням того, що одношарові перцептрони обмежені лінійною роздільністю, важливо встановити, чи можна розділити дані за допомогою ліній. Проте, немає простого способу визначити це, особливо коли кількість змінних велика. Наприклад, нейрон з n бінарними входами може мати 2^n різних комбінацій вхідних станів, а оскільки кожний вхідний набір може мати один з двох виходів, то існує 2^{2^n} можливих функцій з n змінними. Як можна побачити з табл. 1, імовірність випадково вибрати лінійно роздільну функцію є порівняно низькою, навіть за помірної кількості змінних.

Таблиця 1

n	2^{2^n}	Кількість лінійно роздільних функцій
1	4	4
2	16	14
3	256	104
4	65 536	1 882
5	4294967296	94 572
6	$1{,}844 \cdot 10^{19}$	15 028 134

Таким чином, однорівневий перцептрон немає можливості розділити лінійно-нероздільні множини, що підтверджено багатьма дослідження у цій сфері. Так Y. Hu (2007) розглянув обмеження адитивної моделі, використовуваної в ОРП, коли вхідні змінні не є незалежними, і запропонував перцептрон на основі нечіткого інтегралу, який демонструє кращу продуктивність порівняно з традиційними ОРП у задачах класифікації патернів [6]. Y. Zhong (2001) представив інтегровану мережу перцептронів для розпізнавання багатокласових лінійно роздільних шаблонів і підтвердив, що ОРП не можуть розпізнавати нелінійно роздільні шаблони [7]. Y. Hu (2010) зауважив, що припущення щодо традиційних функцій активації про те, що немає взаємозалежності між атрибутами, не завжди відповідає дійсності, тим самим підсилюючи лінійне обмеження ОРП [8]. Додаткові докази подані працях K. Singh та ін. (2014), які представили мемристивний перцептрон для класифікації лінійно роздільних шаблонів, визнаючи, що ОРП непридатні для нелінійно роздільних шаблонів [9].

Більше того, обмеження ОРП до лінійної роздільності вивчаються різними авторами, які пропонують поліпшені або альтернативні моделі для класифікації нелінійно роздільних шаблонів або покращення здатності класифікації ОРП шляхом розгляду поліпшених особливостей або функцій активації, а також використання різних алгоритмів, таких як генетичні алгоритми, для оптимізації процесу навчання та продуктивності ОРП, зокрема:

– Y. Hu і F. Tseng (2007) запропонували використання нечітких інтегралів для поліпшення ОРП [10];

– H. Ramchoun та ін. (2017) та (2019) досліджували вплив вибору архітектури та типів функцій активації на збіжність та продуктивність багатошарових перцептронів, вказуючи на обмеження однорівневих архітектур [11], [12].

– С. Мельничук, С. Яковин, М. Кузь (2018), дослідив використання функції інформаційної ентропії для емуляції логічних функцій за допомогою лише одного ОРП, продемонстрували можливість емулювати повний набір логічних функцій (NOT, AND, OR, і XOR) та вказали на те, що функціональні можливості перцептрона можуть бути розширені за допомогою ймовірнісних індикаторів сигналів, що генеруються сенсорами елементів S перцептрона [13]. Основні поліпшення щодо класифікації нелінійно роздільних шаблонів, запропоновані у матеріалах згаданої статті, включають:

1. Використання функції інформаційної ентропії: автори підходять до використання інформаційної ентропії як способу збільшення розпізнавальної здатності перцептрона. Функція ентропії оцінює розподіл ймовірностей виходів, що дозволяє перцептрону краще адаптуватися до різних шаблонів вхідних даних.

2. Розширення функціональних можливостей за допомогою індикаторів сигналу: додаткове використання ймовірнісних індикаторів сигналів може допомогти в розширенні можливостей перцептрона за рахунок додавання додаткової інформації про стан сенсорів.

3. Емуляція комплексних логічних функцій: автори також показали, що за допомогою запропонованого підходу можна реалізувати складні логічні функції, такі як XOR, які традиційно вважаються складними для класичних однорівневих перцептронів через їхню нелінійну роздільність.

Такі поліпшення можуть бути важливими для вдосконалення наявних чи розробки нових перцепронних структур, здатних виконувати складніші задачі класифікації, ніж ті, для яких вони традиційно використовувалися, і відкривають нові можливості для застосування перцептронів у сучасних системах штучного інтелекту.

Таким чином, результати проведеного аналізу показують, що хоча ОРП є ефективні для аналізу лінійно роздільних об'єктів з бінарними результатами, вони дійсно нездатні розпізнавати лінійно нероздільні шаблони, що потребує використання складніших мереж, як-от багатошарові перцептрони, для таких завдань, або застосування модифікованих перцептронних структур, тобто тих, у яких задіяні степеневі функції замість функцій активації від суми зважених входів або радіальних чи інших типів функцій активацій.

Проте поза увагою, за винятком [13], залишаються інші функції, зокрема статистичні, які можна застосувати для розв'язання задач класифікації [15].

Окремим аспектом функціонування засобів класифікації є спрощена задача, що зводиться до опрацювання бінарних векторів (матриць), яка виникає у разі класифікації зображень, таких як розпізнавання рукописного тексту, об'єктів, явищ тощо [16].

Для кращого розуміння сучасного контексту досліджень, цей розділ включатиме дані про наявні моделі розпізнавання, їхні розміри та точність. Зокрема, нейронні мережі застосовано до розпізнавання 2D-шаблонів на основі їхньої форми, де різні архітектури, включаючи багаторівневі перцептрони (БРП), мережі самоорганізації Кохонена та гібридні структури, призвели до різної точності розпізнавання та класифікації [17]. Зокрема, структуровані нейронні мережі виявилися ефективними для класифікації складних шаблонів та потенційно небезпечних ситуацій у системах спостереження, перевершуючи класичні статистичні процедури класифікації [14]. Щодо статистичного розпізнавання шаблонів, застосування функціональних мереж зі статистичними підходами продемонструвало обнадійливі результати в прикладних застосунках, пропонуючи надійні та ефективні техніки в задачах пошуку даних, web-пошуку а також вилучення мультимедійних даних з контенту [18]. Для оцінки ефективності одношарових перцептронів у задачах опрацювання бінарних сигналів доцільно застосувати такі критерії:

– максимальний відсоток булевих функцій, які вибрана структура зможе відтворити.

– мінімальна кількість параметрів (ваг), необхідна для моделювання.

– обчислювальна складність розрахунку реакції.

На цьому етапі з аналізу виключено оцінювання обчислювальної складності процедури навчання перцептрону, а увагу зосереджено на здатності структур відтворити булеві функції та на обчислювальній складності їхньої реакції.

Для спрощення аналізу використаємо булеві функції, оскільки вони мають невелику вхідну розмірність, проте дозволяють проаналізувати поведінку різних типів перцептронів на здатність апроксимувати всі можливі функції для вибраної кількості вхідних параметрів. Слід зауважити, що у випадку n вхідних змінних, кількість можливих комбінацій становитиме 2^n, а кількість можливих функцій — 2^{2^n}.

$$y = B_l\left(\overline{x}\right), \tag{1}$$

де $\overline{x} = \left(x_1, x_2, \ldots, x_n\right)$ — вектор вхідних значень булевої функції, l — порядковий номер булевої функції, $1 \le l \le 2^{2^n}$.

У цьому контексті важливим аспектом аналізу є визначення мінімальної кількості параметрів, або «ваг», необхідних для перцептронної структури щоб відтворити будь-яку з можливих булевих функцій з означеного вище набору. Це вимагає оптимізації структури перцептрона, для можливості забезпечення точного моделювання з мінімальними вимогами до обчислювальних ресурсів, що є критичним для практичного використання у великих масштабах, наприклад, у домені машинного навчання чи складних системах опрацювання даних.

Моделювання реалізовано за допомогою функції, що приймає як вхідний булевий вектор такі параметри, які зафіксовані для кожної окремої логічної функції B_l.

$$y = F\left(\overline{x}, \overline{p}_l\right), \tag{2}$$

де $\overline{p}_l = \left(p_1, p_2, ..., p_m\right)$ — вектор параметрів «ваг» перцептрону.

Для оцінювання достатньої (мінімально необхідної) кількості параметрів необхідно скористатись такими функціями:

$$E(\overline{x}) = \sum_{i=0}^{n-1} 2^i x_i = \overline{vx} = q\,, \tag{3}$$

де $\overline{v} = \left(2^0, 2^1, ..., 2^{n-1}\right)$;

$$G\left(q, \overline{p}_l\right) = F\left(\overline{x}, \overline{p}_l\right); \tag{4}$$

$$G\left(q, \overline{p}_l\right) = \sum_{j=0}^{2^n-1} p_{j+1} H\left(q - j\right), \tag{5}$$

де $p_j = y_j - y_{j-1}$, $y_0 = 0$, $\overline{y} = \left(y_1, y_2, ..., y_{2^n}\right)$ — усі послідовні значення булевої функції B_l, H — функція Хевісайда.

$$F\left(\overline{x}, \overline{p}_l\right) = \sum_{j=0}^{2^n-1} \left(y_{j+1} - y_j\right) H\left(\overline{vx} - j\right). \tag{6}$$

Як можна побачити з виразу (6), для опису будь-якої булевої функції із n вхідних змінних достатньо 2^n параметрів (ваг). Проте це достатня, але не необхідна умова, тому варто розглянути які можливі спрощення можна застосувати в залежності від кількості вхідних змінних, щоб зменшити кількість параметрів «ваг».

Схематично вираз (6) можна представити як трирівневу нейромережу, що складається з 1-го вхідного, 1-го вихідного та 2^n проміжних (прихованих) нейронів (рис. 2).

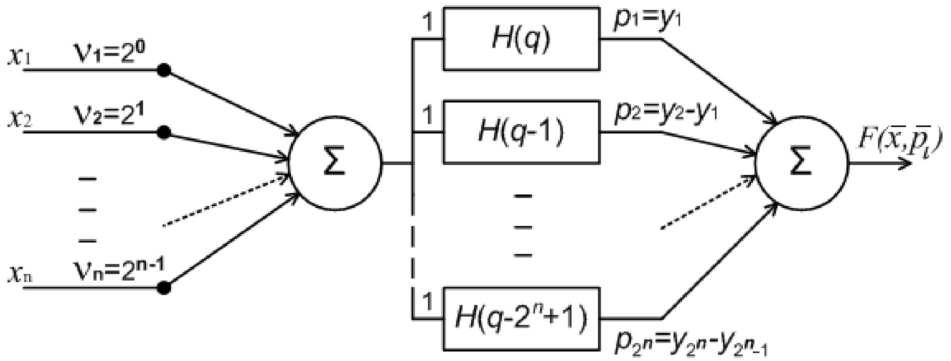

Рис. 2. Трирівнева мережа класичних перцептронів

Він потребує $n + 3 \cdot 2^n$ ваг, проте тільки 2^n з них залежать від булевої функції, що емулюється. Таку мережу, описану в [19], можна перетворити у 2-рівневу, як це продемонстровано на рис. 3, причому H_0 — варіант функції Хевісайда, що при аргументі 0 приймає нульове значення.

Подана дворівнева нейромережа із 2^n вхідних нейронів та 1 вихідного характеризується меншою кількістю шарів та нейронів потребує більшої кількості вагових коефіцієнтів $(n + 2) \cdot 2^n$.

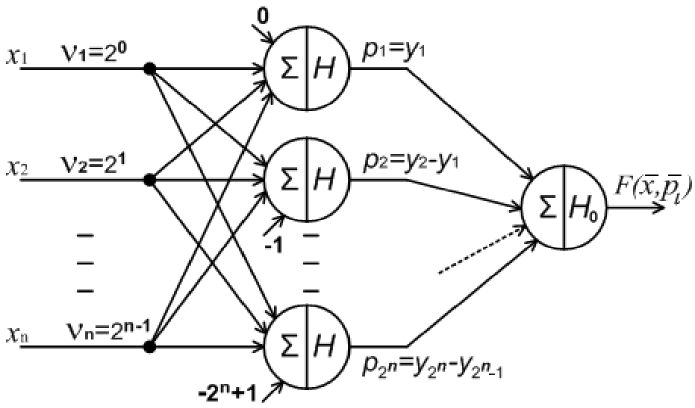

Рис. 3. Дворівнева мережа класичних перцептронів

Причому, як і в попередньому випадку тільки 2^n з них залежать від булевої функції, що емулюється. Таким чином, на відміну від нейромережі, запропонованої в [19], нейромережа, показана на рис. 2:

– потребує менше вагових коефіцієнтів в цілому;

– вимагає менших обчислень, оскільки тільки 2^n коефіцієнтів необхідно обчислити, причому за простими формулами;

– дозволяє реалізувати точну, а не наближену емуляцію булевих функцій.

На першому етапі розглянемо булеві функції порядку 1 (табл. 2).

Таблиця 2

x_1	Y	B_1	B_2	B_3	B_4
0	y_1	0	1	0	1
1	y_2	0	0	1	1

Для булевої функції першого порядку за (6) отримаємо:

$$F\left(\overline{x},\overline{p}_l\right)=y_1H(x)+\left(y_2-y_1\right)H(x-1). \tag{7}$$

На другому етапі розрахуємо булеві функції першого порядку (табл. 3).

Таблиця 3

x_1	Y	B_1	B_2	B_3	B_4
0	y_1	0	1	0	1
1	y_2	0	0	1	1
	p_1	0	1	0	1
	p_2	0	-1	1	0
	$F(\overline{x},\overline{p})$	0	$H(x)-H(x-1)$	$H(x-1)$	$H(x)$

Як можна побачити, для опису будь-якої функції достатньо 2-х параметрів, проте, підібравши функцію певним чином, можна обійтися 1-м параметром:

$$F\left(\overline{x},\overline{p}\right)=H\left(x+2-p\right)-H(x-p). \tag{8}$$

Результати моделювання булевих функцій першого порядку одним параметром див. табл. 4.

Таблиця 4

x_1	Y	B_1	B_2	B_3	B_4
0	y_1	0	1	0	1
1	y_2	0	0	1	1
	p_1	0	1	3	2
	$F(\overline{x},\overline{p})$	$H(x+2)-H(x)$	$H(x+1)-H(x-1)$	$H(x-1)-H(x-3)$	$H(x)-H(x-2)$

Таким чином, одного параметра необхідно та достатньо для моделювання логічних функцій першого порядку. За аналогією, для функцій другого порядку отримаємо:

$$F\left(\overline{x},\overline{p}_l\right)=\sum_{j=0}^{3}\left(y_{j+1}-y_j\right)H\left(\overline{vx}-j\right). \tag{9}$$

В результаті, для представлення всіх можливих булевих функцій n-го порядку необхідно та достатньо 2^n біт (доводиться через перетворення всіх значень булевої функції у число). Таким чином, для задач великої розмірності ($n \geq 30$) необхідно шукати компроміс між повнотою опису та об'ємом необхідної пам'яті.

Висновки

Розглянуті особливості реалізації одношарових перцептроних структур дозволяють означити переваги та недоліки їхнього застосування, зокрема до переваг слід віднести простоту структурної реалізації, у них є чітке геометричне тлумачення як гіперплощина, яка розділяє два класи даних. До одного із суттєвих недоліків таких перцептроних структур варто віднести їхню обмежену виразність та здатність до узагальнення, зокрема вони не здатні налаштовуватись (навчитись) для класифікації лінійно нероздільних шаблонів, таких як XOR, кола чи спіралі.

Більше того, в ході дослідження отримано аналітичні міри (8) та (9), що дозволяють оцінити компонентну складність реалізації пецептроної структури, що моделює логічні функцій першого та другого порядку. Встановлено, що орієнтація на лінійне розділення даних зумовлює схильність таких структур до перенавчання та їхню вразливість до шумів у даних, що зумовлено намаганням адаптуватись до флуктуацій (неідеальностей) даних за допомогою лінійного розділення.

Проте, вдосконалення одношарових перцептроних структур можливе, зокрема шляхом застосування степеневих, логарифмічних та інших нестандартних функцій активації, включно зі статистичними функціями, які дозволяють розширити межі їхніх класифікаційних можливостей. Тобто, актуальною залишається потреба у вдосконалені наявних та створенні нових перцептроних структур, які б дозволили розширити функціонал нейромереж, зокрема на основі застосування ймовірнісних оцінок вхідних сигналів, що може дозволити реалізувати алгоритми та методи, які полегшать імплементацію складних булевих функцій на платформі нейронних мереж, включно з розробкою спеціалізованих структур та алгоритмів навчання, які краще підходять для роботи з функціями вищих порядків. Також, застосування згаданих методів може знизити ризик перенавчання, забезпечуючи ліпше балансування між засвоєнням тренувальних даних та здатністю до узагальнення на нових даних, що в кінцевому випадку розширює перспективи для застосування одношарових перцептроних структур в сучасних задачах машинного навчання, які потребують більшої виразності та гнучкості.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

[1] V. Silaparasetty, Perceptrons. *Deep Learning Projects Using TensorFlow*, 2021. https://doi.org/10.1142/9789814343039_0007.

[2] M. Schuld, I. Sinayskiy, and F. Petruccione, *Simulating a perceptron on a quantum computer*, 2014. ArXiv, abs/1412.3635. https://doi.org/10.1016/j.physleta.2014.11.061.

[3] T. Hong, and S. Tseng, "A probabilistic perceptron learning algorithm," *J. Exp. Theor. Artif. Intell.*, no. 4, pp. 265-279, 1992. https://doi.org/10.1080/09528139208953751 .

[4] J. Clark, K. Gernoth, S. Dittmar, and M. Ristig, "Higher-order probabilistic perceptrons as Bayesian inference engines," *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 59, 5 Pt B, pp. 6161-74, 1999. https://doi.org/10.1103/PHYSREVE.59.6161 .

[5] J. Feldman, *Probabilistic models of perceptual features.* 2015. https://doi.org/10.1093/OXFORDHB/9780199686858.013.049 .

[6] Y. Hu, "Fuzzy integral-based perceptron for two-class pattern classification problems", *Inf. Sci.*, no. 177, pp. 1673-1686, 2007. https://doi.org/10.1016/j.ins.2006.09.009 .

[7] Y. Zhong, "An Integrated Perceptron Network and Learning Algorithm for Multi-Class Patterns Recognition," *Journal of Beijing Institute of Technology*, 2001.

[8] Y. Hu, "Pattern classification by multi-layer perceptron using fuzzy integral-based activation function," *Appl. Soft Comput.*, no. 10, pp. 813-819, 2010. https://doi.org/10.1016/j.asoc.2009.09.011 .

[9] K. Singh, C. Sahu, and J. Singh, "Linearly separable pattern classification using memristive crossbar circuits," *Fifteenth International Symposium on Quality Electronic Design*, pp. 323-329, 2014. https://doi.org/10.1109/ISQED.2014.6783343 .

[10] Y. Hu, and F. Tseng, "Functional-link net with fuzzy integral for bankruptcy prediction," *Neurocomputing*, no. 70, pp. 2959-2968, 2007. https://doi.org/10.1016/j.neucom.2006.10.111 .

[11] H. Ramchoun, M. Idrissi, Y. Ghanou, and M. Ettaouil, "Multilayer Perceptron: Architecture Optimization and training with mixed activation functions," *Proceedings of the 2nd international Conference on Big Data, Cloud and Applications*, 2017. https://doi.org/10.1145/3090354.3090427 .

[12] H. Ramchoun, M. Idrissi, Y. Ghanou, and M. Ettaouil, "Multilayer Perceptron New Method for Selecting the Architecture Based on the Choice of Different Activation Functions," *Int. J. Inf. Syst. Serv. Sect.*, no. 11, pp. 21-34, 2019. https://doi.org/10.4018/ijisss.2019100102 .

[13] S. Melnychuk, M. Kuz, and S. Yakovyn, "Emulation of logical functions NOT, AND, OR, and XOR with a perceptron implemented using an information entropy function," 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), 2018, pp. 878-882. https://doi.org/10.1109/TCSET.2018.8336337 .

[14] V. Murino, "Structured neural networks for pattern recognition. IEEE transactions on systems, man, and cybernetics," *Part B, Cybernetics, a publication of the IEEE Systems, Man, and Cybernetics Society*, vol. 28, no. 4, pp. 553-610, 1998. https://doi.org/10.1109/3477.704294 .

[15] A. Jain, R. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 22, pp. 4-37, 2000. https://doi.org/10.1109/34.824819 .

[16] S.Osowski, and D. Nghia, "Neural networks for classification of 2-D patterns. WCC 2000 - ICSP 2000," 2000 5th International Conference on Signal Processing Proceedings. 16th World Computer Congress 2000, no. 3, pp. 1568-1571 vol. 3. https://doi.org/10.1109/ICOSP.2000.893399 .

[17] E. El-Sebakhy, "Functional networks training algorithm for statistical pattern recognition. Proceedings. ISCC 2004. Ninth International Symposium on Computers And Communications," *IEEE Cat.*, no.04TH8769, 1, pp. 92-97, vol. 1, 2004. https://doi.org/10.1109/ISCC.2004.1358387 .

[18] J. Kittler, "Statistical pattern recognition in image analysis," *Journal of Applied Statistics*, no. 21, pp. 61-75, 1994. https://doi.org/10.1080/757582968 .

[19] X.Shenshu, Z. Zhaoying, Z. Limin, and Z. Wendong, "Approximation to Boolean functions by neural networks with applications to thinning algorithms," *Proceedings of the 17th IEEE Instrumentation and Measurement Technology Conference* [Cat. No. 00CH37066], 2000, vol. 2, no. 2, pp. 1004-1008. https://doi.org/10.1109/IMTC.2000.848892 .

Рекомендована кафедрою автоматизації та інтелектуальних інформаційних технологій ВНТУ

Стаття надійшла до редакції 3.06.2024

Яковин Сергій Васильович — аспірант кафедри комп'ютерних систем і мереж, e-mail: ysv_@ukr.net ;

Мельничук Степан Іванович — д-р техн. наук, професор, завідувач кафедри комп'ютерних систем і мереж, e-mail: stenni@ukr.net .

Івано-Франківський Технічний університет нафти і газу, Івано-Франківськ

S. V. Yakovyn[1]
S. I. Melnychuk[1]

Boolean Functions Implementation in Perception Structures

[1]Ivano-Frankivsk National Technical University of Oil and Gas

The authors consider the peculiarities of Boolean functions realization based on perceptron structures. The limitations of the typical structure of the perceptron are determined, in particular its ability to perform only linearly separable functions, as well as main approaches to overcoming such limitations, which are solved by complicating the structures.

Attention is also paid to probabilistic features that can be used to improve the learning process of perceptrons, using probabilistic learning algorithms and higher-order probabilistic perceptrons that use Bayesian probabilities. The problems of implementing Boolean functions of higher orders are considered, including the need to use polynomials of higher orders, the complexity of learning and computational practicality. Approaches to the decomposition of Boolean functions of higher orders into a series of linearly separable functions are presented.

In addition, the capabilities of single-layer perceptrons for the classification of linearly separable objects with binary results and their limitations are analyzed. Variants of improving single-layer perceptron structures by using power functions and probabilistic signal indicators to expand their classification capabilities are described.

The obtained results confirm the relevance and perspective of further research and development of new perceptron structures for more effective solution of machine learning problems, including the development of specialized structures and learning algorithms to work with higher-order functions.

Keywords: perceptron structures, binary signals, Boolean functions, signal processing, probabilistic characteristics.

Yakovyn Serhiy V. — Post-Graduate Student the Chair of Computer Systems and Networks, e-mail: ysv_@ukr.net ;

Melnychuk Stepan I. — Dr. Sc. (Eng.), Professor, Head of the Chair of Computer Systems and Networks, e-mail: stenni@ukr.net